Algebra
Lecture Notes for MTH 818/819
Fall 12/Spring 13

Ulrich Meierfrankenfeld

May 7, 2012

Chapter 1

Preface

These are the lecture notes for the classes MTH 818 in Fall 2012 and MTH 819 in Spring 2011. The notes are based on Hungerford's Algebra [Hun], but the proofs given here often diverge from Hungerford's.

The lecture will be updated frequently.

Contents

Chapter 2

Group Theory

2.1 Latin Squares

Definition 2.1.1. *Let G be a set and ϕ a function such that $G \times G$ is contained in the domain of G. map.*

(a) If $a, b \in G$ we write ab for $\phi(a, b)$. ϕ is called a binary operation *on G if $ab \in G$ for all $a, b \in G$. In this case the pair (G, ϕ) is called a* magma*.*

(b) $e \in G$ is called an identity element *if $ea = ae = a$ for all $a \in G$.*

(c) We say that (G, ϕ) is a Latin square *if for all a, b in G there exist unique elements x, y in G so that*

$$ax = b \text{ and } ya = b$$

(d) The multiplication table *of (G, ϕ) is the matrix $[ab]_{a \in G, b \in G}$.*

(e) The order *of (G, ϕ) is the cardinality $|G|$ of G.*

We remark that (G, ϕ) is a latin square if and only if each $a \in G$ appears exactly once in each row and in each column of the multiplication table. If there is no confusion about the binary operation in mind, we will just write G for (G, ϕ) and call G a magma.

Definition 2.1.2. *Let G and H be magmas and $\alpha : G \to H$ a map.*

(a) α is called a (magma) homomorphism *if $\alpha(ab) = \alpha(a)\alpha(b)$, for all $a, b \in G$.*

(b) α is called an isomorphism *if α is a homomorphism and there exists a homomorphism $\beta : H \to G$ with $\alpha\beta = \mathrm{id}_H$ and $\beta\alpha = \mathrm{id}_G$.*

(c) α is an automorphism *if $G = H$ and α is an isomorphism.*

Definition 2.1.3. *Let G and H be magmas.*

(a) The opposite *magma* G^{op} *is defined by* $G^{\mathrm{op}} = G$ *as a set and*

$$g \cdot_{\mathrm{op}} h = hg.$$

(b) An magma anti homomorphism$\alpha : G \to H$ *is a magma homomorphism* $\alpha : G \to H^{\mathrm{op}}$. *So* $\alpha(ab) = \alpha(b)\alpha(a)$.

Lemma 2.1.4. *(a) Let* G *be a magma. Then* G *has at most one identity.*

(b) Let $\alpha : G \to H$ *be a magma homomorphism. Then* α *is an isomorphism if and only if* α *is a bijection.*

Proof. (a) Let e and e^* be identities. Then

$$e = ee^* = e^*.$$

(b) Clearly any isomorphism is a bijection. Conversely, assume α is a bijection and let β be its inverse map. We need to show that β is a homomorphism. For this let $a, b \in H$. Then as α is a homomorphism

$$\alpha(\beta(a)\beta(b)) = \alpha(\beta(a))\alpha(\beta(b)) = ab = \alpha(\beta(ab)).$$

Since α is one to one (or by applying β) we get

$$\beta(a)\beta(b) = \beta(ab).$$

So β is an homomorphism. □

2.1.5 (Latin Squares of small order)**.** Below we list (up to isomorphism) all Latin square of order at most 5 which have an identity element e. It is fairly straightforward to obtain this list, although the case $|G| = 5$ is rather tedious). We leave the details to the reader, but indicate a case division which leads to the various Latin squares.

Order 1,2 and 3:

	e
e	e

	e	a
e	e	a
a	a	e

	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Order 4: Here we get two non-isomorphic Latin squares. One for the case that $a^2 \neq e$ for some $a \in G$ and one for the case that $a^2 = e$ for all $a \in G$.

(1)

	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

(2)

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Order 5: This time we get lots of cases:

Case 1: There exists $e \neq a \neq b$ with $a^2 = e = b^2$.

Case 2 There exists $e \neq a$ with $a^2 \neq e$, $aa^2 = e$ and $(a^2a)^2 = e$.

Case 3 There exists $e \neq a$ with $a^2 \neq e$, $aa^2 = e$ and $(a^2a)^2 \neq e$

Case 4 There exists $e \neq a$ with $a^2 \neq e$, $a^2a = e$ and $(aa^2)^2 = e$.

This Latin square is anti-isomorphic but not isomorphic to the one in case 2. Anti-isomorphic means that is there exists bijection α with $\alpha(ab) = \alpha(b)\alpha(a))$.

Case 5 There exists $e \neq a$ with $a^2 \neq e$, $a^2a = e$ and $(aa^2)^2 \neq e$.

This Latin square is isomorphic and anti-isomorphic to the one in case 3.

Case 6 There exists $e \neq a$ with $a^2 \neq e$, $a^2a = aa^2 \neq e$

Case 7 There exists $e \neq a$ with $a^2 \neq e = (a^2)^2$.

Case 8 There exists $e \neq a$ with $(a^2)^2 \neq e$ and $e \neq a^2a \neq aa^2 \neq e$.

In this case put $c = aa^2$. Then $c^2 \neq e$ and either $cc^2 = e$ or $c^2c = e$. Moreover $(c^2c)^2 \neq e$ respectively $(cc^2)^2 \neq e$ and the latin square is isomorphic to the one in Case 3.

(1)

	e	a	b	c	d
e	e	a	b	c	d
a	a	e	c	d	b
b	b	d	e	a	c
c	c	b	d	e	a
d	d	c	a	b	e

(2)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	e	d	c
b	b	c	d	a	e
c	c	d	a	e	b
d	d	e	c	b	a

(3)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	e	d	c
b	b	c	d	e	a
c	c	d	a	b	e
d	d	e	c	a	b

(4)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	e	d	a	c
c	c	d	a	e	b
d	d	c	e	b	a

(5)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	e	d	a	c
c	c	d	e	b	a
d	d	c	a	e	b

(6)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	c	d	e	a
c	c	d	e	a	b
d	d	e	a	b	c

(7)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	d	e	a	c
c	c	e	d	b	a
d	d	c	a	e	b

(8)

	e	a	b	c	d
e	e	a	b	c	d
a	a	b	c	d	e
b	b	d	a	e	c
c	c	e	d	x	y
d	d	c	e	y	x

$\{x, y\} = \{a, b\}$

2.2 Semigroups, monoids and groups

Definition 2.2.1. *Let G be a magma.*

(a) The binary operation on G is called associative *if*

$$(ab)c = a(bc)$$

for all $a, b, c \in G$. If this is the case we call G a semigroup.

(b) G is a monoid *if it is a semigroup and has an identity.*

(c) Suppose that G is a monoid. Then $a \in G$ is called invertible *if there exists $a^{-1} \in G$ with*

$$aa^{-1} = e = a^{-1}a.$$

Such an a^{-1} is called an inverse *of a.*

(d) A group *is a monoid in which every element is invertible.*

(e) G is called abelian *(or* commutative*) if*

$$ab = ba$$

for all $a, b \in G$.

Example 2.2.2. Let $\mathbb{Z}^+$ denote the positive integers and $\mathbb{N}$ the non-negative integers. Then $(\mathbb{Z}^+, +)$ is a semigroup, $(\mathbb{N}, +)$ is a monoid and $(\mathbb{Z}, +)$ is a group. $(\mathbb{Z}, \cdot)$ and $(\mathbb{R}, \cdot)$ are monoids. Let $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$. Then $(\mathbb{R}^*, \cdot)$ is a group. The integers modulo n under addition is another example. We denote this group by $(\mathbb{Z}/n\mathbb{Z}, +)$. All the examples so far have been abelian.

Note that in a group $a^{-1}b$ is the unique solution of $ax = b$ and ba^{-1} is the unique solution of $ya = b$. So every group is a Latin square with identity. But the converse is not true. Indeed of the Latin squares listed in section 2.1 all the once of order less than five are groups. But of Latin squares of order five only the one labeled (6) is a group.

Let $\mathbb{K}$ be a field and V a vector space over $\mathbb{K}$. Let $\mathrm{End}_{\mathbb{K}}(V)$ the set of all $\mathbb{K}$-linear maps from V to V. Then $\mathrm{End}_{\mathbb{K}}(V)$ is a monoid under compositions. Let $\mathrm{GL}_{\mathbb{K}}(V)$ be the set of $\mathbb{K}$-linear bijection from V to V. Then $\mathrm{GL}_{\mathbb{K}}(V)$ is a group under composition, called the general linear group of V. It is easy to verify that $GL_{\mathbb{K}}(V)$ is not abelian unless V has dimension 0 or 1.

Let I be a set. Then the set $\mathrm{Sym}(I)$ of all bijection from I to I is a group under composition, called the symmetric group on I. If $I = \{1, \ldots, n\}$ we also write $\mathrm{Sym}(n)$ for $\mathrm{Sym}(I)$. $\mathrm{Sym}(n)$ is called the symmetric group of degree n. $\mathrm{Sym}(I)$ is not abelian as long as I has at least three elements.

Above we obtained various examples of groups by starting with a monoid and then considered only the invertible elements. This works in general:

Lemma 2.2.3. *Let G be a monoid.*

(a) Suppose that $a, b, c \in G$, a is a left inverse of b and c is right inverse of b. Then $a = c$ and a is an inverse.

(b) An element in G has an inverse if and only if it has a left inverse and a right inverse.

(c) Each element in G has at most one inverse.

(d) If x and y are invertible, then x^{-1} and xy are invertible. Namely x is an inverse of x^{-1} and $y^{-1}x^{-1}$ is an inverse of xy.

(e) Let G^* be the set of invertible elements in G, then G^* is a group.

Proof. (a)

$$a = ae = a(bc) = (ab)c = ec = c$$

(b) and (c) follow immediately from (a).
(d) Clearly x is an inverse of x^{-1}. Also

$$(y^{-1}x^{-1})(xy) = y^{-1}(x^{-1}(xy)) = y^{-1}((x^{-1}x)y) = y^{-1}(ey) = y^{-1}y = e$$

Similarly $(xy)(y^{-1}x^{-1}) = e$ and so $y^{-1}x^{-1}$ is indeed an inverse for xy.
(e) By (d) we can restrict the binary operation $G \times G \to G$ to get a binary operation $G^* \times G^* \to G^*$. Clearly this is associative. Also $e \in G^*$ so G^* is a monoid. By (d) $x^{-1} \in G^*$ for all $x \in G^*$ and so G^* is a group. □

Corollary 2.2.4. *Let G be a group. Then G is isomorphic to its opposite group G^{op}, in fact the map $x \to x^{-1}$ is an anti-automorphism of G and an isomorphism $G \to G^{\mathrm{op}}$.*

Proof. This follows from 2.2.3(d). □

2.2.5 (Products of elements)**.** The associative law says that $(ab)c = (ab)c$ for all a, b, c in a semigroup. Hence also

$$(a(bc))d = ((ab)c)d = (ab)(cd) = a(b(cd)) = a((bc)d)$$

for all a, b, c, d in G. That is for building products of four elements in a given order it does not matter how we place the parenthesis. We will show that this is true for products of arbitrary length. The tough part is to define what we really mean with a product of $(a_1, \ldots, a_n)$ where $a_i \in G$ for some magma G. We do this by induction on n.

For $n = 1$, a_1 is the only product of (a_1).

For $n \geq 2$, z is a product of $(a_1, \ldots, a_n)$ if and only if $z = xy$, where x is a product of $(a_1, \ldots a_m)$ and y is a product of $(a_{m+1} \ldots a_n)$, for some $1 < m < n$.

The only product of (a_1, a_2) is a_1a_2. The products of (a_1, a_2, a_3) are $(a_1a_2)a_3$ and $a_1(a_2a_3)$. Associativity now just says that every 3-tuple has a unique product.

For later use, if G has an identity we define e to be the only product of the empty tuple.

For the proof of next theorem we also define the standard product of $(a_1, \ldots a_n)$. For $n = 1$ this is a_1 while for $n \geq 2$ it is xa_n where x is the standard product of $(a_1, \ldots, a_{n-1})$.

Theorem 2.2.6 (General Associativity Law)**.** *Let G be a semigroup. Then any (non-empty) tuple of elements of G has a unique product.*

Proof. Let $(a_1, a_2, \ldots, a_n)$ be a tuple of elements of G. We will show by induction on n, that any of its products is equal to its standard product. For $n = 1$ this is obvious.

So suppose $n \geq 2$ and that any product of a tuple of length less than n is equal to its standard product. Let z be any product of $(a_1, \ldots, a_n)$. Then by definition of 'product' there exist an integer $1 < m < n$, a product x of $(a_1, \ldots, a_m)$ and a product y of $(a_{m+1}, \ldots, a_n)$ such that $z = xy$.

Suppose first that $m = n - 1$. By induction x is the standard product of $(a_1, \ldots, a_{n-1})$. Also $z = xa_n$ and so by definition z is the standard product of $(a_1, \ldots, a_n)$.

Suppose next that $m < n - 1$. Again by induction y is the standard product of $(a_{m+1}, \ldots, a_n)$ and so $y = sa_n$, where s is the standard product of $(a_{m+1} \ldots, a_{n-1})$. Hence

$$z = xy = x(sa_n) = (xs)a_n$$

As xs is a product of $(a_1, \ldots a_{n-1})$, we are done by the $m = n - 1$ case. □

One of the most common ways to define a group is as the group of automorphism of some object. For example above we used sets and vector spaces to define the symmetric groups and the general linear group.

If the object is a magma G we get a group which we denote by $\mathrm{Aut}(G)$. So $\mathrm{Aut}(G)$ is the set of all automorphisms of the magma G. The binary operation on $\mathrm{Aut}(G)$ is the composition.

We will determine the automorphism for the Latin squares in 2.1. As the identity element is unique it is fixed by any automorphism. It follows that the Latin square of order

1 or 2, have no non-trivial automorphism (any structure as the trivial automorphism which sends every element to itself).

The Latin square of order three has one non-trivial automorphism. It sends

$$e \to e \quad a \to b \quad b \to a.$$

Consider the first Latin square of order 4. It has two elements with $x \not= 2e$, namely a and c. So again we have a unique non- trivial automorphism:

$$e \to e \quad a \to c \quad b \to b \quad c \to a.$$

Consider the second Latin square of order 4. Here is an easy way to describe the multiplication: $ex = x, xx = e$ and $xy = z$ if $\{x, y, z\} = \{a, b, c\}$. It follows that any permutation of $\{e, a, b, c\}$ which fixes e is an automorphism. Hence the group of automorphism is isomorphic to Sym(3),

Consider the Latin square of order 5 labeled (1). The multiplication table was uniquely determine by any pair $x \neq y$ of non-trivial elements with $x^2 = y^2 = e$. But $x^2 = e$ for all x. So every $e \neq x \neq y \neq e$ there exists a unique automorphism with

$$a \to x \quad b \to y$$

Thus the group of automorphisms has order 12. The reader might convince herself that also the set of bijection which are automorphisms or anti-automorphisms form a group. In this case it has order 24. That is any bijection fixing e is an automorphism or anti-automorphism.

Consider the Latins square of order five labeled (2). This multiplication table is uniquely determine by any element with $x^2 \neq e$, $xx^2 = e$ and $(x^2x)^2 = e$. a, b and d have this property and we get two non-trivial automorphism:

$$e \to e, a \to b \quad b \to d, \quad c \to c \quad d \to a \text{ and } e \to e, a \to d \quad b \to a, \quad c \to c \quad d \to b$$

That is any permutation fixing e and c and cyclicly permuting a, b, d is an automorphism. Consider the Latins square of order five labeled (3). This time only a itself has the defining property. It follows that no non-trivial automorphism exists. But it has an anti-isomorphism fixing a, b and d and interchanging a and c.

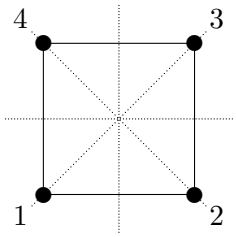
The Latin square (4) and (5) had been (anti-)-isomorphic to (2) and (3). So consider (6). All non-trivial elements have the defining property. So there are 4 automorphisms. They fix e and cyclicly permute (a, b, c, d).

Finally consider the Latin square (7). Here a, c, d have the defining property. So there are 3 automorphism. They fix e and b and cyclicly permuted (a, c, d). Here all bijections fixing a and b are automorphism or anti-automorphism.

It might be interesting to look back and consider the isomorphism types of the groups we found as automorphism of Latin squares. $\mathbb{Z}/n\mathbb{Z}$ for $n = 1, 2, 3, 4$, Sym(3) and a group of order 12. We will later see that Sym(4) has a unique subgroup of order 12 called Alt(4). So the group of order 12 must be isomorphic to Alt(4).

Another class of objects one can use are graphs. We define a graph to be a tuple $(\Gamma, -)$, where Γ is a set and " $-$ " is an anti-reflexive, symmetric relation on Γ. The elements are called vertices, If a and b are vertices with $a - b$ we say that a and b are adjacent. An edge is a pair of adjacent vertices. An automorphism of the graph Γ is an bijection $\alpha \in \mathrm{Sym}(\Gamma)$ such that $a - b$ if and only if $\alpha(a) - \alpha(b)$. In other words a bijection which maps edges to edges. $\mathrm{Aut}(\Gamma)$ is the set of all automorphisms of Γ under composition.

As an example let Γ_4 be a square:



The square has the following automorphisms: rotations by $0, 90, 180$ and 270 degrees, and reflections on each of the four dotted lines. So $\mathrm{Aut}(\Gamma_4)$ has order 8.

To describe $\mathrm{Aut}(\Gamma_4)$ as a subset of $\mathrm{Sym}(4)$ we introduce the cycle notation for elements of $\mathrm{Sym}(I)$ for a finite set I. We say that $\pi \in \mathrm{Sym}(I)$ is a cycle of length if the exists $a_1 \dots a_m \in I$ such that

$$\pi(a_1) = a_2, \pi(a_2) = a_3, \dots, \pi(a_{m-1}) = a_m, \pi(a_m) = a_1$$

and $\pi(j) = j$ for all other $j \in I$.

Such a cycle will be denoted by

$$(a_1 a_2 a_3 \dots a_m)$$

The set $\{a_1, \dots a_m\}$ is called the support of the cycle. Two cycles are called disjoint if their supports are disjoint.

It is clear that every permutations can be uniquely written as a product of disjoint cycle.

$$\pi = (a_1^1 a_2^1 \dots a_{m_1}^1)\,(a_1^2 a_2^2 \dots a_{m_2}^2)\,\dots\,(a_1^k a_2^k \dots a_{m_k}^k)$$

One should notice here that disjoint cycles commute and so the order of multiplication is irrelevant. Often we will not list the cycles of length 1.

So (135)(26) is the permutation which sends 1 to 3, 3 to 5,5 to 1, 2 to 6, 6 to 2 and fixes 4 and any number larger than 6.

With this notation we can explicitly list the elements of $\mathrm{Aut}(\Gamma_4)$:

The four rotations: $e, (1234), (13)(24), (1432)$

And the four reflections: $(14)(23), (13), (12)(34), (24)$.

2.3 The projective plane of order 2

In this section we will look at the automorphism group of the projective plane of order two.

To define a projective plane consider a 3-tuple $\mathcal{E} = (\mathcal{P}, \mathcal{L}, \mathcal{R})$ where $\mathcal{P}$ and $\mathcal{L}$ are non-empty disjoint sets and $\mathcal{R} \subseteq \mathcal{P} \times \mathcal{L}$. The elements of $\mathcal{P}$ are called points, the elements of $\mathcal{L}$ are called lines and we say a point P and a line l are incident if $(P, l) \in \mathcal{R}$. $\mathcal{E}$ is called a projective plane if it has the following three properties

(PP0) Any point is incident with at least 3 lines and any line is incident with at least three points.

(PP1) Any two distinct points are incident with a unique common line.

(PP2) Any two distinct lines are incident with a unique common point.

Note here that the definition of a projective plane is 'symmetric' in points and lines. To be more precise, define $\mathcal{R}^* = \{(l, P) \mid (P, l) \in \mathcal{R}$ and $\mathcal{E}^* = (\mathcal{L}, \mathcal{P}, \mathcal{R}^*)$. If $\mathcal{E}$ is a projective plane, then also $\mathcal{E}^*$ is a projective plane. The points of $\mathcal{E}^*$ are the lines of $\mathcal{E}$ and vice versa. $\mathcal{E}^*$ is called the dual plane of $\mathcal{E}$.

We say that a projective plane has order two if every point is incident with exactly three lines and every line is incident with exactly three points.

Let $\mathcal{E} = (\mathcal{P}, \mathcal{L}, \mathcal{R}$ be a projective plane of order two. Let P be any point. Then any other point lies on exactly one of the three lines through P. Each of whose three lines has 2 points besides P and so we have $1 + 3 \cdot 2 = 7$ points. Note that also $\mathcal{E}^*$ is a projective plane of order 2. So $\mathcal{E}^*$ has seven points, i.e $\mathcal{E}$ has seven lines.

A set of points is called collinear if the points in the set are incident with a common line.

Now let A, B, C be any three points which are not collinear. We will show that the whole projective plane can be uniquely described in terms of the tuple (A, B, C). Given two distinct points P and Q, let PQ be the line incident to P and Q. Also let $P + Q$ be the unique point on PQ distinct from P and Q. Since two distinct lines have exactly on point in common, $A, B, C, A + B, A + C, B + C$ are pairwise distinct. Also the two lines $A(B + C)$ and $B(A + C)$ have a point D in common. It is easy to check that D is not one of the six points we already found. Hence

$$D = A + (B + C) = B + (A + C).$$

Similary, D is the point incident with $B(A + C)$ and $C(A + B)$ and so

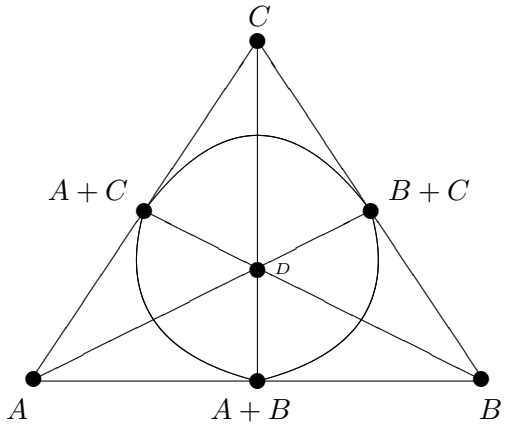
$$D = B + (A + C) = C + (A + B).$$

We have found six pairwise distinct lines:

$$AB, AC, BC, A(B + C), B(A + C), C(A + B).$$

Now $(A + B)(A + C)$ must intersect BC. But B does not lie on $(A + B)(A + C)$ since otherwise $(A + B)(A + C) = (A + B)B = AB$. Similarly C is not on $(A + B)(A + C)$. So

$B+C$ lies on $(A+B)(A+C)$ and $B+C=(A+B)+(A+C)$ So the seventh line is incident with $A+B, A+C$ and $B+C$. So we completely determined the projective plane:



An automorphism of $\mathcal{E}$ is a bijection $\alpha : \mathcal{P} \cup \mathcal{L} \to \mathcal{P} \cup \mathcal{L}$ such that

(i) If P is a point, then $\alpha(P)$ is point.

(ii) If l is a line, then $\alpha(l)$ is a line.

(iii) Let P be a point and l a line. Then P is incident to l if and only if $\alpha(P)$ is incident to$\alpha(l)$.

Note that an automorphism α of $\mathcal{E}$ is uniquely determined by its effect on the points. Namely, if $l = PQ$ is a line, then $\alpha(l)$ is incident with $\alpha(P)$ and $\alpha(Q)$. So $\alpha(l) = \alpha(P)\alpha(Q)$.

Let $\mathrm{Aut}(\mathcal{E})$ be the set of automorphisms of $\mathcal{E}$. If $\alpha, \beta \in \mathrm{Aut}(\mathcal{E})$, then it is easy to see that also $\alpha \circ \beta$ and α^{-1} are also automorphism of $\mathcal{E}$. Moreover, $\mathrm{id}_{\mathcal{P}\cup\mathcal{L}} \in \mathrm{Aut}(\mathcal{E})$ and composition of function is associative. Hence $(\mathrm{Aut}(\mathcal{E}), \circ)$ is a group. We now state an important property of the projective plane of order 2:

Lemma 2.3.1. *Let $\mathcal{E}$ be a projective plane of order two and (A, B, C) and $\tilde{A}, \tilde{B}, \tilde{C}$ be triples of non-collinear points. Then there exists a unique automorphism α of $\mathcal{E}$ with*

$$\alpha(A) = \tilde{A},\ \alpha(B) = \tilde{B} \text{ and } \alpha(C) = \tilde{C}$$

Proof. This follows fairly easily from the fact that the projective plane can be uniquely described in terms of any triple of non-collinear points. We leave the details to the dedicated reader. □

The preceeding lemma shows that $|\mathrm{Aut}(\mathcal{E})|$ is equal to the number of triples $(\tilde{A}, \tilde{B}, \tilde{C})$ of non-collinear points. Now $\tilde{A}$ can be any one of the seven points, $\tilde{B}$ is any of the six

points different from $\tilde{A}$ and $\tilde{C}$ is any of the four points not incident to $\tilde{A}\tilde{B}$. So there are $7 \cdot 6 \cdot 4 = 168$ triples of non-collinear points. Thus

$$|\mathrm{Aut}(\mathcal{E})| = 7 \cdot 6 \cdot 4 = 168.$$

We finish this section with a look at the operation $+$ we have introduced on the points. Let $G = \{e\} \cup \mathcal{P}$. Here e is an arbitrary element not in $\mathcal{P}$. Define a binary operation on G as follows:

$e + g = g + e$, $P + P = e$ and for distinct points P and Q, $P + Q$ is as above.

It is easy to check that G is a group. Also the points correspond to the subgroup of order 2 in G and the lines to the subgroups of order 4. In particular there is an obvious isomorphism between $\mathrm{Aut}(\mathcal{E})$ and $\mathrm{Aut}(G)$.

2.4 Subgroups, cosets and counting

Definition 2.4.1. *Let $(G, *)$ and $(H, \cdot)$ be groups. Then $(H, \cdot)$ is called a* subgroup *of $(G, *)$ provided that:*

(i) $H \subseteq G$.

(ii) $a * b = a \cdot b$ *for all* $a, b \in H$.

Lemma 2.4.2. *Let $(G, *)$ be a group and $(H, \cdot)$ a subgroup of $(G, *)$. Then*

(a) $e_H = e_G$ *where e_H is the identity of H with respect to $\cdot$ and e_G is the identity of G with respect to $*$. In particular, $e_G \in H$.*

(b) $a * b \in H$ *for all* $a, b \in H$.

(c) Let $a \in H$. Then the inverse of a in H with respect to $\cdot$ is the same as the inverse of a in G with respect to $*$. In particular, $a^{-1} \in H$.

Proof. (a)

$$e_H * e_H = e_H \cdot e_H = e_H = e_H * e_G$$

Multiplying with the inverse of e_H in G from the left gives that $e_H = e_G$.

(b) Let $a, b \in H$. Then by definition of a subgroup $a * b = a \cdot b$ and so $a * b \in H$.

(c) Let b be the inverse of a in H with respect to $\cdot$ and c the inverse of a in G with respect to $*$. Then

$$a * b = a \cdot b = e_H = e_G = a * c$$

Multiplying with the inverse of a in G from the left gives $b = c$. □

Lemma 2.4.3. *Let $(G, *)$ be a group and $H \subseteq G$. Suppose that*

(i) $e \in H$.

(ii) H is closed under multiplication, that is for all $a, b \in H$, $ab \in H$

(iii) H is closed under inverses, that is for all $a \in H$, $a^{-1} \in H$.

Define $\cdot : H \times H \to H, (a, b) \to a * b$. Then $(H, \cdot)$ is a subgroup of $(G, *)$.

Proof. We will first verify that $(H, \cdot)$ is a group.

By (ii), $\cdot$ is a well-defined binary operation.

Let $a, b, c \in H$. Then since $H \subseteq G$, a, b, c are in G. Thus since $*$ is associative,

$$(a \cdot b) \cdot c = (a * b) * c = a * (b * c) = a \cdot (b \cdot c)$$

and so $\cdot$ is associative.

By (i), $e \in H$. Let $h \in H$. Then $e \cdot h = e * h = h$ and similarly $h \cdot e = h$ for all $h \in H$. So e is an identity of H with respect to $\cdot$.

Let $h \in H$. Then by (iii), $h^{-1} \in H$. Thus $h \cdot h^{-1} = h * h^{-1} = e$ and similarly $h^{-1} \cdot h = e$. Thus h^{-1} is an inverse of h with respect to $\cdot$.

So $(H, \cdot)$ is a group. By assumption H is a subset of G and by definition of $\cdot$, $a \cdot b = a * b$ for all $a, b \in H$. So $(H, \cdot)$ is a subgroup of $(G, *)$. □

Let $(G, *)$ be a group and $(H, \cdot)$ a subgroup of G. Slightly abusing notation we will often just say that H is a subgroup of G or that $(H, *)$ is a subgroup of $(G, *)$. We also write $H \leq G$ if H is a subgroup of G.

Note that any subgroup of G is itself a group, where the binary operation is given by restricting the one on G. We leave it as an exercise to the reader to verify that a subset H of G is a subgroup if and only if H is not empty and for all $a, b \in H$, $ab^{-1} \in H$. The following lemma is of crucial importance to the theory of groups.

Lemma 2.4.4. *Let H be a subgroup of G. The relation $\sim_H$ on G defined by*

$$a \sim_H b \quad \text{if and only if } a^{-1}b \in H$$

is an equivalence relation.

Proof. Let $a \in G$. Then $a^{-1}a = e \in H$. So $a \sim_H a$ and $\sim$ is reflexive.

Let $a, b \in G$ with $a \sim_H b$. Then $a^{-1}b \in H$ and so also $b^{-1}a = (a^{-1}b)^{-1} \in H$.

Thus $b \sim_H a$. Hence $\sim_H$ is symmetric.

Suppose next that $a, b, c \in G$ with $a \sim_H b$ and $b \sim_H c$. Then $a^{-1}b \in H$ and $b^{-1}c \in H$ and so also

$$a^{-1}c = (a^{-1}b)(b^{-1}c) \in H$$

Thus $a \sim_H c$ and $\sim_H$ is transitive. □

The reader might have noticed that the reflexivity corresponds to $e \in H$, the symmetry to the closure under inverses and the transitivity to the closure under multiplication. Indeed $\sim_H$ can be defined for any subset of G, and it is an equivalence relation if and only if the subset is a subgroup.

Lemma 2.4.5. *Let $\sim$ be a equivalence relation on the set I. For $a \in I$ put $[i] := \{j \in I \mid i \sim j\}$. $[i]$ is called the* equivalence class *of $\sim$ containing i. Let $I/\sim$ be the set of equivalence classes of I.*

(a) Each $i \in I$ lies in a unique equivalence class of $\sim$, namely $[i]$.

(b) $|I| = \sum_{C \in I/\sim} |C|$.

Proof. (a) Let $a \in I$. Since $\sim$ is reflexive, $a \sim a$. So $a \in [a]$ and a is contained in an equivalence class of I. Now let C be an equivalence class of $\sim$ with $a \in C$. We need to show that $C = [a]$. By definition of an equivalence class, $C = [b]$ for some $b \in I$. Since $a \in C = [b]$ we have $b \sim a$

Let $c \in [a]$. Then $a \sim c$. Since $\sim$ is transitive, $b \sim c$ and so $c \in [b]$. Hence $[a] \subseteq [b]$.

We proved that if $a \in [b]$ then $[a] \subseteq [b]$. Since $b \sim a$ and $\sim$ is symmetric we have $a \sim b$ and $b \in [a]$. Thus $[b] \subseteq [a]$.

Hence $[b] = [a]$ and (a) holds.

(b) follows immediately from (b). □

Definition 2.4.6. *Let H be a subgroup of the group G and $g \in G$. Then*

$$gH := \{gh \mid h \in H\}$$

gH is called the (left) coset *of H in G containing g.*

G/H is the set of cosets of H in G.

$|G/H|$ is called the index *of H in G.*

Proposition 2.4.7. *Let H be a subgroup of G and $g \in G$.*

(a) gH is the equivalence class of $\sim_H$ containing g.

(b) g lies in a unique coset of H in G, namely in gH.

(c) $|gH| = |H|$.

Proof. (a) We have

$$\begin{aligned} a \in gH &\iff a = gh \text{ for some } h \in H \iff g^{-1}a = h \text{ for some } h \in H \\ &\iff g^{-1}a \in H \iff g \sim_H a \iff a \in [g] \end{aligned}$$

So $gH = [g]$.

(b) This follows from (a) and 2.4.5.

(c) Define $f : H \to gH, h \to gh$. Then by definition of gH, f is onto. If $gh = gh'$ for some h, h', then $h = h'$. Hence f is 1-1. This gives (c). □

Theorem 2.4.8 (Lagrange)**.** *Let H be a subgroup of G. Then $|G| = |G/H| \cdot |H|$. In particular if G is finite, the order of H divides the order of G.*

Proof.

$$|G| \overset{2.4.5(b)}{=} \sum_{C \in G/H} |C| \overset{2.4.7(c)}{=} \sum_{C \in G/H} |H| = |G/H| \cdot |H|$$

□

2.4.9 (Cycle Notation)**.** Before working out a concrete example for Lagrange's Theorem we introduce the cycle notation for elements of $\mathrm{Sym}(I)$ for a set I. The cycle notation of π can be computes as follows

Pick an element a in I and for $i \in \mathbb{Z}$ put $a_i = \pi^i(a)$. Let $I_a = \{a_i \mid i \in \mathbb{Z}\}$ and define $\pi_a \in \mathrm{Sym}(n)$ by $\pi_a(b) = \pi(a)$ if $b \in I_a$ and $\pi_a(b) = b$ if $b \notin I_a$. Then π_a is called a cycle of length $|I_a|$ of π.

If $a_i \neq a_j$ for all $i \neq j \in \mathbb{Z}$, then $k := |I_a| = \infty$ and we call

$$(\ldots a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, a_3 \ldots)$$

the cycle notatation of π_a. If $a_i = a_j$ for some $i < j \in \mathbb{Z}$ choose such i, j with $k := j - i - 1$ minimal. Applying π^{-i+n} to both sides of $a_i = a_j$ we get $a_{n+1} = a_{n+k}$. By minimality of k, $a_{n+s} \neq a_{n+r}$ for all $1 \leq s < r \leq k$. Thus $I_a = \{a_{n+1}, \ldots a_{n+k}\}$ and $k = |I_a|$. Any of the tuples

$$(a_{n+1}, a_{n+2}, \ldots a_{n+k})$$

is called a cycle notation of π_a. Note that if $m \in \mathbb{Z}$ and r is the remainder of $m - n$ then divided by n, then

$$(a_{m+1}, a_{m+1}, \ldots, a_{m+k}) = (a_{n+1+r}, a_{n+2+r}, \ldots, a_{n+k}, a_{n+1}, a_{n+2}, \ldots, a_{n+r})$$

So all the cycle notations of π_a can be obtained be cyclic permutation of a given one.

Let $b \in I_a$. Suppose that $I_a \cap I_b \neq \emptyset$. Then $a_n = b_m$ for some n, m and so

$$b_l = \pi^{l-m}(b_m) = \pi^{l-m}(a_n) = a_{n_l - m}$$

hence

$$(\ldots a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, a_3 \ldots) = (\ldots b_{-3}, b_{-2}, b_{-1}, b_0, b_1, b_2, b_3 \ldots)$$

It follows that $I_a = I_b$ and so also $\pi_a = \pi_b$. Moreover, any cycle notation of π_a with respect to a is also a cycle notation for $\pi_a = \pi_b$ with respect to b. Pick a family $d_j, j \in J$ of elements of I such that for each $a \in I$ there exists a unique $j \in J$ with $a \in I_{d_j}$. Let c_j be a cycle notation of π_{c_j}. Then $(c_j)_{j \in J}$ is called a cycle notation for π. If I is finite, we choose J to be $\{1, \ldots, l\}$ and call

$$c_1 c_2 \ldots c_l$$

a cycle notation of π. Each c_i is of the form $(a_{i1}, a_{i2}, \ldots, a_{ik_i})$ and so complete cycle notation of π is

$$(a_{11}, \ldots, a_{1k_1})(a_{21}, \ldots, a_{2k_2}) \ldots (a_{l1}, \ldots, a_{lk_l})$$

Since I is the disjoint union the I_a, for each $b \in I$ there exist unique i, j with $1 \leq i \leq l$, $1 \leq j \leq k_i$ with $b = a_{ij}$.

The cycle notation of π is not unique, but any two only differ by the order in which the cycles are listed and by cyclic permutations of each of its cycles. As long as the underlying set I is known, we will also usually do not bother to list the cycles of length 1.

For example (135)(26) is the permutation which sends 1 to 3, 3 to 5,5 to 1, 2 to 6, 6 to 2, and sends 4 to 4 and also sends any number larger than 6 to itself.

Example 2.4.10. Let $G = \mathrm{Sym}(3)$ and $H = \langle (1,2) \rangle = \{(1), (1,2)\}$. Then

$$(1) \circ H = H = \{(1), (1,2)\}$$
$$(1,2,3) \circ H = \{(1,2,3) \circ (1), (1,2,3) \circ (1,2)\} = \{(1,2,3), (1,3)\}$$
$$(1,3,2) \circ H = \{((1,3,2) \circ (1), (1,3,2) \circ (1,2)\} = \{(1,3,2), (2,3)\}$$

Each element of Sym(3) lies in one of these three cosets, so this must be all the cosets. Hence

$$|G| = 6, |G/H| = 3 \text{ and } |H| = 2$$

So by Lagrange's

$$6 = 3 \cdot 2$$

Definition 2.4.11. *Let G be a group. $\mathcal{P}(G)$ denotes the* power set *of G, that is the set of subsets. For $H, K \subset G$ put*

$$HK = \{hk \mid h \in H, k \in K\}.$$

If K is a subgroup then HK is a union of cosets of K, namely $HK = \bigcup_{h \in H} hK$. We write HK/K for the set of cosets of K in HK. In general if $J \subseteq G$ is a union of cosets of H, J/H denotes the sets of all those cosets.

Lemma 2.4.12. *Let G be a group. Then $\mathcal{P}(G)$ is monoid under the binary operation $(A, B) \to AB$. The identity elements is $\{e\}$.*

Proof. We have

$$(AB)C = \{abc \mid a \in A, b \in B, c \in C\} = A(BC)$$

So

$$\mathcal{P}(G) \times \mathcal{P}(G) \to \mathcal{P}(G), (A, B) \to AB$$

is an associative binary operation. Clearly $\{e\}$ is an identity element. □

Lemma 2.4.13. *Let H and K be subgroups of G.*

(a) The map $\alpha : H/H \cap K \to HK/K, h(H \cap K) \to hK$ is a well defined bijection.

(b) $|HK| = |HK/K| \cdot |K| = |H/H \cap K| \cdot |K|$.

(c) If G is finite then $|HK| = \frac{|H||K|}{|H \cap K|}$

Proof. (a) Since $(H \cap K)K = K$, $hK = h((H \cap K)K) = (h(H \cap K))K$ and so is independent of the choice of $h \in h(H \cap K)$. α is clearly onto. Finally if $hK = jK$ for some $h, j \in H$, then $h^{-1}jK = K$, $h^{-1}j \in K$ and so $h^{-1}j \in H \cap K$ and $h(H \cap K) = j(H \cap K)$. Thus α is 1-1.

(b) $|HK| = \sum_{C \in HK/K} |C| = |HK/K| \cdot |K| \overset{\text{(a)}}{=} |H/H \cap K| \cdot |K|$.

(c) By Lagrange's $|H| = |H/H \cap K| \cdot |H \cap K|$. So if G is finite, $|H/H \cap K| = \frac{|H|}{|H \cap K|}$ and thus (c) follows from (b). □

2.5 Normal subgroups and the isomorphism theorem

Just as we have defined (left) cosets one can define right cosets for a subgroup H of G. The right cosets have the form $Hg = \{hg \mid h \in H\}$. In general a left coset of H is not a right coset as the following example shows:

Example 2.5.1. Let $G = \mathrm{Sym}(3)$ and $H = \{(1), (12)\}$. Then

$$(23) \circ H = \{(23), (132)\} \text{ and } H \circ (23) = \{(23), (123)\}$$

So $(23) \circ H \neq H \circ (23)$.

Note that $gH = Hg$ if and only if $gHg^{-1} = H$. We therefore introduce the follwing notation:

Definition 2.5.2. *Let G be a group.*

For $a, b \in G$ put ${}^a b = aba^{-1}$ and for $I \subseteq G$ put ${}^a I = aIa^{-1} = \{{}^a i \mid i \in I\}$. The map $i_h : G \to G, b \to {}^a b$ is the inner automorphism *of G induced by a. It is also called* conjugation *by a and ${}^a b$ is called a conjugate of a.*

Lemma 2.5.3. *Let $N \leq G$. Then the following statements are equivalent:*

(a) ${}^g N = N$ for all $g \in G$.

(b) $gN = Ng$ for all $g \in G$.

(c) Every left coset is a right coset.

(d) Every left coset is contained in a right coset.

(e) ${}^g N \subseteq N$ for all $g \in G$.

(f) ${}^g n \in N$ *for all* $g \in G$, $n \in N$.

Proof. Suppose (a) holds. Then $gNg^{-1} = N$ for all $g \in G$. Multiplying with g from the right we get $gN = Ng$.

Suppose (b) holds. Then the left cosets gN equals the right coset Ng. so (c) holds.

Clearly (c) implies (d)

Suppose that (d) holds. Let $g \in G$. Then $gN \subseteq Nh$ for some $h \in G$. Since $g \in gN$ we conclude $g \in Nh$. By 2.4.7(b), Ng is the unique right coset of N containing g and so and $Ng = Nh$ Thus $gN \subseteq Ng$. Multiplying with g^{-1} from the right we get $gNg^{-1} \subseteq N$. Thus (e) holds.

Clearly (e) implies (f).

Finally suppose that (f) holds. Then $gNg^{-1} \subseteq N$ for all $g \in G$. This statement applied to g^{-1} in place of g gives $g^{-1}Ng \subseteq N$. Multiplying with g from the left and g^{-1} from the right we obtain $N \subseteq gNg^{-1}$. Hence $N \subseteq {}^g N$ and ${}^g N \subseteq N$. So $N = {}^g N$ and (a) holds. □

Definition 2.5.4. *Let G be a group and $N \leq G$. We say that N is* normal *in G and write $N \trianglelefteq G$ if N fulfills one (and so all) of the equivalent conditions in 2.5.3.*

Example 2.5.5. 1. From 2.5.1 we have $(2,3)\mathrm{Sym}(2) \neq \mathrm{Sym}(2)(2,3)$ and so $\mathrm{Sym}(2)$ is not a normal subgroup of $\mathrm{Sym}(3)$.

2. Let $H = \{(1), (123), (132)\}$. Then H is a subgroup of $\mathrm{Sym}(3)$. By Lagrange's

$$|\mathrm{Sym}(3)/H| = \frac{|\mathrm{Sym}(3)|}{|H|} = \frac{6}{3} = 2$$

Hence H has exactly two cosets in H. One of them is

$$H = \{(1), (123), (132)\}$$

Since each element of $\mathrm{Sym}(3)$ lies in a unique coset of H, the other coset must be

$$\mathrm{Sym}(3) \setminus H = \{(12), (13), (23)\}$$

The same argument shows that H and $\mathrm{Sym}(3) \setminus H$ are the only right cosets of $\mathrm{Sym}(3)$. Thus every coset is a right coset and so H is normal in $\mathrm{Sym}(3)$.

3. Let n be a positive integer, let $\mathrm{GL}_n(\mathbb{R})$ the set of invertible $n \times n$-matrices with coefficients in $\mathbb{R}$ and let $\mathrm{SL}_n(\mathbb{R})$ the set of $n \times n$-matrices with coefficients in $\mathbb{R}$ and determinant 1. Note that $GL_n(\mathbb{R})$ is a group under matrix multiplication and $\mathrm{SL}_n(\mathbb{R})$ is a subgroup of $\mathrm{GL}_n(\mathbb{R})$. $\mathrm{GL}_n(\mathbb{R})$ is called a *general linear group* and $\mathrm{SL}_n(\mathbb{R})$ a *special linear group*. Let $A \in \mathrm{GL}_n(\mathbb{R})$ and $B \in \mathrm{SL}_n(\mathbb{R})$. Then

$$\det(ABA^{-1}) = \det(A)\det(B)\det(A^{-1}) = \det(A)\det(B)\det(A)^{-1} = \det B = 1$$

and so $ABA^{-1} \in \mathrm{SL}_n(\mathbb{R})$. Thus $\mathrm{SL}_n(\mathbb{R})$ is a normal subgroup of $\mathrm{GL}_n(\mathbb{R})$.

We will now start to establish a connection between normal subgroups and homomorphism.

Lemma 2.5.6. *Let $\phi : G \to H$ be a group homomorphism.*

(a) $\phi(e_G) = e_H$.

(b) $\phi(a^{-1}) = \phi(a)^{-1}$.

(c) $\phi({}^g a) = {}^{\phi(g)}\phi(a)$.

(d) If $A \leq G$ then $\phi(A) \leq H$.

(e) If $B \leq H$ then $\phi^{-1}(B) \leq G$.

(f) Put $\ker\phi := \{g \in G \mid \phi(g) = e_H\}$. Then $\ker\phi$ is a normal subgroup of G.

(g) ϕ is 1-1 if and only if $\ker\phi = \{e_G\}$.

(h) If $N \trianglelefteq G$, and ϕ is onto, $\phi(N) \trianglelefteq H$.

(i) If $M \trianglelefteq H$, $\phi^{-1}(M) \trianglelefteq G$.

Proof. Except for (c), (f) and (g) this is Exercise 2 on Homework 2.

(c) $\phi({}^g a) = \phi(gag^{-1}) = \phi(g)\phi(a)\phi(g^{-1}) \overset{\text{(b)}}{=} \phi(g)\phi(a)\phi(g)^{-1} = {}^{\phi(g)}\phi(a)$.

(f) This follows from (i) applied to the normal subgroup $M = \{e_H\}$ of H.

(g) Suppose first that ϕ is 1-1 and let $a \in \ker\phi$. Then

$$\phi(a) = e_H = \phi(e_G)$$

and since ϕ is 1-1, $a = e_G$. So $\ker\phi = \{e_G\}$.

Suppose next that $\ker\phi = \{e_G\}$ and let $a, b \in G$ with $\phi(a) = \phi(b)$. Then

$$\phi(a^{-1}b) = \phi(a)^{-1}\phi(b) = \phi(a)^{-1}\phi(a) = e_H.$$

Hence $a^{-1}b \in \ker\phi = \{e_G\}$, $a^{-1}b = e_G$ and so $a = b$. Thus ϕ is 1-1.

□

For $H \subseteq G$ define $H^{-1} = \{h^{-1} \mid h \in H\}$.

Lemma 2.5.7. *Let G be a group and $N \trianglelefteq G$. Let $T, S \in G/N$ and $a, b \in G$ with $T = aN$ and $S = bN$.*

(a) $TS \in G/N$, namely $(aN)(bN) = (ab)N$.

(b) $T^{-1} \in G/N$, namely $(aN)^{-1} = a^{-1}N$.

(c) $TN = T = NT$.

(d) $TT^{-1} = N = T^{-1}T$.

(e) G/N *is a group under the binary operation* $G/N \times G/N \to G/N, (T, S) \to TS$.

(f) The map $\pi_N : G \to G/N, \quad g \to gN$ *is an onto homomorphism with kernel* N.

Proof. (a) $(aN)(bN) = a(Nb)N = a(bN)N = abNN = abN$.

(b) $(aN)^{-1} = N^{-1}a^{-1} = Na^{-1} = a^{-1}N$.

(c) We have $N = eN$ and so by (a) $TN = (aN)(eN) = (ae)N = aN = T$. Similarly $NT = T$.

(d) By (a) and (b) $TT^{-1} = (aN)(a^{-1})N = (aa^{-1})N = eN = N$. Similarly $T^{-1}T = N$.

(f) By (a) the map $G/N \times G/N \to G/N, (T, S) \to TS$ is a well-defined binary operation on G/N. By 2.4.12 multiplication of subsets is associative. By (c) N is an identity element and by (f), T^{-1} is an inverse of T. Thus (e) holds.

(f) We have

$$\pi_N(ab) = abN = (aN)(bN) = \pi_N(a)\pi_N(b)$$

So π_N is a homomorphism. Clearly π_N is onto. We have

$$\ker \pi_N = \{a \in G \mid \pi_N(a) = e_{G/N}\} = \{a \in G \mid aN = N\} = \{a \in G \mid a \in N\} = N$$

□

Theorem 2.5.8 (The Isomorphism Theorem)**.** *Let* $\phi : G \to H$ *be a homomorphism of groups. The map*

$$\bar{\phi} : G/\ker\phi \to \phi(H), \quad g\ker\phi \to \phi(g)$$

is a well-defined isomorphism. Moreover, $\phi = \overline{\phi} \circ \pi_{\ker\phi}$.

Proof. Let $g, k \in G$. Then

$$\begin{aligned} & \phi(g) = \phi(k) \\ \iff \quad & \phi(g)^{-1}\phi(k) = e_H \\ (*) \qquad \iff \quad & \phi(g^{-1}k) = e_H \\ \iff \quad & g^{-1}k \in \ker\phi \\ \iff \quad & g\ker\phi = k\ker\phi \end{aligned}$$

If $g\ker\phi = k\ker\phi$ we get from (*) that $\phi(g) = \phi(k)$ and so $\overline{\phi}$ is well-defined.

If $\overline{\phi}(g\ker\phi) = \overline{\phi}(k\ker\phi)$, we get $\phi(g) = \phi(k)$ and so by (*) $g\ker\phi = k\ker\phi$. Thus $\overline{\phi}$ is 1-1.

Clearly $\bar{\phi}$ is onto.

We have

$$\overline{\phi}((a\ker\phi)(b\ker\phi)) = \overline{\phi}(ab\ker\phi) = \phi(ab) = \phi(a)\phi(b) = \overline{\phi}(a\ker\phi)\overline{\phi}(b\ker\phi)$$

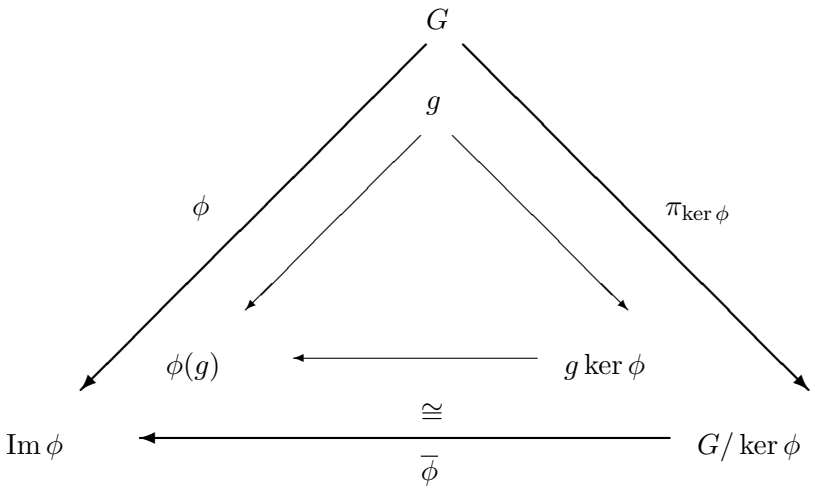
and so $\overline{\phi}$ is a homomorphism.

Also

$$(\overline{\phi}\circ\pi_{\ker\phi})(a) = \overline{\phi}(\pi_{\ker\phi}(a)) = \overline{\phi}(a\ker\phi) = \phi(a)$$

and so $\phi = \overline{\phi}\circ\pi_{\ker\phi}$ □

The Isomomorphism Theorem can be summarized in the following diagram:



Example 2.5.9. Define det : $GL_n(\mathbb{R}) \to (\mathbb{R}\setminus\{0\},\cdot), A \to \det(A)$. Since $\det(AB) = \det(A)\det(B)$, det is a homomorphism. It is easy to see that det is onto. Also $\ker\det = \mathrm{SL}_n(\mathbb{R})$. So 2.5.6(f) gives a new proof that $\mathrm{SL}_n(\mathbb{R})$ is a normal subgroup of $\mathrm{GL}_n(\mathbb{R})$. Moreover the Isomorphism Theorem implies

$$\mathrm{GL}_n(\mathbb{R})/\mathrm{SL}_n(R) \cong (\mathbb{R}\setminus\{0\},\cdot)$$

2.6 Generation of subgroups and cyclic groups

Lemma 2.6.1. *Let G be a group and $(G_i, i\in I)$ a family of subgroups (that is I is a set and for each $i\in I$, G_i is a subgroup of G). Then $\bigcap_{i\in I} G_i$ is a subgroup. If all of the G_i are normal in G, so is $\bigcap_{i\in I} G_i$.*

Proof. Since $e\in G_i$ for all i, $e\in\bigcap_{i\in I} G_i$. Let $a,b\in\bigcap_{i\in I} G_i$. Then $ab\in G_i$ and $a^{-1}\in G_i$ for all $i\in I$. Hence $ab\in\bigcap_{i\in I} G_i$ and $a^{-1}\in\bigcap_{i\in I} G_i$. Thus $\bigcap_{i\in I} G_i$ is a subgroup of G.

Suppose in addition that each G_i is normal in G and let $g\in G$ and $a\in\bigcap_{i\in I} G_i$. Then ${}^g a\in G_i$ and so ${}^g a\in\bigcap_{i\in I} G_i$. Thus $\bigcap_{i\in I} G_i$ is normal in G. □

Definition 2.6.2. *Let G be a group and $J \subseteq G$.*

(a) The subgroup $\langle J\rangle$ of G generated *by J is defined by*

$$\langle J\rangle = \bigcap_{J\subseteq H\leq G} H$$

(b) The normal subgroup $\langle {}^G J\rangle$ of G generated by J is defined by

$$\langle {}^G\rangle = \bigcap_{J\subseteq H\trianglelefteq G} H$$

(c) If $(J_i, i \in I)$ is a family of subsets write $\langle J_i \mid i \in I\rangle$ for $\langle \bigcup_{i\in I} J\rangle$.

(d) $J \subseteq G$ is called normal if ${}^g J = J$ for all $g \in G$.

Lemma 2.6.3. *Let I be a subset of G.*

(a) Let $\alpha : G \to H$ be a group homomorphism. Then $\alpha(\langle I\rangle) = \langle \alpha(I)\rangle$.

(b) Let $g \in G$. Then ${}^g\langle I\rangle = \langle {}^g I\rangle$.

(c) If I is normal in G, so is $\langle I\rangle$.

(d) $\langle I\rangle$ consists of all products of elements in $I \cup I^{-1}$.

(e) $\langle {}^G I\rangle = \langle {}^g I \mid g \in G\rangle$ and consists of all products of elements in $\bigcup_{g\in G} {}^g(I \cup I^{-1})$.

Proof. (a) Let $A = \langle I\rangle$ and $B = \langle \alpha(I)\rangle$. As $\alpha(A)$ is a subgroup of H and contains $\alpha(I)$ we have $B \leq \alpha(A)$. Also $\alpha^{-1}(B)$ is a subgroup of G and contains I. Thus $A \leq \alpha^{-1}(B)$ and so $\alpha(A) \leq B$. Hence $B = \alpha(A)$.

(b) Apply (a) to the homomorphism $i_g : G \to G, x \to {}^g x$.

(c) Follows from (b).

(d) Let H be the subset of G consists of all products of elements in $I \cup I^{-1}$, that is all elements of the form $a_1 a_2 \dots a_n$, with $n \geq 0$ and $a_i \in I \cup I^{-1}$ for all $1 \leq i \leq n$. Here if $n = 0$ we define $a_1 \dots a_n$ to be e. Clearly H is contained in any subgroup of G containing I. Thus $H \subseteq \langle I\rangle$. Now it is readily verified that H is also a subgroup containing I and so $\langle I\rangle \leq H$.

(e) Note that $\bigcup_{g\in G} {}^g I$ is a normal subset of G. Hence by (c) $H := \langle {}^g I \mid g \in G\rangle$ is normal subgroup of G. So $\langle {}^G I\rangle \leq H$. If $I \subseteq K \trianglelefteq G$, then ${}^g I \subseteq K$ for all $g \in G$. Thus also $H \leq K$ and so $H \leq \langle {}^G I\rangle$. It is also contained in every normal subgroup containing I and we get $\langle {}^G I\rangle = H$. The second statement now follows from (d). □

Definition 2.6.4. *Let G be a group and $K, H \subseteq G$. Then*

(a) $N_G(H)$= $\{g \in G \mid {}^g H = H\}$. $N_G(H)$ is called the normalizer *of H in G.*

(b) We say that K normalizes *H provided that $K \subseteq N_G(H)$, that is ${}^kH = H$ for all $k \in K$.*

Lemma 2.6.5. *Let G be a group.*

(a) Let A, B be subgroups of G. Then AB is a subgroup of G if and only if $AB = BA$.

(b) Let $H \subseteq G$. Then $N_G(H)$ is a subgroup of G.

(c) If $K, H \leq G$ and $K \leq N_G(H)$, then $\langle K, H\rangle = KH$.

(d) Let $K_i, i \in I$ be a family of subsets of G. If each K_i normalizes H, so does $\langle K_i \mid i \in I\rangle$.

Proof. (a) If AB is a subgroup of G, then

$$AB = (AB)^{-1} = B^{-1}A^{-1} = BA$$

Conversely suppose that $AB = BA$. The above equation shows that AB is closed under inverses. Also $e = ee \in AB$ and

$$(AB)(AB) = A(BA)B = A(AB)B = A^2B^2 = AB$$

So AB is closed under multiplication.

(b) Readily verified.

(c) Let $k \in K$. Then ${}^kH = H$, $kHk^{-1} = H$, $kH = Hk$ and so $HK = KH$. So by (a) HK is a subgroup of G. Hence $\langle H, K\rangle \leq HK \leq \langle H, K\rangle$ and (c) holds.

(d) Note that $K_i \subseteq N_G(H)$ for all $i \in I$. Since $N_G(H)$ is subgroup of G this implies $\langle K_i \mid i \in I\rangle \leq N_G(H)$ and (d) holds. □

Definition 2.6.6. *Let G be a group and $a, b \in G$ and $A, B \subseteq G$. Then*

$$[a, b] := aba^{-1}b^{-1}$$

and for $A, B \subseteq G$ define

$$[A, B] = \langle [a, b] \mid a \in A, b \in B\rangle.$$

$[a, b]$ is called the commutator *of a and b and $[A, B]$ is called the* commutator group *of A and B.*

Lemma 2.6.7. *Let G be a group and $a, b \in G$.*

(a) $[a, b] = e$ if and only if $ab = ba$.

(b) $[a, b] = {}^ab b^{-1} = a({}^{-b}a)$ where we used the abbreviation ${}^{-b}a = {}^b(a^{-1}) = ({}^ba)^{-1} = ba^{-1}b^{-1}$.

(c) $[a, b]^{-1} = [b, a]$.

(d) $[A, B] = [B, A]$ for any $A, B \subseteq G$.

Proof. (a): $[a,b] = e \iff aba^{-1}b^{-1} = e$. Multipliying with ba from the right the latter equation is equivalent to $ab = ba$.

(b) $[a,b] = (aba^{-1}b^{-1} = {}^{a}bb^{-1}$ and $[a,b] = a(ba^{-1}b^{-1}) = a({}^{-b}a)$.

(c) $[a,b]^{-1} = (aba^{-1}b^{-1})^{-1} = (b^{-1})^{-1}(a^{-1})^{-1}b^{-1}a^{-1} = bab^{-1}a^{-1} = [b,a]$.

Let $a_i \in G$ for $i \in I$. Let H be a subgroup of G. Then $a_i \in H$ if and only if $a_i^{-1} \in H$. Hence

$$\langle a_i \mid i \in I\rangle = \langle a_i^{-1} \mid i \in I\rangle$$

and so

$$[A,B] = \langle [a,b] \mid a \in A, b \in B\rangle = \langle [a,b]^{-1} \mid a \in A, b \in B\rangle \overset{\text{(b)}}{=} \langle [b,a] \mid a \in A, b \in B\rangle = [B,A].$$

□

Lemma 2.6.8. *Let G be a group.*

(a) Let $N \leq G$. Then $N \trianglelefteq G$ if and only if $[G,N] \leq N$.

(b) Let $A, B \trianglelefteq G$. Then $[A,B] \leq A \cap B$.

(c) Let $A, B \trianglelefteq G$ with $A \cap B = \{e\}$. Then $[A,B] = \{e\}$ and $ab = ba$ for all $a \in A, b \in B$.

Proof. (a) ${}^{g}n \in N \iff gng^{-1} \in N \iff gng^{-1}n^{-1} \in N \iff [g,n] \in N$. Thus (b) holds.

(b) By (a) $[A,G] = [G,A] \leq A$ and $[G,B] \leq B$. Thus

$$[A,B] \leq [A,G] \cap [G,B] \leq A \cap B$$

(c) By (b), $[A,B] \leq A \cap B = \{e\}$. Thus for all $a \in A, b \in B$, $[a,b] = e$ and so by 2.6.7(a) we have $ab = ba$. □

Definition 2.6.9. *Let G be a group.*

(a) G is called cyclic *if $G = \langle x\rangle$ for some $x \in G$.*

(b) Let $x \in G$. Then $|x| := |\langle x\rangle|$. $|x|$ is called the order *of x in G.*

We will now determine all cyclic groups up to isomorphism and investigate their subgroups and homomorphisms.

Lemma 2.6.10. *(a) Let H be a subgroup of $(\mathbb{Z},+)$ Then $H = n\mathbb{Z}$ for some $n \in \mathbb{N}$.*

(b) Let $n, m \in \mathbb{N}$. Then $n\mathbb{Z} \leq m\mathbb{Z}$ if and only if m divides n.

Proof. (a) If $H = \{0\}$, then $H = 0\mathbb{Z}$. So we may assume that $H \neq \{0\}$. Since H is a subgroup, $m \in H$ implies $-m \in H$. So H contains some positive integer. Let n be the smallest such. Let $m \in H$ and write $m = rn + s$, $r, s \in \mathbb{Z}$ with $0 \leq s < n$. We claim that $rn \in H$. $rn \in H$ if and only if $-rn \in H$. So we may assume $r > 0$. But then

$$rn = \underbrace{n + n + \ldots + n}_{r-\text{times}}$$

and as $n \in H$, $rn \in H$. So also $s = m - rn \in H$. Since $0 \leq s < n$, the minimal choice of n implies $s = 0$. Thus $m = rn \in n\mathbb{Z}$ and $H = n\mathbb{Z}$.

(b) $n\mathbb{Z} \leq m\mathbb{Z}$ if and only if $n \in m\mathbb{Z}$. So if and only if m divides n. □

Lemma 2.6.11. *Let G be a group and $g \in G$. Then $\phi : \mathbb{Z} \to G, \quad n \to g^n$ is the unique homomorphism from $(\mathbb{Z}, +)$ to G which sends 1 to g.*

Proof. More or less obvious. □

Definition 2.6.12. *For $r \in \mathbb{Z}^+ \cup \{\infty\}$ define* $r^* = \begin{cases} r & \text{if } r < \infty \\ 0 & \text{if } r = \infty \end{cases}$.

This definition is motivated by the following lemma:

Lemma 2.6.13. *Let $n \in \mathbb{N}$. Then $|\mathbb{Z}/n\mathbb{Z}|^* = n$.*

Proof. If $n \neq 0$, then $|\mathbb{Z}/n\mathbb{Z}| = n$ and $n^* = n$. If $n = 0$, then $|\mathbb{Z}/0\mathbb{Z}| = \infty$ and $\infty^* = 0$. □

Lemma 2.6.14. *Let $G = \langle x \rangle$ be a cyclic group and put $n = |G|^*$*

(a) The map

$$\mathbb{Z}/n\mathbb{Z} \to G, \quad m + n\mathbb{Z} \to x^m$$

is a well-defined isomorphism.

(b) Let $H \leq G$ and put $m = |G/H|^*$. Then m divides n, and $H = \langle x^m \rangle$.

Proof. (a) By 2.6.11 the map $\phi : \mathbb{Z} \to G, m \to g^m$ is a homomorphism. As $G = \langle x \rangle$, ϕ is onto. By 2.6.10 $\ker \phi = t\mathbb{Z}$ for some non-negative integer t. By the isomorphism theorem the map

$$\overline{\phi} : \mathbb{Z}/t\mathbb{Z} \to G, m + t\mathbb{Z} \to x^m.$$

is a well defined isomorphism. Hence $\mathbb{Z}/t\mathbb{Z} \cong G$. Thus $t = |\mathbb{Z}/t\mathbb{Z}|^* = |G|^* = n$ and (a) is proved.

(b) By 2.6.10 $\phi^{-1}(H) = s\mathbb{Z}$ for some $s \in \mathbb{N}$. Since $\ker \phi = \phi^{-1}(e) \leq \phi^{-1}(H)$ we have $n\mathbb{Z} \leq s\mathbb{Z}$. Thus 2.6.10 implies that s divides n. As ϕ is onto, $\phi(s\mathbb{Z}) = H$ and so

$$H = \phi(s\mathbb{Z}) = \phi(\langle s \rangle) = \langle \phi(s) \rangle = \langle x^s \rangle$$

It follows that

$$\overline{\phi}(s\mathbb{Z}/n\mathbb{Z}) = \overline{\phi}(\langle s+n\mathbb{Z}\rangle) = \langle\overline{\phi}(s+n\mathbb{Z})\rangle = \langle x^s\rangle = H$$

and since $\overline{\phi}$ is an isomorphism,

$$|G/H| = \left|\mathbb{Z}/n\mathbb{Z}/s\mathbb{Z}/n\mathbb{Z}\right| = |\mathbb{Z}/s\mathbb{Z}|.$$

Thus $s = m$ and (b) is proved. □

Lemma 2.6.15. *Let $G = \langle x\rangle$ be a cyclic group. Let H be any group and $y \in H$. Put $n = |G|^*$ and $m = |y|^*$. Then there exists a homomorphism $G \to H$ with $x \to y$ if and only if m divides n.*

Proof. Exercise. □

2.7 Normal Subgroups of Symmetric Groups

In this section we will investigate the normal subgroups of symmetric group $\mathrm{Sym}(n)$, n a positive integer. We start by defining a particular normal subgroup called the alternating group $\mathrm{Alt}(n)$.

2.7.1 (Alternating Groups)**.** Put

$$e_i = (\delta_{ij})_{j=1}^n \in \mathbb{R}^n.$$

Then $(e_i \mid 1 \le i \le n)$ is a basis of $\mathbb{R}^n$. So for $\pi \in \mathrm{Sym}(n)$ we can define $\alpha(\pi) \in GL_n(\mathbb{R})$ by $\alpha(\pi)(e_i) = e_{\pi(i)}$ for all $1 \le i \le n$. Define $\alpha : \mathrm{Sym}(n) \to \mathrm{GL}_n(\mathbb{R}), \pi \to \alpha(\pi)$. Let $\pi, \mu \in \mathrm{Sym}(n)$ and $1 \le i \le n$. Then

$$\alpha(\mu\circ\pi)(e_i) = e_{\mu(\pi(i))} = \alpha(\mu)(e_{\pi(i)}) = \alpha(\mu)(\alpha(\pi)(e_i)) = (\alpha(\mu)\circ\alpha(\pi))(e_i).$$

So $\alpha(\mu\circ\pi) = \alpha(\mu)\circ\alpha(\pi)$ and α is a homomorphism. Now define $\mathrm{sgn} = \det\circ\alpha : \mathrm{Sym}(n) \to (\mathbb{R}\setminus\{0\}, \cdot), \pi \to \det(\alpha(\pi))$. Since both det and α are homomorphisms, sgn is a homomorphism. Also if $x = (i,j) \in \mathrm{Sym}(n)$ is a 2-cycle it is easy to see that $\det(\alpha(x)) = -1$.

Since

$$(a_1, a_2, \ldots a_k) = (a_1, a_2)(a_2, a_3)\ldots(a_{k-1}, a_k)$$

and sgn is a homomorphism,

$$\mathrm{sgn}((a_1, a_2, \ldots a_k)) = \mathrm{sgn}((a_1, a_2))\mathrm{sgn}((a_2, a_3))\ldots\mathrm{sgn}((a_{k-1}, a_k)) = (-1)^{k-1}$$

Using that sgn is a homomorphism one more time we get

$$\mathrm{sgn}((a_{11}, a_{12}, \ldots a_{1k_1})(a_{21}, a_{22}, \ldots a_{2k_2})\ldots(a_{l1}, a_{l2}, \ldots a_{lk_l})) = (-1)^{k_1-1}(-1)^{k_2-1}\ldots(-1)^{k_l-1}$$

This implies

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \text{ has an even number of even cycles} \\ -1 & \text{if } x \text{ has an odd number of even cycles} \end{cases}$$

An permutation π with $\operatorname{sgn}\pi = 1$ is called an *even permutation* and a permutation with $\operatorname{sgn}(\pi) = -1$ is called an *odd permutation*.

Define $\operatorname{Alt}(n) = \ker \operatorname{sgn}$. Then $\operatorname{Alt}(n)$ is a normal subgroup of $\operatorname{Sym}(n)$, $\operatorname{Alt}(n)$ consists of all permutation which have an even number of even cycles and if $n \geq 2$,

$$\operatorname{Sym}(n)/\operatorname{Alt}(n) \cong \operatorname{sgn}(\operatorname{Sym}(n)) = \{1, -1\} \cong \mathbb{Z}/2\mathbb{Z}.$$

In particular,

$$|\operatorname{Alt}(n)| = \frac{n!}{2}$$

for all $n \geq 2$.

We have $\operatorname{Alt}(2) = \{(1)\}$.

$$\operatorname{Alt}(3) = \{(1), (1,2,3), (1,3,2)\}$$

and

$$\operatorname{Alt}(4) = \{(1), (1,2,3), (1,3,2), (1,2,4), (1,4,2), (1,3,4), (1,4,3), (2,3,4), (2,4,3)$$
$$(1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\}$$

Before continuing to investigate the normal subgroup of $\operatorname{Sym}(n)$ we introduce conjugacy classes in arbitrary groups.

Definition 2.7.2. *We say that two elements x, y in G are* conjugate *in G if $y = {}^g x = gxg^{-1}$ for some $g \in G$. It is an easy exercise to verify that this is an equivalence relation. The equivalence classes are called the* conjugacy classes *of G. The conjugacy class containing x is ${}^G x := \{x^g \mid g \in G\}$.*

Proposition 2.7.3. *A subgroup of G is normal if and only if it is the union of conjugacy classes of G.*

Proof. Let $N \leq G$. The following are clearly equivalent:

$N \trianglelefteq G$

${}^g n \in N$ for all $n \in N, g \in G$

${}^G n \subseteq N$ for all $n \in N$

$N = \bigcup_{n \in N} {}^G n$

N is a union of conjugacy classes

□

To apply this to $\mathrm{Sym}(n)$ we need to determine its conjugacy classes. For this we define

Definition 2.7.4. *Let $\pi \in \mathrm{Sym}(n)$. For $i \in \mathbb{Z}^+$ let λ_i be the number of cycle of length i of π. Then the* cycle type *of π to be sequence $(\lambda_i)_{i=1}^{\infty}$. Alternatively we will write the cycle type as $1^{\lambda_1}2^{\lambda_2}3^{\lambda_3}\dots$ and often will not list terms i^{λ_i} for which $\lambda_i = 0$.*

For example the cycle type of

$$(1,7,3)(2,6)(4)(5,8,10)(9,13,16)(11)(14,15)(16,17)$$

in $\mathrm{Sym}(17)$ is $(2,3,3,0,0,\dots) = 1^2 2^3 3^3$.

Proposition 2.7.5. *(a) Let $\mu, \pi \in \mathrm{Sym}(n)$ and suppose that μ has cycle notation*

$$(a_{11}, a_{12}, \dots, a_{1k_1})(a_{21}, a_{22}, \dots, a_{2k_2}) \dots (a_{l1}, a_{l2}, \dots, a_{lk_l})$$

Then the cycle notation for ${}^{\pi}\mu$ is

$$(\pi(a_{11}), \pi(a_{12}), \dots, \pi(a_{1k_1}))(\pi(a_{21}), \pi(a_{22}), \dots, \pi(a_{2k_2})) \dots (\pi(a_{l1}), \pi(a_{l2}), \dots \pi(a_{lk_l}))$$

(b) Two elements in $\mathrm{Sym}(n)$ are conjugate if and only if they have the same cycle type.

Proof. (a) We have

$${}^{\pi}\mu(\pi(a_{ij})) = \pi(\mu(\pi^{-1}(\pi(a_{ij})) = \pi(\mu(a_{ij})) = \begin{cases} \pi((a_{i,j+1})) & \text{if} j \neq k_i \\ \pi(a_{i,1}) & \text{if} j = k_i \end{cases}$$

So (a) holds.

(b) By (a) μ and ${}^{\pi}\mu$ have the same cycle type. Conversely suppose that μ and σ in $\mathrm{Sym}(n)$ have the same cycle type. Then σ has cycle notation

$$\sigma = (b_{11}, b_{12}, \dots b_{1k_1})(b_{21}, b_{22}, \dots b_{2k_2}) \dots (b_{l1}, b_{l2}, \dots b_{lk_l})$$

Note that for each $1 \leq k \leq n$ there exist unique i, j with $k = a_{i,j}$ and unique s, t with $k = b_{s,t}$. So we can define $\pi \in \mathrm{Sym}(n)$ by $\pi(a_{ij}) = b_{ij}$. Then by (a) ${}^{\pi}\mu = \sigma$ and so elements of the same cycle type are conjugate. □

Example 2.7.6. 1. ${}^{(1,3,5)(2,7)}(1,4,3)(2,6,7)(5,8) = (3,4,5)(7,6,2)(1,8)$

2. Let $\mu = (1,3)(2)(4,7)(5,6,8)$ and $\sigma = (3,5)(8)(1,7)(2,4,6)$

   Define $\pi \in \mathrm{Sym}(8)$ by

$$\pi(1) = 3, \pi(3) = 5, \pi(2) = 8, \pi(4) = 1, \pi(7) = 7, \pi(5) = 2, \pi(6) = 4 \text{ and } \pi(8) = 6$$

   Then ${}^{\pi}\mu = \sigma$.

2.7.7 (Normal sugroups of Sym(3))**.** Lets now investigate the normal subgroups of Sym(3). We start by listing the conjugacy classes

$$\begin{array}{ll} e & 1 \text{ element} \\ (123), (132) & 2 \text{ elements} \\ (12), (13), (23) & 3 \text{ elements} \end{array}$$

Let $e \neq N \trianglelefteq \mathrm{Sym}(3)$. If N contains the 2-cycles, then $|N| \geq 4$. Since $|N|$ divides $|\mathrm{Sym}(3)| = 6$ we get $|N| = 6$ and $N = \mathrm{Sym}(3)$.

If N does not contain the 2-cycles we get $N = \{e, (123), (132)\} = \mathrm{Alt}(3)$.

So the normal subgroups of Sym(3) are

$$(1), \mathrm{Alt}(3), \text{ and } \mathrm{Sym}(3)$$

2.7.8 (Normal subgroups of Sym(4))**.** The conjugacy classes of Sym(4) are:

$$\begin{array}{cl} e & 1 \text{ element} \\ (123), (132), (124), (142), (134), (143), (234), (243) & 8 \text{ elements} \\ (12)(34), (13)(24), (14)(23) & 3 \text{ elements} \\ (12), (13), (14), (23), (24), (34) & 6 \text{ elements} \\ (1234), (1243), (1324), (1342), (1423), (1432) & 6 \text{ elements} \end{array}$$

Let N be a proper normal subgroup of Sym(4). Then $|N|$ divides $24 = |\mathrm{Sym}(4)|$. Thus $|N| = 2, 3, 4, 6, 8$ or 12. So N contains $1, 2, 3, 5, 7$ or 11 non-trivial elements. As $N \setminus \{e\}$ is a union of conjugacy classes, $|N| - 1$ is a sum of some of the numbers $3, 6, 6$ and 8. In particular, $|N| - 1 \geq 3$ and so $|N| - 1 \in \{3, 5, 7, 11\}$. Thus $|N| - 1$ is odd. Since 3 is the only of the possible summands which is odd, we conclude that 3 is one of the summands. So $K \subseteq N$, where $K = \{e, (12)(34), (13)(24), (14)(23)\}$. Then $|N \setminus K| \in \{0, 3, 8\}$ and $|N \setminus K|$ is a sum of some of the numbers $6, 6$ and 8. It follows that $|N \setminus K| = 0$ or 8. In the first case $N = K$ and in the second case, N consist of K and the 3-cycles and so $N = \mathrm{Alt}(4)$. Note also that $(12)(34) \circ (13)(24) = (14)(23)$ and so K is indeed a normal subgroup of Sym(4).

Thus the normal subgroups of Sym(4) are

$$\{(1)\}, \quad \{(1), (12)(34), (13)(24), (14)(23)\}, \quad \mathrm{Alt}(4) \quad \text{and} \quad \mathrm{Sym}(4).$$

Let us determine the quotient group $\mathrm{Sym}(4)/K$. No non-trivial element of K fixes "4". So $\mathrm{Sym}(3) \cap K = \{e\}$ and

$$|\mathrm{Sym}(3)K| = \frac{|\mathrm{Sym}(3)||K|}{|\mathrm{Sym}(3) \cap K|} = \frac{6 \cdot 4}{1} = 24 = |\mathrm{Sym}(4)|.$$

Thus $\mathrm{Sym}(3)K = \mathrm{Sym}(4)$. And

$$Sym(4)/K = \mathrm{Sym}(3)K/K \cong \mathrm{Sym}(3)/(\mathrm{Sym}(3) \cap K) = \mathrm{Sym}(3)/\{e\} \cong \mathrm{Sym}(3)$$

So the quotient of $\mathrm{Sym}(4)$ by K is isomorphic to $\mathrm{Sym}(3)$.

Counting arguments as above can in theory be used to determine the normal subgroups in all the $\mathrm{Sym}(n)$'s, but we prefer to take a different approach.

Lemma 2.7.9. *(a)* $\mathrm{Alt}(n)$ *is the subgroup of* $\mathrm{Sym}(n)$ *generated by all the 3-cycles.*

(b) If $n \geq 5$ *then* $\mathrm{Alt}(n)$ *is the subgroup of* $\mathrm{Sym}(n)$ *generated by all the double 2-cycles.*

(c) Let N *be a normal subgroup of* $\mathrm{Alt}(n)$ *containing a 3-cycle. Then* $N = \mathrm{Alt}(n)$.

(d) Let $n \geq 5$ *and* N *a normal subgroup of* $\mathrm{Alt}(n)$ *containing a double 2-cycle. Then* $N = \mathrm{Alt}(n)$.

Proof. (a) By induction on n. If $n \leq 2$, then $\mathrm{Alt}(n) = \{(1)\}$ and (a) holds . So we may assume $n \geq 3$. Let H be the subgroup of $\mathrm{Sym}(n)$ generated by all the 3-cycles. Then $H \leq \mathrm{Alt}(n)$ and by induction $\mathrm{Alt}(n-1) \leq H$. Let $g \in \mathrm{Alt}(n)$. If $g(n) = n$, $n \in \mathrm{Alt}(n-1) \leq H$. So suppose $g(n) \neq n$. Since $n \geq 3$, there exists $1 \leq a \leq n$ with $a \neq n$ and $a \neq g(n)$. Let h be the 3-cycle $(g(n), n, a)$. Then $(hg)(n) = h(g(n)) = n$. Hence $hg \in Alt(n-1) \leq H$ and so also $g = h^{-1}(hg) \in H$. We proved that $g \in H$ and so $\mathrm{Alt}(n) \leq H$ and $H = \mathrm{Alt}(n)$.

(b) Let $h = (a, b, c)$ be a 3-cycle in $\mathrm{Sym}(n)$. Since $n \geq 5$, there exist $1 \leq d < e \leq n$ distinct from a, b and c. Note that

$$(a, b, c) = (a, b)(d, e) \circ (b, c)(d, e)$$

and so the subgroup generated by the double 2-cycles contains all the 3-cycles. Hence (b) follows from (a).

(c) Let $h = (a, b, c)$ be a 3-cycle in N and g any 3-cycle in $\mathrm{Sym}(n)$. By (a) it suffices to prove that $g \in N$. Since all 3-cycles are conjugate in $\mathrm{Sym}(n)$ there exists $t \in \mathrm{Sym}(n)$ with ${}^t h = g$. If $t \in \mathrm{Alt}(n)$ we get $g = {}^t h \in N$, as N is normal in $\mathrm{Alt}(n)$.

So suppose that $t \notin \mathrm{Alt}(n)$. Then $t(a, b) \in \mathrm{Alt}(n)$. Note that $h^{-1} = (c, b, a) = (b, a, c)$ and so ${}^{(a,b)}(h^{-1}) = {}^{(a,b)}(b, a, c) = (a, b, c) = h$. Thus

$${}^{t(a,b)}(h^{-1}) = {}^t({}^{(a,b)}(h^{-1})) = {}^t h = g$$

As the left hand side is in N we get $g \in N$.

(d) This is very similar to (c) : Let $h = (a, b)(c, d)$ be a double 2-cycle in N and let g be any double 2-cycle in $\mathrm{Sym}(n)$. Then $g = {}^t h$ for some $t \in \mathrm{Sym}(n)$. Note that also $g = {}^{t(a,b)}h$ and either $t \in \mathrm{Alt}(n)$ or $t(a, b) \in \mathrm{Alt}(n)$. Since $N \trianglelefteq \mathrm{Alt}(n)$ we conclude that $g \in N$ and so by (b), $N = \mathrm{Alt}(n)$. □

Definition 2.7.10. *Let G be a group. Then G is called* simple *if $G \neq \{e\}$ and $\{e\}$ and G are the only normal subgroup of G.*

Proposition 2.7.11. *Let $n \geq 5$. Then $\mathrm{Alt}(n)$ is simple.*

Proof. If $n > 5$ we assume by induction that $\mathrm{Alt}(n-1)$ is simple. Let N be a non-trivial normal subgroup of $\mathrm{Alt}(n)$.

Case 1. *N contains an element $g \neq e$ with $g(i) = i$ for some $1 \leq i \leq n$.*

Let $H = \{h \in \mathrm{Alt}(n) \mid h(i) = i\}$. Then $H \cong \mathrm{Alt}(n-1)$, $g \in H \cap N$ and so $H \cap N$ is a non-trivial normal subgroup.

We claim that $H \cap N$ contains a 3-cycle or a double 2-cycle. Indeed if $n = 5$, then $n - 1 = 4$ and the claim holds as every non-trivial element in $\mathrm{Alt}(4)$ is either a 3-cycle or a double 2-cycle. So suppose that $n > 5$. Then by the induction assumption $H \cong \mathrm{Alt}(n-1)$ is simple. Since $H \cap N$ is a non-trivial normal subgroup of H, this implies $H \cap N = H$ and again the claim holds.

By the claim N contains a 3-cyle or a double 2-cycle. So by 2.7.9(c),d we conclude $N = \mathrm{Alt}(n)$.

Case 2. *N contains an element g with a cycle of length at least 3.*

Let $(a, b, c, \ldots)$ be a cycle of g of length at least 3. Let $1 \leq d \leq n$ be distinct from a, b and c. Put $h = {}^{(adc)}g$. Then h has the cycle $(d, b, a, \ldots)$. Also as N is normal in $\mathrm{Alt}(n)$, $h \in N$. So also $hg \in N$.

We compute $(hg)(a) = h(b) = a$ and $(hg)(b) = h(c) \neq h(d) = b$. So $hg \neq (1)$. Hence by (Case 1) (applied to hg in place of g), $N = \mathrm{Alt}(n)$.

Case 3. *N contains an element g with at least two 2-cycles.*

Such a g has the form $(ab)(cd)t$ where t is a product of cycles disjoint from $\{a, b, c, d\}$. Put $h = {}^{(abc)}g$. Then $h = (bc)(ad)t$. Thus

$$gh^{-1} = (ab)(cd)tt^{-1}(bc)(ad) = (ac)(bd).$$

As h and gh^{-1} are in N, (Case 1) (or 2.7.9(d)) shows that $N = \mathrm{Alt}(n)$.

Now let $e \neq g \in N$. As $n \geq 4$, g must fulfill one of the three above cases and so $N = \mathrm{Alt}(n)$. □

Proposition 2.7.12. *Let $N \trianglelefteq \mathrm{Sym}(n)$. Then either $N = \{e\}, \mathrm{Alt}(n)$ or $\mathrm{Sym}(n)$, or $n = 4$ and $N = \{e, (12)(34), (13)(24), (14)(23)\}$.*

Proof. For $n \leq 2$, this is obvious. For $n = 3$ see 2.7.7 and for $n = 4$ see 2.7.8. So suppose $n \geq 5$. Then $N \cap \mathrm{Alt}(n)$ is a normal subgroup of $\mathrm{Alt}(n)$ and so by 2.7.11, $N \cap \mathrm{Alt}(n) = \mathrm{Alt}(n)$ or $\{e\}$.

In the first case $\mathrm{Alt}(n) \leq N \leq \mathrm{Sym}(n)$. Since $|\mathrm{Sym}(n)/\mathrm{Alt}(n)| = 2$, we conclude $N = \mathrm{Alt}(n)$ or $N = \mathrm{Sym}(n)$.

In the second case we get

$$|N| = |N/N \cap \mathrm{Alt}(n)| = |N\mathrm{Alt}(n)/Alt(n)| \leq |\mathrm{Sym}(n)\mathrm{Alt}(n)| \leq 2.$$

Suppose that $|N| = 2$ and let $e \neq n \in N$. As $n^2 = e$, n has a 2-cycle (ab). Let $a \neq c \neq b$ with $1 \leq c \leq n$. The ${}^{(abc)}n$ has cycle (bc) and so $n \neq {}^{(abc)}n$. A contradiction to $N = \{e, n\}$ and $N \trianglelefteq \mathrm{Sym}(n)$. □

Lemma 2.7.13. *The abelian simple groups are exactly cyclic groups of prime order.*

Proof. Let A be an abelian simple group and $e \neq a \in A$. Then $\langle a \rangle \trianglelefteq A$ and so $A = \langle a \rangle$ is cyclic. Hence $A \cong \mathbb{Z}/m\mathbb{Z}$ for some $m \geq 0$. If $m = 0$, $2\mathbb{Z}$ is a normal subgroup. Hence $m > 0$. If m is not a prime we can pick a divisor $1 < k < m$. But then $k\mathbb{Z}/m\mathbb{Z}$ is a proper normal subgroup. □

2.8 Direct products and direct sums

Let G_1 and G_2 be groups. Then $G = G_1 \times G_2$ is a group where the binary operation is given by

$$(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_2 b_2).$$

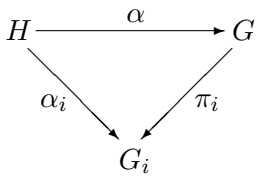
Consider the projection maps

$$\pi_1 : G \to G_1, (a_1, a_2) \to a_1 \quad \text{and} \quad \pi_2 : G \to G_2, (a_1, a_2) \to a_2$$

Note that each $g \in G$ is uniquely determined by its images under π_1 and π_2. Indeed we have $g = (\pi_1(g), \pi_2(g))$. We exploit this fact in the following abstract definition.

Definition 2.8.1. *Let $(G_i, i \in I)$ be a family of groups. A* direct product *of the $(G_i, i \in I)$ is a group G together with a family of homomorphism $(\pi_i : G \to G_i, i \in I)$ such that:*

Whenever H is a group and $(\alpha_i : H \to H_i, i \in I)$ is family of homomorphism, then there exists a unique homomorphism $\alpha : H \to G$ such that the diagram:



commutes for all $i \in I$ (thats is $\alpha_i = \pi_i \circ \alpha$)

Lemma 2.8.2. *Let $(G_i, i \in I)$ be a family of groups then there exists direct product $G, (\pi_i : G \to G_i; i \in I)$. Moreover the direct product is unique up to isomorphism in the follwinh sense: If $(H, (\alpha_i : H \to G_i; i \in I)$ is annother direct product of $(G_i, i \in I)$, then there exists ismorphism $\alpha : H \to G$ with $\alpha_i = \pi_i \circ \alpha$ for all $i \in I$.*

Proof. We will first show the existence. As a set let $G = \times_{i\in} G_i$, the set theoretic direct product of the $G_i's$. That is G consists of all function $f : I \to \bigcup_{i\in I} G_i$ with $f(i) \in G_i$ for all $i \in I$. The binary operation is defined by $(fh)(i) = f(i)h(i)$. It is easy to check that G is a group. Define

$$\pi_i : \prod_{i\in I} G_i \to G_i, f \to f(i).$$

Since $\pi_i(fg) = fg(i) = f(i)g(i) = \pi_i(f)\pi_i(g)$, π_i is a homomorphism.

Let H a group and $\alpha_i : H \to G_i$ a family of homomorphism. Define $\alpha : H \to G$ by

$$\alpha(h)(i) := \alpha_i(h)\text{for all } i \in I$$

But this is equivalent to

$$\pi_i(\alpha(h)) = \alpha_i(h)$$

and so to

$$\pi_i\alpha = \alpha_i.$$

In other words, α is the unique map which makes the above diagram commutative. It easy to verify that α is a homomorphism and so $(\pi_i, i \in I)$ meets the definition of the direct product.

To show uniqueness, let $(\alpha_i : H \to G_i; i \in I)$ be annother direct product of $(G_i, i \in I)$. Since $(\pi_i : G \to G_i, i \in I)$ is a direct product there exists a homomorphism $\alpha : H \to G$ with $\alpha_i = \pi_i \circ \alpha$ for all $i \in I$. Since $(\alpha_i : H \to G_i; i \in I)$ is a direct product there exists $\beta : G \to H$ with $\pi_i = \alpha_i \circ \beta$. Consider the composition $\alpha \circ \beta\alpha : G \to G$. We have

$$\pi_i \circ (\alpha \circ \beta) = (\pi_i \circ \alpha) \circ \beta = \alpha_i \circ \beta = \pi_i$$

also

$$\pi_i = \pi_i \circ \mathrm{id}_G$$

and so by the uniqueness assertion in the defintion of a direct product we conclude that $\alpha \circ \beta = \mathrm{id}_G$. By symmetry also $\beta \circ \alpha = \mathrm{id}_H$. Thus α is an isomorphism. □

Definition 2.8.3. *Let $(G_i \mid i \in I)$ be a family of groups.*

(a) Let $f \in \times_{i\in I}$. Then we view f as the tuple $(f(i))_{i\in I}$. Conversely every tuple $(g_i)_{i\in I}$ with $g_i \in G$ we view as the function $f : I \to \bigcup_{i\in I} G_i, i \to g_i$ in $\times_{i\in I}$.

(b) Let $g = (g_i)_{i\in I} \in \times_{i\in I} G_i$. Then

$$\mathrm{Supp}(g) := \{i \in I \mid g_i = e\}$$

(c) $\bigoplus_{i\in I} G_i := \{g \in \times_{i\in I} G_i \mid \mathrm{Supp}(g)$ is finite $\}$. $\bigoplus_{i\in I} G_i$ is called the direct sum *of $(G_i \mid i \in I)$.*

Definition 2.8.4. *Let G be a group.*

(a) Let $(a_i \mid i \in I)$ be a family of elements in G. We say that almost all of the a_i's equal e if $\{i \in I \mid a_i \neq e\}$ is finite.

(b) Let $(a_i \mid i \in I)$ be a family of elements in G with almost all of the a_i eqaul to one. Suppose also that $a_i a_j = a_j a_i$ for all $i, j \in J$. Then we define

$$\prod_{i \in I} a_i = a_{i_1} a_{i_2} \dots a_{i_k}$$

where $i_1, i_2 \dots i_k$ are the pairwise distinct elemenst of I with $a_{i_j} \neq e$. Note that since $a_i a_j = a_j a_i$, this defintion does not dependent on the order the $i_1, \dots i_k$ are chosen.

Lemma 2.8.5. *let $(G_i \mid i \in I)$ be a family of groups. For $j \in I$ define $\rho_j : G_j \to \bigoplus_{i \in I} G_i$ by*

$$(\rho_j(g)) = (h_i)_{i \in I} \text{ where } h_i = \begin{cases} g & \text{if } i = j \\ e & \text{if } i \neq j \end{cases}$$

(a) $\bigoplus_{i \in I} G_i$ is a subgroup of $\times_{i \in I} G_i$.

(b) For all $j \in I$, ρ_j is a 1-1 homomorphism.

(c) $[\rho_i(G_i), \rho_j(G_j)] = \{e\}$ for all $i \neq j \in I$.

(d) Let $g \in \bigoplus_{i \in I} G_i$. Then there exist uniquely determined $h_i \in G_i, i \in I$, almost all equal to e,with $g = \prod_{i \in I} \rho(h_i)$. Namely $h_i = g_i$.

(e) $\bigoplus_{i \in I} G_i = \langle \rho_j(G_j) \mid j \in J \}$

Proof. (a) This follws since $\mathrm{Supp}(a^{-1}) = \mathrm{Supp}(a)$ and $\mathrm{Supp}(ab) \subseteq \mathrm{Supp}(a) \cup \mathrm{Supp}(b)$.
(b) This is readily verified.
(c) Let $j \neq k \in I$, $g_j \in G_j$ and $g_k \in G_k$. Then

$$(\rho(g_j)\rho(g_k))_i = \begin{cases} g_j & \text{if } i = j \\ g_k & \text{if } i = k \\ e & \text{if } j \neq i \neq k \end{cases} = (\rho(g_k)\rho(g_j))_i$$

Thus $\rho_j(g_j) = \rho_k(g_k)$ and (c) holds.
(d) Just observe that by the definition of $\rho_j(g_j)$

$$(\prod_{i \in I} \rho(h_i))_j = h_j.$$

(e) By (d) $g = \prod_{i \in I} \rho_i(g_i) \in \langle \rho_i(G_i) \mid i \rangle$. Thus (e) holds. □

Definition 2.8.6. *Let G be a group and $(G_i, i \in I)$ a family of subgroups of G.Then we say that G is the* internal direct sum *of $(G_i, i \in I)$ and write*

$$G = \bigoplus_{i\in I}^{\text{int}} G_i$$

provided that

(i) $G_i \trianglelefteq G$ for all $i \in I$.

(ii) $G = \langle G_i \mid i \in I\rangle$.

(iii) For each i, $G_i \cap \langle G_j \mid i \neq j \in I\rangle = \{e\}$.

Proposition 2.8.7. *Let G be a group and $(G_i, i \in I)$ a family of subgroups if G_i. Suppose that G is the internal direct sum of $(G_i, i \in I)$.*

Then the map

$$\alpha : \bigoplus_{i\in I} G_i \to G, (g_i)_{i\in I} \to \prod_{i\in I} g_i$$

is a well-defined isomorphism.

Proof. For $i \in I$ put $G^i := \langle G_j \mid i \neq j \in I\rangle$. Let $g \in G$. Since $G_j \trianglelefteq G$ we have ${}^gG_j = G_j$ and so using 2.6.3(b) we compute

$${}^gG^i = \langle {}^gG_j \mid i \neq j \in I\rangle = \langle G_j \mid i \neq j \in I\rangle = G^i$$

Thus $G^i \trianglelefteq G$. By 2.8.6(iii), $G_i \cap G^i = \{e\}$ and so by 2.6.8(c) $ab = ba$ for all $a \in G_i, b \in G^i$. If $j \neq i \in I$ then $G_j \leq G^i$ and so $g_ig_j = g_jg_i$ for all $g_i \in G_i$ and $g_j \in G_j$. So (see Definition 2.8.4(b)), α is well defined. Moreover this implies that α is a homomorphism. Note that $\alpha(\rho_i(g_i)) = g_i$. So $G_i \leq \operatorname{Im}\alpha$. Since $\operatorname{Im}\alpha$ is a subgroup of G we conclude $\langle G_i \mid i \in I\rangle \leq \operatorname{Im}\alpha$. Hence 2.8.6(ii), $\operatorname{Im}\alpha = G$ and so α is onto.

Suppose that

$$\prod_{i\in I} g_i = \prod_{i\in I} a_i$$

for some $(g_i)_{\in I}, (a_i)_{i\in I} \in \bigoplus_{i\in I}$. Then

$$a_ig_i^{-1} = \prod_{i\neq j\in I} a_j^{-1}g_j$$

Note that the left side is in G_i and the right side in G^i. Since $G_i \cap G^i = \{e\}$ we conclude that $a_ig_i^{-1} = e$ and so $a_i = g_i$. Thus α is 1-1 and the lemma is proved. □

Note that the preceeding lemma implies that if $G = \bigoplus^{\text{int}} G_i$ then G is canonically isomorphic to $\bigoplus_{i\in I} G_i$. For this reason we will slightly abuse language and write $G = \bigoplus_{i\in I} G_i$, that is we drop the superscript int to denote the internal direct sum.

Example 2.8.8. Let $G = \{(1), (1,2)(3,4), (1,3)(2,4), (1,4)(2,3)\} \leq \mathrm{Sym}(4)$. Let $G_1 = \{(1), (1,2)(3,4)\}$ and $G_2 = \{(1), (1,3)(2,4)\}$. Since G is abelian, G_1 and G_2 are normal subgroup of G. Since $(1,2)(3,4) \circ (1,3)(2,4) = (1,4)(2,3)$, $\langle G_1, G_2 \rangle = G$. Moreover, $G_1 \cap G_2 = \{(1)\}$ and so G is the internal direct sum of G_1 and G_2. Note also that $G_i \cong \mathbb{Z}/2\mathbb{Z}$ and so

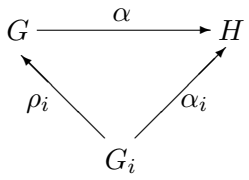
$$G = G_1 \oplus G_2 \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2\mathbb{Z}.$$

2.9 Co-products and free groups

Having looked at the direct product and direct sum of groups we know define the coproduct of a family of groups:

Definition 2.9.1. *Let $(G_i, i \in I)$ be a family of groups. Then an* coproduct *of $(G_i, i \in I)$ is a group G together with a family of homomorphism $(\rho_i : G_i \to G)$ with the following property:*

Whenever H is a group and $(\alpha_i : G_i \to H)$ a family of homomorphisms, then there exists a unique homomorphism $\alpha : G \to G_i \to H$ so that the diagram



commutes for all $i \in I$.

On an intuitive level this group is the largest group which contains the G_i's and is generated by them. Notice also that the defintion of the coproduct is nearly identical to the defintion of the direct product. The difference is that all the arrows are reversed, that is a map fro A to B is replaced by a map from B to A. But it turns out the the coproduct is much harder to construct

Before starting the construction we need a couple of general lemmas.

Lemma 2.9.2. *Let I be a set and $\sim$ a relation on I. Define the relation $\approx$ of I by $a \approx b$ if there exits $n \in \mathbb{N}$ and a sequence of elements $(x_0, x_1, \ldots x_n)$ in I such that $x_0 = a$, $x_n = b$ and for each $1 \leq i \leq n$ either $x_{i-1} \sim x_i$ or $x_i \sim x_{i-1}$. Then*

(a) $\approx$ is an equivalence relation on I, called the equivalence relation *generated by $\sim$.*

(b) If $a, b \in I$ with $a \sim b$ then $a \approx b$.

(c) Let $\approxeq$ be an equivalence relation on I such that $a \sim b$ implies $a \approxeq b$. Then also $a \approx b$ implies $a \approxeq b$.

Proof. Straightforward. □

Lemma 2.9.3. *Let $(G,\cdot)$ be a pregroup, $\sim$ a relation on G and $\approx$ the equivalence relation on G generated by $\sim$. Suppose that if $a,b,c \in G$ with $b \sim c$, then $ab \approx ac$ and $ba \approx ca$. Then*

(a) The map $* : G/\approx \times G/\approx \to G/\approx, \quad ([a],[b]) \to [ab]$ is a welldefined binary operation.

(b) If $\cdot$ is associative, then $*$ is associative.

(c) If e is an identity in G, then $[e]$ is an identity in G.

(d) Suppose G is a monoid and H is a subset of G such

(i) G is generated by H as a monoid.

(ii) For each $h \in H$ there exists $h' \in G$ with $hh' \approx e \approx hh'$.

Then $G/\approx$ is a group.

Proof. (a) Let $a,b,c,d \in G$ with $a \approx c$ and $b \approx d$. We need to show that $ab \approx cd$. By definition of $\approx$ there exist $n,m \in \mathbb{N}$ and $(x_0,\dots x_n)$ and $(y_0,\dots y_m)$ such that $x_0 = a, x_n = c$, $y_0 = b$, $y_m = d$, for all $1 \le i \le n$ either $x_{i-1} \sim x_i$ or $x_i \sim x_{i-1}$ and for all $1 \le j \le m$ either $y_{j-1} \sim x_j$ or $y_j \sim y_{j-1}$.

The proof of $ab \approx cd$ will be by induction on $n+m$. If $n+m=0$, then $n = 0 = m$ and so $a = c$ and $b = d$. So supppose $n+m > 0$.

If $n = 0$ then $m > 0$ and so by induction $ab \approx cy_{m-1}$. We have $y_{m-1} \sim y_m = d$ or $d \sim y_{m-1}$. In the first case by assumption $cy_{m-1} \approx cd$ and in the second $cd \approx cy_{m-1}$. Since $\approx$ is symmetric both cases imply that $cy_{m-1} \approx cd$. As $\approx$ is tranistive and $ab \approx cy_{m-1}$ we get $ab \approx cd$.

So suppose $n > 0$. Then by induction $ab \approx x_{m-1}d$. As above $x_{m-1} \sim c$ implies $x_{m-1}d \approx cd$ and so $ab \approx cd$.

(b) and (c) are obvious.

(d) Let $h \in H$. Then $[hh'] = [e] = [hh']$ and so $[h]$ is invertible in $G/\approx$. Put

$$K := \{g \in G \mid [g] \text{ is invertible in } G/\approx\}$$

Then $H \subseteq K$ and $e \in K$. let $a,b \in K$. Then by 2.2.3(d), $[ab] = [ab]$ is invertible. Hence $ab \in K$ and K is a submonoid of G. Thus (d:i) implies $K = G$. Hence every $[g]$ for $g \in G$ is invertibel. Together wit (b) and (c) we conclude that $G/\approx$ is a group. □

Proposition 2.9.4. *Let I be a set and let M_I be the set of all sequences $(i_1, i_2, \dots, i_n)$ where $n \in \mathbb{N}$ and $i_j \in I$ for all $1 \le j \le n$. For $i = (i_1, i_2, \dots, i_n)$ and $j = (j_1, \dots j_m)$ in M_I define*

$$ij = (i_1, i_2, \dots, i_n, j_1, \dots j_m)$$

Then

(a) M_I is a monoid, called the free monoid *generated by I.*

(b) The map $\rho : I \to M_I, i \to (i)$ is 1-1 and we identify i with (i).

(c) Let M_I be a monoid and $f : I \to M$ a function. Then there exists a unique homorphism of monoids $f^* : M_I \to M$ with $f^*(i) = f(i)$ for all $i \in I$.

Proof. (a) The binary operation is clearly associative and () is an idenity element.

(b) Define $f^*((i_1, \dots, i_n)) = f(i_1)f(i_2)\dots f(i_n))$ there as usually the empty product is defined to be e_M. This is clearly a homomorphism. Conversely if $g : M_I \to M$ is a homomorphism with $g(i) = f(i)$, then

$$g((i_1, i_2, \dots i_n)) = g(i_1 i_2 \dots i_n) = g(i_1)g(i_2)\dots g(i_n) = f(i_1)\dots f(i_n)$$

□

Theorem 2.9.5. *Let $(G_i \mid i \in I)$ be a family of groups. Then there exists a coproduct $\coprod_{I\in I} G_i$ of $(G_i, i \in I)$.*

Proof. To simplify notation we assume without loss that the G_i are pairwise disjoint. Let $X = \bigcup_{i\in I} G_i$ and let W be the free monoid generated by X. Note that if $a, b \in G_i$ we have to ways to multiply a and b, namely once as elements of G_i and once as elements of W. We therefore denote the binary operation on W by $*$ and the one on G_i be $\cdot$. t Define the relation $\sim$ on W by $v \sim w$ if of the follwing holds:

(i) There exist $x, y \in W$, $i \in I$ and $a, b \in G_i$ with $w = x * a * b * y$ and $v = x * (a \cdot b) * y$

(ii) There exists $x, y \in W$ and $i \in I$ with $w = x * e_{G_i} * y$ and $v = x * y$.

Let $\approx$ be the equivalence relation on W generated by $\sim$. Note that the defintion of $\sim$ implies that if $u, c, w \in W$ with $v \sim w$, then also $uv \sim uw$ and $vu \sim uw$. Thus by 2.9.3 $W/\approx$ is a monoid with identity $[()]$.

Let $i \in I$ and $a, b \in G_i$. Then by (i)

$$a * b \sim a \cdot b \text{ and so } [a] * [b] = [a \cdot b] \tag{1}$$

By (ii)

$$e_i \approx () \text{ and so } [e_i] = [()]$$

If follows that $a * a^{-1} \approx a \cdot a^{-1} = e_{G_i} \approx ()$. Thus by 2.9.3(d) $W/\approx$ is a group.

Define $\rho_i : G_i \to W/\approx, g \to [g]$. Then by (1), ρ_i is a homomorphism.

Now let H be a group and $(\alpha_i : G_i \to H)$ a family of homomorphism. Define $\beta : X \to H$ by $\beta(x) = \alpha_i(w)$ if $i \in I$ with $x \in G_i$. Note here that i is uniquely determined since the G_i's are pairwise disjoint. By 2.9.4(c) there exists a unique homomorphism $\gamma : W \to H$ with $\gamma(x) = \beta(x)$ for all $x \in X$ and so $\gamma(a) = \alpha_i(a)$ for all $a \in G_i$.

We claim that $\gamma(v) = \gamma(w)$ whenever $v \approx w$. By definition of $a \approx$ it it suffices to show this if $v \sim w$.

Suppose first that (i) holds. Then $w = x * a * b * y$ and $v = x * (a \cdot b) * y$ for some $x, y \in W$, $i \in I$ and $a, b \in G_i$ Hence

$$\gamma(w) = \gamma(x)\gamma(a)\gamma(b)\gamma(y) = \gamma(x)(\alpha_i(a)\alpha_i(b)\gamma(y) =$$
$$\gamma(x)\alpha_i(a \cdot b)\gamma(y) = \gamma(x)\gamma(a \cdot b)\gamma(y) = \gamma(v).$$

Suppose next that (ii) holds. Then $w = x * e_{G_i} * y$ and $v = x * y$ for some $x, y \in W$ and $i \in I$ Hence

$$\gamma(w) = \gamma(x)\gamma(e_i)\gamma(y) = \gamma(x)\alpha_i(e_{G_i})\gamma(y) = \gamma(x)e_H\gamma(y) = \gamma(x)\gamma(y) = \gamma(v).$$

By the claim we get a well defined map $\alpha : W/\approx \to H, [w] \to \gamma(w)$. Also as γ is a homomorphism, α is, too. The uniqueness of α can be easily deduced from the uniqueness of γ. □

Lemma 2.9.6. *Let $(\rho_i \mid G_i \to G, i \in I)$ be a coproduct of the family of groups $(G_i, i \in I)$. The each $\rho_j, j \in J$ is 1-1.*

Proof. Let $j \in J$. let $\alpha_j = \mathrm{id}_{G_j}$ and for $j \neq i \in j$ define $\alpha_i : G_i \to G_j, g \to e_{G_j}$. Then by defintion of the coproduct there exists a homomorphism $\beta : G \to G_j$ with $\alpha_i = \beta \circ \rho_i$ for all $i \in I$. For $i = j$ we conclude,

$$\mathrm{id}_{G_j} = \beta \circ \rho_j$$

Note that this implies that ρ_j is 1-1. □

2.9.7 (Reduced Words)**.** In this subsection we have a closer look ar the coproduct of a family of groups $G_i, i \in I)$ constructed in the proof of 2.9.5. As where we assume that the G_i's are pairwise disjoint, put $X = \bigcup_{i \in I} G_i$, and $W = M_I$. Let $\sim$ and $\approx$ be the relations on W define defined in 2.9.5. Our goal now is to find canonical representative for each of the equivalence classes of $\approx$.

Let $x \in W$. Then $x = x_1 x_2 \dots x_n$ for some uniquely determined $n \in \mathbb{N}$ and $x_i \in G_{j_i}$ for some $j_i \in I$. n is called the *length* of x We say that x is *reduced* if $x_i \neq e_{G_{j_i}}$ (for all $1 \leq k \leq n$) and $j_i \neq j_{i+1}$ for all $1 \leq i < n$. Comapring with we definition of $\sim$ we see that w is not reduced if and only if there exists $v \in W$ with $v \sim w$. Note that the empty tuple is reduced and any $a \in G_i$ with $a \neq e_{G_i}$ is reduced. We will show that every equivalence class contains a unique reduced word.

For this we consider one further relation on W. Define $v \ll w$ if

$$v = v_0 \sim v_1 \sim v_2 \dots v_{n-1} \sim v_n = w$$

for some $v_k \in W$. Again allow for $n = 0$ and so $v \ll v$. Also note that $v \ll w$ implies $v \approx w$ (but not vice versa).

1°. *For each $w \in W$ there exists reduced word $v \in W$ with $v \ll w$.*

If w is reduced, put $v = w$. If w is not reduced there exists $y \in W$ with $y \sim v$. The length of y is one less than the length of y. So by induction there exists a reduced word v with $v \ll w$. Then also $v \ll w$ and (1°) is proved.

2°. *For each $w \in W$ there exists a unique reduced word $w_r \in W$ with $w_r << w$. w_r is called the reduction of w.*

The existence has been establised in (1°). For the uniqueness let z_1 and z_2 be reduced with $z_i \ll w$.

If $z_1 = w$ then w is reduced. So there does not exists $y \in W$ with $y \sim w$ and so $z_2 \sim w$ implies $z_2 = w = z_1$. So we may assume that $z_1 \neq w \neq z_2$.

By definition of $\ll$ there exist $v_i \in W$ with $z_i \ll v_i \sim w$.

We will show that there exists $v \in W$ with $v \ll v_1$ and $v \ll v_2$. Suppose such an v exists and let z be reduced with $z \ll v$. Then for $i = 1, 2$, $z \ll v_i$. Since also z_i is reduced with $z_i \ll v_i$ and since v_i has length less than w, we conclude by induction that $z = z_i$. Thus $z_1 = z_2$ and we are done once the existence of v is established.

Suppose that for $i = 1$ and 2, $w = x_i e_{G_{k_i}} y_i$ and $v_i = x_i y_i$ for some $k_i \in I$, $x_i, y_i \in W$. $x_1 = x_2$ have $v_1 = v_2$ we can choose $v = v_1$. So may assume that x_1 has length less than x_2. Then $x_2 = x_1 * e_{G_{k_1}} * x$ for some $x \in W$ and we can put $v = x_1 x y_2$.

Suppose next that $w = x_1 e_{G_{k_1}} y_1$, $v_1 = x_1 y_2$, $w = x_2 a b y_2$, $v_2 = x_2(a \cdot b) y_2$ for some $k_1, k_2 \in I$, $x_1, x_2, y_1, y_2 \in W$ and $a, b \in G_{k_1}$. If $x_1 = x_2$ or $x_1 = x_2 a$, then $ab = a$ and $ab = b$ respectively and so $v_1 = v_2$ and $v = v_1$ works. If x_1 has length less than x_2, then $x_2 = x_1 e_{G_{k_1}} x$ for some $x \in W$ and we can choose $v = x_1 x (a \cdot b) y_2$. If x_1 has length larger than $x_2 a$ than $y_2 = x e_{G_{k_2}} y_1$ for some $x \in W$ and we can choose $x_2(a \cdot b) x y_1$.

Suppose finally that for $i = 1, 2$, $w = x_i a_i b_i y_i$, $v_i = x_i(a_i \cdot b_i) y_i$ for some $x_i, y_i \in W$, $a_i, b_i \in G_{k_i}$. Let l_i be the length of x_i. Without loss $l_1 \leq l_2$ and so either $l_1 \leq l_2 - 2$, $l_1 = l_2 - 1$ or $l_1 = l_2$. If $l_1 \leq l_2 - 2$, then $x_2 = x_1 a_1 b_1 x$ for some $x \in W$. Put $v = x_1(a_1 \cdot b_1) x (a_2 b_1) y_2$ in this case. If $l_1 + 1 = l_2$, then $b_1 = a_2$, $k_1 = h_2$ and we can put $v = x_1(a_1 b_1 b_2) y_2$. If $l_1 = l_2$, then $v_1 = v_2$ and we put $v = v_1$.

3°. *let $v, w \in W$. Then $v \approx w$ if and only $v_r = w_r$.*

Let v, w be words. If $v_r = w_r$, then $v \approx v_r = w_r \approx w$ and so $v \approx w$.

Define the relation $\approxeq$ on W by $v \approxeq w$ if $v_r = w_r$. Clearly $\approxeq$ is an equivalence relation. Let $v, w \in W$ with $v \sim w$. Since $v_r \ll v$ we get $v_r \ll w$. Since v_r is reduced, (2°) gives that $v_r = v_w$. We show that $v \sim w$ implies $v \approxeq w$. Thus by 2.9.2(c) says that $v \approx w$ implies $v \approxeq w$. So $v \approx w$ implies $v_r = w_r$.

4°. *Let $g \in W/\approx$. Then there exists a unique reduced word $v \in W$ with $g = [v]$. Moreover, if $g = [w]$ for some $w \in W$, then $v = w_r$.*

Let $g = [w]$ with $w \in W$. Then $w \approx w_r$ and so $g = [w] = [w_r]$. Let $v \in W$ be reduced with $g = [v]$. Then $v \approx w$ and so by (4°), $v = v_r = w_r$.

Let W_r be the set of reduced words. (4°), the map $W_r \to \tilde{W}$, $w \to [w]$ is a bijection. Unfortunately, W_r is not closed under multiplication, that is the product of two reduced words usually is not reduced. But it is not difficult to figure out what the reduction of the product is. Indeed let $x = x_1 x_2 \dots x_n$ and $y = y_1 y_2 \dots y_n)$ be reduced words. Let s be maximal with $y_t^{-1} = x_{n-t}$ for all $1 \leq t < s$. Then

$$xy \approx (x_1 x_2 \dots x_{n-s} y_s y_{s+1} \dots y_m)$$

If x_{n-s} and y_s are not contained in a common G_i this is the reduction of xy.
On the other hand if x_{n-s} and y_s both are contained in G_i, then

$$xy \approx x_1, \dots x_{n-s-1}(x_{n-s} \cdot y_s) y_{s+1} \dots y_m)$$

By maximality of s, $x_{n-s} \cdot y_{s+1} \neq e_{G_i}$ and it is easu to seen that word on the roght hand side of the last equation is reduced, and so is the reduction of $x * y$.

We remark that coproducts also exists for semigroups and for monoids. Indeed, everything we did for groups carries over word for word, with one exception though. In case of semigroups we do not include the empty tuple in the sets of words and omit 2.9.5(i) in the definition of $v \sim w$.

Example 2.9.8. Let $A \cong B \cong \mathbb{Z}/2\mathbb{Z}$. We will compute $D = A \coprod B$. To simply notation we identify $x \in A \cup B$ with its image in D. In particular $e := e_G = e_A = e_B$ and ρ_A and ρ_B are just the inclusion map from A and B respectively to D. Let $e \neq a \in A$ and $b \neq B$ in B. Then every elements in D has one of the following forms:

$$e$$
$$\underbrace{(ab)(a*b)\dots(ab)}_{n \text{ times}}$$
$$\underbrace{(ba)(ba)\dots(ba)}_{n \text{ times}}$$
$$b\underbrace{(ab)(ab)\dots(ab)}_{n \text{ times}}$$
$$a\underbrace{(ba)(ba)\dots(ba)}_{n \text{ times}}$$

Put $z = ab$. Then $z^{-1} = b^{-1}a^{-1} = ba$. So the above list now reads z^0, z^n, z^{-n}, bz^n and az^{-n}. Note that $bz^n = b(ab)^n = a(ab)(ab)^n = a(ab)^{n+1} = az^{n+1}$ and so

$$D = \{z^n, z^n a \mid n \in \mathbb{Z}\}.$$

It is also easy to compute the product of two elements in D: first observe that $z^n a = a(ba)^{n-1}ba = az^{-n}$ and so

$$z^n * z^m = z^{n+m}, z^n * z^m a = z^{n+m}a, az^n * az^m = aaz^{-n}z^m = z^{m-n}$$

This can be combined in one formula: Define $\epsilon(0) = 1$ and $\epsilon(1) = -1$. Then for $n, m \in \mathbb{N}$ and $i, j \in \{0, 1\}$:

$$(a^i z^n) * (a^j z^m) = a^{i+j} z^{\epsilon(j)n+m}$$

By now we determined the complete multiplication table of D. D is called the infinite *dihedral* group.

We will know constructed a second group $\tilde{D}$ and show that it is isomorphic to D. Define

$$\tilde{a}: \quad \mathbb{Z} \to \mathbb{Z} \quad m \to -m$$
$$\tilde{b}: \quad \mathbb{Z} \to \mathbb{Z} \quad m \to 1 - m$$

Then $\tilde{a}, \tilde{b} \in \mathrm{Sym}(\mathbb{Z})$, $\tilde{a}$ is a reflection at 0 and $\tilde{b}$ is the reflection at $\frac{1}{2}$ Put $\tilde{D} = \langle \tilde{a}, \tilde{b} \rangle$. Since both $\tilde{a}$ and $\tilde{b}$ have order two, there exist homomorphism,$\alpha_A : A \to \tilde{D}$ and $\alpha_B : B \to \tilde{D}$ with $\alpha_A(a) = \tilde{a}$ and $\alpha_B(b) = \tilde{b}$. Hence by the definition of the co-product, there exists a homomorphism $\beta : D \to \tilde{B}$ with $\alpha_A = \beta \circ \rho_A$ and $\alpha_B = \beta \circ \rho_B$. Then

$$\beta(a) = \beta(\rho_A(a)) = \alpha_A(a) = \tilde{a}$$

and similarly $\beta(b) = \tilde{b}$.

Hence

$$\beta(D) = \beta(\langle a, b \rangle) = \langle \beta(a), \beta(b) \rangle = \langle \tilde{a}, \tilde{b} \rangle = \tilde{D}$$

So β is onto. Put $\tilde{z} = \tilde{a} \circ \tilde{b}$. Then

$$\tilde{z}(m) = \tilde{a}(\tilde{b}(m)) = \tilde{a}(1 - m) = m - 1$$

So $\tilde{z}$ is the translation by -11. Also $\tilde{z}^j(m) = m - n$ and $(\tilde{a}\tilde{z}^j)(m) = \tilde{a}(m - j) = j - m$. Thus z^j is translation by $-j$ and $\tilde{a}\tilde{z}^n$ is the reflection at $\frac{j}{2}$.

We have $\beta(z) = \beta(ab) = \beta(a)\beta(b) = \tilde{a}\tilde{b} = \tilde{z}$ and so also $\beta(a^i z^j) = \beta(a)^i \tilde{(b)}^j = \tilde{a}^i \tilde{z}^j$. Since the $\tilde{a}^i \tilde{z}^j, i = 0, 1, j \in \mathbb{Z}$ are pairwise distinct, we conclude that β is 1-1. Thus β is an isomorphism and

$$D \cong \tilde{D}$$

Let $Z = \langle z \rangle$. Then $Z \cong (\mathbb{Z}, +)$ and Z has index two in $G_1 * G_2$. In particular, $Z \trianglelefteq D$. Also $z^a = aza = aaba = ba = z^{-1}$. Thus

$$(z^n)^a = z^{-n} \quad \text{and} \quad z^n a = az^{-n}.$$

In particular, if $A \leq Z$ then both Z and a normalize D and $A \trianglelefteq G$.

Here is a property of D which will come in handy later on:

All elements in $D \setminus Z$ are conjugate to a or b.

Indeed $z^n a z^{-n} = z^n z^n a = z^{2n} a$ and $z^n b z^{-n} = z^{2n} b = z^{2n} baa = z^{2n+1} a$. So $z^{2n} a$ is conjugate to a and $z^{2n+1} a$ is conjugate to b.

Fix $n \in \mathbb{Z}$. Consider the relation $z^n = e$. Put $N = \langle z^n \rangle$. Then $N \trianglelefteq D$ and so

$$\coprod_{i \in \{1,2\}} G_i / \langle (ab)^n = e \rangle = D/N$$

Since $Z/D \cong \mathbb{Z}/n\mathbb{Z}$, D/N has order $2n$. D/N is called the dihedral group of order $2n$, or the dihedral group of degree n.

Suppose now that $\bar{D}$ is any group generated by two elements of order two, $\bar{a}$ and $\bar{b}$. Then there exists a homomorphism $\alpha : D \to \bar{D}$ sending a to $\bar{a}$ and b to $\bar{b}$. Let $\bar{z} = \bar{a}\bar{b}$ and $\bar{Z} = \langle \bar{z} \rangle$. Since neither a nor b are in $\ker \alpha$ and all elements in $D \setminus Z$ are conjugate to a or b, $\ker \alpha \leq Z$. Thus $\ker \alpha = \langle z^n \rangle$ for some $n \in \mathbb{N}$ and so $\bar{D} \cong D/\ker \alpha = D/N$. So any group generated by two elements of order 2 is a dihedral group.

Definition 2.9.9. *Let I be a set. A free group generated by I is a group F_I together with a map $\rho : I \to F_I$ with the following property:*

Whenever H is a group and $\alpha : I \to H$ is a function, then there exists a unique homomorphism $\beta : F_I \to H$ with $\alpha = \beta \circ \rho$.

Lemma 2.9.10. *Let I be a set. Then there exists a free generated by I.*

Proof. For $i \in I$ let $G_i = (\mathbb{Z}, +)$ and let $\rho_i : G_i \to F_I, i \in I)$ be a coproduct of $(G_i, i \in I)$. Define $\rho : I \to F_i$, $i \to \rho_i(1)$. Now let H be a group and $\alpha : I \to H$ be function. Define $\alpha_i : G_i \to H, m \to \alpha(i)^m$. Then by definition of the coproduct of $(G_i, i \in I)$ there exists a unique homomorphism $\beta : F_I \to H$, with $\alpha_i = \beta \circ \rho_i$. Then $\alpha(i) = \alpha_i(1) = \beta(\rho_i(1)) = \beta(\rho(i))$ and so $\alpha = \beta \circ \rho$. Suppose alos $\gamma : F_I \to H$ fulfills, $\alpha = \gamma \circ \rho$. Then for all $m \in G_i$,

$$\alpha_i(m) = \alpha(i)^m = \gamma(\rho(i))^m = \gamma(\rho_i(1)^m) = \gamma(\rho_i(m))$$

Hence $\alpha_i = \gamma \circ \rho_i$ and so by the uniqueness assertion on the defintion of the coproduct, $\beta = \gamma$. □

Notation 2.9.11. *Let I be a set. Then F_I is a group with $I \subseteq F_I$ such that $\mathrm{id}_{I,F_I} : I \to F_I, i \to i$ is a free group generated by I.*

2.9.12 (Reduced words in free groups)**.** Let I be a set and $G_i = \langle i \rangle = \{i^m \mid m \in \mathbb{Z}\}$, the subgroup of F_I generated by I. Then by proof of 2.9.10 $G_i \cong \mathbb{Z}$ and F_I is the co-product of the $(G_i, i \in I)$. So by 2.9.7, each element in F_I can be uniquely written as $g_1 g_2 \dots g_n$ where $n \in \mathbb{N}$, $g_j \in G_{i_j}$, $g_j \neq e_{G_{i_j}}$ for all $1 \leq j \leq n$ and $i_j \neq i_{j+1}$ for all $1 \leq j < n$. Since $G_{i_j} = \langle i_j \rangle$ we have $g_j = i_j^{m_j}$ for some $0 \neq m_j \in \mathbb{Z}$. Thus every element w in F_I can be uniquely written as

$$w = i_1^{m_1} i_2^{m_2} \dots i_n^{m_n}$$

where $n \in \mathbb{N}$, $i_j \in I, 0 \neq m_j \in \mathbb{Z}$ and $i_j \neq i_{j+1}$. $i_1^{m_1} i_2^{m_2} \dots i_n^{m_n}$ is called the *reduced form* of g.

Let G is a group and $(g_i, i \in I)$ a tuple of elements in G Define $\alpha : I \to G, i \to g_i$. Then by definition of the free group there exists a unique homomorphism $\beta : F_I \to G$ with $\alpha = \beta \circ \mathrm{id}_{I \to F}$. Note that $\alpha = \beta \circ \mathrm{id}_{I \to F_I}$ just means $\beta(i) = g_i$ for all $i \in I$. Thus

$$\beta(i_1^{m_1} i_2^{m_2} \dots i_n^{m_n}) = g_1^{m_1} g_2^{g_2} \dots g_n^{m_n}$$

Definition 2.9.13. *Let I be a set.*

(a) A group relation *over I is an expression of the form $v = w$, where $v, w \in F_I$.*

(b) Let G be a group and $(g_i, i \in I)$ a family of elements in G. We say that $(g_i, i \in i)$ fulfills the relation $v = w$ provided that $\beta(v) = \beta(w)$, where β is the unique homomorphism from F_I to G with $\beta(i) = g_i$.

If $w = i_1^{m_1} \dots i_n^{m_n}$, then $(g_i, i \in I)$ fulfills the relation $w = e$ if and only if $g_1^{m_1} \dots g_n^{m_n}$.

Definition 2.9.14. *Let I be a set and $\mathcal{R}$ a set of group relations on I. Then a group with generators I and relations $\mathcal{R}$ is a group G together with a family of elements $(g_i, i \in I)$ of elements of G such that*

(a) $(g_i, i \in I)$ fulfills all the relations in $\mathcal{R}$.

(b) Whenever H is a group and $(h_i, i \in I)$ is a family of elements of H fulfilling all the relations in $\mathcal{R}$, then there exists a unique homomorphism $\delta : G \to H$ with $\delta(g_i) = h_i$ for all $i \in I$.

Example 2.9.15. Let $I = \{a, b\}$, $G = \mathrm{Sym}(3)$ and consider the relation $aba^{-1} = b^{-1}$.

Do $g_a = (12)$ and $g_b = (123)$ fulfill the relation? In other words is

$$(12) \circ (123) \circ (12)^{-1} \stackrel{?}{=} (123)^{-1}$$

The left hand side is (213) and the right hand side is (321), both of which are equal to (132). So the answer is yes.

Do $h_a = (12)$ and $h_b = (23)$ fulfill the relation?

$$(12) \circ (23) \circ (12)^{-1} \stackrel{?}{=} (23)^{-1}$$

The left side is (13) the right side is (23), so this time the answer is no.

Lemma 2.9.16. *Let I be a set and $\mathcal{R}$ a set of group relations on I. Then there exists a group G with generators I and relations $\mathcal{R}$.*

Proof. Note that the relation $v = w$ is fulfilled if and only if the relation $vw^{-1} = e$ is fulfilled. So we may assume that $R = \{r = e \mid r \in R\}$ for some subset R of F_I. Put $N := \langle {}^{F_I}R\rangle = \langle {}^{w}r \mid w \in F_I, r \in R\rangle$, so N is the normal subgroup of F_I generated by R. Put $G = F_I/N$ and let $g_i = iN$ for $i \in N$. $\pi_N : F_I \to G, w \to wN$ is a homomorphism fro F_I to G with $\pi_N(i) = g_i = iN$ and $\pi_N(r) = rN = N = e_G$ for all $r \in R$. So $(g_i, i \in I)$ fulfills all the relation in $\mathcal{R}$.

Now let H be a group and $(h_i, i \in I)$ is a family of elements of H fulfilling all the relations in $\mathcal{R}$. Let $\beta : F_I \to H$ be the unique homomorphism with $\beta(i) = h_i$ for all $i \in I$. Let $r \in R$. Then $(h_i, i \in I)$ fulfills the relation $r = e$ and so $\beta(r) = e_H$. Hence $r \in \ker\beta$. Since $\ker\beta \trianglelefteq F_I$, ${}^{w}r \in N$ for all $w \in W$ and so $N \leq \ker\beta$. It follows that the map

$$\delta : G \to G, wN \to \beta(w)$$

is a well-defined homomorphism. Also $\delta(g_i) = \delta(iN) = \beta(i) = h_i$.

It remains to show that uniqueness δ. So let $\alpha : G \to H$ be a homomorphism with $\alpha(g_i) = h_i$. The $(\alpha \circ \pi_N)(i) = \alpha(g_i) = h_i$ and so $\alpha \circ \pi_N = \beta$ by uniqueness of β. Hence for all $w \in F_I$, $\alpha(wN) = (\alpha \circ \pi_N)(w) = \beta(w)$ and so $\alpha = \delta$. □

2.9.17 (Notation in groups with generators and relation)**.** Let I be a set and $\mathcal{R}$ a set of relations on $\mathcal{R}$. Then

$$G = \langle I \mid \mathcal{R}\rangle$$

means that G together with the family of elemensts $(g_i)_{i\in I}$ is is a group with generators I and relations $\mathcal{R}$. So

$$G = \langle g_i \mid i \in I\rangle$$

and if

$$(*) \qquad i_1^{n_1} \dots i_k^{n_k} = j_1^{m_1} \dots j_l^{m_l}$$

is one of the relation in $\mathcal{R}$ then

$$(**) \qquad g_{i_1}^{n_1} \dots g_{j_1}^{n_k} = g_{j_1}^{m_1} \dots g_{j_l}^{m_l}.$$

In practical computation is is often quite cumbersome to work with elements with supscripts. We therefore often just write a for the element g_a in G. This should be only done if this is clearly from the context that the conputation are done in G and that a no longer stands for the element a in F_I. Note also that this is not an identification, since the map $I \to G, a \to g_a$ is (in general) not 1-1. The advantage of this convention is that, replacing all g_a by a, the equation (**) now turns into the easier

$$(***) \qquad i_1^{n_1} \dots i_k^{n_k} = j_1^{m_1} \dots j_l^{m_l}.$$

This seems to be the same as (*). But it is not. (*) is a group relation between elements of F_I, not an actual equality. (***) is an actual equality of elements in G.

Example 2.9.18. 1. The group

$$G = \langle a, b \mid a^2 = e, b^2 = e \rangle$$

is the infinite dihedral group. To see that let $H_a = \langle h_a \rangle$ and $H_b = \langle h_b \rangle$ be cyclic groups of order 2 and $H = H_a \coprod H_b$. Let $G_a = \langle g_a \rangle \leq G$ and $G_b = \langle g_b \rangle \leq G$. Since $g_a^2 = g_b^2 = e$. There exists homomorphism $\alpha_a : H_a \to G_a$ and $\alpha_b : H_b \to G_b$ with $\alpha_a(h_a) = g_a$ and $\alpha_b(h_b) = g_b$. So by defintion of the coproduct, there exists unique homomorphism $\alpha : H \to G$ with $\alpha(h_a) = g_a$ and $\alpha(h_b) = g_b$. Conversely, since (g_a, g_b) fulfills the relation $a^2 = e$ and $b^2 = e$, there exists a unique homomorphism $\beta : G\overline{H}$, with $\beta(g_a) = h_a$ and $\beta(g_b) = h_b$. It is now easy to see that $\alpha \circ \beta = \mathrm{id}_G$ and $\beta \circ \alpha = \mathrm{id}_H$. So $G \cong H$.

Informally what we just proved is that the 'largest' group generated by two elements of order two is same as the 'largest' groups generated by two groups of order two.

2. The group

$$\langle a, b \mid a^2 = 2, b^2 = e, (ab)^n = e \rangle$$

is the called the *dihedral group* Dih_{2n} of degree n or the *dihedral group* of order $2n$

Let $F = F_{\{a,b\}}$ and $K = \langle {}^F\{a^2, b^2\}$ and $N = \langle {}^F\{a^2, b^2, (ab)^n\}\rangle$. Then by (1), F/K is the infinite dihedral group. For $x \in F$ let $\overline{x} = xK$. Put $z = ab$ and $y = (ab)^n$. Then $N = K\langle {}^F y\rangle$. By 2.9.8, $\overline{a}\overline{z} = \overline{z}^{-1}\overline{a}$. If follows that ${}^{\overline{a}}\overline{z} = \overline{z}^{-1}$, ${}^{\overline{a}}\overline{y} = \overline{y}^{-1}$ and ${}^{\overline{a}}\overline{y}\rangle = \langle {}^{\overline{a}}\overline{y}\rangle = \langle \overline{y}^{-1}\rangle = \langle o\ y\rangle$. Hence $\overline{a}$ and $\overline{z}$ normalizes $\langle \overline{y}\rangle$. Since $\overline{F} = \langle \overline{a}, \overline{z}\rangle$, $\langle \overline{y}\rangle$ is normal in $\overline{F}$. Hence $K\langle y\rangle$ is normal in F and $N = K\langle \overline{y}\rangle$. Thus

$$F/N \cong \overline{F}/\overline{N} = \overline{F}/\langle \overline{y}\rangle = \overline{F}/\langle \overline{z}^n\rangle$$

By 2.9.8 $\overline{F} = \{\overline{a}^i\overline{z}^j \mid i \in 0, 1, j \in \mathbb{Z}\}$. Since $(\overline{a}^i\overline{z}^j)\overline{z}^{nm} = \overline{a}^i\overline{z}^{l+nm}$ we see that

$$a^i z^j N = a^k z^l N \quad \Longleftrightarrow i = k \text{ and } j \equiv k \pmod{n}$$

Thus $F/N = \{a^i z^j N \mid 0 \leq i \leq 1, 0 \leq j < n\}$ and so $|F/N$ has order 2.

We will now construct a second group which is isomorphic to F/N. This is similar to construction of the group $\tilde{D}$ in Example 2.9.8. The only difference is that we repalce $\mathbb{Z}$ by $\mathbb{Z}/n\ mbZ$ where n is an integer with $n \geq 2$.

Define $\tilde{a} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, m \to -m$ and $\tilde{b} : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, m \to 1 - m$. Put $\tilde{z} = \tilde{a} \circ \tilde{b}$. Then as in 2.9.8

$$\tilde{z}(m) = m + 1, \quad \tilde{z}^j(m), \quad = m - j, \quad \tilde{a}\tilde{z}^j(m) = j - m$$

Put $\tilde{G} = \langle \tilde{a}, \tilde{b}\rangle$. Since the calculation are done modulo n we conclude that

$$|\tilde{z}| = n, \tilde{D} = \{\tilde{a}^i\tilde{z}^j \mid 0 \leq 1 \leq i, 0 \leq j < n\}, \text{ and } |\tilde{D}| = 2n$$

So $(\tilde{a}, \tilde{b})$ fulfills the relations for G and so there exists a unique homomorhism $\beta : G \to \tilde{G}$ with $\beta(aN) = \tilde{a}$ and $\beta(bN) = \tilde{b}$. Then $\beta(a^iz^jN) = \tilde{a}^i\tilde{b}^j$ and so β is a bijection. Thus $G \cong \tilde{G}$.

Here is a more geometric version of the above. Let $\xi = e^{\frac{2\pi i}{n}} \in \mathbb{C}$ and $U_n = \{\xi^i \mid 0 \leq i < n\}$. So U_n is the set of n-roots of unity in $\mathbb{C}$ and the map $\mathbb{Z}/n\mathbb{Z}, +) \to (U_n, \cdot), i \to \xi^i$ is an isomorphism. $\tilde{z}^j$ corresponds to clockwise rotation by $\frac{j}{n}2\pi$ radiants and $\tilde{a}z^i$ correponds to the reflection at the line through 0 and $\xi^{\frac{j}{2}}$. Note that if j is even $xi^{\frac{j}{2}}$ is in U_n, while if j is odd $xi^{\frac{j}{2}}$ is the midpoint on the unit circle between $\xi^{\frac{j-1}{2}}$ and $\xi^{\frac{j+1}{2}}$.

3. We will show that

$$G := \langle a, b, c \mid ab = c, ab = ba, c^2 = a, c^3 = b, c^5 = e\rangle$$

is the trivial group.

We will follow the conventions of 2.9.17 and just write a for g_a, b for g_b and c for g_c, that is we treat a, b, c as elements of G, rather than elements of $F_{\{a,b,c\}}$. Then the relations defining G become actual equalities and so $c = ab = c^2c^3 = c^5 = e$. Hence also $a = c^2 = e$ and $b = c^2 = e$. Thus $G = \{e\}$.

4.

$$G := \langle a, b \mid a^3 = b^3 = (ab)^2 = e\rangle \cong \text{Alt}(4)$$

To see this let $z = ab$. Then $z^2 = 1$. Put $K = \langle z, z^a\rangle$. Since both z and z^a have order two (or 1), K is a dihedral group. We compute

$$z^{a^2}z^az = (a^2(ab)a^{-2})\, a(ab)a^{-1}\, ab = a^3b(a^{-2}a^2)b(a^{-1}a)b = a^3b^3 = e.$$

Thus $z^az = z^{-a^2} = z^{a^2}$. In particular, $(z^az)^2 = e$. It is easy to see that $\{e, z, z^a, z^{a^2}\}$ is a subgroups of G and so $K = \{e, z, z^a, z^{a^2}\}$. (This also can be deduced from (2). Indeed (z, a^2) fulfills the relation for $\text{Dih}_{2\cdot 2}$ and by (2), $K = \{z^i(zz^a)^j \mid 0 \leq i, j < 2\}$.) Now $(z^{a^2})^a = z^{a^3} = z^e = z$ and so $a \in N_G(K)$. Thus $\langle a\rangle K$ is a subgroup of G. It contains a and $z = ab$ and so also $b = a^{-1}z$. Thus $G = \langle a\rangle K$. As K has order dividing 4 and $\langle a\rangle$ has order dividing 3, G has order dividing 12. Thus to show that G is isomorphic to Alt(4) it suffices to show that Alt(4) is a homomorphic image of G. I.e we need to verify that Alt(4) fulfills the relations.

For this let $a^* = (123)$ and $b^* = (124)$. Then $a^*b^* = (13)(24)$ and so $(a^*b^*)^2 = e$. Thus there exists a homomorphism $\phi : G \to \text{Alt}(4)$ with $\phi(a) = a^*$ and $\phi(b) = b^*$. As a^* and b^* generate Alt(4), ϕ is onto. As $|G| \leq |\text{Alt}(4)|$ we conclude that ϕ is an isomorphism.

2.10 Group Actions

Definition 2.10.1. *An* action *of a group $(G, \cdot)$ on a set S is a function*

$$\diamond : G \times S \to S, (a, s) \to a \diamond s$$

such that

(GA1) $e \diamond s = s$ *for all* $s \in S$.

(GA2) $(a \cdot b)s = a \diamond (b \diamond s)$ *for all* $a, b \in G$, $s \in S$.

A G-set *is a set S together with an action of G on S.*

We wll often just write, as for $a \diamond s$. So the two axioms of a group action then read $es = e$ and $(ab)s = a(bs)$.

Example 2.10.2. 1. Note the similarity between the definition of a groups action and the definition of a group. In particular, we see that the the binary operation of a group $\cdot : G \times G \to G$ defines an action of G on G, called the action by *left multiplication.* Indeed since e is an identity, 2.10.1(*(GA1)*) holds and since $\cdot$ is associative 2.10.1(*(GA2)*) holds.

2. The function
$$G \times G(a, s) \to a * s := sa$$
is not an action (unless G is abelian) since $(ab) * s = sab = (a * s)b = (b * a)s$. For this reason we define the action of G on G by *right multiplication* as
$$\cdot_{\mathrm{r}} G \times G, (a, s) \to sa^{-1}.$$
Then $(ab) \cdot_{\mathrm{r}} s = s(ab)^{-1} = sb^{-1}a^{-1} = a \cdot_{\mathrm{r}} (b \cdot_{\mathrm{r}} s)$ and $\cdot_{\mathrm{r}}$ is indeed an action.

3. Let G be a group and H a subgroup of G. Then H acts on G by left multiplication:
$$H \times G \to G, (h, g) \to hg$$

4. Let G be a group acting on the set I and let $H \leq G$. Then H acts on I via
$$H \times I \to I, (h, i) \to hi$$

5. G acts on G via conjugation:
$$G \times G \to G, (a, g) \to {}^a g$$
Indeed ${}^e g = g$ and ${}^{(ab)}g = {}^a({}^b g)$.

6. Let I be a set. Then $\mathrm{Sym}(I)$ acts on I via

$$\mathrm{Sym}(I) \times I \to I, (\pi, i) \to \pi(i)$$

Indeed, $\mathrm{id}_I(i) = i$ for all i in I and $\alpha(\beta(i)) = (\alpha\beta)(i)$ for all $\alpha, \beta \in \mathrm{Sym}(I), i \in I$.

7. Let G be a group. Then $\mathrm{Aut}(G)$ acts on G via

$$\mathrm{Aut}(G) \times G \to G, (\alpha, g) \to \alpha(g)$$

Indeed by (5), $\mathrm{Sym}(G)$ acts on G and so by (4) alsp the subgroup $\mathrm{Aut}(G)$ of $\mathrm{Sym}(G)$ acts on G.

The next lemma shows that an action of G on S can also by thought of as an homomorphism from G to $\mathrm{Sym}(S)$.

Lemma 2.10.3. *Let G be a group and S a set.*

(a) Let $\diamond : G \times S \to S$ an action of G on S. For $g \in G$ define

$$g^\diamond : S \to S, s \to gs$$

Then $g^\diamond \in \mathrm{Sym}(S)$ and the map

$$\Phi_\diamond G \to \mathrm{Sym}(S)$$

is an homomorphism.

$\Phi_\diamond$ is called the homomorphism *corresponding to $\diamond$,*

(b) Let $\Phi : G \to \mathrm{Sym}(S)$ be a homomorphism. Define

$$\diamond_\Phi : G \times S \to S, (g, s) \to \Phi(g)(s)$$

then $\diamond_\Phi$ is an action of G on S.

$\diamond_\Phi$ is called the action *corresponding to Φ*

(c) $\Phi_{\diamond_\Phi} = \Phi$ and $\diamond_{\Phi_\diamond} = \diamond$.

Proof. To simplify notation we just write ϕ_g for $g^\diamond$. (a) (GA1) into $\phi_e(s) = gs = s$ and so $\phi_e = id_S$. By (GA2)

$$(\phi_g \circ \phi_h)(s) = g(hs) = (gh)s = \phi_{gh}(s)$$

and so

$$\phi_g \circ \phi_h = \phi_{gh}$$

Hence Φ is a homomorphism. We still need to verify that $\phi_g \in \mathrm{Sym}(S)$. But this follows from

$$\mathrm{id}_S = \phi_e = \phi_{gg^{-1}} = \phi_g \circ \circ \phi_{g^{-1}}$$

and so also $\phi_{g^{-1}} \circ \phi_g = \mathrm{id}_S$. So $\phi_{g^{-1}}$ is an inverse for ϕ_g and $\phi_g \in \mathrm{Sym}(S)$.

(b) Since Φ is a homomorphism $\Phi(e) = e_{\mathrm{Sym}(S)} = \mathrm{id}_S$ and so $e \diamond_\Phi s = \Phi(e)(s) = \mathrm{id}_S(s) = s$. So $(GA1)$ holds. Also

$$(gh) \diamond_\Phi s = \Phi(gh)(s) = (\Phi(g) \circ \Phi(h))(s) = \Phi(g)(\Phi(h)(s)) = \gamma \diamond_\Phi (h \diamond_\Phi s)$$

and so also (GA2) holds.

(c) $g \diamond_{\Phi_\diamond} s = \Phi_\diamond(g)(s) = g \diamond s$ and
$\Phi_{\diamond_\Phi}(g)(s) = g \diamond_\Phi s = \Phi(g)(s)$. □

Example 2.10.4. 1. Let G be a group. For $a \in G$, define $\phi_a : G \to G, g \to ag$. Then by 2.10.2(1) and 2.10.3(a) the map

$$\Phi : G \to \mathrm{Sym}(G), a \to \phi_a$$

is a hommorphism. If $\Phi(a) = \mathrm{id}_G$, then $a = ae = \Phi(a)(e) = \mathrm{id}_G(e) = e$ and so Φ is 1-1. Thus $G \cong \Phi(G)$. In particular, G is isomorphic to a subgroup of a symmetric group. This is known as *Cayley's Theorem.*

2. Let G be group. Recall that for $g \in G$, i_g is the map

$$\mathrm{i}_g : G \to G, a \to {}^g a$$

By 2.10.2(1) G acts G by conjugation, the corresponing homomorphism is:

$$G \to \mathrm{Sym}(G), g \to i_g$$

3. The homomorphism corresponding to the action of $\mathrm{Sym}(I)$ on I corresponds is $\mathrm{id}_{\mathrm{Sym}(I)}$. Indeed,

Definition 2.10.5. *Let $\diamond$ by an action of the group G on the set I, $H \subseteq G$,$g \in G$, $s \in S$ and $T \subseteq S$. Then*

(a) $\mathrm{Stab}^\diamond_H(T) = \{h \in H \mid gt = t \text{ for all } t \in T\}$ *and* $\mathrm{Stab}^\diamond_H(s) = \{h \in H \mid hs = s\}$. $\mathrm{Stab}^\diamond_H(T)$ *is called the* stabilizer *of T in H.*

(b) $\mathrm{Fix}_T(H) = \{t \in T \mid ht = t \text{ for all } t \in T\}$ *and* $\mathrm{Fix}_T(g)) = \{t \in T \mid gt = t\}$. *The elements of* $\mathrm{Fix}_T(H)$ *are called the fixed-points of H in T.*

(c) $gT = \{gt \mid t \in T\}$, $Hs = \{hs \mid h \in H\}$, $HT = \{ht \mid h \in H, t \in T$

(d) $\diamond$ *is called a* faithful action *if* $\mathrm{Stab}_G(S) = \{e\}$. *In this case we also say that S is a faithful G-set.*

(e) T *is called* H-invariant *(with respect to* $\diamond$ *if* $hT = T$ *for all* $h \in H$*.* T *is called* g*-invariant if* $gT = T$*.*

(f) $N_H(T) = \{h \in H \mid hT = T\}$*.* $N_H(T)$ *is called the* normalizer *of* T *in* H*.*

(g) $G^\diamond := \operatorname{Im} \Phi_\diamond$*.*

We will often just write $\operatorname{Stab}_H(S)$ in place of $\operatorname{Stab}_H^\diamond(S)$, but of course only if its clear from the context what the underlying action $\diamond$ is.

Lemma 2.10.6. *Let* $\diamond$ *be an action of the group group on the set* S*.*

(a) $\operatorname{Stab}_G(S) = \ker \Phi_\diamond \trianglelefteq G$*.*

(b) $G/\operatorname{Stab}_G(S) \cong G^\diamond \leq \operatorname{Sym}(S)$*.*

(c) S *is a faithful* G*-set if and only if* $\Phi_\diamond$ *is 1-1. So if* S *is faithful,* G *is isomorphic to a subgroups of* $\operatorname{Sym}(S)$*.*

(d) Let $H \leq G$ *and* T *an* H*-invariant subset of* S*, then*

$$\diamond|_{H,T}\ H \times T \to T, (h,t) \to ht$$

ia an action of H *on* T*.*

(e) The map $\diamond_{\mathcal{P}} G \times \mathcal{P}(G) \to \mathcal{P}(G), (g,T) \to T$ *is an action of* G *on* $\mathcal{P}(G)$*.*

Proof. (a) Let $g \in G$, then

$$\begin{aligned}
& g \in \operatorname{Stab}_G(S) \\
\iff \quad & gs = s \text{ for all } g \in G \\
\iff \quad & g^\diamond(s) = s \text{ for all } g \in G \\
\iff \quad & \gamma^\diamond = \mathrm{id}_S \\
\iff \quad & \Phi^\diamond(g) = \mathrm{id}_S \\
\iff \quad & g \in \ker \Phi^\diamond
\end{aligned}$$

(b) Since $G^\diamond = \operatorname{Im} \Phi^\diamond$, this follows from (a) and the First Isomorphism Theorem.
(c) - (e) are readily verified. □

Lemma 2.10.7. *Let* G *be a group acting on the set* S*. Let* $s \in S$ *and* $T \subseteq S$*.*

(a) $\operatorname{Stab}_G(T)$ *is a subgroup of* G*.*

(b) $\operatorname{Stab}_G(s)$ *is a subgroup of* G*.*

(c) $\mathrm{N}_G(T)$ *is a subgroup of* G*.*

Proof. (a) $et = e$ for all $t \in T$ and so $e \in \mathrm{Stab}_G(T)$. Let $g, h \in \mathrm{Stab}_G(T)$. Then $gt = t$ and $ht = t$ for all $t \in T$. Thus

$$(gh)t \overset{(\mathrm{GA2}}{=} g(ht) = gt = t$$

and so $gh \in \mathrm{Stab}_G(T)$.

From $gt = t$ we get $g^{-1}(gt) = g^{-1}t$. So by (GA2), $(g^{-1}g)t = g^{-1}t$ and $et = g^{-1}t$. Thus by (GA1), $t = g^{-1}t$. Hence $g^{-1} \in \mathrm{Stab}_G(T)$. 2.4.3 now implies that $\mathrm{Stab}_G(T)$ is a subgroup of G.

Note that $\mathrm{Stab}_G(s) = \mathrm{Stab}_G(\{s\})$. Thus (b) follows from (c).

(c) We have

$$\mathrm{N}_G^{\diamond}(T) = \{g \in G \mid gT = T\} = \mathrm{Stab}_G^{\diamond_{\mathcal{P}}}(T).$$

(Note that on the left hand side T is treated as a subset of the G-set S, and in the right hand side, T is treated as an element of the G-set $\mathcal{P}(S)$.) Thus (c) follows from (b). □

Example 2.10.8. 1. Let G be a group Let $\diamond$ be the action of G on G be conjugation and let $A \subseteq G$. Let $g \in G$. Then

$$\begin{aligned} & g \in \mathrm{Stab}_G^{\diamond}) \\ \iff \quad & g \diamond a = a \text{ for all } a \in A \\ \iff \quad & {}^g a = a \text{ for all } a \in A \\ \iff \quad & gag^{-1} = a \text{ for all } a \in A \\ \iff \quad & ga = ag \text{ for all } a \in A \end{aligned}$$

Define

$$C_G(A) := \{g \in G \mid ga = ag \text{ for all } a \in A$$

Then we proved $C_G(A) = \mathrm{Stab}^{\diamond}(A)$ and so by 2.10.7(a), $C_G(A) \leq G$.

Recall from the Homework that $\mathrm{Z}(G)$ is defined as $\{g \in G \mid ga = ga \text{ for all } a \in A\}$. So

$$Z(G) = C_G(G) = \mathrm{Stab}_G^{\diamond}(G)$$

and so by 2.10.6(a), $Z(G) \trianglelefteq G$, which of course we already know from the homework.

Lemma 2.10.9. *Let $\diamond : G \times S \to S$ be a group action. Define a relation $\sim_{\diamond}$ on S by $s \sim_{\diamond} t$ if and only $t = as$ for some $a \in G$. Then $\sim_{\diamond}$ is an equivalence relation on S.*

Proof. Since $s = es$, $s \sim s$ and $\sim$ is reflexive.

If $t = as$, then

$$a^{-1}t = a^{-1}(as) = (a^{-1}a)s = es = s$$

Thus $s \sim t$ implies $t \sim s$ and $\sim$ is symmetric.

Finally if $s = at$ and $t = br$ then $s = at = a(br) = (ab)r$. Thus $s \sim t$ and $t \sim r$ implies $s \sim r$ and $\sim$ is reflexive. □

Definition 2.10.10. *Let $\diamond : G \times S \to S$ be a group action.*

(a) The equivalence classes of $\sim_\diamond$ are called the orbits *of G on S with respect to $\diamond$.*

(b) The set of orbits is denoted by $S/^\diamond G$.

(c) We say that G acts transitively *on S if G has exactly one orbit on S.*

Lemma 2.10.11. *Let G be a group acting on the set non-empty set S.*

(a) Let $s \in S$. Then the orbit of G on S containing s is $Gs = \{gs \mid g \in G\}$.

(b) The following are equivalent:

(a) For each $s, t \in S$ there exists $g \in G$ with $t = gs$.

(b) There exists $s \in S$ with $S = Gs$.

(c) G acts transitively on S.

Example 2.10.12. Let G be group and $H \leq G$.

1. The right cosets of H are the orbits for the action of H on G by left multiplication.

2. The left cosets of H are the orbits for the action of H on G by the right multiplication.

3. The conjugacy classes of G are the orbits for the action G on G by conjugation.

4. Let I be a non-empty set. Then $\mathrm{Sym}(I)$ acts transitively on I.

5. $\diamond : G \times G/H \to G/H, (g, T) \to gT$ is a well -defined transitive action. Indeed, if $T = tH$, then $g(tH) = (gT) \in G/H$. So $\diamond$ is well-defined. It is clearly an action and $G \diamond H = \{gH \mid g \in G\} = G/H$ and so G acts transitively on G/H. This action is called the action of G on G/H by left multiplication.

We will show that any transitive action of G is isomorphic to the action on the coset of a suitable subgroup. But first we need to define isomorphism for G-sets.

Definition 2.10.13. *Let G be a group, $\diamond$ an action of G on the set S, $\triangle$ an action of G on the set T and $\alpha : S \to T$ a function.*

(a) α *is called* G-equivariant *if*

$$\alpha(g \diamond s) = g \triangle \alpha(s)$$

for all $g \in G$ *and* $s \in S$.

(b) α *is called a* G-isomorphism *if* α *is* G*-equivariant and an bijection.*

(c) If there exists a G*-isomorphism from* S *to* T *we say that* $\diamond$ *is isomorphic to* $\triangle$ *or that* S *and* T *are isomorphic* G*-sets and write*

$$\diamond \cong \triangle, \quad (S, \diamond) \cong (T, \triangle), \quad \text{or } S \cong_G T$$

Lemma 2.10.14. *Let* S *be a* G*-set,* $s \in S$ *and put* $H = \mathrm{Stab}_G(s)$.

(a) The map

$$\alpha : G/H \to S,\ aH \to as$$

is well defined, G*-equivariant and one to one.*

(b) α *is an* G*-isomorphism if and only if* G *acts transitively on* S

(c) $\mathrm{Stab}(as) = {}^aH$ *for all* $a \in G$.

(d) $|Gs| = |G/\mathrm{Stab}_G(s)|$.

Proof. (a) Let $a, b \in G$. Then

$$\begin{aligned} & aH = bH \\ \iff \quad & a^{-1}b \in H \\ \iff \quad & a^{-1}b \in \mathrm{Stab}_G(s) \\ \iff \quad & (a^{-1}b)s = s \\ \iff \quad & a^{-1}(bs) = s \\ \iff \quad & bs = as \end{aligned}$$

The forward direct shows that α is well-defined and the backward direction shows that α is 1-1.

Also

$$\alpha(a(bH)) = \alpha((ab)H) = (ab)s = a(bs) = a\alpha(bH)$$

So α is G-equivariant.

(b) By (a) α is a G-isomorphism if and only if α is onto. We have

$$\mathrm{Im}\,\alpha = \{\alpha(gH) \mid g \in G\} = \{gs \mid g \in G\} = Gs$$

So α is onto if and only if $S = Gs$ and so if and only if G is transitive on S.

(c)

$$g \in \mathrm{Stab}_G(as) \iff g(as) = as \iff a^{-1}gas = s \iff a^{-1}ga \in H \iff g \in aHa^{-1} = {}^aH$$

(d) Since α is 1-1, $|G/H| = |\operatorname{Im}\alpha| = |Gs|$. □

Lemma 2.10.15. *Suppose that G acts transitively on the sets S and T. Let $s \in S$ and $t \in T$. Then S and T are G-isomorphic if only if* $\mathrm{Stab}_G(s)$ *and* $\mathrm{Stab}_G(t)$ *are conjugate in G.*

Proof. Suppose first that $\alpha : S \to T$ is a G-isomorphism. Let $g \in G$. Since α is 1-1 and G-equivariant:

$$gs = s \iff \alpha(gs) = \alpha(s) \iff g\alpha(s) = \alpha(s)$$

So $\mathrm{Stab}_G(s) = \mathrm{Stab}_G(\alpha(s))$. Since G is transitive on T, there exists $g \in G$ with $g\alpha(s) = t$. Thus

$$\mathrm{Stab}_G(t) = \mathrm{Stab}_G(g\alpha(s)) = {}^g\mathrm{Stab}_G(\alpha(s)) = {}^g\mathrm{Stab}_G(s).$$

Conversely suppose that ${}^g\mathrm{Stab}_G(s) = \mathrm{Stab}_G(t)$ for some $g \in G$. Then $\mathrm{Stab}_G(gs) = {}^g\mathrm{Stab}_G(s) = \mathrm{Stab}_G(t)$ and so by 2.10.14(b) applied to S and to T:

$$S \cong G/\mathrm{Stab}_G(gs) = G/\mathrm{Stab}_G(t) \cong T.$$

□

Definition 2.10.16. *Let G be a group and S a G-set. A subset $R \subseteq S$ is called a set of* representatives *for the orbits of G on S, provided that R contains exactly one element from each G-orbit. In other words if the map $R \to S/G, r \to Gr$ is a bijection.*

An orbit O of G on S is called trivial *if $|O| = 1$.*

Let R be an set of representatives for the orbits of G on S and any trivail orbit $\{s\}$. Then s must be in R. Thus $\mathrm{Fix}_S(G) \subseteq R$ and $R \setminus \mathrm{Fix}_G(R)$ is a set of representatives for the non-trivial G-orbits.

Proposition 2.10.17 (Orbit Equation)**.** *Let G be a group, S a G-set and $R \subseteq S$ be a set of representatives for S/G.*

$$|S| = \sum_{r \in R} |G/\mathrm{Stab}_G(r)| = |\mathrm{Fix}_S(G)| + \sum_{r \in R \setminus \mathrm{Fix}_S(G)} |G/\mathrm{Stab}_G(r)|.$$

Proof. Since the orbits are the equivalemce classes of an equivalence relation S is the disjoint union of its orbit. Thus

$$|S| = \sum_{O \in S/G} |O| = \sum_{r \in R} |Gr|$$

By 2.10.14d, $|Gr| = |G/\mathrm{Stab}_G(r)|$ and so

$$|S| = \sum_{r\in R} |G/\mathrm{Stab}_G(r)|$$

If $r \in \mathrm{Fix}_S(G)$, then $gs = s$ for sll $g \in G$ and so $\mathrm{Stab}_G(r) = G$ and $|G/Stab_G(r)| = 1$. So

$$\begin{aligned} |S| &= \textstyle\sum_{r\in \mathrm{Fix}_S(G)} |G/\mathrm{Stab}_G(r)| + \sum_{r\in R\setminus \mathrm{Fix}_S(G)} |G/\mathrm{Stab}_G(r)| \\ &= \textstyle\sum_{r\in \mathrm{Fix}_S(G)} 1 + \sum_{r\in R\setminus \mathrm{Fix}_S(G)} |G/\mathrm{Stab}_G(r)| \\ &= |\mathrm{Fix}_S(G)| + \textstyle\sum_{r\in R\setminus \mathrm{Fix}_S(G)} |G/\mathrm{Stab}_G(r)|. \end{aligned}$$

□

Corollary 2.10.18 (Class Equation)**.** *Let G be a group and R be a set of representatives for the conjugacy classes of G. Then*

$$G = \sum_{r\in R} |G/\,\mathrm{C}_G(r)| = |\,\mathrm{Z}(G))| + \sum_{r\in R\setminus \mathrm{Z}(G)} |G/\,\mathrm{C}_G(r)|$$

Proof. Let $\diamond$ be the action of G on G be conjugation. Then

$$\mathrm{Fix}_G^{\diamond}(G) = \{g \in G \mid {}^h g = g \text{ for all } h \in G\} = \{g \in G \mid hg = gh \text{ for all } h \in G\} = \mathrm{Z}(G)$$

and by 2.10.8(1) $\mathrm{Stab}_G(a) = \mathrm{C}_G(a)$. So the Class Equation follows from the orbit equation. □

Example 2.10.19. By 2.7.5 has three conjugacy classes corresponding to the cyles types 1^3, $1^1 2^1$ and 3^1. So $R = \{(1), (13), (123)\}$ is a set of representatives for the conjugacy class of Sym(3). A straight forward calculation shows that

$$\mathrm{C}_{\mathrm{Sym}(3)}((1)) = \mathrm{Sym}(3), \quad \mathrm{C}_{\mathrm{Sym}(3)}((13)) = \{(1), (13)\}, \quad C_{\mathrm{Sym}(3)}((123)) = \{(1), (123), (132)\}$$

The orders of these centralizers are

$$6, 2, 3.$$

Sym(3) has order 6 and since $|G/C_G(r)| = \frac{|G|}{|C_G(r)|}$ the class equation now says

$$6 = \frac{6}{6} + \frac{6}{2} + \frac{6}{3} = 1 + 2 + 3$$

These Orbit Equation become particular powerful if G is a finite *p-group*:

Definition 2.10.20. *Let G be finite group and p a prime. Then G is called a* p-group *provided that that is $|G| = p^k$ for some $k \in \mathbb{N}$.*

Proposition 2.10.21 (Fixed-Point Equation)**.** *Let p be a prime and P a p-group acting on a finite set S. Then*

$$|S| \equiv |\mathrm{Fix}_S(P)| \pmod{p}.$$

Proof. Let R be a set of representatives the orbits of P on S and $r \in R \setminus \mathrm{Fix}_S(P)$. Then $\mathrm{Stab}_P(r) \lneq P$. By Lagrange's Theorem $|P/\mathrm{Stab}_P(r)|$ divides $|P|$. Since $|P|$ is a power of p and $|P/\mathrm{Stab}_P(r)| \neq 1$ we get

$$|P/\mathrm{Stab}_P(r)| \equiv 0 \pmod{p}.$$

So by the Orbit Equation 2.10.17

$$|S| = |\mathrm{Fix}_S(P)| + \sum_{r \in R \setminus \mathrm{Fix}_S(P)} |P/\mathrm{Stab}_P(r)| \equiv |\mathrm{Fix}_S(P) \pmod{p}$$

□

Example 2.10.22. Let $\mathcal{E} = (\mathcal{P}, \mathcal{L}, \mathcal{R})$ be a projective plane of order two

1. Let T be a 2-group and S a finite T-set with $|T|$ odd. By the fixed-point equation:

   $$|S| \equiv \mathrm{Fix}_S(T) \pmod{2}$$

   and so $|\mathrm{Fix}_S(T)|$ is odd. In particular, $|\mathrm{Fix}_S(T)| \neq 0$ and so

   $$\mathrm{Fix}_S(T) \neq \emptyset$$

2. Let $t \in \mathrm{Aut}(\mathcal{E})$ with $|t| = 2$. Then $T = \langle t \rangle$ has order two. Since $|\mathcal{P}| = 7$ we conclude from (1) that

   $$\mathrm{Fix}_S(t) = \mathrm{Fix}_S(T) \neq \emptyset$$

   So every element of order 2 in $\mathrm{Aut}(\mathcal{E})$ fixes at least point.

3. Let $T \leq \mathrm{Aut}(\mathcal{E})$ with $|T| = 8$. Since the number of points is odd, (1) implies that T fixes a point P. Also the number of lines is odd and so T also fixes a line l.

   Suppose that P is not incident with l. Let A, B be distinct points on l. Then (P, A, B) is a non-collinear triple. If $\alpha \in T$, then $\alpha(P) = P$, $\alpha(l) = l$, $\alpha(A)$ is one of the three points on l, and $\alpha(B)$ is one of the two points on l distinct from $\alpha(A)$. So there are only six choices for the triple $(\alpha(P), \alpha(A), \alpha(B)$ and so bt 2.3.1 implies that $|T| \leq 6$, a contradiction. Hence P is incident with l. By Homework 2#7 $\mathrm{Stab}_{\{Aut(\mathcal{E})|}(\{P, l\})$ has order eight and so

   $$T = \mathrm{Stab}_{\mathrm{Aut}(\mathcal{E})}(\{P, l\})$$

4. By definition of $\mathrm{Aut}(\mathcal{E})$, if $(P,l) \in \mathcal{R}$ and $\alpha \in \mathrm{Aut}(\mathcal{E})$, the $(\alpha(P), \alpha(l) \in \mathcal{R}$. So $\mathrm{Aut}(\mathcal{E})$ acts on $\mathcal{R}$. Let $(P,L) \in \mathcal{R}$. By Homework 2#7 $\mathrm{Stab}_{\mathrm{Aut}(\mathcal{E})}((P,l))$ has order eight. Let O be the orbit of $\mathrm{Aut}(\mathcal{E})$ on $\mathcal{R}$ containing (P,l). Then

$$|O| = |\mathrm{Aut}(\mathcal{E})/\mathrm{Stab}_{\mathrm{Aut}(\mathcal{E})}((P,l))| = \frac{168}{8} = 21.$$

On the otherhand, $\mathcal{E}$ has seven lines and each line is incident with 3 points and so $|\mathcal{R}| = 21$. Hence $O = \mathcal{R}$ and $\mathrm{Aut}(\mathcal{E})$ acts tranistively on $\mathcal{R}$.

Let $T \leq \mathrm{Aut}(\mathcal{E})$ be 2-group. Then by (1), T fixes some $(P,l) \in \mathcal{R}$ and so

$$T \leq \mathrm{Stab}_{\mathrm{Aut}(\mathcal{E})}((P,l))$$

Definition 2.10.23. *Let H be a group and $H \subseteq G$. Then $\mathrm{N}^*_G(H) = \{a \in G \mid H \subseteq {}^aH\}$. $\mathrm{N}^*_G(H)$ is called the weak normalizer of H in G.*

Lemma 2.10.24. *Let G be a groups and $H \subseteq G$.*

(a) $\mathrm{N}^*_G(H)$ is a submonoid of G and if $H \leq G$, then $H \subseteq \mathrm{N}^*_G(H)$.

(b) Let $g \in G$. Then $g \in \mathrm{N}_G(H)$ if and only if $\{g, g^{-1}\} \subseteq \mathrm{N}^*_G(H)$.

(c) Suppose $H \leq G$. Let $a \in G$ and let $\diamond$ be the action G on G/H by left multiplication. Then

$$\mathrm{Stab}^{\diamond}_G(aH) = {}^aH \quad \textit{and} \quad \mathrm{Fix}^{\diamond}_{G/H}(H) = \mathrm{N}^*_G(H)/H.$$

(d) Let $\Box$ be the action og G on the subsets of G. Then $\mathrm{Stab}^{\Box}(H) = \mathrm{N}_G(H)$.

(e) If H is finite, then $\mathrm{N}^*_G(H) = \mathrm{N}_G(H)$.

Proof. (a) We have ${}^eH = H$ and if $a, b \in \mathrm{N}^*(H)$, then $H \subseteq {}^bH$ and so also ${}^aH \subseteq {}^{ab}H$. Since $H \subseteq {}^aH$ this implies $H \subseteq {}^{ab}H$ and $ab \in \mathrm{N}^*_G(H)$.

If $H \leq H$ then ${}^hH = H$ for all $h \in H$ and so $H \subseteq \mathrm{N}^*_G(H)$.

(b) ${}^gH = H$ if and only if ${}^gH \subseteq H$ and $H \subseteq {}^gH$ and if and only if $H \subseteq {}^{g^{-1}}H$ and $H \subseteq {}^gH$.

(c) Let $a, g \in G$. Then $gH = H$ if and only if $g \in H$. Hence $\mathrm{Stab}^{\diamond}_G(H) = H$ and so by 2.10.14(c) $\mathrm{Stab}^{\diamond}_G(aH) = {}^a\mathrm{Stab}^{\diamond}_G(H) = {}^aH$. Note that H fixes aH if and only if $H \subseteq \mathrm{Stab}^{\diamond}_G(aH)$. That is if and only if $H \leq {}^aH$ and if and only if $a \in \mathrm{N}^*_G(H)$. Thus (c) holds.

(d) $g \in \mathrm{N}_G(H)$ if and only if ${}^gH = H$ and if and only if $g \in \mathrm{Stab}^{\Box}_G(H)$.

(e) As conjugation is an bijection, $|H| = |{}^gH|$. So for finite H, $H \leq {}^gH$ if and only if $H = {}^gH$. □

Lemma 2.10.25. *Let P be a non-trivial p-group.*

(a) $\mathrm{Z}(P)$ is non-trivial.

(b) If $H \lneq P$ then $H \lneq \mathrm{N}_P(H)$.

Proof. (a) Consider first the action by conjugation. By 2.10.21

$$0 \equiv |P| \equiv |\,\mathrm{Z}(P)| \pmod{p}.$$

Thus $|\,\mathrm{Z}(P)| \neq 1$. (b) Consider the action of H on P/H. By 2.10.24 and 2.10.21

$$0 \equiv |P/H| \equiv |\,\mathrm{N}_P(H)/H)| \pmod{p}.$$

So $|\,\mathrm{N}_P(H)/H| \neq 1$. □

Lemma 2.10.26. *Let p be a prime and P a p-group.*

(a) Let $H \leq P$. Then there exists $n \in \mathbb{N}$ and for $0 \leq i \leq n$, $H_i \leq P$ with

$$H = H_0 \trianglelefteq H_1 \trianglelefteq H_2 \trianglelefteq \ldots \trianglelefteq H_{n-1} \trianglelefteq H_n = P$$

and $|H_i/H_{i-1}| = p$ for all $1 \leq i \leq n$.

(b) Let m be a divisor of $|P|$. Then P has a subgroup of order m.

Proof. (a) The proof is by induction on $|P/H|$. If $|P/H| = 1$, then $P = H$ and (a) holds with $n = 0$ and $H_0 = H = P$. So suppose $H \neq P$. The by 2.10.25(b), $H \ntrianglelefteq \mathrm{N}_P(H)$. Hence there exits $e \neq x \in \mathrm{N}_P(H)/H$. Let $|x| = p^l$ and put $y = x^{p^{l-1}}$. Then $|y| = p$. By the Correspondence Theorem (Homework 4#1), there exists $H_1 \leq N_G(H)$ with $H_1/H = \langle y \rangle /$ Then $H \trianglelefteq H_1$ and $|H_1/H| = |\langle y \rangle| = |y| = p$. Since $|P/H_1| < |P/H|$ (a) now follwos by induction.

(b) Apply (a) with $H = \{e\}$. Then $|H_i| = p^i$ and (b) holds. □

As a further example how actions an set can be used we give a second proof that $\mathrm{Sym}(n)$ has normal subgroup of index two. For this we first establish the following lemma.

Lemma 2.10.27. [**equivrep**] *Let Δ be a finite set and $\sim$ a non-trivial equivalence relation on Δ so that each equivalence class has size at most 2. Let*

$$\Omega = \{R \subseteq \Delta \mid R \text{ contains exactly one element from each equivalence class of } \sim\}.$$

Define the relation $\approx$ on Ω by $R \approx S$ if and only if $|R \setminus S|$ is even. Then $\approx$ is an equivalence relation and has exactly two equivalence classes.

Proof. For $d \in \Delta$ let $\tilde{d}$ be the equivalence class of $\sim$ containing d and let $\tilde{\Delta}$ be the set of equivalence classes. Let $A, B \in \Omega$ and define

$$\tilde{\Delta}_{AB} = \{X \in \tilde{\Delta} \mid A \cap X \neq B \cap X\}.$$

Let $d \in A$. Then $d \notin B$ if and only if $A \cap \tilde{d} \neq B \cap \tilde{d}$. So $|A \setminus B| = |\tilde{\Delta}_{AB}|$ and

$$A \approx B \Longleftrightarrow \tilde{\Delta}_{AB} \text{ is even}$$

In particular, $\approx$ is reflexive and symmetric. Let $R, S, T \in \Omega$. Let $X \in \tilde{\Delta}$. Then $X \cap R \neq X \cap T$ exactly if either $X \cap R \neq X \cap S = X \cap T$ or $X \cap R = X \cap S \neq X \cap T$.

Thus

$$\tilde{\Delta}_{RT} = (\tilde{\Delta}_{RS} \setminus \tilde{\Delta}_{ST}) \cup (\tilde{\Delta}_{ST} \setminus \tilde{\Delta}_{RS})$$

Hence

$$(*) \quad |\tilde{\Delta}_{RT}| = |\tilde{\Delta}_{RS}| + |\tilde{\Delta}_{ST}| - 2|\tilde{\Delta}_{RS} \cap \tilde{\Delta}_{ST}|.$$

If $R \approx S$ and $S \approx T$, the right side of $(*)$ is an even number. So also the left side is even and $R \approx T$.

So $\approx$ is an equivalence relation. Let $R \in \Omega$. As $\sim$ is not trivial there exist $r, t \in \Delta$ with $r \sim t$ and $r \neq t$. Exactly one of r and t is in R. Say $r \in R$. Let $T = (R \cup \{t\}) \setminus \{r\}$. Then $T \in \Omega$ and $|T \setminus R| = 1$. Thus R and T are not related under $\approx$. Let $S \in \Omega$. Then the left side of $(*)$ odd and so exactly one of $|\tilde{\Delta}_{RS}|$ and $|\tilde{\Delta}_{ST}|$ is even. Hence $S \approx R$ or $S \approx T$. Thus $\approx$ has exactly two equivalence classes and all the parts of the lemma are proved. □

Back to $\mathrm{Sym}(n)$. Let $\Delta = \{(i,j) \mid 1 \leq i, j \leq n, i \neq j\}$. Then $\mathrm{Sym}(n)$ acts on Δ. Define $(i,j) \sim (k,l)$ iff $(k,l) = (i,j)$ or $(k,l) = (j,i)$. Define Ω as in the previous lemma. Clearly $\mathrm{Sym}(n)$ acts on Ω and also on $\Omega/\approx$ (the set of equivalence classes of " $\approx$ "). Let $R = \{(i,j)\, 1 \leq i < j \leq n\}$. Then $R \in \Omega$. The 2-cycle $(1,2)$ maps R to $(R \cup \{(2,1)\}) \setminus \{(1,2)\}$. Thus R and $(1,2)R$ are not related under $\approx$ and so $\mathrm{Sym}(n)$ acts non trivially on $\Omega/\approx$, which is a set of size 2. The kernel of the action is a normal subgroup of index two.

The following lemma is an example how the actions on a subgroup can be used to identify the subgroup.

Lemma 2.10.28. *Let n be an integer with $n \geq 5$. Let $G = \mathrm{Sym}(n)$ or $\mathrm{Alt}(n)$ and let $H \leq G$ with $|G/H| = n$. Put $I = |G/H$ and $i = H \in I$. Let $\diamond$ be the action of G on I be left multiplication. Then G acts faithfully on I, $|I| = n$ and*

(a) If $G = \mathrm{Sym}(n)$, then $G^\diamond = \mathrm{Sym}(I)$, $H \cong H^\diamond = \mathrm{Stab}_{\mathrm{Sym}(I)}(i) \cong \mathrm{Sym}(n-1)$.

(b) If $G = \mathrm{Alt}(n)$, then $G^\diamond = \mathrm{Alt}(I)$, $H \cong H^\diamond = \mathrm{Stab}_{\mathrm{Alt}(I)}(i) \cong \mathrm{Alt}(n-1)$.

Proof. We will write X_Y for $\mathrm{Stab}_X(Y)$. Let $g \in G_I$. Then $gH = H$ and so $g \in H$ and $G_I \leq H$. Thus $|G/G_I| \geq |G/H| = n > 2$ and so $G_I \neq \mathrm{Alt}(n)$ and $G_I \neq \mathrm{Sym}(n)$. By 2.7.12 and 2.7.11 the only normal subgroups of G are $\{e\}, \mathrm{Alt}(n)$ and $\mathrm{Sym}(n)$. By 2.10.6(a), $G_I \trianglelefteq G$ and so $G_I = \{e\}$, G acts faithfully on I and $G \cong G^\diamond$. In particular, $|G^\diamond| = |G|$. Since $|I| = |G/H| = n$ we have $\mathrm{Sym}(I) \cong \mathrm{Sym}(n)$ and so $|\mathrm{Sym}(I)/G^\diamond| \leq 2$. Therefore $G^\diamond \trianglelefteq G$ and by 2.7.12 $G^\diamond = \mathrm{Sym}(I)$ in (a) and $G^\diamond = \mathrm{Alt}(I)$ in (b) Note that $H^\diamond$ fixes an element i in I, namely $i = H$. Thus $H^\diamond \leq \mathrm{Sym}(I)_i$.

Suppose $G = \mathrm{Sym}(n)$. Then $|\mathrm{Sym}(I)_i| = |Sym(n-1)| = |H| = |H^\diamond|$ and so

$$H \cong H^\diamond = \mathrm{Sym}(I)_i \cong \mathrm{Sym}(n-1)$$

Suppose $G = \mathrm{Alt}(n)$. Then $|\mathrm{Sym}(I)_i / H^\diamond| = 2$, $H^\diamond \trianglelefteq \mathrm{Sym}(I)_i \cong \mathrm{Sym}(n-1)$ and so by 2.7.12 $H \cong H^\diamond = \mathrm{Alt}(I)_i \cong \mathrm{Alt}(n-1)$. □

Example 2.10.29. The goal of this example is to find a subgroup of Sym(6) whihc is isomorphic to Sym(5) but acts tranistively on $\{1,\dots,6\}$. For this let J be the set of subgroups of order five in Sym(5). Let $F \in J$ and $(1) \neq f \in F$. Then $|f|$ divides $\mathbb{F}| = 5$ and so $|f| = 5$, $F = \langle f\rangle$ and any f is a five cycle. So $f = (abcde)$ for some pairwise a, b, c, d, e. Since thare are 5! choices for a, b, c, d, e but $(abcde) = (bcdea)\dots(eabcbde)$ there are $\frac{5!}{5} == 24$ 5-cycle in Sym(4). Each $F \in J$ contains four 5-cycles and $F_1 \cap F_2 = \{(1)\}$ for $F_1 \neq F_2 \in J$. Thus $|J| = \frac{24}{4} = 6$. For $i = 1, 2$ let Let $F_i \in J$ and $(1) \neq f_i\mathbb{F}$. By 2.7.5(b), $f_2 = {}^g f_1$ for some $g \in \text{Sym}(5)$. Thus

$$ {}^g F_1 = {}^g\langle f_1\rangle = \langle {}^g f_1\rangle = f_2\rangle = F_2$$

and so Sym(5) acts tranistively on J. Note that $\text{Sym}(5)_J \leq \text{Sym}(5)_F$ for any $F \in J$. Since $|Sym(5)/\text{Sym}(5)_F = | = |J| = 6$, $\text{Sym}(5)_F$ has order 24 and so $|\text{Sym}(5)_J| \leq 24$. Since $\text{Sym}(5)_J$ is normal in Sym(5) and the only normal subgroups of Sym(5) are $\{(1)\}$, Alt(5) and Sym(5), we conclude that $\text{Sym}(5)_J = \{(1)\}$. So Sym(5) acts faithfully on J. Thus Sym(J) contains a subgroup isomorphic to Sym(5) and acting transitively on J. Since $|J| = 6$ it follows that Sym(6) contains a subgroup H isomorphic to Sym(5) which acts transitively on $\{1, 2, \dots, 6\}$.

On the other hand by 2.10.28 H fixes the point $i = H$ in the Sym(6)-set $I = \text{Sym}(6)/H$. This seems to be contradictory, but isn't. The set I is a set with six elements on which Sym(6) acts but it is not isomorphic to the set $\{1, 2, 3, 4, 5, 6\}$. So Sym(6) has two non-isomorphic action on sets of size six. Indeed this also follows from Homework 4#4: Let $\alpha : \text{Sym}(6) \to \text{Sym}(6)$ be an isomorphism which is not inner. Let $\diamond_\alpha$ be the corresponding action of Sym(6) on $\{1,\dots,6\}$. It is fairly easy to see that since α is not inner, $\diamond_\alpha$ is not isomorphic the standard action of Sym(6) on $\{1,\dots,6\}$. (see Homework 5#1).

2.11 Sylow p-subgroup

Hypothesis 2.11.1. *Throughout this section G is a finite group and p a prime.*

Definition 2.11.2. *A* p-subgroup *of G is a subgroup $P \leq G$ which is a p-group. A* Sylow p-subgroup *S of G is a maximal p-subgroup of G. That is S is a p-subgroup of G and if $S \leq Q$ for some p-subgroup Q, then $S = Q$. Let* $\text{Syl}_p(G)$ *be the set of all Sylow p-subgroups of G.*

Let $n \in \mathbb{Z}^+$ and $n = p^k m$ with $k \in \mathbb{N}$, $m \in \mathbb{Z}^+$ and $p \nmid m$, then $n_p = p^k$. n_p is called the p-part of n. Often a Sylow p-subgroup is defined to be a subgroup of order $|G|_p$. This turns out to be equivalent to our definition (see 2.11.3(b) and 2.11.7(c)), but I prefer the above definition for two reason: 1. It is easy to see that Sylow p-subgroups exists (see the next lemma). 2. The given definition also makes sense for infinite groups (allthough infinite groups may not have a Sylow p-subgroup).

Lemma 2.11.3. *(a) Any p-subgroup of G is contained in a Sylow p-subgroup of G. In particular,* $\text{Syl}_p(G)$ *is not empty.*

(b) Let $S \leq G$ with $|S| = |G|_p$. Then S is a Sylow p-subgroup of G.

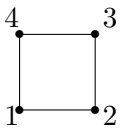
Proof. (a) Let P be a p-sugroups and let S be a p-subgroup of G such that $|S|$ is maximal with respect to $P \leq S$. We claim that $S \in \mathrm{Syl}_p(G)$. For this let Q be a p-subgroup of G with $S \leq Q$. Then also $P \leq Q$ and so by maxiality of $|S|$, $|Q| \leq |S|$. Since $S \leq Q$ this gives $S = Q$ and so $S \in \mathrm{Syl}_p(G)$.

In particular, $\{e\}$ is contained in a Sylow p-subgroup of G and so $\mathrm{Syl}_p(G) \neq \emptyset$.

(b) Let Q be a p-subgroup of G with $S \leq Q$. By Lagrange's, $|Q|$ divides $|G$. Since $|Q|$ is a power of p, $|Q|$ divides $|G|_p = |S|$. Thuse $|Q| \leq |S|$ and $S = Q$. So $S \in \mathrm{Syl}_p(G)$. □

Example 2.11.4. 1. Let $G = \mathrm{Sym}(5)$. Then $|G| = 5! = 120 = 2^3 \cdot 3 \cdot 5$. Thus by 2.11.3(b),

$$\langle (123) \rangle \in \mathrm{Syl}_3(G)$$
$$\langle (12345) \rangle \in \mathrm{Syl}_5(G)$$
$$\mathrm{Dih}_8 \in \mathrm{Syl}_2(G)$$

Here $\mathrm{Dih}_8 = \langle (14)(23), (13) \rangle$ is the automorphism groups of the square 

2. $\mathcal{E}$ be a projective plane of order two and $G = \mathrm{Aut}(\mathcal{E})$. Then $|G| = 168 = 2^3 \cdot 3 \cdot 7$. Let

Proposition 2.11.5 (Cauchy)**.** *If p divides $|G|$, then G has an element of order p*

Proof. Let $X = \langle x \rangle$ be any cyclic group of order p. Then X acts on G^p by

$$x^k * (a_1, \ldots, a_p) = (a_{1+k}, a_{2+k} \ldots, a_p, a_1, \ldots, a_k).$$

Consider the subset

$$S = \{(a_1, \ldots, a_p) \in G^p \mid a_1 a_2 \ldots a_p = e\}.$$

Note that we can choose the first $p-1$ coordinates freely and then the last one is uniquely determined. So $|S| = |G|^{p-1}$.

We claim that S is X invariant. For this note that

$$(a_1 a_2 \ldots a_p)^{a_1^{-1}} = a_1^{-1}(a_1 \ldots a_p)a_1 = a_2 \ldots a_p a_1.$$

Thus $a_1 a_2 \ldots a_p = e$ if and only if $a_2 \ldots a_p a_1 = e$. So X acts on S.

From 2.10.21 we have

$$|S| \equiv |\mathrm{Fix}_S(X)| \pmod{p}$$

As p divides $|G|$, it divides $|S|$ and so also $|\mathrm{Fix}_S(X)|$. Hence there exists some $(a_1, a_2, \ldots a_p) \in \mathrm{Fix}_S(X)$ distinct from $(e, e, \ldots, e)$. But being in $\mathrm{Fix}_S(X)$ just means $a_1 = a_2 = \ldots a_p$. Being in S implies $a_1^p = a_1 a_2 \ldots a_p = e$. Therefore a_1 has order p. □

The following easy lemma is crucial for our approach to the theory of Sylow p-subgroups.

Lemma 2.11.6. *Let $P \in \mathrm{Syl}_p(G)$ and $\alpha \in \mathrm{Aut}(G)$. Then $\alpha(P) \in \mathrm{Syl}_p(G)$. In particular, G acts on $\mathrm{Syl}_p(G)$ by conjugation.*

Proof. Since α is an bijection, $|P| = |\alpha(P)|$ and so $\alpha(P)$ is a p-group. Let Qbe a p-subgroup of G with $\alpha(P) \leq Q$. Then $\alpha^{-1}(Q)$ is a p-subgroup of G with $P \leq \alpha^{-1}(Q)$ and the maximality of P implies $P = \alpha^{-1}(Q)$. Thus $\alpha(P) = Q$ and $\alpha(P)$ is indeed a maximal p-subgroup of G.

Let $g \in G$. Then ${}^gP = \mathrm{i}_g(P) \in \mathrm{Syl}_p(G)$. Thus $\mathrm{Syl}_p(G)$ is subset of $\mathcal{P}(G)$ invariant under the action by conjugation. Therefore G acts on $\mathrm{Syl}_p(G)$ be conjugation. □

Theorem 2.11.7 (Sylow's Theorem)**.** *Let G be a finite group, p a prime and $P \in \mathrm{Syl}_p(G)$.*

(a) All Sylow p-subgroups are conjugate in G.

(b) $|\mathrm{Syl}_p(G)| = |G/\mathrm{N}_G(P)| \equiv 1 \pmod{p}$.

(c) $|P| = |G|_p$.

Proof. Let $\mathcal{S} = {}^GP :== \{{}^gP \mid g \in G\}$. So $\mathcal{S}$ is the set of Sylow p-subgroups conjugate to P. First we show

1°. *P has a unique fixed-point on $\mathcal{S}$ and on $\mathrm{Syl}_p(G)$, namely P itself*

Indeed, suppose that P fixes $Q \in \mathrm{Syl}_p(G)$. Then $P \leq \mathrm{N}_G(Q)$ and PQ is a subgroup of G. Now $|PQ| = \frac{|P||Q|}{|P \cap Q|}$ and so PQ is a p-group. Hence by maximality of P and Q, $P = PQ = Q$.

2°. $\mathcal{S} \equiv 1 \pmod{p}$.

By (1°) $\mathrm{Fix}_{\mathcal{S}}(P) = 1$ and by Fixed-Point Formula 2.10.21 $|\mathcal{S}| \equiv |\mathrm{Fix}_{\mathcal{S}}(G)| \pmod{p}$. So (2°) holds.

3°. $\mathrm{Syl}_p(G) = \mathcal{S}$ *and so (a) holds*

Let $Q \in \mathrm{Syl}_p(G)$. Then $|\mathrm{Fix}_{\mathcal{S}}(Q)| \equiv |\mathcal{S}| \equiv 1 \pmod{p}$. Hence Q has a fixed-point $T \in \mathcal{S}$. By (2°) applied to Q, this fixed-point is Q. So $Q = T \in \mathcal{S}$.

4°. *(b) holds.*

By (2°) and (5°) $|\mathrm{Syl}_p(G)| = |\mathcal{S}| \equiv 1 \pmod{p}$. Note that $\mathrm{N}\, N_G(P)$ is the stabilizer of P in G with respect to conjugation. As G is transitive on $\mathcal{S}$ we conclude from 2.10.14(d) that $|\mathcal{S}| = |G/\mathrm{N}_G(P)|$. Thus (b) holds.

5°. *p does not divides $|\mathrm{N}_G(P)/P|$.*

Suppose it does. Then by Cauchy's theorem there exists $1 \neq a \in \mathrm{N}\,N_G(P)/P$ with $|a| = p$. Put $A = \langle a \rangle$. Then $\mathbb{A}| = p$. By the Correspondence Theorem (Homework 4#1), $A = Q/P$ for some $P \leq Q \leq \mathrm{N}_G(P)$. Since $|Q| = |Q/P| \cdot |P| = |A|cdot|P| = p|P|$, Q is a p-group with $P \lneq Q$, a contradiction to the maximality of P.

6°. *(c) holds.*

By (b) and (5°), p divides neither $|G/\,\mathrm{N}_G(P)|$ nor $|\,\mathrm{N}_G(P)/P|$. Since

$$|G| = |G/\,\mathrm{N}\,N_G(P)| \cdot |\,\mathrm{N}_G(P)/P| \cdot |P|$$

we get that p does not divide $|G/P|$. Hence $|G|_p$ divides $|P|$. By Lagrange's $|P|$ divides $|G|$ and so also $|G|_p$. Thus $|P| = |G|_p$. and (c) holds. □

Corollary 2.11.8. *Let $P \in \mathrm{Syl}_p(G)$ and $M \trianglelefteq G$.*

(a) Let Q be a p-subgroup of G. Then $Q \in \mathrm{Syl}_p(G)$ if and only if $|Q| = |G|_p$ and if and only if p does not divide $|G/Q|$.

(b) $P \trianglelefteq G$ if and only if P is the unique Sylow p-subgroup of G.

(c) $P \cap M \in \mathrm{Syl}_p(M)$ and $PM/M \in \mathrm{Syl}_p(G/M)$.

Proof. (a) Since $|Q|$ is a power of p, $|Q| = |G|_p$ if and only if p does not divide $\frac{|G|}{|Q|}$. If $|Q| = |G|_p$ then by 2.11.3(b), $Q \in \mathrm{Syl}_p(G)$ and if $Q \in \mathrm{Syl}_p(G)$ then by 2.11.7(c), $|Q| = |G|_p$.

(b) Since all Sylow p-subgroups are conjugate, $\mathrm{Syl}_p(G) = \{{}^gP \mid g \in G\}$. Hence $\mathrm{Syl}_p(G) = \{P\}$ if and only if $P = {}^gP$ for all $g \in G$.

(c) By (a) p does not divide $|G/P|$. Also

$$\frac{|G|}{|P|} = \frac{|G|}{|MP|}\frac{|MP|}{P} = |G/M/PM/M| \cdot |M/P \cap M|$$

and so neither

$|G/M/|PM/M|$ nor $|M/P \cap M|$ are divisible by p. So (c) follows from (a) □

As an application of Sylow's theorem we will investigate groups of order 12,15,30 and 120. We start with a couple of general observation.

Lemma 2.11.9. *Let $P \in \mathrm{Syl}_p(G)$ and put $N = \mathrm{Stab}_G(\mathrm{Syl}_p(G))$.*

(a) Let $s_p := |\mathrm{Syl}_p(G)|$. Then s_p divides $\frac{|G|}{|G|_p}$,$s_p \equiv 1 \pmod p$ and $s_p|G_p|$ divides $|G|$.

(b) $P \trianglelefteq PN$. In particular, P is the unique Sylow p-subgroup of PN.

(c) $N \cap P \trianglelefteq G$. In particular $P \leq N$ if and only if $P \trianglelefteq G$ and if and only if $N = G$.

(d) The map

$$\mathrm{Syl}_p(G) \to \mathrm{Syl}_p(G/N) \quad Q \to QN/N$$

is a bijection.

Proof. (a) By 2.11.7(b), $s_p = |G/\mathrm{N}_G(P)| \equiv 1 \pmod{p}$. Also

$$\frac{|G|}{|G|_p} = \frac{|G|}{|P|} = \frac{|G|}{|\mathrm{N}_G(P)|}|\frac{|\mathrm{N}_G(P)|}{|P|} = s_p \cdot |\frac{|\mathrm{N}_G(P)|}{|P|}|$$

So s_p divides $\frac{|G|}{|G|_p}$.

(b) Since $N \leq \mathrm{N}_G(P)$, $P \trianglelefteq PN$. So by 2.11.8(b) (applied to PN in place of G), PN has exactly one Sylow p-subgroup.

(c) Since $P \trianglelefteq PN$, $N \cap P \trianglelefteq N$. So by 2.11.8(b), $N \cap P$ is the only Sylow p-subgroup of N. Let $g \in G$. Then by 2.11.6 ${}^gN \cap P$ is a Sylow p-subgroup of N and so equal to $N \cap P$. Thus $N \cap P \trianglelefteq G$.

(d) By 2.11.8(c) $PN/N \in \mathrm{Syl}_p(G/N)$.

Since every Sylow p-subgroup of G/N is of the form $(PN/N)^{gN} = P^gN/N$ the map is onto. Suppose that $PN/N = QN/N$. Then $Q \leq PN$ and so by (b) $Q = P$. Thus the map is also one to one. □

Lemma 2.11.10. *Let G be a finite group of order $2n$ with n odd. Then G index a normal subgroup of index 2.*

Proof. By Cayley's Theorem 2.10.4(1), G is isomorphic to $G^{\cdot}$, (the image of G in $\mathrm{Sym}(G)$ under the homomorphism $\Phi^{\cdot}$ corresponding the action $\cdot$ of G and G by left multiplication. Let $t \in G$ be an element of order 2. Since $tg \neq g$ for all $g \in G$, t and so also $t^{\cdot} = \Phi^{\cdot}(t)$ has no fixed-points on G. Hence $t^{\cdot}$ has n-cycles of length 2 and so $t^{\cdot}$ is an odd permutation. Thus $G^{\cdot} \not\leq \mathrm{Alt}(G)$ and $G^{\cdot} \cap \mathrm{Alt}(G)$ is normal subgroup of index 2 in $G^{\cdot}$. □

Lemma 2.11.11. *(a) Let G be a group of order 12. Then either G has unique Sylow 3-subgroup or $G \cong \mathrm{Alt}(4)$.*

(b) Let G be group of order 15. Then $G \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$.

(c) Let G be a group of order 30. Then G has a unique Sylow 3-subgroup and a unique Sylow 5-subgroup.

Proof. (a) By 2.11.9a the number of Sylow 3 subgroups divides $\frac{12}{3}$ is 1 (mod 3). Thus $|\mathrm{Syl}_3(G)| = 1$ or 4. In the first case we are done. In the second case let $N = \mathrm{Stab}_G(\mathrm{Syl}_3(G))$. By 2.11.9, G/N still has 4 Sylow 3-subgroups. Thus $|G/N| \geq 4 \cdot 3 = 12 = |G|$, $N = \{e\}$ and G is isomorphic to a subgroup of order 12 in $\mathrm{Sym}(4)$. Such a subgroup is normal and so $G \cong \mathrm{Alt}(4)$ by 2.7.12.

(b) The numbers of Sylow 5-subgroups is 1 (mod 5) and divides $\frac{15}{3} = 5$. Thus G has a unique Sylow 5-subgroup S_5. Also the number of Sylow 3 subgroups is 1 (mod 3) and divides $\frac{15}{3} = 5$. Thus G has a unique Sylow 3-subgroup S_3. Then $S_3 \cap S_5 = 1$, $|S_3S_5| = 15$ and so $G = S_3S_5$. Hence by 2.8.7

$$G \cong S_3 \times S_5 \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}/15\mathbb{Z}$$

where the latter isomorphism holds since we just proved that any group of order 15 is isomorphic to $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$.

(c) By 2.11.10 any group which as order twice an odd number has a normal subgroup of index two. Hence G has a normal subgroup of order 15. This normal subgroup contains all the Sylow 3 and Sylow 5-subgroups of G and so (c) follows from (b). □

Lemma 2.11.12. *Let G be a group of order 120. Then one of the following holds:*

(a) G has a unique Sylow 5-subgroup.

(b) $G \cong \mathrm{Sym}(5)$.

(c) $|\,\mathrm{Z}(G)| = 2$ and $G/\,\mathrm{Z}(G) \cong \mathrm{Alt}(5)$.

Proof. Let $P \leq \mathrm{Syl}_5(G)$ and put $I = \mathrm{Syl}_5(G)$.

If $|I| = 1$, (a) holds.

So suppose that $|I| > 1$. Then by2.11.9(a), $|I| \equiv 1 \pmod 5$ and $|I|$ divides $|G/P| = 24$. The numbers which are larger than 1, are less or equal to 24 and are $1 \pmod 5$ are $1, 6, 11, 16$ and 21. Of these only 6 divides 24. Thus $|I| = 6$. Let $\phi : G \to \mathrm{Sym}(I)$ be the homomorphism corresponding to the action of G on I. Put $N = \ker \phi$ and $H = \phi(G)$. Then H is subgroup of $\mathrm{Sym}(I) \cong \mathrm{Sym}(6)$ and $H \cong G/N$. By 2.11.9(d), G/N (and so also H) has exactly six Sylow 5-subgroups. In particular the order of H is a multiple of 30. By 2.11.11c, $|H| \neq 30$.

Suppose that $|H| = 120$. Then $N = 1$ and so $G \cong H$ in this case. Now $H \leq \mathrm{Sym}(I) \cong \mathrm{Sym}(6)$. Thus 2.10.28(a) implies $G \cong H \cong \mathrm{Sym}(5)$.

Suppose next that $|H| = 60$. If $H \not\leq \mathrm{Alt}(I)$, then $H \cap \mathrm{Alt}(I)$ is a group of order 30 with six Sylow 5-subgroups, a contradiction to 2.11.11. Thus $H \leq \mathrm{Alt}(I) \cong Alt(6)$. So by 2.10.28(b), $H \cong \mathrm{Alt}(5)$. Since $|N| = 2$ and $N \trianglelefteq G$, $N \leq \mathrm{Z}(G)$. Also $\phi(\mathrm{Z}(G))$ is a abelian normal subgroup of $H \cong \mathrm{Alt}(5)$ and so $\phi(\mathrm{Z}(G)) = e$. Hence $N = \mathrm{Z}(G)$ and

$$G/\,\mathrm{Z}(G) = G/N \cong H \cong \mathrm{Alt}(5).$$

□

Chapter 3

Rings

3.1 Rings

Definition 3.1.1. *A* ring *is a tuple* $(R, +, \cdot)$ *such that*

(a) $(R, +)$ *is an abelian group.*

(b) $(R, \cdot)$ *is a semigroup.*

(c) For each $r \in R$ *both left and right multiplication by* r *are homomorphisms of* $(R, +)$

3.1.2 (Ring Axioms)**.** Unwinding definitions we see that a ring is a set R together with two binary operations $+ : R \times R \to R, (a, b) \to a + b$ and $\cdot : R \times R \to R, (a, b) \to ab$ such that

(R1) $(a + b) + c = a + (b + c)$ for all $a, b, c \in R$

(R2) There exists $0_R \in R$ with $0_R + a = a = a + 0_R$ for all $a \in R$.

(R3) For each $a \in R$ there exists $-a \in R$ with $a + (-a) = 0_R = (-a) + a$.

(R4) $a + b = b + a$ for all $a, b \in R$.

(R5) $a(bc) = (ab)c$ for all $a, b, c \in R$.

(R6) $a(b + c) = ab + ac$ for all $a, b, c \in R$.

(R7) $(a + b)c = ab + ac$ for all $a, b, c \in R$.

Definition 3.1.3. *Let* R *and* S *be rings. A* ring homomorphism *is a map* $\phi : R \to S$ *such that* $\phi : (R, +) \to (S, +)$ *and* $\phi : (R, \cdot) \to (S, \cdot)$ *are homomorphisms of semigroups. A bjijetcive ring homomorphism is called a* ring isomorphism *and* R *and* S *are called* isomorphic *and we write* $R \cong S$ *if there exists a ring isomorphism from* R *to* S*.*

Note that $\phi : R \to S$ is an homomorphism if and only if $\phi(r + s) = \phi(r) + \phi(s)$ and $\phi(rs) = \phi(r)\phi(s)$ for all $r, s \in R$.

Definition 3.1.4. *Let $(R, +\cdot)$ be a ring. An* identity *in R is an element 1_R which is an identity for $\cdot$, what is $1_R r = r = r1_R$ for all $r \in R$. If there exists an identity in R we say that R is a ring with identity. R is called* commutative *if $\cdot$ is commutative, that is $rs = sr$ for all $r, s \in R$.*

In the following lemma we collect a few elementary properties of rings.

Lemma 3.1.5. *Let R be a ring.*

(a) $0a = a0 = 0$ *for all* $a \in R$

(b) $(-a)b = a(-b) = -(ab)$ *for all* $a, b \in R$.

(c) $(-a)(-b) = ab$ *for all* $a, b \in R$.

(d) $(na)b = a(nb) = n(ab)$ *for all* $a, b \in R, n \in \mathbb{Z}$.

(e) $(\sum_{i=1}^{n} a_i)(\sum_{j=1}^{n} b_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j$

Proof. This holds since since right and left multiplication by elements in R are homomorphisms of $(R, +)$. For example any homomorphism sends 0 to 0. So (a) holds. We leave the details to the reader. □

Example 3.1.6. 1. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ and $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ are rings.

2. Let A be an abelian group and $\mathrm{End}(A)$ the set of endomorphisms of A, (that is the homomorphisms from A to A). Define $(\alpha+\beta)(a) = \alpha(a)+\beta(a)$ and $(\alpha\circ\beta)(a) = \alpha(\beta(a))$. Then $(\mathrm{End}(A), +, \circ)$ is a ring called the *endomorphism ring* of A.

3. Let V be a vector space over $\mathbb{R}$. Let $\mathrm{End}_R(V)$ be set of $\mathbb{R}$-linear maps from V to V. Then $(\mathrm{End}_{\mathbb{R}}(V), +, \circ)$ is a ring called the *endomorphism ring* of V over $\mathbb{R}$.

4. Let $(A, +)$ be any abelian group. Define $\cdot_0 : A \to A, (a, b) \to 0_R$. Then $(A, +, \cdot_0)$ is a ring, called the ring on A with zero-multiplication.

5. Up to isomorphism there is unique ring with one element:

$+$	0
0	0

$\cdot$	0
0	0

6. Up to isomorphism there are two rings of order two :

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	n

Here $n \in \{0, 1$ for $n = 0$ we have a ring with zero-multiplication For $n = 1$ this is $(\mathbb{Z}/2\mathbb{Z}, +, \cdot)$.

7. Rings of order order 3 up to isomorphism:

$+$	0	1	-1
0	0	1	-1
1	1	-1	0
-1	-1	0	1

$\cdot$	0	1	-1
0	0	0	0
1	0	n	$-n$
-1	0	$-n$	n

Indeed if we define $n = 1 \cdot 1$, then $(-1) \cdot 1 = -(1 \cdot 1) = -n$. Here $n \in \{0, 1, -1\}$. For $n = 0$ this is a ring with zero multiplication. For $n = 1$ this is $(\mathbb{Z}/3\mathbb{Z}, +, \cdot)$. For $n = -1$ we see that -1 is an identity and the ring for $n = -1$ is isomorphic to the ring with $n = 1$ case under the bijection $0 \leftrightarrow 0$, $1 \leftrightarrow -1$.

At the end of this section we will generalize these argument to find all rings whose additive group is cyclic.

8. Direct products and direct sums of rings are rings. Indeed, let $(R_i, i \in I)$ be a family of rings. For $f, g \in \bigtimes_{i\in I} R_i$ define $f + g$ and fg by $(f + g)(i) = f(i) + g(i)$ and $(fg)(i) = f(i)g(i)$. With this definition both $\bigtimes_{i\in I} R_i$ and $\bigoplus_{i\in I} R_i$ are rings. each R_i as an idendity 1_i, then $(1_i)_{i\in I}$ is an indentity of $\bigtimes_{i\in I} R_i$. Note that if I is infinite

$$= (1_i)_{i\in I}$$

usually is not in $\bigoplus_{i\in I} R_i$ and so $\bigoplus_{i\in I} R_i$ usually does not have an indentity.

If each R_i is commutative then both $\bigtimes_{i\in I} R_i$ and $\bigoplus_{i\in I} R_i$ have an identity.

Definition 3.1.7. *Let R be a ring and G be semigroup. The* semigroup ring $R[G]$ *of G over R is defined as follows:*

For $g \in G$ let $R_g = R$. As an abelian group we put $R[G] = \bigoplus_{g\in G} R_g$. Define

$$(r_g)_{g\in G} \cdot (s_g)_{g\in G} = (t_g)_{g\in G}$$
$$\text{where} \quad t_g = \sum_{\{(h,l)\in G\times G | hl=g\}} r_h s_l$$

Note that since the elements in $\bigoplus_{g\in G} R_g$ have finite support all these sums are actual finite sums.

For $r \in R$ and $g \in G$ we denote $\rho_g(r)$ by rg, so (see 2.8.5)

$$(rg)_g = r \text{ and } (rg)_h = 0_R \text{ for } h \neq g$$

Lemma 3.1.8. *Let G be a semigroup and R a ring.*

(a) $(R[G], +, \cdot)$ *is a ring.*

(b) For each $a \in RG$ *ther exists uniquely determined* $r_g \in R$, $g \in G$ *with* $r_g = 0_R$ *for almost all* $g \in G$ *and*

$$a = \sum_{g \in G} r_g g$$

(c) $\sum_{g\in G} r_g g \quad + \quad \sum_{g\in G} s_g g = \sum_{g\in G}(r_g + s_g)g.$

(d) $\sum_{g\in G} r_g g \quad \cdot \quad \sum_{h\in G} s_h h = \sum_{g\in G, h\in G}(r_g s_h)gh.$

(e) If R *and* G *have an identities, then* $1_R e_G$ *is an identity in* $R[G]$.

(f) If R *and* G *are commutative,* $R[G]$ *is too.*

Proof. This is Homework 6#1. □

3.1.9 (Identities in Group Rings)**.** If $R[G]$ has an identity, then R has an identity. Indeed $r = \sum r_g g$ is an identity in $R[G]$. Let $a = \sum r_g$. We will show that a is an identity in R. Let $s \in R^{\#}$ and $g \in G$. Then

$$sh = r(sh) = \sum (r_g s)gh.$$

Summing up the coefficients we see that $s = (sum_{g\in G} r_g)s = as$. Similarly $sa = s$ and so a is an identity in R. If R has an identity 1, we identify g with $1g$.

Here is an example of a semigroup G without an identity so that (for any ring R with an identity) $R[G]$ has an identity. As a set $G = \{a, b, i\}$. Define the multiplication by

$$xy = \begin{cases} x & \text{if } x = y \\ i & \text{if } x \neq y \end{cases}$$

Then

$$(xy)z = (xy)z = \begin{cases} x & \text{if } x = y = z \\ i & \text{otherwise} \end{cases}$$

Hence the binary operation is associative and G is a semigroup. Put $r = a + b - i \in R[G]$. We claim that r is an identity. We compute $ar = ra = aa + ab - ai = a + i - i = a$, $br = rb = ba + bb - bi = i + b - i = b$ and $ir = ri = ia + ib - ii = i + i - i = i$. As $R[G]$ fulfills both distributive laws this implies that r is an identity in $R[G]$.

Suppose $R[G]$ is commutative, then

$$(rs)(gh) = (rg)(sh) = (sh)(rg) = (sr)(hg) = (rs)(hg).$$

So if $rs \neq 0$ for some $r, s \in R$ we get $gh = hg$ and G is commutative.

But if $rs = 0$ for all $r, s \in R$ then also $xy = 0$ for all $x, y \in R[G]$. So $R[G]$ is commutative, regardless whether G is or not.

Example 3.1.10. let R be a ring and G a semigroup.

1. If $G = \{e_G\}$ the $R[G] \cong R$.

2. View $\mathbb{N}$ has a monoid under addition. Then $R[\mathbb{N}]$ is called the the polynomial ring over R in one variable. To avoid confusion between the addition in $\mathbb{N}$ and R we will adopt the following conventions: We write x for the element $1 \in \mathbb{N}$ and use multiplicative notation for $\mathbb{N}$. Then $x^n = n$, $x^n x^m = n + m = x^{n+m}$ and

$$R[\mathbb{N}] = \{\textstyle\sum_{i\in\mathbb{N}} r_i i \mid r_i \in R, \text{almost all } r_i = 0_R\} = \sum_{i=0}^{m} r_i x^i \mid m \in \mathbb{N}, r_i \in R\}$$
$$(\textstyle\sum_{i=0}^{m} r_i x^i) + (\sum_{i=0}^{m} s_i x^i) = \sum_{i=0}^{m} (r_i + s_i) x^i$$
$$(\textstyle\sum_i r_i x^i) \cdot (\sum_j s_j x^j) = \sum_i \sum_j r_i s_j x^{i+j} = \sum_n (\sum_{k=0}^{n} r_k s_{n-k})\, x^k$$

   We will denote this ring by $R[x]$.

Definition 3.1.11. *Let R be a ring and $a \in R$.*

(a) $R^{\#} \setminus \{0\}$.

(b) *a is left (resp. right)* zero divisor *if $a \neq 0_R$ and there there exists $b \in R^{\#}$ with $ab = 0$ (resp. $ba = 0$). a is a zero divisor if a is a left or a right zero divisor.*

Suppose now that R has an identity.

(c) *a is called (left,right) invertible if it is (left,right) invertible in $(R, \cdot)$. An invertible element is also called a unit.*

(d) *R^* is the set of units in R.*

(e) *R is called an* integral domain *if R is commutative, $1_R \neq 0_R$ and R no zero-divisors.*

(f) *R is called a* division ring *if $1_R \neq 0_R$ and all it non-zero elements are invertible. A* field *is a commutative division ring.*

Note that a ring with identity is a zero ring (that is $R = \{0_R\}$ if and only if and only if $1_R = 0 - R$. So in (e) and (f) the condition $1_R \neq 0_R$ can be replaced to $R \neq \{0_R\}$.

Lemma 3.1.12. *Let R be a ring. Then the following statements are equivalent:*

(a) *R has no right zero-divisors.*

(b) *If $a, b \in R$ with $ab = 0_R$, then $a = 0_R$ or $b = 0_R$.*

(c) *R has no left zero-divsors.*

(d) *The* Right Cancellation Law *holds, that is*

   Whenever $a, b, c \in R$ with $c \neq 0_R$ and $ac = bc$, $a = b$.

(e) The left Left Cancellation Law *holds, that is*

Whenever $a, b, c \in R$ with $c \neq 0_R$ and $ca = cb$, $a = b$.

Clearly (a) and (b) are equivalent and similarly (b) and (c) are equivalent.

Suppose that R has no left zero-divisors and $a, b, c \in R$ with $c \neq 0_R$ and $ab = ac$. Then

$$0_R = ac - bc = (a - b)c$$

Since R has no left zero-divisors this implies $a - b = 0_R$ and so $a = b$. Thus the Right Cancellation Law holds.

Suppose the Right Cancellation Law holds and let $a, b \in R$ with $b \neq 0_R$ and $ab = 0_R$. Then $ab = 0_R = 0_R \cdot b$ and so by the Right Cancellation Law, $a = 0_R$. So R has no left zero-divisors. Thus (c) and (d) are equivalent. Similarly (a) and (e) are equivalent.

Example 3.1.13. 1. $\mathbb{R}$ is a field, $\mathbb{Q}$ and $\mathbb{C}$ are fields. $\mathbb{Z}$ is an integral domain.

2. For which $n \in \mathbb{Z}^+$ is $\mathbb{Z}_n$ an integral domain? If $n = 1$, then $\mathbb{Z}_1$ is a zero ring and so not an integral domains. So suppose $n \geq 2$. Then $1 \neq 0$ in $\mathbb{Z}_n$ and thuas Z_n is an integral domain if and only,

$$n \mid kl \Longrightarrow n \mid k \text{ or } n \mid l$$

and so if and only if n is a prime. The following lemma implies that $\mathbb{Z}/p\mathbb{Z}$ is a field for all primes p.

Lemma 3.1.14. *All finite integral domains are fields*

Proof. Let R be a finite integral domain and $a \in R^{\#}$. As R is an integral domain, multiplication by a is a one to one map from $R^{\#} \to R^{\#}$. As R is finite, this map is onto. Thus $ab = 1_R$ for some $b \in R$. Since R is commutative $ba = 1 - R$ and so all non-zero elements are invertible. □

For a ring R we define the *opposite ring* R^{op} by $(R^{\text{op}}, +^{\text{op}}) = (R, +)$, and $a \cdot^{\text{op}} b = b \cdot a$. If R and S are rings then a map $\phi : R \to S$ is called an *anti-homomorphism* if $\phi : R \to S^{\text{op}}$ is ring homomorphism. So $\phi(a + b) = \phi(a) + \phi(b)$ and $\phi(ab) = \phi(b)\phi(a)$.

Let $\text{End}(R)$ be the set of ring homomorphism. Then $\text{End}(R)$ is monoid under composition. But as the sum of two ring homomorphisms usually is not a ring homomorphism, $\text{End}(R)$ has no natural structure as a ring.

The map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \quad m \to m + n\mathbb{Z}$ is a ring homomorphism.

For $r \in R$ let $\mathcal{R}_r : R \to R, s \to sr$ and $\mathcal{L}_r : R \to R, s \to rs$. By definition of a ring $\mathcal{R}_a$ and $\mathcal{L}_r$ are homomorphisms of $(R, +)$. But left and right multiplication usually is not a ring homomorphism. The map $\mathcal{L} : R \to \text{End}((R, +)), r \to \mathcal{L}_r$ is a homomorphism but the map $\mathcal{R} : R \to \text{End}((R, +)), r \to \mathcal{R}_r$ is an anti-homomorphism. Note that if R has an identity, then both $\mathcal{R}$ and $\mathcal{L}$ are one to one.

Definition 3.1.15. *(a) Let G be a group We say that G has finite* exponent *of there exists $n \in \mathbb{Z}^+$ with $g^n = e$ for n for all $g \in G$. If G has finite exponent then exponent $\exp(G)$ of G is the smallest positive integer m with $g^m = e$ for all $g \in G$, otherwise $\exp(G) = \infty$.*

(b) Let $(R, +, \cdot)$ be a ring. If $(R, +)$ is finite then the characteristic $\operatorname{char} R$ *of R is the exponent of $(R, +)$. If $(R, +)$ has infinite exponent then* $\operatorname{char} G = 0$.

Lemma 3.1.16. *Let R be a ring with identity.*

(a) Let $n \in \mathbb{Z}$ then $n1_R = 0_R$ if and only if $nr = 0_R$ for all $r \in R$.

(b) Suppose $1_R \neq 0_R$ and that R has no zero-divisors. Then $\operatorname{char} R$ *is 0 or a prime.*

Proof. (a) If $nr = 0_R$ then clearly $n1_R = 0_R$. So suppose $n1_R = 0_R$. Then for all $r \in R$

$$nr = n(1_R r) = (n1_R)r = 0_R r = 0_R$$

(b) Suppose $n := \operatorname{char} R \neq 0$. If $n = 1$, then $0_R = 1 \cdot 1_R = 1_R$, contrary to the assumptions. So $n > 1$. Let $n = st$ with $s, t \in \mathbb{Z}^+$. Then

$$0_R = n1_R = (st)1_R = st1_R1_R = (s1_R)(t1_R)$$

Since R has no zero divisors we conclude that $s1_R = 0_R$ or $t1_R = 0_R$. The minimality n implies $s = n$ or $t = n$. Hence n is a prime. □

Let $r \in R$. If R has an identity we define $r^0 = 1$. If R does not have an identity we will use the convention $r^0 s = s$ for all $s \in R$.

Lemma 3.1.17 (Binomial Theorem)**.** *Let R be ring, $a_1, a_2 \ldots, a_n \in R$ and $m \in \mathbb{Z}^+$.*

(a)

$$(\sum_{i=1}^{n} a_i)^m = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_m=1}^{n} a_{i_1} a_{i_2} \ldots a_{i_m}$$

(b) If $a_i a_j = a_j a_i$ for all $1 \leq i, j \leq n$, then

$$(\sum_{i=1}^{n} a_i)^m = \sum_{\{(m_i) \in \mathbb{N}^n \mid \sum_{i=1}^{n} m_i = m\}} \binom{m}{m_1, m_2, \ldots, m_n} a_1^{m_1} a_2^{m_2} \ldots a_n^{m_n}$$

Proof. (a) Follows form 3.1.5e and induction on m.

For (b) notice that $a_{i_1} \ldots a_{i_m} = a_1^{m_1} a_2^{m_2} \ldots a_n^{m_n}$, where $m_k = |\{j \mid i_j = k\}|$. So (b) follows from (b) and a simple counting argument. □

Lemma 3.1.18. *Let $n, m, k \in \mathbb{Z}^+$.*

(a) If $\gcd(m,k) = 1$ *or* $\gcd(n,m) = 1$*, then* $\gcd(f,k) = 1$ *for some* $f \in \mathbb{Z}$ *with* $f \equiv n \pmod{m}$.

(b) There exists $f \in \mathbb{Z}$ *so that* $\gcd(f,k) = 1$ *and* $fn \equiv \gcd(n,m) \pmod{m}$

Proof. (a) Suppose first that $\gcd(m,k) = 1$. Then $1 - n = lm + sk$ for some integers l, s. Thus $1 = (n + lm) + sk$. Put $f = n + lm$, then $\gcd(n + lm, k) = 1$.

Suppose next that $gcd(n,m) = 1$. Write $k = k_1k_2$ where $\gcd(k_1, m) = 1$ and all primes dividing k_2 also divide m. By the first part there exists $l \in \mathbb{Z}$ with $\gcd(n + lm, k_1) = 1$. Now any prime dividing k_1, divides m and (as $\gcd(n,m) = 1$), does not divide m. Hence it also does not divide $m + lm$. Thus $\gcd(n + lm, k) = \gcd(n + lm, k_1) = 1$.

(b) Let $d = \gcd(n,m)$. Replacing n be $\frac{n}{d}$ and m by $\frac{m}{d}$ we may assume that $d = 1$. Then $n^*n \equiv 1 \pmod{m}$ for some $n^* \in \mathbb{Z}$. Since $\gcd(n^*, m) = 1$ we can apply (a) to n^*, m and k. So there exists f with $\gcd(f,k) = 1$ and $f \equiv n^* \pmod{m}$. Then also $fn \equiv 1 \pmod{m}$. □

Lemma 3.1.19. *Let* R *be a ring with* $(R,+)$ *cyclic. Then* R *is isomorphic to exactly one of the following rings:*

1. $\mathbb{Z}$ *with regular addition but zero-multiplication.*

2. $(n\mathbb{Z}/nm\mathbb{Z}, +, \cdot)$*, where* $m \in \mathbb{N}, n \in \mathbb{Z}^+$ *and* n *divides* m.

Proof. Let $m \in \mathbb{N}$ so that $(R,+) \cong (\mathbb{Z}/m\mathbb{Z}, +)$ and let a be generator for $(R,+)$. So $a \cdot a = na$ for some $n \in \mathbb{Z}$. Then for all $k, l \in \mathbb{Z}$, $(ka) \cdot (la) = klna$ and so the multiplication is uniquely determine by n. Note that $(-a)(-a) = na = (-n)(-a)$. So replacing a be $-a$ we may assume that $n \in \mathbb{N}$. Also if $m > 0$ we may choose $0 < n \leq m$.

Suppose first that $n = 0$. Then by our choice $m = 0$ as well. So $(R,+) \cong (\mathbb{Z}, +)$ and $rs = 0$ for all $r, s \in R$.

Suppose next that $n > 0$. Then the map

$$n\mathbb{Z}/nm\mathbb{Z} \to R, \quad nk + nm\mathbb{Z} \to ka$$

is an isomorphism. If $m = 0$, these rings are non-isomorphic for different n. Indeed $R^2 = nR$ and so $|R/R^2| = n$. Therefore n is determined by the isomorphism type R.

For $m > 0$, various choices of n can lead to isomorphic rings. Namely the isomorphism type only depends on $d = \gcd(n,m)$. To see this we apply 3.1.18 to obtain $f \in \mathbb{Z}$ with $\gcd(f,m) = 1$ and $fn \equiv d \pmod{m}$. Then $1 = ef + sm$ for some $e, s \in \mathbb{Z}$ and so $f + m\mathbb{Z}$ is invertible. Hence also fa is a generator for $(R,+)$ and

$$(fa) \cdot (fa) = f^2na = (fn)(fa) = d(fa).$$

Also $R^2 = dR$ and $|R/R^2| = \frac{m}{d}$. So d is determined by the isomorphism type of R. □

3.2 Ideals and homomorphisms

Definition 3.2.1. *Let $(R, +, \cdot)$ be a be a ring.*

(a) A subring *of R is a ring $(S, \triangle, \square)$ such that S is a subset of R and*

$$s \triangle t = s + t \quad \text{and} \quad s \square r = s \cdot t$$

for all $s, t \in R$.

(b) An left (right) ideal *in R is a subring I of R so that $rI \subseteq I$ $Ir \subseteq I$ for all $r \in R$.*

(c) An ideal in R is a left ideal which is also a right ideal.

Lemma 3.2.2. *Let R be a ring and $S \subseteq R$ such that*

(i) $0_R \in S$.

(ii) $a + b \in S$ *for all* $a, b \in S$.

(iii) $-a \in S$ *for all* $a \in S$.

(iv) $ab \in S$ *for all* $a, b \in S$.

Define $+_S : S \times S \to S, (a, b) \to a + b$ and $\cdot_S : S \times S, (a, b) \to a \cdot b$. Then $(S, +_S, \cdot_S)$ is a subring of S.

If S fulfills (i),(ii), (iii) and

(iv′) $rb \in S$ *for all* $r \in R, b \in R$,

then S is a left ideal.

If S fulfills (i),(ii), (iii) and

(iv″) $ar \in S$ *for all* $r \in R, a \in R$,

then S is a right ideal in G

If S fulfills, (i),(ii), (iii) , ((iv′)) and ((iv″)) then S is an ideal.

Proof. Straightforward and we leave the few details to the reader. □

Example 3.2.3. 1. Let $n \in \mathbb{Z}$. Then $n\mathbb{Z}$ is an ideal in $\mathbb{Z}$.

2. Let V be a vector space over $\mathbb{R}$. Let W be any subset Define

$$\mathrm{Ann}(W) = \{\alpha \in \mathrm{End}_{\mathbb{R}}(V) \mid \alpha(w) = 0_V \text{ for all } w \in W\}.$$

$\mathrm{Ann}(W)$ is called the annihilator of W in $\mathrm{End}(W)$. We will show that $\mathrm{Ann}(W)$ is left ideal in $\mathrm{End}_{\mathbb{R}}(V)$.

Let $\alpha, \beta \in \mathrm{Ann}(W)$, $\gamma \in \mathrm{End}_{\mathbb{R}}(V)$ and $w \in W$, then

$$0_{\mathrm{End}_{\mathbb{R}}(V)}(w) = 0_W$$
$$(\alpha + \beta)(w) = \alpha(w) + \beta(w) = 0_V + 0_V = 0_V$$
$$(\gamma \circ \alpha)(w) = \gamma(\alpha(w) = \gamma(0_V) = 0_V$$

and so by 3.2.2 $\mathrm{Ann}(W)$ is an ideal.

Lemma 3.2.4. *Let $\phi : R \to S$ be a ring homomorphism.*

(a) If T is a subring of R, $\phi(T)$ is a subring of S.

(b) If T is a subring of S then $\phi^{-1}(T)$ is a subring of R.

(c) $\ker \phi$ an ideal in R.

(d) If I is an (left,right) ideal in R and ϕ is onto, $\phi(I)$ is a (left,right) ideal in S.

(e) If J is a (left,right) ideal in S, then $\phi^{-1}(J)$ is an (left,right) ideal on R.

Proof. Straight forward. □

Theorem 3.2.5. *Let R and S be rings. Suppose that one of the following holds.*

(i) $\alpha : R \to S$ is a ring homomorphism and $\beta : G \to (S, \cdot)$ a semigroup homomorphism such that
$$\alpha(r)\beta(g) = \beta(g)\alpha(r) \text{ for all } r \in R, g \in G$$

(ii) $S = R^*[G^*]$ for some ring S^* and semigroup G^*, $\alpha : R \to R^*$ is a ring homomorphism, and $\beta : G \to G^*$ is semigroup homomorphism.

Then
$$\gamma : R[G] \to S \quad \sum_{g \in G} r_g g \to \sum \alpha(r_g)\beta(g)$$

is a ring homomorphism.

Proof.

$$\gamma\left(\sum_{g\in G} r_g g + \sum_{g\in G} s_g g\right) = \gamma\left(\sum_{g\in G}(r_g + s_g)g\right) = \sum_{g\in G}\alpha(r_g + s_g)\beta(g) =$$
$$= \sum_{g\in G}(\alpha(r_g) + \alpha(s_g))\beta(g) = \sum_{g\in G}\alpha(r_g)\beta(g) + \sum_{g\in G}\alpha(s_g)\beta(g) =$$
$$\gamma\left(\sum_{g\in G} r_g g\right) + \gamma\left(\sum_{g\in G} s_g g\right)$$

and

$$\gamma\left(\sum_{g\in G} r_g g \cdot \sum_{h\in G} s_h h\right) = \gamma\left(\sum_{k\in G}\left(\sum_{(g,h)\in G, gh=k} r_g r_h\right) k\right) = \sum_{k\in G} \alpha\left(\sum_{(g,h)\in G, gh=k} r_g r_h\right)\beta(k)$$
$$= \sum_{k\in G}\left(\sum_{(g,h)\in G, gh=k} \alpha(r_g r_h)\right)\beta(k) = \sum_{k\in G}\left(\sum_{(g,h)\in G, gh=k} \alpha(r_g r_h)\beta(gh)\right)$$
$$= \sum_{g\in G}\sum_{k\in G} \alpha(r_g s_k)\beta(gk) = \sum_{g\in G}\sum_{k\in G} \alpha(r_g)\alpha(s_k)\beta(g)\beta(k) =$$
$$\left(\sum_{g\in G} \alpha(r_g)\beta(g)\right)\cdot\left(\sum_{k\in G} \alpha(s_k)\beta(k)\right) = \gamma\left(\sum_{g\in G} r_g g\right)\cdot\gamma\left(\sum_{k\in G} s_k k\right)$$

□

Example 3.2.6. 1. Let S be a commutaive ring, R a subring of S and $s \in S$. Define $\alpha : R \to S, r \to r$ and $\beta : (\mathbb{N}, +) \to (S, i), n \to s^i$. Then α is a ring homomorphism and β a semigroup homomorphism. Then by 3.2.5

$$\gamma_s : R[x] \to S, \sum_{i=0}^{n} r_i x^i \to r_i s^i$$

is a ring homomorphism. For $f \in R[x]$ we denote $\gamma_s(f)$ by $f(s)$. Then

For all $f, g \in R[x]$ and $s \in S$:

$$(f+g)(s) = \gamma_s(f+g) = \gamma_s(f) + \gamma_s(g) = f(s) + g(s)$$
$$(f\cdot g)(s) = \gamma_s(f\cdot g) = \gamma_s(f)\gamma_s(g) = f(s)\cdot f(g)$$

2. Let $\alpha : R \to S$ be a ring homomorphism and $\beta : G \to H$ a semigroup homomorphism. Then by 3.2.5

$$\gamma : R[G] \to S[H] \quad \sum_{g\in G} r_g g \to \sum \alpha(r_g)\beta(g)$$

is a ring homomorphism. What is the image and the kernel of γ? Clearly $\gamma(R[G]) = \alpha(R)[\beta(G)]$. Let $I = \ker\alpha$. To compute $\ker\gamma$ note that

$$\gamma(\sum_{g\in G} r_g g) = \sum_{h\in H} \alpha(\sum_{g\in\beta^{-1}(h)} r_g)h$$

and so

$$\sum_{g\in G} r_g g \in \ker\gamma \iff \sum_{t\in\beta^{-1}(h)} r_t \in I \text{ for all } h \in \beta(G).$$

If β is a group homomorphism we can describe $\ker\gamma$ just in terms of $I = \ker\alpha$ and $N := \ker\beta$. Indeed the $\beta^{-1}(h)$'s ($h \in \beta(G)$) are just the cosets of N and so

$$\sum_{g\in G} r_g g \in \ker g \Longleftrightarrow \sum_{t\in T} r_t \in I \text{ for all } T \in G/N.$$

Let as consider the special case where $R = S$, $\alpha = \mathrm{id}_R$ and $H = \{e\}$. Identify $R[e]$ with R via $re \leftrightarrow r$. Then γ is the map

$$R[G] \to R, \sum r_g g \to \sum r_g.$$

The kernel of γ is the ideal

$$R^\circ[G] = \left\{\sum r_g g \mid \sum r_g = 0\right\}$$

$R^\circ[G]$ is called the *augmentation ideal* of $R[G]$.

Definition 3.2.7. *Let R be a ring and $A, B \subseteq R$. Then*

(a) $A + B := \{a + b \mid a \in A, b \in B\}$.

(b) $\langle A\rangle$ *is subgroup of* $(R, +)$ *generated by* A.

(c) For $1 \le i \le n$ *let* $A_i \subseteq R$. *Then*

$$A_1A_2\ldots A_n := \langle a_1a_2\ldots a_n \mid a_i \in A_i, 1 \le i \le n\rangle$$

(d) $(A) = \bigcap\{I \mid I$ *is an ideal in* $R, A \subseteq I\}$. (A) *is called the ideal in* R *generated by* A.

Lemma 3.2.8. *Let R be a ring and $A, B \subseteq R$.*

(a) If If A is a left ideal, then AB is a left ideal.

(b) If B is a right ideal, then AB is a right ideal.

(c) If A is a left ideal in R and B is right ideal, then AB is a ideal in R.

(d) If A and B are (right,left) ideals then $A + B$ is a (left,right) ideal.

(e) Let $(A_i, i \in I)$ be a family if (left,right) ideals of R the $\bigcap_{i\in I} A_i$ is a (left,right) ideal.

(f) (A) is an ideal.

(g) $RA + \langle A\rangle = \bigcap\{I \mid A \subseteq I, I$ *is a left ideal in* $R\}$. *If R has an identity then* $RA + \langle A\rangle = RA$.

(h) $AR + \langle A\rangle = \bigcap\{I \mid A \subseteq I, I$ *is a right ideal in* $R\}$. *If R has an identity then* $AR + \langle A\rangle = AR$.

(i) If $(A) = RA + AR + RAR + \langle A\rangle$. *If* R *has an identity, then* $(A) = RAR$.

Proof. Let $r \in R, a \in A$ and $b \in B$.

(a) Since A is a left ideal, $ra \in A$ and so $r(ab) = (ra)b \in AB$. Since left multiplication by r is a homomorphism, we conclude from 2.6.3(c) that

$$rAB = r\langle ab \mid a \in A, b \in b\rangle = \langle rab \mid a \in A, b \in B\rangle \leq AB$$

So AB is left ideal in R. (b) Similar to (a).

(c) Follows from (a) and (b).

(d) Suppose A and B are left ideals. Then $r(a + b) = ra + rb \in A + B$ and so $A + B$ is a left ideal. The remaining statements are proved similarly.

(e) Suppose each A_i is an left ideal. By 2.6.1 $\bigcap_{i\in I} A_i$ is subgroup of $(R, +)$. Let $a \in \bigcap_{i\in I} A_i$. The $a \in A_i$ and so $ra_i \in A_i$ for all $i \in I$. Thus $ra_i \in \bigcap_{i\in I} A_i$ and so $\bigcap_{i\in I} A_i$ is a left ideal. The remaining statements are proved similarly.

(f) (A) is the intersection of the ideals containing A and so by (e), is an ideal.

(g) Clearly $RA + \langle A\rangle$ is contained in every left ideal containing A, and so also in the intersection of this ideals. So it suffices to show that $RA + \langle A\rangle$ is a left ideal. We have $r\langle A\rangle = \langle rA\rangle \leq RA$ and by (a) also $rRA \leq RA$. Hence $r(RA + \langle A\rangle \leq RA \leq RA + \langle A\rangle$ and $RA + \langle A\rangle$ is a left ideal.

If R has an identity, the $\langle A\rangle = 1_R A \leq RA$.

(h) Similar to (g).

(i) As in (g) for the first statement it suffices to show that $RAR + RA + AR + \langle A\rangle$ is an ideal. By (g) $RA + \langle A\rangle$ is a left ideal. Also $AR = \langle AR\rangle$ and so by (g) $RAR + AR = R(AR) + \langle AR\rangle$ is a left ideal. Hence by (d), $RAR + RA + AR + \langle A\rangle$ is a left ideal. Similary it is a right ideal and so an ideal.

If R has an identity, then $RA = RA1_R$, $AR = 1_R RA$ and $\langle A\rangle = 1_R A1_R$ all all are contained in RAR. Thus $(A) = RAR$. □

Lemma 3.2.9. *Let* I *be an ideal in the ring* R.

(a) The binary operations

$$\begin{aligned} +_{R/I} &: R/I \times R/I \to R/I, \quad (a + I, b + I) \to (a + b) + I \quad \textit{and} \\ \cdot_{R/I} &: R/I \times R/I \to R/I, \quad (a + I, b + I) \to ab + I \end{aligned}$$

are well-defined.

(b) $(R/I, +_{R/I}, \cdot_{R/I})$ *is a ring.*

(c) The map

$$\pi : R \to R/I, \quad r \to r + I$$

is a ring homomorphism with kernel I.

Proof. (a) That $+_{R/I}$ is well-defined follows from 2.5.7. $i, j \in I$. Then $(a+i)(b+j) = ab + ib + aj + ij$. As I is an ideal, $ib + aj + ij \in I$ and so $(a+i)(b+j) + I = ab + I$. Thus also $\cdot_{R/N}$ is well-defined.

(b) By 2.5.7 $(R/I, +)$ is a group. The remaining axiom of a ring are readily verified.

(c) By 2.5.7 is a well-defined homomorphism of abelian groups with $\ker \pi = I$. Since

$$\phi(ab) = ab + I = (a + I) \cdot_{R/I} (b + I) = \pi(a) \cdot_{R/N} \pi(b)$$

and so π is ring homomorphism. □

A little warning: Let $a, b \in R/I$. Then $a \cdot_{R/I} b$ is usually not equal to $a \cdot b$. (Note that a, b are subsets of R and so $a \cdot b = \langle xy \mid x \in a, y \in b \rangle$.) For example consider $R = \mathbb{Z}$ and $a = b = I = 2\mathbb{Z}$. Then

$$2\mathbb{Z} \cdot 2\mathbb{Z} = 4\mathbb{Z} \text{ and } 2\mathbb{Z} \cdot_{Z/2\mathbb{Z}} 2\mathbb{Z} = (0 + 2\mathbb{Z}) \cdot_{\mathbb{Z}/2\mathbb{Z}} (0 + 2\mathbb{Z}) = 0 + 2\mathbb{Z}$$

Nevertheless we still usually just write ab for $a \cdot_{R/N} b$ if its clear from the context that are viewing a, b as elements of R/I and not as subsets of R.

Theorem 3.2.10 (The Isomomorphism Theorem for Rings)**.** *Let $\phi : R \to S$ be a ring homomorphism. Then the map*

$$\overline{\phi} : R/\ker \phi \to \phi(R), \quad r + \ker \phi \to \phi(r)$$

is a well-defined isomorphism of rings.

Proof. By the Isomorphism Theorem for groups 2.5.8, this is a well-defined isomorphism for the additive groups. We have

$$\overline{\phi}((a + \ker \phi)(b + \ker \phi)) = \overline{\phi}(ab + \ker \phi) = \phi(ab) = \phi(a)\phi(b) = \overline{\phi}(a + \ker \phi)\overline{\phi}(b + \ker \phi)$$

and $\overline{\phi}$ is a ring isomorphism. □

We will see below that any ring R can be embedded into a ring S with an identity. This embedding is somewhat unique. Namely suppose that $R \leq S$ and S has an identity. Then for $n, m \in \mathbb{Z}$ and $r, s \in R$ we have $(n1 + r)+)(m1 + s) = (n + m)1 + (r + s)$ and $(n1 + r)(m1 + s) = (nm)1 + (mr + ns + rs)$. So already $\mathbb{Z}1 + R$ is a ring with 1, contains R and the addition and multiplication on $\mathbb{Z}1 + R$ is uniquely determined. But there is some degree of freedom. Namely $\mathbb{Z}1 + R$ does not have to be a direct sum.

Let $\hat{R} = \mathbb{Z} \times R$ as abelian groups. We make $\hat{R}$ into a ring by defining

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

Then $(1, 0)$ is an identity in $\hat{R}$. The map $\phi : \hat{R} \to S, (n, r) \to n1 + r$ is a homomorphism with image $\mathbb{Z}1 + R$. Let us investigate $\ker \phi$. $(n, r) \in \ker \phi$ iff $r = -n1$. Let $k\mathbb{Z}$ be the inverse

image of $\mathbb{Z}1 \cap R$ in $\mathbb{Z}$. Also put $t = k1$ and $D_{k,t} = \{(lk, -lt) \mid l \in \mathbb{Z}\}$. Then $\ker \phi = D_{k,t}$. Hence $\hat{R}/D_{k,t} \cong \mathbb{Z}1 + R$.

Now which choices of $k \in \mathbb{Z}$ and $t \in R$ can really occur? Note that as $t = -n1$, $tr = kr = rt$. This necessary condition on k and t turns out to be sufficient:

Let $k \in \mathbb{Z}$. $t \in R$ is called a k-element if $tr = rt = kr$ for all $r \in R$. Note that a 1-element is an identity, while a 0-element is an element with $tR = Rt = 0$. Also if a and b are k-elements, then $a - b$ is a 0-element. So if a k-elements exists it unique modulo the zero elements.

Suppose now that t is a k-element in R. Define $D_{t,k}$ has above. We claim that $D_{k,t} = \mathbb{Z}(k, -t)$ is an ideal in R. For this we compute(using $rt = kr$)

$$(n, r) \cdot (k, -t) = (nk, kr - nt - rt) = (nk, kr - nt - kr) = (nk, -nt) = n(k, -t).$$

So $D_{k,t}$ is a left ideal. Similarly, $D_{t,k}$ is a right ideal. Put $R_{k,t} = \hat{R}/D_{k,t}$. Then $R_{k,t}$ is a ring with identity, contains R (via the embedding $r \to (0, r) + D_{k,t}$) and fulfills $\mathbb{Z}1 \cap R = k\mathbb{Z}1 = \mathbb{Z}t$.

Note that if t is an k-element and s an l-element, then $-t$ is an $-k$ element and $t + s$ is an $(k + l)$-element. Therefore the sets of $k \in \mathbb{Z}$ for which there exists a k-element is a subgroup of $\mathbb{Z}$ and so of the form $i\mathbb{Z}$ for some $i \in \mathbb{N}$. Let u be a i-element. $R_{i,u}$ is in some sense the smallest ring with a identity which contains R. Also if R has no 0-elements, u and so $R_{i,u}$ is uniquely determined.

For example if $R = n\mathbb{Z}$, then $i = n = u$ and $R_{i,u} \cong \mathbb{Z}$. Indeed $\hat{R} = \mathbb{Z} \times n\mathbb{Z}$, $D_{nn} = \{(jn, -jn) \mid j \in \mathbb{Z}\}$, $\hat{R} = \mathbb{Z}(1, 0)\mathrm{op}lusD_{n,n}$ and the map $R_{n,n} \to \mathbb{Z}$, $(j, r) + D_{n,n} \to j + r$ is an isomorphism between $R_{n.n}$ and $\mathbb{Z}$.

Next we will show that R can be embedded into a ring with identity which has same characteristic as R. Put $n = \operatorname{char} R$, then 0 is an n-element. Also $D_{n,0} = n\mathbb{Z} \times \{0\}$ and $R_{n,0} \cong \mathbb{Z}/n\mathbb{Z} \times R$ as abelian groups. So $R_{n,0}$ has characteristic n. On the other hand $\hat{R} = R_{0,0}$ always has characteristic 0.

Definition 3.2.11. *Let I be an ideal in the ring R with $I \neq R$.*

(a) I is prime ideal *if for all ideals A, B in R*

$$AB \leq I \quad \Longleftrightarrow \quad A \leq I \text{ or } B \leq I$$

(b) I is a maximal ideal *if for each ideal A of R.*

$$I \leq A \subseteq R \Longleftrightarrow A = I \text{ or } A = R.$$

Example 3.2.12. Let I be an ideal in $\mathbb{Z}$ with $I \neq \mathbb{Z}$. Then I is a subgroup of $\mathbb{Z}$ and so $I = n\mathbb{Z}$ for some $n \in \mathbb{N}$ with $n \neq 1$. Let $A = a\mathbb{Z}$ and $B = b\mathbb{Z}$ with $a, b \in \mathbb{N}$. Then $AB = ab\mathbb{Z}$ and so $AB \leq I$ if and only if $n \mid ab$. Also $n\mathbb{Z} \leq a\mathbb{Z}$ if and only if $n \mid a$ and so I is a prime ideal if and only if

$$\iff n \mid ab \iff n \mid a \text{ or } n \mid a$$

This is this is case if and only if $n = 0$ or n is a prime. So the prime ideals in $\mathbb{Z}$ are $\{0\}$ and $p\mathbb{Z}$, p a prime.

I is a maximal ideal if and only if $n\mathbb{Z} \leq a\mathbb{Z}$ implies $n\mathbb{A} = a\mathbb{Z}$ or $a\mathbb{Z} = \mathbb{Z}$. So if and only if $a \mid n$ implies $n = a$ or $n = 1$. This is the case if and only if n is a prime. So the maximal ideals in $\mathbb{Z}$ are $p\mathbb{Z}$, p a prime.

Lemma 3.2.13. *Let P be an ideal in the ring R with $P \neq R$.*

(a) If for all $a, b \in R$,

$$ab \in P \iff a \in P \text{ or } b \in P$$

then P is a prime ideal

(b) If R is commutative, the converse of (a) holds.

(c) If R is commutative, then P is a prime ideal if and only R/P has no zero divisors.

Proof. (a) Let A and B are ideals in R with $AB \leq P$. We need to show that $A \leq P$ or $B \leq P$. So suppose $A \not\leq P$ and pick $a \in A \setminus P$. Since $ab \in P$ for all $b \in B$ we conclude $b \in P$ and $B \leq P$.

(b) Suppose that P is prime ideal and $a, b \in R$ with $ab \in P$. Note that $(a) = \langle a \rangle + Ra = \{na + ra \mid n \in a, r \in R\}$. Let $n, m\mathbb{Z}$ and $r, s \in R$. Then

$$(na + ra)(mb + sa) = (nm)ab + (ns + mr + rs)ab$$

and so $(a)(b) \subseteq (ab) \subseteq P$. As P is prime ideal , $(a) \subseteq P$ or $(b) \subseteq P$. Hence $a \in P$ or $b \in P$.

(c) Just note that the condition in (a) is equivalent to saying that R/P has no zero divisors. □

Lemma 3.2.14. *Let R be a non-zero commutative ring with identity and P an ideal in R. Then P is prime ideal if and only if R/P is an integral domain.*

Proof. If P is a prime ideal or if R/P is an integral domain we have that $R \neq P$. So the lemma follows from 3.2.13c. □

Theorem 3.2.15. *Let R be a ring with identity and $I \neq R$ be an ideal. Then I is contained in a maximal ideal. In particular every non-zero ring with identity has a maximal ideal.*

Proof. The second statement follows from the first applied to the zero ideal. To prove the first we apply Zorn's lemma A.6. For this let $\mathcal{M}$ be the set of ideals J of R with $I \subseteq J \subsetneq R$. Order $\mathcal{M}$ by inclusion and let $\mathcal{C}$ be a nonempty chain in $\mathcal{M}$. Let $M = \cup\mathcal{C}$. By A.7 M is an ideal. Clearly $I \subseteq M$. Also 1_R is not contained in any member of $\mathcal{C}$ and so $1_R \notin M$. Hence $M \neq R$ and $M \in \mathcal{M}$. Thus every chain has an upper bound and so by Zorn's Lemma $\mathcal{M}$ has a maximal element M. If $M \subset A$ for some ideal $A \neq R$, then $I \subseteq A$, $A \in \mathcal{M}$ and so by maximality of M in $\mathcal{M}$, $A = M$. Thus M is a maximal ideal. □

Theorem 3.2.16. *Let M be a maximal ideal in the ring R. Then M is a prime ideal if and only $R^2 \subseteq M$. In particular if R is a ring with $R^2 = R$ or a ring with identity then every maximal ideal is a prime ideal.*

Proof. We will show that $R^2 \subseteq M$ if and only if M is not a prime ideal.

Suppose $R^2 = RR \subseteq M$. Since R is an ideal in R and $R \nsubseteq M$, we conclude that M is not a prime ideal.

Suppose that M is not a prime ideal. Then $AB \subseteq M$ for some ideals A and B with $A \nsubseteq M$ and $B \nsubseteq M$. By 3.2.8(d), $A + M$ and $B + M$ are ideals in R. So he maximality of M implies $R = A + M = B + M$. Thus $R^2 = (A + M)(B + M) \subseteq AB + M \subseteq M$.

If R has an identity, then $R^2 = R$ and if $R^2 = R$, then $R^2 \nsubseteq M$. So the second statement follows from the first. □

Definition 3.2.17. *Let R be a ring.*

(a) A subring of S of R is called proper, if $S \neq \{0_R\}$ and $S \neq R$.

(b) R is called simple *if $R^2 \neq \{0_R\}$ and R has no proper ideals.*

Lemma 3.2.18. *(a) Let R be a division ring. Then R has non proper left or right ideals. In particular, R is simple.*

(b) Let R be commutative ring with identity. The R is simple if and only if R is a field.

Proof. (a) Let I be an non-zero left ideal in R and pick $0_R \neq i \in I$. Then $1_R = i^{-1}i \in RI \subseteq R$ and so $R = R1_R \subseteq I$. Similarly R has no -proper right ideals. Since $0_r \neq 1_R = 1_R^2 \in R^2$, $R^2 \neq \{0_R\}$ and so R is simple.

(b) Let R be simple commutative ring with identity. $0_R \neq a \in R$. Since R has an identity, Ra is non-zero ideal. As R is simple $Ra = R$. Thus $ra = 1_R$ for some r. As R is commutative, $ar = 1_R$ and so r has an inverse. Since R is simple, $R \neq\neq \{0_R\}$ and so also $1_R \neq 0_R$. Hence R is a field.

If R is a field, then by (a), R is simple. □

Lemma 3.2.19. *Let R be a ring and M an ideal in R. Then R/M is simple if and only if M is a maximal ideal with $R^2 \nsubseteq M$.*

Proof. In both cases $M \neq R$ and so we may assume $R \neq M$. $(R/M)^2 \neq 0_{R/M}$ if and only if $R^2 \nsubseteq M$. R/M has no proper ideals if and only if there does not exists an ideal J with $I \nsubseteq J \nsubseteq R$ and so if and only if M is a maximal ideal. □

If I is an ideal we will sometimes write $a \equiv b \pmod{I}$ if $a+I = b+I$, that is if $a-b \in I$. If $R = \mathbb{Z}$ and $I = n\mathbb{Z}$ then $a \equiv b \pmod{n\mathbb{Z}}$ is the same as $a \equiv b \pmod{n}$.

Theorem 3.2.20 (Chinese Remainder Theorem)**.** *Let $(A_i, i \in I)$ be a family of ideals in the ring R.*

(a) The map θ :

$$R/\bigcap_{i\in I} A_i \to \prod_{i\in A_i} R/A_i$$

$$r + \bigcap_{i\in I} A_i \to (r + A_i)_{i\in I}$$

is a well defined monomorphism.

(b) Suppose that I *is finite,* $R = R^2 + A_i$ *and* $R = A_i + A_j$ *for all* $i \neq j \in I$*. Then*

(a) If $|I| > 1$*, then* $R = A_i + \bigcap_{i\neq j\in I} A_j$*.*

(b) θ *is an isomorphism.*

(c) For $i \in I$ *let* $b_i \in R$ *be given. Then there exists* $b \in R$ *with*

$$b \equiv b_i \pmod{A_i} \text{ for all } i \in I$$

Moreover, b *is unique* $\pmod{\bigcap_{i\in I} A_i}$.

Proof. (a) The map $r \to (r + A_i)_{i\in I}$ is clearly a ring homomorphism with kernel $\bigcap_{i\in I} A_i$. So (a) holds.

(b:a) For $\emptyset \neq J \subseteq I$ put $A_J = \bigcap_{j\in J} A_j$. We will show by induction on $|J|$ that

$$R = A_i + A_J$$

for all $\emptyset \neq J \subseteq I \setminus \{i\}$. Indeed if $|J| = 1$ this is part of the assumptions. So suppose $|J| > 1$, pick $j \in J$ and put $K = J \setminus \{j\}$. Then by induction $R = A_i + A_K$ and $R = A_i + A_j$. Note that as A_j and A_K are ideals, $A_jA_K \subseteq A_j \cap A_K = A_J$ Thus

$$R^2 = (A_i + A_j)(A_i + A_K) \subseteq A_i + A_jA_K \subseteq A_i + A_J$$

Hence $R = A_i + R^2 = A_i + A_J$.

(b:b) By (a) we just need to show that θ is onto. For $|I| = 1$, this is obvious. So suppose $I| \geq 2$. Let

$$x = (x_i)_{i\in I} \in \prod_{i\in A_i} R/A_i.$$

We need to show that $x = \theta(b)$ for some $b \in R$. Let $x_i = b_i + A_i$ for some $b_i \in R$. By (ba), we may choose $b_i \in \bigcap_{j\in i\neq I} A_j$. So $b_i \in A_j$ for all $j \neq i$. Thus

$$\theta(b_i)_j = \begin{cases} x_i & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$$

Put $b \sum_{i\in I} b_i$. Then $\theta(b)_j = x_j$ and so $\theta(b) = x$.

(b:c) This is clearly equivalent to (b:b) □

The special case $R = \mathbb{Z}$ is an elementary result from number theory which was know to Chinese mathematicians in the first century A.D. To state this result we first need to observe a couple of facts about ideals in $\mathbb{Z}$.

Let n, m be positive integers. $\gcd(n, m)$ denotes the greatest common divisor and $\operatorname{lcm}(n, m)$ the least common multiple of n and m. Then

$$n\mathbb{Z} \cap m\mathbb{Z} = \operatorname{lcm}(n, m)\mathbb{Z}$$

and

$$n\mathbb{Z} + m\mathbb{Z} = \gcd(n, m)\mathbb{Z}$$

In particular n and m are relatively prime if and only if $n\mathbb{Z} + m\mathbb{Z} = \mathbb{Z}$. So part (b:c) of the Chinese Remainder Theorem translates into:

Corollary 3.2.21. *Let $m_1, \dots m_n$ be positive integers which are pairwise relatively prime. Let $b_1, \dots, b_n$ be integers. Then there exists an integer b with*

$$b \equiv b_i \pmod{m_i} \text{ for all } 1 \leq i \leq n$$

Moreover, b is unique $\pmod{m_1 m_2 \dots m_n}$

3.3 Factorizations in commutative rings

Definition 3.3.1. *Let R be a commutative ring and $a, b \in R$.*

(a) We say that a divides *b and write $a \mid b$, if $(b) \subseteq (a)$.*

(b) We say that a and b are associate *and write $a \sim b$, if $(a) = (b)$*

(c) We say that a is proper if $\{0\} \neq (a) \neq R$.

Lemma 3.3.2. *Let R be a commutative ring and $a, b \in R$.*

(a) $a \sim b \iff a \mid b$ and $b \mid a$.

(b) The relation $\mid$ on R is reflexive and transitive.

(c) The relation $\sim$ on R is an equivalence relation.

(d) If R has an identity, then $a \mid b$ if and only if $b = ra$ for some $r \in R$.

Proof. Obvious. □

Lemma 3.3.3. *Let R be a commutative ring with identity and $u \in R$. The following are equivalent*

(a) u is a unit.

(b) $u \mid 1_R$

(c) $u \mid r$ *for all* $r \in R$

(d) $r \sim ur$ *for all* $r \in R$.

(e) $u \sim 1_R$

(f) $(u) = R$

(g) u *is not contained in any maximal ideals of* R.

Proof. (a) $\Longrightarrow$ (b): If u is a unit, then $ru = 1_R$ for some $r \in R$ and so $u \mid 1_R$.
(b) $\Longrightarrow$ (c): Since $1_R \mid r$, $u \mid 1_R$ implies $u \mid r$.
(c) $\Longrightarrow$ (d): We have $u \mid 1_R$. Hence $1_R = su$ for some $s \in R$. Thus $r = 1_R r = (su)r = s(ur)$ and so $ur|r$. Clearly $r|ur$ and so $u \sim ur$.
(d) $\Longrightarrow$ (e): Choosing $r = 1_R$ in (d) we see that $1_r = \sim u1_R = u$.
(e) $\Longrightarrow$ (f): Suppose $u \sim 1_R$, then $(u) = (1_R) = R$.
(f) $\Longrightarrow$ (g): Suppose $R = (u)$ and let M be a maximal ideal of R. If $u \in M$, then also $R = (u) \subseteq M$, a contradiction. Hence $u \notin M$.
(g) $\Longrightarrow$ (a): Suppose u is not a unit. The $1_R \not\in Ru = (u)$ and so by 3.2.15 there exists a maximal ideal M of R with $(u) \subseteq M$. A contradiction. □

Lemma 3.3.4. *Let* R *be an integral domain. Let* $a, b, u \in R^{\#}$ *with* $b = ua$. *Then* $b \sim a$ *if and only if* u *is a unit.*

Proof. The "if" part follows from 3.3.3(d). So suppose that $b \sim a$. Then $a = vb$ for some $v \in R$. Thus $1_R b = b = ua = u(vb) = (uv)b$. Since the cancellations law hold in integral domains we conclude that $uv = 1_R$. So u is a unit. □

3.3.5 (R^*)**.** Note that by 2.2.3(e), $(R^*, \cdot)$ is a group. Also R^* acts on R by left multiplication. The previous lemma now says that in an integral domain the orbits of R^* on R are exactly the equivalence classes of $\sim$.

Definition 3.3.6. *Let* R *be a ring.*

(a) An ideal I *is called a* principal ideal *if its generated by one element, that is* $I = (r)$ *for some* $r \in R$.

(b) R *is called a* principal ideal ring *if every ideal is a principal ideal.*

(c) R *is* principal ideal domain *(*PID*), if* R *is an integral domain and a principal ideal ring.*

Definition 3.3.7. *Let* R *be a commutative ring with identity and* c *a proper element.*

(a) c *is called* irreducible *if for all* $a, b \in R$

$$c = ab \Longrightarrow a \text{ or } b \text{ is a unit}$$

(b) c is called a prime *if for all $a, b \in R$*

$$p \mid ab \Longrightarrow p|a \text{ or } p \mid b.$$

Lemma 3.3.8. *Let p be nonzero element in the integral domain R. Then following are equivalent:*

(a) p is a prime

(b) (p) is a prime ideal

(c) $R/(p)$ is an integral domain

Proof. Note that in each case p is not a unit, that is $R \neq (p)$. So we may assume that p is not a unit.

Observe that for $d \in R$, $p \mid d$ if and only if $d \in (p)$. Let $a, b \in R$. Then $p \mid ab$ implies $p \mid a$ or $p \mid b$ if and only if $ab \in (p)$ implies $a \in (p)$ or $b \in (p)$. Using 3.2.13 we conclude that p is a prime if and only if (p) is a prime ideal. By 3.2.14 (p) is a prime ideal if and only if $R/(p)$ is an integral domain. □

Lemma 3.3.9. *Let c be a proper element in the integral domain R. Then the following are equivalent*

(a) c is irreducible.

(b) For all $a \in R$,

$$a \mid c \Longrightarrow a \sim c \text{ or } a \text{ is a unit}$$

(c) (c) is maximal in the set of proper principal ideals.

Proof. (a) $\Longrightarrow$ (b): Let $a \in R$ with $a \mid c$. Then $c = ab$ for some $b \in R$. Since c is irreducible, a is a unit or b is a unit. If b is a unit then by 3.3.3 $ab \sim a$ and so $a \sim c$. Hence (b) holds

(b) $\Longrightarrow$ (c): Since c is proper, $(c) \neq R$. Suppose $(c) \subseteq (a)$ for some $a \in R$ with $(a) \neq R$. Then $c \mid a$. Thus by (b), $a \sim c$ or a is a unit. if a is a unit $(a) = R$, a contradiction. Hence $a \sim c$ and so $(a) = (c)$ and (c) holds.

(c) $\Longrightarrow$ (a): Let $a, b \in R$ with and $c = ab$. Then $(c) \subseteq (a)$ and so by (c), $(a) = (c)$ or $(a) = R$. In the first case $a \sim c$ and so by 3.3.4 b is a unit. In the second case a is a unit. Hence c is irreducible. □

Lemma 3.3.10. *Let R be an integral domain.*

(a) Every prime element in R is irreducible.

(b) Every associate of a prime is a prime and every associate of an irreducible element is irreducible.

Proof. (a) Let p be a prime and suppose $p = ab$ for some $a, b \in R$. Since $p \mid p$, $p \mid ab$. Since p is a prime, $p \mid a$ or $p \mid b$. Without loss $p \mid a$. Since also $a \mid p$ we get $p \sim a$. Thus 3.3.4 implies that b is unit. Hence p is irreducible.

(b) By definition of proper, the property 'a is proper" only dependents on the ideal generated by a. By 3.3.8 and 3.3.9 also the properties "prime" and "irreducible" also only depends on the ideal generated by a. This gives (b). □

Lemma 3.3.11. *Let R is principal ideal domain. Then the following are equivalent*

(a) p is a prime

(b) p is irreducible.

(c) (p) is a maximal ideal.

(d) $R/(p)$ is a field.

Proof. (a) $\Longrightarrow$ (b): This is 3.3.10

(b) $\Longrightarrow$ (c): By 3.3.9is (p) is a maximal proper principal ideal. Since every ideal in a PID is a principal ideal, (p) is a maximal ideal. So (c) holds. holds.

(c) $\Longrightarrow$ (d): This follows from 3.2.18 3.

(d) $\Longrightarrow$ (a): . By 3.2.16, (p) is a prime ideal. So by 3.3.8 p is a prime. □

Lemma 3.3.12. *Let R be an integral domain and $a \in R$. Suppose that $a = p_1 \cdot \ldots \cdot p_n$ with each p_i a prime in R.*

(a) If $q \in R$ is a prime with $q \mid a$, then $q \sim p_i$ for some $1 \leq i \leq n$.

(b) If $a = q_1 \ldots q_m$ with each q_i a prime. Then $n = m$ and there exists $\pi \in \mathrm{Sym}(n)$ with $q_i \sim p_{\pi(i)}$.

Proof. (a) Suppose first that $n = 1$. Then $q \mid p_1$. Since p_1 is a prime, p_1 is irreducible and so by 3.3.9, q is a unit or $q \sim p_1$. q is a prime and so not a unit. hence $q \sim p_1$ and (a) holds for $n = 1$.

Suppose $n > 1$ and put $b = p_2 \ldots p_n$. Then $a = p_1 b$ and so $q \mid p_1 b$. Since q is a prime, $q \mid p_1$ or $q \mid b$. If $q \mid p_1$, the $n = 1$ case gives $q \sim p_1$. If $q \mid b$, then induction on n implies $q \sim p_j$ for some $j \leq 2 \leq n$.

(b) We may assume that $n \leq m$. Suppose first that $m = 1$. Then also $n = 1$ and $p_1 = q_1$.

Suppose now that $m > 1$. Since $q_m \mid a$ we conclude from (a) that $q_m \sim p_i$ for some $1 \leq i \leq n$. Without loss, $i = n$. Then $p_n \sim q_m$ and so $up_n = q_m$ for some unit $u \in R$. hence $q_{m-1}q_m = q_{m-1}up_n = (uq_{m-1}p_n$ and q_m by p_n and q_{m-1} by uq_{m-1} we may assume that $p_n = q_m$ Thus

$$(p_1 \ldots p_{n-1})p_n = a = (q_1 \ldots q_{m-1})p_n$$

Suppose that $n = 1$, then by 3.3.4 $(q_1 \dots q_{m-1})$ is a unit. It follows that $q_1 q_2 \dots q_{m-1} u = 1_R$ for some $u \in R$. But then q_1 is a unit, a contradiction to the definition of a prime. Thus $n > 1$.

Since R is an integral domain, R has no non-zero zero-divisors and so the Cancellation Law 3.1.12 implies

$$p_1 \dots p_{n-1} q_1 \dots q_{m-1}$$

So by induction on $n-1 = m-1$ and there exists $\mu \in Sym(n-1)$ with $q_i \sim p_{\mu(i)}$ for all $1 \leq i \leq n-1$. Thus the lemma holds if we define $\pi \in \mathrm{Sym}(n)$ by $\pi(n) = n$ and $\pi(i) = \mu(i)$ for $1 \leq i \leq n-1$. □

Definition 3.3.13. *A* unique factorization domain *(*UFD*) is an integral domain in which every proper element is a product of primes.*

Lemma 3.3.14. *Let R be a UFD and $r \in R$. Then r is a prime if and only if r is irreducible.*

Proof. By 3.3.10 each prime in R is irreducible. Now let r be irreducible. Then by definition of a UFD, $r = p_1 \dots p_n$ where each p_i is a prime. Then $p_1 \mid r$. Since r is irreducible we conclude from 3.3.9 that p_1 is a unit or $p_1 \sim r$. p_1 is not a unit and so $p_1 \sim r$. Since associates of primes are primes (3.3.10(b)) r is a prime. □

Our next goals is to show that every PID is a UFD. For this we need a couple of preparatory lemmas.

Lemma 3.3.15. *Let $\mathcal{I}$ be chain of ideals in the ring R. If $\bigcup \mathcal{I}$ is finitely generated as an ideal, then $\bigcup \mathcal{I} \in \mathcal{I}$.*

Proof. Suppose that $\bigcup \mathcal{I} = (F)$ for some finite $F \subseteq \bigcup \mathcal{I}$. For each $f \in F$ there exists $I_f \in \mathcal{I}$ with $f \in I_f$. Since $\mathcal{I}$ is totally ordered, the finite set $\{I_f \mid f \in F\}$ has a maximal element I then $I \in \mathcal{I}$, $F \subseteq I$ and so

$$\bigcup \mathcal{I} = (F) \subseteq I \subseteq \bigcup \mathcal{I}$$

Thus $I = \bigcup \mathcal{I}$. □

Lemma 3.3.16. *Let R be an integral domain and $\mathcal{I}$ a non-empty set of principal ideals. Then one of the following holds:*

1. $\bigcap \mathcal{I} = \{0_R\}$.

2. $\mathcal{I}$ has a minimal element.

3. There exists an infinite strictly ascending series of principal ideals.

Proof. Assume that (2) does not hold. We claim that there exists an infinite descending series

$$Ra_1 \supsetneq Ra_2 \supsetneq \dots Ra_n \supsetneq Ra_{n+1} \supsetneq \dots$$

with $Ra_i \in \mathcal{I}$ for all i. Indeed since $\mathcal{I}$ is not empty there exists $I_1 \in \mathcal{I}$. Since I_1 is principal, $I_1 = Ra_1$ for some a_1. Suppose that we already found $Ra_i \in \mathcal{I}$, $1 \leq i \leq n$ with $Ra_i \supsetneq Ra_{i+1}$ for all $1 \leq i < n$. By assumption Ra_n is not a minimal element of $\mathcal{I}$ and so $Ra_n \supsetneq Ra_{n+1}$ for some $Ra_{n+1} \in \mathcal{I}$. So our claim holds by induction.

If $\bigcap_{i=1}^{\infty} Ra_i = \{0_R\}$, then also $\bigcap \mathcal{I} = \{0_R\}$ and (1) holds. Hence we may assume that there exists $0_R \neq a$ with $a \in Ra_n$ for all $n \in \mathbb{Z}^+$. Then $a = r_n a_n$ for some $r_n \in R$. Since $a_{n+1} \in Ra_n$, $a_{n+1} = sa_n$ for some $s \in R$. Thus

$$r_n a_n = a = r_{n+1} a_{n+1} = r_{n+1} s a_n$$

As R is an integral domain, $r_n = r_{n+1}s$ and so $Rr_n \subseteq Rr_{n+1}$. If $Rr_{n+1} = Rr_n$, then $Rr_{n+1} = Rsr_{n+1}$. So $r_{n+1} \sim sr_{n+1}$ and by 3.3.4, s is a unit. As $a_{n+1} = sa_n$ we conclude that $Ra_n = Ra_{n+1}$, a contradiction. Thus

$$Rr_1 \subsetneq Rr_2 \subsetneq \dots Rr_n \subsetneq Rr_{n+1} \subseteq \dots$$

is an infinite strictly ascending series of ideal and (3) holds. □

Lemma 3.3.17. *Let R be a ring in which every ideal is finitely generated.*

(a) Any nonempty set of ideals in R has a maximal member.

(b) Suppose in addition that R is an integral domain. Then every non empty set of principal ideals with nonzero intersection has a minimal member.

Proof. (a) By 3.3.15, R has no strictly ascending series of ideals. Thus (a) holds.
(b) follows from (a) and 3.3.16. □

Lemma 3.3.18. *Every principal ideal domain is a unique factorization domain.*

Proof. Let S be the set of proper elements in R which can be written as a product of primes. Let a be proper in R. We will first show

1°. *a is divisible by a prime.*

Indeed, by 3.2.15 there exists be a maximal ideal I with with $(a) \subset I$. Since R is a PID, $I = (s)$ for some $s \in R$. Then by 3.3.9 s is irreducible and so by 3.3.11 s is a prime. Since $(a) \subseteq (s)$, $s \mid a$ and $(1°)$ holds.

2°. *Put $\mathcal{S} = \{(s) \mid s \in S, s \mid a\}$. Then $\mathcal{S} \neq \emptyset$.*

By (1°) there exists a prime s with $s \mid a$. Then $s \in S$ and so (2°) holds.

By (2°) and 3.3.17b, $\mathcal{S}$ has a minimal member, say (b) with $b \in S$. Since $b \mid a$, $a = ub$ for some $u \in R$. Suppose that u is not a unit. Then by (1°) applied to u, there exists a prime p dividing a. Then pb divides a and $pb \in S$. Thus $(pb) \in \mathcal{S}$ and $(pb) \subseteq (b)$. The minimality of (b) gives $(b) = (pb)$. Hence $pb \mid b$ and so $pb \mid b$. But then by 3.3.4, p is a unit, a contradiction.

Thus u is a unit. Since b is a product of primes and any associate of a prime is a prime, we conclude that a is a product of primes. □

3.4 Euclidean Rings

Definition 3.4.1. *Let R be a commutative ring.*

(a) A pre-Euclidean *function on R is a function $d : R \to \Lambda$, where Λ is a well-ordered set, such that for all $a, b \in R$ with $b \neq 0_R$*

(i) $d(0_R) < d(b)$ and

(ii) if $d(b) \leq d(a)$, then there exists $t \in R$ with $d(a - tb) < d(a)$

(b) R is called an Euclidean domain *if R is an integral domain and there exists an pre-Euclidean function on R.*

Example 3.4.2. 1. Let $d : \mathbb{Z} \to Z, m \to |m|$ be the absolute value function. Let $a, b \in \mathbb{Z}$ and $0 < |b| \leq |a|$. If a and b are both positive or both negative, then $|a - b| < |a|$. If one of a, b is positive and the other negative, then $|a + b| > |a|$. So d is a pre-Euclidean function. Thus $\mathbb{Z}$ is an Euclidean domain.

2. Let $\mathbb{F}$ be any field, $\Lambda = \{-\infty\} \cup \mathbb{N}$. Let $0_F \neq f, g \in \mathbb{F}[x]$ of degree n and m respectively. Suppose that $n < m$. Let a and b be the leading coefficients of f and g, respectively. $ba^{-1}x^{m-n}f$ is a polynomial of degree m and leading coefficient b. Thus $g - ba^{-1}x^{m-n}f$ has degree less than g and so d is a pre-Euclidean function.

   Note also that fg is a polynomial of degree x^{n+m} with leading coefficient ab. Thus $fg \neq 0_{\mathbb{K}}$ and so $\mathbb{F}[x]$ is an integral domain. Hence $\mathbb{F}[x]$ is a Euclidean domain.

Lemma 3.4.3. *Let d be a pre-Euclidean function on a ring R. Let $a, b \in R$ with $b \neq 0_R$. Then there exist $q, r \in R$ with*

$$a = qb + r \text{ and } d(r) < d(b).$$

Proof. Since the co-domain of d is well ordered we can choose $q \in R$ with

$$(*) \qquad d(a - qb) = \min\{d(a - sb) \mid s \in R\}$$

Put $r = a - qb$ and suppose that $d(r) \geq d(b)$. Then $r \neq 0_R$ and by the definition of a pre-Euclidean function there exists $t \in R$ such that $d(r - tb) < d(r)$. But $r - tb = a - qb - tb = a - (q + t)b$ and we obtain a contradiction (*). Hence $d(r) < d(b)$ and the lemma is proved. □

Definition 3.4.4. *Let R be an ring, Λ a well-ordered set and $d : R \to \Lambda$ a function such that for all $a, b \in R$ with $b \neq 0_R$:*

(i) $d(0_R) < d(b)$.

(ii) If $ab \in R^\sharp$, *then* $d(b) \leq d(ab)$.

(iii) There exists q, r *in* R *with*

$$a = sb + r \text{ and } d(r) < d(b).$$

Then d is called a Euclidean function

Lemma 3.4.5. *Let R be a ring with identity and d a pre-Euclidean function on R. Let $a \in R$. If $a = 0_R$ define $d^*(a) = d(a)$, otherwise put*

$$d^*(a) = \min\{d(b) \mid 0_R \neq b \in Ra\}.$$

Then d^* is a Euclidean function.

Proof. Since R has an identity, $a = 1_R a \in Ra$ and $d^*(a) \leq d(a)$. We need to verify the conditions i- iii in the definition of an Euclidean function. For $x \neq 0_R$ in R we choose $\tilde{x}^* \in Rx$ with $d \neq 0_R$ and $d^*(x) = d(x^*)$.

(ii): By definition of a pre-Euclidean function $d(0_R) < d(a^*)$ and so $d^*(0_R) < d^*(a)$.

(ii): Suppose $ab \neq 0_R$. We have $(ab)^* \in Rab \subseteq Rb$ and so by definition of $d^*(b)$, $d^*(b) \leq d((ab)^*) = d^*)ab)$.

(iii): By 3.4.3 there exists $\tilde{q}$ and $\tilde{r}$ with

$$a = \tilde{q}b^* + r^* \text{ and } d(\tilde{r}) < d(b^*).$$

Since $b^* \in Rb$, $b^* = tb$ for some $t \in R$. Put $s = \tilde{q}t$ and $r = \tilde{q}$. Then

$$a = \tilde{s}b^* + \tilde{r} = \tilde{q}ta + \tilde{r} = qb + r$$

and

$$d^*(r) \leq d(r) < d(b^*) = d^*(b).$$

So d^* is indeed an Euclidean function. □

Theorem 3.4.6. *Let d be a pre-Euclidean function on the ring R and I a non-zero left ideal in R. Let $0_R \neq a \in I$ with $d(a)$ minimal, then $I = Ra$. In particular every Euclidean domain is a PID.*

Proof. So suppose that $I \neq \{0_R\}$. Let $0_R \neq b \in I$ with $d(b)$ minimal. Let $a \in I$. By 3.4.3 there exist $q, r \in R$ such that $a = qb + r$ with

$$d(r) < d(b)$$

Since $r = a - qb$ and both a, b are in I we get $r \in I$. So the minimal choice of $d(b)$ implies $r = 0_R$. Thus $a = qb$ and so $a \in Rb$. Thus $I = Rb$. □

Next we introduce greatest common divisor in arbitrary commutative rings. But the reader should be aware that often no greatest common divisor exist.

Definition 3.4.7. *Let X be a subset of the commutative ring R and $d \in R$*

(a) We say d is a common divisor *of X and write $d \mid X$ if $X \subseteq (d)$, that is if $d \mid x$ for all $d \in X$.*

(b) We say that d is greatest common divisor *of X and write $d \sim \gcd(X)$ if $d|X$ and $e|d$ for all $e \in R$ with $e|X$.*

(c) We say that X is relatively prime *if all commons divisors of X are units.*

Note that if a greatest common divisor exists it is unique up to associates. A common divisor exists if and only if (X) is contained in a principal ideal. A greatest common divisor exists if and only if the intersection of all principal ideals containing X is a principal ideal. (Here we define the intersection of the empty set of ideals to be the ring itself). The easiest case is then (X) itself is a principal ideal. Then the greatest common divisors are just the generators of (X). An element in (X) generates (X) if and only if it's a common divisor. So if the ring has an identity, (X) is a principal ideal if and only if X has a common divisor of the form $\sum_{x \in X} r_x x$, where as usually all but finitely many r_x's are supposed to be 0.

Note that from the above we have the following statement:

Every subset of R has a greatest common divisor if and only if any intersection of principal ideals is a principal ideal. That is if and only if the set of principal ideals in R is closed under intersections.

In particular, greatest common divisors exists in PID's and can be expressed as a linear combination of the X.

Greatest common divisors still exists in UFD's, but are no longer necessarily a linear combination of X. Indeed let $\mathcal{P}$ be a set of representatives for the associate classes of primes. For each $0 \neq r \in R$,

$$x = u_r \prod_{p \in \mathcal{P}} p^{m_p(r)}$$

for some $m_p(r) \in \mathbb{N}$ and a unit u_r. Let

$$m_p = \min_{x \in X} m_p(x).$$

Since $m_p \leq m_p(x)$ only finitely many of the m_p are nonzero. So we can define

$$d = \prod_{p \in \mathcal{P}} p^{m_p}.$$

A moments thought reveals that d is a greatest common divisor.

Here are a couple of concrete examples which might help to understand some of the concepts we developed above.

First let $R = \mathbb{Z}[i]$, the subring of $\mathbb{C}$ generated by i. R is called the ring of *Gauian integers.*

Note that $R = \mathbb{Z} + \mathbb{Z}i$. We will first show that R is an Euclidean ring. Indeed, put $\phi(a_1 + a_2 i) = a_1^2 + a_2^2$. Then $\phi(xy) = \phi(x)\phi(y)$ and $\phi(x) \in \mathbb{Z}^+$. So $(ER1)$ holds. Let $x, y \in R$ with $x \neq 0$. Put $z = \frac{y}{x} \in \mathbb{C}$. Then $y = zx$. Also there exists $d = d_1 + d_2 i \in \mathbb{C}$ with $q := z - d \in R$ and $|d_i| \leq \frac{1}{2}$. In particular, $\phi(d) \leq \frac{1}{2}^2 + \frac{1}{2}^2 = \frac{1}{2}$. Put $r = y - qx$ then $r = zx - qx = (z - q)x = dx$. So $\phi(r) = \phi(d)\phi(x) \leq \frac{1}{2}\phi(x)$. Hence also $(ER2)$ holds.

Let a be a prime in R and put $P = (a)$. Since $\phi(a) = \bar{a}a \in P$, $P \cap \mathbb{Z} \neq 0$. Also $1 \notin P$ and so $P \cap Z$ is a proper ideal in $\mathbb{Z}$. Since R/P has no zero divisors, $\mathbb{Z} + P/P \cong \mathbb{Z}/P \cap \mathbb{Z}$ has no zero divisors. Thus $P \cap \mathbb{Z} = p\mathbb{Z}$ for some prime integer p. Let $Q = pR$. Then $Q \leq P \leq R$. We will determine the zero divisors in R/Q. Indeed suppose that $ab \in Q$ but neither a nor b are in Q. Then p^2 divides $\phi(ab)$. So we may assume that p divides $\phi(a)$. Hence $a_1^2 = -a_2^2 \pmod{p}$. If p divides a_1 it also divides a_2, a contradiction to $a \notin Q$. Therefore we can divide by $a_2 \pmod{p}$ and conclude that the equation $x^2 = -1$ has a solution in $\mathbb{Z}/p\mathbb{Z}$. Conversely, if $n^2 \equiv -1 \pmod{p}$ for some integers n we see that (up to associates) $n + i + Q$ and $n - i + Q$ are the only zero divisors.

Suppose that no integer n with $n^2 \equiv -1 \pmod{p}$ exists. Then R/Q is an integral domain and so a field. Hence $Q = P$ and $a \sim p$ in this case.

Suppose that n is an integer with $n^2 \equiv -1 \pmod{p}$. As P is a prime ideal and $(n + i)(n - i) \in Q \leq P$, one of $n \pm i$ is in P. We conclude that $a \sim n \pm i$.

Next let $R = \mathbb{Z}[\sqrt{10}]$. We will show that R has some irreducible elements which are not primes. In particular, R is neither UFD, PID or Euclidean. Note that $R = \mathbb{Z} + \mathbb{Z}\sqrt{10}$. For $r \in R$ define $r_1, r_2 \in \mathbb{Z}$ by $r = r_1 + r_2\sqrt{10}$. Define $\tilde{r} = r_1 - r_2\sqrt{10}$ and $N(r) = r\tilde{r} = r_1^2 - 10r_2^2$. $N(r)$ is called the *norm* of r. We claim that $r \to \tilde{r}$ is a ring automorphism of R. Clearly it is an automorphism of $(R, +)$. Let $r, s \in R$. Then

$$rs = (r_1 + r_2\sqrt{10})(s_1 + s_2\sqrt{10}) = (r_1 s_1 + 10 r_2 s_2) + (r_1 s_2 + r_2 s_2)\ \sqrt{10}$$

It follows that $\widetilde{rs} = \tilde{r}\tilde{s}$. In particular,

$$N(rs) = rs\widetilde{rs} = rs\tilde{r}\tilde{s} = r\tilde{r}s\tilde{s} = N(r)N(s)$$

and $N : R \to \mathbb{Z}$ is a multiplicative homomorphism. Let r be a unit in R. Since $N(1) = 1$, we conclude that $N(r)$ is unit in $\mathbb{Z}$ and so $N(r) = \pm 1$. Conversely, if $N(r) = \pm 1$, then $r^{-1} = \frac{\tilde{r}}{N(r)} = N(r)\tilde{r} \in R$ and r is a unit. For example $3 + \sqrt{10}$ is unit with inverse $-3 + \sqrt{10}$. As $\sqrt{10}$ is not rational, $N(r) \neq 0$ for $r \in R^{\#}$.

We claim that all of $2, 3$, $f := 4 + \sqrt{10}$ and $\tilde{f}$ are irreducible. Indeed suppose that ab is one of those numbers and neither a nor b are units. Then $N(a)N(b) \in \{4, 9, 6\}$ and so $N(a) \in \{\pm 2, \pm 3\}$ and

$$N(a) \equiv 2, 3 \pmod{5}$$

But for any $x \in R$ we have

$$N(a) \equiv a_1^2 \equiv 0, 1, 4 \pmod{5}$$

So indeed $2, 3, f$ and $\tilde{f}$ are primes. Note that $2 \cdot 3 = 6 = -f\tilde{f}$. Hence 2 divides $f\tilde{f}$ but (as f and $\tilde{f}$ are irreducible) 2 divides neither f nor $\tilde{f}$. So 2 is not a prime. With the same argument none of 3, f and $\tilde{f}$ are not primes.

We claim that every proper element in R is a product of irreducible. Indeed let a be proper in R and suppose that a is not irreducible. Then $a = bc$ with neither b nor c units. Then as $N(a) = N(b)N(c)$ both b and c have smaller norm as a. So by induction on the norm, both b and c can be factorized into irreducible.

Since R has irreducibles which are not primes, we know that R can not be a PID. But let us verify directly that $I = (2, f) = 2R + fR$ is not a principal ideal. First note that $f\tilde{f} = -6 \in 2R$. Since also $2f \in 2R$ we $I\tilde{f} \in 2R$. Since 4 does not divide $N(f)$, $f \notin 2R$ and so I does not contain a unit. Suppose now that h is a generator for I. Then h is not a unit and divides f. So as f is irreducible, $h \sim f$ and $I = (f)$. But every element in (f) has norm divisible by $N(f) = 6$, a contradiction to $2 \in I$ and $N(2) = 4$.

3.5 Localization

Let R be a commutative ring and $\emptyset \neq S \subseteq R$.In this section we will answer the following question:

Does there exists a commutative ring with identity R' so that R is a subring of R' and all elements in S are invertible in R ?

Clearly this is not possible if $0 \in S$ or S contains zero divisors. It turns out that this condition is also sufficient. Note that if all elements in S are invertible in R', also all elements in the subsemigroup of $(R, \cdot)$ generated by S are invertible in R'. So we may assume that S is closed under multiplication:

Definition 3.5.1. *A* multiplicative subset *of the ring R is a nonempty subset S with $st \in S$ for all $s, t \in S$.*

Lemma 3.5.2. *Let S be a multiplicative subset of the commutative ring R. Define the relation $\sim$ of $R \times S$ by*

$$(r_1, s_1) \sim (r_2, s_2) \quad \text{if} \quad t(r_1 s_2 - r_2 s_1) = 0 \text{ for some } t \in S.$$

Then $\sim$ is an equivalence relation.

Proof. $\sim$ is clearly reflexive and symmetric. Suppose now that $(r_1, s_1) \sim (r_2, s_2)$ and $(r_2, s_2) = (r_3, s_3)$. Pick t_1 and t_2 in S with

$$t_1(r_1 s_2 - r_2 s_1) = 0 \text{ and } t_2(r_2 s_3 - r_3 s_2) = 0.$$

Multiply the first equation with $t_2 s_3$ and the second one with $t_1 s_1$. Then both equation contain the term $t_1 t_2 r_2 s_1 s_3$ but with opposite sign. So adding the two resulting equations we see:

$$0 = t_2 s_3 t_1 r_1 s_s - t_1 s_1 t_2 r_3 s_2 = t_1 t_2 s_2 (r_1 s_3 - r_3 s_1)$$

Thus $(r_1, s_1) \sim (r_3, s_3)$. □

Let S,R and $\sim$ as in the previous lemma. Then $S^{-1}R$ denotes the set of equivalence classes of $\sim$. $\frac{r}{s}$ stands for the equivalence class containing (r, s). Note that if $0 \in S$, $\sim$ has exactly one equivalence class. If R has no zero divisors and $0 \notin S$ then $\frac{r_1}{s_1} = \frac{r_2}{s_2}$ if and only if $r_1 s_2 = r_2 s_1$.

Proposition 3.5.3. *Let S be a multiplicative subset of the commutative ring R and $s \in S$*

(a) The binary operations

$$\frac{r}{s} + \frac{r'}{s'} = \frac{rs' + r's}{ss'} \quad \text{and} \quad \frac{r}{s} \cdot \frac{r'}{s'} = \frac{rr'}{ss'}$$

on $S^{-1}R$ are well-defined.

(b) $(S^{-1}R, +, \cdot)$ is an ring.

(c) $\frac{s}{s}$ is an identity.

(d) The map

$$\phi_S : R \to S^{-1}R, \quad r \to \frac{rs}{s}$$

is a ring homomorphism and independent from the choice of s.

(e) $\phi_S(s)$ is invertible.

Proof. (a) By symmetry it suffices to check that the definition of $+$ and $\cdot$ does not depend on the choice of (r, s) in $\frac{r}{s}$. Let $\frac{r_1}{s_1} = \frac{r}{s}$ so $t(rs_1 - r_1 s) = 0$ for some $t \in S$.

Then

$$t[(rs' + r's)s_1 s' - (r_1 s' + r' s_1)ss'] = t(rs_1 - r_1 s)s's' = 0$$

and so $+$ is well defined. Also

$$t(rr's_1 s' - r_1 r' ss') = t(rs_1 - r_1 s)r's' = 0$$

and so $\cdot$ is well defined.

(b) It is a routine exercise to check the various rules for a ring. Maybe the least obvious one is the associativity of the addition:

$$(\frac{r_1}{s_1} + \frac{r_2}{s_2}) + \frac{r_3}{s_3} = \frac{r_1 s_2 + r_2 s_1}{s_1 s_2} + \frac{r_3}{s_3} = \frac{r_1 s_2 s_3 + r_2 s_1 s_3 + r_3 s_1 s_2}{s_1 s_2 s_3} = \frac{r_1}{s_1} + \frac{r_2 s_3 + r_3 s_2}{s_2 s_3} = \frac{r_1}{s_1} + (\frac{r_2}{s_2} + \frac{r_3}{s_3})$$

We leave it to the reader to check the remaining rules.

(c) and (d) are obvious. For (e) note that $\phi_S(s) = \frac{s^2}{s}$ has $\frac{s}{s^2}$ as its inverse. □

The ring $S^{-1}R$ is called the ring if fraction of R by S. It has the following universal property:

Proposition 3.5.4. *Let R be a commutative ring, S_0 a non-empty subset and S the multiplicative subset of R generated by S. Suppose that R' is a commutative ring with identity and $\alpha(R) \to R'$ is a ring isomorphism so that $\alpha(s_0)$ is a unit for all $s_0 \in S_0$. Then there exists a unique homomorphism*

$$\alpha^* : S^{-1}R \to R' \text{ with } \phi_S(r) \to \alpha(r).$$

Moreover,

$$\alpha^*(\frac{r}{s}) = \alpha(r)\alpha(s)^{-1}.$$

Proof. Note that as $\alpha(S_0)$ consists of units so does $\alpha(S)$. So once we verify that

$$\alpha^*(\frac{r}{s}) = \alpha(r)\alpha(s)^{-1}.$$

is well defined, the remaining assertion are readily verified.

So suppose that $\frac{r_1}{s_1} = \frac{r_2}{s_2}$. Then

$$t(r_1s_2 - r_2s_1) = 0$$

for some $t \in S$. Applying α we conclude

$$\alpha(t)(\alpha(r_1)\alpha(s_2) - \alpha(r_2)\alpha(s_1)) = 0.$$

As $\alpha(t), \alpha(s_1)$ and $\alpha(s_2)$ are units, we get

$$\alpha(r_1)\alpha(s_1)^{-1} = \alpha(r_2)\alpha(s_2)^{-1}$$

Hence α^* is indeed well-defined. □

When is $\frac{r}{s} = 0$? The zero element in $S^{-1}R$ is $\frac{0}{s}$ Hence $\frac{r}{s} = 0$ if and only if there exists $t \in S$ with $0 = t(rs - 0s) = trs$ for some $t \in S$. This is true if and only if $tr = 0$ for some $t \in S$. Put

$$R_S = \{r \in R \mid tr = 0 \text{ for some } t \in S\}.$$

So $\frac{r}{s} = 0$ if and only $r \in R_S$.

What is the kernel of $\phi = \phi_S$? $\phi(r) = \frac{rs}{s}$. Hence $r \in \ker \phi$ if and only if $rs \in R_S$ and so if and only if $r \in R_S$. Thus $\ker \phi = R_S$. So ϕ is one to one if and only if $R_S = 0$. This in turn just means that S contains no zero divisors. In this is the case we will identify R with it image in $S^{-1}R$

Let $\check{R}$ be the set of all non-zero, non zero divisors. and assume $\check{R} \neq \emptyset$. We claim that $\check{R}$ is a multiplicative set. Indeed let $s, t \in \check{R}$. Suppose that $rst = 0$ for some $r \in R$. Then as $t \in \check{R}$, $rs = 0$ and as $s \in \check{R}$, $r = 0$ so $st \in \check{R}$. The $\check{R}^{-1}R$ is called the *complete ring of fraction* of R.

If R has no zero divisors, then $\check{R} = R^{\#}$ and the complete ring of fraction is a field. This field is called the *field of fraction* of R and denoted by $\mathbb{F}_R$.

The standard example is $R = \mathbb{Z}$. Then $\mathbb{F}_{\mathbb{Z}} = \mathbb{Q}$.

If $\mathbb{K}$ is a field then $\mathbb{F}_{\mathbb{K}[x]} = \mathbb{K}(x)$, the field of rational functions over K. Slightly more general if R has no-zero divisors then $\mathbb{F}_{R[x]} = \mathbb{F}_R(x)$, the field of rational function over the field of fractions of R.

We will now spend a little but of time to investigate the situation where S does contain some zero divisors.

Define

$$\phi^* : S^{-1}R \to \phi(S)^{-1}\phi(R), \quad \frac{r}{s} \to \frac{\phi(r)}{\phi(s)}$$

We claim that ϕ^* is a well defined isomorphism. For this we prove the following lemma.

Lemma 3.5.5. *Let $\alpha : R \to R'$ be a homomorphism of commutative rings and S and S' multiplicative subsets of R and R' respectively. Suppose that $\alpha(S) \subseteq S'$.*

(a) $\alpha(S)$ is a multiplicative subset of R'.

(b)

$$\alpha^* : S^{-1}R \to S'^{-1}R', \quad \frac{r}{s} \to \frac{\alpha(r)}{\alpha(s)}$$

is a well defined homomorphism.

(c) Suppose that $S' = \alpha(S)$. Then

$$\ker \alpha^* = \{\frac{r}{s} \mid r \in R, s \in S, Sr \cap \ker \alpha \neq \emptyset\} \text{ and } \alpha^*(S^{-1}R) = \alpha(S)^{-1}\alpha(R)$$

Proof. (a) Just note that $\alpha(s)\alpha(t) = \alpha(st)$ for all $s, t \in S$.

(b) Note that $\phi_{S'}(\alpha(s))$ is invertible. Hence α^* is nothing else as the homomorphism given by 3.5.4 applied to the homomorphism:

$$\phi_{S'} \circ \alpha : R \to S'^{-1}R'$$

(c) Let $\frac{r}{s} \in \ker \alpha^*$. As seen above this means $t'\alpha(r) = 0$ for some $t' \in S'$. By assumption $t' = \alpha(t)$ for some $t \in T$. Thus $\frac{r}{s} = 0$ if and only if $tr \in \ker \alpha$ for some $t \in S$.

That $\alpha^*(S^{-1}R) = \alpha(S)^{-1}\alpha(R)$ is obvious. □

Back to the map ϕ^*. By the previous lemma ϕ^* is a well defined homomorphism and onto. Let $\frac{r}{s} \in \ker \phi^*$. Then $tr \in \ker \phi$ for some and $t \in S$. As $\ker \phi = R_S$, $\tilde{t}tr = 0$ for some $\tilde{t} \in S$. Hence $r \in R_S$ and $\frac{r}{s} = 0$. Therefore ϕ^* is one to one and so an isomorphism.

Note also that $\phi(R) \cong R/R_S$. Let $\bar{R} = R/R_S$ and $\bar{S} = S + R_S/R_S$. As ϕ^* is an isomorphism we get

$$S^{-1}R \cong \bar{S}^{-1}\bar{R}$$

We have $\bar{R}_{\bar{S}} = 0$. So in some sense we can always reduce to the case where S has no zero divisors.

In the next lemma we study the ideals in $S^{-1}R$. For $A \subset R$ and $T \subseteq S$ put

$$T^{-1}A = \{\frac{a}{t} \mid a \in A, t \in T\}$$

Proposition 3.5.6. *Let S be a multiplicative subset of the commutative ring R.*

(a) If I is an ideal in R then $S^{-1}I$ is an ideal in $S^{-1}R$

(b) If J is an ideal in R then $I = \phi_S^{-1}(J)$ is an ideal in R with $J = S^{-1}I$.

(c) The map $I \to S^{-1}I$ is a surjection from the set of ideals in I to the set of ideals to $S^{-1}R$.

Proof. Put $\phi = \phi_S$.

(a) is readily verified.

(b) Inverse images of ideals are always ideals. To establish the second statement in (b) let $j = \frac{r}{s} \in J$. As J is an ideal

$$\frac{s^2}{s}\frac{r}{s} = \frac{rs^2}{s^2} = \frac{rs}{s} \in J$$

Thus $\phi(r) \in J$ and $r \in I$. So $j = \frac{r}{s} \in S^{-1}I$.

Conversely, if $r \in I$ and $s \in S$ then since $\phi(r) \in J$ an dJ is an ideal:

$$\frac{r}{s} = \frac{rs^2}{s^3} = \frac{s}{s^2}\frac{rs}{s} = \frac{s}{s^2}\phi(r) \in \frac{s}{s^2}J \subseteq J$$

So (b) holds.

(c) follows from (a) and (b). □

If R has an identity the previous proposition can be improved:

Proposition 3.5.7. *Let R be a commutative ring with identity and S a multiplicative subset R.*

(a) Suppose R has an identity. Let I be an ideal in R. Then

$$\phi_S^{-1}(S^{-1}I) = \{r \in R \mid Sr \cap I \neq \emptyset\}.$$

(b) Define an ideal I in R to be S^{-1}-closed if $r \in I$ for all $r \in R$ with $rS \cap I \neq \emptyset$. Then

$$\tau : I \to S^{-1}I$$

is a bijection between the S^{-1}-closed ideals and the ideals in $S^{-1}R$. The inverse map is given by

$$\tau^{-1} : J \to \phi_S^{-1}J.$$

(c) $I \cap S = \emptyset$ for all S^{-1} closed ideals with $I \neq R$.

(d) A prime ideal P in R is S^{-1} closed if an only if $P \cap S = \emptyset$.

(e) τ induces a bijection between the S^{-1} closed prime ideals in R and the prime ideals in $S^{-1}R$.

Proof. (a) Let $r \in R$ then the following are equivalent:

$\phi(r) \in S^{-1}I$.
$\phi(r) = \frac{i}{s}$ for some $i \in I, s \in S$
$\frac{r}{1} = \frac{i}{s}$ for some $i \in I, s \in S$
$t(rs - i) = 0$ for some $i \in I, s, t \in S$
$tsr = ti$ for some $i \in I, s, t \in S$
$sr \in I$ for some $s \in S$
$Sr \cap I \neq \emptyset$.

So (a) holds.

We write ϕ for ϕ_S and say "closed" for S^{-1} closed.

(b) follows from (a) and 3.5.6b,c.

(c) Suppose I is closed and $s \in S \cap I$. Then $S^{-1}I$ contains the unit $\phi(s)$ and so $S^{-1}I = S^{-1}R$. Thus $I = \phi^1(S^{-1}R$ and $I = R$.

(d) Let P be a prime ideal in R. Suppose that $S \cap P = \emptyset$ and let $r \in R$ and $s \in S$ with $rs \in P$. Then by 3.2.13b, $r \in P$ or $s \in P$. By assumption $s \notin P$ and so $r \in P$. Thus P is closed. Conversely, if P is closed, (c) implies $P \cap S = \emptyset$.

(e) Let P be a closed prime ideal. We claim the $S^{-1}R$ is a prime ideal in R. First since τ is an bijection, $S^{-1}P \neq R$. Suppose that $\frac{r}{s}\frac{r'}{s'} \in S^{-1}P$. Then also $\phi(rr') \in S^{-1}P$. As P is closed, $rr' \in P$. As P is prime we may assume $r \in P$. But then $\frac{r}{s} \in S^{-1}P$ and so $S^{-1}P$ is a prime.

Suppose next that I is closed and $S^{-1}I$ is a prime. If $rr' \in I$, then $\phi(r)\phi(r') \in S^{-1}I$. As $S^{-1}I$ is prime we may assume that $\phi(r) \in S^{-1}I$. As I is closed this implies $r \in I$ and so I is a prime ideal. □

Let R be a commutative ring with identity. By 3.2.13 an ideal P in R is a prime ideal if and only if $R \setminus P$ is a multiplicative subset of R. Let P be the prime ideal. The ring

$$R_P := (R \setminus p)^{-1}R$$

is called the *localization* of R at the prime P. For $A \subseteq R$ write A_P for $(R \setminus p)^{-1}A$.

Theorem 3.5.8. *Let P be a prime ideal in the commutative ring with identity R.*

(a) The map $Q \to Q_P$ is a bijection between the prime ideals of R contained in P and the prime ideals in R_P.

(b) P_P is the unique maximal ideal in R_P. $r \in R$ is a unit if and only if $r \notin P_P$.

Proof. (a) Put $S = R \setminus P$ and let Q a prime ideal in R. Then $Q \cap S = \emptyset$ if and only if $Q \subset P$. Thus (a) follows from 3.5.7.

(b) Let I be a maximal ideal in R_P. Then by 3.2.16 I is prime ideal. Thus by (a) $I = Q_P$ for some $Q \subseteq P$. Thus $I \subseteq P_P$ and $I = P_P$. The statement about the units now follows from 3.3.3.

Actually we could also have argued as follows: all elements in $R_P \setminus P_P$ are of the form $\frac{s}{s'}$ and so invertible. Hence by 3.3.3 P_P is the unique maximal ideal in R. □

Definition 3.5.9. *A* local ring *is a commutative ring with identity which as a unique maximal ideal.*

Using 3.3.3 we see

Lemma 3.5.10. *Let R be a commutative ring with identity. The the following are equivalent:*

(a) R is a local ring.

(b) All the non-units are contained in an ideal $M \lneq R$.

(c) All the non-units form an ideal. □

We finish this section with some examples.

Let p be an prime integer. Then $\mathbb{Z}_{(p)} = \{\frac{n}{m} \in \mathbb{Q} \mid p \nmid m\}$. Since 0 and (p) are the only prime ideals of $\mathbb{Z}$ contained in (p), O and $\frac{n}{m} \in \mathbb{Q} \mid p \nmid m, p \mid n\}$ are the only prime ideal in $\mathbb{Z}_{(p)}$. What are the ideals? Every non zero ideal of $\mathbb{Z}$ is of the form (t) for some $t \in \mathbb{Z}^+$. Write $t = ap^k$ with $p \nmid a$. Suppose (t) is closed. As $a \in S = \mathbb{Z} \setminus (p)$ and $ap^k \in (t)$ we conclude that $p^k \in (t)$. Thus $a = 1$ and $t = p^k$. It is easy to see that (p^k) is indeed closed.

So we conclude that the ideals in $\mathbb{Z}_{(p)}$ are

$$p^k\mathbb{Z}_{(p)} = \{\frac{n}{m} \in \mathbb{Q} \mid p \nmid m, p^k \mid n\}$$

In particular $\mathbb{Z}_{(p)}$ is a PID.

We reader might have notice that in the above discussion $\mathbb{Z}$ can be replaced by any PID R, p by any prime in R and $\mathbb{Q}$ by $\mathbb{F}_R$.

3.6 Polynomials rings, power series and free rings

Let R be a ring and $(A, +)$ a semigroup written additively. Since $(A, +)$ is a subgroup of the multiplicative group of R it is convenient to the following *exponential notation*'. Denote $a \in A$ be x^a and define $x^a x^b = x^{a+b}$. The the elements in $R[A]$ can be uniquely written as $f = \sum_{a \in A} f_a x^a$ where $f_a \in R$, almost all $f_a = 0$. Also

$$fg = \sum_{a \in A} \sum_{b \in A} = f_a f_b x^{a+b} = \sum_{c \in A} (\sum_{a+b=c} f_a f_b) x^c$$

Let R be a ring and I a set. Put $\mathbb{I} = \bigoplus_{i\in I} \mathbb{N}$, the free abelian monoid on I. The semigroup ring

$$R[\mathbb{I}].$$

is called the *polynomial ring* over R in the variables I and is denoted by $R[I]$

We will use the above exponential notation. Let $i \in I$. Recall that $\phi_i(1) \in \mathbb{I}$ is defined as

$$(\rho_i(1))_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

We write x_i for $x^{\rho_i(1)}$. Let $\alpha \in \mathbb{I}$, then α is a tuple $(\alpha_i)_{i\in I}$, $\alpha_i \in \mathbb{N}$, where almost all α_i are zero. So

$$x^\alpha = \prod_{i\in I} x_i^{\alpha_i}$$

. Every element $f \in R[I]$ now can be uniquely written as

$$f = \sum_{\alpha\in\mathbb{I}} f_\alpha x^\alpha$$

where $f_\alpha \in R$ and almost all f_α are zero. The element of $R[I]$ are called *polynomials*. Polynomials of the form rx^α, with $r \neq 0$, are called *monomials*. As almost all α_i are zero

$$rx^\alpha = rx_{i_1}^{\alpha_{i_1}} x_{i_2}^{\alpha_{i_2}} \dots x_{i_n}^{\alpha_{i_n}}$$

for some $i_k \in I$.

The $f_\alpha x^\alpha$ with $f_\alpha \neq 0$ are called the monomials of f. So every polynomial is the sum of its monomials.

Note that the map $r \to rx^0$ is monomorphism. So we can and do identify r with rx^0.

If $I = \{1, 2, \dots m\}$ we also write $R[x_1, \dots, x_m]$ for $R[I]$. Then every element in $R[x_1, \dots, x_m]$ can be uniquely written as

$$\sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} r_{n_1,\dots,n_m} x_1^{n_1} x_2^{n_2} \dots x_m^{n_m}$$

If I has a unique element i we write x for x_i and $R[x]$ for $R[I]$. Then each $f \in R[x]$ has the form $f = r_0 + r_1 x + r_2 x^2 + \ldots + r_n x^n$, where n is maximal with repect to $r_n \neq 0$. r_n is called the *leading coeficient* of f and n the *degree* of f,

Note that if I happens to have the structure of a semigroup, the symbol $R[I]$ has a double meaning, the polynomial ring or the semigroup ring. But this should not lead to any confusions.

If $y = (y_i, i \in I)$ is a family of pairwise commuting elements is some semigroup, and $\alpha \in \mathbb{I}$ we define

$$y^\alpha = \prod_{i\in I} y_i^{\alpha_i}$$

Note that as almost all α_i are zero and the y_i pairwise commute, this is well-defined. Also if we view the symbol x as the family $(x_i, i \in I)$, this is consistent with the x^α notation.

The polynomial ring has the following universal property:

Proposition 3.6.1. *Let $\Phi : R \to S$ be a ring homomorphism and $y = (y_i)_{i\in I}$ a family of elements in S such that*

(a) For all $r \in R$ and all $i \in I$

$$\Phi(r)y_i = y_i\Phi(r)$$

(b) for all $i, j \in I$

$$y_iy_j = y_jy_i$$

Then there exists a unique homomorphism

$$\Phi_y : R[I] \to S; \qquad with \qquad rx_i \to \Phi(r)y_i \text{ for all } r \in R, i \in I.$$

Moreover,

$$\Phi_y : \sum_{\alpha\in\mathbb{I}} f_\alpha x^\alpha \to \sum_{\alpha\in\mathbb{I}} \Phi(f_\alpha)y^\alpha.$$

Proof. As $\mathbb{I}$ is the free abelian monoid on I, (b) implies that there exists a unique homomorphism $\beta : \mathbb{I} \to (S, \cdot)$ with $\alpha(x_i) = y_i$. The existence of Φ_y now follows from 3.2.5. The uniqueness is obvious. □

The reader should notice that the assumption in the previous proposition are automatically fulfilled if S is commutative. So each $f \in R[I]$ gives rise to a function $f^\Phi : S^I \to S$ with $f^\Phi(y) = \Phi_y(f)$.

Example: Suppose $I = \{1, 2, \dots m\}$, $R = S$ is commutative and $\alpha = \mathrm{id} = \mathrm{id}_R$ and

$$f = \sum r_{n_1,\dots,n_m} x_1^{n_1} \dots x_m^{n_m}$$

Then

$$f^{\mathrm{id}}(y_1, \dots, y_n) = \sum r_{n_1,\dots,n_m} y_1^{n_1} \dots y_m^{n_m}.$$

The reader should be careful not to confuse the polynomial f with the function f^{id}. Indeed the following example shows that f^{id} can be the zero function without f being zero.

Let $R = \mathbb{Z}/p\mathbb{Z}$, $I = 1$, p a prime integer,and

$$f = x(x-1)(x-1)\dots(x-(p-1))$$

Then f is a polynomial of degree p in $(\mathbb{Z}/p\mathbb{Z})[x]$. But $f^{\mathrm{id}}(y) = 0$ for all $y \in \mathbb{Z}/p\mathbb{Z}$.

Lemma 3.6.2. *Let R be a ring and I and$J$ disjoint sets. Then there exists a unique isomorphism*

$$R[I][J] \to R[I \cup J] \text{ with } rx_i \to rx_i \text{ and } rx_j \to rx_j$$

for all $r \in R, i \in I, j \in J$.

Proof. Use 3.6.1 to show the existence of such a homomorphism and its inverse. We leave the details to the reader. □

Let R be a ring and G a semigroup. In the definition of the semigroup ring $R[G]$ we had to use the direct sum rather than the direct product since otherwise the definition of the products of two elements would involve infinite sums. But suppose G has the following property

$$(FP) \quad |\{(a,b) \in G \times G \mid ab = g\}| \text{ is finite for all } g \in G.$$

Then we can define the *power semigroup ring* of G over R, $R[[G]]$by

$$(R[[G]], +) = (\prod_{g \in G} R, +)$$

and

$$(r_g)_{g\in G} \cdot (s_g)_{g\in G} = (\sum_{(h,k)\in G\times G | hk=g} r_h s_k)_{g\in G}$$

If G is a group then it fulfills (FP) if and only if G is finite. So we do not get anything new. But there are lots of infinite semigroups with (FP). For example $G = \mathbb{N}$. $R[[\mathbb{N}]]$ is isomorphic to $R[[x]]$ the ring of formal power series. Other semigroups with (FP) are the free (abelian) monoids (or semigroups) over a set

Let I be a set. Then the power semigroup ring

$$R[[\bigoplus_{i\in I} \mathbb{N}]]$$

is called the ring of *formal power series* over R in the variables I and is denoted by $R[[I]]$. The elements of $R[[I]]$ are called formal power series. We use the same exponential notation as for the ring of polynomials. Every formal power series can be uniquely written as a formal sum

$$f = \sum_{\alpha \in \mathbb{I}} f_\alpha x^\alpha$$

Here $f_\alpha \in R$. But in contrast to the polynomials we do not require that almost all f_α are zero.

If $I = \{1\}$ the formal power series have the form:

$$f = \sum_{n=0}^{\infty} f_n x^n = f_0 + f_1 x + f_2 x^2 \dots f_n x^n \dots$$

with $f_n \in R$. Note that there does not exist an analog for 3.6.1 for formal power series, since the definition of $\Phi_y(f)$ involves an infinite sum.

Lemma 3.6.3. *Let R be ring with identity and $f \in R[[x]]$.*

(a) f is a unit if and only if f_0 is.

(b) If R is commutative and f_0 is irreducible, then f is irreducible.

Proof. (a) Note that $(fg)_0 = f_0 g_0$ and $1_0 = 1$ so if f is a unit so is f_0. Suppose now that f_0 is a unit. We define $g \in R[[x]]$ by defining its coefficients inductively as follows $g_0 = f_0^{-1}$ and for $n > 0$,

$$g_n = -f_0^{-1} \sum i = 0^{n-1} f_{n-i} g_i$$

. Note that this just says $\sum_{i=0}^{n} f_{n-i} g_i = 0$ for all $n > 0$. Hence $fg = 1$. Similarly f has a left inverse h by2.2.3 $g = h$ is a left inverses.

(b) Suppose that $f = gh$. Then $f_0 = g_0 h_0$. So as f_0 is irreducible, one of g_0, f_0 is a unit. Hence by (a) g or h is a unit. □

As an example we see that $1 - x$ is a unit in $R[[x]]$. Indeed

$$(1 - x)^{-1} = 1 + x + x^2 + x^3 + \dots.$$

Lemma 3.6.4. *Let $\mathbb{D}$ be a division ring.*

(a) $(x) = \{f \in \mathbb{D}[[x]] \mid f_0 = 0\}$

(b) The elements of (x) are exactly the non-units of $\mathbb{D}[[x]]$.

(c) Let I be a left ideal in $\mathbb{D}[[x]]$. Then $I = x^k \mathbb{D}[[x]] = (x^k)$ for some $k \in \mathbb{N}$.

(d) Every left ideal in $\mathbb{D}[[x]]$ is a right ideal and $\mathbb{D}[[x]]$ is a principal ideal ring.

(e) (x) is the unique maximal ideal in $\mathbb{D}[[x]]$.

(f) If $\mathbb{D}$ is a field, $\mathbb{D}[[x]]$ is a PID and a local ring.

Proof. (a) is obvious and (b) follows from 3.6.3.

(c) Let $k \in \mathbb{N}$ be minimal with $x^k \in I$. Let $f \in I$ and let n be minimal with $f_n \neq 0$. Then $f = x^n g$ for some $g \in \mathbb{D}[[x]]$ with $g_0 \neq 0$. Hence g is unit and $x^n = g^{-1} f \in I$. So $k \leq n$ and $f = (x^{n-k} g) x^k \in \mathbb{D}[[x]] x^k = (x^k)$. Thus $I = (x^k)$.

(d),(e) and (f) follow immediately form (c). □

3.7 Factorizations in polynomial rings

Definition 3.7.1. *Let R be a ring and I a set and $f \in R[I]$. We define the degree function*

$$\deg : R[I] \to \mathbb{N} \cup \{-\infty\}$$

as follows :

(a) if f is a monomial rx^α, then $\deg f = \sum_{i \in I} \alpha_i$,

(b) if $f \neq 0$ then $\deg f$ is the maximum of the degrees of its monomials.

(c) if $f = 0$ then $\deg f = -\infty$

Sometimes it will be convenient to talk about the degree $\deg_J f$ with respect to subset of J of I. This is defined as above, only that

$$\deg_J(rx^\alpha) = \sum_{j \in J} \alpha_j$$

Alternatively, $\deg_J f$ is the degree of f as a polynomial in $R'[J]$, where $R' = R[I \setminus J]$.

A polynomial is called *homogeneous* if all its monomials have the same degree. Let $f \in R[X]$ then f can be uniquely written as

$$f = \sum_{i=0}^{\infty} h(f, i)$$

where $h(f, i)$ is zero or a homogenous polynomial of degree i. Note here that almost all $h(f, i)$ are zero. Let $h(f) = h(f, \deg f)$.

Lemma 3.7.2. *Let R be a ring, I a set and $f, g \in R[I]$.*

(a) $\deg(f + g) \leq \max(\deg f, \deg g)$ with equality unless $h(g) = -h(f)$.

(b) If f and g are homogeneous, then fg is homogeneous. Also either $\deg(fg) = \deg(f) + \deg(g)$ or $fg = 0$.

(c) $h(fg) = h(f)h(g)$ unless $h(f)h(g) = 0$.

(d) $R[I]$ has no zero divisors if and only if R has no zero divisors.

(e) $\deg fg \leq \deg f + \deg g$ with equality if R has no zero divisors.

Proof. (a),(b) and (c) are readily verified.

(d) If R has zero divisors, then as R is embedded in $R[I]$, $R[I]$ has zero divisors.

Suppose next that R has no zero divisors. Let $f, g \in R[I]^\#$. We need to show that $fg \neq 0$. By (c) we may assume that f and g are homogeneous.

Consider first the case that $|I| = 1$. Then $f = ax^n$, $g = bx^m$ and $fg = (ab)x^{n+m}$. Here $a, b \in R^\#$ and so $ab \neq 0$. Thus also $fg \neq 0$. If I is finite, $R[I] = R[I \setminus \{i\}][i]$ and so by induction $R[I]$ has no zero divisors.

For the general case just observe that $f, g \in R[J]$ for some finite subset J of I.

(e) If R has no zero divisors, (d) implies $h(f)h(g) \neq 0$. Thus by (b) and (c),

$$\deg f = \deg h(fg) = \deg h(f)h(g) = \deg h(f) + \deg h(g) = \deg f + \deg g.$$

□

Lemma 3.7.3. *Let R be a ring, P an ideal in R and I a set.*

(a) Let $P[I] = \{f \in R[I] \mid f_\alpha \in P \text{ for all } \alpha \in \mathbb{I}\}$. *Then* $P[I]$ *is an ideal in* $R[I]$ *and*

$$R[I]/P[I] \cong (R/P)[I]$$

(b) If R *has an identity,* $P[I] = P \cdot R[I]$ *is the ideal in* $R[I]$ *generated by* P.

Proof. (a) Define $\phi : R[I] \to (R/P)[I], \sum_{\alpha\in\mathbb{I}} f_\alpha x^\alpha \to \sum_{\alpha\in\mathbb{I}}(f_\alpha + P)x^\alpha$. By 3.6.1 ϕ is a ring homomorphism. Clearly ϕ is onto and $\ker\phi = P[I]$ so (a) holds.

(b) Let $p \in P$ then $px^\alpha \in P \cdot R[I]$. Thus $P[I] \leq P \cdot R[I]$. The other inclusion is obvious. □

Corollary 3.7.4. *Let* R *be a commutative ring with identity,* I *a set and* $p \in R$. *Then* p *is a prime in* R *if and only if* p *is a prime in* $R[I]$.

Proof. R is a prime if and only if R/pR is an integral domain. So by 3.7.2d if and only if $(R/pR)[I]$ is an integral domain. So by 3.7.3 if and only if $R[I]/pR[I]$ is a prime ideal and so if and only if p is a prime in $R[I]$. □

Theorem 3.7.5 (Long Divison)**.** *Let* R *be a ring and* $f, g \in R[x]$. *Suppose that the leading coefficient of* g *is a unit in* R. *Then there exist uniquely determined* $q, r \in R$ *with*

$$f = qg + r \text{ and } \deg r < \deg g$$

Proof. Let $h(f) = ax^n$ and $h(g) = bx^m$. If $n < m$, we conclude that $q = 0$ and $r = f$ is the unique solution.

So suppose that $m \leq n$. Then any solution necessarily has $h(f) = h(q)h(g)$ and so $s(q) = ab^{-1}x^{n-m}$. Now $f = qg - r$ if and only if

$$f - ab^{-1}x^{n-m}g = (q - ab^{-1}x^{n-m})g + r$$

So uniqueness and existence follows by induction on $\deg f$. □

Let R be a ring and $f \in R[x]$. Define the function

$$f^r : R \to R, c \to \sum_{\alpha\in\mathbb{N}} f_\alpha c^\alpha$$

The function f^r is called the *right evaluation* of f. Note here that as R is not necessarily commutative , $f_\alpha c^\alpha$ might differ from $c^\alpha f^\alpha$. If R is commutative $f^r = f^{\mathrm{id}}$.

The map $f \to f^r$ is an additive homomorphism but not necessarily a multiplicative homomorphism. That is we might have $(fg)^r(c) \neq f^r(c)g^r(c)$. Indeed let $f = rx$ and $g = sx$. Then $fg = (rs)x^2$, $(fg)^r(c) = rsc^2$ and $f^r(c)g^r(c) = rcsc$.

Lemma 3.7.6. *Let* R *be a ring,* $f, g \in R[x]$ *and* $c \in R$. *If* $g^r(c)c = cg^r(c)$ *then*

$$(fg)^r(c) = f^r(c)g^r(c).$$

Proof. As $f \to f^r$ is a additive homomorphism we may assume that $f = rx^m$ for some $r \in R$, $m \in \mathbb{N}$. Thus

$$fg = \sum_{\alpha \in \mathbb{N}} r g_\alpha x^{\alpha+m}$$

and so

$$(fg)^r(c) = \sum_{\alpha \in \mathbb{N}} r g_\alpha c^{\alpha+m} =$$

$$= r(\sum_{\alpha \in \mathbb{N}} g_\alpha c^\alpha) c^m = r g^r(c) c^m = r c^m g^r(c) = f^r(c) g^r(c)$$

□

Corollary 3.7.7. *Let R be a ring with identity, $c \in R$ and $f \in R[x]$.*

(a) Then there exists a unique $q \in R[x]$ with

$$f = q(x - c) + f^r(c).$$

(b) $f^r(c) = 0$ if and only if $f = q(x - c)$ for some $q \in R[x]$.

Proof. (a) By3.7.5 $f = q \cdot (x - c) + r$ with $\deg r < \deg(x - c) = 1$. Thus $r \in R$. By 3.7.6

$$f^r(c) = q^r(c)(c - c) + r = r$$

Hence $r = f^r(c)$. The uniqueness follows from 3.7.5

(a) follows from (b). □

Corollary 3.7.8. *Let R be an commutative ring with identity and $c \in R$.*

(a) $R[x]/(x - c) \cong R$.

(b) $x - c$ is a prime if and only R is an integral domain.

Proof. Proof. Consider the ring homomorphism $\mathrm{id}_c : R[x] \to R, f \to f(c)$ (see 3.6.1 Clearly id_c is onto. By 3.7.7b $\ker \mathrm{id}_c = (x - c)$ so (b) follows from the Isomorphism Theorem for rings.

(b) Note that $x - c$ is a prime if and only if $R[x]/(x - c)$ has non-zero divisors. Thus (a) follows from (b). □

Corollary 3.7.9. *Let $\mathbb{F}$ be a field. Then $\mathbb{F}[x]$ is an Euclidean domain. In particular, $\mathbb{F}[x]$ is a PID and a UFD. The units in $\mathbb{F}[x]$ are precisely the nonzero elements in $\mathbb{F}$.*

Proof. Just note that by 3.7.5 $\mathbb{K}[x]$ is a Euclidean domain. □

Let R be a subring of the commutative ring S. Write $R \to S$ for the inclusion map from R to S. Let I be a set, $f \in R[I]$ and $c \in S^I$. We say that c is a *root* of f if

$$f^{R\to S}(c) = 0.$$

Let R be any ring, $f \in R[x]$ and $c \in R$. We say that c is a root of f if $f^r(c) = 0$. Note that for R commutative this agrees with previous definition of a root for f in R.

Theorem 3.7.10. *Let D be an integral domain contained in the integral domain E. Let $0 \neq f \in D[x]$. Let $m \in \mathbb{N}$ be maximal so that there exists $c_1, \dots c_m \in E$ with*

$$\prod_{i=1}^{m} x - c_i \quad | \quad f$$

in $E[x]$. Let c be any root of f in E. Then $c = c_i$ for some i. In particular, f has at most $\deg f$ *distinct roots in E.*

Proof. Let $f = g\prod_{i=1}^{m} x - c_i$ with $g \in E[x]$. By maximality of m, $x - c \nmid g$. By 3.7.8 $x - c$ is a prime in $E[x]$ and so

$$x - c \mid \prod_{i=1}^{m} x - c_i$$

By 3.3.12, $x - c \sim x - c_i$ for some i. Thus $x - c = x - c_i$ and $c = c_i$. □

We remark that the previus theorem can be false for non-commuative divison rings. For example the polynomial $x^2 + 1 = 0$ has at infinitely many roots in the division ring $\mathbb{H}$ of quaternions, namely any $ai + bj + ck$ with $a^2 + b^2 + c^2 = 1$.

Let R be a ring, $f \in R[x]$ and c b a root of f in D. Then by 3.7.10 We can write f has $f = g(x - c)^m$ with $m \in \mathbb{Z}^+$, $g \in R[x]$ and so that c is not a root of g. m is called the *multiplicity* of the root g. If $m \geq 2$ we say that c is a *multiple root*

As a tool to detect multiple roots we introduce the formal *derivative* f' of a polynomial $f \in R[x]$.

$$f' := \sum_{\alpha\in\mathbb{Z}^+} nf_\alpha x^{\alpha-1}$$

Put $f^{[0]} = f$ and inductively, $f^{[k+1]} = (f^{[k]})'$ for all $k \in \mathbb{N}$.

Lemma 3.7.11. *Let R be a ring, $f, g \in R[x]$ and $c \in R$. Then*

(a) $(cf)' = cf'$

(b) $(f + g)' = f' + g'$.

(c) $(fg)' = f'g + fg'$.

(d) If $ff' = f'f$, $(f^n)' = nf^{n-1}f'$.

Proof. (a) and (b) are obvious.

(c) By (b) we may assume that $f = rx^m$ and $g = sx^n$ are monomials. We compute

$$(fg)' = (rsx^{n+m})' = (n+m)rsx^{n+m-1}$$
$$f'g + fg' = mrx^{m-1}sx^n + rx^m nsx^{n-1} = (n+m)rsx^{m+n-1}$$

Thus (c) holds.

(d) follows from (c) and induction on n. □

Lemma 3.7.12. *Let R be a ring with identity, $f \in R[x]$ and $c \in R$ a root of f.*

(a) Suppose that $f = g(x-c)^n$ for some $n \in \mathbb{N}$ and $g \in R[x]$. Then

$$f^{[n]}(c) = n!g(c).$$

(b) c is a multiple root of f if and only if $f'(c) = 0$.

(c) Suppose that $(\deg f)!$ is neither zero nor a zero divisor in R. Then the multiplicity of the root c is smallest number $m \in \mathbb{N}$ with $f^{[m]}(c) \neq 0$.

Proof. (a) We will show that for all $0 \leq i \leq n$, there exists $h_i \in R[x]$ with

$$f^{[i]} = \frac{n!}{(n-i)!}g(x-c)^{n-i} + h_i(x-c)^{n-i+1}$$

For $i = 0$ this is true with $h_0 = 0$. So suppose its true for i. Then using 3.7.11

$$f^{[i+1]} = (f^{[i]})' = \frac{n!}{(n-i)!}(g'(x-c)^{n-i} + g(n-i)(x-c)^{n-i-1}) + h_i'(x-c)^{n-i+1} + h_i(n-i+1)x^{n-i}$$

This is of the form $\frac{n!}{(n-i-1)!}g(x-c)^{n-i-1}$ plus a left multiple of $(x-c)^{n-i}$. So the statements holds for $i+1$.

For $i = n$ we conclude $f^{[n]} = n!g + h_n(x-a)$ Thus (a) holds.

(b) Since c is a root, $f = g(x-a)$ for some $g \in R[x]$. So by (a) applied to $n = 1$, $f'(c) = g(c)$. Thus (b) holds.

(c) Let m the multiplicity of c has a root of f. So $f = g(x-c)^m$ for some $g \in R[x]$ with $g(c) \neq 0$. Let $n < m$. Then $f = (g(x-c)^{m-n})(x-c)^n$ and (a) implies $f^{[n]}(c) = 0$. Suppose that $f^{[m]}(c) = 0$. Then by (a), $m!g(c) = 0$. As $m \leq \deg f$ we get $(\deg f)!\, g(c) = 0$. Thus by assumption $g(c) = 0$, a contradiction. This $f^{[m]}(c) \neq 0$ and (c) holds. □

Consider the polynomial x^p in $\mathbb{Z}/p\mathbb{Z}[x]$. Then $(x^p)' = px^{p-1} = 0$. This shows that the condition on $(\deg f)!$ in part (c) of the previous theorem is necessary.

Let D be an UFD, I a set and $f \in D[I]$. We say that f is *primitive* if 1 is a greatest common divisor of the coefficents of f.

Lemma 3.7.13. *Let D be a UFD, $\mathbb{F}$ its field of fractions and I a set. Let $f \in \mathbb{F}[I]$. Then there exists $a_f, b_f \in D$ and $f^* \in \mathbb{D}[I]$ so that*

(a) f^* is primitive in $D[I]$.

(b) a_f and b_f are relatively prime.

(c) $f = \frac{a_f}{b_f} f^*$.

Moreover a_f, b_f and f^* are unique up to associates in D.

Proof. We will first show the existence. Let $f = \sum_{\alpha \in \mathbb{I}} f_\alpha x^\alpha$ with $f_\alpha \in \mathbb{F}$. Then $f_\alpha = \frac{r_\alpha}{s_\alpha}$ with $r_\alpha, s_\alpha \in D$. Here we choose $s_\alpha = 1$ if $f_\alpha = 0$. Let $s = \prod_{\alpha \in \mathbb{I}} s_\alpha$. Then $sf \in D[I]$. Let $r = \gcd_{\alpha \in \mathbb{I}} sf_\alpha$ and $f^* = r^{-1} sf$. Then $f^* \in D[I]$, f^* is primitive and $f = \frac{r}{s} f^*$. Let e the a greates common divisor of r and s and put $a_f = \frac{r}{e}$ and $b_f = \frac{s}{e}$. Then (a),(b) and (c) hold.

To show uniqueness suppose that $f = \frac{a}{b} \tilde{f}$ with $a, b \in D$ relative prime and $\tilde{f} \in D[I]$ primitive. Then

$$ba_f^* = b_f a \tilde{f}$$

Taking the greatest common divisor of the coefficents on each side of this equation we see that ba_f and $b_f a$ are associate in D. In particular, a divides ba_f and as b is realtively prime to a, a divides a_f. By symmetry a_f divides a and so $a = ua_f$ for some unit u in D. Similarly $b = vb_f$ for some unit $v \in D$. Thus $vb_f a_f f^* = ub_f a_f \tilde{f}$. As D is an integral domain we conclude $\tilde{f} = u^{-1} v f^*$. □

Let f be as in the previuos theorem. The fraction $c_f = \frac{a_f}{b_f}$ is called the content of f. Note that $c_f \in \mathbb{F}$ and $f = c_f f^*$.

Lemma 3.7.14. *Let D be a UFD, $\mathbb{F}$ its field of fraction, I a set and $f, g \in \mathbb{F}[I]^\#$.*

(a) $c_{fg} = u c_f c_g$ for some unit $u \in D$.

(b) $(fg)^* = u^{-1} f^* g^*$

(c) The product of primitive polynomials is primitive.

(d) If $f \mid g$ in $\mathbb{F}[I]$, then $f^* \mid g^*$ in $D[I]$.

(e) Suppose f is primitive. Then f is irreducible in $D[I]$ if and only if its irreducible in $\mathbb{F}[I]$

(f) Suppose f is primitive.Then f is a prime in $D[I]$ if and only if it is a prime in $\mathbb{F}[I]$.

Proof. Note that $fg = c_f c_g f^* g^*$. So (a), (b) and (c) will follow once the show that the product of two primitive polynomials is primitive. Suppose not. Then there exist primitive $f, g \in D[I]$ and a prime p in D dividing all the coefficients of fg. But then $p \mid fg$ in $D[I]$.

By 3.7.4 p is prime in $D[I]$ and so p divides f or g in $D[I]$. A contradiction as f and g are primitive.

(d) Suppose that $f \mid g$. Then $g = fh$ for some $h \in \mathbb{F}[I]$. By (b) $g^* = f^*h^*$ and so (d) holds.

(e) Suppose that f is irreducible in $\mathbb{F}[I]$ and $f = gh$ with $g, h \in D[x]$ Then by (a) both g and h are primitive. On the other hand since f is irreducible in $\mathbb{F}[I]$, one of g or h is a unit in $F[I]$ and so in $\mathbb{F}$. It follows that one of g and h is a unit in D. So f is also irreducible in $D[I]$.

Suppose that f is irreducible in $D[I]$ and $f = gh$ for some $g, h \in \mathbb{F}[x]$. Then $f = f^* \sim g^*h^*$ and as f is irreducible in $D[I]$, one of g^*, h^* is a unit in D. But then one of g and h is in $\mathbb{F}$ and so a unit in $\mathbb{F}[I]$.

(f) Suppose that f is prime in $D[I]$ and that $f \mid gh$ in $\mathbb{F}[I]$. By (d) $f = f^* \mid g^*h^*$ and as f is a prime in $D[I]$ we may assume $f \mid g^*$. As g^* divides g in $F[I]$ f does too. So f is a prime in $F[I]$.

Suppose that f is a prime in $F[I]$ and $f \mid gh$ in $D[I]$ for some $g, h \in D[I]$. Then as f is a prime in $F[I]$ we may assume that $f \mid g$ in $F[I]$. But (d) $f = f^* \mid g^*$ in $D[I]$. As g^* divides g in $D[I]$, f does too . So f is a prime in $D[I]$. □

Theorem 3.7.15. *Let D be a UFD and I a set, then $D[I]$ is a UFD.*

Proof. Let f be in $D[I]$. We need to show that f is the product of primes. Now $f \in D[J]$ for some finite f and by 3.7.4 a prime factorization in $D[J]$ is a prime factorization in $D[I]$. So we may assume that J is finite and then by induction that $|I| = 1$.

Note that $f = c_f f^*$ with $f^* \in D[x]$ primitive and $c_f \in D$. As D is a UFD, c_f is a product of primes in D and by3.7.4 also a prodcut of primes in $D[x]$. So we may assume that f is primitive. Suppose that $f = gh$ with $g, h \in D[x]$ with neither g nor h a unit. As f is primitive, g and h both have positive degree smaller than f. So by induction on $\deg f$ both g and h are a product of primes. So we may assume that f is irreducible. Let $\mathbb{F} = \mathbb{F}_D$. By 3.7.13 f is irreducible in $\mathbb{F}[x]$. As $\mathbb{F}[x]$ is Euclidean, f is a prime in $\mathbb{F}[x]$. Hence by 3.7.13 f is a prime in $D[x]$. □

Chapter 4

Modules

4.1 Modules and Homomorphism

In this section we introduce modules over a ring. It corresponds to the concept of group action in the theory of groups.

Definition 4.1.1. *Let $(R, +, \cdot)$ be a ring and $(M, +)$ an abelian group. A* ring action *of R on M is function*

$$\diamond : R \times M \to M, \quad (r, m) \to rm$$

such that for all $r, s \in R$ and $a, b \in M$:

(a) $r \diamond (a + b) = ra + rb$.

(b) $\diamond(r + s)a = ra + sa$.

(c) $r \diamond (s \diamond a) = (r \cdot s)a$.

*In this case $(M, +, \diamond)$ is called an R-*module.

Abusing notation we will call M an R-module and write ra for $r \diamond a$.

Lemma 4.1.2. *Let R be a ring and M an abelian group*

(a) Let $\diamond : R \times M \to M$ an ring action of R on M. For $r \in R$ define

$$r^{\diamond} : M \to M, m \to r^{\diamond}$$

Then $r^{\diamond} \in \mathrm{End}(M)$ and the map

$$\Phi^{\diamond} G \to \mathrm{End}(S)$$

is ring homomorphism.

$\Phi^{\diamond}$ is called the ring homomorphism *associated to $\diamond$,*

(b) Let $\Phi : R \to \mathrm{End}_R(M)$ *be a homomorphism. Define*

$$\diamond^{\Phi} : R \times M \to M, (r, m) \to \Phi(R)(m)$$

then $\diamond_{\Phi}$ *is an ring action of* R *on* M*.*

$\diamond^{P}hi$ *is called the* action *associated to* Φ

(c) $\Phi^{\diamond^{\Phi}} = \Phi$ *and* $\diamond^{\Phi^{\diamond}} = \diamond$*.*

Proof. This is very similar to 2.10.3 and we leave the details to the reader. □

Example 4.1.3. Let R be a ring and A an abelian group.

1. A is a $\mathbb{Z}$-module via $n \diamond a = na$ for all $n \in \mathbb{Z}$ and $a \in A$.
2. A is an $\mathrm{End}(A)$-module via $\phi m = \phi(m)$ for all $\phi \in \Phi$, $m \in M$.
3. A is an R-module via, $ra = 0_R$ for all $r \in R$, $a \in A$.
4. R is an R-module via left multiplication.
5. Let $(M_i, i \in I)$ a family of R-modules. Then $\bigtimes_{i \in I} M_i$ and $\oplus_{i \in I} M_i$ are R-modules via

$$r \diamond (m_i)_{i \in I} = (rm_i)_{i \in I}$$

6. Let R be ring, M an R-module, G a group and Ω an G-set. Let $M^{\Omega} = \bigtimes_{\omega \in \Omega} M = \{f : \Omega \to R\}$. Then G acts an M^{Ω} via $(gf)(\omega) = f(g^{-1}\omega)$ and M is an R-module via $(rf)(\omega) = rf(\omega)$. Finally M^{Ω} is a $R[G]$-module via

$$(\sum_{g \in G} r_g g)f = \sum_{g \in G} r_g gf$$

Definition 4.1.4. *Let* V *and* W *be* R*-modules and* $f : V \to W$ *be a function. Then* f *is called* R*-linear or an* R*-homomorphism if*

$$f(a + c) = f(a) + f(c) \text{ and } f(ra) = rf(a)$$

for all $a, c \in R, r \in R$*.*

Definition 4.1.5. *Let* R *be a ring with identity and* M *an* R*-modules.*

(a) M *is a* unitary R*-module provide that*

$$1_R m = m$$

for all $m \in M$*.*

(b) If R *is a division ring and* M *is unitary then* M *is called a vector space over* R*.*

We often will say that $f : V \to W$ is *R-linear* instead of $f : V \to W$ is a R-modules homomorphism. Terms like R-module monomorphism, R-module isomorphism, $\ker f$ and so on are defined in the usual way. If V and W are R-modules, $\mathrm{Hom}_R(V, W)$ denotes the set of R-linear maps from V to W. Since sums of R-linear maps are R-linear, $\mathrm{Hom}_R(V, W)$ is an abelian group. $\mathrm{End}_R(V)$ denotes set of R-linear endomorphsims of V. Since compositions of R-linear maps are R-linear, $\mathrm{End}_R(V)$ is a ring. Note that V is also a module for $\mathrm{End}_R(V)$ via $\phi v = \phi(v)$.

Definition 4.1.6. *Let R be a ring and $(V, +\diamond)$ an R-module. An R-submodule of R is a R-module $(W, \triangle, \square)$ such that*

(i) $W \subseteq V$.

(ii) $a \triangle b = a + b$ *for all* $a, b \in W$.

(iii) $r \square a = r \diamond a$ *for all* $r \in R$, $a \in W$.

Lemma 4.1.7. *Let R be a ring, V an R-module and W an R-submodule of W. Then*

$$\diamond_{V/W} : R \times V/W \to V/W, (r, v + W) \to rv + W$$

is a well-defined ring action of R on $(V/W, +_{V/W})$. Moreover the map

$$\pi : V \to V/W, v \to v + W$$

is an onto R-homomorphism with $\ker \pi = W$.

Proof. Let $v, v' \in V$ with $v + W = v' + W$. Then $v' = v + w$ for some $w \in W$. Let $r \in R$. Since W is an R-submodule, $rw \in W$ thus $rv + W = rv + rw + W = r(v + w) + W$ and so $\diamond_{V/W}$ is well-defined. Straight forward calculations show that $\diamond_{V/W}$ is a ring action.

By 2.5.7(f), π is a well-defined onto homomorphism of abelian groups with $\ker \pi = W$. We have

$$\pi(rv) = rv + W = r(v + W) = r\pi(v)$$

and so π is R-linear. □

Theorem 4.1.8 (Isomorphism Theorem for Modules)**.** *Let R be a ring and $f : V \to W$ an R-linear map. Then$f(V)$ is an R-submodule of W and*

$$\overline{f} : V/ + \ker f \to f(W), v + \ker f \to f(v)$$

is a well-defined R-linear isomorphism.

Proof. By the isomorphism theorem for groups 2.5.8, this is a well defined isomorphism of abelian groups. We just need to check that it is R-linear. So let r and $v \in V$. Then

$$\overline{f}(r(v + \ker f)) = \overline{f}(rv + W) = f(rv) = rf(v) = r\overline{f}(v + \ker f).$$

□

Definition 4.1.9. *Let R be a ring, M an R-module, $S \subseteq R$ and $X \subset M$.*

(a) $\langle X \rangle$ *is the subgroup of* $(M, +)$ *generated* X.

(b) $SX = \langle sx \mid s \in S, x \in X \rangle$

(c) $\mathrm{Ann}_S(X) = \{s \in S \mid sx = 0_M \text{ for all } x \in X\}$. $\mathrm{Ann}_S(X)$ *is called the* annihilator *of* X *in* S

(d) $\mathrm{Ann}_X(S) = \{x \in X \mid sx = 0_M \text{ for alls} \in S\}$. $\mathrm{Ann}_X(S)$ *is called the* annihilator *of* X *in* S.

(e) $\langle X \rangle_R := \{W \mid W \text{ is an } R \text{ submodule of } M, X \subseteq M\}$. $\langle X \rangle_R$ *is called* R*-submodule of* M generated *by* X.

(f) M *is called* finitely generated *if* $M = \langle I \rangle_R$ *for some finite subset* I *of* R.

Lemma 4.1.10. *Let R be a ring, M an R-module, $S, T \subseteq R$ and $X \subseteq M$.*

(a) $SX = \langle S \rangle \langle X \rangle$.

(b) $S(TX) = (ST)X$.

(c) If S is left ideal in R, then SX is an R-submodule of M.

(d) $\langle X \rangle + RX = \langle X \rangle_R$.

(e) If M is unitary, then $RX = \langle X \rangle_R$

(f) If S is a subgroup of $(R, +)$ and $(x_i, i \in I)$ is a family of elements with $X = \langle x_i \mid i \in I \rangle$ then then $SX = \{\sum_{i \in I} s_i x_i \mid s_i \in S, s_i = 0_R \text{ for almost all } i\}$

Proof. (a) Clearly $SX \subseteq \langle S \rangle \langle X \rangle$. Since ST is a subgroup of $(M, +)$ it is easy to verify that $U = \{r \in R \mid rX \subseteq SX\}$ is a subgroup of $(R, +)$. Since $S \subseteq U$, we get $\langle S \rangle \subseteq U$ and so $\langle S \rangle X \subseteq SX$. Similarly $V = \{m \in M \mid \langle S \rangle m \subseteq ST\}$ is a subgroup of M, $\langle X \rangle \subseteq V$ and $\langle S \rangle \langle X \rangle \subseteq ST$.

(b) $S(TX) = \langle s(tx) \mid s \in S, t \in T, x \in X \rangle = \langle (st)x \mid s \in S, t \in T, x \in X \rangle = (ST)X$.

(c) Since S is a left ideal, $RS \subseteq S$. Thus using (b)

$$R(SX) = (RS)X \subseteq SX$$

By definition SX is a subgroup of $(M, +)$ and so SX is an R-submodule of M.

(d) By (a), $R\langle X \rangle = RX$ and by (c) $R(RX) \subseteq RX$. Then $R(\langle X \rangle + RX) \subseteq RX$ and $\langle X \rangle + RX$ is an R-submodule of M containing, X. Thus $\langle X \rangle_R \subseteq \langle X \rangle + RX$.

Clearly $\langle X \rangle + RX \subseteq W$ for each R-submodule of W of M with $X \subseteq W$ and so $\langle X \rangle + RX \subseteq \langle X \rangle_R$.

(e) If M is unitary, then $X \subseteq 1_R X \subseteq RX$ and so $\langle X \rangle + RX = RX$. So (e) follows from (d).

(f) Put $Y = \{\sum_{i\in I} s_i x_i \mid s_i \in S, s_i = 0_R \text{ for almost all } i\}$.

It is easy to verify that Y is a subgroup of $(M,+)$ and so $Y = S\{x_i \mid i \in I\}$. Thus by (a), $Y = SX$. □

Lemma 4.1.11. *Let R be ring, M a R-module, $S \subseteq R$ and $X \subseteq M$. Then*

(a) $\mathrm{Ann}_R(X)$ *is a left ideal in R.*

(b) If X is a R-submodule of M, then $\mathrm{Ann}_R(X)$ *is an ideal in R*

(c) Let I be a right ideal in R. Then $\mathrm{Ann}_M(I)$ *is R-submodule in M.*

(d) $S \subseteq \mathrm{Ann}_R(X)$ *if and only if* $X \subseteq \mathrm{Ann}_M(X)$.

(e) Suppose that one of the following holds:

1. *R is commutative.*
2. *All left ideals in R are also right ideals.*
3. $\mathrm{Ann}_R(X)$ *is a right ideal.*

Then $\mathrm{Ann}_R(X) = \mathrm{Ann}_R(\langle X\rangle_R)$.

(f) Let $m \in M$. Then the map

$$R/\mathrm{Ann}_R(m) \to Rm, r + \mathrm{Ann}_R(m) \to rm$$

is a well defined R-isomorphism. (Note here that by (a), $\mathrm{Ann}_R(x)$ *is an R-submodule. So $R/Ann_R(m)$ is an R-module. Also by 4.1.10(c), Rm is an R-module.)*

Proof. (a) and (b) Let $r, s \in \mathrm{Ann}_R(X)$, $t \in R$ and $x \in X$. Then

$$\begin{gathered} 0_R x = 0_R \\ (r+s)x = rx + sx = 0_M + 0_M = 0_M \\ (-r)x = -(rx) = -0_M \\ (tr)x = t(rx) = t0_M = 0_M \end{gathered}$$

Hence $0_M, r+s, -r$ and tr all are in $\mathrm{Ann}_R(X)$. Thus by 3.2.2 $\mathrm{Ann}_R(X)$ is a left ideal in R.

If X is an R-submodule of M, then $tx \in X$ and so $(rt)x = r(tx) = r0_M = 0_M$. Thus $rt \in \mathrm{Ann}_R(X)$ and and $\mathrm{Ann}_R(X)$ is an ideal.

(c) Let $x, y \in Ann_M(I)$, $r \in R$ and $i \in I$. Since I is a right ideal, $ir \in I$.

$$\begin{gathered} i0_M = 0_M \\ i(x+y) = ix + iy = 0_M + 0_M = 0_M i(-x) = -(ix) = -0_M = 0_M \\ i(rx) = (ir)x = 0_M \end{gathered}$$

Hence 0_M, $x+y$, $-x$ and rx all are contained in $\mathrm{Ann}_M(I)$. Thus $\mathrm{Ann}_M(I)$ is an R-submodule of M.

(d) Both statements are equivalent to t $sx = 0_M$ for all $s \in S, x \in X$.

(e) Note that (e:1) implies (e:2) and by (a), (e:2) implies (e:3) So in any case $\mathrm{Ann}_R(X)$ is a right ideal in R. Hence by (c)

$$W := \mathrm{Ann}_M(\mathrm{Ann}_R(X))$$

is an R-submodule. By (d), $X \subseteq W$ and so since W is an R-submodule, $\langle X\rangle_R \leq W$. Thus by (d), So $\mathrm{Ann}_R(X) \leq \mathrm{Ann}_R(\langle X\rangle_R$..

Since $X \subseteq \langle X\rangle_R$, $\mathrm{Ann}_R(\langle X\rangle_R) \subseteq \mathrm{Ann}_R(X)$. Thus (e) holds.

(f) Consider the map

$$f : R \to M, \quad r \to rm.$$

Let $r, s \in R$. Then $f(r+s) = (r+s)m = rm + sm = f(r) + f(s)$. Also for $r, s \in R$

$$f(rs) = (rs)m = r(sm) = rf(s)$$

So f is R-linear. Since $\mathrm{Ann}_R(m) = \ker f$, (f) follows from the Isomorphism Theorem 4.1.8 □

Example 4.1.12. Let $\mathbb{K}$ be a field. Let $R = \mathrm{M}_{nn}(\mathbb{K})$ be the ring of $n \times n$ matrices with coefficients in $\mathbb{K}$. Then $\mathbb{K}^n$ is a module for R via $(m_{ij})_{ij}(k_j)_j = (\sum_{j=1}^n m_{ij}k_j)_i$. Let $e_1 = (\delta_{1i})_i \in \mathbb{K}^n$, where $\delta_{1i} = 1_{\mathbb{K}}$ if $i = 1$ and $0_{\mathbb{K}}$ if $i \neq 1$. If $A \in \mathrm{M}_{nn}(\mathbb{K})$, then Ae_1 is the first columns of A. Hence $\mathrm{Ann}_R(e_1)$ consists of all matrices whose first column is zero and $(e_1) = Re_1 = \mathbb{K}^n$. Thus $\mathrm{Ann}_R((e_1)) = 0$ and $\mathrm{Ann}_R(e_1) \neq \mathrm{Ann}_R((e_1))$. So the conclusion of 4.1.11(e) does not hold in general.

Moreover, by 4.1.11(f), $\mathbb{K}^n \cong R/\mathrm{Ann}_R(e_1)$ as an R-module.

4.2 Free modules and torsion modules

In this section R is a ring with identity and all modules are assumed to be unitary. $\delta_{ij} = 1_R$ if $i = j$ and $\delta_{ij} = 0_R$ if $i \neq j$.

Definition 4.2.1. *Let V be an R-module and $(v_i, i \in I)$ family if elements in V*

(a) V is called free with respect to $(v_i, i \in I)$ if for all unitray R-modules W and all family of elements $(w_i, i \in I)$ there exists a unique R-linear map $f : V \to W$ with $f(v_i) = w_i$ for all $i \in I$.

(b) $(v_i, i \in I)$ is called R-linearly independent, if for all $r_i \in R_i$, almost all zero, we have

$$\sum_{i\in I} r_i v_i = 0_V \Longleftrightarrow r_i = 0_R$$

(c) $(v_i, i \in I)$ is called a R-spanning set of $V = \langle v_i \mid i \in I\rangle_R$.

(d) $(v_i, i \in I)$ *is called a* R*-basis if is a* R*-linearly independent* R*-spanning set.*

(e) Let c *be a cardinality. The we say that* V *is free of rank* c *if* V *is free with respect to some* $(w_j, j \in J)$ *with* $|J| = c$*.*

Lemma 4.2.2. *Let* V *be an* R*-module and* $(v_i, i \in I)$ *a family of elements in* V*. Then the following statements are equivalent.*

(a) $(v_i, i \in I)$ *is a basis for* V*.*

(b) The map $f : \bigoplus_{i\in I} R, (r_i)_{i\in I} \to \sum_{i\in I} r_i v_i$ *is an* R*-isomorphism.*

(c) For each $v \in V$ *there exists uniquely determined* $r_i \in R, i \in I$ *almost zero with* $v = \sum_{i\in I} r_i v_i$*.*

(d) V *is free with respect to* $(v_i)_{i\in I}$*.*

Proof. (a) $\implies$ (b): . Since $(v_i, i \in I$ is a linearly independent, $\ker f = \{(0_R)_i\}$ and so f is 1-1. So $(v_i, i \in I)$ is a spanning set,

$$V = \{\sum_{i\in I} r_i v_i \mid r_i \in R, \text{ almost all } 0_R\} = \operatorname{Im} f =$$

and so f is onto. It is easy to check that f is R-linear and so (b) holds.

(b) $\implies$ (c): (c) is true since f is a bijection.

(c) $\implies$ (d): Let W be a unitary R-module and $(w_i)_{i\in I}$ a family of R. If $f : V \to W$ is R linear with $f(v_i) = w_i$, then

$$(*) \qquad f(\sum_{i\in I} r_i v_i = \sum_{i\in I} r_i w_i$$

By (c) (*) uniquely defines a defines a function from V to W. It is easy to check that it's R-linear. Since V and W are unitary, $f(v_i) = f(1_R v_i) = 1_R w_i = w_i$. and so (d) holds.

(d) $\implies$ (a): Let $r_i \in R$, $i \in I$, almost all zero, with $\sum_{i\in I} r_i v_i = 0_V$. Fix $j \in I$. Since V is free with respect to $(v_i, i \in I)$ there exist an R-linear map $V \to R$ with $f(v_i) = \delta_{ij}$. Then

$$0_R = f(0_V) = f(\sum_{i\in I} r_i v_i) = \sum_{i\in I} r_i f(v_i) = \sum_{i\in I} r_i \delta_{ij} = r_j$$

So $(v_i, i \in I)$ is linearly independent. Let $W = \langle v_i \mid i \in I\rangle_R$. Consider R-linear maps

$$\pi : V \to V/W, v \to v + W \tag{4.1}$$

$$\alpha : V \to V/W, v \to 0_{V/W} \tag{4.2}$$

Then $\pi(v_i) = v_i W = W = 0_{V/W} = \alpha(v_i)$ for all $i \in I$. The uniqueness assertion in the definition of a free module shows that $\pi = \alpha$. Since $\operatorname{Im}\alpha = \{0_{V/W}\}$ and $\operatorname{Im}\pi = V/W$ we conclude $V/W = \{0_{V/W}$ and so $V = W$. Hence $V/W = \{0_V/W\}$ and $V = W$. Thus $(v_i, i \in I)$ is a spanning set. □

We will now investigate when all submodules of free R-modules are free. First an example.

Example 4.2.3. Let $R = \mathbb{Z}_n$ with $n \in \mathbb{Z}^+$, n not a prime. Let $V = R$, viewed as an $\mathbb{Z}_n$-module by left multiplication. Let m is a proper divisor of m. Then $W = m\mathbb{Z}/n\mathbb{Z}$ is an submodule of R. W which is not free. A obvious necessary condition for all submodules of all free modules for a ring R to be free is that all submodules of R itself are free. The next theorem shows that this condition is also sufficient.

Theorem 4.2.4. *(a) Suppose that all left ideals in the ring R are free as R-modules. Then every submodule of a free module is free.*

(b) Suppose R is a PID and V is a free R-module of rank r. Then every R-submodule of V is free of rank less or equal to r.

Proof. (a) Let M be a free module with basis B and A a R-submodule in M. According to the well ordering principal (A.10) we can choose a well ordering $\leq$ on B. For $b \in B$ define

$$M_b^* = \sum_{e \in B, e<b} Re \text{ and } M_b = \sum_{e \in B | e \leq b} Re$$

Note that $M_b = B_b^* \oplus Rb$. Put $A_b = M_b \cap A$ and $A_b^* = M_b^* \cap A$. Then

$$A_b/A_b^* = A_b/A_b \cap M_b^* \cong A_b + M_b^*/M_b^* \leq M_b/M_b^* \cong Rb \cong R$$

.

By assumption every submodule of R is free and so A_b/A_b^* is free. Let $E_b \subset A_b$ such that $(e + A_b^*, e \in E_b)$ is a basis for A_b/A_b^*. Let $E = \bigcup_{b \in B} E_b$. We claim that E is a basis for A.

Let $m \in M$. Then $m = \sum_{b \in b} r_b b$ with $r_b =\in R$ and almost all $m_b = 0$. So we can choose $b_m \in B$ maximal with respect $r_{b_m} \neq 0$. Clearly for $e \in E_b$, $b_e = b$. In general, b_m is minimal in B with $m \in M_{b_m}$.

Now suppose that $\sum_{e \in E} r_e e = 0$ so that almost all, but not all $r_e = 0$. Let b be maximal with $b = b_e$ and $r_e \neq 0$ for some $e \in E$. The $e \in E_b$ for all e with $b_e = e$ and $e \in A_b^*$ for all e with $r_e \neq 0$ and $b_e \neq b$. Thus

$$0 = \sum_{e \in E} r_e e + A_b^* = \sum_{e \in E_b} r_e e + A_b^*$$

But this contradicts the linear independence of the $(e + A_b^*, e \in E_b)$.

Hence E is linear independent. Suppose that $A \not\leq RE$ and pick $a \in A \setminus RE$ with $b = b_a$ miniaml. Then $a \in A_b$. Hence

$$a + A_b^* = \sum_{e \in E_b} r_e e + A_b^*$$

Put $d = \sum_{e \in E_b} r_e e$. Then $d \in RE \leq A$, $a - d \in A_b^*$. By minimality of b, $a - d \in RE$ and so $a \in d + RE = RE$.

So $A \leq RE$, $A = RE$ and E is a basis for A.

(b) Let I be a left ideal in R. Then $I = Ri$ for some $i \in R$. Since R is an integral domain, $\mathrm{Ann}_R(i) = \{0_R\}$ and so by 4.1.11(f), $R \cong R/\mathrm{Ann}_R(i) \cong Ri$. Then I is free of rank at most 1. Hence $|E_b| \leq 1$ for all $b \in \mathcal{B}$ so (b) holds. □

Among commutative rings R (with identity) only PID's have the property that every left ideal in R is free. Indeed if $a, b \in R^{\#}$, then $ba - ab = 0$ and so a and b are linear dependent. Hence every ideal is cyclic as a module and so a principal ideal. Suppose that $ab = 0$ for some non-zero $a, b \in R$. Then $bRa = 0$ and so Ra is not free, a contradiction. Thus R is also an integral domain and so a PID.

Corollary 4.2.5. *Let R be a PID and M an R-module and W an R-submodule of M. If $M = RI$ for some $I \subseteq M$, then $W = RJ$ for some $J \subseteq W$ with $|J| \leq |I|$. In particular, if M is finitely generated as an R-module, so is M.*

Proof. Let V be a free R-module with basis $(v_i)_{i \in I}$ and let $\phi : V \to M$ be R-linear with $\phi(v_i) = i$ for all $i \in I$. Then

$$\Phi(V) = \Phi(\langle v_i \mid i \in I \rangle) = \langle \Phi(v_i) \mid i \in I \rangle_R = \langle i \mid i \in I \rangle_R = RI = M.$$

Let $A = \Phi^{-1}(W)$. By 4.2.4 A has a basis $(a_k)_{k \in K}$ with $|K| \leq |I|$. Put $J = \Phi(K)$. Since Φ is onto, $\Phi(A) = M$. Thus

$$M = \Phi(A) = \Phi(\langle a_k, k \in K \rangle_R) = \langle \Phi(a_k), k \in K \rangle = R\Phi(K) = RJ$$

and $|J| \leq |K| \leq |I|$. □

Definition 4.2.6. *Let M be a unitary R-module and $m \in M$.*

(a) m is called a torsion element *if $rm = 0_M$ for some $r \in R^{\sharp}$.*

(b) M is called a torsion module *if all elements are torsion elements.*

(c) M is called torsion free *if 0_M is the only torsion element.*

(d) M is a bounded R-module if there exists $r \in R^{\sharp}$ with $rm = 0_M$ for all $m \in M$.

Note that m is not a torsion element if and only if (m) is linearly independent.

Example 4.2.7. Let $R = \mathbb{Z}$. For $n \in \mathbb{Z}$ let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.

Consider $M = \mathbb{Z}_2 \oplus \mathbb{Z}_3$. Since $2(1,0) = (2,0) = (0,0)$, $(1,0)$ is a torsion element. Also $3(0,1) = (0,3) = (0,0)$ and so $(0,1)$ is a torsion element. In fact $6(a,b) = (2(3a), 3(2b)) = (0,0)$ for all $(a,b) \in M$ and so M is bounded.

Consider $M = \mathbb{Z} \oplus \mathbb{Z}$. If $r \in \mathbb{Z}^{\sharp}$ and $(a,b) \in M$ with $r(a,b) = (0,0)$, then $ra = rb = 0$ and so $a = b = 0$. Thus M is torsion free.

Consider $M = \bigoplus_{i \in \mathbb{Z}^+} \mathbb{Z}_i$. Let $m = (m_i)_{i \in \mathbb{Z}^+} \in M$. By definition of the direct sum there exists $k \in \mathbb{Z}^+$ with $m_i = 0$ for all $i > k$. We claim that $k!m = 0_M$. Indeed of $i \leq k$, then $i \mid k!$ and so $k!m_i = 0$ in $\mathbb{Z}_i$. And if $i > k$, then $m_i = 0$ and again $k!m_i = 0$. Thus M is a torsion module. But M is not bounded, indeed suppose that $r \in \mathbb{Z}^+$ with $rm = 0_M$ for all $m \in M$. Let $i \in \mathbb{Z}^+$ and pick $m \in M$ with $m_i = 1$. From $rm = 0$ we get $0 = rm_i = r1 = r$ in $\mathbb{Z}_i$ and so $i \mid r$. Hence $|r| \geq i$ for all $i \in \mathbb{Z}^+$, a contradiction.

Lemma 4.2.8. *Let M be a module for the integral domain R.*

(a) Let I be a finite set of torsion elements in M. Then RI is a bounded R-submodule of M.

(b) Let $\mathrm{T}(M)$ be the set of torsion elements in M. Then $\mathrm{T}(M)$ is R-submodule of M.

(c) $M/\mathrm{T}(M)$ is torsion free.

Proof. (a) For $i \in I$ pick $r_i \in R^{\#}$ with $r_i i = 0_M$. Put $r = \prod_{i\in I} r_i$. Since R is commutative, r is well defined and since R has no zero-divisors, $r \neq 0_R$. Let $j \in I$, then $rj = (\prod_{j\neq i\in I} r_i)r_j j = 0_M$. Since $\mathrm{Ann}_M(r)$ is an R-submodule of M we conclude that $RI \leq \mathrm{Ann}_M(r)$ and so RI is bounded.

(b) Since $1_R 0_M = 0_M$, 0_M is a torsion element. If $x, y \in \mathrm{T}(M)$ and $r \in R$, then by (a), $x + y \in \mathrm{T}(M), -x \in T(M)$ and $rx \in T(M)$. Thus $\mathrm{T}(M)$.

(c) Let $x \in M/\mathrm{T}(M)$ be a torsion element. Pick $m \in M$ with $x = m + \mathrm{T}(M)$ and $r \in R^{\sharp}$ with $rx = 0_{M/\mathrm{T}(M)}$. Then $rm \in \mathrm{T}(M)$ and so $s(rm) = 0_M$ for some $s \in R^{\sharp}$. Hence $(sr)m = 0_M$ and as R is an integral domain, $sr \neq 0_R$. So $m \in T(M)$, $x = m + \mathrm{T}(M) = 0_{M/\mathrm{T}(M)}$ and $M/\mathrm{T}(M)$ is torsion free. □

Theorem 4.2.9. *Let M be an R-module.*

(a) Any linearly independent subset of M lies in a maximal linear independent subset.

(b) Let L be a maximal linear independent subset of M. Then RL is a free R-module with basis L and M/RL is a torsion module.

Proof. Here a subset E of M is called linearly independent if $(e)_{e\in E}$ is linearly independent.

(a) Let E be a linearly independent subset of M. Let $\mathcal{L}$ be the set of linearly independent subsets of M containing E. Since $E \in \mathcal{L}$, $\mathcal{L} \neq \emptyset$. Order $\mathcal{L}$ by inclusion. Let $\mathcal{C}$ be a non-empty chain in $\mathcal{L}$ and put $D = \bigcup \mathcal{C}$. We will show that D is linearly independent. For this let $r_d \in R, d \in D$, almost all zero with $\sum_{d\in D} r_d d = 0_V$. Since almost all r_d's are zero we can choose pairwise distinct $d_i, 1 \leq i \leq n \in D$ such that $r_d = 0_R$ for all $d \in D \setminus \{d_1, d_2, \ldots d_n\}$. Since $d_i \in D = \bigcup \mathcal{C}$, there exists $C_i \in \mathcal{C}$ with $d_i \in C_i$. Since $\mathcal{C}$ is a chain we may assume that $C_1 \subseteq C_2 \subseteq \ldots C_n$ and so $d_i \in C_n$ for all $1 \leq i \leq n$. Note that $\sum_{i=1} r_{d_i} d_i = \sum_{d\in D} r_d d = 0_V$. Since C_n is linearly independent we conclude that $r_{d_i} = 0_R$ for all $1 \leq i \leq n$ and so also $r_d =)_R$ for all $d \in D$. Thus D is linearly independent, $D \in \mathcal{L}$ and D is an upper bound for $\mathcal{C}$.

Thus the assumptions of Zorn's Lemma A.6 are fulfilled and we conclude that $\mathcal{L}$ contains a maximal element L. Then L is a maximal linearly independent subset of M containing E.

(b) Let $v \in V$. We will show that $v + RL$ is a torsion element in V/RL. If $v \in RL$, then $v + RL = 0_{V/RL}$ and so $v + TL$ is torsion element. So suppose $v \notin RL$. Then v $\mathcal{L}$ and by maximality of $\mathcal{L}$, $\{v\} \cup \mathcal{L}$ is linearly dependent. Hence there exist $r \in R$ and $r_l \in R, l \in \mathcal{L}$ almost all but not all equal zero to 0_R such that

$$rv + \sum_{l\in\mathcal{L}} r_l l = 0_V$$

If $r = 0_R$, then since $\mathcal{L}$ is linearly independent, $r_l = 0_R$ for all $l \in \mathcal{L}$, a contradiction. Thus $r \in R^\sharp$ and $rv = -\sum_{l\in\mathcal{L}} r_l l \in RL$. Hence $r(v + RL) = 0_{V/RL}$ and V/RL is a torsion module. □

We remark that if L is a maximal linear independent subset of M, then RL does not have to be a maximal free submodule. Indeed the following example shows that M does not even have to have a maximal free submodule. (Zorn's lemma does not apply as the union of a chain of free submodules might not be free)

Example 4.2.10. Let $R = \mathbb{Z}$ and $M = \mathbb{Q}$ with $\mathbb{Z}$ acting on $\mathbb{Q}$ by left multiplication. As $\mathbb{Q}$ has no zero divisors, $\mathbb{Q}$ is torsion free. In particular, every non-zero element a is linearly independent. We claim $\{a\}$ is a maximal linearly independent subset. Indeed, let $a, b \in \mathbb{Q}^\sharp$. Then $a = \frac{n}{m}$ and $b = \frac{p}{q}$ with $n, p \in \mathbb{Z}$ and $m, q \in \mathbb{Z}^\sharp$. Then

$$(mp)a + (-nq)b = mp\frac{n}{m} - nq\frac{p}{q} = pn - np = 0$$

and $\{a, b\}$ is linearly dependent.

We conclude that every non-zero free submodule of $\mathbb{Q}$ is of the form $\mathbb{Z}a, a \in \mathbb{Q}^\sharp$. Since $\mathbb{Z}a \lneq \mathbb{Z}\frac{a}{2}$ we see that $\mathbb{Q}$ has no maximal free $\mathbb{Z}$-submodule. In particular, $\mathbb{Q}$ is not free $\mathbb{Z}$-module.

$\mathbb{Q}$ as a $\mathbb{Z}$ module has another interesting property: every finitely generated submodules is cyclic (that is generated by one element). Indeed, if A is generated by $\frac{n_i}{m_i}, 1 \le i \le k$, put $m = \mathrm{lcm}_{1\le i\le k} m_i$ and Then $mA \cong A$ and $mA \le \mathbb{Z}$. So mA and A are cyclic.

Corollary 4.2.11. *Let D be a division ring and V a D-module.*

(a) V is torsion free.

(b) Any torsion D-module is a zero module.

(c) Every linear independent subset of V is contained in a basis of V.

(d) V has a basis and so is a free D-module.

Proof. (a) Let $d \in D^\sharp$ and $v \in V$ with $dv = 0_V$. Since D is a division ring $ed = 1_D$ for some $e \in D$. Thus $v = 1_D v = edv = e0_V = 0_V$ and so V is torsion free.

(b) This follows from (a).

(c) Let $\mathcal{L}$ be linearly dependent subset of V. By 4.2.9 $\mathcal{L}$ is contained in a maximal linearly dependent subset $\mathcal{B}$. Also by 4.2.9, $V/R\mathcal{B}$ is a torsion module. By (b) applied to $V/R\mathcal{B}$ we conclude that $V/R\mathcal{B}$ is a zero-module and so $V = R\mathcal{B}$. Hence $\mathcal{B}$ is a basis.

(d) By (c) applied to $\mathcal{L} = \emptyset$, V has a basis and so by 4.2.2 V is free. □

Lemma 4.2.12. *Let R be a ring V an R-module and W be an R-submodule of V. If V/W is a free R-module, then there exists a free R-submodule F of V with $V = F \oplus W$.*

Proof. Let $(v_i)_{i\in I}$ be a tuple of elements in V such that $(v_i + W)_{i\in I}$ is a basis for V/W. Put $F = \langle v_i, i \in I\rangle_R$. Let $v \in V$. Since $(v_i + W)_{i\in I}$ spans V/W:

$$v + W = \sum_{i\in R} r_i(v_i + W)$$

for some $r_i \in R$ almost all zero. Hence $v \in \left(\sum_{i\in I} r_i v_i\right) + W \subseteq F + W$ and so $V = F + W$.

Now let $w \in F \cap W$. Then $w = \sum_{i\in R} r_i v_i$ for some $r_i \in R$ almost all zero. Then

$$\sum_{i\in I} r_i(v_i + W) = w + W = W = 0_{V/W}.$$

Since $(v_i + W)_{i\in I}$ is linearly independent in V/W we conclude $r_i = 0_R$ for all $i \in I$. Thus $w = 0_V$ and $F \cap W = \{0_V\}$ and so $V = F \oplus W$.

Since $V/W = (F + W)/W \cong F/F \cap W \cong F$, F is a free R-module. □

Lemma 4.2.13. *Let M be a torsion free R-module for the integral domain R. Suppose that one of the following holds:*

1. M is finitely generated.

2. If N is a submodule of M such that M/N is a torsion module, then M/N is bounded.

Then there exists a free R-submodule W such that M is isomorphic to a submodule W.

Proof. Note that by 4.2.8 condition (1) implies condition (2). So we may assume that (2) holds. By 4.2.9 there exists a free submodule W of V such that M/W is torsion. By (2) there exists $r \in R^\sharp$ with $rx = 0_{M/W}$. Hence $rM \leq W$.

Consider the map

$$\alpha : M \to W, m \to rm.$$

Since R is commutative, α is a R-linear. As M is torsion free, α is 1-1. Thus $M \cong \alpha(M) = rM \leq W$. □

Lemma 4.2.14. *Let D be a division ring, V a D-module and W a D-submodule. Then there exists a D-submodule K of V with $V = K \oplus W$.*

Proof. By 4.2.11(b), V/W is a free D-module and so the lemma follows from 4.2.12. □

Lemma 4.2.15. *Let $f : A \to B$ be group homomorphism and $C \subseteq A$ with $f(C) = B$. Then $A = \ker f \cdot C$.*

Proof. Let $a \in A$. Then $f(a) = f(c)$ for some $c \in C$ and so $f(ac^{-1}) = e_B$. Thus $ac^{-1} \in \ker f$ and $a = (ac^{-1})c \in \ker f \cdot C$. □

4.3 Modules over PIDs

We continue to assume that R is a ring with identity. Moreover, with module we always mean unitary module.

Definition 4.3.1. *Let M be an R-module. Then M is called* cyclic *if $M = \langle m \rangle_R$ for some $m \in M$.*

Lemma 4.3.2. *Let R be a PID, M an R-module, $m \in M$ and p a prime in R. Suppose that $p^k m = 0_M$ for some $k \in \mathbb{N}$ and let $l \in \mathbb{N}$ be minimal with $p^l m = 0_M$. Then $\mathrm{Ann}_R(m) = p^l R$ and $Rm \cong R/p^k R$ as an R-module.*

Proof. Put $S = \mathrm{Ann}_R(m)$. Then S is an ideal in R and since R is a PID, $S = Rs$ for some $s \in R$. Since $p^l m = 0_M$, $p^l \in S$ and $s \mid p^k$. Thus $s \sim p^t$ for some $t \in \mathbb{N}$ with $t \leq l$. Then $S = Rp^t$ and $p^t m = 0_M$. By minimality of l, $l = t$. Thus $S = Rp^l$ and by 4.1.11(d), $Rm \cong R/Rp^l$. □

Theorem 4.3.3. *Let R be a PID and $p \in R$ a prime. Suppose that M is an R-module with $p^k M = \{0_M\}$ for some $k \in \mathbb{N}$. Then M is a direct sum of non-zero cyclic submodules of M.*

Proof. The proof is by induction on k. If $k = 0$, then, since M is unitary, $M = \{0_M\}$ and the theorem holds.

So suppose $k > 0$. Since $p^{k-1}(pM) = p^k M = \{0_M\}$ we conclude by induction on k that there exist non-zero cyclic submodules $A_i, i \in I$ of pM with $M = \bigoplus_{i \in I} A_i$. Since A_i is cyclic $A_i = \langle a_i \rangle_R = Ra_i$ for some $a_i \in A_i$. Thus

1°. $pM = \bigoplus_{i \in I} Ra_i$

Since A_i is non-zero, $a_i \neq 0_M$. Since $a_i \in pM$ there exists $b_i \in B$ with $a_i = pb_i$. Put $B = \langle b_i, i \in I \rangle_R = \sum_{i \in I} Rb_i$. We claim that

2°. $B = \bigoplus_{i \in I} Rb_i$

For this let $r_i \in R, i \in I$, almost all zero with

$$(*) \qquad \sum_{i \in I} r_i b_i = 0_M.$$

We need to show that $r_i b_i = 0_M$ for all $i \in I$. From (*) we have

$$\sum_{i \in I} r_i a_i = \sum_{i \in I} r_i p b_i = p \sum_{i \in I} r_i b_i = p 0_M = 0_M.$$

Thus (1°) implies that $r_i a_i = 0_M$ for all $i \in I$. Let $S_i = \mathrm{Ann}_R(a_i)$. By 4.3.2 $S_i = Rp^{l_i}$ for some $l_i \in \mathbb{N}$. Since $a_i \neq 0_M$, $l_i \neq 0_R$. Since $r_i \in S_i$ we get $p^{l_i} \mid r_i$. Thus $p \mid r_i$ and $r_i = t_i p$ for some $t_i \in R$. Then $r_i b_i = t_i p b_i = t_i a_i$. Thus

$(**)$ $$r_i b_i = t_i a_i.$$

Substitution into (*) gives:

$$\sum_{i\in I} t_i a_i = 0_M.$$

Thus by (1°), $t_i a_i = 0_M$ and by $(**)$ $r_i b_i = 0_M$. Thus (2°) holds.

3°. $M = \mathrm{Ann}_M(p) + B$.

We have $pB = p\sum_{i\in I} Rb_i = \sum_{i\in I} Rpb_i = \sum_{i\in I} Ra_i = pM$. Define $\alpha : M \to pM, m \to pm$. Then α is R-linear and $\alpha(B) = pM$. Thus by 4.2.15 $M = \ker\alpha + B = \mathrm{Ann}_M(p) + B$.

4°. *R/Rp is a field and $\mathrm{Ann}_M(p)$ is module for R/Rp.*

Since p is a prime, R/Rp is a field by 3.3.11. Since $Rp \le \mathrm{Ann}_R(\mathrm{Ann}_M(p))$, $\mathrm{Ann}_M(p)$ is an R/Rp module via $(r + Rp)m = rm$.

5°. *There exists an R-submodule D of $\mathrm{Ann}_M(p)$ with $\mathrm{Ann}_M(p) = D \oplus \mathrm{Ann}_B(p)$ and $M = D \oplus B$.*

Since R/Rp is a field we conclude from 4.2.14 that $\mathrm{Ann}_M(p) = D \oplus \mathrm{Ann}_B(p)$ for some R/Rp submodule D of $\mathrm{Ann}_M(p)$. Then D is also an R-submodule of $\mathrm{Ann}_M(p)$. We have $M = \mathrm{Ann}_M(p) + B = D + \mathrm{Ann}_B(p) + B = D + B$ and $D \cap B = D \cap \mathrm{Ann}_M(p) \cap B = D \cap \mathrm{Ann}_B(p) = \{0_M\}$. So $M = D \oplus M$.

We now can complete the proof of the theorem. By 4.2.11(b), the R/Rp-module D has a basis $(d_j)_{j\in J}$. Then $D = \bigoplus_{j\in J} R/pR \cdot b_j = \bigoplus_{j\in J} Rd_j$. Together with (2°) and (5°) we get

$$M = D \oplus B = \bigoplus_{j\in J} Rd_j \oplus \bigoplus_{i\in I} Rb_i$$

□

Theorem 4.3.4. *Let M be a finitely generated module for the PID R. Then there exists a free submodule $F \le M$ with $M = F \oplus \mathrm{T}(M)$.*

Proof. By 4.2.8, $M/\mathrm{T}(M)$ is torsion free, so by 4.2.13 $M/\mathrm{T}(M)$ is isomorphic to a submodule of a free module. Hence by 4.2.4 $M/\mathrm{T}(M)$ is free. Thus by 4.2.12 $M = F \oplus \mathrm{T}(M)$ for a free R-submodule F of M. □

Definition 4.3.5. *Let R be a commutative ring, $r \in R$ and $A \subseteq R$.*

(a) We say that r is a common divisor *of A and write $r \mid A$ if $r \mid a$ for all $a \in A$.*

(b) We say that r is a greatest common divisor *and write $r \sim \gcd A$ if r is common divisor of A and $s \mid r$ for all common divisor s of A.*

We remark that in a general commutative ring a given set of elements might not have a greatest common divisor.

Lemma 4.3.6. *Let R be a commutative ring, $r \in R$ and $A \subseteq R$.*

(a) $r \mid A$ if and only if $(A) \subseteq (r)$.

(b) Suppose $s \in R$ is a gcd of A, then r is a gcd of A if and only if $r \sim s$.

(c) The following two statement are equivalent:

(a) r is a gcd of A.

(b) For all $s \in R$: $\quad s \mid A \iff s \mid r$.

Proof. (a) By definition of dividing, $r \mid a$ if and only if $(a) \subseteq (r)$. Since (r) is an ideal, $(a) \subseteq (r)$ for all $a \in A$ if and only if $(A) \subseteq (r)$. Thus (a) holds.

(b) If $s \sim r$, then $(s) = (r)$. Since the definition of a gcd only depends (s) and not on s, we conclude that r is a gcd of A.

Suppose r is a gcd of A. Then since s is a common divisor of A, $s \mid r$. By symmetry $r \mid s$ and so $r \sim s$.

(c) Suppose r is a gcd. If $s \mid A$, then $s \mid r$ by definition of a gcd. If $s \mid r$, then since $r \mid A$ also $s \mid A$.

Suppose for all $s \in R$ we have $s \mid A \iff s \mid r$. Since $r \mid r$ we get $r \mid A$. Also $s \mid r$ for all s with $s \mid A$ and so r is a gcd of A. □

Definition 4.3.7. *Let R be commutative ring, $b \in R$, $a \in R$ $A \subseteq R$. Define*

$$a_b := \sup\{n \in \mathbb{N} \,\big|\, b^n \mid a\} \quad \text{and} \quad A_b := \sup\{n \in \mathbb{N} \,\big|\, b^n \mid A\}$$

Note here that if $b^n \mid a$ for all $n \in \mathbb{N}$, then $a_b = \infty$. For example if $a = 0_R$ then $a_p = \infty$ for all $b \in R$.

Lemma 4.3.8. *Let R be a UFD and $\mathcal{P}$ a set of representatives for the primes in R, that is $\mathcal{P}$ is a set of primes and each prime in R is associate to exactly one element in $\mathcal{P}$.*

(a) Let p be a prime and $a \in R$. If $a = 0_R$, then $a_p = \infty$. If a is a unit then $a_p = 0$. If a is proper and $a = p_1 p_2 \dots p_n$ with each p_i a prime in R, then

$$a_p = |\{i \mid 1 \le i \le n, p_i \sim p\}|.$$

(b) Let $a \in R^\sharp$. Then $a \sim \prod_{p \in \mathcal{P}} p^{a_p}$.

(c) Let $a, b \in R$ and p a prime. Then $(ab)_p = a_p + b_p$.

(d) Let $a, b \in R$. Then $a \mid b$ if and only if $a_p \leq b_p$ for all $p \in \mathcal{P}$.

(e) Let $a \in R$ and $B \subseteq R$. Then $a \mid B$ if and only if $a_p \leq B_p$ for all $p \in \mathcal{P}$.

(f) Let $A \subseteq R$. If $A \subseteq \{0_R\}$, then 0_R is a gcd of A. If $A \nsubseteq \{0_R\}$, then $\prod_{p\in\mathcal{P}} p^{A_p}$ is a gcd of A.

Proof. Let $a, b \in R$ with $a \mid b$, let p be a prime in R and let $k \in \mathbb{N}$.

Suppose a is proper and let $a = p_1 \dots p_n$ there each p_i is a prime. For each $p \in \mathcal{P}$, let $a_p^* = |\{i \mid p_i \sim p\}|$. Then it is easy to see that $a \sim \prod_{p\in\mathcal{P}} p^{a_p^*}$. So $p^{a_p^*} \mid a$ and $a_p^* \leq a_p$. Since $a \mid b$, $b = ac$ for some $c \in R$. Thus

$$\prod_{p\in\mathcal{P}} p^{b_p^*} \sim b = ac \sim \prod_{p\in\mathcal{P}} p^{a_p^*} \prod_{p\in\mathcal{P}} p^{c_p^*} = \prod_{p\in\mathcal{P}} p^{a_p^*+c^*}$$

So by the uniqueness of prime factorizations 3.3.12(b) we get $b_p^* = a_p^* + c_p^*$. Thus $a_p^* \leq b_p^*$. Since $p^{a_p} \mid a$, we conclude that $a_p \leq a_p^*$ and so $a_p = a_p^*$. Hence (a), (b), (c) and the forward direction of (d) are proved.

Suppose $a_p \leq b_p$ for all $p \in \mathcal{P}$. Put $c = \prod_{p\in\mathcal{P}} p^{b_p - a_p}$. Then by (b), $ac \sim b$ and so $a \mid b$. Thus (d) holds.

If follows immediately from the definition of B_p that $B_p = \min\{b_p \mid b \in B\}$. So we have

$$a \mid B \iff a \mid b, \forall b \in B \iff a_p \leq b_p, \forall b \in B, p \in \mathcal{P} \iff a_p \leq B_p \forall p \in \mathcal{P}$$

So (e) holds.

Put $r = \prod_{p\in\mathcal{P}} p^{A_p}$ and let $s \in R$. Then by (a) $r_p = A_p$. Hence

$$s \mid r \iff s_p \leq r_p \iff s_p \leq A_p \iff s \mid A.$$

Hence by 4.3.6(c), r is a gcd of A. □

Lemma 4.3.9. *Let R be a commutative ring, $A \subseteq R$ and $r \in R$. Suppose that A is a principal ideal in R. Then r is a gcd for A if and only of $(A) = (r)$.*

Proof. By assumption there exists $t \in R$ with $(A) = (t)$. Let $s \in R$. Then

$$s \mid A \overset{4.3.6(a)}{\iff} (A) \subseteq (s) \iff (t) \subseteq (s) \iff s \mid t.$$

So by 4.3.6(c), t is a gcd of A. Hence by 4.3.6(b), r is a gcd of A if and only if $(r) = (t)$ and so if and only if $(r) = (A)$. □

Theorem 4.3.10. *Let R be a PID and M a torsion module for R. Let $\mathcal{P}$ a set of representatives for the primes in R. For $p \in \mathcal{P}$ put $M_p = \bigcup_{k\in\mathbb{Z}^+} \mathrm{Ann}_M(p^k)$. Then*

$$M = \bigoplus_{p\in\mathcal{P}} M_p.$$

Proof. Let $m \in M$. Since M is a torsion modules there exists $r \in R^{\#}$ with $rm = 0_M$. By 3.3.18 PIDs are UFDs and so by 4.3.8(b) there exist pairwise distinct $p_i \in \mathcal{P}$ and $k_i \in \mathbb{Z}^+$, $1 \leq k \leq n$, with $r \sim p_1^{k_1} \dots p_n^{k_n}$. So we may assume that

$$r = p_1^{k_1} \dots p_n^{k_n}.$$

Put $a_i = \prod_{j \neq i} p_j^{k_j}$. Then $r = p_i^{k_i} a_i$. Since $p_i \nmid a_i$ we conclude that $\gcd_{i=1}^n a_i \sim 1_R$ and so by 4.3.9 $1_R = \sum s_i a_i$ for some $s_i \in R$. Put $m_i = s_i a_i m$. Then

$$m = 1_R\, m = \left(\sum_{i=1}^n s_i a_i \right) m = \sum_{i=1}^n s_i a_i m = \sum_{i=1}^n m_i$$

and

$$p_i^{k_i} m_i = p_i^{k_i} s_i a_i m = s_i (p_i^{k_i} a_i) m = s_i (rm) = 0_M.$$

Thus $m_i \in \mathrm{Ann}_M(p_i^{k_i}) \leq M_{p_i}$ and $m = \sum_{i=1}^n m_i \in \sum_{p \in \mathcal{P}} M_p$. Therefore

$$M = \sum_{p \in \mathcal{P}} M_p.$$

Let $p \in \mathcal{P}$. Put

$$M_{p'} := \sum_{p \neq q \in \mathcal{P}} M_q.$$

In remains to show that $M_{p'} \cap M_p = \{0_M\}$. For this let $k \in \mathbb{Z}^+$ and $0_M \neq m \in M_{p'}$. Then there exist primes $p_1, p_2, \dots p_n$ and $m_i \in M_{p_i}$ such that $p \neq p_i$ and $m = \sum_{i=1}^n m_i$. Then $p_i^{l_i} m_i = 0_M$ for some $l_i \in \mathbb{N}$. Put $a = \prod_{i=1}^n p_i^{l_i}$. Then $am = 0_M$ and $p \nmid a$. Hence by 4.3.9 $1_R = ra + sp^k$ for some $r, s \in R$ and $m = ram + sp^k m = sp^k m$. Thus $p^k m \neq 0_M$ and so $m \notin M_p$. □

Lemma 4.3.11. *Let R be a ring, and $(M_i)_{i \in I}$ a family of non-zero R-modules. If $\bigoplus_{i \in I} M_i$ is finitely generated, then I is finite.*

Proof. Let A be a finite subset of $M := \bigoplus_{i \in I} M_i$ with $M = \langle A \rangle_R$. By definition of "direct sum" each m is a tuple $(m_i)_{i \in I}$ with almost all m_i zero. So for $a \in A$ we can choose a finite subset J_a of I with $a_k = 0$ for all $k \in I \setminus J_a$. Put $J = \bigcup_{a \in A} J_a$. Then J is finite. We will show that $J = I$. For this let $i \in I$ and put $W = \{m \in M \mid m_i = 0\}$. Then W is a R-submodule of M and since $M_i \neq 0$, $W \neq M$. Since $M = \langle A \rangle_R$ we get $A \nsubseteq W$ and so $a_i \neq 0$ for some $a \in A$. Thus $i \in J_a \subseteq J$, $I = J$ and I is finite. □

Theorem 4.3.12. *Let M be a finitely generated module for the PID R. Then M is direct sum of finitely many cyclic R-modules. Moreover, each of the summand can be chosen be isomorphic to R or $R/p^k R$ for some prime ideal $p \in R$ and some $k \in \mathbb{Z}^+$. In other words,*

$$M \cong \underbrace{R \oplus R \oplus \dots \oplus R}_{k\text{-times}} \oplus R/p_1^{k_1} R \oplus R/p_2^{k_2} R \oplus \dots \oplus R/p_n^{k_n} R$$

for some $k, n \in \mathbb{N}$, $k_1, k_2 \ldots, k_n \in \mathbb{Z}^+$ *and primes* $p_1, p_2, \ldots, p_n \in R$.

Proof. By 4.3.4, $M = F \oplus T(M)$, where F is a free R-module. So F is a direct sum of copies of R. Also by 4.3.10 $T(M) = \bigoplus_{p \in \mathcal{P}} M_p$, where $\mathcal{P}$ is set of representatives for the associate classes of primes in R. Let $p \in \mathcal{P}$. Since M is finitely generated and M_P is a homomorphic image of M, M_p is finite generated. Thus $M_p = \langle I \rangle_R$ for some finite subset I of M_P. For $i \in I$ pick $l_i \in \mathbb{N}$ with $p^{l_i} i = \{0_M\}$ and put $l = \max_{i \in I} l_i$. Then $p^l M_p = \{0_M\}$. Thus by 4.3.3 M_p is the direct sum of non-zero cyclic submodules. By 4.3.2 each of these cyclic submodules is isomorphic to $R/p^k R$ for some $k \in \mathbb{Z}^+$.

It follows that M is a direct sum modules of the form R or $R/p^k R$, $p \in \mathcal{P}$, $k \in \mathbb{Z}^+$. Since M is finitely generated, 4.3.11 this direct sum is a finite direct sum. □

Corollary 4.3.13. *(a) Let A be a finitely generated abelian group. Then A is the direct sum of cyclic groups.*

(b) Let A be an elementary abelian *p-group for some prime p. (That is A is abelian and $pA = 0$). Then A is the direct sum of copies of $\mathbb{Z}/p\mathbb{Z}$.*

Proof. Note that an abelian group is nothing else as a module over $\mathbb{Z}$. So (a) follows from 4.3.12 and (b) follows from 4.3.3 and 4.3.2

(b) can also by proved by observing that A is also a module over the field $\mathbb{Z}/p\mathbb{Z}$ and so has a basis. □

4.4 Exact Sequences

Definition 4.4.1. *A (finite or infinite)* sequence *of R-linear maps*

$$\ldots \xrightarrow{f_{i-2}} A_{i-2} \xrightarrow{f_{i-1}} A_{i-1} \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \xrightarrow{f_{i+2}} \ldots$$

is called exact *if for all suitable $j \in \mathbb{Z}$*

$$\operatorname{Im} f_j = \ker f_{j+1}$$

We denote the zero R-module with 0. Then for all R-modules M there exists unique R-linear maps, $0 \to M$ and $M \to 0$.

The sequence

$$0 \to A \xrightarrow{f} B$$

is and only if f is one to one.

$$A \xrightarrow{f} B \to 0$$

is exact if and only if f is onto.

$$0 \to A \xrightarrow{f} B \to 0$$

is exact if and only if f is an isomorphism.

A sequence of the form

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is called a short sequence. If it is exact we have that f is one to one, $\ker g = \operatorname{Im} f$ and g is onto. Since f is one to one we have $\operatorname{Im} f \cong A$ and so $\ker g \cong A$. Since g is onto the isomorphisms theorem says $B/\ker g \cong C$. So the short exact sequence tells us that B has a submodule which isomorphic to A and whose quotient is isomorphic to C.

Given two exact sequences

$$\mathcal{A} : \xrightarrow{f_{i-1}} A_{i-1} \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} \text{ and } \mathcal{B} : \xrightarrow{g_{i-1}} B_{i-1} \xrightarrow{g_i} B_i \xrightarrow{g_{i+1}}$$

A *homomorphism* of exact sequences $\varphi : \mathcal{A} \to \mathcal{B}$ is a tuple of R-linear maps $(h_i : A_i \to B_i)$ so that the diagram

$$\begin{array}{ccccccccc}
\xrightarrow{f_{i-1}} & A_{i-1} & \xrightarrow{f_i} & A_i & \xrightarrow{f_{i+1}} & A_{i+1} & \xrightarrow{f_{i+2}} \\
 & \downarrow{\scriptstyle h_{i-1}} & & \downarrow{\scriptstyle h_i} & & \downarrow{\scriptstyle h_{i+1}} & \\
\xrightarrow{g_{i-1}} & B_{i-1} & \xrightarrow{g_i} & B_i & \xrightarrow{f_{i+1}} & B_{i+1} & \xrightarrow{g_{i+2}}
\end{array}$$

commutes. $\mathrm{id}_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}$ is defined as (id_{A_i}). φ is called as isomorphism if there exists $\vartheta : \mathcal{B} \to \mathcal{A}$ with $\vartheta\varphi = \mathrm{id}_\alpha$ and $\varphi\vartheta = \mathrm{id}_{\mathcal{B}}$. It is an easy exercise to show that φ is an isomorphism and if and only if each h_i is.

Theorem 4.4.2 (Short Five Lemma)**.** *Given a homomorphism of short exact sequences:*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \beta} & & \downarrow{\scriptstyle \gamma} & & \\
0 & \longrightarrow & A' & \xrightarrow{f'} & B' & \xrightarrow{g'} & C' & \longrightarrow & 0
\end{array}$$

Then

(a) If α and γ are one to one, so is β.

(b) If α and γ are onto, so is β.

(c) If α and γ are isomorphisms, so is β.

Proof. (a) Let $b \in B$ with $\beta(b) = 0$. Then also $g'(\beta(b)) = 0$ and as the diagram commutes $\gamma(g(b)) = 0$. As γ is one to one $g(b) = 0$. As $\ker g = \operatorname{Im} f$, $b = f(a)$ for some $a \in A$. Thus $\beta(f(a)) = 0$ and so $f'(\alpha(a)) = 0$. As f' is one to one, $\alpha(a) = 0$. As α is one to one, $a = 0$. So $b = f(a) = 0$ and β is one to one.

(b) Let $b' \in B'$. As γ and g are onto, so is $\gamma \circ g$. So there exists $b \in B$ with $g'(b') = \gamma(g(b))$. As the diagram commutes $\gamma(g(b)) = g'(\beta(b))$. Thus $d := b' - \beta(b) \in \ker g'$. As $\ker g' = \operatorname{Im} f'$ and α is onto, $\ker g' = \operatorname{Im}(f' \circ \alpha)$. So $d = f'(\alpha(a))$ for some $a \in A$. As the diagram commutes, $d = \beta(f(a))$. So

$$b' - \beta(b) = d = \beta(f(a))$$

Hence $b' = \beta(b + f(a))$ and β is onto.

(c) follows directly from (a) and (b). □

Theorem 4.4.3. *Given a short exact sequence* $0 \to A \xrightarrow{f} B \xrightarrow{g} C \to 0$. *Then the following three statements are equivalent:*

(a) There exists a R-linear map $\gamma : C \to B$ *with* $g \circ \gamma = \mathrm{id}_C$.

(b) There exists a R-linear map $\eta : B \to A$ *with* $\eta \circ f = \mathrm{id}_A$.

(c) There exists $\tau : B \to A \oplus C$ *so that*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle \tau} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\rho_1} & A \bigoplus C & \xrightarrow{\pi_2} & C & \longrightarrow & 0 \end{array}$$

is an isomorphism of short exact sequences.

Proof. (a)⇒ (c) Consider

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{\rho_1} & A \bigoplus C & \xrightarrow{\pi_2} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle (f,\gamma)} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \end{array}$$

Here $(f, \gamma) : A \oplus C \to B, (a, c) \to f(a) + \gamma(c)$. It is readily verified that this is a homomorphism. The Short Five Lemma 4.4.2 implies that is an isomorphism.

(b)⇒ (c) This time consider

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \| & & \downarrow{\scriptstyle (\eta,g)} & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\rho_1} & A \bigoplus C & \xrightarrow{\pi_2} & C & \longrightarrow & 0 \end{array}$$

(c)⇒ (a)& (b) Define $\eta = \pi_1 \circ \tau$ and $\gamma = \tau^{-1}\rho_2$. Then

$$\eta \circ f = \pi_1 \circ (\tau \circ f) = \pi_1 \circ \rho_1 = \mathrm{id}_A$$

and

$$g \circ \gamma = (g \circ \tau^{-1}) \circ \rho_2 = \pi_1 \circ \rho_2 = \mathrm{id}_C$$

□

An exact sequence which fulfills the three equivalent conditions in the previous theorem is called *split*.

To make the last two theorems a little more transparent we will restate them in an alternative way. First note that any short exact sequence can be viewed as pair of R modules $D \leq M$. Indeed, given $D \leq M$ we obtain a short exact sequence

$$0 \longrightarrow D \longrightarrow M \longrightarrow M/D \longrightarrow 0$$

Here $D \to M$ is the inclusion map and $M \to M/D$ is the canonical epimorphism. Conversely, every short exact sequence is isomorphic to one of this kind:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \xrightarrow{f} & B & \xrightarrow{g} & C & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \| & & \downarrow{\scriptstyle \bar{g}^{-1}} & & \\ 0 & \longrightarrow & \operatorname{Im} f & \longrightarrow & B & \longrightarrow & B/\operatorname{Im} f & \longrightarrow & 0 \end{array}$$

Secondly define a homomorphism $\Phi : (A \leq B) \to (A' \leq B')$ to be a homomorphism $\Phi : B \to B'$ with $\Phi(A) \leq A'$

Such a Φ corresponds to the following homomorphism of short exact sequences:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & B/A & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \Phi_A} & & \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \Phi_{B/A}} & & \\ 0 & \longrightarrow & A' & \longrightarrow & B' & \longrightarrow & B'/A' & \longrightarrow & 0 \end{array}$$

Here $\Phi_A : A \to A' : a \to \Phi(a)$ and $\Phi_{B/A} : B/A \to B'/A' : b + A \to \Phi(b) + A'$. Since $\Phi(A) \leq A'$ both of these maps are well defined.

Lets translate the Five Lemma into this language:

Lemma 4.4.4. *Let $\Phi : (A \leq B) \to (A' \leq B')$ be a homomorphism.*

(a) If Φ_A and $\Phi_{B/A}$ are one to one, so is Φ.

(b) If Φ_A and $\Phi_{B/A}$ are onto so is Φ.

(c) If Φ_A and $\Phi_{B/A}$ are isomorphism, so is Φ.

Proof. This follows from the five lemma, but we provide a second proof.
(a) As $\ker \Phi_{B/A} = 0$, $\ker \Phi \leq A$. So $\ker \Phi = \ker \Phi_A = 0$.
(b) As $\Phi_{B/A}$ is onto, $B' = \Phi(B) + A'$. As $\Phi(A) = A'$ we conclude $B' = \Phi(B)$.
(c) Follows from (a) and (b). □

The three conditions on split exact sequences translate into:

Lemma 4.4.5. *Given a pair of R-modules $A \leq B$. The following three conditions are equivalent.*

(a) There exists a homomorphism $\gamma : B/A \to B$ with $\bar{b} = \gamma(\bar{b}) + A$ for all $\bar{b} \in B$.

(b) There exists a homomorphism $\eta : B \to A$ with $\eta(a) = a$ for all $a \in A$.

(c) There exists a R-submodule K of B with $B = A \bigoplus K$.

Proof. Again this follows from 4.4.3 but we give a second proof:

(a)⇒ (c): Put $K = \gamma(B/A)$. Then clearly $K + A = B$. Also if $\gamma(b + A) \in A$ we get $b + A = A = 0_{B/A}$. Thus $\gamma(b + A) = 0$ and $K \cap A = 0$.

(b)⇒ (c) Put $K = \ker \eta$. The clearly $K \cap A = 0$. Also if $b \in B$. Then $\eta(b) \in A$ and $\eta(b - \eta(b)) = \eta(b) - \eta(b) = 0$. Thus $b = \eta(b) + (b - \eta(b)) \in A + B$. Thus $B = A + K$.

(c) ⇒ (a): Define $\gamma(k + A) = k$ for all $k \in K$.

(c) ⇒ (b): Define $\eta(a + k) = a$ for all $a \in A, k \in K$ □

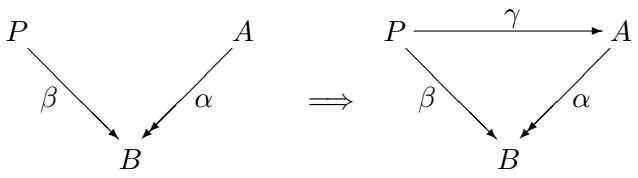
Finally if A is a R-submodule of B we say that B splits over A if the equivalent conditions in the previous lemma hold.

4.5 Projective and injective modules

In this section all rings are assumed to have an identity and all R-modules are assumed to by unitary.

We write $\phi : A \twoheadrightarrow B$ if $\phi : A \to B$ is onto. And $\phi : A \rightarrowtail B$ if ϕ is one to one.

Definition 4.5.1. *Let P be a module over the ring R. We say that P is projective provided that*



where both diagrams are commutative.

Lemma 4.5.2. *Any free module is projective.*

Proof. Given $\alpha : A \twoheadrightarrow B$ and $\beta : F_R(I) \to B$. Let $i \in I$. Since α is onto, $\beta(i) = \alpha(a_i)$ for some $a_i \in A$. By the definition of a free module there exists $\gamma : F_R(I) \to A$ with $\gamma(i) = a_i$. Then

$$\alpha(\gamma(i) = \alpha(a_i) = \beta(i).$$

So by the uniqueness assertion in the definition of a free module $\alpha \circ \gamma = \beta$. □

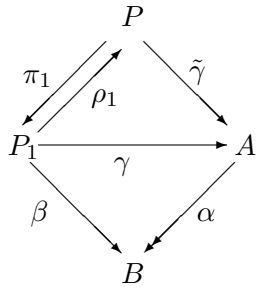
Let A and B be R-modules. We say that A is a *direct summand* of B if $A \leq B$ and $B = A \oplus C$ for some $C \leq B$.

Note that if A is a direct summand of B and B is direct summand of C then A is a direct summand of C. Also if A_i is a direct summand of B_i, then $\bigoplus_{i \in I} A_i$ is a direct summand of $\bigoplus_{i \in I} B_i$.

Lemma 4.5.3. *Any direct summand of a projective module is projective.*

Proof. Let P be projective and $P = P_1 \bigoplus P_2$ for some submodules P_i of P. We need to show that P_1 is projective. Given $\alpha : A \twoheadrightarrow B$ and $\beta : P_1 \to B$. Since P is projective there exists $\tilde{\gamma} : P \to A$ with

$$\alpha \circ \tilde{\gamma} = \beta \circ \pi_1$$



Put $\gamma = \tilde{\gamma}\rho_1$. Then

$$\alpha \circ \gamma = \alpha \circ \tilde{\gamma} \circ \rho_1 = \beta \circ \pi_1 \circ \rho_1 = \beta$$

□

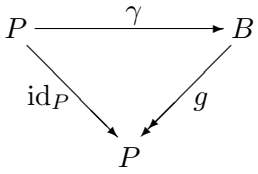
Theorem 4.5.4. *Let P be a module over the ring R. Then the following are equivalent:*

(a) P is projective.

(b) Every shot exact sequence $0 \to A \xrightarrow{f} B \xrightarrow{g} P \to 0$ splits.

(c) P is (isomorphic to) a direct summand of a free module.

Proof. (a)⇒ (b): Since P is projective we have



So the exact sequence is split by 4.4.3a.

(b) ⇒ (c): Note that P is the quotient of some free module F. But then by (b) and 4.4.3c, P is isomorphic to a direct summand of F.

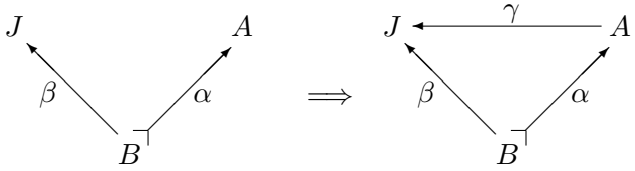
(c)⇒ (a): Follows from 4.5.2 and 4.5.3. □

Corollary 4.5.5. *Direct sums of projective modules are projective.*

Proof. Follows from 4.5.4c. □

Next we will dualize the concept of projective modules.

Definition 4.5.6. *A module J for the ring R is called* injective *if*

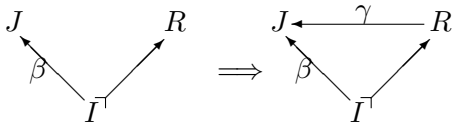


where both diagrams are commutative.

Above we showed that free modules are projective and so every module is the quotient of a projective module. To dualize this our first goal is to find a class of injective R-modules so that every R-modules is embedded into a member of the class. We do this into step steps: First we find injective modules for $R = \mathbb{Z}$. Then we use those to define injective modules for an arbitrary ring (with identity).

To get started we prove the following lemma, which makes it easier to verify that a given module is injective.

Lemma 4.5.7. *Let J be a module over the ring R. Then J is injective if and only if for all left ideals I in R:*



where both diagrams are commutative.

Proof. Given $\alpha : B \rightarrowtail A$ and $\beta : B \to J$, we need to find $\gamma : B \to J$ with $\beta = \gamma\alpha$. Without loss, $B \leq A$ and α is the inclusion map. $\beta = \gamma\alpha$ now just means $\gamma \mid_B = \beta$.

That is we are trying to extend β to A. We will use Zorn's lemma find a maximal extension of β. Indeed let

$$\mathcal{M} = \{\delta : D \to J \mid B \leq D \leq A, \delta \mid_B = \beta\}$$

Order $\mathcal{M}$ by $\delta_1 \leq \delta_2$ if

$$D_1 \subseteq D_2 \text{ and } \delta_2 \mid_{D_1} = \delta_1$$

We claim that every chain $\{\delta_i : D_i \to J \mid i \in I\}$ in $\mathcal{M}$ has an upper bound. Let $D = \bigcup_{i\in I} D_i$ and defone $\delta : D \to J$ by $\delta(d) = \delta_i(d)$ if $d \in D_i$ for some $i \in I$. It is easy to verify that δ is well defined, $\delta \in \mathcal{M}$ and δ is an upper bound for $delta_i : D_i \to J \mid i \in I$.

Hence by Zorn's lemma, $\mathcal{M}$ has a maximal element $\delta : D \to J$.

The reader might have noticed that we did not use our assumptions on J yet. Maximal extensions always exists.

Suppose that $D \neq B$ and pick $b \in B \setminus D$.

Consider the R-linear map:

$$\mu : D \oplus R \to A, \quad (d, r) \to d + rb$$

Let I be the projection of $\ker \mu$ onto D. Then as $\ker \mu$ is a submodule of $D \oplus R$, I is a submodule of R, that is a left ideal. Moreover, $\ker \mu = \{(-ib, i) \mid i \in I$ and I consists of all $r \in R$ with $rb \in D$. Consider the map $\xi : I \to J, i \to \delta(ib))$. By assumption ξ can be extended to a map

$$\xi : R \to J \text{ with } \xi(i) = \delta(ib).$$

Define $\Xi : D \oplus R \to J, (d, r) \to \delta(d) + \xi(r)$. Then Ξ is R-linear. Also $\Xi(-ib, i) = -\delta(ib) + \xi(i) = -\delta(ib) + \delta(ib) = 0$. Hence $\ker \mu \leq \ker \Xi$ and we obtain a R-linear map

$$\bar{\Xi} : (D \bigoplus R)/\ker \mu \to J.$$

So by the Isomorphism Theorem we conclude that

$$D + Rb \to J, \quad d + rb \to \delta(d) + \xi(r)$$

is a well defined R-linear map. Clearly its contained in $\mathcal{M}$, a contradiction to the maximal choice of δ.

Thus $D = B$ and J is injective. The other direction of the lemma is obvious. □

Lemma 4.5.8. *Let R be a ring and M an R-module. Then*

$$\Delta : \mathrm{Hom}_R(R, M) \to M, \, \phi \to \phi(1)$$

is a $\mathbb{Z}$-isomorphism.

Proof. Clearly Δ is $\mathbb{Z}$-linear. To show that Δ is an bijective we will find an inverse. Let $m \in M$. Define

$$\Gamma(m) : R \to M, \, r \to rm$$

. The claim that $\Gamma(m)$ is R-linear. Indeed its $\mathbb{Z}$-linear and

$$\Gamma(m)(sr) = (sr)m = s(rm)$$

for all $s, t \in R$. So $\Gamma(m) \in \mathrm{Hom}_R(R, M)$. Also

$$\Delta(\Gamma(m)) = \Gamma(m)(1) = 1m = m$$

and for $\phi \in \mathrm{Hom}_R(R, M)$,

$$(\Gamma(\Delta(\phi))(r) = r\Delta(\phi) = r\phi(1) = \phi(r1) = \phi(r)$$

So $\Gamma(\Delta(\phi)) = \phi$ and Γ is the inverse of Δ. □

Let R be an integral domain. We say that the R-module M is *divisible* if $rM = M$ for all $r \in R^{\#}$. Note that every quotient of a divisible module is divisible. Also direct sums and direct summand of divisible modules are divisible

If R is divisible as an R-modules if and only if R is a field. The field of fraction, $\mathbb{F}_R$ is divisible as an R-module.

Lemma 4.5.9. *Let R be an integral domain and M an R-module.*

(a) If M is injective, then M is divisible.

(b) If R is a PID, M is injective if and only of M is divisible.

Proof. (a) Let $0 \neq t \in R$ and $m \in M$ Consider the map

$$Rt \to M, rt \to rm$$

As I is an integral domain this is well defined and R-linear. As M is injective this homomorphism can be extended to a homomorphism $\gamma : R \to M$. Then $t\gamma(1) = \gamma(t1) = \gamma(t) = m$. Thus $m \in tR$ and $R = tR$ so M is divisible.

(b) Suppose that M is divisible. Let I be a ideal in R and $\beta : I \to M$ a R-linear map. As R is a PID, $I = Rr$ for some $t \in R$. As M is divisible, $\beta(t) = tm$ for some $m \in M$. Define

$$\gamma : R \to M, r \to rm$$

Then γ is R-linear and $\gamma(rt) = rtm = \beta(rt)$. We showed that the condition of 4.5.7 are fulfilled. So M is injective. □

Proposition 4.5.10. *Let R be a integral domain.*

(a) Every R module can be embedded into an divisible R-module.

(b) If R is a PID, then every R-module can be embedded into a injective module.

Proof. (a) Let M a R module. Then

$$M \cong A/B$$

where $A = \bigoplus_{i\in I} R$for some set I and B is a submodule of A. Let $D = \bigoplus_{i\in I} \mathbb{F}_R$. Then D is divisible and $B \leq A \leq D$. Also D/B is divisible and A/B is a submodule of D/B isomorphic to M.

(b) follows from (a) and 4.5.9. □

An abelian group A is called *divisible* if it is divisible as $\mathbb{Z}$-module.

Let R be a ring and A, B and T be R-modules. Let $\phi : A \to B$ be R-linear. Then the maps

$$\phi^* : \mathrm{Hom}_R(B, T) \to \mathrm{Hom}_R(A, T),\ f \to f \circ \phi$$

and

$$\check{\phi} : \mathrm{Hom}_R(A, T) \to \mathrm{Hom}_R(B, T),\ f \to \phi \circ f$$

are $\mathbb{Z}$ linear. Suppose that $\psi : B \to C$ is R-linear. Then

$$(\psi \check{\circ} \phi) = \check{\phi} \circ \check{\psi} \text{ and } (\phi \circ \psi)^* = \psi^* \circ \phi^*.$$

Lemma 4.5.11. *Let R be a ring, M a R-module, D a right R-module and A an abelian group.*

(a) $\mathrm{Hom}_{\mathbb{Z}}(D, A)$ *is an R-module via $r\phi(d) = \phi(dr)$.*

(b) The map

$$\Xi = \Xi(M, A) : \mathrm{Hom}_R(M, \mathrm{Hom}_{\mathbb{Z}}(R, A)) \to \mathrm{Hom}_{\mathbb{Z}}(M, A), \quad \Xi(\Phi)(m) = \Phi(m)(1)$$

is an $\mathbb{Z}$-isomorphism.

(c) $\Xi(M,A)$ *depends naturally on* M *and* A. *That is*

(a) Let $\beta : A \to B$ *be* $\mathbb{Z}$*-linear. Then the following diagram is commutative:*

$$\begin{array}{ccc} \mathrm{Hom}_R(M,\mathrm{Hom}_{\mathbb{Z}}(R,A)) & \xrightarrow{\Xi(M,A)} & \mathrm{Hom}_{\mathbb{Z}}(M,A) \\ \downarrow \check{\check{\beta}} & & \downarrow \check{\beta} \\ \mathrm{Hom}_R(M,\mathrm{Hom}_{\mathbb{Z}}(R,B)) & \xrightarrow{\Xi(M,B)} & \mathrm{Hom}_{\mathbb{Z}}(M,B) \end{array}$$

That is $\Xi(\check{\beta} \circ \Phi) = \beta \circ \Xi(\Phi)$ *for all* $\Phi \in \mathrm{Hom}_R(M,\mathrm{Hom}_{\mathbb{Z}}(R,A))$.

(b) Let $\eta : M \to N$ *be* R*-linear. Then the following diagram is commutative:*

$$\begin{array}{ccc} \mathrm{Hom}_R(M,\mathrm{Hom}_{\mathbb{Z}}(R,A)) & \xrightarrow{\Xi(M,A)} & \mathrm{Hom}_{\mathbb{Z}}(M,A) \\ \uparrow \eta^* & & \uparrow \eta^* \\ \mathrm{Hom}_R(N,\mathrm{Hom}_{\mathbb{Z}}(R,A)) & \xrightarrow{\Xi(N,A)} & \mathrm{Hom}_{\mathbb{Z}}(N,A) \end{array}$$

That is $\Xi(\Psi) \circ \eta = \Xi(\Psi \circ \eta)$ *for all* $\Psi \in \mathrm{Hom}_R(N,\mathrm{Hom}_{\mathbb{Z}}(R,A))$.

(d) If A *is divisible,* $\mathrm{Hom}_{\mathbb{Z}}(R,A)$ *is an injective* R*-module.*

Proof. (a) Let $r,s \in R$, $\phi,\psi \in \mathrm{Hom}_{\mathbb{Z}}(D,A)$ and $d,e \in D$.

$$(r\phi)(d+e) = \phi((d+e)r) = \phi(dr+er) = \phi(dr) + \phi(er) = (r\phi)(d) + (r\phi)(e)$$

Thus $r\phi \in \mathrm{Hom}_{\mathbb{Z}}(D,A)$.

$$(r(\phi+\psi))(d) = (\phi+\psi)(dr) = \phi(dr) + \psi(dr) = (r\phi + r\psi)(d)$$

$$((r+s)\phi)(d) = \phi(d(r+s)) = \phi(dr) + \phi(ds) = (r\phi + s\phi)(d)$$

$$((rs)\phi)(d) = \phi(d(rs)) = \phi((dr)s) = (s\phi)(dr) = (r(s\phi))(d)$$

So $\mathrm{Hom}_{\mathbb{Z}}(D,A)$ is indeed a R-module.

(b) Clearly Ξ is $\mathbb{Z}$-linear. Suppose that $\Xi(\Phi) = 0$. Then $\Phi(m)(1) = 0$ for all $m \in M$. Let $r \in R$. Then

$$0 = \Phi(rm)(1) = (r\Phi(m))(1) = (\Phi(m))(1r)) = \Phi(m)(r)$$

Thus $\Phi(m)(r) = 0$ for all r. So $\Phi(m) = 0$ for all m and so $\Phi = 0$. So Ξ is on to one.

To show that Ξ is onto, let $\alpha \in \mathrm{Hom}_{\mathbb{Z}}(M,A)$.

Define $\Phi \in \mathrm{Hom}_R(M,\mathrm{Hom}_{\mathbb{Z}}(R,A))$ by

$$\Phi(m)(r) = \alpha(rm)$$

Clearly $\Phi(m)$ is indeed in $\mathrm{Hom}_{\mathbb{Z}}(R, A)$. We need to verify that Φ is R-linear. Let $s \in R$. Then $(Phi(sm))(r) = \alpha(rsm)$ and $(s\Phi(m))(r) = \Phi(m)(rs) = \alpha(rsm)$. So $\Phi(sm) = s\Phi(m)$ and Φ is R-linear.

Also

$$(\Xi(\Phi)(m) = (\Phi(m))(1) = \alpha(1m) = \alpha(m)$$

and so $\Xi(\Phi) = \alpha$ and Ξ is onto.

(ca)

$$(\check{\beta} \circ \Xi)(\Phi)(m) = \beta(\Xi(\Phi)(m)) = \beta(\Phi(m)(1))$$

and

$$(\Xi \circ \check{\check{\beta}})(\Phi)(m) = \check{\check{\beta}}(\Phi))(m)(1) = (\check{\beta}(\Phi(m))(1) = \beta(\Phi(m)(1)).$$

(cb) Let $\Psi \in \mathrm{Hom}_R(N, \mathrm{Hom}_{\mathbb{Z}}(R, A))$. Then

$$(\eta^* \circ \Xi)(\Psi)(m) = \eta^*(\Xi(\Psi))(m) = \Xi(\Psi)(\eta(m) = \Psi(\eta(m))(1)$$

and

$$(\Xi \circ \eta^*)(\Psi)(m) = \Xi(\eta^*(\Psi))(m) = (\eta^*(\Psi(m))(1) = \Psi(\eta(m))(1).$$

(d) Let I be a left ideal in R and $\beta : I \to \mathrm{Hom}_{\mathbb{Z}}(R, A)$. By 4.5.7 we need to show that β extends to $\gamma : R \to \mathrm{Hom}_{\mathbb{Z}}(R, A)$. Let $\Xi = \Xi(I, A)$ be given by (b). Put $\tilde{\beta} = \Xi(\beta)$. Then

$$\tilde{\beta} : I \to A$$

is $\mathbb{Z}$-linear. Since A is divisible, it is injective as an $\mathbb{Z}$-module. So $\tilde{\beta}$ extends a $\mathbb{Z}$-linear map $\tilde{\gamma} : .R \to A$. That is $\tilde{\beta} = \tilde{\gamma} \circ \rho$, where $\rho : I \to M$ is the inclusion map.. By (b) there exists an R-linear $\gamma : M \to \mathrm{Hom}_{\mathbb{Z}}(R, A)$ with $\Xi(\gamma) = \tilde{\gamma}$. By (cb)

$$\Xi(\gamma \circ \rho) = \Xi(\gamma) \circ \rho = \tilde{\gamma} \circ \rho = \tilde{\beta} = \Xi(\beta)$$

As Ξ is one to one, we conclude $\beta = \gamma \circ \rho$ and so γ is the wanted extension of β. □

Theorem 4.5.12. *Let R be a ring. Every R-module can be embedded into an injective R-module.*

Proof. Let M be a R-module. By 4.5.10 M is a subgroup of some divisible abelian group A. Let $\rho : M \to A$ be the inclusion map. Then $\check{\rho} : \mathrm{Hom}_R(R, M) \to \mathrm{Hom}_{\mathbb{Z}}(R, A)$, $\phi \to \rho \circ \phi$ is a R-momomorphism. By 4.5.8 $M \cong \mathrm{Hom}_R(R, M)$ and so M is isomorphic to an R-submodule of $\mathrm{Hom}_{\mathbb{Z}}(R, A)$. By 4.5.11 $\mathrm{Hom}_{\mathbb{Z}}(R, A)$ is injective. □

Lemma 4.5.13. *(a) Direct summands of injective modules are injective.*

(b) Direct products of injective modules are injective.

Proof. (a) Let $J = J_1 \oplus J_2$ with J injective. Given $\alpha : B \rightarrowtail A$ and $\beta : B \to J_1$. As J is injective there exists $\tilde{\gamma} : A \to J$ with

$$\tilde{\gamma} \circ \alpha = \rho_1 \circ \beta.$$

Put $\gamma = \pi_1 \circ \tilde{\gamma}$ Then

$$\gamma \tilde{\circ} \alpha = \pi_1 \circ \tilde{\gamma} \circ \alpha = \pi_1 \circ \rho_1 \circ \alpha = \alpha.$$

(b) Suppose that $J_i, i \in I$ is a family of injective modules. Given $\alpha : B \to A$ and $\beta : B \to \prod_{i \in I} J_i$. Since J_i is injective there exists $\gamma_i : A \to J_i$ with

$$\gamma_i \circ \alpha = \pi_i \circ \beta$$

Put $\gamma = (\gamma_i)_{i \in I}$.

Then

$$\pi_i \circ \gamma \circ \alpha = \gamma_i \circ \alpha = \pi_i \circ \beta$$

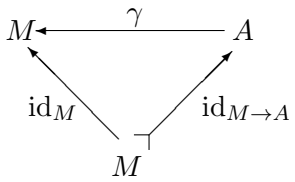
and so $\gamma \circ \alpha = \beta$. Hence $\prod_{i \in I} J_i$ is injective. □

Theorem 4.5.14. *Let M be an R-module. Then the following are equivalent:*

(a) M is injective.

(b) If A is a R-module with $M \leq A$, then A splits over M.

Proof. (a)⇒ (b) Since M is injective we obtain



Hence by 4.4.5, A splits over M.

(b)⇒ (a) By 4.5.12, M is a submodule of an injective module. So by assumption, M is a direct summand of this injective module. Thus by 4.5.13 M is injective. □

4.6 The Functor Hom

If $A \leq B$, $\mathrm{id}_{A \to B}$ denotes the inclusion map $A \to B, a \to a$.

Lemma 4.6.1. *Let R be a ring. Given a sequence $A \xrightarrow{f} B \xrightarrow{g} C$ of R-modules. Then the following two statements are equivalent:*

(a)

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is exact and A splits over $\ker f$.

(b) For all R-modules D,

$$\mathrm{Hom}_R(D,A) \xrightarrow{\check{f}} \mathrm{Hom}_R(D,B) \xrightarrow{\check{g}} \mathrm{Hom}_R(D,C)$$

is exact.

Proof. We first compute $\ker \check{g}$ and $\mathrm{Im}\, \check{f}$. Let $\beta =\in \mathrm{Hom}_R(D,B)$. Then $g \circ \beta = 0$ if and only if $\mathrm{Im}\, \beta \leq \ker g$. Thus

$$\ker \check{g} = \mathrm{Hom}_R(D, \ker g).$$

Also

$$\mathrm{Im}\, \check{f} = f \circ \mathrm{Hom}_R(D,A) := \{f \circ \alpha \mid \alpha \in \mathrm{Hom}_R(D,A)\} \leq \mathrm{Hom}_R(D, \mathrm{Im}\, f).$$

Suppose first that (a) holds. Then $\ker g = \mathrm{Im}\, f$ and $A = \ker f \oplus K$ for some R-submodule K of A. It follows that $f\mid_K : K \to \mathrm{Im}\, f$ is an isomorphisms. Let $\phi \in \mathrm{Hom}_R(D, \mathrm{Im}\, f)$. Let

$$\alpha = \mathrm{id}_{K \to A} \circ (f\mid_K)^{-1} \circ \phi$$

Then $\alpha \in \mathrm{Hom}_R(D,A)$ and $f \circ \alpha = \phi$. So

$$Im\check{f} = \mathrm{Hom}_R(D, \mathrm{Im}\, f) = \mathrm{Hom}_R(D, \ker g) = \ker \check{g}.$$

Suppose next that (b) holds. Let $D = \ker g$. Then

$$\mathrm{id}_{\ker g \to B} \in \ker \check{g} = \mathrm{Im}\, \check{g} \leq \mathrm{Hom}_R(D, \mathrm{Im}\, f)$$

thus $\ker g \leq \mathrm{Im}\, f$. Next choose $D = A$. Then

$$f = f \circ \mathrm{id}_A \in \ker \check{g} = \mathrm{Hom}_R(D, \ker g)$$

Hence $\mathrm{Im}\, f \leq \ker g$ and so $0 \to A \to B \to C \to 0$ is exact.

Finally choose $D = \mathrm{Im}\, f$. Then $\mathrm{id}_{\mathrm{Im}\, f \to B} \in \ker \check{g}$ and so

$$\mathrm{id}_{\mathrm{Im}\, f \to B} = f \circ \gamma$$

for some $\gamma \in \mathrm{Hom}(\mathrm{Im}\, f, A)$. So by 4.4.5, A splits over A. □

Here is the dual version of the previous lemma:

Lemma 4.6.2. *Let R be a ring. Given a sequence $A \xrightarrow{f} B \xrightarrow{g} C$ equivalent.*

(a)

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is exact and C splits over $\mathrm{Im}\, g$.

(b) For all R-modules D,

$$\mathrm{Hom}_R(A, D) \xleftarrow{f^*} \mathrm{Hom}_R(B, D) \xleftarrow{g^*} \mathrm{Hom}_R(C, D)$$

is exact.

Proof. Dual to the proof of4.6.1. We leave the details to the reader. □

The following three theorem are immediate consequences of the previous two:

Theorem 4.6.3. *Let R be ring. Then the following are equivalent*

(a)

$$0 \to A \xrightarrow{f} B \xrightarrow{g} C$$

is exact.

(b) For every R module D,

$$0 \to \mathrm{Hom}(D, A) \xrightarrow{\check{f}} \mathrm{Hom}(D, B) \xrightarrow{\check{g}} \mathrm{Hom}(D, C)$$

is exact.

Proof. □

Theorem 4.6.4. *Let R be ring. Then the following are equivalent*

(a)

$$A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

is exact.

(b) For every R module D,

$$\mathrm{Hom}_R(A, D) \xleftarrow{f^*} \mathrm{Hom}_R(B, A) \xleftarrow{g^*} \mathrm{Hom}_R(C, A) \leftarrow 0$$

is exact.

Proof. □

Theorem 4.6.5. *Let R be a ring. Given a sequence of R-modules* $0 \to A \xrightarrow{f} B \xrightarrow{g} C \to$ *. Then the following three statements are equivalent:*

(a)

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0$$

is split exact.

(b) For all R-modules D,

$$0 \longrightarrow \mathrm{Hom}_R(D, A) \xrightarrow{\check{f}} \mathrm{Hom}_R(D, B) \xrightarrow{\check{g}} \mathrm{Hom}_R(D, C) \longrightarrow o$$

is exact.

(c) For all R-modules D,

$$0 \longleftarrow \mathrm{Hom}_R(A, D) \xleftarrow{f^*} \mathrm{Hom}_R(B, D) \xleftarrow{g^*} \mathrm{Hom}_R(C, D) \longleftarrow 0$$

is exact.

Proof. □

Theorem 4.6.6. *Let A and $B_i, i \in I$ be R-modules. Then*

(a) $\mathrm{Hom}_R(\bigoplus_{i\in I} B_i, A) \cong \prod_{i\in I} \mathrm{Hom}_R(B_i, A)$

(b) $\mathrm{Hom}_R(A, \prod_{i\in I} B_i) \cong \prod_{i\in I} \mathrm{Hom}_R(A, B_i)$

(c) $\mathrm{Hom}_R(A, \bigoplus_{i\in I} B_i) \cong \bigoplus_{i\in I} \mathrm{Hom}_R(A, B_i)$

Proof. Pretty obvious, the details are left to the reader. □

Let R and S be rings. An (R, S)-*bimodule* is abelian group M so that M is a left R-module, a right S module such that

$$(rm)s = r(ms)$$

for all $r \in R$, $s \in S$ and $m \in M$.

For example R is a (R, R) modules if we view R as a left R- and right R-module by multiplication from the left and right, respectively.

Lemma 4.6.7. *Let R and S be rings. Let $\phi : A \to A'$ be R-linear and let B a (R, S)-bimodule. Then*

(a) $\mathrm{Hom}_R(A, B)$ *is a right S-module by*

$$(fs)(a) = f(a)s.$$

(b)

$$\phi^* : \mathrm{Hom}_R(A', B) \to \mathrm{Hom}_R(A, B),\ f \to f \circ \phi$$

is S-linear.

(c) $\mathrm{Hom}_R(B, A)$ *is a left S-module with action of S given by*

$$(sf)(b) = f(bs)$$

(d)

$$\check{\phi} : \mathrm{Hom}_R(B, A) \to \mathrm{Hom}_R(B, A'), f \to \phi f$$

is S linear.

Proof. Straightforward. □

Let R be a ring and M a R-module. The *dual* of M is the module

$$M^* := \mathrm{Hom}_R(M, R)$$

As R is an (R, R)-bimodule, M^* is a right R-module. The elements of M^* are called *linear functionals* on M.

From 4.6.6 we have

$$(\bigoplus_{i \in I} M_i)^* \cong \prod_{i \in I} M_i^*$$

By 4.5.8 $R^* = \cong R$, (but the reader should be aware that here R is a right R-module that is the action is given by right multiplication.)

We conclude

$$F(I)^* \cong R^I$$

and so if I is finite then $F(I)^*$ is isomorphism to the free right-module on I.

An R-module M is called *cyclic* of $M = Rm$ for some $m \in M$.

Lemma 4.6.8. *Let R be a ring and $M = Rm$ a cyclic R modules. Let $I = \mathrm{Ann}_R(m)$ and $J = \{r \in R \mid Ir = 0\}$.*

(a) J is an right ideal in R.

(b)

$$\tau : M^* \to J, \quad f \to f(m)$$

is an isomorphism of right R-modules.

Proof. (a) Let $j \in J$, $r \in R$ and $i \in I$. Then $i(jr) = (ij)r = 0r = 0$ and so $jr \in J$. Thus (a) holds.

(b) Let $a \in \mathrm{Ann}_R(m)$. Then $af(m) = f(am) = f(0) = 0$ and so $f(m) \in J$. So τ is well defined. It is clearly $\mathbb{Z}$-linear and

$$(fr)(m) = f(m)r$$

So $\tau(fr) = \tau(f)r$ and τ is right R-linear.

Let $j \in J$. Then $Ij = 0$ and so the map

$$\xi(j) : M \to R, rm \to rj$$

is well defined and R-linear.

$$\tau(\xi(j) = \xi(j)(m) = \xi(j)(1m) = 1j = j$$

and

$$(\xi(\tau(f)))(rm) = r\tau(f) = rf(m) = f(rm)$$

and so $\xi(\tau(f)) = f$ and τ is a bijection. □

If R is commutative, left and right modules are the same. So we might have that $M \cong M^*$ as R-modules. In this case M is called *self-dual*. For example free modules of finite rang over a commutative ring are self-dual. Let R be a ring, the *double dual* of a module M is $M^{**} := (M^*)^*$. Define

$$\vartheta : M \to M^{**}, \vartheta(m)(f) = f(m).$$

It is readily verified that ϑ is R-linear. If $M = F_R(I)$ is free of finite rang we see that ϑ is an isomorphism. If $M = F_R(I)$ is free of infinite rang, then ϑ is a monomorphism but usually not an isomorphism.

In general ϑ does not need to be one to one. For example if $R = \mathbb{Z}$, $n \in \mathbb{Z}^+$ and $M = \mathbb{X}/n\mathbb{Z}$, then it is easy to see that $M^* = 0$. Indeed let $\phi \in M^*$ and $m \in M$. Then $nm = 0$ and so $n\phi(m) = \phi(nm) = 0$. Thus $\phi(m) = 0$ Since $M^* = 0$, also $M^{**} = 0$.

Let us investigate $\ker \vartheta$ in general. Let $m \in M$ then $\vartheta(m) = 0$ if and only if $\phi(m) = 0$ for all $\phi \in M^*$.

4.7 Tensor products

Let R be a commutative ring and A, B, C R-modules. A function $f : A \times B \to C$ is called *R-bilinear* if for each a in A and b in B the maps

$$f(a, *) : B \to C, y \to f(a, y) \text{ and } f(*, b) : A \to C, x \to f(x, b)$$

are R-linear.

For example the ring multiplication is R-linear. Also if M is any R-module. Then $M^* \times M \to R, (f, m) \to f(m)$

Let R be any ring, A a right and B a left R-module. Let C be any abelian group. A map $f : A \times B \to C$ is called R-balanced, if is $\mathbb{Z}$ bilinear and

$$f(ar, b) = f(a, rb)$$

for all $a \in A, b \in B, r \in R$. $M^* \times M \to R, (f, m) \to f(m)$ is an example of a R-balanced map.

Definition 4.7.1. *Let A be a right and B a left module for the ring R. A tensor product for (A, B) is an R-balanced map:*

$$\otimes : (A \times B) \to A \otimes_R B, (a, b) \to a \otimes b$$

such that for all R-balanced maps $f : A \times B \to C$ there exists a unique $\mathbb{Z}$-linear

$$\bar{f} : A \otimes B \to C \text{ with } f(a, b) = \bar{f}(a \otimes b).$$

Theorem 4.7.2. *Let R be a ring, A be a right and B a left R-module. Then there exits a tensor product for (A, B).*

Proof. Let $A \otimes_R B$ the abelian group with generators $\{x(a,b) \mid a \in A, b \in B$ and relations
$x(a,b) + x(a',b) = x(a+a',b), a, a' \in A, b \in B,$
$x(a,b) + x(a,b') = x(a,b+b'), a \in A, b, b' \in B$
and
$x(ar,b) = x(a,rb), a \in A, b \in B, r \in R$
Write $a \otimes b$ for $x(a,b)$ and define

$$\otimes A \times B \to A \otimes B, (a,b) \to a \otimes b$$

. We leave it as any easy exercise to verify that this is indeed an tensor product. □

Let R be a ring. Then $\otimes : R \times_R R \to R, (r,s) \to rs$ is a tensor product. Indeed given any R-balanced map, $f : R \times R \to C$. Define

$$\bar{f} : R : R \to C, r \to f(r,1)$$

As f is $\mathbb{Z}$-bilinear, $\bar{f}$ is $\mathbb{Z}$-linear. Also

$$\bar{f}(r \otimes s) = \bar{f}(rs) = f(rs,1) = f(r,s)$$

So indeed $\otimes$ is a tensor product. With the same argument we have:

Lemma 4.7.3. *Let R be a ring, A a right and B a left R-module. Then*

$$A \otimes_R R \cong A \quad \text{and} \quad R \otimes_R B \cong B$$

Proof. □

With a little bit more work we will prove

Lemma 4.7.4. *Let R be a ring, J a right ideal in R and I a left ideal in R . Then*

$$\otimes : R/J \times_R R/I \to R/J + I, (r + J, s + I) \to (rs + (I + J)$$

is a tensor product for $(R/J, R/I)$.

Proof. Note here that $I + J$ is neither a left nor a right ideal in R. It is just an additive subgroup, $R/I + J$ is an abelian group but in general not a ring. First we need to check that $\otimes$ is well defined:

$$(r + j)(s + i) + (I + J) = rs + (js + ri + ji) + (I + J) = rs + (I + J)$$

Note here that as J is a right ideal $js + ji \in J$ and as I is a left ideal $ri \in I$.

Clearly $\otimes$ is R-balanced. Suppose now that $f : R/J \times R/I \to C$ is R-balanced.

Define

$$\bar{f} : R/(I+J) \to C, r+(I+J) \to f(r+J, 1+I)$$

Again we first need to verify that this is well-defined.

$$f(r+i+j+J, 1+I) = f((r+J)+(i+J), 1+I) = f(r+I, 1+I) + f((1+J)i, 1+I) =$$

$$= f(r+I, 1+I) + f(1+J, i(1+I)) = f(r+I, 1+I) + f(1+I, 0_{R/I}) = f(r+J, 1+I)$$

So $\bar{f}$ is well defined and clearly $\mathbb{Z}$ linear.

$$\bar{f}(r+J \otimes s+I) = \bar{f}(rs+I+J) = f(rs+J, 1+I) = f((r+J)s, 1+I) = f(r+j, s+I)$$

and $\otimes$ is indeed a tensor product. □

If R is PID we conclude

$$R/Rm \otimes_R R/Rn \cong R/\gcd(n,m)R$$

In particular, if n and m are relative prime $R/Rm \otimes R/Rn = 0$

Let M be a finite dimensional vector space over the division ring $\mathbb{D}$. Let $x \in M, \phi \in M^*$, $R = \mathrm{End}_D(M), I = \mathrm{Ann}_R(x)$ and $J = \mathrm{Ann}_R(y)$. Then $M \cong R/I$ and $M^* = R/J$. Thus $M^* \otimes_R M \cong R/(I+J)$. We leave it as an exercise to verify that $R/I+J \cong D$. We conclude that

$$M^* \times M \to D.(f,m) \to f(m)$$

is a tensor product of (M^*, M).

Lemma 4.7.5. *Let R, S, T be rings, $\alpha : A \to A'$ (R,S)- linear and $\beta : B \to B'$ (S,T)-linear.*

(a) $A \otimes_S B$ is an (R,T) bimodule in such a way that

$$r(a\otimes)b = (ra \otimes bt)$$

for all $r \in R, a \in A, b \in B, s \in S$.

(b) There exists a unique Z-linear map

$$\alpha \otimes \beta : A\otimes_S \to B \to A' \otimes B' \text{ with } a \otimes b \to \alpha(a) \otimes \beta(b)$$

for all $a \in A, b \in B$. Moreover, $\alpha \otimes \beta$ is (R,T)-linear.

Proof. (a) Let $r \in R$, and $t \in Y$. We claim that

$$\phi(r,t) : A \times B \to A \otimes_S B, (a,b) \to ra \otimes bt$$

is S-balanced. Indeed its clearly $\mathbb{Z}$-bilinear and

$$r(as) \otimes bt = (ra)s \otimes bt = ra \otimes s(bt) = ra \otimes (sb)t$$

So its S-balanced. Hence we obtain a map a $\mathbb{Z}$-linear

$$\Phi(r,t) : A \otimes_S B \to A \otimes_S B, \text{ with } a \otimes b \to ra \otimes bs.$$

Let $r, r\prime \in R$ and $t \in T$. It is easy to verify that

$$\Phi(r + r', t)(a \otimes b) = (\Phi(r,t) + \Phi(r',t)(a \otimes b)$$

and

$$\Phi(rr', 1)(a \otimes b) = (\Phi(r,1) \circ \phi(r',1))(a \otimes b)$$

Thus by the uniqueness assertion in the definition of the tensor product,

$$\Phi(r + r', t) = \Phi(r,t) + \Phi(r',t) \text{ and } \Phi(rr',1) = \Phi(r,1) \circ \Phi(r',1).$$

Thus $A \otimes B$ is a left R-module by $rv = \Phi(r,1)(v)$. Similarly$A \otimes B$ is a right T-module by $vt = \Phi(1,t)v$. Also $r((a \otimes b)t = ra \otimes bs = (r(a \otimes b))t$ So $(rv)t = r(vt)$ for all $\in R, t \in T, v \in A \otimes_R B$. Thus (a) holds.

(b) The map

$$A \times B \to A' \otimes_S B', (a,b) \to \alpha(a) \otimes \beta(b)$$

is S-balanced. So $\alpha \otimes \beta$ exists. That its (R,T)-linear is easily verified using arguments as in (a). □

Proposition 4.7.6. *Let D be a right R-module and*

$$A \xrightarrow{f} B \xrightarrow{g} C \to 0$$

an exact sequence of R. Then

$$D \otimes_R A \xrightarrow{\mathrm{id}_D \otimes f} D \otimes_R B \xrightarrow{\mathrm{id}_D \otimes g} D \otimes C \to 0$$

is exact sequence of $\mathbb{Z}$-modules.

Proof. As $D \times C$ is generated by the $d \otimes c$ and g is onto, id)$D \otimes g$ is onto. Also

$$((\mathrm{id}_D \otimes g) \circ (\mathrm{id}_D \otimes f)(d \otimes a) = d \otimes (g(f(a)) = 0$$

So

$$\mathrm{Im}(\mathrm{id}_D \otimes f) \leq \ker g(\mathrm{id}_D \otimes g)$$

.

Let $E = \mathrm{Im}\, f$ and

$$H = \mathrm{Im}\, \mathrm{id}_D \otimes \mathrm{id}_{E \to B} = \langle d \otimes e, d \in D, e \in E \rangle \leq D \otimes_R B$$

Note that $H = \mathrm{Im}(\mathrm{id}_D \otimes f)$. We will show that $H = F := \ker(\mathrm{id}_D \otimes g)$. Without loss $C = B/H$ and g is the canonical epimorphism. We claim that the map

$$D \times B/E \to (D \otimes B)/H, (d, b + E) \to d \otimes b + H$$

is well defined and R-balanced.

Indeed $d \otimes (b+e) + H = (d \otimes b) + (d \otimes e) + H = d \otimes b + H$ So it well defined. Its clearly R-balanced.

Hence we obtain an onto $\mathbb{Z}$-linear map:

$$D \otimes_R B/E \to (D \otimes B)/H, \text{ with } d \otimes (b+E) \to (d \otimes b) + H.$$

$\mathrm{id}_D \otimes g$ induces an isomorphism

$$(D \otimes B)/F \to D \otimes_R B/E, \text{ with } (d \otimes b) + F \to d \otimes (b+E)$$

The composition of these two maps give on onto map

$$\tau : (D \otimes B)/F \to (D \otimes B)/H \text{ with } (d \otimes b) + F \to (d \otimes b) + E.$$

As $D \otimes B$ is generated by the $d \otimes b$ we get $\tau(v+F) = v + E$ for all $v \in D \otimes B$. Since $\tau(0) = 0$ we conclude that $f \in E$ for all $f \in F$. Thus $F \leq E$ and $E = F$. □

Lemma 4.7.7. *(a) Let $(A_i, i \in I)$ be a family of right R modules and $(B_j, j \in J)$ a family of left R-modules. Then*

$$\bigoplus i \in I \otimes_R \bigoplus j \in J \cong \sum_{i \in I} \sum_{j \in J} A_i \otimes B_j$$

(b) Let R be a ring and I, J sets. Then

$$F(I) \otimes F(J) \cong F(I \times J)$$

as a $\mathbb{Z}$-modules.

(c) Let R and S be rings with $R \leq S$. Let I be a set and view S as an (S,R)-bimodule. Then

$$S \otimes_R F_R(I) \cong F_S(I)$$

as S-module.

Proof. (a) is readily verified using the universal properties of direct sums and tensor products.

(b) Since $R \otimes_R R \cong R$, (b) follows from (a).

(c) As $S \otimes_R R \cong S$, (c) follows from (a). □

Lemma 4.7.8. *Let A be a right R module, B a (R,S)-bimodule and C a left S-module. Then there exists $\mathbb{Z}$-linear isomorphism*

$$(A \otimes_R B) \otimes SC \to A \otimes_R (B \otimes SC) \text{ with } (a \otimes b) \otimes c \to a \otimes (b \otimes c)$$

for all $a \in A, b \in B, c \in C$.

Proof. Straightforward form the universal properties of tensor products. □

In future we will just write $A \otimes_R B \otimes_S C$ for any of the two isomorphic tensor products in the previous lemma. A similar lemma holds for more than three factors. $A \otimes_R B \otimes_S C$ can also be characterized through (R, S)-balanced maps from $A \times B \times C \to T$, where T is an abelian group. We leave the details to the interested reader.

Proposition 4.7.9. *Let A be a right R module, B a (R, S)-bimodule and C a right S-module. Then the map:*

$$\Xi : \mathrm{Hom}_S(A \otimes_R B, C) \to\to \mathrm{Hom}_R(A, \mathrm{Hom}_S(B, C), \Xi(f)(a)(b) = f(a \otimes b)$$

is a $\mathbb{Z}$-isomorphism.

Proof. Note that $\Xi(f)(a) : B \to C, b \to f(a, b)$ is indeed S-linear. Also $\Xi(f) : A \to \mathrm{Hom}_S(B, C)$ is R-linear and Ξ is $\mathbb{Z}$-linear. It remains to show that Ξ is a bijection. We do this by defining an inverse. Let $\alpha : A \to \mathrm{Hom}_S(B, C)$ be R-linear. We claim that the map

$$A \times B \to C; (a, b) \to \alpha(a)(b)$$

is R balanced. Indeed it is $\mathbb{Z}$-bilinear and

$$\alpha(ar)(b) = (\alpha(a)r)(b) = \alpha(a)(rb).$$

So there exist

$$\Theta(\alpha) : A \otimes B \to C \text{ with } \Theta(\alpha)(a \times b) = \alpha(a)(b)$$

for all $a \in A$, $b \in B$. It is readily verified that $\Theta(\alpha)$ is S-linear. So

$$\Theta : \mathrm{Hom}_R(A, \mathrm{Hom}_S(B, C) \to \mathrm{Hom}_S(A \otimes_R B, C)$$

We claim that Θ and Ξ are inverses:

$$\Xi(\Theta(\alpha))(a)(b) = \Theta(\alpha)(a \otimes b) = \alpha(a)(b)$$

So $\Xi(\Theta(\alpha)) = \alpha$.

$$\Theta(\Xi(f))(a \otimes b) = \Xi(f)(a)(b) = f(a \otimes b)$$

and so $\Theta(\Xi(f)) = f$. □

Here is a special case of the previous proposition. Suppose R is a commutative ring and A and B, R-modules. Applying 4.7.9 with $C = R$. We get that

$$(A \otimes B)^* \cong \mathrm{Hom}_R(A, B^*)$$

Suppose that R is commutative and A, B are R-modules we obtain a $\mathbb{Z}$-linear map:

$$\sigma : A^* \otimes B^* \to (A \otimes B)^* \text{ with } \gamma(\alpha \otimes \beta))(a \otimes b) = \alpha(a)\beta(b).$$

Indeed this follows from $\alpha \otimes \beta : A \otimes B \to R \otimes R \cong R$.

If A and B are free of finite rang it is easy to see that this is an isomorphism. If A and B are free, σ is still one to one, but not necessarily onto. Suppose that $A_k = R/I_k$, $k \in \{1, 2\}$, is a cyclic R-module. Put $J_k = \mathrm{Ann}_R(I_k)$. Then by 4.6.8, $A_k^* \cong J_k$. Also $A_1 \otimes A_2 \cong R/(I_1 + I_2)$. Now $\mathrm{Ann}_R(I_1 + I_2) = Ann_R(I_1) \cap \mathrm{Ann}_R(I_2) = J_1 \cap J_2$. Thus $(A_1 \otimes_R A_2)^* \cong J_1 \cap J_2$. σ from above (with $A = A_1, B = A_2$) now reads:

$$\sigma : J_1 \otimes_R J_2 \to J_1 \cap J_2 \quad (j_1, j_2) \to j_1 j_2$$

We will know give an example where $\sigma = 0$ but $J_1 \otimes J_2 \neq 0$. Let S be a ring and M an (S, S)-bimodule. Define $M \rtimes S$ to be the ring with additive group $M \oplus S$ and multiplication

$$(m_1, s_1) \cdot (m_2, s_2) = (m_1 s_2 + s_1 m_2, s_1 s_2)$$

It is easy to verify that $M \times S$ is a ring. As an example we verify that the multiplication is associative

$$((m_1, s_1)\cdot(m_2, s_2))\cdot(m_3, s_3) = (m_1 s_2 + s_1 m_2, s_1 s_2)\cdot(m_3, s_3) = (m_1 s_2 s_3 + s_1 m_2 s_3 + s_2 s_2 m_3, s_1 s_2 s_3)$$

A similar calculation shows that the right side is also equal to $(m_1, s_1) \cdot ((m_2, s_2) \cdot (m_3, s_3))$. Identify $(m, 0)$ with m and $(0, s)$ with s. Then

$$M \rtimes S = M + S, s_1 \cdot s_2 = s_1 s_2, s \cdot m = sm, m \cdot s = ms \text{ and } m_1 \cdot m_2 = 0$$

for all $s, s_1, s_2 \in S$ and $m, m_1, m_2 \in m$. Also M is an ideal in $M \rtimes S$ and $(M \rtimes S)/M \cong S$. Indeed the map $M \rtimes S \to S, (m, s) \to s$ is an onto ring homomorphism with kernel M.

Suppose now that S is commutative and M a faithful S module. Put $R = M \rtimes S$. Then R is commutative. As $M^2 = 0$ and $\mathrm{Ann}_S(M) = 0$, $\mathrm{Ann}_R(M) = M$. Also $M \cap M = M = M + M$. We conclude $(R/M)^* \cong M$, $R/M \otimes R/M \cong R/M$, $(R/M \otimes R/M)^* \cong M$ and

$$\sigma : M \otimes_R M \to M, (m_1, m_2) \to m_1 m_2 = 0$$

Suppose that $M = F_S(I)$ is a free S-module. Then as an R-module,

$$M \cong \bigoplus_{i \in I} R/M$$

Thus

$$M \otimes_R M \cong \bigoplus_{i \in I} \bigoplus_{j \in I} R/M$$

and so $M \otimes_R M \neq 0$ (unless $I \neq \emptyset$).

4.8 Composition series

Definition 4.8.1. *Let R be a ring, M an R-module and $\mathcal{C}$ a set of R-submodules in R. We say that $\mathcal{C}$ is a* R-series *on M provided that*

(a) $\mathcal{C}$ is a chain, that is for any $A, B \in \mathcal{C}$, $A \leq B$ or $B \leq A$.

(b) $0 \in \mathcal{C}$ and $M \in \mathcal{C}$.

(c) $\mathcal{C}$ is closed under unions and intersections, that is if $\mathcal{D} \subseteq \mathcal{C}$, then

$$\bigcup \mathcal{D} \in \mathcal{C} \text{ and } \bigcap \mathcal{D} \in \mathcal{C}.$$

For example any finite chain

$$0 = M_0 < M_1 < M_2 < M_3 < \ldots < M_{n-1} < M_n = M$$

of R-submodules of M is an R-series.

If $R = M = \mathbb{Z}$ and p is a prime then

$$0 < \ldots < p^{k+1}\mathbb{Z} < p^k\mathbb{Z} < p^{k-1}\mathbb{Z} < \ldots < p\mathbb{Z} < \mathbb{Z}$$

is a $\mathbb{Z}$-series. More generally, if $n_1, n_2, n_3, \ldots$ is any sequence of integers larger than 1, then

$$0 < n_1 \ldots n_{k+1}\mathbb{Z} < n_1 \ldots n_k\mathbb{Z} < \ldots < n_1 n_2\mathbb{Z} < n_1\mathbb{Z} < \mathbb{Z}$$

is a $\mathbb{Z}$ series on $\mathbb{Z}$.

Definition 4.8.2. *Let R be a ring, M an R-module and $\mathcal{C}$ an R-series on M.*

(a) A jump *in $\mathcal{C}$ is a pair (A, B) with $A, B \in \mathcal{C}$, $A \lneq B$ and so so that*

$$D \leq A \text{ or } B \leq D \text{ for all } D \in \mathcal{C}.$$

Jump($\mathcal{C}$) *is the set of all jumps of $\mathcal{C}$.*

(b) If (A, B) is a jump of $\mathcal{C}$ then B/A is called a factor *of $\mathcal{C}$.*

(c) $\mathcal{C}$ is a R-composition series *on M provided that all the factors of $\mathcal{C}$ are simple R-modules.*

Let $\mathcal{C}$ be R-series on M. For $B \in \mathcal{C}$ define

$$B^- = \bigcup \{A \in \mathcal{C} \mid A \lneq B\}.$$

Note that $B^- \in \mathcal{C}$ and $B^- \leq B$.

Suppose that $B^- \neq B$. Let $D \in \mathcal{C}$. Then $B \leq D$ or $D \lneq B$. In the latter case, $D \leq B^-$ and so (B^-, B) is a jump of $\mathcal{C}$.

Conversely, if (A, B) is a jump it is easy to see that $A = B^-$. Thus

$$\mathrm{Jump}(\mathcal{C}) = \{(B^-, B) \mid B \in \mathcal{C}, B^- \neq B\}.$$

Consider the series

$$0 < n_1 \dots n_{k+1}\mathbb{Z} < n_1 \dots n_k\mathbb{Z} < \dots < n_1 n_2 \mathbb{Z} < n_1\mathbb{Z} < \mathbb{Z}.$$

As $n_1 \dots n_{k+1})\mathbb{Z}/n_1 \dots n_k\mathbb{Z} \cong \mathbb{Z}/n_k\mathbb{Z}$ as R-modules, this series is a composition series if and only if each n_k is a prime. If we chose $n_k = p$ for a fixed prime p we get a composition series all of whose factors are isomorphic. On the other hand we could choose the n_k to be pairwise distinct primes and obtain a composition series so that now two factors are isomorphic.

Proposition 4.8.3. *Let R be a ring and M a R-module. Let $\mathcal{M}$ be the set of chains of R-submodules in M. Order $\mathcal{M}$ by inclusion and let $\mathcal{C} \in \mathcal{M}$. Then $\mathcal{C}$ is a composition series if and only if $\mathcal{C}$ is a maximal element in $\mathcal{M}$.*

Proof. $\Longrightarrow$ Suppose that $\mathcal{C}$ is a composition series but is not maximal in $\mathcal{M}$. Then $\mathcal{C} \subsetneq \mathcal{D}$ for some $\mathcal{D} \in \mathcal{M}$. Hence there exists $D \in \mathcal{D} \setminus \mathcal{C}$. We will show that there exists a jump of $\mathcal{C}$ so that the corresponding factor is not simple, contradicting the assumption that $\mathcal{C}$ is a composition series. Define

$$D^+ = \bigcap\{E \in \mathcal{C} \mid D \leq E\} \text{ and } D^- = \bigcup\{E \in \mathcal{C} \mid E \leq D\}.$$

As $\mathcal{C}$ is closed under unions and intersections both D^+ and D^- are members of $\mathcal{C}$. In particular, $D^- \neq D \neq D^+$. From the definition of D^+, $D \leq D^+$, also $D^- \leq D$ and so

$$D^- \lneq D \lneq D^+.$$

Thus D/D^+ is a proper R-submodule of D^+/D^- and it remains to verify that (D^-, D^+) is a jump. For this let $E \in \mathcal{C}$. As $\mathcal{D}$ is totally ordered, $E \leq D$ or $D \leq E$. In the first case $E \leq D^-$ and in the second $D^+ \leq E$.

$\Longleftarrow$ Let $\mathcal{C}$ be a maximal element of $\mathcal{M}$. We will first show that

(*) Let E be an R-submodule of G such that for all $C \in \mathcal{C}$, $E \leq C$ or $C \leq E$. Then $E \in \mathcal{C}$.

Indeed, under these assumptions, $\{E\} \cup \mathcal{C}$ is a chain of submodules and so the maximality of $\mathcal{C}$ implies $E \in \mathcal{C}$.

From (*) we conclude $0 \in \mathcal{C}$ and $M \in \mathcal{C}$. Let $\mathcal{D} \subseteq \mathcal{C}$ and put $E = \bigcup \mathcal{D}$. We claim that E fulfills the assumptions of (*). For this let $C \in \mathcal{C}$. If $C \leq D$ for some $D \in \mathcal{D}$ then $C \leq D \leq E$. So suppose that $C \not\leq D$ for each $D \in \mathcal{D}$. As $\mathcal{C}$ is totally ordered, $D \leq C$ for each $D \in \mathcal{D}$. Thus $E \leq D$. So we can apply (*) and $E \in \mathcal{C}$. Thus $\mathcal{C}$ is closed under unions.

Similarly, $\mathcal{C}$ is closed under intersections. Thus $\mathcal{C}$ is a series and it remains to show that all its factors are simple. So suppose that (A, B) is a jump of $\mathcal{C}$ so that B/A is not

simple. Then there exists a proper R-submodule $\bar{E}$ of B/A. Note that $\bar{E} = E/A$ for some R-submodule E of M with

$$A \lneq E \lneq B.$$

As (A, B) is a jump, $E \notin \mathcal{C}$. Let $C \in \mathcal{C}$. Then $C \leq A$ or $B \leq C$. So $C \leq E$ or $E \leq C$. Thus by (*), $E \in \mathcal{C}$, a contradiction □

Corollary 4.8.4. *Every R-modules has a composition series.*

Proof. Let $\mathcal{M}$ be as in 4.8.3. We leave it as an routine application of Zorn's Lemma A.6 to show that $\mathcal{M}$ has a maximal element. By 4.8.3 any such maximal element is a composition series. □

In the next lemma we will find series for direct sums and direct products of modules. For this we first need to introduce the concept of cuts for a totally ordered set $(I, \leq)$.

We say that $J \subseteq I$ is a *cut* of I if for all $j \in J$ and all $i \in I$ with $i \leq j$ we have $i \in J$. Let $\mathrm{Cut}(I)$ be the set of all cuts of I. Note that $\emptyset \in \mathrm{Cut}(I)$ and $I \in \mathrm{Cut}(I)$. Order $\mathrm{Cut}(I)$ by inclusion. We claim that $\mathrm{Cut}(I)$ is totally ordered. Indeed, let $J, K \in \mathrm{Cut}(I)$ with $K \not\subseteq J$. Then there exists $k \in K \setminus J$. Let $j \in J$. Since $k \notin J$ and J is a cut, $k \not\leq j$. Since I is totally ordered, $j < k$ and since K is a cut, $j \in K$. So $J \subseteq K$ and $\mathrm{Cut}(I)$ is totally ordered.

Let $i \in I$ and put $i^+ = \{j \in I \mid j \leq i\}$. Note that i^+ is a cut of I. The map $I \to \mathrm{Cut}(I)$, $i \to i^+$ is an embedding of totally ordered sets. Put $i^- = \{j \in I \mid j < i\}$. Then also i^- is a cut.

We leave it as an exercise to verify that unions and intersection of arbitrary sets of cuts are cuts.

As an example consider the case $I = \mathbb{Q}$ ordered in the usual way. Let $r \in \mathbb{R}$ and define $r^- = \{q \in \mathbb{Q} \mid q < r\}$. Clearly r^- is a cut. We claim that every cut of $\mathbb{Q}$ is exactly one of the following cuts:

$$\emptyset; \quad \mathbb{Q}; \quad q^+ \, (q \in \mathbb{Q}); \quad r^- \, (r \in \mathbb{R})$$

Indeed, let be J be a non-empty cut of $\mathbb{Q}$. If J has no upper bound in $\mathbb{Q}$, then $J = \mathbb{Q}$. So suppose that J has an upper bound. By a property of the real numbers, every bounded non-empty subset of $\mathbb{R}$ has a least upper bound. Hence J has a least upper bound a. Then $J \subseteq r^+$.

If $r \in J$, then $r \in \mathbb{Q}$ and $r^+ \subseteq J \subseteq r^+$. So $J = r^+$.

If $r \notin J$ we have $J \subseteq r^-$. We claim that equality holds. Indeed let $q \in r^-$. As r is a least upper bound for J, q is not an upper bound for J and so $q < j$ for some $j \in J$. Thus $q \in J$ and $J = r^-$.

Lemma 4.8.5. *Let $(I, \leq)$ be a totally ordered set and R a ring. For $i \in I$ let M_i be a non zero R-module. Let $M \in \{\bigoplus i \in I M_i, \prod_{i \in I} M_i$. For J a cut of I define*

$$M_J^+ = \{m \in M \mid m_i = 0 \, \forall i \in I \setminus J\}$$

and if $J \neq \emptyset$,

$$M_J^- = \{m \in M \mid \exists j \in J \text{ with } m_i = 0 \, \forall i \geq j\}.$$

Put $M_\emptyset^- = 0$.

(a) For all $k \in I$, $M_{k^+}^- = M_{k^-}^+$ *and* $M_{k^+}^+ / M_{k^-}^+ \cong M_k$.

(b) Let $M = \bigoplus_{i \in I} M_i$. *Then*

(a) $\mathcal{C} := \{M_J^+ \mid J \in J \in \mathrm{Cut}(I)\}$ *is an R-series on* M.

(b) $\mathrm{Jump}(\mathcal{C}) = \{(M_{k^-}^+, M_{k^+}^+) \mid k \in I\}$.

(c) $\mathcal{C}$ *an R-composition series if and only if each* $M_k, k \in I$ *is a simple R-module.*

(c) Let $M = \prod_{i \in I} M_i$ *Then*

(a) $\mathcal{C} := \{M_J^+, M_J^- \mid J \in J \in \mathrm{Cut}(I)\}$ *is an R-series on* M.

(b) $\mathrm{Jump}(\mathcal{C}) := \{(M_J^-, M_J^+) \mid \emptyset \neq J \in \mathrm{Cut}(I)\}$.

(c) $\mathcal{C}$ *is an R-composition series if and only if each non-empty subset ofI has a maximal element and each* $M_k, k \in I$ *is a simple R-module.*

Proof. (a) The first statement follows directly from the definitions. For the second note that the map $M_{k^+} \to M_k, m \to m_k$ is onto with kernel M_{k^-}.

(b) & (c) Note that $M_J^- \leq M_J^+$.

Let $\mathrm{Cut}^*(I)$ be the set of cuts without a maximal element. So

$$\mathrm{Cut}(I) = \{k^+ \mid k \in K\} \cup \mathrm{Cut}^*(I).$$

Let $J \in \mathrm{Cut}^*(I)$. We claim that $M_J^- = M_J^+$ if $M = \bigoplus_{i \in I} M_i$ and $M_J^- \neq M_J^+$ if $M = \prod_{i \in I} M_i$.

So suppose first that $M = \bigoplus_{i \in I} M_i$ and let $0 \neq m \in M_J^+$ and pick $k \in J$ maximal with $m_k \neq 0$ (this is possible as only finitely many m_i's are not 0). Since J has no maximal element there exists $j \in J$ with $k < j$. Then $m_i = 0$ for all $i \geq j$ and so $m \in M_J^-$.

Suppose next that $M = \prod_{i \in I} M_i$. For $j \in J$ pick $0 \neq m_j \in M_j$. For $i \in I \setminus J$ let $m_i = 0$. Then $(m_i) \in M_J^+$ but $(m_i) \notin M_J^-$.

From the claim we conclude that in both cases

$$\mathcal{C} := \{M_J^+, M_J^- \mid J \in \mathrm{Cut}(I)\}$$

We will show now that $\mathcal{C}$ is a chain. For this let J and K be distinct cuts. Since $\mathrm{Cut}(I)$ is totally ordered we may assume $J \subset K$. Then

$$M_J^- \leq M_J^+ \leq M_K^- \leq M_K^+.$$

and so $\mathcal{C}$ is totally ordered.

Also $0 = M_\emptyset^+$ and $M = M_I^+$.

Let $\mathcal{D}$ be a subset of $\mathcal{C}$. We need to show that both $\bigcap \mathcal{D}$ and $\bigcup \mathcal{D}$ are in $\mathcal{D}$. Let $D \in \mathcal{D}$. Then $D = M_{J_D}^{\epsilon_D}$ for some $J_D \in \mathrm{Cut}(I)$ and $\epsilon_D \in \{\pm\}$.

Put $J = \bigcap_{D\in\mathcal{D}} J_D$. Suppose first that $M_J^- \in \mathcal{D}$.

Then $M_J^- \subseteq D$ for all $D \in \mathcal{D}$ and

$$\bigcap \mathcal{D} = M_J^-.$$

So suppose that $M_J^- \notin \mathcal{D}$. Then $M_J^+ \leq D$ for all $D \in \mathcal{D}$ and so $M_J^+ \subseteq \bigcap \mathcal{D}$. We claim that

$$\bigcap \mathcal{D} = M_J^+.$$

Indeed, let $m \in \bigcap \mathcal{D}$ and $i \in I \setminus J$. Then $i \notin J_D$ for some $D \in \mathcal{D}$. As

$$m \in D = M_{J_D}^{\epsilon_D} \leq M_{J_D}^+$$

we get $m_i = 0$. Thus $m \in M_J^+$, proving the claim.

So $\mathcal{C}$ is closed under arbitrary unions.

Let $K = \bigcup\{J_D \mid D \in \mathcal{D}\}$.

Suppose that $M_K^+ \in \mathcal{D}$. Then $M \subseteq M_K^+$ for all $D \in \mathcal{D}$ and

$$\bigcup \mathcal{D} = M_K^+.$$

So suppose that $M_K^+ \notin \mathcal{D}$. Then $\bigcup \mathcal{D} \subseteq M_K^-$. We claim that

$$\bigcup \mathcal{D} = M_K^-.$$

If $K = \emptyset$ each J_D is the empty set. So we may assume $K \neq \emptyset$. Let $m \in M_K^-$. Then by definition there exists $k \in K$ with $m_i = 0$ for all $i \geq k$. Pick $D \in \mathcal{D}$ with $i \in J_D$. Then

$$m \in M_{J_D}^- \leq M_{J_D}^{\epsilon_D} = D \leq \bigcup \mathcal{D}.$$

So the claim is true and $\mathcal{C}$ is closed under unions.

Hence $\mathcal{C}$ is an R-series on M.

Next we investigate the jumps of $\mathcal{C}$. As seen above every cut is of the form (B^-, B) for some $B = M_J^\epsilon \in \mathcal{C}$ with $B \neq B^-$.

Suppose first that $J = k^+$ for $k \in I$. As $M_{k^+}^- = M_{k^-}^+$ we may and do assume $\epsilon = +$. Thus $M_{k^+}^- = M_{k^-}^+ = (M_{k^+}^+)^-$ and $M_{k^-}^+, M_{k^+}^+)$ is a jump with factor isomorphic to M_k.

Suppose next that $J \in \mathrm{Cut}^*(I)$. Then $M_J^- = \bigcup_{j\in J} M_{j^+} \leq (M_J^-)^-$. We conclude that $(M_J^+)^- = (M_J^-)_= M_J^-$. If $M = \bigoplus_{i\in I} M_i$ then as seen above $M_J^- = M_J^+$. So we only get a jump if $\epsilon = +$ and $M = M = \prod_{i\in I} M_i$.

The factor M_J^+/M_J^- can be describes as follows. Identify M_J^+ with $\prod_{j\in J} M_j$. Define $x, y \in \prod_{j\in J} M_j$ to be equivalent if and only if there exists $j \in J$ with $x_i = y_i$ for all $i \in J$ with $j \leq i$. It is easy to check that this is an equivalence relation, indeed x and y are

equivalent if and only if $y - x \in M_J^-$. In particular, M_J^+/M_J^- is the set of equivalence classes. We claim that M_J^+/M_J^- is never a simple module. For this let $J = J_1 \cup J_2$ with $J_1 \cap J_2 = \emptyset$ so that for each $j_1 \in J_1$ there exists $j_2 \in J_2$ with $j_1 < j_2$, and vice versa. (We leave the existence of J_1 and J_2 as an exercise). Then M_J^+/M_J^- is the direct sum of the images of $\prod_{j\in J_i} M_j$ in M_J^+/M_J^-.

Finally we claim that every non-empty subset of I has a maximal element if and only if every non-empty cut of I has a maximal element. One direction is obvious. For the other let J be a non-empty subset of I and define $J^* = \{i \in I \mid i \leq j \text{ for some } j \in J\}$. Clearly J^* is a cut and $J \subseteq J^*$. Suppose J^* has a maximal element k. Then $k \leq j$ for some $j \in J$. As $j \in J^*$ we conclude $j \leq k$ and so $j = k$ and k is the maximal element of J.

It is now easy to see that (bc) and(cc) hold and all parts of the lemma are proved. □

Corollary 4.8.6. *Let R be a ring and I a set. Let M be one of $F_R(I)$ and R^I. Then there exists an R-series $\mathcal{C}$ of on M so that all factors of $\mathcal{C}$ are isomorphic to R and $|\text{Jump}(\mathcal{C})| = |I|$. Moreover, if R is a division ring $\mathcal{C}$ is a composition series.*

Proof. By the well-ordering principalA.10 there exists a well ordering $\leq^*$ be a well ordering on I. Define a partial order $\leq$ on I by $i \leq j$ if and only if $j \leq^* i$. Then every non-empty subset of I has a maximal element and all non empty cuts of I are of the form k^+, $k \in K$. The result now follows from 4.8.5 □

As an example let $R = \mathbb{Q}$. If $I = \mathbb{Q}$ we see that the countable vector space $F_{\mathbb{Q}}(\mathbb{Q})$ as an uncountable composition series. But note that the number of jumps is countable. If $I = \mathbb{Z}^-$ we conclude that uncountable vector space $\mathbb{Q}^{\mathbb{Z}^-}$ as a countable composition series. So the number of jumps in a composition series can be smaller than the dimensions of the vector space. But the next proposition shows that the number of jumps never exceeds the dimension.

Proposition 4.8.7. *Let $\mathbb{D}$ be a division ring and V a vector space over $\mathbb{D}$. Let $\mathcal{C}$ be a $\mathbb{D}$ series on V, and $\mathcal{B}$ a $\mathbb{D}$-basis for V. Then*

$$|\text{Jump}\mathcal{C}| \leq \mathcal{B}.$$

In particular, any two basis for V have the same cardinality.

Proof. Choose some well ordering on $\mathcal{B}$. Let $0 \neq v \in V$. Then $v = \sum_{b\in\mathcal{B}} d_b(v)b$ with $d_b(v) \in \mathbb{D}$, where almost all $d_b(v), b \in \mathcal{B}$ are zero. So we can choose $h(v) \in \mathcal{B}$ maximal with respect to $d_{h(v)}(v) \neq 0$.

Define a map

$$\phi : \text{Jump}(\mathcal{C}) \to \mathcal{B}$$

$$(A, B) \to \min\{h(v) \mid v \in A \setminus B\}$$

We claim that ϕ is one to one. Indeed suppose that (A, B) and (E, F) are distinct jumps with $b = \phi((A, B)) = \phi((E, F))$. As $\mathcal{C}$ is totally ordered and (A, B) and (E, F) are jumps we may assume $A \leq B \leq E \leq F$. Let $v \in B \setminus A$ with $h(v) = b$ and $d_b(v) = 1$. Let $w \in F \setminus E$

with $h(w) = b$ and $d_b(w) = 1$. Since $v \in A \in E$, $w - v \in F \setminus E$. Also $d_b(w - v) = 1 - 1 = 0$ and so $h(w - v) < b$ a contradiction to $b = \phi(E, F)$.

So ϕ is one to one and $|\text{Jump}(\mathcal{C})| \leq |\mathcal{B}|$.

The second statement follows from the first and 4.8.6. □

Lemma 4.8.8. *Let $\mathcal{C}$ be a series for R on M.*

(a) Let $0 \neq m \in M$. Then there exists a unique jump (A, B) of $\mathcal{C}$ with $m \in B$ and $m \notin A$.

(b) Let $D, E \in \mathcal{C}$ with $D < E$. Then there exists a jump (A, B) in $\mathcal{C}$ with

$$D \leq A < B \leq E$$

Proof. (a) Let $B = \bigcap\{C \in \mathcal{C} \mid m \in C\}$ and $A = \bigcup\{C \in \mathcal{C} \mid m \notin C\}$.

(b) Let $m \in E \setminus D$ and let (A, B) be as in (a). □

The following lemma shows how a series can be reconstructed from its jumps.

Lemma 4.8.9. *Let R be a ring, M an R-module and $\mathcal{C}$ an R-series on M. Let $\hat{\mathcal{C}} = \{C \in \mathcal{C} \mid C \neq C^-$. Then the map*

$$\alpha : \text{Cut}(\hat{\mathcal{C}}) \to \mathcal{C},\ K \to \bigcup K$$

is a bijection.

Proof. Note first that as $\mathcal{C}$ is closed under unions $\alpha(K)$ is indeed in $\mathcal{C}$. We will show that the inverse of α is

$$\beta : \mathcal{C} \to \text{Cut}(\hat{\mathcal{C}}),\ D \to \{A \in \hat{\mathcal{C}} \mid A \leq D\}.$$

It is easy to verify that $\beta(D)$ is a cut.

Clearly, $K \subseteq \beta(\alpha(K))$. Let $E \in \hat{\mathcal{C}}$ with $E \notin K$. Then as K is a cut, $A < E$ for all $A \in K$. But then $A \leq E^-$ and so $\alpha(K) \leq E^- < E$. Thus $E \not\leq \alpha(K)$ and $E \notin \beta(\alpha(K))$. Hence $\beta(\alpha(K)) = K$.

Clearly $\alpha(\beta(D) \leq D$. Suppose that $\alpha(\beta(D)) < D$. Then by 4.8.8b there exists a jump (A, B) of $\mathcal{C}$ with $\alpha(\beta(D)) \leq A < B \leq D$. But then $B \in \beta(D)$ and so $B \leq \alpha(\beta(D))$, a contradiction. □

Lemma 4.8.10. *Let $\mathcal{C}$ be a series for R on M and W an R-submodule in M. Then*

(a)

$$\mathcal{C} \cap W := \{D \cap W \mid D \in \mathcal{C}\}$$

is an R-series on M.

(b) Let

$$\mathrm{Jump}^W(\mathcal{C}) = \{(A, B) \in \mathrm{Jump}(\mathcal{C}) \mid A \cap W \neq B \cap W\}.$$

Then the map

$$\mathrm{Jump}^W(\mathcal{C}) \to \mathrm{Jump}(\mathcal{C}) \cap W, \quad (A, B) \to (A \cap W, B \cap W)$$

is a bijection. Moreover,

$$B \cap W/A \cap W \cong (B \cap W)+A/A \leq B/A$$

(c) If $\mathcal{C}$ is a R-composition series on M then $\mathcal{C} \cap W$ is a R-composition series on W. Moreover, there exists an embedding $\phi : \mathrm{Jump}(\mathcal{C} \cap W) \to \mathrm{Jump}(\mathcal{C})$, so that so corresponding factors are R-isomorphic. The image of ϕ consists of all the jumps (A, B) of $\mathcal{C}$ with $B = A + (B \cap W)$.

Proof. (a) Clearly $\mathcal{C} \cap W$ is a chain of R-submodules in W. Also $0 = 0 \cap W \in \mathcal{C} \cap W$, $W = M \cap W \in \mathcal{C} \cap W$ and its is easy to verify that $\mathcal{M} \cap W$ is closed under unions and intersections.

(b) Let $(A, B) \in \mathrm{Jump}^W(\mathcal{C})$. We will first verify that $(A \cap W, B \cap W)$ is a jump of $\mathcal{C} \cap W$. Let $D \in \mathcal{C} \cap W$. Then $D = E \cap W$ for some $E \in \mathcal{C}$. As (A, B) is a jump, $E \leq A$ or $B \leq E$. Thus $D = E \cap W \leq A \cap W$ or $B \cap W \leq E \cap W = D$. To show that the map is bijective we will construct its inverse. For $D \in \mathcal{C} \cap W$ define

$$D^- = \bigcup \{C \in \mathcal{C} \mid C \cap W \leq D\} \text{ and } D^+ = \bigcap \{C \in \mathcal{C} \mid D \leq C \cap W\}.$$

Then it easy to verify that $D^+ \cap W = D = D^- \cap W$. Let (D, E) be a jump in $\mathcal{C} \cap W$. Let $C \in \mathcal{C}$. Since (D, E) is a jump in $\mathcal{C} \cap W$, $C \cap W \leq D$ or $E \leq C \cap W$. In the first case $C \leq D^+$ and in the second $E^- \leq C$. So (D^+, E^-) is a jump of $\mathcal{C}$. It is readily verified that maps $(D, E) \to (D^+, E^-)$ is inverse to the map $(A, B) \to (A \cap W, B \cap W)$.

The last statement in (b) follows from

$$B \cap W/A \cap W = (B \cap W)/(B \cap W) \cap A \cong (B \cap W)+A)/A.$$

(c) Note that $A \cap W \neq B \cap W$ if and only if $(B \cap W)+A/A \neq 0$. Since $\mathcal{C}$ is a composition series, B/A is simple. Thus $(B \cap W)+A/A \neq 0$ if and only if $B = (B \cap W) + A$. Thus by (b) all factors of $\mathcal{C} \cap W$ are simple and $\mathcal{C} \cap W$ is a R-composition series on W. □

Theorem 4.8.11 (Jordan-Hölder)**.** *Let R be a ring and M a module. Suppose R has a finite composition series $\mathcal{C}$ on M and that $\mathcal{D}$ is any composition series for R on M. Then $\mathcal{D}$ is finite and there exists a bijection between the set of factors of $\mathcal{C}$ and the set of factors of $\mathcal{D}$ sending a factor of $\mathcal{C}$ to an R-isomorphic factor of $\mathcal{D}$.*

Proof. Let W be the maximal element of $\mathcal{D} - M$. Then $\mathcal{D} - M$ and (by 4.8.10 $\mathcal{C} \cap W$ are composition series for W. By induction on $|\mathcal{D}|$, $\mathcal{D} \cap W$ is finite and has the same factors as $\mathcal{D} - M$.

For $E \in \mathcal{C} \cap W$ define E^+ and E^- as in 4.8.10. Let $calE = \{E^+, E^- \mid E \in \mathcal{D} \cap W$. Then $\mathcal{E}$ is a finite series on M. Since $W^+ = M \not\leq W$ we can choose $L \in \mathcal{E}$ minimal with respect to $L \not\leq W$. Then $L = E^\epsilon$ for some $E \in \mathcal{C} \cap W$ and $\epsilon \in \{\pm\}$. Suppose first that $L = E^-$. Since $0^- = 0 \leq W$, $E \neq 0$ and so there exists $F \in \mathcal{C} \cap W$ such that (F, E) is a jump in $\mathcal{C} \cap W$. But then $(F^+, E^-) \in \mathrm{Jump}^W(\mathcal{C})$, $F^+ \leq W$ and by 4.8.10c, $E^- = F^+ + (E^- \cap W) \leq W$ a contradiction. So $E^+ = L \neq E^-$. By 4.8.8b there exists a jump (A, B) of $\mathcal{C}$ with $E^- \leq A < B \leq E^+$. Then $E = E^- \cap W \leq A \cap W \leq B \cap W \leq E^+ \cap W = E$ and so $E = A \cap W = B \cap W$. So by definition (see 4.8.8b), $(A, B) \notin \mathrm{Jump}^W(\mathcal{C})$. Also $B \not\leq W$ and so as M/W is simple, $M = B + W$. If $A \not\leq W$, then also $M = A + W$ and $B = B \cap M = B \cap (A + W) = A + (B \cap W) \leq A$ a contradiction. Hence $A \leq W$ and $A = B \cap W$. Thus

$$B/A = B/B \cap W \cong B + W/W = M/W$$

We claim that $\mathrm{Jump}(\mathcal{C}) = \mathrm{Jump}^W(\mathcal{C}) \cup \{(A, B)\}$. So let (X, Y) be a jump of $\mathcal{C}$ not contained in $\mathrm{Jump}^W(\mathcal{C})$. By 4.8.10c, $Y \not\leq X + (Y \cap W)$ and so also $Y \not\leq X + W$. Thus $Y \not\leq W$ and $X \leq W$. As $A \leq W$, $Y \not\leq A$. As (A, B) is a jump $B \leq Y$. As $B \not\leq W$, $B \not\leq X$ and so $X \leq A$. Thus $X \leq A < B \leq Y$ and as (X, Y) is a jump, $(A, B) = (X, Y)$.

By 4.8.10c, the factors of $\mathrm{Jump}^W(\mathcal{C})$ are isomorphic to the factors of $\mathcal{C} \cap W$ and so with the factors of $\mathcal{D} - M$. As $B/A \cong M/W$ it only remains to show that $\mathcal{D}$ is finite. But thus follows from 4.8.9. □

4.9 Matrices

Let R be a ring and I, J sets. Define

$$\mathcal{M}_R(I, J) = \{(m_{ij})_{i \in I, j \in J} \mid m_{ij} \in R\}$$

$M = (m_{ij})_{i \in I, j \in J}$ is called an $I \times J$-matrix over R. For $j \in J$, put $M^j = (m_{ij})_{i \in I}$, $M^j \in R^I$ is called the j'th column of M. For $i \in I$ put $M_i = (m_{ij})_{j \in J}$, M_i is called the i'th row of M. Note that as abelian groups, $\mathcal{M}_R(I, J) \cong R^{I \times J}$. Define

$$\mathcal{M}_R^c(I, J) = \{M \in \mathcal{M}_R(I, J) \mid \forall j \in J, \{i \in I, m_{ij} \neq 0\} \text{ is finite}\}$$

Let $M \in \mathcal{M}_R(I, J)$. Then $M \in \mathcal{M}_R^c(I, J)$ if and only if each column of M lies in $\bigoplus_J R$.

If I, J, K are sets we define a multiplication

$$\mathcal{M}^c(I, J) \times \mathcal{M}^c(J, K) \to \mathcal{M}^c(I, K)$$

by

$$(a_{ij})(b_{jk}) = (\sum_{j \in J} a_{ij} b_{jk})_{i \in I, k \in K}$$

Fix $k \in K$. Then there exists only finitely $j \in J$ with $b_{jk} \neq 0$, so the above sum is well defined. Also for each of these j's there are only finitely many $i \in I$ for which a_{ij}

is not 0. Hence there exists only finitely many i's for which $\sum_{j\in J} a_{ij}b_{jk}$ is not 0. So $(a_{ij})(b_{jk}) \in \mathcal{M}_R^c(I,K)$.

Put $\mathcal{M}_R(I) = \mathcal{M}_R(I,I)$ and $\mathcal{M}_R^c(I) = \mathcal{M}_R^c(I,I)$

Lemma 4.9.1. *Let R be a ring and V, W, Z free R-modules with bases I, J and K, respectively.*

(a) Define $m_A(j,i) \in R$ by $A(i) = \sum_{j\in J} m_A(i,j)j$ and put $M_A(J,I) = (m_A(j,i))$. Then the map

$$\mathrm{M}(J,I) : \mathrm{Hom}_R(V,W) \to \mathcal{M}_{R^{\mathrm{op}}}^c(J,I)$$

$$A \to \mathrm{M}_A(J,I)$$

is an isomorphism of abelian groups.

(b) Let $A \in \mathrm{Hom}_R(V,W)$ and $B \in \mathrm{Hom}_R(W,Z)$. Then

$$\mathrm{M}_B(K,J)\mathrm{M}_A(J,I) = \mathrm{M}_{BA}(K,I).$$

(c) Let $\mathrm{M}(I) := \mathrm{M}(I,I)$. Then $M(I) : \mathrm{End}_R(V) \to \mathcal{M}_{R^{\mathrm{op}}}^c(I)$ is ring isomorphism.

Proof. (a) Note first that as J is a basis of W, the $m_A(j,i)$'s are well defined. To show that $M_A(J,I)$ is a bijection we determine it inverse. Let $M = (m_{ji}) \in \mathcal{M}_{R^{\mathrm{op}}}(J,I)$. Define $A_M \in \mathrm{Hom}_R(V,W)$ by

$$A_M(\sum_{i\in I} r_i i) = \sum_{j\in J}(\sum_{i\in I} r_i m_{ji})j$$

It is easy to check that A_M is R-linear and that the map $M \to A_M$ is inverse to $\mathrm{M}_A(J,I)$.
(b)

$$(BA)(i) = B(A(i)) = B(\sum_{j\in J} m_A(j,i)j) = \sum_{j\in J} m_A(j,i)B(j) =$$

$$= \sum_{j\in J} m_A(j,i)(\sum_{k\in K} m_B(k,j)k) = \sum_{k\in K}(\sum_{j\in J} m_A(j,i)m_B(k,j))k$$

Thus

$$m_{BA}(k,i) = \sum_{j\in J} m_A(j,i)m_B(k,j) = \sum_{j\in J} m_B(k,j) \cdot^{\mathrm{op}} m_A(j,i)$$

So (b) holds.
(c) Follows from (b) and (c). □

Definition 4.9.2. *Let R be a ring, V and W R-modules, $A \in \mathrm{End}_R(V)$ and $B \in \mathrm{End}_R(W)$. We say that A and B are* similar *over R if there exists a R-linear isomorphism $\Phi : V \to W$ with $\Phi \circ A = B \circ \Phi$.*

We leave it as an exercise to show that "similar" is an equivalence relation. Also the condition $\Phi \circ A = B \circ \Phi$ is equivalent to $B = \Phi \circ A \circ \Phi^{-1}$.

Let I and J be sets and $\phi : I \to J$ a function. If $M = (m_{\tilde{j}j})$ is a $J \times J$ matrix, let M^ϕ be the $I \times I$ matrix $(m_{\phi(\tilde{i})\phi(i)})$.

Lemma 4.9.3. *Let R be a ring, V and W R-modules, $A \in \mathrm{End}_R(V)$ and $B \in \mathrm{End}_K(V)$. Suppose that V is free with basis I. Then A and B are similar if and only if there exists a basis J for W and a bijection $\phi : I \to J$ with*

$$M_A(I) = M_B^\phi(J)$$

Proof. Suppose first that A and B are similar. Then there exists an R-linear isomorphism $\Phi : V \to W$ with $\Phi \circ A = B \circ \Phi$. Let $J = \Phi(I)$. As I is a basis for V and Φ is an isomorphism, J is a basis for W. Let $\phi = \Phi \mid_I$. We compute

$$B(\phi(i)) = \Phi(A(i)) = \Phi(\sum_{\tilde{i} \in i} M_A(\tilde{i}, i)\tilde{i}) = \sum_{\tilde{i} \in i} M_A(\tilde{i}, i)\phi(\tilde{i})$$

Hence $M_B(\phi(\tilde{i}), \phi(i)) = M_A(\tilde{i}, i)$ and $M_A(I) = M_B^\phi(J)$.

Suppose conversely that there exist a basis J for W and a bijection $\phi : I \to J$ with $M_A(I) = M_B^\phi(J)$. Then $m_A(\tilde{i}, i) = m_B(\phi(\tilde{i}), \phi(i))$.

Let $\Phi : V \to W$ be the unique R-linear map from V to W with $\Phi(i) = \phi(i)$ for all $i \in I$. As I and J are bases, Φ is an isomorphism. Moreover,

$$\Phi(A(i)) = \Phi(\sum_{\tilde{i} \in I} m_A(\tilde{i}, i)\tilde{i}) = \sum_{\tilde{i} \in I} m_A(\tilde{i}, i)\phi(\tilde{i}) =$$

$$= \sum_{\tilde{i} \in I} m_B(\phi(\tilde{i}), \phi(i))\phi(\tilde{i}) = \sum_{j \in J} m_B(j, \phi(i))j = B(\phi(i))$$

Hence $\Phi \circ A$ and $B \circ \Phi$ agree on I and so $\Phi \circ A = B \circ \Phi$. □

Let R be a ring and V a module over R. Let $A \in \mathrm{End}_R(V)$. Define $\alpha : R \to \mathrm{End}_{\mathbb{Z}}(V)$ by $\alpha(r)v = rv$, we will usually right $r\mathrm{id}_V$ for $\alpha(r)$. Note that A commutes with each $r\mathrm{id}_V$ and so by 3.6.1 there exists a ring homomorphism $\alpha_A : R[x] \to \mathrm{End}_{\mathbb{Z}}(V)$ with $r \to r\mathrm{id}_V$ and $x \to A$. Let $f = \sum_{i=0}^n r_i x^i \in R[x]$. We will write $f(A)$ for $\alpha_A(f)$. Then $f(A) = \sum_{i=0}^n r_i A^i$. It follows that V is a $R[x]$-module with

$$fv = f(A)(v) = \sum_{i=0}^n r_i A^i(v)$$

To indicate the dependence on A we will sometimes write V_A for the $R[x]$ module V obtain in this way.

Lemma 4.9.4. *Let R be a ring and V and W R-modules. Let $A \in \mathrm{End}_R(V)$. and $B \in \mathrm{End}_R(V)$. Then the $R[x]$-modules V_A and W_B are isomorphic if and only if A and B are similar over R.*

Proof. Suppose first that V_A and V_B are isomorphic. Then there exists an $R[x]$-linear isomorphism $\Phi : V \to W$. In particular Φ is R-linear and $\Phi(xv) = x\Phi(v)$ for all $v \in V$. By definition of V_A and W_B thus means $\Phi(A(v)) = B(\Phi(v)$ and so A and B and are similar.

Conversely, if A and B are similar there exists an R-linear isomorphism $\Phi : V \to W$ with $\Phi \circ A = B \circ \Phi$. Hence $\Phi(rv) = r\Phi(v)$ and $\Phi(xv) = x\Phi(v)$ for all $r \in R$ and $v \in V$. Since Φ is $\mathbb{Z}$-linear this implies $\Phi(fv) = f\Phi(v)$ for all $f \in R[x]$. Hence Φ is an $R[x]$-linear isomorphism. □

Lemma 4.9.5. *Let R be a ring and $f = \sum_{i=0}^{n} a_i x^i$ a monic polynomial of degree $n > 0$. Let $I = R[x]f$ be the left ideal in $R[x]$ generated by f.*

(a) $\{x^i \mid i \in \mathbb{N}\}$ is a basis for $R[x]$ as a left R-module.

(b) For $0 \leq i < n$ let h_i be a monic polynomial of degree i. Then $\{h_i + I \mid 0 \leq i < n\}$ is basis for $R[x]/I$.

(c) Let $A \in \mathrm{End}_R(R[x]/I)$ be defined by $A(h + I) = hx + I$.

(a) The matrix of A with respect the basis

$$1 + I, x + I, \ldots x^{n-1} + I$$

is

$$M(f) := \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & -a_2 \\ 0 & 0 & 1 & \ldots & 0 & 0 & -a_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -a_{n-1} \end{pmatrix}$$

(b) Suppose that $f = g^m$ for some monic polynomial g of degree s and some $m \in \mathbb{Z}^+$. Let E^{1s} be the matrix $k \times k$ with $e_{ij} = 0$ if $(i, j) \neq (1, s)$ and $e_{1s} = 1$. Then the matrix of A with respect to the basis

$$1+I, x+I, \ldots x^{s-1}, g+I, xg+I, \ldots, x^{s-1}g+I, \ldots, g^{m-1}+I, xg^{m-1}+I, x^{s-1}g^{m-1}+I,$$

has the form

$$M(g,m) := \begin{pmatrix} M(g) & 0 & 0 & \dots & 0 & 0 & 0 \\ E^{1s} & M(g) & 0 & \ddots & 0 & 0 & 0 \\ 0 & E^{1s} & M(g) & \ddots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & M(g) & 0 & 0 \\ 0 & 0 & 0 & \ddots & E^{1s} & M(g) & 0 \\ 0 & 0 & 0 & \dots & 0 & E^{1s} & M(g) \end{pmatrix}$$

Proof. (a) is obvious as any polynomial can be uniquely written as R-linear combination of the x^i.

(b) We will first show by induction on $\deg h$ that every $h + I, h \in R[x]$ is a R linear combination of the $h_i, 0 \leq i < n$. By 3.7.5 $h = qf + r$ for some $q, r \in R[x]$ with $\deg r < \deg f = n$. Since $h + I = r + I$ we may assume that $h = r$ and so $i := \deg h < n$. Let a be the leading coefficient of h. Then $\deg h - ah_i < \deg h$ and so by induction is a linear combination of the h_i's.

Suppose now that $\sum_{i=0}^{n-1} \lambda_i(h_i + I) = 0 + I$ for some $\lambda_i \in \mathbb{K}$, not all 0. Then $h := \sum_{i=0}^{n-1} \lambda_i h_i \in I$. Let j be maximal with $\lambda_j \neq 0$. Then clearly $j = \deg h$ and the leading coefficient of h is λ_j. In particular $h \neq 0$.

Note that all non-zero polynomials in I have degree larger or equal to n. But this contradicts $0 \neq h \in I$ and $\deg h = j < n$. Thus (b) holds.

(ca) is the special case $g = f$ and $m = 1$ of (cb). So it remains to prove (cb). Note that $\deg x^i g^j = i + js$. Hence by (b) $\{x^i g^j + I \mid 0 \leq i < s, 0 \leq j < m\}$ is a basis for $R[x]/I$.

Let $y_{i,j} := x^i g^j + I$. Then

$$A(y_{i,j}) = x^{i+1} g^j + I.$$

Thus

$$A(y_{i,j}) = y_{i+1,j} \text{ for all } 0 \leq i < s-1, 0 \leq j < m.$$

Let $g = \sum_{i=0}^{s} b_i x^i$. As g is monic $b_s = 1$ and so $x^s = g - \sum_{i=0}^{s-1} b_i x^i$.

Hence

$$A(y_{s-1,j}) = x^s g^j + I = (g^{j+1} - \sum_{i=0}^{s-1} b_i x^i g^j) + I = (g^{j+1} + I) - \sum_{i=0}^{s-1} b_i y_{i,j}.$$

If $j < m - 1$, $g^{j+1} + I = y_{0,j+1}$ and so

$$A(y_{s-1,j}) = y_{0,j+1} - \sum_{i=0}^{s-1} b_i y_{i,j}$$

If $j = m-1$ then $g^{j+1} = g^m = f \in I$ and so

$$A(y_{s-1,m-1}) = -\sum_{i=0}^{s-1} b_i y_{s-1,m-1}$$

Thus (cb) holds. □

Theorem 4.9.6 (Jordan Canonical Form)**.** *Let $\mathbb{K}$ be a field, V a non-zero finite dimensional vector space over $\mathbb{K}$ and $A \in \mathrm{End}_{\mathbb{K}}(V)$. Then there exist irreducible monic polynomials $f_1, \ldots, f_t \in \mathbb{K}[x]$, positive integers $m_1, \ldots m_t$ and a basis*

$$y_{ijk}, 0 \le i < \deg f_k, 0 \le j < m_k, 1 \le k \le t$$

of V so that the matrix of A with respect to this basis has the form

$$M(f_1, m_1 \mid \ldots \mid f_t, m_t) := \begin{pmatrix} M(f_1,m_1) & 0 & \ldots & 0 & 0 \\ 0 & M(f_2,m_2) & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & M(f_{t-1},m_{t-1}) & 0 \\ 0 & 0 & \ldots & 0 & M(f_t,m_t) \end{pmatrix}$$

Proof. View V as a $\mathbb{K}[x]$-module by $fv = f(A)(v)$ for all $f \in \mathbb{K}[x]$ and $v \in V$ (see before 4.9.4). Since $\mathbb{K}[x]$ is a PID (3.4.6) we can use Theorem 4.3.12. Thus V_A is the direct sum of modules V_k, $1 \le k \le t$ with $V_k \cong K[x]/(f_k^{m_k})$, where $f_k \in \mathbb{K}[x]$ is either 0 or prime, and $m_k \in \mathbb{Z}^+$. By 4.9.5(a) $\mathbb{K}[x]$ is infinite dimensional over $\mathbb{K}$. As V is finite dimensional, $f_k \neq 0$. So we may choose f_k to be irreducible and monic. By 4.9.5(cb), V_k has a basis y_{ijk}, $0 \le i < \deg f_k, 0 \le j < m_k$ so that the matrix of $A\mid_{V_k}$ with respect to this basis is $M(f_k, m_k)$. Combining the basis for V_k, $1 \le k \le t$, to a basis for V we see that the theorem is true. □

The matrix $M(f_1, m_1 \mid f_2, m_2 \mid \ldots \mid f_t, m_t)$ from the previous theorem is called the *Jordan canonical form* of A. We should remark that our notion of the Jordan canonical form differs slightly from the notion found in most linear algebra books. It differs as we do not assume that all the roots of the minimal polynomial (see below) of A are in $\mathbb{K}$. Note that if $\mathbb{K}$ contains all the roots then $f_k = x - \lambda_k$ and $M(f_k)$ is the 1×1 matrix (λ_k) and E^{1s} is the 1×1 identity matrix. So the obtain the usual Jordan canonical form.

We remark that the pairs $(f_k, m_k), 1 \le k \le t$ are unique up to ordering. Indeed let f be an irreducible monic polynomial of degree s and m a positive integer. Then the number of k's with $(f_k, m_k) = (f, m)$ is $\frac{d}{s}$ where d is the dimension of the $\mathbb{K}$-space

$$\ker f^m(A) / \ker f^m(A) \cap \mathrm{Im}\, f(A)$$

We leave the details of this computation to the dedicated reader.

The following two polynomials are useful to compute the Jordan canonical form of A. The *minimal polynomial* m_A and the *characteristic polynomial* χ_A.

m_A is defined has the monic polynomial of minimal degree with $m_A(A) = 0$. i.e m_A is monic and (m_A) is the kernel of the homomorphism $\alpha_A : \mathbb{K}[x] \to \mathrm{End}_{\mathbb{K}}(V)$. m_A can be computed from the Jordan canonical form. For each monic irreducible polynomial let e_f be maximal so that (f, e_f) is one of the (f_k, m_k) (with $e_f = 0$ if f is not one of the f_k.)Then

$$m_A = \prod f^{e_f}$$

The characteristic polynomial is defined as

$$\chi_A = (-1)^n f_1^{m_1} f_2^{m_2} \dots f_k^{m_k}$$

where n is the dimension of V. The importance of the characteristic polynomials comes from the fact that χ_A can be computed without knowledge of f_k's. Indeed

$$\chi_A = \det(A - x\mathrm{id}_V).$$

To see this we use the Jordan canonical form of f. Note that

$$\det(A - x\mathrm{id}_V) = \prod_{k=1}^{t} \det(M(f_k, m_k) - xI)$$

and

$$\det(M(f, m) - xI) = (\det(M(f) - xI))^m.$$

Finally its is easy to verify that

$$\det(M(f) - xI) = (-1)^{\deg f} f.$$

Chapter 5

Fields

5.1 Extensions

Definition 5.1.1. *Let F be an integral domain, $\mathbb{K}$ a subfield of F and $a \in F$.*

(a) F is called an extension *of $\mathbb{K}$. We will also say that $\mathbb{K} \leq F$ is an extension.*

(b) If F is a field, F is called field extension *of $\mathbb{K}$*

(c) A vector space *over $\mathbb{K}$ is a unitary $\mathbb{K}$-module. A vector space over $\mathbb{K}$ is also called a* $\mathbb{K}$-space.

(d) If V is a $\mathbb{K}$-space, then $\dim_{\mathbb{K}} V$ is the cardinality of a $\mathbb{K}$ basis for V. (Note here that 4.2.11(d), V has a basis and by Homework 1#3, any two basis have the same cardinality.)

(e) The extension $\mathbb{K} \leq \mathbb{F}$ is called a finite *if $\dim_{\mathbb{K}} F$ finite, where F is viewed as a $\mathbb{K}$ space by left multiplication.*

(f) If S is a ring, R a subring if S and $I \subseteq R$, then

$$R[I] := \bigcap \{T \mid T \text{ is a subring of } S \text{ with } R \cup I \subseteq S\}$$

$R[I]$ is called the subring of S generated by R and I.

(g) If F is a field and $I \subseteq F$, then

$$\mathbb{K}(I) := \bigcap \{T \mid T \text{ is a field of } F \text{ with } \mathbb{K} \cup I \subseteq F\}$$

$\mathbb{K}(I)$ is called the subfield of $\mathbb{F}$ generated by $\mathbb{K}$ and I.

(h) A polynomial $f \in \mathbb{K}[x]$ is called monic if its leading coefficient is $1_{\mathbb{K}}$.

(i) $\Phi_a = \Phi_a^{\mathbb{K}}$ *denotes the unique ring homomorphism*

$$\Phi_a : \mathbb{K}[x] \to \mathbb{K}[a], \qquad \text{with } \Phi_a(x) = a \text{ and } \Phi_a(k) \text{ for all } k \in \mathbb{K}.$$

(See Example 3.2.6(1).)

(j) The unique zero or monic polynomial $m_a = m^{\mathbb{K}}(a) \in \mathbb{K}[x]$ *with* $\ker \Phi_a = \mathbb{K}[x]m_a$ *is called the* minimal polynomial *of* a *over* $\mathbb{K}$.

(k) a *is called* algebraic *over* $\mathbb{K}$ *if* $m_a \neq 0_F$.

(l) The extension $\mathbb{K} \subseteq F$ *is called* algebraic *if all* $b \in F$ *are algebraic over* $\mathbb{K}$.

(m) a *is called* transcendental *over* $\mathbb{K}$ *if* $m_a = 0_F$.

Note that we used the symbol $\mathbb{K}[I]$ also to denoted the polynomial ring in the variables I. To avoid confusion we will from now denote polynomials ring by $\mathbb{K}[x_i, i \in I]$. The field of fraction of $\mathbb{K}[x_i, i \in I]$ is denoted by $\mathbb{K}(x_i, i \in I)$.

Lemma 5.1.2. *Let* $\mathbb{K} \leq F$ *be an extension and* $a \in F$. *Then one of the following holds*

1. Φ_a *is not 1-1,* $\dim_{\mathbb{K}} \mathbb{K}[a] = \deg m_a$ *is finite,* m_a *is monic and irreducible,* $\mathbb{K}[a] = \mathbb{K}(a)$ *is a field,* a *is algebraic over* $\mathbb{K}$, *and* $(a^i, 0 \leq i < \deg m_a)$ *is a basis for* $K[a]$.

2. Φ_a *is an isomorphism,* $\dim_{\mathbb{K}} \mathbb{K}[a] = \infty$, $m_a = 0_{\mathbb{K}}$, a *is not invertible in* $\mathbb{K}[a]$, a *is transcendental over* $\mathbb{K}$, $(a^i, 0 \leq i < \infty)$ *is a basis for* $\mathbb{K}[a]$.

Proof. Since F is an integral domain, $\mathbb{K}[a]$ is an integral domain. Clearly Φ_a is onto and so $\mathbb{K}[x]/\mathbb{K}[x]m_a \cong \mathbb{K}[x]/\ker \Phi_a \cong \mathbb{K}[a]$. Thus by 3.3.8 $\mathbb{K}[x]m_a$ is a prime ideal.

Suppose first that $m_a \neq 0$. Then a is algebraic over $\mathbb{K}[a]$ and Φ_a is not 1-1. Note that by 3.3.8 m_a is a prime. By Example 3.4.2(2), $\mathbb{K}[x]$ is am Euclidean domain and so also a PID. So we conclude from 3.3.11 that m_a is irreducible and $\mathbb{K}[a] \cong \mathbb{K}[x]/\mathbb{K}[x]m_a$ is a field. Let $f \in \mathbb{K}[x]$. As $\mathbb{K}[x]$ is a Euclidean domain, $f \equiv g \pmod{m_a}$ for a unique polynomial $g \in K[x]$ with $\deg g < \deg m_a$. Also g is a unique $\mathbb{K}$-linear combination of $(x^i, 0 \leq i < \deg m_a$ and so $(x^i + \mathbb{K}[x]m_a, 0 \leq i < \deg m_a)$ is a basis for $\mathbb{K}[x]/\mathbb{K}[x]m_a$. Hence $(a^i, 0 \leq i < \deg m_a)$ is basis for $\mathbb{K}[a]$. Thus (1) holds.

Suppose next that $m_a = 0_{\mathbb{K}}$. Then a is transcendental. Moreover, Φ_a is 1-1 and so an isomorphism. Since x is not invertible in $\mathbb{K}[x]$ and $(x^i, i \in \mathbb{N})$ is a basis for $\mathbb{K}$ we conclude that a is not invertible in $\mathbb{K}[a]$ and $(a^i, i \in \mathbb{N})$ is a basis for $\mathbb{K}[a]$. So (2) holds in this case. □

In this case m_a^K is the monic, irreducible polynomial of minimal degree with respect to $m_a^{\mathbb{K}}(a) = 0$. So any algebraic extension is a field extension. Also by the previous theorem any finite extension is algebraic and so a field extension. Note that every finite integral domain is a finite extension of some $\mathbb{Z}/p\mathbb{Z}$. So we obtain a second proof that every finite integral domain is a field (cf. 3.1.14).

Lemma 5.1.3. *(a) Let S be a ring, R a subring of S and $I \subseteq S$. Then*

$$R[I] = \bigcup \{R[J] \mid J \subseteq I, J \text{ is finite}\}.$$

(b) Let $\mathbb{K} \leq \mathbb{F}$ be a field extension and $I \subseteq \mathbb{F}$. Then

$$\mathbb{K}(I) = \bigcup \{\mathbb{K}(J) \mid J \subseteq I, J \text{ is finite}\}.$$

Proof. (a) Let $T = \bigcup \{R[J] \mid J \subseteq I, J \text{ is finite}\}$. Clearly $R \cup I \subseteq T \subseteq \mathbb{K}[I]$ and we just need to verify that T is a subring of F. For this let $a, b \in T$. Then $a \in R[J_a]$ and $b \in R[J_b]$ for some finite subset J_a and J_b of I. Then 0_S, $a + b$, $-a$, and ab all are contained in $\mathbb{K}[J_a \cup J_b] \subseteq T$. So T is indeed a subring and $T = R[I]$.

(b) Similar to (a). □

Lemma 5.1.4. *(a) Let R be a ring with identity, M a unitary R-module and S a subring of R. Let $\mathfrak{r} = (r_i)_{i \in I}$ be a tuple of elements in R and $\mathfrak{m} = (m_j)_{j \in J}$ tuple of elements in M. Put $\mathfrak{rm} = (r_i m_j)_{(i,j) \in I \times J}$.*

- *(a) If $R = \langle \mathfrak{r} \rangle_S$ and $M = \langle \mathfrak{m} \rangle_R$, then $M = \langle \mathfrak{rm} \rangle_S$*
- *(b) If $\mathfrak{r}$ is linearly independent over S and $\mathfrak{m}$ is linearly independent over R, then $\mathfrak{rm}$ is linearly independent over S.*
- *(c) If $\mathfrak{r}$ is an S-basis for R and $\mathfrak{m}$ is an R-basis for M, then $\mathfrak{rm}$ is an S-basis for M.*

(b) Let $\mathbb{K} \leq \mathbb{E}$ be a field extension and V a vector space over $\mathbb{E}$. Then

$$\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} \mathbb{E} \cdot \dim_{\mathbb{E}} V.$$

(c) Let $\mathbb{K} \leq \mathbb{E}$ be a field extension and $\mathbb{E} \leq F$ an extension. Then

$$\dim_{\mathbb{K}} F = \dim_{\mathbb{K}} \mathbb{E} \cdot \dim_{\mathbb{E}} F.$$

In particular, if $\mathbb{K} \leq \mathbb{E}$ and $\mathbb{E} \leq F$ are finite, also $\mathbb{K} \leq F$ is finite.

Proof. (a:a) Let $m \in M$. Since $\langle \mathfrak{m} \rangle_R = M$, $m = \sum_{j \in J} k_j m_j$ for some, almost all 0, $k_j \in R, j \in J$. $j \in J$. Since $\langle \mathfrak{r} \rangle_S = R$, $k_j = \sum_{i \in I} s_{ij} r_i$ for some, almost all zero, $s_{ij} \in S$, where we choose $s_{ij} = 0$ if $k_j = 0$. Then

$$m = \sum_{j \in J} k_j m_j = \sum_{j \in J} \left(\sum_{i \in I} s_{ij} r_i \right) m_j = \sum_{(i,j) \in I \times j} s_{ij} r_i m_j$$

Thus (a:a) holds.

(a:b) Suppose that $\sum_{(i,j) \in I \times J} s_{ij} r_i m_j = 0$, for some almost all zero $s_{ij} \in S$. Then

$$\sum_{j\in J}\left(\sum_{i\in I} s_{ij}r_i\right)m_j = 0$$

Since $\mathfrak{m}$ is linearly independent over R, we conclude $\sum_{i\in I} s_{ij}r_i = 0$ for all j in J. As $\mathfrak{r}$ is linearly independent over S we get $s_{ij} = 0$ for all $(i,j) \in I \times J$. Thus (a:b) holds.

(a:c) follows from (a:a) and (a:b). (b) follows from (a:c) and (c) follow from (c) . □

Lemma 5.1.5. *Let $\mathbb{K} \leq F$ be an extension, let $a \in F$ be algebraic over $\mathbb{K}$ and let $f \in \mathbb{K}[x]$.*

(a) $f(a) = 0$ if and only if $m_a \mid f$ in $\mathbb{K}[x]$.

(b) If f is irreducible then $f(a) = 0$ if and only if $f \sim m_a$ in $\mathbb{K}[x]$. That is if and only if $f = km_a$ for some $k \in \mathbb{K}^\sharp$.

(c) m_a is the unique monic irreducible polynomial in $\mathbb{K}[x]$ with a as a root.

Proof. (a) Since $f(a) = \Phi_a(f)$, $f(a) = 0$ if and only if $a \in \ker \Phi_a$. Since $\ker \Phi_a = \mathbb{K}[x]m_a$, this holds if and only if $m_a \mid f$.

(b) Let f be irreducible with $f(a) = 0$, then $m_a \mid f$. Since f is irreducible we get $m_a \sim f$. By 3.3.3 this means $f = km_a$ for some unit k in $\mathbb{K}[x]$. It is easy to see that the units in $\mathbb{K}[x]$ are exactly the non-zero constant polynomials. So $k \in \mathbb{K}^\sharp$

(c) If in addition f is monic, then since also m_a is monic we conclude $k = 1$ and $f = m_a$. □

Lemma 5.1.6. *Let $\mathbb{K} \leq \mathbb{E}$ be a field extension, $\mathbb{E} \leq F$ an extension and $b \in F$. If b is algebraic over $\mathbb{K}$, then b is algebraic over $\mathbb{E}$ and $m_b^{\mathbb{E}}$ divides m_b^K in $\mathbb{E}[x]$.*

Proof. Note that $m_b^{\mathbb{K}}(b) = 0$ and $m_b^{\mathbb{K}} \in \mathbb{E}[x]$. So by 5.1.5 $m_a^{\mathbb{E}}$ divides $m_b^{\mathbb{K}}$ in $\mathbb{E}[x]$. Since b is algebraic over $\mathbb{K}$, $m_b^{\mathbb{K}} \neq 0$ and so also $m_b^{\mathbb{E}} \neq 0$. Hence b is algebraic over $\mathbb{E}$. □

Definition 5.1.7. *Let $\mathbb{K}$ be a field and $f \in \mathbb{K}[x]$. We say that f* splits *over $\mathbb{K}$ if*

$$f = k_0(x - k_1)(x - k_2)\dots(x - k_n)$$

for some $n \in \mathbb{N}$ and $k_i \in \mathbb{K}, 0 \leq i \leq n$..

Lemma 5.1.8. *Let $\mathbb{K}$ be a field and $f \in K[x]^\sharp$.*

(a) If $\mathbb{K} \leq \mathbb{E}$ is a field extension, $f \in \mathbb{K}[x]$ is irreducible and $\mathbb{E} = \mathbb{K}[a]$ for some root a of f in $\mathbb{E}$, then the map

$$\mathbb{K}[x]/f\mathbb{K}[x] \to \mathbb{E}, h + f\mathbb{K}[x] \to h(a)$$

is ring isomorphism.

(b) If f is not constant, then there exists a finite field extension $\mathbb{K} \leq \mathbb{E}$ such that f has a root in $\mathbb{E}$ and $\dim_{\mathbb{K}} \mathbb{E} \leq \deg f$.

(c) There exists a finite field extension $\mathbb{K} \leq \mathbb{F}$ such that f splits over and $\dim_{\mathbb{K}} \mathbb{F} \leq (\deg f)!$.

Proof. (a) By 5.1.5(b), $f \sim m_a$. Thus $\ker \Phi_a = m_a\mathbb{K}[x]$. Also $h(a) = \Phi_a(h)$ and (a) follows from Isomorphism Theorem of Rings.

(b) Let g be an irreducible divisor of f in $\mathbb{K}[x]$. Put $\mathbb{E} = \mathbb{K}[x]/g\mathbb{K}[x]$. Then $\mathbb{E}$ is a field For $h \in \mathbb{K}[x]$. Let $\overline{h} = h + g\mathbb{K}[x] \in \mathbb{E}$. Note that the map $h \to \overline{h}$ is a ring homomorphism. Put $a = \overline{x}$. We identify $k \in \mathbb{K}$ with $\overline{k} \in \mathbb{E}$. Then $\mathbb{K}$ is a subfield of $\mathbb{E}$, $a^i, 0 \leq i < \deg g$ is a $\mathbb{K}$ basis for $\mathbb{E}$ and so $\dim_{\mathbb{K}} \mathbb{E} = \deg g \leq \deg f$. Let $f = \sum_{i=0}^{n} k_i x^i$ with $k_i \in g$. Then

$$f(a) = \sum_{i=0}^{n} k_i a^i = \sum_{i=0}^{n} \overline{k_i}\overline{x}^i = \overline{\sum_{i=0}^{n} k_i x^i} = \overline{f}.$$

Since $g \mid f$, $f \in g\mathbb{K}[x]$ and so $\overline{f} = 0_{\mathbb{E}}$. Thus $f(a) = 0_{\mathbb{E}}$ and a is a root of f in $\mathbb{E}$.

(c) Let $\mathbb{E}$ be as in (b) and e a root of f in $\mathbb{E}$. Then $f = (x - e)g$ for some $g \in \mathbb{E}[x]$ with $\deg g = \deg f - 1$. By induction on $\deg f$ there exists a field extension $\mathbb{E} \leq \mathbb{F}$ such that g splits over $\mathbb{F}$ and $\dim_{\mathbb{E}} \mathbb{F} \leq (\deg g)! = (\deg f - 1)!$. Then f splits over $\mathbb{F}$ and

$$\dim_{\mathbb{K}} \mathbb{F} = \dim_{\mathbb{K}} \mathbb{E} \cdot \dim_{\mathbb{E}} \mathbb{F} \leq (\deg f - 1)! \deg f = \deg f!.$$

□

Example 5.1.9. Let $f = x^2+1 \in \mathbb{R}[x]$. Then f has no root in $\mathbb{R}$ and so is irreducible over $\mathbb{R}$. Thus $\mathbb{E} = \mathbb{R}[x]/(x^2 + 1)\mathbb{R}[x]$ is a field. For $h \in \mathbb{R}[x]$ let $\overline{h} = h + f\mathbb{R}[x] \in \mathbb{E}$. We also identify $r \in \mathbb{R}$ with $\overline{r}$ in $\mathbb{E}$. Put $i = \overline{x}$. Then i is a root of f in $\mathbb{E}$ and so $i^2 + 1 = 0$ and $i^2 = -1$. Moreover $1, i$ is an $\mathbb{R}$ basis for $\mathbb{F}$. Let $a, b, c, d \in \mathbb{R}$. Then $(a+bi)+(c+di) = (a+b)+(c+d)i$ and

$$(a + bi)(c + di) = ac + bdi^2 + (ad + bc)i = (ac - bd) + (ad + bc)i$$

Hence $\mathbb{E}$ is isomorphic the field $\mathbb{C}$ of complex numbers.

Lemma 5.1.10. *(a) Any finite extension is algebraic.*

(b) If $\mathbb{K} \leq \mathbb{E}$ is a finite field extension and $\mathbb{E} \leq F$ is a finite extension, then $\mathbb{K} \leq F$ is a finite extension.

Proof. (a) Let $\mathbb{K} \leq F$ a finite extension and $a \in F$. Then by 4.2.5 $\dim_{\mathbb{K}} \mathbb{K}[a] \leq \dim_{\mathbb{K}} F < \infty$. Thus by 5.1.2 a is albegraic over $\mathbb{K}$.

(b) follows from 5.1.4c □

‘

Lemma 5.1.11. *Let $\mathbb{K} \leq F$ be an extension and $A \subseteq F$ be a set of elements in F algebraic over $\mathbb{K}$.*

(a) If A is finite, $\mathbb{K} \leq \mathbb{K}[A]$ is a finite extension

(b) $\mathbb{K} \leq \mathbb{K}[A]$ *is an algebraic extension.*

(c) The set of elements in F *algebraic over* $\mathbb{K}$ *form a subfield of* F*.*

Proof. (a) By induction on $|A|$. If $|A| = 0$, $\mathbb{K}[A] = \mathbb{K}$. So suppose $A \neq \emptyset$ and let $a \in A$. Put $B = A \setminus \{a\}$. By induction $\mathbb{K} \leq \mathbb{K}[B]$ is finite. As a is algebraic over $\mathbb{K}$, a is algebraic over $\mathbb{K}[B]$ (see 5.1.6) Thus by 5.1.2 $\mathbb{K}[B] \leq \mathbb{K}[B][a]$ is finite. Hence by 5.1.10(b) also $\mathbb{K} \leq \mathbb{K}[B][a]$ is finite. Since $\mathbb{K}[B][a] = \mathbb{K}[A]$ we conclude that (a) holds.

(b) Let $b \in \mathbb{K}[A]$. By 5.1.3(a), $b \in \mathbb{K}[B]$ for some finite $B \subseteq A$. By (a) $\mathbb{K} \leq \mathbb{K}[B]$ is finite and so also algebraic (5.1.10(b)). So b is algebraic over $\mathbb{K}$.

(c) Follows from rf b applied to A being the set of all algebraic elements in L. □

Proposition 5.1.12. *Let* $\mathbb{K} \leq \mathbb{E}$ *and* $\mathbb{E} \leq \mathbb{F}$ *be algebraic field extensions. Then* $\mathbb{K} \leq \mathbb{F}$ *is algebraic.*

Proof. Let $b \in F$ and $m = m_b^{\mathbb{E}}$. Let $m = \sum_{i=0}^{n} e_i x^i$ and $A = \{e_0, e_2 \ldots, e_n\}$. Then A is a finite subset of $\mathbb{E}$.

Since $\mathbb{K} \leq \mathbb{E}$ is algebraic, 5.1.11 implies that $\mathbb{K} \leq \mathbb{K}[A]$ is finite. Also $m \in \mathbb{K}[A][x]$ and so b is algebraic over $\mathbb{K}[A]$. Hence (by 5.1.2) $\mathbb{K}[A] \leq \mathbb{K}[A][b]$ is finite. By 5.1.4c, $\mathbb{K} \leq \mathbb{K}[A][b]$ is finite and so by 5.1.10 also algebraic. Thus b is algebraic over $\mathbb{K}$. □

Proposition 5.1.13. *Let* $\mathbb{K}$ *be a field and* P *a set of non constant polynomials over* $\mathbb{K}$*. Then there exists an algebraic extension* $\mathbb{K} \leq \mathbb{F}$ *such that each* $f \in P$ *has a root in* $\mathbb{F}$*.*

Proof. Suppose first that P is finite. Put $f = \prod_{g \in P} g$. 5.1.8(c), there exists a finite extension $\mathbb{E}$ of $\mathbb{K}$ such that f splits over $\mathbb{E}$. Then each $g \in P$ has a root in $\mathbb{E}$.

In the general case let $R = \mathbb{K}[x_f, f \in P]$ be the polynomial ring of P over $\mathbb{K}$. Let I be the ideal in R generated by $f(x_f), f \in P$. We claim that $I \neq R$. So suppose that $I = R$, then $1 \in I$ and so $1 = \sum_{f \in P} r_f f(x_f)$ for some $r_f \in R$, almost all 0. Note that each r_f only involves finitely many $x_g, g \in P$. Hence there exists a finite subset J of I such that $r_f = 0$ for $f \notin J$. Thus $r_f \in \mathbb{K}[x_g, g \in J$ for all $f \in J$. Therefore

$$(*) \qquad 1 = \sum_{f \in J} r_f f(x_f).$$

On the other hand by the finite case there exists a field extension $\mathbb{K} \leq \mathbb{E}$ such that each $f \in J$ has a root $e_f \in \mathbb{E}$. Let

$$\Phi : \mathbb{K}[x_g, g \in J] \to \mathbb{E}$$

be the unique ring homomorphism with $\Phi(x_f) = e_f$ for all $f \in J$ and $\Phi(k) = k$ for all $k \in K$. Since $f(x_f) = \sum_{i=0}^{n} k_i x_f^i$ for some $k_i \in \mathbb{K}$ we have $\Phi(f(x_f)) = \sum_{i=0}^{n} k_i e_f^i = f(e_f) = 0$. So applying Φ to (*) we get

$$1 = \Phi(1) = \sum_{f \in J} \Phi(r_f) f(a_f) = 0$$

a contradiction.

Hence $I \neq R$ and by 3.2.15 I is contained in a maximal ideal M of R. Put $\mathbb{F} = R/M$. Then by 3.2.18 $\mathbb{F}$ is a field. Since $M \neq R$, M contains no units. Thus $\mathbb{K} \cap M = 0$. Thus the map $\mathbb{K} \to \mathbb{F}, k \to k + M$ is a 1-1 ring homomorphism. So we may view $\mathbb{K}$ as a subfield of $\mathbb{F}$ by identifying k with $k + M$. Put $a_f = x_f + M$. Then $f(a_f) = f(x_f) + M$. But $f(x_f) \in I \subseteq M$ and so $f(a_f) = M = 0_{\mathbb{F}}$. □

Lemma 5.1.14. *Let $\mathbb{K}$ be a field. Then the following statements are equivalent.*

(a) Every polynomial over $\mathbb{K}$ splits over $\mathbb{K}$.

(b) Every non-constant polynomial over $\mathbb{K}$ has a root in $\mathbb{K}$.

(c) $\mathbb{K}$ has no proper algebraic extension (that is if $\mathbb{K} \leq F$ is an algebraic extension, then $\mathbb{K} = F$.)

(d) $\mathbb{K}$ has no proper finite extension (that is if $\mathbb{K} \leq F$ is a finite extension, then $\mathbb{K} = F$.)

Proof. (a) $\Longrightarrow$ (b): Obvious.

(b) $\Longrightarrow$ (c): Let $\mathbb{K} \leq \mathbb{E}$ be algebraic and $e \in \mathbb{E}$. Then by (b), $m_e^{\mathbb{K}}$ has a root $k \in \mathbb{K}$. Since $m_e^{\mathbb{K}}$ is irreducible we get $m_e^{\mathbb{K}} = x - a$. Since e is a root of $m_e^{\mathbb{K}}$, $e = k \in \mathbb{K}$. Thus $\mathbb{K} = \mathbb{E}$.

(c) $\Longrightarrow$ (d): Just observe that by 5.1.10(a), every finite extension is algebraic.

(d) $\Longrightarrow$ (a): Let $f \in K$. By 5.1.8 f has a root a in some finite extension $\mathbb{E}$ of K. By assumption $\mathbb{E} = \mathbb{K}$. So $a \in \mathbb{K}$ and (a) holds. □

Definition 5.1.15. *Let $\mathbb{K}$ be a field.*

(a) $\mathbb{K}$ is algebraically closed *if $\mathbb{K}$ fulfills one (and so all) of four equivalent statement in 5.1.14.*

(b) An algebraic closure *of $\mathbb{K}$ is a algebraically closed, algebraic extension of $\mathbb{K}$.*

Lemma 5.1.16. *Let $\mathbb{K} \leq \mathbb{E}$ be an algebraic field extension. Then the following two statements are equivalent.*

(a) $\mathbb{E}$ is an algebraic closure of $\mathbb{K}$.

(b) Every polynomials over $\mathbb{K}$ splits in $\mathbb{E}$.

Proof. If $\mathbb{E}$ is algebraic closed, every polynomial over $\mathbb{E}$ and so also every polynomial over $\mathbb{K}$ splits over $\mathbb{E}$. Thus (a) implies (b).

So suppose (a) holds. Let $\mathbb{F}$ be an algebraic extension of $\mathbb{E}$. Let $a \in \mathbb{F}$. Since $\mathbb{K} \leq \mathbb{E}$ and $\mathbb{E} \leq \mathbb{K}$ are algebraic we conclude from 5.1.12 that $\mathbb{K} \leq \mathbb{F}$ is algebraic. Thus $m_a^{\mathbb{K}}$ is not zero and has a as a root. By assumption, $m_a^{\mathbb{K}}$ splits over $\mathbb{E}$ and so $a \in \mathbb{E}$. Thus $\mathbb{E} = \mathbb{F}$. Hence by 5.1.14, $\mathbb{E}$ is algebraically closed. □

Theorem 5.1.17. *Every field has an algebraic closure.*

Proof. Let $\mathbb{K}_0$ be a field. By 5.1.13 applied to the set of non-constant polynomials there exists an algebraic field extension $\mathbb{K}_1$ of $\mathbb{K}_0$ such that every non-zero polynomial over $\mathbb{K}_0$ has a root in $\mathbb{K}_1$. By induction there exist fields

$$\mathbb{K}_0 \leq \mathbb{K}_1 \leq \mathbb{K}_2 \leq \mathbb{K}_3 \leq \dots$$

such that that every non zero polynomial in $\mathbb{K}_i$ has a root in $\mathbb{K}_{i+1}$. Let $\mathbb{E} = \bigcup_{i=0}^{\infty} \mathbb{K}_i$. By A.7 $\mathbb{E}$ is a field. By 5.1.12 each $\mathbb{K}_i$ is algebraic over $\mathbb{K}_0$. So also $\mathbb{K}_0 \leq \mathbb{E}$ is algebraic. Let $f \in \mathbb{E}[x]$. Then $f \in \mathbb{K}_i[x]$ for some i. Hence f has a root in $\mathbb{K}_{i+1}$ and so in $\mathbb{E}$. Thus by 5.1.14 $\mathbb{E}$ is algebraically closed. □

Definition 5.1.18. *Let $\mathbb{K}$ be a field and P a set of polynomials over $\mathbb{K}$. A* splitting field *for P over $\mathbb{K}$ is an extension $\mathbb{E}$ of $\mathbb{K}$ such that*

(a) Each $f \in P$ splits over $\mathbb{E}$.

(b) $\mathbb{E} = \mathbb{K}[A]$, where $A := \{a \in \mathbb{E} \mid f(a) = 0$ for some $0 \neq f \in P\}$.

Corollary 5.1.19. *Let $\mathbb{K}$ be a field and P a set of polynomials over $\mathbb{K}$. Then there exists a splitting field for P over $\mathbb{K}$.*

Proof. Let $\overline{\mathbb{K}}$ be a algebraic closure for $\mathbb{K}$, $B := \{a \in \bar{\mathbb{K}} \mid f(a) = 0$ for some $f \in P\}$ and put $\mathbb{E} = \mathbb{K}[B]$. Then $\mathbb{E}$ is a splitting field for P over $\mathbb{K}$. □

5.2 Splitting fields, Normal Extensions and Separable Extensions

Lemma 5.2.1. *Let $\phi : \mathbb{K}_1 \to \mathbb{K}_2$ be a 1-1 homomorphism of fields. Then*

(a) There exists a unique homomorphism $\tilde{\phi} : \mathbb{K}_1[x] \to \mathbb{K}_2[x]$ with $\tilde{\phi}(k) = \phi(k)$ and $\tilde{\phi}(x) = x$.

(b) $\tilde{\phi}\left(\sum_{i=0}^{\infty} a_i x^i\right) = \sum_{i=0}^{\infty} \phi(a_i) x^i$ for all $\sum_{i=0} a_i x^i \in \mathbb{K}_1[x]$

(c) $\tilde{\phi}$ is 1-1 and if ϕ is an isomorphism, $\tilde{\phi}$ is an isomorphism.

We will usually just write $\tilde{\phi}$ for ϕ.

Proof. (a) and (b) follow from 3.2.5. (c) is readily verified. □

Lemma 5.2.2. *Let $\phi : \mathbb{K}_1 \to \mathbb{K}_2$ be an isomorphism of fields and for $i = 1$ and 2 let $\mathbb{K}_i \leq \mathbb{E}_i$ be a field extension. Let $f_1 \in \mathbb{K}_1[x]$ be irreducible and put $f_2 = \phi(f_1)$. Suppose e_i is a root of f_i in $\mathbb{K}_i$. Then there exists a unique isomorphism $\psi : \mathbb{K}_1[e_1] \to \mathbb{K}_2[e_2]$ with $\psi \mid_{\mathbb{K}_1} = \phi$ and $\psi(e_1) = e_2$.*

Proof. Using 5.1.8(a) we have the following three isomorphism:

$$\begin{array}{ccccccc} \mathbb{K}_1[e_1] & \cong & \mathbb{K}_1[x]/f_1\mathbb{K}_1[x] & \cong & \mathbb{K}_2[x]/f_2\mathbb{K}_2[x] & \cong & \mathbb{K}_2[e_2] \\ g(e_1) & \to & g+f_1\mathbb{K}_1[x] & \to & \phi(g)+f_2\mathbb{K}_2[x] & \to & \phi(g)(e_2) \end{array}$$

Let ψ be the composition of these three isomorphism. Then

$$\psi : e_1 \to x + f_1\mathbb{K}_1[x] \to x + f_2\mathbb{K}_2[x] \to e_2$$

and for $k \in \mathbb{K}_1$,

$$\psi : k \to k + f_1\mathbb{K}_1[x] \to \phi(k) + f_2\mathbb{K}_2[x] \to \phi(k)$$

Thus shows the existence of ψ. If $\tilde{\psi}$ is any such ring homomorphism then

$$\tilde{\psi}\left(\sum_{i=0}^{\deg f-1} a_i e_1^i\right) = \sum_{i=0}^{\deg f-1} \phi(a_i)e_2^i$$

and so ψ is unique. □

Definition 5.2.3. *Let $\mathbb{K}$ be a field and F and E extensions of $\mathbb{K}$.*

(a) A $\mathbb{K}$-homomorphism *from F to E is a $\mathbb{K}$-linear ring homomorphism from F to E.* $\mathbb{K}$-homomorphism. $\mathbb{K}$-*isomorphism and* $\mathbb{K}$-*automorphism are defined similarly.*

(b) $\mathrm{Aut}(\mathbb{K})$ *is the set of automorphism of $\mathbb{K}$ and* $\mathrm{Aut}_{\mathbb{K}}(F)$ *is the set of $\mathbb{K}$-automorphism of $\mathbb{F}$.*

(c) $\mathbb{E}$ is an intermediate field of the extension $\mathbb{K} \leq F$ if $\mathbb{K}$ is a subfield of $\mathbb{E}$ and $\mathbb{E}$ is a subfield of F.

Lemma 5.2.4. *Let $\phi : \mathbb{F} \to \mathbb{K}$ a non-zero homomorphisms of fields. Then ϕ is 1-1 and $\phi(1_{\mathbb{F}}) = 1_{\mathbb{E}}$.*

Proof. Since ϕ is non-zero, $\ker\phi \neq 0$. Since $\ker\phi$ is an ideal and $\mathbb{F}$ has no proper ideals, $\ker\phi = 0$ and so ϕ is 1-1.

We have

$$\phi(1_{\mathbb{F}})\phi(1_F) = \phi(1_{\mathbb{F}}\,1_{\mathbb{F}}) = \phi(1_{\mathbb{F}}) = 1_{\mathbb{E}}\phi(1_{\mathbb{F}}).$$

Since ϕ is 1-1, $\phi(1_{\mathbb{F}}) \neq 0_{\mathbb{E}}$ and so the Cancellation Law implies $\phi(1_{\mathbb{F}}) = 1_{\mathbb{E}}$ □

Lemma 5.2.5. *Let $\mathbb{K} \leq \mathbb{F}$ and $\mathbb{K} \leq \mathbb{E}$ be field extensions and $\phi : \mathbb{F} \to \mathbb{K}$ a non-zero ring homomorphism. Then ϕ is $\mathbb{K}$-linear if and only if $\phi(k) = k$ for all $k \in \mathbb{K}$.*

Proof. Let $k \in \mathbb{K}$ and $a \in \mathbb{F}$. If ϕ is $\mathbb{K}$-linear

$$\phi(k) = \phi(k1_F) = k\phi(1_F) = k1_E = k$$

and if $\phi(k) = k$ then

$$\phi(ka) = \phi(k)\phi(a) = k\phi(a).$$

□

Lemma 5.2.6. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension. Then $\mathrm{Aut}(\mathbb{F})$ is a subgroup of $\mathrm{Sym}(\mathbb{F})$ and $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ is a subgroup of $\mathrm{Aut}(\mathbb{F})$.*

Proof. Readily verified. □

Lemma 5.2.7. *Let $\mathbb{F}$ be a field and $f \in \mathbb{F}[x]$ a non-zero polynomial.*

Then there an integer m with $0 \leq m \leq \deg f$, $a_1, \ldots a_m \in F$ and $q \in \mathbb{F}[x]$ such that

(a) $f = q \cdot (x - a_1) \cdot (x - a_2) \cdot (x - a_m)$.

(b) q has no roots in F.

(c) $\{a_1, a_2, \ldots a_m\}$ is the set of roots of f.

In particular, the number of roots of f is at most $\deg f$.

Proof. Suppose that f has no roots. Then the theorem holds with $q = f$ and $m = 0$.

The proof is by induction on $\deg f$. Since polynomials of degree 0 have no roots, the theorem holds if $\deg f = 0$.

Suppose now that theorem holds for polynomials of degree k and let f be a polynomial of degree $k + 1$. If f has no root we are done by the above. So suppose f has a root a. By 3.4.3 there exists $g, r \in \mathbb{F}[x]$ with $f = g \cdot (x - a) + t$ and $\deg r < \deg(x - a) = 1$. Thus $r \in \mathbb{F}$ and $0 = f(a) = g(a) \cdot (a - a) + r$. Thus $r = 0$ and

$$(*) \qquad f = g \cdot (x - a)$$

Then $\deg g = k$ and so by the induction assumption there exists an integer n with $0 \leq n \leq \deg g$, $a_1, \ldots a_n \in F$ and $q \in F[x]$ such that

(i) $g = q \cdot (x - a_1) \cdot (x - a_2) \cdot (x - a_n)$

(ii) q has no roots in F.

(iii) $\{a_1, a_2, \ldots a_n\}$ is the set of roots of g.

Put $m = n + 1$ and $a_m = a$. From $f = g \cdot (x - a) = g \cdot (x - a_m)$ and (i) we conclude that (a) holds. By (ii), (b) holds.

Let $b \in F$. Then b is a root if and only if $f(b) = 0_R$ and so by (*) if and only $g(b)(b - a) = 0_F$. Since F is an integral domain this holds if and only if $g(b) = 0$ or $b - a = 0_F$. From $a = a_m$ and (iii) we conclude that the roots of f are $\{a_1, a_2 \ldots, a_m\}$. So also (c) holds. □

Lemma 5.2.8. *Let $\mathbb{K}$ be a field field and P a set of polynomials. Let $\mathbb{E}_1$ and $\mathbb{E}_2$ be splitting fields for P over $\mathbb{K}$*

(a) For $i = 1, 2$ let $\mathbb{L}_i$ be an intermediate field of $\mathbb{K} \leq \mathbb{E}_i$ and let $\delta : \mathbb{L}_1 \to \mathbb{L}_2$ be a $\mathbb{K}$-isomorphism. Then there exists a $\mathbb{K}$-isomorphism $\psi : \mathbb{E}_1 \to \mathbb{E}_2$ with $\psi|_{\mathbb{L}_i} = \delta$.

(b) $\mathbb{E}_1$ and $\mathbb{E}_2$ are $\mathbb{K}$-isomorphic.

(c) Let $f \in \mathbb{K}[x]$ be irreducible and suppose that, for $i = 1$ and 2, e_i is a root of f in $\mathbb{E}_i$. Then there exists a $\mathbb{K}$-isomorphism $\psi : \mathbb{E}_1 \to \mathbb{E}_2$ with $\psi(e_1) = \psi(e_2)$.

(d) Let $f \in \mathbb{K}[x]$ be irreducible and let e and d be roots of f in $\mathbb{E}_1$. Then there exists $\psi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{E}_1)$ with $\psi(e) = d$.

(e) Any two algebraic closures of $\mathbb{K}$ are $\mathbb{K}$-isomorphic.

Proof. Let $\mathcal{M}$ be the set of all $\mathbb{K}$-linear isomorphism $\phi : \mathbb{F}_1 \to \mathbb{F}_2$ where, for $i = 1$ and 2, $\mathbb{F}_i$ is an intermediate field of $\mathbb{K} \leq \mathbb{E}_i$. Order $\mathcal{M}$ by $\phi \leq \psi$ if ϕ is the restriction of ψ to a subfield. Let $\mathcal{M}^* = \{\phi \in \mathcal{M} \mid \delta \leq \phi\}$. Since $\delta \in \mathcal{M}^*$, $\mathcal{M}^*$ is not empty.

It is easy to verify that $\leq$ is a partial ordering. Let $\mathcal{C} = \{\psi_s : \mathbb{F}_{s1} \to \mathbb{F}_{s2} \mid s \in S\}$ be a chain in $\mathcal{M}^*$. Define $\mathbb{F}_i = \bigcup_{s \in S} \mathbb{F}_{si}$ and $\phi : \mathbb{F}_1 \to \mathbb{F}_2, a \to \phi_s(a)$, whenever $a \in \mathbb{F}_{s1}$. It is straightforward to check that $\mathbb{F}_i$ is a field, ϕ is well-defined and ϕ is a isomorphism. Moreover $\phi_s \leq \phi$ for all $s \in S$ and so ϕ is an upper bound for $\mathcal{C}$.

Thus by Zorn's Lemma A.6 $\mathcal{M}^*$ has a maximal element $\phi : \mathbb{F}_1 \to \mathbb{F}_2$. It remains to show that $\mathbb{F}_i = \mathbb{E}_i$. For this put

$$A_i = \{e_i \in \mathbb{E}_i \mid f(e_i) = 0 \text{ for some } 0 \neq\in P\}$$

By definition of a splitting field, $\mathbb{E}_i = \mathbb{K}[A_i]$. Since $\mathbb{K} \leq \mathbb{F}_i \leq \mathbb{E}_i$ we just need to show that $A_i \subseteq \mathbb{F}_i$.

So let $e_1 \in A_1$ and $0 \neq f \in P$ with $f(e_1) = 0$. Let $f_1 \in \mathbb{F}_1[x]$ be an irreducible divisor of f with $f_1(e_1) = 0$. Put $f_2 = \phi(f_1)$. Since f_1 divides f in $\mathbb{F}_1[x]$, f_2 divides $\phi(f)$ in $\mathbb{F}_2[x]$. Since $f \in \mathbb{K}[x]$ and ϕ is a $\mathbb{K}$-homomorphism, $\phi(f) = f$. Thus f_2 divides f. Since f splits over $\mathbb{E}_2$, also f_2 splits over $\mathbb{E}_2$ and so f_2 has a root $e_2 \in \mathbb{E}_2$. By 5.2.2 there exists a field isomorphism $\psi : \mathbb{F}_1[e_1] \to \mathbb{F}_2[e_2]$ with $\psi|_{\mathbb{F}_1} = \phi$. Hence by maximality of ϕ, $\mathbb{F}_1 = \mathbb{F}_1[e_1]$. Thus $e_1 \in \mathbb{F}_1$. So $A_1 \subseteq \mathbb{F}_1$ and $\mathbb{F}_1 = \mathbb{E}_1$. Hence $\mathbb{F}_1$ is a splitting field for P over $\mathbb{K}$. Since ϕ is a $\mathbb{K}$-isomorphism we conclude that $\mathbb{F}_2$ is a splitting field for $P = \phi(P)$ over $\mathbb{K}$. Since $\mathbb{F}_2 \subseteq \mathbb{E}_2$ this implies $A_2 \subseteq \mathbb{F}_2$ and $\mathbb{F}_2 = \mathbb{E}_2$.

(b) Apply (b) to $\delta = \mathrm{id}_{\mathbb{K}}$.

(c) By 5.2.2 there exists a $\mathbb{K}$-linear isomorphisms $\delta : \mathbb{K}[e_1] \to \mathbb{K}[e_2]$ with $\delta(e_1) = e_2$. By (a) δ can be extended to an isomorphism $\psi : \mathbb{E}_1 \to \mathbb{E}_2$. So (a) holds.

(d) Follows from (c) with $\mathbb{E}_2 = \mathbb{E}_1$.

(e) Follows from (b) with P the set of all polynomials. □

Definition 5.2.9. *Let $\mathbb{E} \leq \mathbb{F}$ and $H \leq \mathrm{Aut}(\mathbb{F})$.*

(a) $\mathbb{E}$ is called H-stable *if $h(e) \in \mathbb{E}$ for all $h \in H, e \in \mathbb{E}$.*

(b) 'stable' means G-stable.

(c) If $\mathbb{E}$ is H-stable, then $H^{\mathbb{E}} := \{h|_{\mathbb{E}} \mid h \in H\}$.

(d) $\mathbb{E} \leq \mathbb{F}$ is called normal *if it is algebraic and each irreducible $f \in \mathbb{E}[x]$, which has a root in $\mathbb{F}$, splits over $\mathbb{F}$.*

Lemma 5.2.10. *(a) Let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ be field extensions and suppose that $\mathbb{E}$ is the splitting field for some set P of polynomials over $\mathbb{K}$. Then $\mathbb{E}$ is $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ stable.*

(b) $\mathbb{K} \leq \mathbb{E}$ is normal if and only if $\mathbb{E}$ is a splitting field of some set of polynomials over $\mathbb{K}$.

Proof. (a) Let $A = \{e \in \mathbb{E} \mid f(e) = 0 \text{ for some } 0 \neq f \in P\}$. Let $f \in P$, e a root of f in $\mathbb{E}$ and $\phi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$. Then $\phi(e)$ is a root of $\phi(f) = f$ and as f splits over $\mathbb{E}$, $\phi(e) \in \mathbb{E}$. Thus $\Phi(A) \subseteq \mathbb{E}$. By definition of a splitting field, $\mathbb{E} = \mathbb{K}[A]$ and so $\phi(\mathbb{E}) = \phi(\mathbb{K})[\phi(A)] = \mathbb{K}[\phi(A)] \leq \mathbb{E}$. So $\mathbb{E}$ is $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$-stable.

(b) If $\mathbb{K} \leq \mathbb{E}$ is normal, $\mathbb{E}$ is the splitting field of the set of polynomial over $\mathbb{K}$ with roots in $\mathbb{E}$.

So suppose that $\mathbb{E}$ is the splitting field for some set P of polynomials over $\mathbb{K}$. Let $e \in \mathbb{E}$ and $f = m_e^{\mathbb{K}}$. Let $\mathbb{F}$ be a splitting field for f over $\mathbb{E}$ and let d be a root of f in $\mathbb{F}$. Note that $\mathbb{F}$ is splitting field for $P \cup \{f\}$ over $\mathbb{K}$ and so by 5.2.8(d) there exists $\phi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ with $\phi(e) = d$. By (a) $d = \phi(e) \leq \mathbb{E}$ and so $d \in \mathbb{E}$. Hence f splits over $\mathbb{E}$ and $\mathbb{K} \leq \mathbb{E}$ is normal. □

Lemma 5.2.11. *Let $\mathbb{K} \leq \mathbb{L}$ be an algebraic field extension and $\mathbb{E}$ and $\mathbb{F}$ intermediate fields. Suppose that $\mathbb{K} \leq \mathbb{E}$ is normal, then $m_b^{\mathbb{F}} = m_b^{\mathbb{F} \cap \mathbb{E}}$ for all $b \in \mathbb{E}$.*

Proof. Let $g = m_b^{\mathbb{F}}$ and $f = m_b^{\mathbb{K}}$. As $\mathbb{K} \leq \mathbb{E}$ is normal and b is a root of f in $\mathbb{E}$, f splits over $\mathbb{E}$. Since f is monic we conclude

$$f = (x - e_1)(x - e_2) \dots (x - e_n)$$

with $e_i \in \mathbb{E}$. By 5.1.6, g divides f in $\mathbb{L}[x]$ and so $g = (x - e_{i_1})(x - e_{i_2}) \dots (x - e_{i_k})$ for some $1 \leq i_1 < \dots i_l \leq n$. This implies $g \in \mathbb{E}[x]$ and so $g \in (\mathbb{E} \cap \mathbb{F})[x]$. Since g is irreducible in $\mathbb{F}[x]$ it is also irreducible in $(\mathbb{E} \cap \mathbb{F})[x]$. Since g is monic and has b as a root we conclude that $g = m_b^{\mathbb{F} \cap \mathbb{E}}$.

□

Definition 5.2.12. *Let $\mathbb{K}$ be a field and $f \in \mathbb{K}[x]$.*

(a) Let $\mathbb{E}$ a splitting field of f over $\mathbb{K}$ and e a root of f in $\mathbb{E}$. Let $m \in \mathbb{N}$ be maximal such that $(x-e)^m$ divides f $\mathbb{E}[x]$. Then m is called the multiplicity *of e as a root of f. If $m > 1$, then e is called a* multiple root *of f.*

(b) If $f = \sum_{i=0}^{n} k_i x^i$, then $f' := \sum_{i=1}^{n} i k_i x^{i-1}$. f' is called the derivative *of f.*

(c) For $n \in \mathbb{N}$ define $f^{[n]}$ inductively by $f^{[0]} = f$ and $f^{[n+1]} = (f^{(n)})'$. $f^{[n]}$ is called the n-th derivative of f.

Lemma 5.2.13. *Let $\mathbb{K}$ be a field and $f, g \in \mathbb{K}[x]$*

(a) The derivative function $\mathbb{K}[x] \to \mathbb{K}[x], f \to f'$ is $\mathbb{K}$-linear.

(b) $(fg)' = f'g + fg'$.

(c) For all $n \in \mathbb{Z}^+$, $(f^n)' = nf^{n-1}f'$.

Proof. (a) is obvious. (b) By (a) we may assume that $f = x^m$ and $g = x^n$

$$(fg)' = (x^{n+m})' = (n+m)x^{n+m-1}$$
$$f'g + fg' = mx^{m-1}x^n + x^m nsx^{n-1} = (n+m)x^{m+n-1}$$

Thus (b).

(c) Follows from (b) and induction on n. □

Lemma 5.2.14. *Let $\mathbb{K}$ be a field, $f \in \mathbb{K}[x]$ and c a root of f.*

(a) Suppose that $f = g \cdot (x-c)^n$ for some $n \in \mathbb{N}$ and $g \in \mathbb{K}[x]$. Then $f^{[n]}(c) = n!g(c)$.

(b) c is a multiple root of f if and only if $f'(c) = 0$.

(c) Suppose that $\operatorname{char}\mathbb{K} = 0$ or $\deg f < \operatorname{char}\mathbb{K}$. Then the multiplicity of c as a root of f is smallest $m \in \mathbb{N}$ with $f^{[m]}(c) \neq 0$.

Proof. (a) We will show that for all $0 \leq i \leq n$, there exists $h_i \in \mathbb{K}[x]$ with

$$(*) \qquad f^{[i]} = \frac{n!}{(n-i)!} \cdot g \cdot (x-c)^{n-i} + h_i \cdot (x-c)^{n-i+1}$$

For $i = 0$ this is true with $h_0 = 0$. So suppose its true for i. Then using 3.7.11

$$f^{[i+1]} = (f^{[i]})' =$$
$$\tfrac{n!}{(n-i)!}\left(g' \cdot (x-c)^{n-i} + g \cdot (n-i) \cdot (x-c)^{n-i-1}\right) + h_i' \cdot (x-c)^{n-i+1} + h_i \cdot (n-i+1)(x-c)^{n-i}$$
$$= \tfrac{n!}{(n-i-1)!} \cdot g \cdot (x-c)^{n-i-1} + \left(\tfrac{n!}{(n-i)!} \cdot g \cdot (n-i) + h_i' \cdot (x-c) + (n-i+1) \cdot h_i\right) \cdot (x-c)^{n-i}$$

Thus (*) also hold for $i+1$ and so for all $i \in \mathbb{N}$.

For $i = n$ we conclude $f^{[n]} = n!g + h_n \cdot (x-a)$. Thus (a) holds.

(b) Since c is a root, $f = g \cdot (x-a)$ for some $g \in \mathbb{K}[x]$. Then c is a multiple root of f if and only if $g(c) = 0$. By (a) applied with $n = 1$, $f'(c) = g(c)$. Thus (b) holds.

(c) Let m the multiplicity of c as a root of f. So $f = g \cdot (x-c)^m$ for some $g \in \mathbb{K}[x]$ with $g(c) \neq 0$. Let $n < m$. From $f = g \cdot (x-c)^{m-n} \cdot (x-c)^n$ and (a) we get $f^{[n]}(c) = 0$. As $m \leq \deg f$ we have $\operatorname{char} \mathbb{K} = 0$ or $m < \operatorname{char} \mathbb{K}$. Since $g(c) \neq 0$ we get $m!g(c) \neq 0$. Hence by (a) $f^{[m]}(c) = m!g(c) \neq 0$ and (c) holds. □

Example 5.2.15. Consider the polynomial $f = x^p$ in $\mathbb{Z}_p[x]$. Then 0 is a root of multiplicity p of f. But then $f' = px^{p-1} = 0$. This shows that the condition on $(\deg f)!$ in part c the previous theorem is necessary.

Definition 5.2.16. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension.*

(a) An irreducible polynomial $f \in \mathbb{F}[x]$ is called separable *over $\mathbb{E}$ if f has no multiple roots. An arbitrary polynomial in $\mathbb{F}[x]$ is called* separable *over $\mathbb{F}$ if all its irreducible factors are separable over $\mathbb{F}$.*

(b) $b \in \mathbb{F}$ is called separable *over $\mathbb{K}$, if b is algebraic over $\mathbb{K}$ and $m_b^{\mathbb{K}}$ is separable over $\mathbb{K}$. $\mathbb{K} \leq \mathbb{F}$ is called separable if each $b \in \mathbb{F}$ is separable over $\mathbb{K}$.*

Lemma 5.2.17. *Let $\mathbb{K}$ be a field, $\overline{\mathbb{K}}$ an algebraic closure of $\mathbb{K}$ and suppose that $\operatorname{char} \mathbb{K} = p$ with $p \neq 0$.*

(a) For each $n \in \mathbb{Z}^+$, the map $\operatorname{Frob}_{p^n} : \mathbb{K} \to \mathbb{K}, k \to k^{p^n}$ is a 1-1 ring homomorphism.

(b) For each $b \in \mathbb{K}$ and $n \in \mathbb{Z}^+$ there exists a unique $d \in \bar{K}$ with $d^{p^n} = b$. We will write $b^{p^{-n}}$ for d.

(c) If $\mathbb{K}$ is algebraically closed then each $\operatorname{Frob}_{p^n}$, $n \in \mathbb{Z}$ is an automorphism.

(d) For each $n \in \mathbb{Z}$, the map $\operatorname{Frob}_{p^n} : \mathbb{K} \to \overline{\mathbb{K}}, k \to k^{p^n}$ is a 1-1 ring homomorphism.

(e) If $f \in \mathbb{K}[x]$ and $n \in \mathbb{N}$, then $f^{p^n} = \operatorname{Frob}_{p^n}(f)(x^{p^n})$.

Proof. (a) Clearly $(ab)^p = a^p b^p$. Note that p divides $\binom{p}{i}$ for all $1 \leq i < p$. So by the Binomial Theorem 3.1.17 $(a+b)^p = a^p + b^p$. So Frob_p is a field homomorphism. If $a \in \mathbb{K}$ with $a^p = 0$, then $a = 0$. So $\ker \operatorname{Frob}_p = 0$ and Frob_p is 1-1. Since $\operatorname{Frob}_{p^n} = \operatorname{Frob}_p^n$, (a) holds.

(b) Let d be a root of $x^{p^n} - b = 0$. Then $d^{p^n} = b$. The uniqueness follows from (a).

(c) For all $n \in \mathbb{Z}$, $\operatorname{Frob}_{p^{-n}}$ is the inverse of $\operatorname{Frob}_{p^n}$. So $\operatorname{Frob}_{p^n}$ is a bijective. For $m \in \mathbb{N}$, $\operatorname{Frob}_{p^m}$ is a ring homomorphism and all $\operatorname{Frob}(p^n)$ are automorphism.

(d) Follows from (c).

(e) Let $f = \sum a_i x^i$. Then $\operatorname{Frob}_{p^n}(f) = \sum a_i^{p^n} x^i$ and so

$$\operatorname{Frob}_{p^n}(f)(x^{p^n}) = \sum a_i^{p^n} x^{p^n i} = \left(\sum a_i x^i\right)^{p^n} = f^{p^n}$$

□

Example 5.2.18. Let $\mathbb{K} = \mathbb{Z}_p(x)$, the field of fractions of the polynomial ring $\mathbb{Z}_p[x]$. If $a \in \mathbb{Z}_p$, then $a^p = a$. (Indeed since $(\mathbb{Z}_p^\sharp, \cdot)$ is a group of order $p-1$, $a^{p-1} = 1$ for all $a \in \mathbb{Z}_p^\sharp$. Thus $a^p = a$). It follows that $f^p = f(x^p)$ for all $f \in \mathbb{Z}_p[x]$. Hence

$$\mathrm{Frob}_p(\mathbb{K}) = \left\{ \frac{f(x^p)}{g(x^p)} \middle| f, g \in \mathbb{Z}_p[x], g \neq 0 \right\} = \mathbb{Z}_p(x^p)$$

So $\mathbb{Z}_p(x^p)$ is a proper subfield of $\mathbb{Z}_p(x)$ isomorphic to $\mathbb{Z}_p(x)$.

Let $\overline{\mathbb{K}}$ be an algebraic closure of $\mathbb{K}$. Consider the polynomial ring $\mathbb{K}[t]$ over $\mathbb{K}$ in the indeterminate t and $f = t^p - x \in \mathbb{K}[t]$. We claim that f is irreducible. Note that $x^{\frac{1}{p}}$ is a root of f in $\overline{\mathbb{K}}$ and $f = (t - x^{\frac{1}{p}})^p$. Let g be a non-constant monic polynomial in $\mathbb{K}[x]$ dividing f. Then $g = (t - x^{\frac{1}{p}})^k$ for some $1 \leq k \leq p$. Then $x^{\frac{k}{p}} = \pm g(0) \in \mathbb{K}$ and so $k = p$. Thus f is irreducible. Since $x^{\frac{1}{p}}$ is a root of multiplicity p of f, f is not separable.

Lemma 5.2.19. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension such that $p := \mathrm{char}\,\mathbb{K} \neq 0$ and $b \in \mathbb{F}$. Suppose that $b^{p^n} \in \mathbb{K}$ for some $n \in \mathbb{N}$. Then*

(a) b is the only root of m_b^K (in any algebraic closure of $\mathbb{K}$.)

(b) If b is separable over $\mathbb{K}$, $b \in \mathbb{K}$.

(c) $d^{p^n} \in \mathbb{K}$ for all $d \in \mathbb{K}[b]$.

Proof. Put $q = p^n$.

(a) Note that b is a root of $x^q - b^q$, so by 5.1.5 $m_b^{\mathbb{K}}$ divides $x^q - b^q = (x-b)^q$. Thus (a) holds.

(b) If m_b^K is separable, we conclude from (a) that $m_b^{\mathbb{K}} = x - b$. Thus $b \in \mathbb{K}$.

(c) Let $\phi = \mathrm{Frob}_q$. Then $\{d^q \mid, d \in \mathbb{K}[b]\} = \phi([\mathbb{K}[b]) = \phi(\mathbb{K})[\phi(b)] \leq \mathbb{K}[b^q] \leq \mathbb{K}$. □

Lemma 5.2.20. *Let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ be field extensions and $b \in \mathbb{F}$. If $b \in \mathbb{F}$ is separable over $\mathbb{K}$, then b is separable over $\mathbb{E}$*

Proof. By 5.1.6 $m_b^{\mathbb{E}}$ divides $m_b^{\mathbb{K}}$. As b is separable over $\mathbb{K}$, $m_b^{\mathbb{K}}$ has no multiple roots. So also $m_b^{\mathbb{E}}$ has no multiple roots and b is separable over $\mathbb{E}$. □

Lemma 5.2.21. *Let $\mathbb{K}$ be a field and let $f \in \mathbb{K}[x]$ be monic and irreducible.*

(a) f is separable if and only if $f' \neq 0$.

(b) If $\mathrm{char}\,\mathbb{K} = 0$, all polynomials over $\mathbb{K}$ are separable.

Proof. (a) By 3.7.12 b is a multiple root of f if and only if $f'(b) = 0$. Since f is irreducible, $f = m_b^{\mathbb{K}}$. So b is a root of f' if and only if f divides f'. As $\deg f' < \deg f$, this the case if and only if $f' = 0$.

(b) follows from (a). □

Lemma 5.2.22. *Let $\mathbb{K}$ be a field and $f \in \mathbb{K}[x]$ monic and irreducible. Suppose $p := \operatorname{char} \mathbb{K} \neq 0$ and let $b_1, b_2, \dots b_d$ be the distinct roots of f in an algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$. Let b be any root of f. Then there exist an irreducible separable polynomial $g \in \mathbb{K}[x]$, $n \in \mathbb{N}$ and a polynomial $h \in \operatorname{Frob}_{p^{-n}}(\mathbb{K})[x]$ such that*

(a) $f = g(x^{p^n}) = h^{p^n}$.

(b) $g = \operatorname{Frob}_{p^n}(h)$.

(c) $g = (x - b_1^{p^n})(x - b_2^{p^n}) \dots (x - b_d^{p^n})$.

(d) $h = (x - b_1)(x - b_2) \dots (x - b_d) \in \mathbb{K}[b_1, \dots, b_d][x]$.

(e) $f = (x - b_1)^{p^n}(x - b_2)^{p^n} \dots (x - b_d)^{p^n}$.

(f) f *is separable over* $\mathbb{K}$ *if and only if* $n = 0$.

(g) $\dim_{\mathbb{K}[b^{p^n}]} \mathbb{K}[b] = p^n$.

(h) b *is separable over* $\mathbb{K}$ *if and only if* $\mathbb{K}[b] = \mathbb{K}[b^p]$.

(i) b^{p^n} *is separable over* $\mathbb{K}$.

Proof. We will first show that $f = g(x^{p^n})$ for some separable $g \in \mathbb{K}[x]$ and $n \in \mathbb{N}$. If f is separable, this is true with $g = f$ and $n = 0$. So suppose f is not separable. By 5.2.21(a) $f' = 0$. Let $f = \sum a_i x^i$. Then $0 = f' = \sum i a_i x^{i-1}$ and so $i a_i = 0$ for all i. Thus p divides i for all i with $a_i \neq 0$. Put $\tilde{f} = \sum a_{pi} x^i$. Then $\tilde{f}(x^p) = \sum a_{pi} x^{pi} = f$. By induction on $\deg f$, $\tilde{f} = g(x^{p^m})$ for some irreducible and separable $g \in \mathbb{K}[x]$. Let $n = m + 1$, then $f = g(x^{p^n})$.

Since f is irreducible, g is irreducible.

Let $h = \operatorname{Frob}_{p^{-n}}(g) \in \bar{\mathbb{K}}[x]$. Then $g = \operatorname{Frob}_{p^n}(h)$. Thus by 5.2.17(e), $h^{p^n} = g(x^{p^n}) = f$. Let $b \in \bar{K}$. Then b is a root of f if and only if b^{p^n} is a root of g. So $\{b_1^{p^n}, \dots, b_d^{p^n}\}$ is the set of roots of g. As $\operatorname{Frob}_{p^n}$ is one to one, the $b_i^{p_n}$ are pairwise distinct. Since g is separable, $g = \prod\{x - e \mid e \text{ a root of } g\}$ and so (a) holds.

(d) now follows from $h = \operatorname{Frob}_{p^{-n}}(g)$.

(d) By (a) $f = h^{p^n}$ and so (d) implies (e).

(f) follows from (e)

(g) Note that g is the minimal polynomial of $b_i^{p^n}$ over $\mathbb{K}$, f is the minimal polynomial of b_i over $\mathbb{K}$ and $\deg f = p^n \deg g$. Thus

$$\dim_{\mathbb{K}[b^{p^n}]} \mathbb{K}[b] = \frac{\dim_{\mathbb{K}} \mathbb{K}[b]}{\dim_{\mathbb{K}} \mathbb{K}[b^{p^n}]} = \frac{\deg f}{\deg g} = p^n$$

(h) Suppose b is not separable. Then $n > 0$ and so b^p is a root of $g(x^{p^{n-1}})$. So $\dim_{\mathbb{K}} \mathbb{K}[b^p] \leq p^{n-1}$ and $\mathbb{K}[b] \neq \mathbb{K}[b^p]$.

Suppose that b is separable over. Then by 5.2.20 b is separable over $\mathbb{K}[b^p]$. So by 5.2.19, $b \in \mathbb{K}[b^p]$. Thus $\mathbb{K}[b] = \mathbb{K}[b^p]$.

(i) follows since $b_i^{p^n}$ is a root of the separable g. □

Definition 5.2.23. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension Let $b \in \mathbb{F}$. Then b is* purely inseparable *over $\mathbb{K}$ if b is algebraic over $\mathbb{K}$ and b is the only root of $m_b^{\mathbb{K}}$. $\mathbb{K} \leq \mathbb{F}$ is called purely inseparable if all elements in $\mathbb{F}$ are purely inseparable over $\mathbb{K}$.*

Lemma 5.2.24. *Let $\mathbb{K} \leq \mathbb{F}$ be an algebraic field extension with* $\operatorname{char} K = p \neq 0$.

(a) Let $b \in \mathbb{K}$ then b is purely inseparable if and only if $b^{p^n} \in \mathbb{K}$ for some $n \in \mathbb{N}$

(b) Put $\mathbb{S} := \{b \in \mathbb{F} \mid b$ is separable over $\mathbb{K}\}$. $\mathbb{K} \leq \mathbb{F}$ is purely inseparable if and only if $\mathbb{K} = \mathbb{S}$.

(c) Let $\mathbb{P} := \{b \in \mathbb{F} \mid b^{p^n} \in \mathbb{K}$ for some $n \in \mathbb{N}\}$.. Then $\mathbb{P}$ is the of elements in $\mathbb{F}$ purely inseparable over $\mathbb{K}$ and $\mathbb{P}$ is a subfield of $\mathbb{F}$.

(d) If $\mathbb{K} \leq \mathbb{F}$ is normal, then $\mathbb{P} \leq \mathbb{F}$ is separable.

(e) If $b \in \mathbb{F}$ is separable over $\mathbb{K}$, then $m_b^{\mathbb{P}} = m_b^{\mathbb{K}}$.

(f) $\mathbb{P} \leq \operatorname{Fix}_{\mathbb{F}}\operatorname{Aut}_{\mathbb{K}}(\mathbb{F})$ with equality if $\mathbb{K} \leq \mathbb{F}$ is normal.

Proof. Let $b \in \mathbb{F}$. Let $f := m_b^{\mathbb{K}}$. Then by 5.2.22 $f = g(x^{p^n})$ with $g \in \mathbb{K}[x]$ separable. Moreover, if $b_1, b_2, \ldots, b_k$ are the distinct roots of f in an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$,, then $g = (x - b_1^q)(x - b_2^q)\ldots(x - b_k^q)$, where $q = p^n$.

(a) If $b^{p^m} \in \mathbb{K}$ for some $m \in \mathbb{K}$, then by 5.2.19(a), b is the only root of f and so b is purely inseparable over $\mathbb{K}$. If

Suppose b is the only root of f. Then $k = 1$ and $g = x - b^q$. Since $g \in \mathbb{K}[x]$, $b^q \in \mathbb{K}$ So (a) holds.

(b) Suppose first that $\mathbb{K} = \mathbb{S}$. Since b^q is a root of the separable, g, $b^q \in \mathbb{S} = \mathbb{K}$. Thus by (a) , b is purely inseparable over $\mathbb{K}$

Suppose next that $\mathbb{K} \leq \mathbb{F}$ is purely inseparable. Then as seen above $b^q \in \mathbb{K}$. If b is separable over $\mathbb{K}$ then by 5.2.22(f), $n = 0$ and $b = b^q \in \mathbb{K}$. Thus $\mathbb{S} = \mathbb{K}$.

(c) Let $c, d \in \mathbb{P}$ with $d \neq 0$. Then there exists $r, s \in \mathbb{N}$ with $c^{p^r} \in \mathbb{K}$ and $d^{p^s} \in \mathbb{K}$. Put $t = \max(r, s)$. Then

$$(c \pm d)^{p^t} = c^{p^t} \pm d^{p^t} \in \mathbb{K} \text{ and } (cd^{\pm 1})^{p^t} = c^{p^t}(d^{pt})^{\pm 1} \in \mathbb{K}$$

In follows that $c \pm d$ and $cd^{\pm 1} \in \mathbb{P}$ and so ı is a subfield of $\mathbb{F}$.

(d) Since b is a root of $f \in \mathbb{F}$ and $\mathbb{K} \leq \mathbb{F}$ is normal, f splits over $\mathbb{F}$. So the distinct roots $b_1, \ldots b_k$ of f all are contained in $\mathbb{F}$. Put $h = \operatorname{Frob}_{\frac{1}{q}}(g)$. By 5.2.22(d) $h^q = f$ and $h = (x - b_1)(x - b_2)\ldots(x - b_k)$. Thus h splits over $\mathbb{F}$ and $h \in \mathbb{F}[x]$. Since $h^q = f \in \mathbb{K}[x]$ we see that $d^q \in \mathbb{K}$ for each coefficient d of f. Hence $d \in \mathbb{P}$ and $h \in \mathbb{P}[x]$. Since h has no multiple roots and $h(b) = 0$ we conclude that h is separable over $\mathbb{P}$.

(e) $t = m_b^{\mathbb{P}}$. As $t \in \mathbb{P}[x]$, $t^q \in \mathbb{K}[x]$ for some power q of p. Since $t^q[b] = 0$ we conclude f divides t^q. As f is separable, f divides t and so $f = t$.

(f) Let $b \in \mathbb{P}$ and $\phi \in \operatorname{Aut}_{\mathbb{K}}\mathbb{F}$. The $b^q \in \mathbb{K}$ and so $\phi(b)^q = \phi(b^q) = b^q$. Thus $b = \phi(b)$ and $\mathbb{P} \leq \operatorname{Fix}_{\mathbb{F}}(\operatorname{Aut}_{\mathbb{K}}F)$.

Next let $b \notin \mathbb{P}$ and suppose that $\mathbb{K} \leq \mathbb{F}$ is normal. Then f splits over $\mathbb{F}$ and by (a) there exists a root $d \neq b$ of f in $\mathbb{F}$. By 5.2.10(b) $\mathbb{F}$ is a splitting field over $\mathbb{K}$. So 5.2.8(d) implies that there exists $\phi \in \mathrm{Aut}_{\mathbb{K}}\mathbb{F}$ with $\phi(b) = d$. Hence $b \notin \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{K}}F)$. □

Lemma 5.2.25. *(a) Let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ be field extensions. Then $\mathbb{K} \leq \mathbb{F}$ is separable if and only $\mathbb{K} \leq \mathbb{E}$ and $\mathbb{E} \leq \mathbb{F}$ are separable.*

(b) $\mathbb{K} \leq \mathbb{F}$ is separable if and only if $\mathbb{F} = \mathbb{K}[S]$ for some $S \subseteq \mathbb{F}$ such that each $b \in S$ is separable over $\mathbb{K}$.

Proof. Put $p := \mathrm{char}\,\mathbb{K}$. If $p = 0$ then by 5.2.21(a)ll algebraic extensions are separable. So 5.1.11 and 5.1.12 show that (a) and (a) holds a.

So suppose $p > 0$. Before proving (a) and (b) we prove

(*) Let $\mathbb{K} \leq \mathbb{L}$ be a field extension, $I \subset \mathbb{L}$ and $b \in \mathbb{L}$. If all elements in I are separable over $\mathbb{K}$ and b is separable over $\mathbb{K}[I]$, then b is separable over $\mathbb{K}$.

Let $s = m_b^{K(I)}$. By 5.1.3 $\mathbb{K}(I) = \bigcup\{\mathbb{K}(J) \mid J \subseteq I, J \text{ finite}\}$. Hence there exists a finite subset J of I with $s \in \mathbb{K}[J][x]$. So b is separable over $\mathbb{K}[J]$. We know proceed by induction on $|J|$. If $J = \emptyset$, b is separable over $\mathbb{K}$ and $(*)$ holds. So suppose $J \neq \emptyset$ and let $a \in J$. Then b is separable over $\mathbb{K}[a][J - a]$ and so by induction b is separable over $\mathbb{K}[a]$. Hence by 5.2.22(h), $\mathbb{K}[a, b] = \mathbb{K}[a, b^p]$. Let $\mathbb{E} = \mathbb{K}[b^p]$. Then $b \in \mathbb{K}[a, b] = \mathbb{K}[a, b^p] = \mathbb{E}[a]$. By 5.2.20 a is separable over $\mathbb{E}$. Since $b^p \in \mathbb{E}$, $\mathbb{E} \leq \mathbb{E}[b]$ is purely inseparable. So by 5.2.24e $m_a^{\mathbb{E}} = m_a^{\mathbb{E}[b]}$. Thus $\dim_{\mathbb{E}} \mathbb{E}[a] = \dim_{\mathbb{E}[b]} \mathbb{E}$. Hence $\mathbb{E} = \mathbb{E}[b]$. So $K[b] = \mathbb{K}[b^p]$ and by 5.2.22(h), b is separable over $\mathbb{K}$.

(a) Suppose that $\mathbb{K} \leq \mathbb{E}$ and $\mathbb{E} \leq \mathbb{F}$. are separable. Let $b \in \mathbb{F}$ and let $I = \mathbb{E}$. Then by (*), b is separable over $\mathbb{K}$. So $K \leq \mathbb{F}$ is separable.

Conversely suppose $\mathbb{K} \leq \mathbb{F}$ is separable. Then clearly $\mathbb{K} \leq \mathbb{E}$ is separable. By 5.2.20 also $\mathbb{E} \leq \mathbb{F}$ is separable.

(b) If $\mathbb{K} \leq \mathbb{F}$ is separable, then $\mathbb{F} = \mathbb{K}[S]$ with $S = \mathbb{F}$. So suppose $\mathbb{F} = \mathbb{K}[S]$ with all elements in S separable over $\mathbb{K}$. Let $b \in \mathbb{F} = \mathbb{K}[S]$. Then b is separable over $\mathbb{K}[S]$ and so by (*), b is separable over $\mathbb{K}$. Thus $\mathbb{K} \leq \mathbb{F}$ is separable. □

Lemma 5.2.26. *Let $\mathbb{K} \leq \mathbb{F}$ be an all algebraic field extension and $\mathbb{E}$ and $\mathbb{L}$ intermediate field. Then $\mathbb{EL}$ is a subfield of $\mathbb{F}$.*

Proof. Since $\mathbb{F}$ is commutative, $\mathbb{EL} = \mathbb{LE}$ and so $\mathbb{EL}$ is a subring of $\mathbb{F}$. Let $0 \neq a \in \mathbb{EL}$. Since $\mathbb{K} \leq \mathbb{F}$ is algebraic and $\mathbb{K} \leq \mathbb{EL}$, $a^{-1} \in \mathbb{K}[a] \leq \mathbb{EL}$ for all $0 \neq a \in \mathbb{EL}$. Thus $\mathbb{EL}$ is a subfield of $\mathbb{F}$. □

Definition 5.2.27. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension. Then $\mathbb{S}(\mathbb{K}, \mathbb{F})$ consist of the elements in $\mathbb{F}$, which are separable over $\mathbb{K}$ and $\mathbb{P}(\mathbb{K}, \mathbb{F})$ of the elements in $\mathbb{F}$ which are purely inseparable over $\mathbb{K}$.*

Lemma 5.2.28. *Let $\mathbb{K} \leq \mathbb{F}$ be an algebraic field extension with $\mathrm{char}\,K = p \neq 0$. Put $\mathbb{P} = \mathbb{P}(\mathbb{K}, \mathbb{F})$ and $\mathbb{S} = \mathbb{S}(\mathbb{K}, \mathbb{F})$.*

(a) $\mathbb{S}$ *is a subfield of* $\mathbb{F}$.

(b) $\mathbb{S} \leq \mathbb{F}$ *is purely inseparable.*

(c) If $\mathbb{K} \leq \mathbb{F}$ *is normal, then* $\mathbb{F} = \mathbb{SP}$.

Proof. (a) Follows from 5.2.25(a).

(b) Let $b \in \mathbb{F}$. By 5.2.22(i),b^{p^n} is separable over $\mathbb{K}$ for some $n \in \mathbb{N}$. Thus $b^{p^n} \in \mathbb{S}$ and so by 5.2.24(a)m b is purely inseparable over $\mathbb{S}$.

(c) By 5.2.26 $\mathbb{SP}$ is a subfield of $\mathbb{F}$. By (b), $\mathbb{SP} \leq \mathbb{F}$ is purely inseparable and since $\mathbb{K} \leq \mathbb{F}$ is normal, $\mathbb{SP} \leq \mathbb{F}$ is separable 5.2.24(d) Thus $\mathbb{F} = \mathbb{SP}$. □

5.2.29 (Polynomial Rings)**.** Let R be a ring and I a set. Define

$$J := \bigoplus_{i \in I} \mathbb{N} = \{(n_i)_{i \in I} \mid n_i \in \mathbb{N}, \text{ almost all } 0\}$$

Note that J is a monoid under the addition $(n_i) + (m_i) = (n_i + m_i)$. Let S be the semi group ring of J over R. Since we use addition notation for $\mathbb{N}$ but multiplicative notation in R its is convenient to introduce the following notation: we write x^n for $n \in J$ and define $x^n x^m = x^{n+m}$. Then every elements in S can be unique written as

$$\sum_{n \in J} r_n x^n$$

where $r_n \in R$ for $n \in J$, almost all 0.

We denote S be $R[x_i, i \in I]$.

Suppose that R has an identity 1. Let $n(i) = (\delta_{ij})$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Define $x_i := x^{n(i)}$. (Note that x_i also depends on I and so on rare occasion we will write $x_{i,I}$ for x_i.). Then for $n = (n_i)_{i \in I} \in J$, $x^n = \prod_{i \in I} x_i^{n_i}$. Since $r \to r x^{0_J}$ is a 1-1 ring homomorphism, we can and do identify r with $r x^{0_J}$.

Suppose now that $\phi : R \to T$ is a ring homomorphism, that T is commutative ring with identity and $t = (t_i)_{i \in i}$ is a family of elements in T. For $n = (n_i)_{i \in I} \in J$ define $t^n = \prod_{i \in I} t_i^{n_i}$. Then the maps

$$J \to (T, \cdot), n \to t^n$$

is a homomorphism of semigroups and so by 3.2.5 the map

$$\gamma : \mathbb{R}[x_i, i \in I] \to T, \sum_{n \in J} r_n x^n \sum_{n \in J} \phi(r_n) t^n$$

is a ring homomorphism.

If R has an identity, then $\gamma(r) = \beta(r)$ and $\gamma(x_i) = t_i$ for all $i \in I$. Moreover γ is the unique homomorphism with this property. In the case that R is a subring of T $\phi(r) = r$, γ is called the evaluation homomorphism and we denote $\gamma(f)$ by $f(t)$. So if $f = \sum_{n \in J} r_n x^n$ then

$$f(t) = \sum_{n \in J} r_n t^n$$

Suppose now that $I = I_1 \cup I_2$ with $I_1 \cap I_2 = \emptyset$.Put $J_k := \bigoplus_{i \in I_k} \mathbb{N}$. Give $a = (a_i)_{i \in I_1} \in J_1$ and $b = (b_i)_{i \in I_2}$ we can define $c = (c_i)_{i \in I}$ by $c_i = a_i$ if $i \in J_1$ and $c_i = b_i$ if $i \in J_2$. We denote c by (a, b). Let $0_k = (0)_{i \in J_i} = 0_{J_i}$. Then the maps $J_1 \to J, a \to (a, 0_2)$ and $J_2 \to J, b \to (0_1, b)$ are 1-1 homomorphism of semigroups. Using 3.2.5 we obtain ring homomorphism

$$\alpha : R[x_i, i \in I_1] \to R[x_i, i \in I], \sum_{a \in J_1} r_a x^a \to \sum_{a \in J_1} r_a x^{(a,0_2)}$$

Note that $x^{(a,0_2)} x^{(0_1,b)} = x^{(a,b)}$. Hence using 3.2.5 one more time a ring homomorphism.

$$\beta : R[x_i, i \in I_1][x_i, i \in I_2] \to R[x_i, i \in I], \sum_{b \in J_2} \left(\sum_{a \in J_1} r_{a,b} x^a \right) x^b \to \sum_{(a,b) \in J_1 \times J_2} r_{a,b} x^{(a,b)}$$

β is clearly 1-1 and onto and so an isomorphism.

(We remark that isomorphism is special case of $R[G_1][G_2] \cong R[G_1 \times G_2]$, whenever, G_1, G_2 are semigroups) Thanks to this canonical isomorphism we will usually identify $R[x_i, i \in I_1][x_i, i \in I_2]$ with $R[x_i, i \in I]$ and view $R[x_i, i \in I_1]$ as a subring of $R[x_i, i \in I]$. In particular we identify x^a with $x^{(a,0_2)}$ and x_{i,I_1} with $x_{i,I}$. (So writing x_i for x_{i,I_1} creates no harm).

Suppose now that $R = \mathbb{K}$ is a field. The it is easy to see that $\mathbb{K}[x_i, i \in I]$ is an integral domain. (For $|I| = 1$ see Example 3.4.2, the same argument works in general, alternatively reduced to the case $I|$ finite and proceed by by induction) We denote the field of fraction of $\mathbb{K}[x_i, i \in I]$ by $\mathbb{K}(x_i, i \in I)$. Consider the case $|I_2| = \{k\}$. From Homework 3#1 we obtain canonical isomorphism between the field of fractions of $\mathbb{K}[x_i, i \in I_1][x_k]$ and $\mathbb{K}(x_i, i \in I_1)(x_k)$. The former is canonical isomorphic to the field of fraction of $\mathbb{K}[x_i, i \in I]$, that is to $\mathbb{K}(x_i, i \in I)$. A similar argument gives a canonical isomorphism

$$\mathbb{K}(x_i, i \in I_1)(x_i, i \in I_2) \cong \mathbb{K}(x_i, i \in I)$$

Example 5.2.30. 1. Let $\mathbb{K}$ be a field with char $K = p \neq 0$, I a set and $\mathbb{F} = \mathbb{K}(x_i, i \in I)$, the field of fractions of the polynomial ring $R = \mathbb{K}[x_i, i \in I]$ in the indeterminates $(x_i, i \in I)$. Put $\mathbb{E} = \mathbb{K}(x_i^p, i \in I)$. The each x_i is purely inseparable over $\mathbb{E}$. Note that the elements in $\mathbb{F}$ with $a^p \in \mathbb{E}$ form a subfield of $\mathbb{F}$. This contains $\mathbb{K}$ and all x_i and so is equal to $\mathbb{F}$. Thus $f^p in \mathbb{E}$ for all $f \in \mathbb{F}$ and $\mathbb{E} \leq \mathbb{F}$ is purely inseparable. We will show that $\dim_{\mathbb{E}} \mathbb{F} = \infty$ if $|I|$ is infinite. Put

$$J = \{(n_i)_{i \in I} \mid n_i \in \mathbb{N}, \text{ almost all } 0\} \text{ and } J_p = \{(r_i)_{i \in I} \mid r_i < p\}$$

For $n = (n_i) \in I$ define $x^n := \prod_{\in I} x_i^{n_i} \in R$. Then $(x^n)_{n\in J}$ is $\mathbb{K}$-basis for R. We will show that $(x^r)_{r\in J_p}$ is a $\mathbb{E}$ basis for $\mathbb{F}$.

For this let $n = (n_i) \in J$. For $l \in \mathbb{N}$ define $ln = (ln_i)_{i\in I} \in J$. Pick $q_i, r_i \in I$ with $n_i = pq_i + r_i$ with $q_i, r_i \in \mathbb{Z}$, $0 \leq r_i < p$. Put $q = (q_i)$, $r = (r_i)$. Then $r \in J_p$, $n = pq + r$ and

$$x^n = x^{pq}x^r$$

Since $x^{pq} \in \mathbb{E}$, we conclude that x^n is in the $\mathbb{E}$-span of $(x^r)_{r\in J_p}$. Hence also R is in the $\mathbb{E}$-span of $(x^r)_{r\in J_p}$. Let $f \in F$. Then $f = \frac{g}{h}$ with $f, g \in R$, $g \neq 0$. Since $(\frac{1}{g})^p \in \mathbb{E}$ and

$$\frac{f}{g} = \left(\frac{1}{g}\right)^p fg^{p-1}$$

we conclude that $(x^r)_{r\in J_p}$ spans $\mathbb{F}$ over $\mathbb{E}$.

Suppose now that $e_r \in \mathbb{E}$ for $r \in J_p$ almost all 0 and

$$\sum_{r\in J_p} e_r x^r = 0$$

We need to show that $e_r = 0$ for all $r \in J_p$. Multiplying with an appropriate element of $R_p := \mathbb{K}[x_i^p, i \in I]$ we may assume that $e_r \in R_p$ for all $r \in R$. Thus $e_r = \sum_{q\in J} k_{q,r}x^{pq}$ for some $k_{q,r} \in \mathbb{K}, q \in J$ almost all 0. Thus

$$\sum_{r\in J_p}\sum_{q\in J} k_{q,r}x^{pq+r} = 0$$

Each element in $n \in$ J can by uniquely written as $pq + r$ with $q \in J$ and $r \in J_p$. Thus the linear independence of the $(x^n)_{n\in J}$ shows that

$$k_{q,r} = 0$$

for all q and r. Hence also $e_r = 0$ for all $r \in J_p$. Thus $(x^r)_{r\in J_p}$ is linearly independent.

2. Next we give an example of a field extension $\mathbb{K} \leq \mathbb{F}$ such that $\mathbb{P}(\mathbb{K}, \mathbb{F}) \leq \mathbb{F}$ is not separable and $\mathbb{P}(\mathbb{K}, \mathbb{F})\mathbb{S}(\mathbb{K}, \mathbb{F}) \neq \mathbb{F}$. So 5.2.24(e) and 5.2.28(c) may be false if $\mathbb{K} \leq \mathbb{F}$ is not normal.

   Let $\mathbb{F}_4$ be a splitting field for $x^2 + x + 1$ over $\mathbb{Z}_2$ and a a root of $x^2 + x + 1$ in $\mathbb{F}_4$. Then $a \neq 0, 1$ and so $\mathbb{F}_4 \neq \mathbb{Z}_2$. Let y and z be indeterminates over $\mathbb{F}_4$. Put $\mathbb{E} = \mathbb{F}_4(y, z)$, $\mathbb{K} = \mathbb{Z}_2(y^2, z^2)$, $\mathbb{S} = \mathbb{F}_4(y^2, z^2)$ and $\mathbb{P} = \mathbb{F}_2(y, z)$. Since $(x^2 + x + 1)' = 1 \neq 0$, $x^2 + x + 1$ is separable. Thus a is separable over $\mathbb{K}$ and so $\mathbb{K} \leq \mathbb{S}$ is separable. Since $d^2 \in \mathbb{K}$ for all $d \in \mathbb{P}$ and $d^2 \in \mathbb{S}$ for all $\mathbb{K} \leq \mathbb{P}$ and $\mathbb{S} \leq \mathbb{E}$ are purely inseparable. If d is separable over $\mathbb{K}$, then $\mathbb{K}(d) = \mathbb{K}(d^2) \subseteq \mathbb{S}$. So $\mathbb{S}$ consist of all the elements in $\mathbb{E}$, separable over $\mathbb{K}$. Since $\mathbb{S} \neq \mathbb{K}$, $\mathbb{K} \leq \mathbb{E}$ is not purely inseparable. Thus the field consisting of the purely

inseparable elements in $\mathbb{E}$ over $\mathbb{K}$ is not equal to $\mathbb{K}$. It contains $\mathbb{P}$ and since $\dim_{\mathbb{P}} \mathbb{K} = 2$, it is equal to $\mathbb{K}$. Let $b = y + az$. Then $b^2 = y^2 + a^2z^2 = (y^2 + z^2) + az$. Thus $\mathbb{K}(b^2) = \mathbb{S}$, $x^2 + (y^2 + z^2 + az)$ is the minimal polynomial of b over $\mathbb{S}$ and so $(1, b)$ is an s-basis for $\mathbb{F} := \mathbb{S}[b]$. Thus $\mathbb{F} = \{s + tb \mid s, t \in \mathbb{S}\} = \{s + ty + taz \mid s, t \in \mathbb{S}\}$. So $(1, y, z, yz)$ is a $\mathbb{K}$-basis for $\mathbb{P}$ and an $\mathbb{S}$ basis for $\mathbb{E}$. Let $d \in \mathbb{F} \cap \mathbb{P}$. Then there exists $s, t \in \mathbb{S}$ and $k_1, k_2, k_3.k_4 \in \mathbb{K}$ with

$$s + ty + taz = d = k_1 + k_2y + k_3z + k_4yz$$

Since $\{1, y, z, yz\}$ is linearly independent over $\mathbb{S}$ we conclude that $s = k_1, t = k_2$ and $az = k_3$. So s, t and at are in $\mathbb{K}$. If $t \neq 0$ we get $a = att^{-1} \in \mathbb{K}$, a contradiction. Thus $t = 0$ and $f = s \in \mathbb{K}$. Thus $\mathbb{F} \cap \mathbb{P} = \mathbb{K}$. Hence

$$\mathbb{P}(\mathbb{K}, \mathbb{F}) = \mathbb{F} \cap \mathbb{P} = \mathbb{K} \text{ and } \mathbb{S}(\mathbb{K}, \mathbb{F}) = \mathbb{F} \cap \mathbb{S} = \mathbb{S} \neq \mathbb{F}$$

It follows that $\mathbb{P}(\mathbb{K}, \mathbb{F}) \leq \mathbb{F}$ is not separable and $\mathbb{P}(\mathbb{K}, \mathbb{F})\mathbb{S}(\mathbb{K}, \mathbb{F}) = \mathbb{K}\mathbb{S} = \mathbb{S} \neq \mathbb{F}$.

5.3 Galois Theory

Hypothesis 5.3.1. *Throughout this section $\mathbb{F}$ is a field and $G \leq \mathrm{Aut}(\mathbb{F})$.*

Definition 5.3.2. *Let $H \leq G$ and $\mathbb{E}$ a subfield of $\mathbb{F}$.*

(a) $\mathcal{F}H := \mathrm{Fix}_{\mathbb{F}}(H)$.

(b) $\mathcal{G}E := G \cap \mathrm{Aut}_{\mathbb{E}}(F)$.

(c) We say that H is $(G, \mathbb{F})$-closed *or that H is closed in G if $H = \mathcal{G}\mathcal{F}H$.*

(d) $\mathbb{E}$ is $(G, \mathbb{F})$-closed *if $\mathbb{E} = \mathcal{F}\mathcal{G}\mathbb{E}$.*

(e) "closed" means $(G, \mathbb{F})$-closed.

Lemma 5.3.3. *Let $T \leq H \leq G$ and $\mathbb{L} \leq \mathbb{E} \leq \mathbb{F}$. Then*

(a) $\mathcal{F}H$ is a subfield of $\mathbb{F}$ containing $\mathcal{F}G$

(b) $\mathcal{G}E$ is a subgroup of G

(c) $\mathcal{F}H \leq \mathcal{F}T$.

(d) $\mathcal{G}\mathbb{E} \leq \mathcal{G}\mathbb{L}$.

(e) $H \leq \mathcal{G}\mathcal{F}H$

(f) $\mathbb{E} \leq \mathcal{F}\mathcal{G}\mathbb{E}$.

(g) $\mathcal{F}H$ is closed.

(h) $\mathcal{G}\mathbb{E}$ is closed.

Proof. (a) and (b) are readily verified. (c) Let $t \in T$ and $a \in \mathcal{F}H$. Since $t \in H$, $t(a) = a$ and so $\mathcal{F}H \subseteq \mathcal{F}T$.

(d) Similar to (c).

(e) Note that $h(a) = a$ for all $a \in \mathcal{G}H$, $h \in H$. So $H \leq \mathcal{G}\mathcal{F}H$.

(f) Similar to (e).

(g) By (e) $H \leq \mathcal{G}\mathcal{F}H$ and so by (c) $\mathcal{F}\mathcal{G}\mathcal{F}H \leq \mathcal{F}H$, On the other hand, by (f) applied to $\mathbb{E} = \mathcal{F}H$, $\mathcal{F}H \leq \mathcal{F}\mathcal{G}\mathcal{F}H$. So (g) holds.

(h) Similar to (h) □

Proposition 5.3.4. *$\mathcal{F}$ induces an inclusion reversing bijection between the closed subgroups of G and the closed subfields of $\mathbb{F}$. The inverse is induced by $\mathcal{G}$.*

Proof. By 5.3.3(g), $\mathcal{F}$ sends a closed subgroup to a closed subfields and by 5.3.3(h) , $\mathcal{G}$ sends a closed subfields to a closed subgroup. By definition of closed, $\mathcal{F}$ and $\mathcal{G}$ are inverse to each other, then restricted to closed objects. Finally by 5.3.3(c) and (d), $\mathcal{F}$ and $\mathcal{G}$ are inclusion reversing. □

Lemma 5.3.5. *Let $H \leq T \leq G$ with T/H finite. Then* $\dim_{\mathcal{F}T} \mathcal{F}H \leq |T/H|$.

Proof. Let $k \in \mathcal{F}H$ and $W = tH \in T/H$. Define $W(k) := t(k)$. Since $(th)(k) = t(h(k)) = t(k)$ for all $h \in H$, this is well defined. Define

$$\Phi : \mathcal{F}H \to \mathbb{F}^{T/H},\ k \to (W(k))_{W \in T/H}$$

Let $L \subseteq \mathcal{F}H$ be a basis for $\mathcal{F}H$ over $\mathcal{F}T$. We claim that $(\Phi(l))_{l \in L}$ is linear independent over $\mathbb{F}$. Otherwise choose $I \subseteq L$ minimal such that $(\Phi(i))_{i \in I}$ is linear dependent over $\mathbb{F}$. Note that $|I|$ is finite. Then there exists $0 \neq k_i \in \mathbb{F}$, with

$$(*) \qquad \sum_{i \in I} k_i \Phi(i) = 0.$$

Fix $b \in I$. Dividing by k_b we may assume that $k_b = 1$.

Note that (*) means

$$(**) \qquad \sum_{i \in I} k_i W(i) = 0, \quad \text{for all } W \in T/H.$$

Suppose that $k_i \in \mathcal{F}T$ for all i. Note that $H(i) = \mathrm{id}_{\mathbb{F}}(i) = i$ for all $i \in I$ So using $W = H$ in (**) we get $\sum_{i \in I} k_i i = 0$, a contradiction to the linear independence of I over $\mathcal{F}T$. So there exists $d \in I$ and $\mu \in T$ with $\mu(k_d) \neq k_d$. Note that $\mu(t(k)) = (\mu t)(k)$ and so $\mu(W(k)) = (\mu W)(k)$. Thus applying μ to $(**)$ we obtain.

$$\sum_{i\in I} \mu(k_i)(\mu W)(i) = 0, \quad \text{for all } W \in T/H.$$

As every $W \in T/H$ is of the form $\mu W'$ for some $W' \in T/H$, (namely $W' = \mu^{-1}W$) we get

$$(***) \quad \sum_{i\in I} \mu(k_i)W(i) = 0, \quad \text{for all } W \in T/H.$$

Subtracting (**) form (***) we conclude:

$$\sum_{i\in I} (\mu(k_i) - k_i)W(i) = 0, \quad \text{for all } W \in T/H.$$

and so

$$\sum_{i\in I} (\mu(k_i) - k_i)\Phi(i) = 0.$$

The coefficient of $\Phi(b)$ in this equation is $\mu(1) - 1 = 0$. The coefficient of $\Phi(d)$ is $\mu(k_d) - k_d \neq 0$. We conclude that $(Phi(j)_{j\in I\setminus\{b\}}$ is linear dependent over $\mathbb{F}$, a contradiction the minimal choice of $|I|$.

This contradiction proves that $(\Phi(l))_{l\in L}$ is linear independent over $\mathbb{F}$.

Thus

$$\dim_{\mathcal{F}T} \mathcal{F}H = |L| \leq \dim_{\mathbb{F}} \mathbb{F}^{T/H} = |T/H|$$

So the theorem is proved. □

We remark that the last equality in the last equation is the only place where we used that $|T/H|$ is finite.

Lemma 5.3.6. *Let $b \in \mathbb{F}$ and $H \leq G$.*

(a) b is algebraic over $\mathcal{F}H$ if and only if $Hb := \{\phi(b) \mid \phi \in H\}$ is finite.

(b) Suppose that b is algebraic over $\mathcal{F}H$ and let m_b be the minimal polynomial of b over $\mathcal{F}H$. Then

(a) $m_b = \prod_{e\in Hb} x - e$.

(b) m_b is separable and b is separable over $\mathcal{F}H$.

(c) m_b splits over $\mathbb{F}$.

(d) Put $H_b := \{\phi \in H \mid \phi(b) = b\}$. Then

$$|H/H_b| = \deg m_a = |Hb| = \dim_{\mathcal{F}H} \mathcal{F}H[b]$$

Proof. Suppose b is algebraic over $\mathcal{F}H$ and let m_a be the minimal polynomial of b over $\mathcal{F}H$. Let $\phi \in H$. Then $\phi(b)$ is a root of $\phi(m_a) = m_a$. Since m_a has only finitely many roots, Hb is finite.

Suppose next H_b holds. Let $f := \prod_{e\in Hb} x - e$. Since the map $Hb \to Hb, e \to \phi(e)$ is a bijection with inverse $e \to \phi^{-1}(e)$,

$$\phi(f) = \prod_{e\in Hb} x - \phi(b) = \prod_{e\in Hb} x - ef.$$

Hence all coefficient of f are fixed by ϕ and so $f \in \mathcal{F}H[x]$. Clearly b is a root of f and so b is algebraic over $\mathcal{F}H$. Thus (a) holds.

Moreover m_b divides f. Let $e \in Hb$. Then $e = \phi(b)$ for some $\phi \in H$. Then $\phi(m_b) = m_b$ and as b is a root of m, $\phi(b)$ is a root of m_b. Hence f divides m_b and $f = m_b$. Thus (b:a) hold. Since f is separable and splits over $\mathbb{F}$ also (b:b) and (b:c) holds.

By 2.10.14 $|H/H_b| = |Hb|$, By 5.1.2(1) $\dim_{\mathcal{F}H} \mathcal{F}H[b] = \deg m = \deg f = |Hb|$ and so also (b:d) holds. □

Lemma 5.3.7. *Let $\mathbb{L} \le \mathbb{E} \le \mathbb{F}$ with $\mathbb{L} \le \mathbb{E}$ finite. Then*

$$|\mathcal{G}\mathbb{L}/\mathcal{G}\mathbb{E}| \le \dim_{\mathbb{L}} \mathbb{E}$$

Proof. If $\mathbb{E} = \mathbb{L}$, this is obvious. So we may assume $\mathbb{E} \neq \mathbb{L}$. Pick $e \in \mathbb{E} \setminus \mathbb{L}$. Then e is algebraic over $\mathbb{L}$ and since $\mathbb{L} \le \mathcal{F}\mathcal{G}\mathbb{L}$, e is also algebraic over $\mathcal{F}\mathcal{G}\mathbb{L}$. Moreover, $g = m_e^{\mathcal{F}\mathcal{G}\mathbb{L}}$ divides $f = m_e^{\mathbb{L}}$. Put $H = \mathcal{G}\mathbb{L}$. By 5.3.6 $|H/H_e| = \deg g$. Note that $H_e = \mathcal{G}(\mathbb{L}[e])$ and so

$$|\mathcal{G}\mathbb{L}/\mathcal{G}(\mathbb{L}[e])| = |H/H_e| = \deg g \le \deg f = \dim_{\mathbb{L}} \mathbb{L}[e].$$

By induction on $\dim_{\mathbb{L}} \mathbb{E}$,

$$|\mathcal{G}(\mathbb{L}[e])/\mathcal{G}\mathbb{E}| \le \dim_{\mathbb{L}[e]} \mathbb{E}.$$

Multiplying the two inequalities we obtain the result. □

Theorem 5.3.8. *(a) Let $H \le T \le G$ with H closed and T/H finite. Then T is closed and*

$$\dim_{\mathcal{F}T} \mathcal{F}H = |T/H|.$$

(b) Let $\mathbb{L} \le \mathbb{E} \le \mathbb{F}$ with $\mathbb{L}$ closed and $\mathbb{L} \le \mathbb{E}$ finite. Then $\mathbb{E}$ is closed and

$$|\mathcal{G}\mathbb{L}/\mathcal{G}\mathbb{E}| = \dim_{\mathbb{L}} \mathbb{E}.$$

Proof. (a) Using 5.3.3,5.3.5, 5.3.7 and $H = \mathcal{G}\mathcal{F}H$ we compute

$$|T/H| \ge \dim_{\mathcal{F}T} \mathcal{F}H \ge |\mathcal{G}\mathcal{F}T/\mathcal{G}\mathcal{F}H| = |\mathcal{G}\mathcal{F}T/H| \ge |T/H|.$$

So all the inequalities are equalities. Thus $T = \mathcal{G}\mathcal{F}T$ and

$$\dim_{\mathcal{F}T} \mathcal{F}H = |T/H|.$$

(b) This time we have

$$\dim_{\mathbb{L}} \mathbb{E} \geq |\mathcal{G}L/\mathcal{G}\mathbb{E}| \geq \dim_{\mathcal{F}\mathcal{G}\mathbb{L}} \mathcal{F}\mathcal{G}\mathbb{E} = \dim_{\mathbb{L}} \mathcal{F}\mathcal{G}\mathbb{E} \geq \dim_{\mathbb{L}} \mathbb{E}$$

So all the inequalities are equalities. Thus $\mathbb{E} = \mathcal{F}\mathcal{G}\mathbb{E}$ and

$$\dim_{\mathbb{L}} \mathbb{E} = |\mathcal{G}\mathbb{L}/\mathcal{G}\mathbb{E}|$$

□

Proposition 5.3.9. *(a) Let $H \leq G$ with H finite. Then H is closed and* $\dim_{\mathcal{F}H} \mathbb{F} = |H|$.

(b) Put $K = \mathcal{F}G$ and let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ with $\mathbb{K} \leq \mathbb{E}$ finite. Then $\mathbb{E}$ is closed and $\dim_{\mathbb{K}} \mathbb{E} = |G/\mathcal{G}\mathbb{E}|$.

Proof. (a) Note that $\mathcal{F}\{\mathrm{id}_{\mathbb{F}}\} = \mathbb{F}$ and so $\mathcal{G}\mathcal{F}\{\mathrm{id}_{\mathbb{F}}\} = \{\mathrm{id}_{\mathbb{F}}\}$. Hence trivial group is closed and has finite index in H. So (a) follows from 5.3.8a

(b) By 5.3.3(g), $\mathbb{K} = \mathcal{F}G$ is closed. Moreover, $\mathcal{G}\mathbb{K} = G \cap \mathrm{Aut}_{\mathbb{K}}(\mathbb{F}) = G$. Thus by 5.3.8(b), applied with $\mathbb{L} = \mathbb{K}$, $\mathbb{E}$ is closed and

$$\dim_{\mathbb{K}} \mathbb{E} = |\mathcal{G}\mathbb{K}/\mathcal{G}\mathbb{E}| = |G/|\mathcal{G}\mathbb{E}|$$

□

Definition 5.3.10. *A field extension $\mathbb{L} \leq \mathbb{E}$ is called* Galois *if $\mathbb{L} = \mathrm{Fix}_{\mathbb{E}}(\mathrm{Aut}_{\mathbb{L}}(\mathbb{E}))$.*

Lemma 5.3.11. *Put $\mathbb{K} = \mathcal{F}G$. Then $\mathbb{K} \leq \mathbb{F}$ is a Galois Extension. Moreover, if $\mathbb{K} \leq \mathbb{F}$ is finite, then $G = \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$.*

Proof. Since $G \leq \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ we have

$$\mathbb{K} \leq \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{F}}(\mathbb{K})) \leq \mathrm{Fix}_{\mathbb{F}}(G) = \mathbb{K}$$

Thus equality holds everywhere and $\mathbb{K} = \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ and $\mathbb{K} \leq \mathbb{F}$ is Galois.
Moreover, if $\mathbb{K} \leq \mathbb{F}$, then by 5.3.9 applied to G and to $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ in place of H.

$$|\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})| = \dim_{\mathbb{K}} \mathbb{F} = |G|$$

Since $G \leq \mathrm{Aut}_{\mathbb{F}}(\mathbb{K})$, this implies $G = \mathrm{Aut}_{\mathbb{F}}(\mathbb{K})$. □

Theorem 5.3.12 (Fundamental Theorem Of Galois Theory)**.** *Let $\mathbb{K} \leq \mathbb{F}$ be a finite Galois extension and put $G = \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$. Then*

(a) $\mathcal{F}$ is inclusion reversing bijection from the set of subgroups of G to the set of intermediate field of $\mathbb{K} \leq \mathbb{F}$.

(b) Let $H \leq G$ and $\mathbb{E} = \mathcal{F}H$. Then $\dim_{\mathbb{E}} \mathbb{F} = |H|$ and $H = \mathrm{Aut}_{\mathbb{E}}(\mathbb{F})$.

Proof. (a) Since $\mathbb{K} \leq \mathbb{F}$ is Galois, $\mathcal{G} = \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})) = \mathbb{K}$. Since $\mathbb{K} \leq \mathbb{F}$ is finite, 5.3.9(b) implies that G is finite and so by 5.3.9 all intermediate field of $\mathbb{K} \leq \mathbb{F}$ and all subgroups of G are closed. So by 5.3.4, $\mathcal{F}$ induces a inclusion reversing bijection between the subgroups of G and intermediate fields of $\mathbb{K} \leq \mathbb{F}$.

(b) By 5.3.9(a) $\dim_{\mathbb{E}} \mathbb{F} = |H|$. By 5.3.11 applied to H in place of G, $H = \mathrm{Aut}_{\mathbb{F}}(K)$. □

Lemma 5.3.13. *Put $\mathbb{K} = \mathrm{Fix}_{\mathbb{F}}(G)$ and let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ with $\mathbb{E}$ stable.*

(a) $\mathrm{Fix}_{\mathbb{E}}(G^{\mathbb{E}}) = \mathbb{K}$ and $\mathbb{K} \leq \mathbb{E}$ is Galois.

(b) If $\mathbb{K} \leq \mathbb{E}$ is finite, then $G^{\mathbb{E}} = \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$.

Proof. (a) $\mathrm{Fix}_{\mathbb{E}}(G^{\mathbb{E}}) = \mathrm{Fix}_{\mathbb{F}}G \cap \mathbb{E} = \mathbb{K} \cap \mathbb{E} = \mathbb{K}$

(b) Follows from 5.3.12 applied to $\mathbb{K} \leq \mathbb{E}$ in place of $\mathbb{K} \leq \mathbb{F}$ and $H = G^{\mathbb{E}}$. □

Lemma 5.3.14. *(a) Let $\mathbb{E} \leq \mathbb{F}$ and $g \in G$. Then ${}^{g}(\mathcal{G}\mathbb{E}) = \mathcal{G}(g(\mathbb{E}))$.*

(b) Let $H \leq G$ and $g \in G$. Then $\mathcal{F}({}^{g}H) = g(\mathcal{F}H)$.

(c) Let $H \trianglelefteq G$. Then $\mathcal{F}H$ is stable.

(d) Let $\mathbb{E} \leq \mathbb{F}$ and suppose $\mathbb{E}$ is stable. Then $\mathcal{G}\mathbb{E} \trianglelefteq G$ and $G^{\mathbb{E}} \cong G/\mathcal{G}\mathbb{E}$.

(e) Let $H \leq G$ be closed. Then $H \trianglelefteq G$ if and only if $\mathcal{F}H$ is stable.

(f) Let $\mathbb{E}$ be a closed subfield of $\mathbb{F}$. Then $\mathbb{E}$ is stable if and only if $\mathcal{G}\mathbb{E}$ is normal in G.

Proof. Let $g \in G$ and $b \in \mathbb{F}$. By 2.10.14(c)

$$(*) \qquad \mathrm{Stab}_G(g(b)) = {}^{g}\mathrm{Stab}_G(b)$$

(a) Intersecting (*) over all $b \in \mathbb{E}$ gives (a).

(b) By (*) $H \leq \mathrm{Stab}_G(b)$ if and only if ${}^{g}H \leq \mathrm{Stab}_G(g(b))$. Thus $b \in \mathcal{F}H$ if and only if $g(b) \in \mathcal{F}({}^{g}H)$. This gives (b).

(c) If $H \trianglelefteq G$ then by (b), $\mathcal{F}H = g(\mathcal{F}H)$.

(d) Suppose $\mathbb{E}$ is stable. Note the $\mathcal{G}\mathbb{E}$ is exactly the kernel of the restriction map, $\phi \to \phi|_{\mathbb{E}}$. Hence $\mathcal{G}\mathbb{E} \trianglelefteq G$ and (d) follows from the First Isomorphism Theorem 2.5.8.

(e) The forward direction follows from (c). By (d), if $\mathcal{F}H$ is stable , then $\mathcal{G}\mathcal{F}H \trianglelefteq G$. If H is closed, then $\mathcal{G}\mathcal{F}H = H$ and so the backward direction holds.

(f) Follows from (e) applied to $H = \mathcal{G}\mathbb{E}$. □

Lemma 5.3.15. *Put $\mathbb{K} = \mathrm{Fix}_{\mathbb{F}}(G)$ and let $\mathbb{L} \leq \mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$. Suppose that $\mathbb{L} \leq \mathbb{E}$ algebraic, and $\mathbb{L} \leq \mathbb{K}$ is purely inseparable. Then $\mathbb{E}$ is stable if and only if $\mathbb{L} \leq \mathbb{E}$ is normal.*

Proof. Suppose first that $\mathbb{E}$ is stable. Let $e \in E$ and $f = m_e^{\mathbb{K}}$.By 5.3.6, f splits over $\mathbb{F}$ and Ge is the set of roots of f. As $\mathbb{E}$ is stable, $Ge \leq \mathbb{E}$ and so f splits over $\mathbb{E}$. Since $\mathbb{K}/\mathbb{L}$ is pure inseparable, $f^q \in \mathbb{L}[x]$ for some $q \in \mathbb{N}$. So $m_e^{\mathbb{L}}$ divides f^q. We conclude that $m_e^{\mathbb{L}}$ splits over $\mathbb{E}$ and so $\mathbb{L} \leq \mathbb{E}$ is normal.

If $\mathbb{L} \leq \mathbb{E}$ is normal, then by 5.2.10 $\mathbb{E}$ is $\mathrm{Aut}_{\mathbb{L}}(\mathbb{F})$-stable. Since $G \leq \mathrm{Aut}_{\mathbb{F}}(\mathbb{K}) \leq \mathrm{Aut}_{\mathbb{L}}(\mathbb{F})$, $\mathbb{E}$ is stable. □

Lemma 5.3.16. *Put $\mathbb{K} = \mathcal{F}G$ and let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ with $\mathbb{K} \leq \mathbb{E}$ algebraic. Then*

(a) $\mathbb{K} \leq \mathbb{E}$ is separable.

(b) If $\mathbb{E}$ is closed, then the following are equivalent:

- *(a) $\mathcal{G}E \trianglelefteq G$.*
- *(b) $\mathbb{E}$ is stable*
- *(c) $\mathbb{K} \leq \mathbb{E}$ is normal.*

Proof. (a) follows from 5.3.6(b:b) (applied to $H = G$ and so $\mathcal{F}H = \mathbb{K}$). (b) By 5.3.14(f) (b:a) and (b:b) are equivalent. By 5.3.15 (b:b) and (b:c) are equivalent. □

Theorem 5.3.17. *Let $\mathbb{K} \leq \mathbb{F}$ be an algebraic field extension. Then the following are equivalent:*

(a) $\mathbb{K} \leq \mathbb{F}$ is Galois

(b) $\mathbb{K} \leq \mathbb{F}$ is separable and normal.

(c) $\mathbb{F}$ is the spitting field of a set over separable polynomials over $\mathbb{K}$.

Proof. (a) $\Longrightarrow$ (b): Suppose first that $\mathbb{K} \leq \mathbb{F}$ is Galois and put $G = \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$. Then $\mathbb{K} = G$. So by 5.3.16(a) $\mathbb{K} \leq \mathbb{F}$ is separable. Since $\mathbb{F}$ is closed and $\mathcal{G}F = \{id_{\mathbb{F}}\} \trianglelefteq G$, 5.3.16(b) gives that $\mathbb{K} \leq \mathbb{F}$ is normal.

(b) $\Longrightarrow$ (a): Suppose next that $\mathbb{K} \leq \mathbb{F}$ is normal and separable. Since $\mathbb{K} \leq \mathbb{F}$ is normal 5.2.24(f), shows that $\mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})) = \mathbb{P}(\mathbb{K}, \mathbb{F})$. Since $\mathbb{K} \leq \mathbb{F}$ is separable, $\mathbb{P}(\mathbb{K}, \mathbb{F}) = \mathbb{K}$ and so $\mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})) = \mathbb{K}$ and $\mathbb{K} \leq \mathbb{F}$ is Galois.

(b) $\Longrightarrow$ (c): Let $P = \{m_b^{\mathbb{K}} \mid b \in \mathbb{F}\}$. Since $\mathbb{K} \leq \mathbb{F}$ is normal, each $m_b^{\mathbb{K}}$ splits over $\mathbb{K}$. Since $\mathbb{K} \leq \mathbb{F}$ is separable, each $m_b^{\mathbb{K}}$ is separable. Hence $\mathbb{F}$ is a splitting field for P over $\mathbb{K}$ and (c) holds.

(b) $\Longrightarrow$ (c): Suppose $\mathbb{F}$ is the the splitting field of a set P of separable polynomials over $\mathbb{K}$. Then by 5.2.10b $\mathbb{K} \leq \mathbb{F}$ is normal. Let A consists of all $a \in \mathbb{F}$ such that a is a root of some non-zero $p \in P$. By definition of a splitting field, $\mathbb{F} = \mathbb{K}[A]$. Since each $p \in P$ is separable, each $a \in A$ is separable over $\mathbb{F}$. Thus by 5.2.25(b), $\mathbb{K} \leq \mathbb{F}$ is separable. □

Proposition 5.3.18. *Suppose that $\mathbb{K} \leq \mathbb{F}$ is algebraic and Galois, and suppose $G = \mathrm{Aut}_{\mathbb{F}}(\mathbb{K})$. Let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$.*

(a) $\mathcal{G}E = \mathrm{Aut}_{\mathbb{F}}(\mathbb{E})$, $\mathbb{K} \leq \mathbb{E}$ *is Galois and* $\mathbb{E}$ *is closed .*

(b) $\mathbb{K} \leq \mathbb{E}$ *is Galois if and only of* $\mathcal{G}\mathbb{E}$ *is normal in* G*.*

(c) $\mathrm{N}_G(\mathcal{G}E)/\mathcal{G}E \cong \mathrm{N}_G(\mathcal{G}E)^{\mathbb{E}} = \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$

Proof. (a) $\mathcal{G}E = G \cap \mathrm{Aut}_{\mathbb{E}}(\mathbb{F}) = \mathrm{Aut}_{\mathbb{E}}(\mathbb{F})$. By 5.3.17(a),(b) $\mathbb{K} \leq \mathbb{F}$ is normal and separable. Hence also $\mathbb{E} \leq \mathbb{F}$ is normal and separable. So by 5.3.17, $\mathbb{E} \leq \mathbb{F}$ is Galois. This implies that

$$\mathbb{E} = \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{E}}(\mathbb{F})) = \mathcal{F}\mathcal{G}\mathbb{E}$$

and so $\mathbb{E}$ is closed.

(b) As $\mathbb{K} \leq \mathbb{F}$ is separable, $\mathbb{K} \leq \mathbb{E}$ is separable. Hence by 5.3.17 $\mathbb{K} \leq \mathbb{E}$ is Galois if and only if $\mathbb{K} \leq \mathbb{E}$ is normal. Since $\mathbb{E}$ is closed, (b) now follows from 5.3.16(b).

(c) As $\mathcal{F}\mathcal{G}\mathbb{E} = \mathbb{E}$ we conclude from 5.3.14(b) that $\mathbb{E}$ is $\mathrm{N}_G(\mathcal{G}E)$-stable. So $\mathrm{N}_G(\mathcal{G}E)/\mathcal{G}E \cong \mathrm{N}_G(\mathcal{G}E)^{\mathbb{E}}$ by the 2.5.8. Clearly $\mathrm{N}_G(\mathcal{G}E)^{\mathbb{E}} \leq \mathrm{Aut}_{\mathbb{E}}(\mathbb{K})$. Let $h \in \mathrm{Aut}_{\mathbb{E}}(\mathbb{K})$. By 5.2.8 $h = g|_{\mathbb{E}}$ for some $g \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$. Then $g \in \mathrm{N}_G(\mathcal{G}E)$ and (c) holds. □

Definition 5.3.19. *Let* $\mathbb{K} \leq \mathbb{E}$ *be an algebraic field extension. A* normal closure *of* $\mathbb{K} \leq \mathbb{E}$ *is an extension* $\mathbb{L}$ *of* $\mathbb{E}$ *such that* $\mathbb{K} \leq \mathbb{L}$ *is normal and no proper proper subfield of* $\mathbb{L}$ *containing* $\mathbb{E}$ *is normal over* $\mathbb{K}$*.*

Lemma 5.3.20. *Let* $\mathbb{K} \leq \mathbb{E}$ *be an algebraic field extension.*

(a) Suppose $\mathbb{E} = \mathbb{K}(I)$ *for some* $I \subseteq \mathbb{E}$ *and let* $\mathbb{E} \leq \mathbb{L}$ *be a field extension. Then the following are equivalent:*

  (a) $\mathbb{L}$ *is a normal closure of* $\mathbb{K} \leq \mathbb{E}$ *.*

  (b) $\mathbb{L}$ *is a splitting field for* $\{m_b^{\mathbb{K}} \mid b \in I\}$ *over* $\mathbb{K}$*.*

  (c) $\mathbb{L}$ *is a splitting field for* $\{m_b^{\mathbb{K}} \mid b \in I\}$ *over* $\mathbb{E}$*.*

(b) There exists a normal closure $\mathbb{L}$ *of* $\mathbb{K} \leq \mathbb{E}$ *and* $\mathbb{L}$ *is unique up to* $\mathbb{E}$*-isomorphism.*

(c) Let $\mathbb{L}$ *be a normal closure of* $\mathbb{K} \leq \mathbb{E}$*. Then*

  (a) $\mathbb{K} \leq \mathbb{L}$ *is finite if and only if* $\mathbb{K} \leq \mathbb{E}$ *is finite.*

  (b) $\mathbb{K} \leq \mathbb{L}$ *is Galois if and only if* $\mathbb{K} \leq \mathbb{L}$ *is separable and if and only if* $\mathbb{K} \leq \mathbb{E}$ *is separable.*

(d) Let $\overline{\mathbb{E}}$ *be an algebraic closure of* $\mathbb{E}$*. Then*

  (a) $\overline{\mathbb{E}}$ *is an algebraic closure of* $\mathbb{K}$*.*

  (b) $\overline{\mathbb{E}}$ *contains a unique normal closure* $\mathbb{L}$ *of* $\mathbb{K} \leq \mathbb{E}$*.* $\mathbb{L}$ *is called the normal closure of* $\mathbb{K} \leq \mathbb{E}$ *in* $\overline{\mathbb{E}}$*.*

Proof. (a) Put $P = \{m_b^{\mathbb{K}} \mid b \in I\}$, $A = \{b \in \mathbb{L} \mid f(b) = 0_{\mathbb{K}} \text{ for some } b \in \mathbb{L}\}$ and $\mathbb{D} = \mathbb{K}(A)$. Note that $b \in A$ for all $b \in I$ and so

$$(*) \qquad \mathbb{E} = \mathbb{K}(I) \leq \mathbb{D}$$

Next we show:

(**) Let $\mathbb{K} \leq \mathbb{F} \leq \mathbb{L}$ such that $\mathbb{K} \leq \mathbb{F}$ is normal. Then $\mathbb{D} \subseteq \mathbb{F}$ and $\mathbb{K} \leq \mathbb{D}$ is normal.

Note that each $m_b^{\mathbb{K}}, b \in I$ has a root in $\mathbb{L}$, namely b. Since $\mathbb{K} \leq \mathbb{F}$ is normal each $m_b^{\mathbb{K}}$ splits over $\mathbb{L}$. So $A \subseteq \mathbb{F}$, $\mathbb{D} = \mathbb{K}[A] \leq \mathbb{F}$ and $\mathbb{D}$ is a splitting field for P over $\mathbb{K}$. Thus by 5.2.10(b), $\mathbb{K} \leq \mathbb{D}$ is normal.

(a:a) $\Longrightarrow$ (a:b): Suppose first that $\mathbb{L}$ is a normal closure of $\mathbb{K} \leq \mathbb{E}$. Then $\mathbb{K} \leq \mathbb{L}$ is normal and so by (**) $\mathbb{K} \leq \mathbb{D}$ is normal. By (*) $\mathbb{E} \leq \mathbb{D}$ and so the definition of a normal closure implies $\mathbb{L} = \mathbb{D}$.

(a:b) $\Longrightarrow$ (a:c): Suppose next $\mathbb{L}$ is a splitting field of P over $\mathbb{K}$. Then $\mathbb{L} = \mathbb{K}[\mathbb{A}] = \mathbb{E}[A]$ and $\mathbb{L}$ is also a splitting field for P over $\mathbb{E}$.

(a:c) $\Longrightarrow$ (a:a): Suppose next that $\mathbb{L}$ is a splitting field of P over $\mathbb{E}$. Then $\mathbb{L} = \mathbb{E}[A]$ and $\mathbb{D}$ is a splitting field for P over $\mathbb{E}$. Hence by rf[normalstable] b, $\mathbb{K} \leq \mathbb{D}$ is normal. By (*) $\mathbb{E} \leq \mathbb{D}$ and so $\mathbb{L} = \mathbb{E}[A] \leq \mathbb{D} \leq \mathbb{L}$. Thus $\mathbb{L} = \mathbb{D}$ and $\mathbb{K} \leq \mathbb{L}$ is normal. If $\mathbb{K} \leq \mathbb{F} \leq \mathbb{L}$ and $\mathbb{K} \leq \mathbb{F}$ is normal, then by (**) $\mathbb{D} \leq \mathbb{F}$. Since $\mathbb{D} = \mathbb{L}$ and $\mathbb{F} \leq \mathbb{L}$ we get $\mathbb{F} = \mathbb{L}$ and so $\mathbb{L}$ is a normal closure of $\mathbb{K} \leq \mathbb{E}$.

(b) By (a) applied with $I = \mathbb{E}$ a normal closure of $\mathbb{K} \leq \mathbb{E}$ is the same as splitting field of $\{m_b^{\mathbb{K}} \mid b \in \mathbb{E}\}$. Thus by 5.1.19 $\mathbb{K} \leq \mathbb{E}$ has a normal closure and by 5.2.8(b), the normal closure is unique up to $\mathbb{K}$-isomorphism.

(c:a) If $\mathbb{K} \leq \mathbb{E}$ is finite, then $\mathbb{E} = \mathbb{K}(I)$ for some finite subset I of $\mathbb{E}$. Note the splitting field of a finite set of polynomials over $\mathbb{K}$ is a finite extension of $\mathbb{K}$ So $\mathbb{K} \leq \mathbb{L}$ is finite by (a).

(c:b) Suppose that if $\mathbb{K} \leq \mathbb{L}$ is Galois. Then by 5.3.17 $\mathbb{K} \leq \mathbb{L}$ is separable. If $\mathbb{K} \leq \mathbb{L}$ is separable, then also $\mathbb{K} \leq \mathbb{E}$ is separable. So suppose $\mathbb{K} \leq \mathbb{E}$ is separable, then by (a), $\mathbb{L}$ is the splitting field of the set of separable polynomials $\{\{m_b^{\mathbb{K}} \mid b \in \mathbb{E}\}$ and so by 5.3.17, $\mathbb{K} \leq \mathbb{L}$ is Galois.

(d:a) Since $\mathbb{K} \leq \mathbb{E}$ and $\mathbb{E} \leq \overline{\mathbb{E}}$ are algebraic, $\mathbb{K} \leq \overline{\mathbb{E}}$ is algebraic (5.1.12. Also $\overline{\mathbb{E}}$ is algebraicly closed and thus (d:a) holds.

(d:b) Let $\mathbb{E} \leq \mathbb{L} \leq \overline{\mathbb{E}}$. Then by (a) $\mathbb{L}$ is a normal closure of $\mathbb{K} \leq \mathbb{E}$ if and only if $\mathbb{L}$ is generated by $\mathbb{K}$ and all the roots of the $m_b^{\mathbb{K}}, b \in \mathbb{E}$. So (d:b) holds. □

Lemma 5.3.21. *Let $\mathbb{K} \leq \mathbb{E}$ be a normal field extension. Put $\mathbb{P} = \mathbb{P}(\mathbb{K}, \mathbb{E})$ and $\mathbb{S} = \mathbb{S}(\mathbb{K}, \mathbb{E})$. Then $\mathbb{P} = \mathrm{Fix}_{\mathbb{E}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{E}))$, $\mathbb{K} \leq \mathbb{P}$ is purely inseparable, $\mathbb{P} \leq \mathbb{E}$ is Galois, $\mathbb{K} \leq \mathbb{S}$ is Galois and the map*

$$\tau : \mathrm{Aut}_{\mathbb{K}}(\mathbb{E}) \to \mathrm{Aut}_{\mathbb{K}}(\mathbb{S}), \phi \to \phi|_{\mathbb{E}}$$

is an isomorphism of groups.

Proof. By definition $\mathbb{K} \leq \mathbb{P}$ is purely inseparable. By 5.2.24(c), 5.2.24 f, $\mathbb{P} = \mathrm{Fix}_{\mathbb{E}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{E}))$ and so by 5.3.11 (applied with $\mathbb{F} = \mathbb{E}$ and $G = \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$) $\mathbb{P} \leq \mathbb{E}$ is Galois.

Let $\phi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$ and $s \in \mathbb{S}$. Then $\phi(s)$ is a root of $m_s^{\mathbb{K}}$. Since s is separable over $\mathbb{K}$, we conclude that $\phi(s)$ is separable over $\mathbb{K}$. So $\phi(s) \in \mathbb{K}$. Thus $\mathbb{S}$ is $\mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$ stable and so by 5.3.15, $\mathbb{K} \leq \mathbb{S}$ is normal. $\mathbb{K} \leq \mathbb{S}$ is separable and thus by 5.3.17, $\mathbb{K} \leq \mathbb{S}$ is Galois. Let $\phi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$ with $\phi\,|_{\mathbb{S}} = \mathrm{id}_{\mathbb{S}}$. Then $\mathbb{S} \subseteq \mathrm{Fix}_{\mathbb{F}}(\phi)$. By definition of $\mathbb{P}$, $\mathbb{P} \leq \mathrm{Fix}_{\mathbb{F}}(\phi)$. By 5.2.28(c), $\mathbb{E} = \mathbb{S}\mathbb{P}$. Hence $\mathbb{E} \leq \mathrm{Fix}_{\mathbb{E}}(\phi)$ and $\phi = \mathrm{id}_{\mathbb{E}}$. Thus τ is 1-1. Let $\psi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{S})$. Since $\mathbb{K} \leq \mathbb{E}$ is normal is normal, we conclude from 5.2.8(a) that $\psi = \phi\,|_{\mathbb{S}}$ for some $\phi \in \mathrm{Aut}_{\mathbb{K}}(\mathbb{E})$. So τ is onto. □

5.4 The Fundamental Theorem of Algebra

In this section we show that the field $\mathbb{C}$ of complex numbers is algebraically closed. Our proof is based on the following well known facts from analysis which we will not prove:

Every polynomial $f \in \mathbb{R}[x]$ of odd degree has a root in $\mathbb{R}$.
Every polynomials of degree 2 over $\mathbb{C}$ is reducible.
$\dim_{\mathbb{R}} C = 2$.

Some remarks on this assumptions. The first follows from the intermediate value theorem and the fact that any odd polynomial has positive and negative values. The second follows from the quadratic formula and the fact that every complex number has a complex square root ($\sqrt{re^{\phi i}} = \sqrt{r}e^{\frac{\phi}{2}i}$). The last property follows from $\mathbb{C} = \mathbb{R} + \mathbb{R}i$.

Definition 5.4.1. *Let s be a prime, $\mathbb{K} \leq \mathbb{F}$ a finite field extension and $f \in \mathbb{K}[x]$.*

(a) f is a s'-polynomial if s does not divide $\deg f$

(b) $\mathbb{K} \leq \mathbb{E}$ is a s'-extension s does not divide $\dim_{\mathbb{K}} \mathbb{F}$.

Lemma 5.4.2. *Let $\mathbb{K}$ be a field and s a prime. Then the following are equivalent.*

(a) Every irreducible s'-polynomial over $\mathbb{K}$ has degree 1.

(b) Every s'-polynomial over $\mathbb{K}$ has a root in $\mathbb{K}$

(c) If $\mathbb{K} \leq \mathbb{E}$ is a s' extension then $\mathbb{K} = \mathbb{E}$.

Proof. (a) $\Longrightarrow$ (b): Let $f \in \mathbb{K}[x]$ with $s \nmid \deg f$. Let $f = f_1 \dots f_k$ with f_i irreducible. Then $\deg f = \sum_{i=1}^{k} \deg f_i$ and so $s \nmid f_i$ for some $1 \leq i \leq k$. By (a) , f_i has degree 1. Hence f_i and so also f has a root in $\mathbb{K}$.

(b) $\Longrightarrow$ (c): Let $\mathbb{K} \leq \mathbb{E}$ be an s'-extension and $b \in \mathbb{E}$. Then $\deg m_b^{\mathbb{K}} = \dim_{\mathbb{K}} \mathbb{K}[b]$ divides $\dim_{\mathbb{K}} \mathbb{E}$. Hence $m_b^{\mathbb{K}}$ is an irreducible s' polynomial and so by (a) has a root d in $\mathbb{K}$. As f is irreducible we get $b = d \in \mathbb{K}$ and $\mathbb{E} = \mathbb{K}$.

(c) $\Longrightarrow$ (a): Let f be irreducible s'-polynomial. Then $\mathbb{K}[x]/f\mathbb{K}[x]$ is an extension of degree $\deg f$. So its is an s'-extension of $\mathbb{K}$ and by (c), $\deg f = 1$. □

Lemma 5.4.3. *Let $\mathbb{K} \leq \mathbb{F}$ be a finite purely inseparable extension. Put $p = \operatorname{char} \mathbb{K}$. If $p = 0$, then $\mathbb{K} = \mathbb{F}$ and if $p \neq 0$, then $\dim_{\mathbb{K}} \mathbb{F} = p^m$ for some $m \in \mathbb{N}$.*

Proof. If $p = 0$, then $\mathbb{K} \leq \mathbb{F}$ is separable and so $\mathbb{K} = \mathbb{F}$. So suppose $p \neq 0$. We proceed by induction on $\dim_{\mathbb{K}} \mathbb{F}$. If $\dim_{\mathbb{K}} \mathbb{F} = 1$, then $\mathbb{K} = \mathbb{F}$. So suppose $\dim_{\mathbb{K}} \mathbb{F} > 1$ and let $b \in \mathbb{F} \setminus \mathbb{K}$. By 5.2.22 there exists $n \in \mathbb{N}$ such that b^{p^n} is separable over $\mathbb{K}$ and $\dim_{\mathbb{K}[b^{p^n}]} \mathbb{K}[b] = p^n$. Since $b^{p^n} \in \mathbb{F}$ and $\mathbb{K} \leq \mathbb{F}$ is purely inseparable, $b^{p^n} \in \mathbb{K}$ and so $\dim_{\mathbb{K}} \mathbb{K}[b] = p^n$. By Homework 3#6 $\mathbb{K}[b] \leq \mathbb{F}$ is purely inseparable and so by induction $\dim_{\mathbb{K}[b]} \mathbb{E} = p^l$ for some $l \in \mathbb{N}$. Thus by the dimension formula 5.1.4(c), $\dim_{\mathbb{K}} \mathbb{E} = p^n p^l = p^{k+l}$. □

Proposition 5.4.4. *Let $\mathbb{K} \leq \mathbb{F}$ be an algebraic extension and s a prime. Suppose that*

(i) Every s'-polynomial over $\mathbb{K}$ has a root in $\mathbb{K}$.

(ii) All polynomials of degree s over $\mathbb{F}$ are reducible.

Then $\mathbb{F}$ is algebraically closed.

Proof. Let $\overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}$ and $b \in \overline{\mathbb{F}}$. We need to show that $b \in \mathbb{F}$. For this let $\mathbb{E}$ a normal closure of $\mathbb{K} \leq \mathbb{K}[b]$ in $\overline{F}$. By 5.3.20(c:a), $K \leq \mathbb{E}$ is finite.

Put $\mathbb{P} = \mathbb{P}(\mathbb{K}, \mathbb{E})$. By 5.3.21 $\mathbb{K} \leq \mathbb{P}$ is purely inseparable and $\mathbb{P} \leq \mathbb{E}$ is Galois. We will show that

$$(*) \qquad \mathbb{P} \leq \mathbb{F}$$

If $\operatorname{char} p = 0$, then $\mathbb{P} = \mathbb{K}$. So suppose $\operatorname{char} K = p$, p a prime. Assume first that $p \neq s$, then by 5.4.3 $\mathbb{K} \leq \mathbb{P}$ is an s'-extension and so by 5.4.2 $\mathbb{P} = \mathbb{K}$. Assume next that $p = s$ and suppose first exits $b \in \mathbb{P} \setminus \mathbb{F}$. By 5.2.24(a),$b^{p^n} \in \mathbb{K}$ for some $n \in \mathbb{N}$. Hence we can choose $n \in \mathbb{Z}^+$ minimal with $b^{p^n} \in \mathbb{F}$. Put $a = b^{p^{n-1}}$. Then $a \in \mathbb{P} \setminus \mathbb{F}$ and $a^p \in \mathbb{F}$. 5.2.19 implies $\deg m_a^{\mathbb{K}} = p = s$, a contradiction to (ii). Thus (*) holds.

Put $\mathbb{S} = \mathbb{S}(\mathbb{K}, \mathbb{E})$. Then by 5.3.21, $\mathbb{K} \leq \mathbb{S}$ is Galois. Put $G = \operatorname{Aut}_{\mathbb{K}}(\mathbb{S})$. Since $\mathbb{K} \leq \mathbb{E}$ is finite, G is finite.

By 2.11.7 there exists a Sylow s-subgroup S of G. Put $\mathbb{L} = \operatorname{Fix}_{\mathbb{S}}(S)$. Then by the FTGT, 5.3.12 $\dim_{\mathbb{L}} \mathbb{S} = |S|$ and so

$$\dim_{\mathbb{K}} \mathbb{L} = \frac{\dim_{\mathbb{K}} \mathbb{S}}{\dim_{\mathbb{K}} \mathbb{L}} = \frac{|G|}{|S|}$$

Since S is a Sylow s-subgroup we conclude that $\mathbb{K} \leq \mathbb{L}$ is a s' extension. Thus by 5.4.2 $\mathbb{L} = \mathbb{K}$ and so $G = S$. Thus G is a s-group. Since $\mathbb{K} \leq \mathbb{S} \cap \mathbb{F} \leq \mathbb{S}$, 5.3.12 implies $\mathbb{S} \cap \mathbb{F} = \operatorname{Fix}_{\mathbb{S}}(H)$ for some $H \leq G$.

Suppose for a contradiction that $H \neq \{\mathrm{id}_{\mathbb{S}}\}$. Then by 2.10.26(a), there exists $T \trianglelefteq H$ with $|H/T| = s$. Put $\mathbb{D} = \operatorname{Fix}_{\mathbb{S}}(T)$. Then $\dim_{\mathbb{S} \cap \mathbb{F}} \mathbb{D} = |H/T| = p$. Let $d \in \mathbb{D} \setminus (\mathbb{S} \cap \mathbb{F})$. Then $\deg m_d^{\mathbb{S} \cap \mathbb{F}} = s$. By 5.2.11 $m_d^{\mathbb{S} \cap \mathbb{F}} = m_d^{\mathbb{F}}$ and so $\deg m_d^{\mathbb{F}} = s$ a contradiction to (ii).

Thus $H = \{\mathrm{id}_{\mathbb{S}}\}$ and so $\mathbb{S} \cap \mathbb{F} = \mathrm{Fix}_{\mathbb{S}}(\mathrm{id}_{\mathbb{S}}) = \mathbb{S}$. Thus $\mathbb{S} \leq \mathbb{F}$. Together with (*) we get $\mathbb{SP} \leq \mathbb{F}$. By 5.2.28(c), $\mathbb{E} = \mathbb{PS}$ and so $\mathbb{E} \leq \mathbb{F}$. Since $b \in \mathbb{E}$ we have $b \in \mathbb{F}$. As $b \in \overline{\mathbb{F}}$ was arbitrary this means, $\mathbb{F} = \overline{\mathbb{F}}$ and so $\mathbb{F}$ is algebraically closed. □

Theorem 5.4.5. *The field of complex numbers is algebraically closed.*

Proof. By the three properties of $\mathbb{R} \leq \mathbb{C}$ listed above we can apply 5.4.4 with $s = 2$. Hence $\mathbb{C}$ is algebraically closed. □

Lemma 5.4.6. *Let $\mathbb{K} \leq \mathbb{E}$ be algebraic and $\overline{\mathbb{K}}$ an algebraic closure of $\mathbb{K}$. Then $\mathbb{E}$ is $\mathbb{K}$-isomorphic to some intermediate field $\tilde{\mathbb{E}}$ of $\mathbb{K} \leq \overline{\mathbb{K}}$.*

Proof. Let $\overline{\mathbb{E}}$ be an algebraic closure of $\mathbb{E}$. Then by 5.3.20(d:a) $\overline{\mathbb{E}}$ is an algebraic closure of $\mathbb{K}$. By 5.2.8(e) there exists an $\mathbb{K}$-isomorphism $\phi : \overline{\mathbb{E}} \to \overline{\mathbb{K}}$. Put $\tilde{\mathbb{E}} = \phi(\mathbb{E})$. □

Lemma 5.4.7. *Up to $\mathbb{R}$-isomorphisms, $\mathbb{C}$ is the only proper algebraic extension of $\mathbb{R}$*

Proof. Note that $\mathbb{C}$ is an algebraic closure of $\mathbb{R}$. So by 5.4.6 any algebraic extension of $\mathbb{R}$ is $\mathbb{R}$-isomorphic to an intermediate field $\mathbb{E}$ of $\mathbb{R} \leq \mathbb{C}$. As $\dim_{\mathbb{R}} \mathbb{C} = 2$, we get $\mathbb{E} = \mathbb{R}$ or $\mathbb{E} = \mathbb{C}$. □

5.5 Finite Fields

In this section we study the Galois theory of finite fields.

Lemma 5.5.1. *Let $\mathbb{F}$ be a finite field and $\mathbb{F}_0$ the subring generated by 1. Then $\mathbb{F}_0 \cong \mathbb{Z}_p$ for some prime p. In particular, $\mathbb{F}$ is isomorphic to a subfield of the algebraic closure of $\mathbb{Z}_p$.*

Proof. Let $p = \mathrm{char}\,\mathbb{F}$. Then $p\mathbb{Z}$ is the kernel of the homomorphism $\mathbb{Z} \to \mathbb{F}$, $n \to n1_{\mathbb{F}}$. Also $\mathbb{F}_0$ is its image and so $\mathbb{F}_0 \cong \mathbb{Z}_p$. □

Theorem 5.5.2. *Let p be a prime, $\mathbb{F}_0$ a field of order p, $\mathbb{F}$ an algebraic closure of $\mathbb{F}_0$ and* $G := \{\mathrm{Frob}_{p^n}^{\mathbb{F}} \mid n \in \mathbb{Z}\}$

(a) G is subgroup of $\mathrm{Aut}(\mathbb{F})$.

(b) $\mathbb{F}_0 = \mathrm{Fix}_{\mathbb{F}}(G) = \mathrm{Fix}_{\mathbb{F}}(\mathrm{Frob}_p)$.

(c) A proper subfield of $\mathbb{F}$ is G-closed if and only if its finite.

(d) All subgroups of G are closed.

(e) $\mathcal{G}$ is a inclusion reversing bijection between the finite subfields of $\mathbb{F}$ and the non-trivial subgroups of G.

(f) Let $q \in \mathbb{Z}, q > 1$. Then $\mathbb{F}$ has a subfield of order q if and only if q is a power of p. If q is a power of p then F has a unique subfield of order q, $\mathbb{F}_q$.

(g) Let $n \in \mathbb{Z}^+$ and $q = p^n$. Then

$$\mathbb{F}_q = \mathrm{Fix}_{\mathbb{F}}(\mathrm{Frob}_q) = \{a \in \mathbb{F} \mid a^q = a\}$$

So $\mathbb{F}_q$ consists exactly of the roots of $x^q - x$.

(h) $\mathbb{F}_{p^m} \leq \mathbb{F}_{p^n}$ if and only if m divides n.

(i) Let $n \in \mathbb{Z}^+, m \in \mathbb{N}$ and $q = p^n$. Then $\mathbb{F}_q \leq \mathbb{F}_{q^m}$ is a Galois extension and

$$\mathrm{Aut}_{\mathbb{F}_q}(\mathbb{F}_{q^m}) = \{\mathrm{Frob}_{q^i} \mid 0 \leq i < m.\}$$

In particular, $\mathrm{Aut}_{\mathbb{F}_q}\mathbb{F}_{q^m}$ is cyclic of order m.

Proof. Note first that $G = \langle \mathrm{Frob}_p \rangle$ is a cyclic subgroup of $\mathrm{Aut}(\mathbb{F})$. Let H be a non-trivial subgroup of G. Then $H = \langle \mathrm{Frob}_q \rangle$ where $q = p^n$ for some $n \in \mathbb{Z}^+$. Put $\mathbb{F}_q = \mathcal{F}H = \mathrm{Fix}(\mathrm{Frob}_q$. Let $b \in \mathbb{F}$. Then $b \in \mathbb{F}_q$ if and only if $b^q = b$. $\mathbb{F}_q$ consist exactly of the roots of $x^q - x$. Note that $(x^q - x)' = qx^{q-1} - 1 = -1$ has no roots and so by 3.7.12 $x^q - x$ has no multiple roots. Hence $|\mathbb{F}_q| = q$. In particular $\mathbb{F}_0 = \mathbb{F}_p$ and (c) holds. Also $\mathbb{F}_0$ is closed and every proper closed subfield of $\mathbb{F}$ is finite.

Since $\mathbb{F}_q = \mathcal{F}H$, $\mathbb{F}_q$ is closed by 5.3.3e. So $\mathcal{F}\mathcal{G}\mathbb{F}_q = \mathbb{F}_q$. Since H is the only subgroup of G with fixed field $\mathbb{F}_q$, $\mathcal{G}\mathbb{F}_q = H = \langle \mathrm{Frob}_q \rangle$. Thus H is closed.

By 5.3.9b, every finite extension of $\mathbb{F}_0$ in $\mathbb{F}$ is closed. So every finite subfield of $\mathbb{F}$ is closed. Thus (c) to (g) are proved.

(g)

$$\mathbb{F}_{p^m} \leq F_{p^n} \Longleftrightarrow \mathcal{G}\mathbb{F}_{p^n} \leq \mathcal{G}\mathbb{F}_{p^m} \Longleftrightarrow \langle \mathrm{Frob}_{p^n} \rangle \leq \langle \mathrm{Frob}_{p^m} \rangle.$$

By 2.6.10b this is the case if and only if m divides n.

(i) Since H is abelian, all subgroups of H are normal. Hence by 5.3.16 (applied to $(\mathbb{F}, \mathbb{F}_q, H)$ in place of $(\mathbb{F}, \mathbb{K}, G)$) $\mathbb{F}_{q^m}$ is H-stable. Thus by 5.3.13 (again applied with H in place of G) $\mathbb{F}_q \leq \mathbb{F}_{q^m}$ is Galois and $\mathrm{Aut}_{\mathbb{F}_q}F_{q^m} = H^{\mathbb{F}_{q^m}}$. By 5.3.14b,

$$H^{q^m} \cong H/\mathcal{F}\mathbb{F}_{q^m} = \langle \mathrm{Frob}_q \rangle / \langle \mathrm{Frob}_{q^m} \rangle \cong \mathbb{Z}/m\mathbb{Z}.$$

Thus (i) holds. □

5.6 Transcendence Basis

Definition 5.6.1. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension and $s = (s_i)_{i \in I}$ a family of elements in $\mathbb{F}$. We say that $(s_i, i \in I)$ is* $\mathbb{K}$-algebraically independent *if the evaluation homomorphism:*

$$\mathbb{K}[x_i, i \in I] \to \mathbb{K}[s_i, i \in I], f \to f(s)$$

is 1-1.

A subset S of $\mathbb{F}$ is called algebraically independent over $\mathbb{K}$, if $(s)_{s \in S}$ is linearly independent.

Recall here that if $f = \sum_{n\in J} f_n x^n$, then $f(s) = \sum_{n\in J} f_n s^n$. Here $J = \bigoplus_{i\in I} \mathbb{N}$, $x^n = \prod_{i\in I} x_i^{n_i}$, $s^n = \prod_{i\in I} s_i^{n_i}$.

Note that s is algebraically dependent if and only if there exists $0 \neq f \in \mathbb{K}[x_i, i \in I]$ with $f(s) = 0$. In particular, if $|I| = 1$, s is algebraically dependent over $\mathbb{K}$ if and only if s is algebraic over $\mathbb{K}$. Also since each $f \in \mathbb{K}[x_i, i \in I]$ only involves finitely many variables, s is algebraically independent if and only if every finite subfamily is algebraically independent.

If $s_i = s_j$ for some $i \neq j$ then s is a root of $x_i - x_j$ and so s is algebraically dependent. On the other hand, if the $s_i, i \in I$ are pairwise distinct, s is linear independent if and only if the underlying set $S = \{s_i \mid i \in I\}$ is algebraically independent over $\mathbb{K}$.

Lemma 5.6.2. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension and $s = (s_i)_{i\in I}$ be algebraically independent over $\mathbb{K}$. Then there exists a unique $\mathbb{K}$-isomorphism $\tilde{\Phi}_s : \mathbb{K}(x_i, i \in I) \to \mathbb{K}(s_i, i \in I)$ with $\alpha(x_i) = s_i$ for all $i \in I$. Moreover, $\tilde{\Phi}_s(\frac{f}{g}) = f(s)g(s)^{-1}$ for all $f, g \in \mathbb{K}[x_i, i \in I]$, $g \neq 0$.*

Proof. By definition of algebraic independent, the map $\Phi_s : \mathbb{K}[x_i, i \in I] \to K[s_i, i \in I]$ 1-1. It is easy to see that Φ_s it is onto and then that

$$\tilde{\Phi}_s : \mathbb{K}(x_i, i \in I) \to \mathbb{K}(s_i, i \in I), \frac{f}{g} \to f(s)g(s)^{-1}$$

is well-defined isomorphism. Clearly $\tilde{\Phi}_s(x_i) = s_i$ and $\tilde{\Phi}_s(k) = k$. So Φ_s exists. The uniqueness is readily verified. □

Lemma 5.6.3. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension.*

(a) Let S and T disjoint subsets of $\mathbb{F}$. Then $S \cup T$ is algebraically independent over $\mathbb{K}$ if and only if S is algebraically independent over $\mathbb{K}$ and T is algebraically independent over $\mathbb{K}(S)$.

(b) Let $S \subseteq \mathbb{F}$ be algebraically independent over $\mathbb{K}$ and let $b \in \mathbb{F} \setminus S$. Then $S \cup \{b\}$ is algebraically independent over $\mathbb{K}$ if and only if b is transcendental over $\mathbb{K}$.

Proof. (b) By 5.6.2 $S \cup T$ is algebraically independent if and only if the there exists an $\mathbb{K}$-isomorphism $\mathbb{K}(x_r, r \in S \cup T) \to \mathbb{K}(S \cup T)$ with $x_r \to r, \forall r \in S \cup T$. S is algebraically independent over $\mathbb{K}$ and T algebraically independent over $\mathbb{K}(S)$ is equivalent to the existence of a $\mathbb{K}$-isomorphism $\mathbb{K}(x_s, s \in S)(x_t, t \in T) \to \mathbb{K}(S)(T)$ with $x_s \to s, \forall s \in S$ and $x_t \to t, \forall t \in T$. Since $\mathbb{K}(S \cup T) = \mathbb{K}(S)(T)$ we conclude from 5.2.29 that the two property are equivalent.

(b) Follows from (a) applied to $T = \{b\}$. □

Definition 5.6.4. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension. A* transcendence basis *for $\mathbb{K} \leq \mathbb{F}$ is a algebraically independent subset S of $\mathbb{K} \leq \mathbb{F}$ such that $\mathbb{F}$ is algebraic over $\mathbb{K}(S)$.*

Lemma 5.6.5. *Let $\mathbb{K} \leq \mathbb{F}$ be field extension ,$S \subseteq \mathbb{F}$ and suppose that S algebraically independent over $\mathbb{K}$.*

(a) S is a transcendence basis if and only if S is a maximal $\mathbb{K}$-algebraically independent subset of $\mathbb{F}$.

(b) S is contained in a transcendence basis for $\mathbb{K} \leq \mathbb{F}$.

(c) $\mathbb{K} \leq \mathbb{F}$ has a transcendence basis.

Proof. (a) S is a maximal algebraically independent set if and only if $S \cup \{b\}$ is algebraically dependent for all $b \in \mathbb{F} \setminus S$. By 5.6.3b, this is the case if and only if each $b \in \mathbb{F}$ is algebraic over $\mathbb{K}(S)$.

(b) Let $\mathcal{M}$ be the set of $\mathbb{K}$-algebraically independent subsets of $\mathbb{F}$ containing S. Since $S \in \mathcal{M}$, $\mathcal{M}$ is not empty. Order $\mathcal{M}$ by inclusion. Then $\mathcal{M}$ is a partially ordered set. We would like to apply Zorn's lemma. So we need to show that every chain $\mathcal{D}$ of $\mathcal{M}$ has an upper bound. Note that the elements of $\mathcal{D}$ are subsets on $\mathbb{F}$. So we can build the union $D := \bigcup \mathcal{D}$. Then $E \subseteq D$ for all $E \in \mathcal{D}$. Thus D is an upper bound for $\mathcal{D}$ once we establish that $D \in \mathcal{M}$. That is we need to show that D is algebraically independent over $\mathbb{K}$. As observed before we just this amounts to showing that each finite subset $J \subseteq D$ is algebraically independent. Now each $j \in J$ lies in some $E_j \in \mathcal{D}$. Since $\mathcal{D}$ is totally ordered, the finite subset $\{E_s \mid j \in J\}$ of $\mathcal{D}$ has a maximal element E. Then $j \in E_j \subseteq E$ for all $j \in J$. So $J \subseteq E$ and as E is algebraically independent, J is as well.

Hence every chain in $\mathcal{M}$ has an upper bound. By Zorn's Lemma A.6 $\mathcal{M}$ has a maximal element T. By (a) T is a transcendence basis and by definition of $\mathcal{M}$, $S \subseteq T$.

(c) follows from (b) applied to $S = \emptyset$. □

Proposition 5.6.6. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension and S and T transcendence basis for $\mathbb{K} \leq \mathbb{F}$. Then $|S| = |T|$. $|S|$ is called the* transcendence degree *of $\mathbb{F} \leq \mathbb{K}$ and is denoted by* $\mathrm{tr} - \deg_{\mathbb{K}} \mathbb{F}$.

Proof. Well order S and T. For $s \in S$ define $s^- := \{b \in S \mid b < s\}$ and $s^+ := \{b \in S \mid b \leq s\}$. Similarly define $t^{\pm}$ for $t \in T$. Let $s \in S$. As $\mathbb{K}(I) \leq \mathbb{F}$ is algebraic there exists a finite subset $J \subseteq T$ such that s is algebraic over $\mathbb{K}(J)$ and so also algebraic over $\mathbb{K}(s^-, J)$. Let j be the maximal element of J. Then $J \subseteq j^+$ and so s is algebraic over $\mathbb{K}(s^-, j^+)$. Hence we can define a function $\phi : S \to T$, where $\phi(s) \in T$ is minimal with respect to s being algebraic over $\mathbb{K}(s^-, \phi(s)^+)$. Similarly define $\psi : T \to S$.

We will show that ϕ and ψ are inverse to each other. Let $s \in S$ and put $t = \phi(s)$. Put $\mathbb{L} := \mathbb{K}(s^-, t^-)$ We claim that s is transcendental over $\mathbb{L}$. If not, we can choose J as above with $J \subseteq t^-$. But then s is algebraic over $\mathbb{K}(s^-, j^+)$. Since $j < t$ this contradicts the minimal choice of t.

Thus s is transcendental over $\mathbb{L}$. Note that s is algebraic over $\mathbb{K}(s^-, t^+) = \mathbb{L}(t)$, So if t would be algebraic over $\mathbb{L}$ it also would be algebraic over $\mathbb{L}$, a contradiction. Hence t is transcendental over $\mathbb{L} = \mathbb{K}(t^-, s^-)$. Since t is algebraic over $\mathbb{K}(t^-, \psi(t)^+)$ we get $\psi(t)^+ \not\subseteq s^-$ and so $\psi(t) \not< s$.

Since s is algebraic over $\mathbb{L}(t)$, 5.6.5(b) implies that t, s are algebraic dependent over $\mathbb{L}$. Since s is transcendental over $\mathbb{L}$ another application of 5.6.5(b) shows that t is algebraic over $\mathbb{L}(s) = \mathbb{K}(s^+, t^-)$. Thus by definition of ψ, $\psi(t) \leq s$. Together with $\psi(t) \not< s$ this gives, $\psi(t) = s$. Therefore $\psi \circ \phi = \mathrm{id}_S$. By symmetry $\phi \circ \psi = \mathrm{id}_T$ and so ϕ is a bijection. □

Example 5.6.7. Let $\mathbb{K}$ be a field and let s be transcendental over $\mathbb{K}$. Let $\mathbb{F}$ be an algebraic closure of $\mathbb{K}(s)$. Put $s_0 = s$ and inductively let s_{i+1} be a root of $x^2 - s_i$ in $\mathbb{F}$. Then $s_i = s_{i+1}^2$ and so $\mathbb{K}(s_i) \leq \mathbb{K}(s_{i+1})$. Note that s_{i+1} is transcendental over $\mathbb{K}$ and so $\mathbb{K}(s_i) = \mathbb{K}(s_{i+1}^2) \neq \mathbb{K}(s_i)$. Put $\mathbb{E} = \bigcup_{i=0}^{\infty} \mathbb{K}(s_i)$. Then $\mathbb{K}(s_i) \leq \mathbb{E}$ is algebraic. Thus each $\{s_i\}$ is a transcendence basis for $\mathbb{K} \leq \mathbb{E}$. We claim that that $\mathbb{K}(b) \neq \mathbb{E}$ for all $b \in \mathbb{E}$. Indeed, $b \in \mathbb{K}(s_i)$ for some i and so $\mathbb{K}(b) \leq \mathbb{K}(s_i) \subseteq \mathbb{E}$.

5.7 Algebraically Closed Fields

In this section we study the Galois theory of algebraically closed field.

Lemma 5.7.1. *Let $\phi : \mathbb{K}_1 \to \mathbb{K}_2$ be a field isomorphism and $\mathbb{F}_i$ an algebraically close field with $\mathbb{K}_i \leq \mathbb{F}_i$. Suppose that $\operatorname{tr\,deg}_{\mathbb{K}_1} \mathbb{F}_1 = \operatorname{tr\,deg}_{\mathbb{K}_2} \mathbb{F}_2$. Let S_i be a transcendence basis for $\mathbb{F}_i$ over $\mathbb{K}_i$ and $\lambda : S_1 \to S_2$ a bijection. Then there exists an isomorphism $\psi : \mathbb{F}_1 \to \mathbb{F}_2$ with $\psi|_{\mathbb{K}_1} = \phi$ and $\psi|_{S_1} = \lambda$.*

Proof. Let S_i be a transcendence basis for $\mathbb{F}_i : \mathbb{K}_i$. By assumption there exists a bijection $\lambda : S_1 \to S_2$. By 5.6.2 there exists a unique isomorphism

$$\delta : \mathbb{K}_1(S_1) \to \mathbb{K}_2(S_2)$$

with $\delta(k) = \phi(k), \forall k \in K_1$ and $\delta(s) = \phi(s), \forall s \in S_1$. Since $\mathbb{F}_i : K_i(S_i)$ is algebraically closed, $\mathbb{F}_i$ is an algebraically closure of $\mathbb{K}_i(S_i)$. Hence by 5.2.8(a), δ extends to an isomorphism $\psi : \mathbb{F}_1 \to F_2$. □

Lemma 5.7.2. *Let $\mathbb{K} \leq \mathbb{F}$ be a field extension and suppose that $\mathbb{F}$ is algebraically closed. Then $\operatorname{Aut}_{\mathbb{K}}(\mathbb{F})$ acts transitively on the set of elements in $\mathbb{F}$ transcendental over $\mathbb{K}$.*

Proof. Let $s_i \in \mathbb{F}$, i=1,2, be transcendental over $\mathbb{K}$. By 5.6.5b there exists a transcendence basis S_i for $\mathbb{K} \leq \mathbb{F}$ with $s_i \in S_i$. Let $\lambda : S_1 \to S_2$ be a bijection with $\lambda(s_1) = s_2$. By 5.7.1 there exists $\psi \in \operatorname{Aut}_{\mathbb{K}}\mathbb{F}$ with $\psi(s) = \lambda(s)$ for all $s \in S_1$. Thus $\psi(s_1) = s_2$. □

Example 5.7.3. By results from analysis, both π and e are transcendental over $\mathbb{Q}$. Since $\mathbb{C}$ is algebraically closed we conclude from 5.7.2 that there exists $\alpha \in \operatorname{Aut}_{\mathbb{Q}}(\mathbb{C})$ with $\alpha(\pi) = e$.

Definition 5.7.4. *Let $\mathbb{K}$ be the field and $\mathbb{K}_0$ the intersection of all the subfield. Then $\mathbb{K}_0$ is called the* base field *of $\mathbb{K}$. of $\mathbb{K}$.*

Lemma 5.7.5. *Let $\mathbb{K}$ be the field and $\mathbb{K}_0$ the base field of $\mathbb{K}$. Put $p = \operatorname{char} \mathbb{K}$. If $\operatorname{char} p = 0$ then $\mathbb{K}_0 \cong \mathbb{Q}$ and if p is a prime then $\mathbb{K}_0 \cong \mathbb{Z}/p\mathbb{Z}$*

Proof. Let $Z = \{n1_F \mid n \in \mathbb{Z}\}$. The Z is a subring and $\mathbb{K}_0$ is the field of fraction of Z. If $p = 0$, then $Z \cong \mathbb{Z}$ and so $\mathbb{K}_0 \cong \mathbb{Q}$ and if $p > 0$, then $Z \cong Z_p$ and $\mathbb{K}_0 = Z$. □

Corollary 5.7.6. *(a) Let $\mathbb{K}$ be a field. Then for each cardinality c there exists a unique (up to $\mathbb{K}$-isomorphism) algebraically closed $\mathbb{F}$ with $\mathbb{K} \leq \mathbb{F}$ and $\operatorname{tr\,deg} \mathbb{F} : \mathbb{K} = c$. Moreover, $\mathbb{F}$ is isomorphic to the algebraic closure of $\mathbb{K}(x_i, i \in I)$, where I is a set with $|I| = c$.*

(b) Let $p = 0$ or a prime and c a cardinality. Then there exists a unique (up to isomorphism) algebraically closed field $\mathbb{F}$ with characteristic p and transcendence degree c over its base field. Moreover, the algebraic closure of $\mathbb{F}_p(x_i, i \in I)$, where I is a set with cardinality I and $\mathbb{F}_p$ is $\mathbb{Q}$ respectively $\mathbb{Z}_p$, is such a field.

Proof. Follows immediately from 5.7.1 □

Lemma 5.7.7. *Let $\mathbb{K}$ be a field. Then the following are equivalent.*

(a) $\mathbb{K}$ has no proper purely inseparable field extension.

(b) Let $\overline{\mathbb{K}}$ be an algebraic closure of $\mathbb{K}$. Then $\mathbb{K} \leq \overline{\mathbb{K}}$ is Galois.

(c) All polynomials over $\mathbb{K}$ are separable.

(d) $\operatorname{char} \mathbb{K} = 0$ or $\operatorname{char} \mathbb{K} = p \neq 0$ and for each $b \in \mathbb{K}$ there exists $d \in \mathbb{K}$ with $d^p = b$.

(e) $\operatorname{char} \mathbb{K} = 0$ or $\operatorname{char} \mathbb{K} = p \neq 0$ and Frob_p is an automorphism.

Proof. (a) $\Longrightarrow$ (b): Since $\overline{\mathbb{K}}$ is the algebraic closure of $\mathbb{K}$, $\mathbb{K} \leq \overline{\mathbb{K}}$ is algebraic and normal. Put $\mathbb{P} := \mathbb{P}(\mathbb{K}, \overline{\mathbb{K}})$. Since $\mathbb{K} \leq \overline{\mathbb{K}}$ is normal, 5.2.24(e) implies that $\mathbb{P} \leq \overline{\mathbb{K}}$ is separable. Since $\mathbb{K} \leq \mathbb{P}$ is purely inseparable (a) gives $\mathbb{K} = \mathbb{P}$. Hence by $\mathbb{K} \leq \overline{\mathbb{K}}$ is normal and separable and thus by 5.3.17 $\mathbb{K} \leq \overline{\mathbb{K}}$ is Galois.

(b) $\Longrightarrow$ (c): Since $\mathbb{K} \leq \overline{\mathbb{K}}$ is Galois, 5.3.17 implies that $\mathbb{K} \leq \overline{\mathbb{K}}$ is separable. Let $f \in \mathbb{K}[x]$ be irreducible. Then f has root in $\overline{\mathbb{K}}$. This root is separable over $\mathbb{K}$ and so f is separable.

(c) $\Longrightarrow$ (d): Let $b \in \mathbb{K}$ and f an irreducible monic factor of $x^p - b$. Then f has a unique root in $\bar{\mathbb{K}}$ and is separable. Thus $f = x - d$ for some $d \in \mathbb{K}$ with $d^p = b$.

(d) $\Longrightarrow$ (e): By 5.2.17 Frob_p is a monomorphism. By (d) Frob_p is onto.

(e) $\Longrightarrow$ (a): Let $\mathbb{K} \leq \mathbb{F}$ be purely inseparable. Let $b \in \mathbb{F}$. Then $d := b^{p^n} \in \mathbb{K}$ for some $n \in \mathbb{N}$. Thus $b = \operatorname{Frob}_p^{-n}(d) \in \mathbb{K}$ and $\mathbb{F} = \mathbb{K}$. □

Definition 5.7.8. *A field $\mathbb{K}$ which fulfills one and so all of the equivalent conditions in 5.7.7 is called* perfect.

Lemma 5.7.9. *Finite fields and algebraically closed fields are perfect.*

Proof. Let $\mathbb{K}$ be a field. If $\mathbb{K}$ is finite, then as Frob_p is one to one, its onto. If $\mathbb{K}$ is algebraically closed, Frob_p is an automorphism by 5.2.17d. □

Proposition 5.7.10. *Let $\mathbb{K} \leq \mathbb{F}$ field extension with $\mathbb{F}$ algebraically closed. Put $G := \operatorname{Aut}_{\mathbb{K}}(\mathbb{F})$, $\mathbb{P} := \mathbb{P}(\mathbb{K}, \mathbb{F})$ and let $\mathbb{A}$ be the set elements in $\mathbb{F}$ which are algebraic over $\mathbb{K}$. Let $\mathbb{K} \leq \mathbb{E} \leq \mathbb{F}$ with $\mathbb{E} \neq \mathbb{F}$*

(a) $\mathbb{A}$ is an algebraic closure of $\mathbb{K}$ and $\mathbb{K} \leq \mathbb{A}$ is normal.

(b) If $\mathbb{E}$ is G-stable then $\mathbb{E}^G = \operatorname{Aut}_{\mathbb{K}}(\mathbb{E})$.

(c) $\mathbb{E}$ *is* G*-stable if and only* $\mathbb{K} \leq \mathbb{E}$ *is normal.*

(d) $\mathrm{Fix}_{\mathbb{F}}(G) = \mathbb{P}$.

(e) $\mathbb{E}$ *is* G*-closed if and only if* $\mathbb{E} \leq \mathbb{F}$ *is Galois and if and only if* $\mathbb{E}$ *is perfect.*

(f) Suppose $\mathbb{A} \neq \mathbb{F}$. *Then* $\mathrm{Aut}_{\mathbb{A}}\mathbb{F}$ *is the unique minimal non-trivially closed normal subgroup of* G.

Proof. (a) Note that $\mathbb{K} \leq \mathbb{A}$ is algebraic. Let $f \in \mathbb{K}[x]$ be a non-constant polynomial. Since $\mathbb{F}$ is algebraically closed, f has a root $b \in \mathbb{F}$. Then b is algebraic over $\mathbb{K}$ and so $b \in \mathbb{A}$. Thus f has a root in $\mathbb{A}$ and so by definition (see 5.1.15), $\mathbb{A}$ is an algebraic closure of $\mathbb{K}$. If g is a polynomial with a root on $\mathbb{A}$, then g splits over $\mathbb{A}$, since $\mathbb{A}$ is algebraically closed. Thus $\mathbb{K} \leq \mathbb{A}$ is normal.

(b) By 5.7.1 every $\phi \in \mathrm{Aut}_{\mathbb{K}}\mathbb{E}$ can be extended to some $\psi \in \mathrm{Aut}_{\mathbb{K}}\mathbb{F}$. So (b) holds.

(c) Suppose $\mathbb{K} \leq E$ is normal, then by 5.2.10a, $\mathbb{E}$ is G-stable.

Suppose that $\mathbb{K} \leq \mathbb{E}$ is G-stable. We will first show that $\mathbb{E} \leq \mathbb{A}$. Suppose not and pick $e \in \mathbb{E}$ such that e is transcendental over $\mathbb{K}$. By 5.7.2 Ge consists of all the transcendental elements in $\mathbb{F}$. As $\mathbb{E}$ is G-stable, $Ge \subseteq \mathbb{F}$. So $\mathbb{E}$ contains all the transcendental elements.

Let $b \in \mathbb{A} \cap \mathbb{E}$. Suppose for a contradiction that $b+e$ is algebraic over $\mathbb{K}$. Then $\mathbb{K} \leq \mathbb{K}(b)$ and $\mathbb{K}(b) \leq \mathbb{K}(b, b+e)$ are algebraic.Thus by 5.1.11 also $\mathbb{K} \leq \mathbb{K}(b, b+e)$ is algebraic and so $e = (e+b) - b$ is algebraic over $\mathbb{K}$, a contradiction. Hence $b+e$ is transcendental over $\mathbb{K}$. Thus $b + e \in \mathbb{E}$ and $b \in \mathbb{K}(b+e, e) \leq \mathbb{E}$. It follows that $\mathbb{F} = \mathbb{E}$, a contradiction to the assumptions.

Hence $\mathbb{E} \leq \mathbb{A}$. . So by (b)

$$(*) \qquad G^{\mathbb{A}} = \mathrm{Aut}_{\mathbb{K}}(\mathbb{A}).$$

Hence $\mathbb{E}$ is $\mathrm{Aut}_{\mathbb{K}}(\mathbb{A})$-stable and so by 5.3.15 $\mathbb{K} \leq \mathbb{E}$ is normal.

(d) Let $b \in \mathbb{F} \setminus \mathbb{A}$. Then also $b + 1 \in \mathbb{F} \setminus \mathbb{A}$ and by 5.7.2 there exists $\sigma \in G$ with $\sigma(b) = b + 1 \neq b$. Thus $b \notin \mathrm{Fix}_{\mathbb{F}}(G)$ and so $\mathrm{Fix}_{\mathbb{F}}(G) \leq \mathbb{A}$. Thus

$$(**) \qquad \mathrm{Fix}_{\mathbb{F}}(G) = \mathrm{Fix}_{\mathbb{A}}(G^{\mathbb{A}}) \overset{(*)}{=} \mathrm{Fix}_{\mathbb{A}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{A})) \overset{5.2.24(\mathrm{f})}{=} \mathbb{P}$$

(e) $\mathbb{E}$ is G-closed if and only if

$$(1) \qquad \mathrm{Fix}_{\mathbb{F}}(\mathrm{Aut}_{\mathbb{E}}(\mathbb{F})) = \mathbb{E}$$

and so if and only if $\mathbb{E} \leq \mathbb{F}$ is Galois. Since (1) does not dependent on $\mathbb{K}$ we may replace $\mathbb{K}$ by $\mathbb{E}$. Then $G = \mathrm{Aut}_{\mathbb{E}}(\mathbb{F})$ and (1) becomes

$$(2) \qquad \mathrm{Fix}_{\mathbb{F}}(G) = \mathbb{K}$$

By (**) $\mathrm{Fix}_{\mathbb{F}}(G) = \mathrm{Fix}_{\mathbb{A}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{A}))$ and so (2) is equivalent to

$$\mathrm{Fix}_{\mathbb{A}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{A})) = \mathbb{K} \tag{3}$$

By definition of a Galois extension (3) holds if and only if $\mathbb{K} \leq \mathbb{A}$ is Galois. Since $\mathbb{A}$ is an algebraic closure of $\mathbb{K}$, 5.7.7 implies that $\mathbb{K} \leq \mathbb{A}$ is Galois if and only if $\mathbb{K}$ is perfect. Since $\mathbb{E} = \mathbb{K}$, (e) is proved.

(f) Let H be a closed normal subgroup of G with $H \neq \{\mathrm{id}_{\mathbb{F}}\}$. Then $\mathrm{Fix}_{\mathbb{F}}(H)$ is 5.3.14(e), $\mathrm{Fix}_{\mathbb{F}}(H)$ is G-stable. Since $H \neq G$, $\mathrm{Fix}_{\mathbb{F}}(H) \neq \mathbb{F}$. So by (c), $\mathbb{K} \leq \mathrm{Fix}_{\mathbb{F}}(H)$ is normal and so algebraic. Hence $\mathrm{Fix}_{\mathbb{F}}(H) \leq \mathbb{A}$ and $\mathrm{Aut}_{\mathbb{A}}(\mathbb{F}) \leq \mathrm{Aut}(\mathrm{Fix}_{\mathbb{F}}(H))$. Since H is closed, $\mathrm{Aut}(\mathrm{Fix}_{\mathbb{F}}(H)) = H$ and so $\mathrm{Aut}_{\mathbb{A}}(\mathbb{F}) \leq H$.

By 5.3.3(h), $\mathrm{Aut}_{\mathbb{A}}(\mathbb{K})$ is closed in G. By (a) $\mathbb{K} \leq \mathbb{A}$ is normal and so by 5.3.14(e) $\mathrm{Aut}_{\mathbb{A}}(\mathbb{F})$ is a normal subgroup of G. By 5.7.9 $\mathbb{A}$ is perfect and so by (e), $\mathbb{A} \leq \mathbb{F}$ is Galois. Thus $\mathrm{Fix}_{\mathbb{A}}(\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})) = \mathbb{A} \neq \mathbb{F}$ and so $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F}) \neq \{\mathrm{id}_{\mathbb{F}}\}$. Thus $\mathrm{Aut}_{\mathbb{K}}(\mathbb{F})$ is a non-trivial, closed subgroup of G. □

Chapter 6

Multilinear Algebra

Throughout this chapter ring means commutative ring with identity $1 \neq 0$. All modules are assumed to be unitary. We will write (non)-commutative ring for a ring which might not be commutative.

6.1 Multilinear functions and Tensor products

Let $(M_i, i \in I)$ be a family of sets. For $J \subseteq I$ put $M_J = \prod_{j \in J} M_j$ and for $m = (m_i)_{i \in I} \in M_I$ put $m_J = (m_j)_{j \in J} M_J$. If $I = J \cup K$ with $L \cap K = \emptyset$, the map $M_I \to M_J \times M_K, m \to (m_J, m_K)$ is a bijection. We use this canonical bijection to identify M_I with $M_J \times M_K$.

Let W be a set and $f : M_I \to W$ a function. Let $b \in M_K$. Then we obtain a function a function $f_b : M_J \to W, a \to f(a, b)$.

Definition 6.1.1. *Let R a ring, $M_i, i \in I$ a family of R-modules and W an R-module. Let $f : M_I \to W$ be a function. f is* R-multilinear *if for all $i \in I$ and all $b \in M_{I-i}$ the function*

$$f_b : M_i \to W, a \to f(a, b)$$

is R-linear.

Note here that f_b R-linear just means $f(ra, b) = rf(a, b)$ and $f(a+\tilde{a}, b) = f(a, b)+f(\tilde{a}, b)$ for all $r \in R, a \in M_i, b \in M_{I-i}$ and $i \in I$.

The function $f : R^n \to R, \quad (a_1, a_2, \ldots, a_n) \to a_1 a_2 \ldots a_n$ is multilinear. But the function $g : R^n \to R, \quad (a_1, \ldots, a_n) \to a_1$ is not R-linear.

Lemma 6.1.2. *Let $M_i, i \in I$ be a family of R-modules, $f : M_I \to W$ an R-multilinear map, $I = J \uplus K$ and $b \in M_K$. Then $f_b : M_J \to W$ is R-multilinear.*

Proof. Let $j \in J$ and $a \in M_{J-j}$. Then $(a, b) \in M_{I-j}$ and $(f_b)_a = f_{(a,b)}$ is R-linear. So f_b is R-multilinear. □

Lemma 6.1.3. *Let R a ring, $M_i, i \in I$ a finite family of R-modules, W an R-module and $f : M_I \to W$ be a function. Then f is multilinear if and only if*

$$f((\sum_{j\in J_i} r_{ij}m_{ij})_{i\in I}) = \sum_{\alpha\in J_I} (\prod_{i\in I} r_{i\alpha(i)}) f((m_{i\alpha(i)})_{i\in I})$$

whenever $(J_i, i \in I)$ is a family of sets, $m_{ij} \in M_i$ and $r_{ij} \in R$ for all $i \in I$ and $j \in J_i$.

Proof. Suppose first that f is multilinear. If $|I| = 1$ we need to show that $f(\sum_{j\in J} r_j m_j) = \sum_{j\in J} r_j f(m_j)$ But this follows easily from the fact that f is linear and induction on J. So suppose that $|I| \geq 2$, let $s \in I$, $K = I - s$. Then by induction

$$f((\sum_{j\in J_i} r_{ij}m_{ij}) \overset{\text{definition of } f_b}{=} f_{\sum_{j\in J_s} r_{sj}m_{sj}}((\sum_{j\in J_i} (r_{ij}m_{ij})_{i\in K}$$

$$= \sum_{\alpha\in J_K} (\prod_{i\in K} r_{i\alpha(i)} f_{\sum_{j\in J_s} r_{sj}m_{sj}}(m_{i\alpha(i)})_{i\in K}$$

$$= \sum_{\alpha\in J_K} (\prod_{i\in K} r_{i\alpha(i)} f(\sum_{j\in J_s} r_{sj}m_{sj}, (m_{i\alpha(i)})_{i\in K}$$

$$\sum_{\alpha\in J_I} \prod_{i\in I} r_{i\alpha(i)} f(m_{i\alpha(i)})$$

The other direction is obvious. □

Example: Suppose $f : M_1 \times M_2 \times M_3 \to W$ is multilinear.
Then

$$f(m_{11} + 2m_{12}, 4m_{21}, 3m_{31} + m_{32} =$$

$$= 12f(m_{11}, m_{21}, m_{31}) + 4f(m_{11}, m_{21}, m_{32}) + 24f(m_{12}, m_{21}, m_{31}) + 8f(m_{12}, m_{21}, m_{32})$$

Definition 6.1.4. *Let R be a ring and $M_i, i \in I$ a family of R-modules. A* tensor product *for $(M_i, i \in I)$ over R is a R-multilinear map $f : M_I \to W$ so that for each multilinear map $g : M_I \to \tilde{W}$ there exists a unique R-linear $\breve{g} : W \to \tilde{W}$ with $g = \breve{g} \circ f$.*

Lemma 6.1.5. *Let R be a ring and $(M_i, i \in I)$ a family of R-modules. Then $(M_i, i \in I)$ has a tensor product over R. Moreover, it is unique up to isomorphism, that is if $f_i : M_I \to W_i$, i=1,2, are tensor products, than there exists a R-linear isomorphism $g : W_1 \to W_2$ with $f_2 = g \circ f_1$.*

Proof. Let $F = F_R(M_I)$, the free module on the set M_I. So F has a basis $z(m), m \in M_I$. Let D be the R-submodule if F generated by the all the elements in F of the form

$$z(ra, b) - rz(a, b)$$

and

$$z(a, b) + z(\tilde{a}, b) - z(a + \tilde{a}, b)$$

where $r \in R$, $a \in M_i$, $b \in M_{I-i}$ and $i \in I$.

Let $W = F/D$ and define $f : M_I \to W, m \to z(m) + D$.

To check that f is multilinear we compute

$$f(ra, b) - rf(a, b) = (z(ra, b) + D) - r(z(a, b) + D) = (z(ra, b) - rz(a, b)) + D = D = 0_W$$

and

$$f(a+\tilde{a}, b)-f(a, b)-f(\tilde{a}, b) = (z(a+\tilde{a}, b)+D)-(z(a, b)+D)-z(\tilde{a}, b)+D) = (z(a+\tilde{a}, b)-z(a, b)-z(\tilde{a}, b))+D = D =$$

So f is R-.multilinear.

To verify that f is a tensor product let $\tilde{f} : M_I \to \tilde{W}$ by R-multilinear. Since F is a free with basis $z(m), m \in M$. There exists a unique R-linear map $\tilde{g} : F \to \tilde{W}$ with $\tilde{g}(z(m)) = \tilde{f}(m)$ for all $m \in M_I$. We claim that $D \leq \ker \tilde{g}$. Indeed

$\tilde{g}(z(ra, b) - rz(a, b)) = \tilde{g}(z(ra, b) - r\tilde{g}(z(a, b) = \tilde{f}(ra, b) - r\tilde{f}(a, b)$,

Here the first equality holds since $\tilde{g}$ is R-linear and the second since $\tilde{f}$ is multilinear.

Similarly $\tilde{g}(z(a + \tilde{a}) - z(a, b) - z(\tilde{a}, b)) = \tilde{g}(z(a + \tilde{a})) - \tilde{g}(z(a, b)) - \tilde{g}(z(\tilde{a}, b)) = \tilde{f}(a + \tilde{a}) - \tilde{f}(a, b) - \tilde{f}(\tilde{a}, b) = 0$.

Hence $\ker \tilde{g}$ contains all the generators of D and since $\ker \tilde{g}$ is an R-submodule of F, $D \leq \ker tildeg$. Thus the map $g : W \to \tilde{W}, e + D \to \tilde{g}(e)$ is well defined and R-linear. Note that $g(f(m)) = \tilde{g}(f(m)) = \tilde{g}(z(m)) = \tilde{f}(m)$ and so $\tilde{f} = g \circ f$. To show the uniqueness of g suppose that $h : W \to \tilde{W}$ is R-linear with $\tilde{f} = h \circ f$. Define $\tilde{h} : F \to \tilde{W}$ by $\tilde{h}(e) = h(e + D)$. Then h is R linear and $\tilde{h}(z(m)) = h(z(m) + D) = h(f(m)) = \tilde{f}(m) = \tilde{g}(z(m))$. Since $z(m)$ is a basis for F this implies $\tilde{h} = \tilde{g}$. Thus $g(e + D) = \tilde{g}(e) = \tilde{h}(e) = h(e + D)$ and $g = h$, as required.

So f is indeed a tensor product.

Now suppose that $f_i : M_I \to W_i$,i=1,2 are tensor products for $(M_i, i \in I$ over R. Let $\{1, 2\} = \{i, j\}$. Since f_i is a tensor product and f_j is multilinear, there exists $g_i : W_i \to W_j$ with $f_j = g_i f_i$. Then $(g_j g_i)f_i = g_j(g_i f_i) = g_j f_j = f_i$. Note that also $\mathrm{id}_{W_i} f_i = f_i$ and so the uniqueness assertion in the definition of the tensor product implies $g_j g_i = \mathrm{id}_{W_i}$. Hence g_1 and g_2 are inverse to each other and g_1 is a R-linear isomorphism. □

Let $(M_i, i \in I)$ be a family of R-modules and $f : M_I \to W$ a tensor product. We denote W by $\bigotimes_R^{i \in I} M_i$ $f((m_i)_{i \in I}$ by $\otimes_{i \in I} m_i$. Also if there is no doubt about the the ring R and the set I in question, we just use the notations $\bigotimes M_i$, $\otimes m_i$ and (m_i)

If $I == \{1, 2 \ldots, n\}$ we also write $M_1 \otimes M_2 \otimes \ldots \otimes M_n$ for $\bigotimes M_i$ and $m_1 \otimes m_2 \otimes \ldots \otimes m_n$ for $\otimes m_i$.

With this notation we see from the proof of 6.1.5 $\bigotimes M_i$ is as an R-module generated by the elements of the form $\otimes m_i$ But these elements are not linear independent. Indeed we have the following linear dependence relations:

$(ra) \otimes b = r(a \otimes b)$ and $(a + \tilde{a}) \otimes b = a \otimes b + \tilde{a} \otimes b$.

Here $r \in R, a \in M_i$, $b = \otimes_{j \in J} b_j$ with $b_j \in M_j$ and $i \in I$.

Lemma 6.1.6. *Let I be finite. Then $\bigotimes^I R = R$. More precisely, $f : R^I \to R, (r_i) \to \prod_{i\in I} r_i$ is a tensor product of $(R, i \in I)$.*

Proof. We need to verify that f meets the definition of the tensor product. Let $\tilde{f} : R^I \to \tilde{W}$ be R-multilinear. Define $g : R \to tildeW, r \to r\tilde{f}((1)))$, where (1) denotes the element $r \in R^I$ with $r_i = 1$ for all $i \in I$. Then clearly g is R-linear. Moreover,

$$\tilde{f}((r_i)) = \tilde{f}((r_i 1)) = (\prod_{i\in I} r_i)\tilde{f}((1)) = g(\prod_{i\in I} r_i) = g(f((r_i))$$

Thus $\tilde{f} = gf$.

Next let $\tilde{g} : R \to \tilde{W}$ be linear with $\tilde{f} = \tilde{g}f$. Then $\tilde{g}(r) = \tilde{g}(r1) = r\tilde{g}(1) = r\tilde{g}(\prod_{i\in I} 1) = rg(f((1)) = r\tilde{f}((1)) = g(r)$ and so g is unique. □

Lemma 6.1.7. *Let $(M_i, i \in I)$ be a family of R-modules. Suppose that I is the disjoint union of subsets $I_j, j \in J$. For $j \in J$ let $f_j : M_{I_j} \to W_j$ be R-multilinear. Also let $g : W_J \to W$ be R-multilinear. Then*

$$g \circ (f_j) : M_I \to W, m \to g((f_j(m_j))$$

is R-multilinear.

Proof. Let $f = g \circ (f_j)$. Let $m \in M_I$ and put $w_j = f_j(m_J)$. Let $w = (w_j) \in W_J$. Then $f(m) = g(w)$.

Let $i \in I$ and pick $j \in J$ with $i \in I_j$. Put $b = (m_k)_{k\in I-i}$ and $v = (w_k)_{k\in J}$. Then $w = (w_j, v)$, $m = (m_i, b)$ and $f_b(m_i) = f(m) = g(w_j, v) = g_v(w_j)$. Let $d = (m_l)_{k\in I_j - i)}$. Then $m_{I_J} = (m_i, d)$. Thus $w_j = f_j(m_{I_j}) = f_j(m_i, d) = (f_j)_d(m_i)$.

Hence $f_b(m_i) = g_v(w_j) = g_v((f_j)_d(m_i))$. So $f_b = g_v \circ (f_j)_d$. Since g is multilinear, g_v is R linear. Since f_j is a multilinear product, $(f_j)_d$ is R-linear. Since the composition of R-linear maps are R-linear, f_b is R-linear. So f is R-multilinear. □

Lemma 6.1.8. *Let $M_i, i \in I$ be a family of R-modules, $f : M_I \to W$ an R-multilinear map, $I = J \uplus K$ and $b \in M_K$.*

(a) There exists a unique R-linear map $\breve{f}_b : \otimes^J \to W$ with $\breve{f}_b(\otimes^J m_j) = f_b((m_j))$.

(b) The function $f_K : M_K \to \mathrm{Hom}_R(\bigotimes^J M_j, W), b \to \breve{f}_b$ is R-multilinear.

(c) There exists a unique R-linear map $\breve{f}_K : \bigotimes^K M_k \to \mathrm{Hom}_R(\bigotimes^J M_j, W)$ with $\breve{f}_K(\otimes^K m_k)(\otimes^J m_j) = f((m_i))$.

(d) There exists a unique R-bilinear map, $f_{K,J} : \bigotimes^K M_k \times \bigotimes^J M_j \to W$ with $f_{K,J}(\otimes^K m_k, \otimes^J m_j) = f((m_i))$

Proof. (a) Follows from 6.1.2 and the definition of a tensor product.

(b) Let $k \in K$, $a, \tilde{a} \in M_k$, $r \in R$, $b \in M_{K-a}$ and $d \in M_J$. The $(a,b) \in M_K$ and $(a,b,d) \in M_I$. We compute

$$(rf_{(a,b)})(\otimes^J d_j) = rf(a,b,d) = f(ra,b,d) = f_{(ra,b)}(\otimes^J d_j).$$

By the uniqueness assertion in (b), $rf_{(a,b)} = f_{(ra,b)}$. Thus $f_K(ra,b) = rf_K(a,b)$
Similarly

$$(f_{(a,b)} + f_{(\tilde{a},b)})(\otimes^J d_j) = f(a,b,d) + f(\tilde{a},b,d) = f(a+\tilde{a},b,d) = f_{(a+\tilde{a},b)})(\otimes^J d_j)$$

and $f_{(a,b)} + f_{(\tilde{a},b)} = f_{(a+\tilde{a},b)}$. Hence $f_K(a\tilde{a},b) = f_K(a+\tilde{a},b)$ and f_K is R-multilinear.

(c Follows from (b) and the definition of a tensor product.

(d) Define $f_{K,J}(a,b) = \breve{f}_K(a)(b)$. Since $\breve{f}_K$ and $\breve{f}_K(a)$ are R-linear and $f_{K,J}$ is bilinear. Thus (d) follows from (c). □

Lemma 6.1.9. *Let R be a ring and A, B and C R-modules. Then there exists an R-isomorphism*

$$A \otimes B \otimes C \to A \otimes (B \otimes C)$$

which sends $a \otimes b \otimes c \to a \otimes (b \otimes c)$ for all $a \in A, b \in B, c \in C$.

Proof. Define $f : A \times B \times C \to A \otimes (B \otimes C)$, $(a,b,c) \to a \otimes (b \otimes c)$. By 6.1.7, f is multilinear. So there exists an R-linear map $\breve{f} : A \otimes B \otimes C \to A \otimes (B \otimes C)$ with $g(a \otimes b \otimes c) = a \otimes (b \otimes c)$.

By 6.1.8 there exists an R-linear map $g = \otimes_{\{1\},\{2,3\}}$: $A \otimes (B \otimes C) \to A \otimes B \otimes C$ with $g(a \otimes (b \otimes c)) = a \otimes c$.

Note that $(g\breve{f})(a \otimes b \otimes c) = g(a \otimes (b \otimes c)) = a \otimes b \otimes c$. Since $A \otimes B \otimes C$ is generated by the $a \otimes b \otimes c$, we get $g\breve{f} = \mathrm{id}$. Similarly $\breve{f}g = \mathrm{id}$ and so $\breve{f}$ is an R-isomorphism. □

Lemma 6.1.10. *Let I be a finite set and for $i \in I$ let $(M_{ij}, j \in J_i)$ be a family of R-modules. Then there exists an R-isomorphism,*

$$\bigotimes_{i\in I}(\bigoplus_{j\in I_j} M_{ij}) \to \bigoplus_{\alpha\in J_I}(\bigotimes_{i\in I} M_{i\alpha_i}).$$

with

$$\otimes_{i\in I}\,(m_{ij})_{j\in J_i} \to (\otimes_{i\in I}\, m_{i\alpha_i})_{\alpha\in J_I}$$

Proof. Let $M_i = \bigoplus_{j\in J_i} M_{ij}$ and let $\pi_{ij} : M_i \to M_{ij}$ the projection map of M_i onto M_{ij}. Note here if $m_i \in M_i$, then $m_i = (m_{ij})_{j\in J_i}$ with $m_{ij} \in M_{ij}$ and $\pi_{ij}(m_i) = m_{ij}$. Let $\alpha \in J_I = \prod_{i\in I} J_i$. Define

$$f_\alpha : M_I \to \bigotimes_{i\in I} M_{i\alpha(i)}, \quad (m_i) \to \otimes_{i\in I}\, m_{i\alpha_i}.$$

Since $\otimes$ is multilinear and π_{ij} is linear, 6.1.7 implies that f_α is multilinear. Hence there exists a unique R-linear map

$$\breve{f}_\alpha : \bigotimes_{i\in I} M_i \to \bigotimes_{i\in I} M_{i\alpha(i)}$$

with $\breve{f}_\alpha(\otimes m_i) = \otimes m_{i\alpha_i}$. We claim that for a given $m = (m_i)$ there exists only finitely many $\alpha \in I_J$ with $f_\alpha(m) \neq 0$. Indeed there exists a finite subset $K_i \subseteq J_i$ with $m_{ij} = 0$ for all $j \in J_i \setminus K_i$. Thus $\alpha(m) = 0$ for all $\alpha \in J_I \setminus K_I$. Since I and K_i are finite, K_I is finite. Thus

$$\breve{f} = (\breve{f}_\alpha)_{\alpha\in J_I} : \bigotimes_{i\in I}(\bigoplus_{j\in I_j} M_{ij}) \to \bigoplus_{\alpha\in J_I}(\bigotimes_{i\in I} M_{i\alpha_i}).$$

is R-linear with

$$(*) \quad \breve{f}(\otimes_{i\in I}(m_{ij})_{j\in J_i}) = (\otimes_{i\in I} m_{i\alpha_i})_{\alpha\in J_I}$$

To show that $\breve{f}$ is an isomorphism, we define its inverse. For $j \in J_i$ let $\rho_{ij} : M_{ij} \to M_i$ be the canonical embedding. So for $a \in M_{ij}$, $\rho_{ij}(a) = (a_k)_{k\in I_j}$, where $a_k = 0$ of $k \neq j$ and $a_j = a$. Let $\alpha \in J_I$ and define

$$\rho_\alpha : \prod_{i\in I} M_{i\alpha_i} \to \bigotimes_{i\in I} M_i, \quad (m_{i\alpha_i}) \to \otimes_{i\in I} \rho_{i\alpha_i}(m_{i\alpha_i}).$$

Then ρ_α is R-multilinear and we obtain an R linear map

$$\breve{\rho}_\alpha : \bigotimes_{i\in I} M_{i\alpha_i} \to \bigotimes_{i\in I} M_i$$

with

$$\breve{\rho}_\alpha(\otimes_{i\in I} m_{i\alpha_i}) = \otimes_{i\in I} \rho_{i\alpha_i}(m_{i\alpha_i}).$$

Define

$$\breve{\rho} : \bigoplus_{\alpha\in J_I}(\bigotimes_{i\in I} M_{i\alpha_i}) \to \bigotimes_{i\in I} M_i, \quad (d_\alpha) \to \sum_{\alpha\in I_J} \rho_\alpha(d_\alpha).$$

Then $\breve{\rho}$ is R linear. We claim that $\breve{\rho} \circ \breve{f} = \mathrm{id}$ and $\breve{f} \circ \breve{\rho} = \mathrm{id}$.
Let $m = (m_i) = ((m_{ij})) \in M_i$. Then $m_i = \sum_{j\in J_i} \rho_{ij}(m_{ij})$ and by multilinearity of $\otimes$.

$$\otimes_{i\in I} m_i = \sum_{\alpha\in J_I} \otimes_{i\in I} \rho_{i\alpha_i}(m_{i\alpha_i})$$

By (*) and the definition of $\breve{\rho}$.

$$\breve{\rho}(\breve{f}(\otimes_{i\in I} m_i)) = \sum_{\alpha\in I_J} \breve{\rho}\alpha(\otimes_{i\in I} m_{i\alpha_i}) = \sum_{\alpha\in I_J} \otimes_{i\in I} \rho_{i\alpha_i}(m_{ialpha_i}) = \otimes_{i\in I} m_i.$$

Hence $\breve{\rho}\breve{f} = \mathrm{id}$.

Let $d = (d_\alpha) \in \bigoplus_{\alpha \in J_I} (\bigotimes_{i \in I} M_{i\alpha_i})$. To show that $(\breve{f}\breve{\rho})(d) = d$ we may assume that $d_\alpha = 0$ for all $\alpha \neq \beta$ and that $d_\beta = \otimes_{i \in I} m_{i\beta_i}$ with $m_{i\beta_i} \in M_{i\beta_i}$. Put $m_{ij} = 0$ for all $j \neq \beta_i$. Then $m_i := (m_i j) = \rho_{i\beta i}(m_{i\beta_i}$

Then

$$\breve{\rho}(d) = \sum_{\alpha \in J_I} \breve{\rho}_\alpha(d_\alpha) = \breve{\rho}_\beta(\otimes_{i \in I} m_{i\beta_i}) = \otimes_{i \in I} \rho_{i\beta_i}(m_{i\beta_i}) = \otimes_{i \in I} m_i$$

Let $\alpha \in J_I$ with $\alpha \neq 0$. Then $\alpha_i \neq \beta_i$ for some $i \in I$ and so $m_{i\alpha_i} = 0$. Hence

$\breve{f}_\alpha(\breve{\rho}(d)) = 0 = d_\alpha$ if $\alpha \neq \beta$ and $\breve{f}_\alpha(\breve{\rho}(d) = \otimes_{i \in I} m_{i\beta_i} = d_\beta$ if $\beta = \alpha$.

Thus $\breve{f}(\breve{\rho}(d)) = (\breve{f}_\alpha(\breve{\rho}(d)) = (d_\alpha) = d$. Hence $\breve{f}\breve{\rho} = \mathrm{id}$ and $\breve{f}$ is an isomorphism with inverse ρ. □

Corollary 6.1.11. *Let $(M_i, i \in I)$ be a finite family of R-modules. Suppose that M_i is a free R-module with basis $\mathcal{A}_i, i \in I$. Then $\bigotimes_{i \in I} M_i$ is a free R-module with basis*

$$(\otimes_{i \in I} a_i \mid a \in \mathcal{A}_I\}$$

.

Proof. For $j \in \mathcal{A}_i$ let $M_{ij} = R_j$. Then $M_i = \bigoplus_{j \in \mathcal{A}_i} M_{ij}$. For $a \in \mathcal{A}_i$, put $T_a = \bigotimes_{i \in I} M_{ia_i}$. Since each $M_{ij} \cong R$, 6.1.6 implies $T_a \cong R$. More precisely, $\otimes_{i \in I} a_i$ is a basis for T_a. By 6.1.10 $\bigotimes_{i \in I} M_i \cong \bigoplus_{a \in \mathcal{A}_I} T_a$. Hence $(\otimes_{i \in I} a_i \mid a \in \mathcal{A}_I\}$ is indeed a basis for $\bigotimes_{i \in I} M_i$. □

We will denote the basis from the previous theorem by $\bigotimes_{i \in I} \mathcal{A}_i$. If $I = \{1, \ldots, n\}$ and $\mathcal{A}_i = \{a_{i1}, a_{i2}, \ldots, a_{im_i}\}$ is finite we see that $\bigotimes_{i \in I} M_i$ has the basis

$$a_{1j_1} \otimes a_{2j_2} \otimes \ldots \otimes a_{nj_n}, \quad 1 \leq j_1 \leq m_1, \ldots, 1 \leq j_n \leq m_n.$$

Lemma 6.1.12. *(a) Let $(\alpha_i : A_i \to B_i, i \in I)$ a family of R-linear maps. Then there exists a unique R-linear map.*

$$\otimes \alpha_i : \bigotimes A_i \to \bigotimes B_i$$

with

$$(\otimes \alpha_i)(\otimes a_i) = \otimes \alpha_i(a_i)$$

(b) Let $(\alpha_i : A_i \to B_i, i \in I)$ and $(\beta_i : B_i \to C_i, i \in I)$ families of R-linear maps. Then $\otimes(\beta_i \circ \alpha_i) = (\otimes \beta_i) \circ \otimes(\alpha_i)$.

Proof. (a) Define $f : A_I \to \bigotimes B_i, \quad (a_i) \to \otimes \alpha_i(a_i)$. By 6.1.7 f is R-multilinear. So (b) follows from the definition of the tensor product.

(b) Both these maps send $\otimes a_i$ to $\otimes(\beta_i(\alpha_i(a_i))$. □

6.2 Symmetric and Exterior Powers

Let I be a finite set, R a ring and M an R-module. Let $M_i = M$ for all $i \in M$. Then $M_I = M^I$. Let $\pi \in \mathrm{Sym}(I)$ and $m = (m_i) \in M^I$. Define $m\pi \in M$ by $(m\pi)_i = m_{\pi(i)}$. (So if we view m as a function from $I \to M$, $mpi = m \circ \pi$) For example if $\pi = (1,2,3)$, then $(m_1, m_2, m_3)\pi = (m_2, m_3, m_1)$. Note that for $\pi, \mu \in \mathrm{Sym}(I)$, $m(\pi\mu) = (m\pi)\mu$.

Definition 6.2.1. *Let I be a finite set, R a ring and M an R-modules. Let $f : M^I \to W$ be R-multilinear.*

(a) f is symmetric *if $f(m\pi) = f(m)$ for all $m \in M, \pi \in \mathrm{Sym}(I)$.*

(b) f is skew symmetric *if $f(m\pi) = (\mathrm{sgn}\pi)f(m)$ for all $m \in M, \pi \in \mathrm{Sym}(I)$.*

(c) f is alternating *if $f(m) = 0$ for all $m \in M^I$ with $m_i = m_j$ for some $i \neq j \in I$.*

Lemma 6.2.2. *(a) Let $f : M^I \to W$ be alternating. Then f is skew symmetric.*

(b) Suppose that $f : M^I \to W$ is skew symmetric and that $w \neq -w$ for all $0 \neq w \in W$. Then f is alternating.

(c) Let $f : M^n \to W$ be multilinear with $f(m) = 0$ for all $m \in M^n$ with $m_i = m_{i+1}$ for some $1 \leq i < n$. Then f is alternating.

Proof. (a) Let $\pi \in \mathrm{Sym}(I)$ and $m \in M$ we need to show that $f(\pi m) = \mathrm{sgn} f(\pi m)$. Since π is the product of two cycles we may assume that π itself is a 2-cycle. So $\pi = (i,j)$ for some $i \neq j \in I$. Let $a = m_i, b = m_j$, $d = m_{I\setminus\{i,j\}}$ and $g = f_d$. Then $m = (a,b,d)$, $f(m) = g(a,b)$ and $(\pi f)(m) = f(b,a,d) = g(b,a)$.

Since f and so also g is alternating we compute

$$0 = g(a+b, a+b) = g(a,a) + g(a,b) + g(b,a) + g(b,b) = g(a,b) + g(b,a)$$

Thus $f(\pi m) = g(b,a) = -g(a,b) = (\mathrm{sgn}\pi)f(m)$

(b) Suppose that $m_i = m_j$ for some $i \neq j$ and let $\pi = (i,j)$. Then $m = \pi m$ and so $f(m) = f(\pi m) = (\mathrm{sgn}\pi)f(m) = -f(m)$ Thus by assumption on W, $f(m) = 0$ and f is alternating.

(c) By induction on n. Let $m \in M$ with $m_i = m_j$ for some $1 \leq i < j \leq n$. Let $m = (a,b)$ with $a \in M^{n-1}$, $b \in M$. Let $g = f_b$, that is $g(d) = f(d,b)$ for $d \in M^{n-1}$. By induction g is alternating. So if $j \neq n$, $f(m) = g(a) = 0$. So suppose $j = n$. Let $\pi = (i, n-1)$. By induction and (b), $f(m\pi) = g(a\pi) = -g(a) = -f(m)$. But $(m\pi)_{n-1} = m_i = m_j = m_n = (m\pi)_n$ and so by assumption $f(m\pi) = 0$. Hence also $f(m) = 0$. □

Definition 6.2.3. *Let R be a ring, M an R-module, I a finite set and $f : M^I \to W$ an R-multilinear function.*

(a) f is called an Ith symmetric power *of M over R provided that f is symmetric and for every symmetric function $g : M^I \to \tilde{W}$, there exists a unique R-linear map $\breve{g} : W \to \tilde{W}$ with $g = \breve{g} \circ f$.*

(b) f is called an Ith exterior power *of M over R provided that f is alternating and for every alternating function $g : M^I \to \tilde{W}$, there exists a unique R-linear map $\breve{g} : W \to \tilde{W}$ with $g = \breve{g} \circ f$.*

Lemma 6.2.4. *Let R be a ring, M an R-module and I a finite set. Then an I-th symmetric and an I-th exterior power of M over R exist. Moreover they are unique up to R-isomorphism.*

Proof. Let A be the R-submodule of $\bigotimes^I M$ generated by the elements $\otimes m - \otimes m\pi$, $m \in M_I, \pi \in \mathrm{Sym}(I)$. Let $W = (\bigotimes^I M)/A$ and define $f : M_I \to W$ by $f(m) = \otimes m + A$. We claim that f is an I-th symmetric power for M over R. So let $g : M_I \to \tilde{W}$ be symmetric. Then g is multilinear and so by the definition of a tensor product there exists a unique R-linear map $\grave{g} : \bigotimes^I M \to \tilde{W}$ with $\grave{g}(\otimes m) = g(m)$. Since $g(m) = g(m\pi)$ for all $m \in M, \pi \in \mathrm{Sym}(I)$ we have $\grave{g}(\otimes m) = \grave{g}(\otimes m\pi)$. Thus $\otimes m - \otimes m\pi \in \ker \grave{g}$. Hence also $A \leq \ker \grave{g}$. So there exists a uniquely determined and well defined R-linear map $\breve{g} : W \to \tilde{W}, d + A \to \grave{g}(d)$ for all $d + A \in W$. So f is an I-symmetric power of M over R.

Next let B be the R-submodule of $\bigotimes^I M$ generated by the elements $\otimes m$ where $m \in M$ with $m_i = m_j$ for some $i \neq j \in I$. Let $W = \bigotimes_I M/B$ and define $f : M_I \to W$ by $f(m) = \otimes m + B$. As above it is now a routine exercise to verify that f is an R-exterior power of M over R.

Finally the uniqueness of the symmetric and alternating powers are verified in the usual way. □

We will denote the I-th symmetric power of M over R by $M^I \to S^I M, (m_i) \to \prod_{i \in I} m_i$. The exterior power is denoted by $M^I \to \bigwedge^I M, (m_i) \to \wedge_{i \in I} m_i$.

Lemma 6.2.5. *(a) $S^n R \cong R$ for all $n \geq 1$*

(b) $\bigwedge^1 R \cong R$ and $\bigwedge^n R = 0$ for all $n \geq 2$.

Proof. (a) By 6.1.6 $R^n \to R, (r_i) \to \prod r_i$ is the n-th tensor power of R. Since the map is symmetric and is also the n-th symmetric power.

(b) An alternating map in one variable is just a linear map. So $\bigwedge^R = R$. Now suppose $n \geq 2$, $a, b \in R$, $c \in R^{n-2}$ and $f : R^n \to W$ is alternating. Then $f(a, b, c) = abf(1, 1, c) = 0$. Hence $\Lambda^n R = 0$. □

Lemma 6.2.6. *Let $(M_i, i \in I)$ be an R modules, I a finite set and suppose that I is the disjoint unions of the subsets $I_k \in K$ and M_k is an R-module with $M_i = M_k$ for all $i \in I_k$. Let $g : M_I \to W$ be multilinear. Then*

(a) Suppose that for all $k \in K$ and $b \in I \setminus I_k$, $g_b : M_k^{I_k} \to W$ is alternating. Then there exists a unique R-linear map

$$\breve{g} : \bigotimes_{k \in K} (\bigwedge^{I_k} M_k) \to W$$

with

$$\breve{g}(\otimes_{j\in J}(\wedge_{i\in I_k} m_i) = g((m_i))$$

for all $(m_i) \in M^I$.

(b) Suppose that for all $k \in K$ *and* $b \in I \setminus I_k$, $g_b : M_k^{I_k} \to W$ *is symmetric Then there exists a unique* R*-linear map*

$$\breve{g} : \bigotimes_{k\in K}(S^{I_k}M) \to W$$

with

$$\breve{g}(\otimes_{j\in J}(\wedge_{i\in I_k} m_i) = g((m_i))$$

for all $(m_i) \in M^I$.

Proof. This is easily proved using the methods in 6.1.2 and 6.1.9 □

Lemma 6.2.7. *Let* R *be a ring,* I *a finite set and* $(M_j, j \in J)$ *a family of* R*-modules. Let* $\Delta = \{d \in \mathbb{N}^J \mid \sum_{j\in J} d_j = |I|\}$. *For* $j \in J$ *let* $\{I_d^j \mid j \in J\}$ *be a partition of* I *with* $|I_d^j| = d_j$ *for all* $j \in J$. *For* $d \in \Delta$ *put* $A(d) = \{\alpha \in J^n \mid |\alpha^{-1}(j)| = d_j\}$. *For* $\alpha \in A(d)$ *and* $j \in J$ *put* $I_\alpha^j = \alpha^{-1}(j) = \{i \in I \mid \alpha_i = j\}$. *Let* $\pi_\alpha \in \mathrm{Sym}(I)$ *with* $\pi_\alpha(I_d^j) = I_\alpha^j$. *Then*

(a) The function

$$f : (\bigoplus_{j\in J} M_j)^I \to \bigoplus_{d\in\Delta}(\bigotimes_{j\in J} S^{I_d^j}M_j)$$

$$((m_{ij})_{j\in J})_{i\in I}) \to (\sum_{\alpha\in A(d)} \bigotimes_{j\in J}(\prod_{i\in I_d^j} m_{\pi_\alpha(i)j}))_{d\in\Delta}$$

is an I*-th symmetric power of* $\bigoplus_{j\in J} M_j$ *over* R.

(b) The function

$$f : (\bigoplus_{j\in J} M_j)^I \to \bigoplus_{d\in\Delta}(\bigotimes_{j\in J} \bigwedge^{I_d^j} M_j)$$

$$((m_{ij})_{j\in J})_{i\in I}) \to (\sum_{\alpha\in A(d)} \mathrm{sgn}_{\pi_\alpha} \bigotimes_{j\in J}(\bigwedge_{i\in I_d^j} m_{\pi_\alpha(i)j}))_{d\in\Delta}$$

is an I*-th exterior power of* $\bigoplus_{j\in J} M_j$ *over* R.

Proof. (b) View each $\alpha = (\alpha_i)_{i\in I} \in J^n$ as the function $I \to J, i \to \alpha_i$. Since $\{I_d^j \mid j \in J$ of I is a partition of I, each I_d^j is a subset of I and each $i \in I$ is contained I_d^j for a unique $j \in J$. Define $\alpha_d \in J^I$ by $(\alpha_d)_i = j$ where $i \in I_d^j$.

Let $\alpha \in J^I$. Note that $\{I_\alpha^j \mid j \in J\}$ is a partition of I. Define $d = d_\alpha \in \Delta$ by $(d_\alpha)_j = |I_\alpha^j|$. So d is unique in Δ with $\alpha \in A(d)$. Note that $I_{\alpha_d}^j = I_d^j$. We will now verify that there exists a $\pi_\alpha \in \mathrm{Sym}(I)$ with $\pi_\alpha(I_d^j) = I_\alpha^j$. Since $|I_\alpha^j| = |I_d^j|$, there exists a bijection $\pi_\alpha^j : I_d^j \to I_\alpha^j$.

Define $\pi_\alpha \in \mathrm{Sym}(I)$ by $\pi_\alpha(i) = \pi^j_\alpha(i)$, where $i \in I^j_d$. Since $\pi^j_\alpha(i) \in I_{\alpha^j}$, $\alpha(\pi^j_\alpha(i)) = j$. But $j = \alpha_d(i)$ and so $\alpha \circ \pi_\alpha = \alpha_d$.

Conversely if $\pi \in \mathrm{Sym}(I)$ with $\alpha \circ \pi = \alpha_d$ then $\pi^j : I^j_d \to I^j_\alpha, i \to \pi(i)$ is a well defined bijection.

Define

$$f^j_d : M^n \to \bigwedge^{d_j} M_j, \quad m \to \wedge_{i \in I^j_d} m_{ij}$$

and

$$f_d : M^n \to \bigotimes_{j \in J} (\bigwedge^{d_j} M_j), \quad m \to \otimes_{j \in J} f^j_d(m))$$

We will now show $\mathrm{sgn}\pi_\alpha f_d \circ \pi_a lpha$ does not depend on the particular choice of π_α. For this let $\pi \in \mathrm{Sym}(n)$ with $\alpha_d = \alpha \circ \pi$. Put $\sigma = \pi - 1\pi_\alpha$ and $\sigma^j = (\pi^j) - 1\pi^j_\alpha$. So Then $\sigma^j \in \mathrm{Sym}(I^j_d)$ and and

$$(f^j_d \circ \pi_\alpha(m) = f^j_d(m\pi_\alpha) = \ \wedge_{i \in I^j_d}(m\pi_\alpha)_{ij} = \wedge_{i \in I^j_d}(m_{\pi^j_\alpha(i)j} =$$

$$= \wedge_{i \in I^j_d} m_{\pi^j(\sigma^j(i))j} = (\mathrm{sgn}\sigma^j)(\wedge_{i \in I^j_d} m_{\pi^j(i)j}) = (\mathrm{sgn}\sigma^j)(f^j_d \circ \pi)(m)$$

Thus $f^j_d \circ \pi_\alpha = (\mathrm{sgn}\sigma^j) f^j_d \circ \pi$ Taking the tensor product over all $j \in J$ and using $\mathrm{sgn}\sigma = \prod_{j \in J} \mathrm{sgn}\sigma^j$ we get $f_d \circ \pi_\alpha = \mathrm{sgn}\sigma f_d \circ \pi$. But $\mathrm{sgn}\pi = \mathrm{sgn}\pi_\alpha \mathrm{sgn} sigma$ and so

$$\mathrm{sgn}_{\pi_\alpha} f_d \circ \pi_\alpha \mathrm{sgn}\pi f_d \circ \pi$$

So we can define $f_\alpha = \mathrm{sgn}_\pi f_d \circ \pi$, where $\pi \in \mathrm{Sym}(n)$ with $\alpha_d = \alpha\pi$.

Let $\mu \in \mathrm{Sym}(n)$ and $j \in J$. Then $(\alpha\mu)(i) = j$ if and only if $\alpha(\mu(i)) = j$. Thus $\mu(I^j_{\alpha\mu}) = I^j_\alpha$. Hence $d_{\alpha\mu} = d_\alpha = d$. Put $\rho = \pi_{\alpha\mu}$ Then

$$\alpha_d = (\alpha\mu) \circ \rho = \alpha \circ (\mu \circ \rho$$

So by definition of f_α

$f_\alpha(m) = (\mathrm{sgn}(\mu \circ \rho))(f_d \circ (\mu \circ \rho) = (\mathrm{sgn}\mu)(\mathrm{sgn}\rho)(f_d \circ \rho)(m\mu) = \mathrm{sgn}\mu f_{\alpha\mu}(m\mu)$. So we proved:

(**) $f_{\alpha\mu}(m\mu) = (\mathrm{sgn}\mu) f_\alpha(m)$

For $d \in \Delta$ define $\bar{f}_d = \sum_{\alpha \in A(d)} f_\alpha$ We will show that $\bar{f}_d$ is alternating. By 6.1.7, $f^a_d lpha$ is multilinear. Hence also $\bar{f}_d$ is multilinear.

Now suppose that $m_k = m_l$ for some $k \neq l \in I$. Put $\mu = (k, l) \in \mathrm{Sym}(I)$.

Let $\alpha \in A(d)$. Suppose that $\alpha = \alpha\mu$, that is $\alpha_k = \alpha_l$. Let $j = \alpha(i)$. Then $k, l \in I^j_\alpha)$. Since $m_l = m_k$, $m_{lj} = m_{ij}$ Thus $\wedge_{i \in I^j_\alpha} m_{ij} = 0$, $f^j_\alpha(m) = 0$ and so also $f_\alpha(m) = 0$.

Suppose next that $\alpha \neq \alpha\mu$. Since $m_k = m_l$, $m = m\mu$. So by (**)

$$f_{\alpha\mu}(m) = f_{\alpha\mu}(m\mu) = \mathrm{sgn}\mu f_\alpha(m) = -f_\alpha(m)$$

Hence $f_{\alpha\mu}(m) + f_\alpha(m) = 0$. It follows that $\bar{f}_d(m) = \sum_{\alpha\in A(d)} f_\alpha(m) = 0$ and $\bar{f}_d$ is alternating.

Now define

$$f = (\bar{f}_d) : M^n \to \bigoplus_{d\in\Delta}(\bigotimes_{j\in J}\bigwedge^{d_j} M_j), \quad m \to (\bar{f}_d(m))_{d\in\Delta}.$$

To complete the proof of (b) it remains to verify that f is an I-th exterior power of M. Since each f_d is alternating, f is alternating. Let $g : M^n \to W$ be alternating.

By 6.2.6 there exists a unique R-linear map

$$\breve{g}_d : \bigotimes_{j\in J}(\bigwedge^{I_d^j} M_j) \to W$$

with

$$\breve{g}_d(\otimes j \in J \wedge_{i\in I_d^j} m_i = g(m)$$

where $m \in M^I$ with $m_i \in M_j$ for all $i \in I_d^j$.

Define

$$\breve{g} : \bigoplus_{d\in\Delta}(\bigotimes_{j\in J}\bigwedge^{d_j} M_j) \to W, \quad (u_d)_{d\in\Delta} \to \sum_{d\in\Delta}\breve{g}_d(u_d)$$

Let $m \in M^I$. Since g is multilinear,

$$g(m) = \sum_{\alpha\in J^I} w_\alpha$$

where $w_\alpha = g(m_{i\alpha(i)})$.

Let $\pi = \pi_\alpha$. Since g is alternating and $\alpha_d = \alpha\pi$,

$$w_\alpha = \mathrm{sgn}\pi g(m_{\pi i,\alpha_d(i)})$$

Note that $\otimes j \in J \wedge_{i\in I_d^j} m_i \wedge_{i\in I_d^j} m_{\pi i\alpha_d(i)} = f_d(m\pi))$ and so by definition of $\breve{g}_d$ and the previous equation

$$w_\alpha = \mathrm{sgn}\pi\breve{g}_d(f_d(m\pi)) = \breve{g}_d(f_\alpha(m))$$

Thus

$$g(m) = \sum_{\alpha\in J^I} w_\alpha) = \sum_{d\in\Delta}\sum_{\alpha\in A(d)} \breve{g}_d(f_\alpha(m)) ==$$

$$= \sum_{d\in\Delta} \breve{g}_d(\sum_{\alpha\in A(d)} f_\alpha(m)) = \sum_{d\in\Delta} \breve{g}_d\bar{f}_d(m)) = \breve{g}((\bar{f}_d(m)_{d\in\Delta}) = \breve{g}(f(m))$$

Thus $g = \breve{g}\circ f$. So f is indeed an exterior power and (b) is proved.

(a) To prove (a) we change the proof for (b) as follows: Replace $\bigwedge$ by S. Replace $\wedge$ by $\cdot$. Replace every sgn*lambda* by 1. Finally the following argument needs to be added:

Let $\mu \in \mathrm{Sym}(I)$. Then using (**) and $A(d) = \{\alpha\mu \mid \alpha \in A(d)$ we get

$$\bar{f}_d(m) = \sum_{\alpha\in A_d} f_\alpha(m) = \sum_{\alpha\in A(d)} f_{\alpha\mu}(m\mu) = \sum_{\alpha\in A_d} f_\alpha(m\mu) = \bar{f}_d(m\mu).$$

Thus $\bar{f}_d$ is symmetric. □

A remark on the preceding theorem. The proof contains an explicit isomorphism. But this isomorphism depends on on the choice of the partitions I_d^k. And the computation of the isomorphism depends on the choice of the π_α. Here is a systematic way to make these choices. Assume $I = \{1,\dots,n\}$ and choose some total ordering on J. Let $d \in \Delta$ and let $J_d = \{j \in J \mid d_j \neq 0$. Note that $|J_d| \leq |I|$ and so J_d is finite. Hence $J_d = \{j_1,\dots j_u$ with $j_1 < j_2 < \dots < j_u$. To simplify notation we write k for j_k. Choose $I_d^1 = \{1,2,\dots,d_1$, $I_d^2 = \{d_1+1, d_1+2,\dots,d_1+d_2\}$ and so on. Now let $\alpha \in A(d)$. So $I_d^j = \{s+1, s+2,\dots s+d_j$, where $s = \sum_{k<j} d_k$Define π_α as follows. Send 1 to the smallest i with $\alpha(i) = 1$, 2 to the second smallest element with $\alpha(i) = 1$, d_1 to the largest element with $\alpha(i) = 2$, $d_1 + 1$ to the smallest element with $\alpha(i) = 2$ and so on.

Finally we identify $\bigwedge^{I_d^j} M_j$ with $\bigwedge^{d_j} M_j$ by identifying $\wedge_{i\in I_{d^j}} v_i \in \bigwedge^{I_d^j} M_j$ with $\wedge_{t=1}^{d_j} v_{s+t} \in \bigwedge^{d_j} M_j$, where $s = \sum_{k<j} d_k$.

Let $m = (m_i) \in M^I$ such that for all $i \in I$ there exists a unique $j \in J$ with $m_{ij} \neq 0$. So $m_i = m_{ij}$ for a unique $j \in J$. Denote this j by $\alpha(i)$. Then $\alpha \in J^I$. Note that $\bar{f}_d(m) = 0$ for all $d \neq d_\alpha$. So suppose that $\alpha \in A(d)$. Let $I_\alpha^j = \{i_1^j, i_2^j,\dots i_{d_j}^j\}$ with $i_1^j < i_2^j < \dots < i_{d_j}$. Then since $\wedge$ is skew symmetric there exists $\epsilon \in \{1,-1\}$ with

$$\wedge m = m_{1,\alpha(1)} \wedge m_{2,\alpha(2)} \wedge \dots \wedge m_{n,\alpha(n)} =$$

$$= \epsilon m_{i_1^1,1} \wedge m_{i_2^1,1} \wedge \dots \wedge m_{i_{d_1},1} \wedge m_{i_2^2,2} \dots \wedge m_{i_{d_2},2} \wedge \dots \wedge m_{i_1^u,u} \wedge \dots \wedge m_{i_{d_u}^u,u}$$

Then $\epsilon = \mathrm{sgn}\pi_\alpha$ and $\bar{f}_d(m)$ is

$$\epsilon(m_{i_1^1,1} \wedge m_{i_2^1,1} \wedge \dots \wedge m_{i_{d_1},1}) \otimes (m_{i_2^2,2} \dots \wedge m_{i_{d_2},2}) \otimes \dots \otimes (m_{i_1^u,u} \wedge \dots \wedge m_{i_{d_u}^u,u})$$

For example suppose that $|I| = 3$ and $|J| = 2$. We want to compute $f(m_{11}+m_{12}, m_{21}+m_{22}, m_{31}+m_{32})$. Since f is multilinear we need to compute $f(m_{1\alpha(1)}, m_{2\alpha(2)}, m_{3\alpha(3)}$ where $\alpha(i) \in J = \{1,2\}$.

If $\alpha = (1,1,1)$ then $d_\alpha = (3,0)$ and

$$\bar{f}_{(3,0)}(m_{11}, m_{21}, m_{31}) = m_{11} \wedge m_{21} \wedge m_{31}$$

If $\alpha = (1,1,2)$ then $d_\alpha = (2,1)$ and

$$\bar{f}_{(2,1)}(m_{11}, m_{21}, m_{32}) = (m_{11} \wedge m_{21}) \otimes m_{32}$$

If $\alpha = (1,2,1)$ then $d_\alpha = (2,1)$ and

$$\bar{f}_{(2,1)}(m_{11}, m_{22}, m_{31}) = -(m_{11} \wedge m_{31}) \otimes m_{22}$$

If $\alpha = (1,2,2)$ then $d_\alpha = (1,2)$ and

$$\bar{f}_{(1,2)}(m_{11}, m_{22}, m_{32}) = m_{11} \otimes (m_{22} \wedge m_{32})$$

If $\alpha = (2,1,1)$ then $d_\alpha = (2,1)$ and

$$\bar{f}_{(2,1)}(m_{11}, m_{21}, m_{32}) = (m_{21} \wedge m_{31}) \otimes m_{12}$$

If $\alpha = (2,1,2)$ then $d_\alpha = (1,2)$ and

$$\bar{f}_{(1,2)}(m_{12}, m_{21}, m_{32}) = -m_{21} \wedge (m_{12} \otimes m_{32})$$

If $\alpha = (2,2,1)$ then $d_\alpha = (1,2)$ and

$$\bar{f}_{(1,2)}(m_{12}, m_{22}, m_{31}) = m_{31} \otimes (m_{12} \wedge m_{22})$$

If $\alpha = (2,2,2)$ then $d_\alpha = (0,3)$ and

$$\bar{f}_{(0,3)}(m_{12}, m_{22}, m_{32}) = m_{12} \wedge m_{22} \wedge m_{32}.$$

Thus the four coordinates of $f(m)$ are:

$d = (3,0)$:

$$m_{11} \wedge m_{21} \wedge m_{31}$$

$d = (2,1)$:

$$(m_{11} \wedge m_{21}) \otimes m_{32} - (m_{11} \wedge m_{31}) \otimes m_{22} + (m_{21} \wedge m_{31}) \otimes m_{12}$$

$d = (1,2)$:

$$m_{11} \otimes (m_{22} \wedge m_{32}) - m_{21} \wedge (m_{12} \otimes m_{32}) + m_{31} \otimes (m_{12} \wedge m_{22})$$

$d = (0,3)$:

$$m_{12} \wedge m_{22} \wedge m_{32}$$

Lemma 6.2.8. *Let R be a ring, n a positive integer and M a free R-modules with basis $\mathcal{B}$. Let " $\leq$ " be a total ordering on $\mathcal{B}$.*

(a) $(b_1 b_2 \dots b_n \mid b_1 \leq b_2 \leq \dots b_n \in \mathcal{B})$ *is a basis for* $S^n M$.

(b) $(b_1 \wedge b_2 \wedge \dots \wedge b_n \mid b_1 < b_2 < \dots b_n \in \mathcal{B})$ *is a basis for* $S^n M$).

Proof. For $b \in \mathcal{B}$ put $M_b = Rb$. Then $M_b \cong R$ and $M = \bigoplus_{b\in\mathcal{B}} M_b$. We will apply 6.2.7 with $I = \{1, \dots, n\}$ and $J = \mathcal{B}$. Let Δ be as in the statement of that theorem. Let $d \in \Delta$.

(a) By 6.2.5, $S^t M_b \cong R$ with basis b^t. By 6.1.6

$$\bigotimes_{b\in\mathcal{B}} (S^{d_b} M_b) \cong R$$

and has $\otimes_{b\in\mathcal{B}} b^{d_b}$ has a basis. (a) now follows from 6.2.7(a)

(b) By 6.2.5 $\bigwedge^t M_b = 0$ for all $t \geq 2$. So

$$\bigotimes_{b\in\mathcal{B}} (\bigwedge^{d_b} M_b) \cong R = 0$$

if $d_b \geq 2$ for some $b \in \mathcal{B}$ and

$$\bigotimes_{b\in\mathcal{B}} (\bigwedge^{d_b} M_b) \cong R$$

if $d_b \leq 1$ for all $b \in \mathcal{B}$. Moreover, it has basis $\otimes_{b\in\mathcal{B}, d_b=1} b$. (b) now follows from 6.2.7(b). □

Example: Suppose M has basis $\{a, b, c, d\}$. Then $S^3 M$ has basis

$$d^3, cd^2, c^2 d, c^3, bd^2, bcd, bc^2, b^2 d, b^2 c, b^3, ad^2, acd, ac^2, abd, abc, ab^2, a^2 d, a^2 c, a^2 b, a^3$$

and $\bigwedge^3 M$ has basis

$$b \wedge c \wedge d, a \wedge c \wedge d, a \wedge b \wedge d, a \wedge b \wedge c$$

Corollary 6.2.9. *Let R be a ring and n, m positive integer. Then*

(a) $S^m R^n \cong R^{\binom{n+m+1}{m}}$

(b) $\bigwedge^m R^n \cong R^{\binom{n}{m}}$.

Proof. This follows from 6.2.8 □

Lemma 6.2.10. *Let R be a ring and M an free R-module with finite basis $\mathcal{A}$ and $\mathcal{B}$. Then $|\mathcal{A}| = |\mathcal{B}|$.*

Proof. Let $n = |\mathcal{A}|$. Then $M \cong R^n$. So by 6.2.9(b), n is the smallest non-negative integer with $\bigwedge^{n+1} M = 0$. So n is uniquely determined by M and $n = |\mathcal{B}|$. □

Definition 6.2.11. *Let R be a ring and M and free R-module with a finite basis $\mathcal{B}|$. Then $|\mathcal{B}|$ is called the rank of M.*

6.3 Determinants and the Cayley-Hamilton Theorem

Lemma 6.3.1. *Let I be finite set and R a ring.*

(a) Let $\alpha : A \to B$ be R-linear. Then there exists a unique R-linear map

$$\wedge^I \alpha : \bigwedge^I A \to \bigwedge^I B$$

with

$$\wedge^I \alpha(\wedge a_i) = \wedge \alpha(a_i).$$

(b) Let $\alpha : A \to B$ and $\beta : B \to C$ be R-linear. Then

$$\wedge^I(\beta \circ \alpha) = \wedge^I \beta \circ \wedge^I \alpha.$$

Proof. (a) Define $g : A^I \to \bigwedge^I B, (a_i) \to \wedge \alpha(a_i)$. If $a_i = a_j$ for some $i \neq j$ then also $\alpha(a_i) = \alpha(b_i)$ and so $g(a) = 0$. Thus g is alternating and (a) follows from the definition of an exterior power.

(b) Both these maps send $\wedge a_i$ to $\wedge\beta(\alpha(a_i))$. □

Theorem 6.3.2. *Let R be a ring and n a positive integer.*

(a) Let R be a ring, $0 \neq M$ a free R-module of finite rank n, and $\alpha \in \mathrm{End}_R(V)$. Then there exists a unique $r \in R$ with $\wedge^n \alpha = r\mathrm{id}_{\bigwedge^n M}$. We denote this r by $\det \alpha$.

(b)

$$\det : \mathrm{End}_R(V) \to R, \alpha \to \det \alpha$$

is a multiplicative homomorphism.

(c) There exists a unique function $\det : \mathcal{M}_R(n) \to R$ *(called* determinant*) with the following two properties:*

(a) When viewed as a function in the n columns, det *is alternating.*

(b) Let I_n be the $n \times n$ idendity matrix. Then $\det I_n = 1$.

(d) Let $A = (a_{ij}) \in \mathcal{M}_R(n)$. Then

$$\det A = \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}\pi \prod_{i=1}^{n} a_{i\pi i}$$

(e) Let $A = (a_{ij}) \in \mathcal{M}_R(n)$ and $a_j = (a_{ij})$ the j-th column of A. Then $\wedge a_j = \det A \wedge e_j$, where $e_j = (\delta_{ij}) \in R^n$.

(f) Let R be a ring, $0 \neq M$ a free R-module of finite rank n, $\alpha \in \mathrm{End}_R(V)$. and $\mathcal{B}$ a basis for M. Let $A = \mathcal{M}^{\mathcal{B}}(\alpha)$ be the matrix for α with respect to $\mathcal{B}$. Then

$$\det \alpha = \det A$$

(g) Let $A \in \mathcal{M}_R(n)$. Then

$$\det A = \det A^T$$

where $a_{ij}^T = a_{ji}$.

Proof. (a) By 6.2.9, $\bigwedge^I M \cong R$. Thus by 4.5.8, $\mathrm{End}_R(\bigwedge^I M) = R_{\mathrm{id}}$. So (a) holds.

(b) follows from 6.3.1.

(c) Let $e_i = (\delta_{ij}) \in R^n$. Then by 6.2.8, $e := \wedge_{i=1}^n e_i$ is a basis for $\bigwedge^n R^n$. Define $\tau : \bigwedge^n R^n \to R, re \to r$. Let $A \in \mathcal{M}_R(n)$ a view A as $(a_i)_{1 \leq i \leq n}$ with $a_i \in R^n$. Define $\det A = \tau(\wedge_{i \in I} a_i$. Since $I_n = (e_i)$, $\det I_n = 1$. So det fulfills **(Det Alt)** and **Det I**. Suppose now $f : (R^n)^n \to R$ is alternating with $f((e_i)) = 1$. Then by definition of an I-th exterior power there exists an R-linear map $\breve{f} : \bigwedge^n R^n \to R$ with $f = \breve{f} \circ \wedge$. Then $\breve{f}(e) = \breve{e}(\wedge e_i) = f((e_i) = 1$ and so $\breve{f} = \tau$ and $f = \det$. Thus (c) holds.

(d) We will apply 6.2.7 with $I = J = \{1, \ldots, n\}$ and $M_j = Re_j$. So $\bigoplus_{j \in J} = R^n$. Let $\delta \in \Delta$. If $d_j \geq 2$ for some $j \in J$ then $\bigwedge^{I_d^j} M_j = 0$. If $d_j \leq 1$ for all j, then $\sum_{j \in J} d_j = n = |I|$ forces $d_j = 1$ for all $j \in J$. Let $d \in \Delta$ with $d_j = 1$ for all $j \in J$. Also $R_{e_j} \to R, re_j \to R$ is an 1-st exterior power. Let $\alpha \in J^I$. Then $\alpha \in A(d)$ if and only if $|\alpha^{-1}(j)| = 1$ for all $j \in J$. This is the case if and only of $\alpha \in \mathrm{Sym}(n)$. Also $\pi_\alpha = \alpha$. Hence 6.2.7 implies that

$$f : (R^n)^n \to R \quad (m_{ij}) \to \sum_{\alpha \in \mathrm{Sym}(n)} \prod_{i=1}^n m_{i\pi i}$$

is an n-th exterior power of R^n. Note that $f((e_i)) = 1$. So this this choice of $\bigwedge^n R^n$ we have $e = 1$, $\tau = \mathrm{id}_R$ and $\det = f$. so (d) holds.

(e) was proved in (c).

(f) For $A \in \mathcal{M}_{\mathcal{B}}(R)$ let $\alpha = \alpha_A$ be the corresponding elements of $\mathrm{End}_R(M)$. So $\alpha(b) = \sum_{d \in \mathcal{B}} a_{db} d$. Let $a_b = (a_{db}$, the b-th column of A. Suppose that $a_b = a_c$ with $b \neq c$. Then $\alpha(b) = \alpha(c)$ and so $(\wedge\alpha)(\wedge b) = \wedge\alpha(b) = 0$. Hence $\det \alpha = 0$. Also $\det I_n = \det \mathrm{id} = 1$ and so $A \to \det(\alpha_A)$ fulfilled **(Det Alt)** and **(Det I)**. Thus the uniqueness of $\det A$ implies $\det A = \det \alpha$.

(g) Using (d) we compute

$$\det A^T = \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}\pi \prod_{i \in I} a_{i\pi(i)}^T = \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}\pi \prod_{i \in I} a_{\pi(i)i} =$$

$$= \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}\pi \prod_{i \in I} a_{i\pi^{-1}(i)} = \sum_{\pi \in \mathrm{Sym}(n)} \mathrm{sgn}\pi \prod_{i \in I} a_{i\pi(i)} = \det A$$

□

Definition 6.3.3. *Let R be a ring and $s : A \to B \to C$ R-bilinear.*

(a) An s-basis for is a triple $((a_d \mid d \in D), (b_d \mid d \in D), c)$ such that D is a set, $(a_d \mid d \in D)$ is a basis for A, $(b_d, d \in D)$ is a basis for B and $\{c\}$ is a basis for C with $s(a_d, b_e) = \delta_{de}c$ for all $d, e \in D$.

(b) We say that is s is a pairing *if there exists an s-basis. s is a finite pairing if s is pairing and* $\mathrm{rank}A = \mathrm{rank}B$ *is finite.*

Note that if $s : A \to B \to C$ is a pairing, then A, B and C are free R-modules and $C \cong R$ as an R-module. Also s is non-degenerate, that is $s(a, b) = 0$ for all $b \in B$ implies $a = 0$, and $s(a, b) = 0$ for all $a \in A$ implies $b = 0$.

The converse is only true in some special circumstances. For example if R is a field, $s : A \to B \to C$ is bilinear, $\dim_R C = 1$ and $\dim_R A$ is finite, then it is not to difficult to see that s is a pairing.

But if $\dim_R A$ is not finite this is no longer true in general. For example let $B = A^* = \mathrm{Hom}_R(A, R)$ and $s(a, b) = b(a)$. Then $\dim_R B > \dim_R A$ and so s is not a pairing.

For another example define $s : \mathbb{Z} \to \mathbb{Z} \to \mathbb{Z}(a, b) \to \mathbb{Z}, (a, b) \to 2ab$. The s is not a pairing. Indeed suppose $(\{a\}, \{b\}, c)$ is an s basis. Then $c = s(a, b) = 2ab$, a contradiction to $\mathbb{Z} = \mathbb{Z}c$.

Lemma 6.3.4. *Let R be a ring, I, J, K finite sets with $K = I \uplus J$ and let $s : A \times B \to R$ be R-bilinear. Let $\Delta = \{E \subseteq K \mid |E| = |J|\}$ and for $E \in \Delta$ choose $\pi_E \in \mathrm{Sym}(K)$ with $\pi_E(J) = E$.*

(a) There exists a unique R-bilinear map

$$s_K^J : \bigwedge^K A \times \bigwedge^J B \to \bigwedge^I A$$

with

$$s_K^J(\wedge a_k, \wedge b_j) = \sum_{E \in \Delta} \det(s(a_{\pi_E(j)}, b_{j'})_{j,j' \in J}) \bigwedge_{i \in I} a_{\pi_E(i)}$$

(b) s_K^J is independent form the choice of the π_E.

(c) Let $\alpha \in \mathrm{End}_R(A)$ and $\beta \in \mathrm{End}_R(B$ with $s(\alpha(a), b) = s(a, \beta(b))$ for all $a \in A, b \in B$. Then

$$(\wedge^I \alpha)\,(s_K^J(\,u\,,\,(\wedge^J \beta)(v)\,) = s_K^J(\,(\wedge^K \alpha)(u)\,,\,v\,)$$

for all $u \in \bigwedge^K A$ andv $\in \bigwedge^J B$.

(d) Suppose there exists a basis $\mathcal{E} = (e_d, d \in D)$ for A and a basis $\mathcal{F} = (f_d, d \in D)$ for B such that $s(e_d, f_{d'}) = \delta_{dd'}$. Let $\alpha \in D^K$ and $\beta \in D^J$ be one to one. Then

$$s_K^J((\bigwedge_{k \in K} e_{\alpha(k)}, \bigwedge_{j \in J} f_{\beta(k)}) = \begin{cases} \pm \bigwedge_{k \in K \setminus \alpha^{-1}(\beta(J))} e_{\alpha(k)} & \text{if } \beta(J) \subseteq \alpha(K) \\ = 0 & \text{if } \beta(J) \not\subseteq \alpha(K) \end{cases}.$$

Proof. (a) and (b) We first show that

$$f_E(a,b) := \mathrm{sgn}\pi_E \det(s(a_{\pi_E(j)}, b_{j'})_{j,j'\in J} \bigwedge_{i\in I} a_{\pi_E(i)}$$

is independent from the choice of π_E. Indeed let $\pi \in \mathrm{Sym}(K)$ with $\pi(J) = E$. Let $\sigma = \pi^{-1}\pi_E$. Let $\sigma_J \in \mathrm{Sym}(J)$ be defined by $\sigma_J(j) = \sigma(j)$. Similarly define σ_I. Then

$$\det(s(a_{\pi_E}(j), b_{j'})) = \det(s(a_{(\pi\sigma_J(j)}, b_{j'}) = \mathrm{sgn}\sigma_J \det(s(a_{\pi i}, b_{j'})$$

and

$$\bigwedge_{i\in I} a_{\pi_E(i)} = \bigwedge_{i\in I} a_{\pi\sigma_I(i)} = \mathrm{sgn}\sigma_I \bigwedge_{i\in I} a_{\pi(i)}.$$

Using that $\mathrm{sgn}\pi = \mathrm{sgn}\sigma\mathrm{sgn}\pi_E = \mathrm{sgn}\sigma_I\mathrm{sgn}\sigma_j\mathrm{sgn}\pi_E$ and multiplying the last two equations together we obtain the claimed independence from the choice of π_E.

Define

$$f : A^K \times B^J \to \bigwedge^J A, \quad (a,b) \to \sum_{E\in\Delta} f_E(a,b)$$

In view of 6.2.6 it remains to show that f_b and f_a are alternating for all $a \in A^K$ and $b \in B^J$. That f_a is alternating is obvious. So suppose $b \in B^J$ and $a \in A^K$ with $a_k = a_l$ for distinct $k, l \in K$. Let $E \in \Delta$ and put $\pi = \pi_E$. If k and l are both in $\pi(J)$ then $\det(s(a_{\pi j}, b_{j'})) = 0$. If k, l are both in I then $\bigwedge_{i\in I} a_{\pi(i)} = 0$. So in both these cases $f_E(a,b) = 0$. Suppose now that $k \in \pi(I)$ and $l \in \pi(J)$. Let $\sigma = (k,l) \in \mathrm{Sym}(K)$ and $E' = \sigma(E) \neq E$. We may choose $\pi_{E'} = \sigma\pi$. $a_k = a_l$ now implies $f_{E'}(a,b) = \mathrm{sgn}\sigma f_E(a,b)$ and so $f_{E'}(a,b) + f_E(a,b) = 0$. If follows that $f_b(a) = f(a,b) = 0$ and f_b is alternating.

(c) Let $a \in A^K$, $b^{\in}B^J$. Note that $\beta \circ b = (\beta(b_j))$. Let $E \in \Delta$. Then

$$(\bigwedge^I \alpha)(f_E(a, \beta \circ b)) = (\bigwedge^I \alpha)(\mathrm{sgn}\pi_E \det(s(a_{\pi_E(j)}, \beta(b_{j'})) \bigwedge_{i\in I} a_{\pi_E(i)}) =$$

$$= \mathrm{sgn}\pi_E \det(s(\alpha(a_{\pi_E(j)}, b_{j'}) \bigwedge_{i\in I} \alpha(a_{\pi_E(i)})) = f_E(\alpha \circ a, b)$$

Thus (c) holds.

(d) Suppose $E \in \Delta$ and $f_E(a,b) \neq 0$ where $a = (e_{\alpha(k)})$ and $b = (f_{\beta(j)})$. Let $A = s(e_\alpha(\pi_E(j)), f_{\beta(j')}$. Then $\det A \neq 0$. Let $t \in E$. Then $t = \pi_E(j)$ for some $\in J$ and so $(s(e_\alpha(t), t, \alpha f_{\beta(j')})_{j'\in J}$ is a row of A. This row cannot be zero and $s(e_\alpha(t), t, \alpha f_{\beta(t')} \neq 0$ for some $t' \in J$. But then $\alpha(t) = \beta(t')$. It follows that $\beta(J) \subseteq \alpha(I)$ and $E = \alpha^{-1}\beta(I)$. Also $\det A = \pm 1$ and so (ca) holds. □

Proposition 6.3.5. *Let R be a ring and M an R-module.*

(a) Let I, J and K finite sets with $K = I \uplus J$ Then there exists a unique bilinear map

$$\wedge : \bigwedge^I M \times \bigwedge^J M \to \bigwedge^K M, (a, b) \to a \wedge b$$

with

$$(\wedge_{i\in I} m_i) \wedge (\wedge_{j\in J} m_j) = \wedge_{k\in K} m_i$$

for all $(m_i) \in M^{k+l}$.

(b) Define

$$\bigwedge M = \bigoplus_{i=0}^{\infty} \bigwedge^i M$$

and

$$\wedge : \bigwedge M \times \bigwedge M \to \bigwedge M, \quad (a_i)_{i=o}^{\infty} \wedge (b_j)_{j=0}^{\infty} = (\sum_{i=0}^{k} a_i \wedge b_{k-i})_{k=0}^{\infty}.$$

Then $(\bigwedge M, +, \wedge)$ is a (non)-commutative ring with $R = \bigwedge^0 M \leq Z(\bigwedge M)$.

Proof. (a) Define $f : M^I \times M^J \to \bigwedge^K M, ((a_i), (a_j)) \to \wedge_{k\in K} a_k$. Clearly $f_{(a_i)}$ and $f_{(a_j)}$ is alternating and so (a) follows from 6.2.6.

(b) First of all $(\bigwedge M, +)$ is an abelian group. By (a) $\wedge$ is bilinear. So the distributive laws hold. Let l, m, n be non-negative integers and $m_k \in M$ for $1 \leq k \leq l + m + n$. Then

$$(\bigwedge_{i=1}^{l} m_i \wedge \bigwedge_{i=l+1}^{l+m} m_i) \wedge \bigwedge_{i=l+m+1}^{l+m+n} m_i = \bigwedge_{i=1}^{l+m+m} m_i = \bigwedge_{i=1}^{l} m_i \wedge (\bigwedge_{i=l+1}^{l+m} m_i \wedge \bigwedge_{i=l+m+1}^{l+m+n} m_i)$$

and so $\wedge$ is associative.

So $(\bigwedge M, +, \wedge)$ is indeed a (non)-commutative ring. That $R \leq Z(\bigwedge M)$ follows from the fact that $\wedge$ is R-linear. □

Lemma 6.3.6. *Let R be a ring and $s : A \times B \to C$ a finite pairing.*

(a) The functions

$$s_A : A \to \mathrm{Hom}_R(B, C), a \to s_a$$

and

$$s_B : B \to \mathrm{Hom}_R(A, C), b \to s_b$$

are R-linear isomorphism.

(b) Let $f \in \mathrm{End}_R(B)$. Then there exists a unique $f^s \in \mathrm{End}_R(A)$ with $s(f^s(a), b) = s(a, f(b))$ for all $a \in A$, $b \in B$.

(c) Suppose $(a_d, d \in D)$, $(b_d, d \in D)$ *and* (c) *are s-basis for* (A, B, C). *Let* $M_D(f^s) = M_D(f)^T$

Proof. Let $((a_d \mid, d \in D), (b_d \mid d \in D), c)$ be an s basis. (a) For $e \in D$ define $\phi_e \in \mathrm{Hom}_R(B, C)$ by $\phi_e(\sum_{r_d b_d} = r_e c$. Then $(\phi_d, d \in D)$ is a basis for $\mathrm{Hom}_R(B, C)$. Since $s(a_e, b_d) = \delta_{ed} c$. $s_A(e) = \phi_e$. Hence (a) holds.

(b) Define $\tilde{f} \in \mathrm{End}_R(\mathrm{Hom}_R(B, C)$ by $\tilde{f}(\phi) = \phi \circ f$. Let $g \in \mathrm{End}_R(A), a \in A$ and $b \in B$. Then

$$s(a, f(b)) = s_A(a)(f(b)) = ((\tilde{f})(s_A))(a)(b)$$

and

$$s(g(a), b) = s_A(g(a))(b)$$

Hence $s(a, f(b) = s(g(a), b)$ for all $a \in A$, $b \in B$ if and only if $\tilde{f} \circ s_A = s_A \circ g$. By (a), s_A has an inverse so $f^s = s_A^{-1} \tilde{f} s_A$ is the unique element fulfilling (c).

(c) Let $g \in \mathrm{End}_R(B)$. Put $U = M_f(D)$ and $V = M_g(D)$. So $g(a_d) = \sum_{h \in D} v_{hd} a_h$ and $f(b_d) = \sum_{h \in D} u_{hd} b_h$. Thus

$$s(a_e, f(b_d)) = \sum_{h \in D} u_{hd} s(a_e, b_h) = u_{ed} c$$

and

$$s(g(a_e), b_d) = sum_{h \in D} v_{he} s(a_h, b_d) = v_{de} c$$

Hence $s(a, f(b)) = s(g(a), f)$ for all $a \in A$, $b \in B$ if and only if $v_{de} = u_{ed}$ for all $d, e \in D$. So (c) holds (and we have a second proof for (b)). □

Recall that for an R-module M, M^* denote the dual module, so $M^* = \mathrm{Hom}_R(M, R)$.

Lemma 6.3.7. *Let* R *be a ring,* M *a free module of finite rank over* R *and* I *a finite set*

(a) There exists a unique R*-bilinear function* $s^I : \bigwedge^I M^* \times \bigwedge^I M \to R$ *with* $s^I(\bigwedge \phi_i, \bigwedge m_i) = \det(\phi_i(m_j))_{i,j \in I}$.

(b) s_I *is a finite pairing.*

(c) $\bigwedge^I M^* \cong (\bigwedge^I M)^*$ *as* R*-modules.*

Proof. Define $s : M^* \times M \to R, (\phi, m) \to \phi(m)$. (a) follows from 6.3.4(a) applied with $A = M^*, B = M, K = I, J = I$ and "$I = \emptyset$". And (b) follows from part (d) of the same lemma. Finally (c) is a consequence of (b) and 6.3.6(a). □

Proposition 6.3.8. *Let* R *be a ring and* M *a* R*-module of finite rank. Let* $f \in \mathrm{End}_R(M)$. *Then there exists* $f^{\mathrm{ad}} \in \mathrm{End}_R(M)$ *with* $f \circ f^{\mathrm{ad}} = f^{\mathrm{ad}} \circ f = \det f \mathrm{id}_M$.

Proof. Consider $t : M \times \bigwedge M^{n-1} \to \bigwedge M^n, (m, b) \to m \wedge b$. We claim that t is a finite pairing. For this let $(a_i, 1 \leq i \leq n)$ be a basis for M. Put $b_i = a_1 \wedge a_2 \wedge a_{i-1} \wedge a_{i+1} \wedge a_n$. Let $c = a_1 \wedge \ldots a_n$. By 6.2.8, $(b_i, 1 \leq i \leq n)$ is a basis for $\bigwedge^{n-1} M$ and $\{c\}$ is a basis for $\bigwedge^n M$. Also $a_i \wedge b_j = 0$ for $i \neq j$ and $a_i \wedge b_i = (-1)^{i_1} c$. $((a_i), ((-1)^{i-1} b_i), c)$ is a t basis. Let $f^a d = (bigwedge^{n-1} f)^t$ be given by 6.3.6(b). So $f^{\mathrm{ad}} \in \mathrm{End}_R(M)$ is uniquely determined by

$$f^a d(m) \wedge b = m \wedge (\bigwedge^{n-1} f)(b)$$

for all $m \in M, b \in \bigwedge^{n-1} M$.

In particular,

$$(f^{\mathrm{ad}}(f(m) \wedge b) = f(m) \wedge (\bigwedge^{n-1} f)(b) = (\bigwedge^{n} f)(m \wedge b) = (\det f)(m \wedge b) = m \wedge (\det f)b$$

Note that also $(\det f)m \wedge b = m \wedge (\det f)b$ and so by 6.3.6

$$f^{\mathrm{ad}} \circ f = ((\det f)\mathrm{id}_{\bigwedge^{n-1} M})^t = (\det f)\mathrm{id}_M$$

To show that also $f \circ f^{\mathrm{ad}} = \mathrm{id}_M$ we use the dual M^* of M. Recall that $f^* \in \mathrm{End}_R(M^*)$ is define by $f^*(\phi) = \phi \circ f$. It might be interesting to note that $f^* = f^s$, where s is the pairing $s : M^* \to M, (\phi, m) \to \phi(m)$.

Applying the above results to f in place of f^* we have

$$f^{*\mathrm{ad}} \circ f^* = (\det f^*)\mathrm{id}_{M^*}$$

By 6.3.2(g) we have $\det f* = \det f$. So dualizing the previous statement we get

$$f \circ (f^{*\mathrm{ad}*} = \det f \mathrm{id}_M$$

So the proposition will be proved once we show that $f^{*\mathrm{ad}*} = f^{\mathrm{ad}}$ or $f^{*\mathrm{ad}} = f^{\mathrm{ad}^*}$.

To do this we will compute that matrix of f^{ad} with respect to the basis (a_i). Let D be the matrix of f with respect to (a_i) and E the matrix of $\bigwedge^{n-1} f$ with respect to $((-1)^{i-1} b_i)$. We compute

$$(\bigwedge^{n-1} f)(b_i) = \wedge_{h \neq i} f(a_h) = \wedge_{h \neq i} (\sum_{k=1}^{n} d_{hk} a_k$$

Let D_{ij} be the matrix $(d_{kl})_{k \neq i, l \neq j}$. Then the coefficient of b_j in $\wedge_{h \neq i} (\sum_{k=1}^n d_{hk} a_k$ is readily seen to be $\det D_{ij}$.

It follows that

$$E_{ij} = (-1)^{i-1} - 1^{j-1} \det D_{ij} = (-1)^{i+j} \det D_{ij}$$

Let (ϕ_i) be the basis of M^* dual to (a_i). So $\phi_i(a_j) = \delta_{ij}$. Then the matrix for f^* with respect to (ϕ_i) is D^T. Note that $(D^T)_{ij} = (D_{ji})^T$ and so the (i, j) coefficient of the matrix of $f^{*\mathrm{ad}}$ is

$$(-1)^{i+j}\det(D^T)_{ij} = (-1)^{i+j}\det(D_{ji})^T = (-1)^{i+j}\det D_{ji}$$

Thus $f^{*\mathrm{ad}}$ has the matrix E^T with respect to (ϕ_i). So does $(f^{\mathrm{ad}*}$. Hence $f^{*\mathrm{ad}} = f^{\mathrm{ad}*}$ and the proposition is proved. □

Lemma 6.3.9. *Let R and S be rings with $R \leq S$. Let M be an R module. Then there exists bilinear function*

$$\cdot : S \times S \otimes_R M \to S \otimes M, (s, \tilde{m}) \to s\tilde{m}$$

with

$$s(t \otimes m = st \otimes m$$

for all $s, t \in S$ and $m \in M$. Moreover, $(S \otimes_R M, c \cdot 0$ is an S-module.

Proof. Let $s \in S$. By 6.1.12 there exists a unique $s\mathrm{id}_S \otimes id_M \in \mathrm{End}_R(S \otimes_R M)$ which sends $t \otimes m$ to $st \otimes m$. We will write $s \otimes 1$ for $s\mathrm{id}_S \otimes \mathrm{id}_M$. It is readily verified that $s \to s \otimes 1$ is a ring homomorphism. So the lemma is proved. □

Lemma 6.3.10. *Let R and S be rings with $R \leq S$.Let M be a free R-module with basis $\mathcal{B}$.*

(a) $S \otimes_R M$ is a free S-module with basis $1 \otimes \mathcal{A} := \{1 \otimes b \mid b \in \mathcal{B}$.

(b) Let $\alpha \in \mathrm{End}_R(M)$, A the matrix of α with respect to $\mathcal{B}$ and $s \in S$. Then sA is the matrix of $s \otimes \alpha$ with respect to $1 \otimes \mathcal{B}$.

Proof. (a) Note that $M = \bigoplus_{b\in\mathcal{B}} Rb$ and $Rb \cong R$. By 6.1.10 $S \otimes_R M \cong \bigoplus_{b\in\mathcal{B}} S \otimes_R Rb$. Also by 6.1.6 $S \otimes_R b \cong S$.

(b) Let $d \in \mathcal{B}$ Then

$$(s \otimes \alpha)(1 \otimes d) = s \otimes \alpha(d) = s \otimes (\sum_{e\in\mathcal{B}} b_{ed}e) = \sum_{e\in\mathcal{B}} (sb_{ed}(1 \otimes e)$$

So (b) holds. □

Definition 6.3.11. *Let R be a ring, M a free R-module of finite rank and $\alpha \in \mathrm{End}_R(M)$.*

(a) Let S be a ring with R as a subring. Let $s \in S$. Then $s \otimes \alpha$ denotes the unique R-endomorphism of $S \otimes_R M$ with

$$(s \otimes 1)(t \otimes m) = (st \otimes \alpha(m)$$

for all $t \in S, m \in M$.

(b) Consider $x \otimes 1 - 1 \otimes \alpha \in \mathrm{End}_{R[x]}(R[x] \otimes_R M)$. Then

$$\chi_\alpha = \det(x \otimes 1 - 1 \otimes \alpha) \in R[x]$$

is called the characteristic polynomial *of α.*

(c) Let n be positive integer and $A \in \mathcal{M}_R(n)$. Consider the matrix $xI_n - A \in \mathcal{M}_{R[x]}(n)$. Then $\chi_A = \det(xI_n - A)$ is called the characteristic polynomial *of A.*

Lemma 6.3.12. *Let R be a ring, M an R-module with finite basis I, $n = |I|$, $\alpha \in \mathrm{End}_R(M)$ and A the matrix of α with respect to A.*

(a) $\chi_\alpha = \chi_A$.

(b) For $J \subset I$ let $A_J = (a_{ij})i, j \in J$. The coefficient of x^m in χ_A is

$$(-1)^{n-m} \sum_{J \subset I, |J| = n-m} \det A_J$$

(c) χ_α is monic of degree n.

Proof. (a) By 6.3.10(b) the matrix for $x \otimes 1 - 1 \otimes \alpha$ with respect to $xI_n - A$. Thus (a) follows from 6.3.2(f)

(b) Let $D = xI_n - A$. Let a_i be the i column of A_i. Let $e_i = (\delta_{ij})$. The $D = (xe_i - a_i)$. For $J \subset I$ let A^*_J be the matrix with whose k-column is a_k if $k \in J$ and e_k if $k \notin J$. Then since det is multilinear

$$\det D = \sum_{J \subseteq I} x^{|I|-|J|}(-1)^{|J|} \det A^* J$$

.

Let $T(J)$ be the matrix with

$$t(J)_{ij} = \begin{cases} a_{ij} & \text{if } i, j \in J \\ 1 & \text{if } i = j \notin J \\ 0 & \text{otherwise} \end{cases}$$

Then it is easy to see that $\det A^*(J) = \det T(J) = \det A(J)$ and (b) follows.

(c) Follows from (b) □

Theorem 6.3.13. *Let R be a ring, M be a free R-module of finite rank. Let $\alpha \in \mathrm{End}_R(M)$. Then*

$$\chi_\alpha(\alpha) = 0.$$

Proof. Define

$$\phi : R[x] \times M \to M, \quad (f, m) \to f(\alpha)(m).$$

Since ϕ is bilinear there exists a unique R-linear map

$$\Phi : R[x] \otimes_R M \to M \text{ with } \Phi(f \otimes m) = f(\alpha)(m).$$

Let $\beta = x \otimes 1 - 1 \otimes \alpha \in \mathrm{End}_{R[x]}(R[x] \otimes_R M)$.

Let $f \in R[x]$ and $m \in M$. Then

$$\beta(f \otimes m) = xf \otimes m - f \otimes \alpha(m) = fx \otimes m - f \otimes \alpha(m)$$

and so

$$\Phi(\beta(f \otimes m)) = (f(\alpha)\alpha)(m) - (f(\alpha)(\alpha(m)) = 0$$

Hence $\Phi\beta = 0$.
By 6.3.8 there exists $\beta^{\mathrm{ad}} \in \mathrm{End}_{R[x]}(R[x] \otimes_R M)$ with $\beta \circ \beta^{\mathrm{ad}} = \det\beta \otimes 1$.
It follows that

$$0 = (\Phi \circ \beta) \circ \beta^{\mathrm{ad}} = \phi \circ (\beta \circ \beta^{\mathrm{ad}}) = \Phi \circ (\det\beta \otimes 1)$$

So

$$0 = \phi((\det\beta \otimes 1))(1 \otimes m) = \phi(\det\beta \otimes m) = (\det\beta)(\alpha)(m)$$

By definition $\chi_\alpha = \det\beta$ and so the Cayley Hamilton Theorem is proved. □

Theorem 6.3.14. *Let M be a finitely generated R-module and $\alpha \in \mathrm{End}_R(M)$. Then there exists a monic polynomial $f \in R[x]$ with $f(A) = 0$.*

Proof. Let I be a finite subset of M with $M = RI$. Let $F = F_R(I)$ be the free R-module on I. So F has a basis $(a_i, i \in I)$. Let π be the unique R-linear map from F to M with $a_i \to i$ for all $i \in I$. Since $M = RI$, $M = \pi(F)$. By 4.5.2 there exists $\beta \in \mathrm{End}_R(F)$ with

$$\pi \circ \beta = \alpha \circ \pi$$

We claim that (*) $\pi \circ f(\beta) = f(\alpha) \circ \pi$ for all $f \in R[x]$

For this let $S = \{f \in R[x] \mid \pi \circ f(\beta) = f(\alpha) \circ \pi\}$. Let $f, g \in S$. Then

$$\pi \circ (fg)(\alpha) = \pi \circ (f(\alpha \circ g(\alpha)) = (\pi \circ f(\alpha) \circ g(\alpha) = (f(\alpha) \circ \pi) \circ g(\alpha) =$$
$$= f(\alpha) \circ (\pi \circ g(\alpha)) = f(\alpha) \circ (g(\alpha) \circ \pi) = (f(\alpha) \circ g(\alpha)) \circ \pi) = (fg)(\alpha) \circ \pi$$

Since π is $\mathbb{Z}$-linear, also $f - g \in S$. Thus S is a subring of $R[x]$. Since R and x are in S, $S = R[x]$ and (*) is proved. Let $f = \chi_\beta$. The f is monic and by 6.3.13 $f(\beta) = 0$. By (*)

$$f(\alpha) \circ \pi = \pi \circ f(\beta) = 0$$

Since π is onto this implies $f(\alpha) = 0$. □

Chapter 7

Hilbert's Nullstellensatz

Throughout this chapter ring means commutative ring with identity and R is a ring. All R-modules are assumed to be unitary.

7.1 Multilinear Maps

Definition 7.1.1. *Let $(V_i, i \in I)$ a family of R-modules, an R module and $f : \bigtimes_{i\in I} M_i \to M$ a function. Let $I = J \cup K$ with $J \cap K = \emptyset$,*

(a) $V_J := \bigtimes_{j\in J} V_j$.

(b) If $u = (u_j)_{j\in J} V_J$ and $v = (v_k)_{k\in K} \in V_K$, then we identify $(u, v) \in V_J \times V_K$ with the tuple $w = (w_i)_{i\in I}$ where $w_i = u_i$ if $i \in J$ and $w_i = v_i$ if $i \in K$. We also write $f(u, v)$ for $f((u, v))$.

(c) For $u \in V_J$ define $f_u : V_K \to W, v \to f(u, v)$. Sometimes we will write f_u^J for f_u.

(d) f is called R-multilinear if for all $i \in I$ and all $u \in V_{I\setminus i},$, $f_u : V_i \to W$ is R-linear.

(e) An R-multilinear map is called bilinear of $|I| = 2$ and trilinear if $|I| = 3$.

(f) $W^I = \bigtimes_{i\in I} W$.

Example 7.1.2. (a) If $|I| = 1$ a R-multilinear map is R-linear map.

(b) $R^n \times R^n \to R, \left((r_i)_{i=1}^n, (s_i)_{i=1}^n\right) \to \sum_{i=1}^n r_i s_i$ is R-bilinear.

(c) Let V be an R-module. Note that $\mathrm{End}_R(V)$ is an R module via $(r\phi)(v) = r\phi(m)$. Then $\mathrm{End}_R(V) \times V \to V, (\phi, v) \to \phi(v)$ is R-bilinear.

Definition 7.1.3. *Let V and W be R-modules and I a set. Put $V_i = V$ for all $i \in I$ and so $V_J = V^J$ for all $J \subseteq I$. An R-multilinear function $f : V^I \to V$ is called*

(a) symmetric *if $f_u(v, w) = f_u(v, v)$,*

(b) skew-symmetric *if* $f_u(v,w) = -f_u(w,v)$*;*

(c) alternating *if* $f_u(v,v) = 0$.

for all $i \neq j \in I$, $u \in V^{I\setminus\{i,j\}}$ *and* $v, w \in V$.

Lemma 7.1.4. *(a) Every alternating map is skew-symmetric.*

(b) If $f : M^I \to N$ *is skew symmetric and* $2n \neq 0_N$ *for all* $n \in N^\sharp$*, then* f *is alternating.*

Proof. Let $i \neq j \in I$, $u \in M^{I\setminus\{i,j\}}$, $v, w \in M$ and put $g = f_u$.

(a) $0 = g(v+w, v+w) = g(v,v) + g(v,w) + g(w,v) + g(w,w) = g(v,w) + g(w,v)$.

(b) From $g(v,w) = -g(w,v)$ applied with $v = w$ we get $g(v,v) = -g(v,v)$. So $2g(u,u) = 0$ and $g(u,u) = 0$ □

Lemma 7.1.5. *Let* I *be a set,* $(I_j)_{j\in J}$ *a partition of* I*,* $(V_i)_{i\in I}$ *and* $(W_j, j \in J)$ *families of* R*-modules and* Z *an* R*-module. Let* $f : W_J \to Z$ *be* R*-multilinear and for each* $j \in J$ *let* $g_j : V_{I_j} \to W_j$ *be* R*-multilinear. Then*

$$V_I \to Z, (v_i)_{i\in I} \to f\left(\left(g_j\big((v_i)_{i\in I_j}\big)\right)_{j\in J}\right)$$

is R*-multilinear.*

Proof. Readily verified. □

Lemma 7.1.6. *Let* I *be finite set,* $(V_i, i \in I)$ *be family of* R*-modules and* W *an* R*-module. Suppose that for each* $i \in I$*,* V_i *is a free* R*-module with basis* $\mathcal{B}_i \subseteq V_i$*. Put* $\mathcal{B}_i = \bigtimes_{i\in I} \mathcal{B}_i \subseteq V_I$ *and let* $g : \mathcal{B}_I \to W$ *be a function. Then there exists a unique* R*-multilinear map* $f : V_I \to W$ *with* $f\,|_{\mathcal{B}_I} = g$*.*

Proof. Suppose first that $f : V_I \to W$ is R-multilinear with $f\,|_{\mathcal{B}_I} = g$. Let $v = (v_i)_{i\in I} \in V_I$. Then since $\mathcal{B}_i$ is a basis for V_i there exists uniquely determined $r_{ib_i} \in R, b_i \in \mathcal{B}_i$ with

$$v_i = \sum_{b_i\in\mathcal{B}_i} r_{ib_i} b_i$$

Since f is R-multilinear we conclude that

$$(*) \qquad \begin{aligned} f(v) &= f\left(\left(\textstyle\sum_{b_i\in\mathcal{B}_i} r_{ib_i} b_i\right)_{i\in I}\right) \\ &= \textstyle\sum_{(b_i)_{i\in I}\in\mathcal{B}_i} \left(\prod_{i\in I} r_{ib_i}\right) f\big((b_i)_{i\in I}\big) \\ &= \textstyle\sum_{(b_i)_{i\in I}\in\mathcal{B}_i} \left(\prod_{i\in I} r_{ib_i}\right) g\big((b_i)_{i\in I}\big) \end{aligned}$$

So f is uniquely determined. Conversely it is readily verified that (*) defines an R-multilinear map. If $v \in \mathcal{B}_I$, then $r_{ib_i} = \delta_{v_i b_i}$ and so $\prod_{i\in I} r_{ib_i} = \prod_{i\in I} \delta_{v_i b_i}$, which is 0 unless $v_i = b_i$ for all $i \in I$, in which case it is 1. So $f(v) = g(v)$ for $v \in \mathcal{B}_I$. □

Proposition 7.1.7. *Let $n \in \mathbb{Z}^+$, $\mathrm{M}_n(R)$ the ring of $n \times n$-matrices with coefficients in R and $A \in M_n(R)$. We define the* determinant $\det(A)$ *of A inductively on as follows: If $n = 1$ and $A = (a)$, define $\det(A) = a$. Suppose next $n > 1$ and that $\det(B)$ has been defined for all $(n-1) \times (n-1)$-matrices. For $1 \leq i, j \leq n$ let A_{ij} be the $n \times n$ matrix defined obtained from A by deleting row i and column j. Define*

$$\det_j(A) := \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}$$

Then $\det_j(A) = \det_l(A)$ for all $1 \leq j, l \leq n$ and we define $\det(A) = \det_j(A)$. View det *function in the n-columns :*

$$\det : (R^n)^n \to R, ((a_{ij})_{i=1}^n)_{j=1}^n \to \det\left((a_{ij})\right)$$

Then det *is alternating and R-linear. Also $\det(I_n) = 1$,*

Proof. We will first show that $\det_j(A) = \det_l(A)$. Without loss $j < k$. We have

$$\begin{aligned} \det_j(A) &= \textstyle\sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) \\ &= \textstyle\sum_{i=1}^n \sum_{k=1, k\neq i} (-1)^{i+j} (-1)^{\epsilon} a_{ij} a_{kl} \det((A_{ij})_{kl}) \end{aligned}$$

where $(A_{ij})_{kl})$ is the matrix obtained by deleting row i and k and columns j and l from A and ϵ is as follows:

Observe that column l of A is column $l-1$ of A_{ij}. If $k < i$, then row k of A is row k of A_{ij}. If $k > i$, then row k of A is row $k-1$ of A_{ij}. Hence

$$\epsilon = \begin{cases} k + l - 1 & \text{if } k < i \\ k + l - 2 & \text{if } k > i \end{cases}$$

Similarly

$$\begin{aligned} \det_l(A) &= \textstyle\sum_{k=1}^n (-1)^{k+l} a_{kl} \det(A_{kl}) \\ &= \textstyle\sum_{k=1}^n \sum_{i=1, i\neq k} (-1)^{k+l} (-1)^{\eta} a_{kl} a_{ij} \det(A_{kl})_{ij}) \end{aligned}$$

where $(A_{kl})_{il})$ is the matrix obtained by deleting row k and i and columns l and j from A and η is as follows:

Observe that column j of A is column j of A_{kl}. If $k < i$, then row i of A is row $i-1$ of A_{kl}. If $k > i$, then row i of A is row i of A_{kl}. Hence

$$\eta = \begin{cases} i + j - 1 & \text{if } k < i \\ i + j & \text{if } k > i \end{cases}$$

If $i < k$ we conclude

$$(-1)^{i+j}(-1)^{k+l+\epsilon} = (-1)^{i+j+k+l-1} = (-1)^{k+l}(-1)^{\eta}$$

and if $k > i$ then

$$(-1)^{i+j}(-1)^{k+l+\epsilon} = (-1)^{i+j+k+l-2} = (-1)^{i+j+k+l} = (-1)^{k+l}(-1)^{\eta}$$

Thus show that $\det_j(A) = \det_l(A)$ and so we can define $\det(A) = \det_j(A)$ for any $1 \leq j \leq n$. Clearly $\det_j$ is R-linear as a function in column j of A (with the remaining columns fixed). Since $\det = \det_j$ we conclude that det is R-multilinear as a functions of its columns.

To show that det is alternating suppose that column r and column s of A are equal for some $1 \leq r < s$ of A. Suppose $n \geq 3$. Then we may choose $j \neq r$ and $j \neq s$. Then for each i, A_{ij} has two equal columns. Thus by induction $\det(A_{ij}) = 0_R$ for all i and so $\det(A) = \det_j(A) = 0_R$.

So suppose $n = 2$. Then $A = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$ and so $\det A = ab - ba = 0_R$. Thus det is alternating.

Suppose $A = I_n$. Then for $i \neq j$, A_{ij} has a zero column and so $\det(A_{ij} = 0_R$. For $i = j$, $A_{ii} = I_{n-1}$ and so by induction $\det A_{ii} = 1_R$. So $\det I_n = 1_R$. □

Lemma 7.1.8. *Let $n \in \mathbb{N}$ and $A = (a_{ij}) \in \mathrm{M}_n(R)$. Define the $A^{\mathrm{ad}} = (b_{ij}) \in \mathrm{M}_n(R)$ by $b_{ij} = (-1)^{ij} \det(A_{ji})$. Then*

$$\mathbb{A}^{\mathrm{ad}} A = \det(A) I_n$$

A^{ad} is called the adjoint *of A.*

Proof. Fix $1 \leq i, j \leq n$ and let $D = (d_{rs})$ be $n \times n$ obtained from A by replacing column i of A by column j if A. So

$$d_{rs} = \begin{cases} a_{rs} & \text{if } s \neq i \\ a_{rj} & \text{if } s = i \end{cases}$$

Note that $A_{ki} = D_{ki}$. The (i, j)-coefficient of $A^{\mathrm{ad}} A$ is

$$\begin{aligned} \textstyle\sum_{k=1}^n b_{ik} a_{kj} &= \textstyle\sum_{k=1}^n a_{kj} (-1)^{i+k} \det(A_{ki}) \\ &= \textstyle\sum_{k=1}^n d_{ki} (-1)^{i+k} \det(D_{ki}) \end{aligned}$$

The definition of $\det = \det_i$ shows that this is equal to $\det D$. If $i = j$, then $D = A$ and so $\det D = \det A$. If $i \neq j$, then the i and j columns of D are equal and so $\det A = 0_R$. Thus $A^{\mathrm{ad}} A = \det(A) I_n$. □

Proposition 7.1.9. *Let V and W be R-modules and I a finite set. Suppose V is free of finite rank and $\mathcal{B}$ is a finite R-basis for V. Choose a total order on I and a total order on $\mathcal{B}$. Let*

$$\mathcal{B}^I_< = \{(b_i)_{i\in I} \mid b_i < b_j \text{ for all } i < j \in I\}$$

Let $g : \mathcal{B}^I_< \to W$ be any function. Then there exists unique alternating R-multilinear function with $f : V^I \to W$ with $f|_{\mathcal{B}^I_<} = g$.

Proof. Let $f : V^I \to I$ be a alternating R-multilinear function with $f|_{\mathcal{B}^I_<} = g$. To show that f is unique it suffices to show that $f(b)$ is uniquely determined for all $b = (b_i)_{i\in I} \in \mathcal{B}^I$, (see 7.1.6). If $b_i = b_j$ for some $i \neq j \in I$, then since f is alternating $f(b) = 0_R$. So suppose that $b_i \neq b_j \in i$. Then there exists a unique $\pi \in \mathrm{Sym}(I)$ such that $b \circ \pi \in \mathcal{B}^I_<$ (note here that $b \circ \pi = (b_{\pi(i)})_{i\in I}$). Observe that there exist 2-cycles $\pi_j = (a_j, b_j) \in \mathrm{Sym}(I), 1 \leq j \leq k$ such that $\pi = \pi_1\pi_2 \dots \pi_k$. By 7.1.4(a), $f(c \circ \mu) = -f(c)$ for all $c \in V^I$ and any two cycle $\mu \in \mathrm{Sym}(I)$. $f(b) = (-1)^k f(b \circ \pi) = (-1)^k g(b \circ \pi)$ and so also $f(b)$ is uniquely determined.

To show the existence of f we assume without loss that $I = \{1, 2, \dots, n\}$ with the usual ordering. Let $v = (v_i)_{i\in I} \in V^I$. Then $v_i = \sum_{b\in\mathcal{B}} a_{ib}b$ for some unique $a_{ib} \in R, i \in I, b \in \mathcal{B}$. Let $A = (a_{ib}) \in M_{I\times\mathcal{B}}(R)$. For $b = (b_i)_{i\in I} \in \mathcal{B}^I_<$ let A_b be the $n \times n$ submatrix $(a_{ib_j})_{1\leq i,j\leq n}$ of A. Define

$$f(v) := \sum_{b\in\mathcal{B}^I_<} \det(A_b)g(b)$$

Since det is an alternating it is easy to see that f is alternating and R-multilinear. Suppose $v \in \mathcal{B}^I_<$ and $b \in \mathcal{B}^I_<$. Then $r_{ib_j} = \delta_{d_ib_j}$ Thus A_b has a zero column unless each b_j is equal to some d_i. Since both b and d are increasing, this shows that $\det(A_b) = 0_R$ for all $b \neq v$. For $b = v$, $A_b = I_n$ and so $\det(A_v) = 1$. So $f(v) = g(v)$ and $f|_{\mathcal{B}^I_<} = g$. □

Lemma 7.1.10. *Let V and W be R-modules and I a set.*

(a) Let $\mathrm{L}_I(V, W)$ is the set of R-multilinear map from $V^I \to W$. Then $\mathrm{L}_I(V, W)$ is an R-module via:

$$(f + g)(v) = f(v) + g(v) \text{ and } (rf)(v) = rf(v)$$

for all $f, g \in \mathrm{L}_I(V, W)$, $r \in R$ and $v \in V^I$.

(b) V^I is an R and an $\mathrm{End}_R(V)$-module via $u + v = (u_i + v_i)_{i\in I}$, $rv = (ru_i)_{i\in I}$ and $sv = (s(v_i))_{i\in I}$ for all $u = (u_i)_{I\in I}, v = (v_i)_{i\in I} \in V^I$, $r \in R$ and $s \in \mathrm{End}_R(V)$.

(c) The monoid $(\mathrm{End}_R(V), \circ)$ is acting on $\mathrm{L}_I(V, W)$ on the right via

$$(fs)(v) = f(sv)$$

for all $f \in \mathrm{L}_I(V, W)$, $s \in \mathrm{End}_R(V)$ and $v \in V^I$.

(d) $\mathrm{L}_I(V,W)$ *is a* $\mathrm{End}_R(W)$*-module via*

$$(tf)(v) = t(f(v))$$

for all $f \in \mathrm{L}_I(V,W)$, $t \in \mathrm{End}_R(W)$ *and* $v \in V^I$.

(e) Let $\bigwedge_I(V,W)$ *be the the set of alternating* R*-multilinear map from* $V^I \to W$. *Then* $\bigwedge_I(V,W)$ *is an* $\mathrm{End}_R(V)$*-invariant* R*-submodule of* $\mathrm{L}_I(V,W)$.

Proof. Readily verified. □

Corollary 7.1.11. *Let* V *and* W *be free* R*-modules with basis* $\mathcal{B}$ *and* $\mathcal{D}$ *and* I *a set. Suppose* I *and* $\mathcal{B}$ *are finite and choose a total orderings on* I *and a total ordering on* $\mathcal{B}$. *For* $b \in \mathcal{B}^I_<$ *and* $d \in \mathcal{D}$ *let* $f^{bd} : V^I \to W$ *be the unique alternating* R*-multilinear map with*

$$f^{bd}(c) = \begin{cases} d & \text{if } b = c \\ 0_W & \text{if } b \neq c \end{cases}$$

Then $\bigwedge_I(V,W)$ *is a free* R*-module with basis* $(f^{bd})_{(b,d)\in\mathcal{B}^I_<\times\mathcal{D}}$.

Proof. Let $f \in \bigwedge_I(V,W)$ and let $a_{bd} \in R$ for $b \in \mathcal{B}^I_<$ and $d \in \mathcal{D}$, almost all 0. Then

$$\begin{array}{rrcll} & f & = & \sum_{(b,d)\in\mathcal{B}^I_<\times\mathcal{D}} a_{bd} f^{bd} & \\ \iff & f(c) & = & \sum_{(b,d)\in\mathcal{B}^I_<\times\mathcal{D}} a_{bd} f^{bd}(c) & \text{for all } c \in \mathcal{B}^I_< \\ \iff & f(c) & = & \sum_{d\in\mathcal{D}} a_{cd} d & \text{for all } c \in \mathcal{B}^I_< \end{array}$$

Since $\mathcal{D}$ is a R-basis for $\mathcal{B}$ we see that there exists uniquely determined a_{cd} fulfilling the last of these equations. □

Definition 7.1.12. *Let* V *and* W *be* R*-modules,* $n \in \mathbb{N}$ *and* I *a set. Then* $\bigwedge_I(V) = \bigwedge_I(V,R)$, $\bigwedge_n(V,W) = \bigwedge_{\{1,2,\ldots,n\}}(V,W)$ *and* $\bigwedge_n(V) = \bigwedge_{\{1,2,\ldots,n\}}(V)$.

Lemma 7.1.13. *Let* V *be a free* R*-module of finite rank* n. *Let* $\alpha \in \mathrm{End}_R(V)$.

(a) There exists a unique $r_\alpha \in R$ *with*

$$f\alpha = rf \text{ for all } f \in \bigwedge_n(V)$$

(b) The map $\det : \mathrm{End}_R(V) \to R, \alpha \to r_\alpha$ *is a multiplicative homomorphism, that is* $\det(\alpha\beta) = \det(\alpha)\det(\beta)$ *for all* $\alpha, \beta \in \mathrm{End}_R(V)$.

(c) If A *is the matrix of* α *with respect to some* R*-basis of* V, *then* $\det(\alpha) = \det(A)$.

Proof. (a) Let $\mathcal{B} = \{b_1, b_2, \dots b_n\}$ be basis for V. Order $\mathcal{B}$ by $b_1 < b_2 \dots b_n$ and put $b = (b_1, b_2, \dots b_n)$. Put $I = \{1, 2, 3 \dots, n\}$ order in the usual way. Then clearly $\mathcal{B}^I_< = \{b\}$ and 1 is an R-basis for R. Thus by 7.1.11 f^{b1} is an R-basis for $\bigwedge_n(V)$. Hence

$$f^{b1}\alpha = r_\alpha f^{bl}$$

for a unique $r_\alpha \in R$. Also each $f \in \bigwedge_n(V)$ is of the form $f = rf^{b1}$ for some $r \in R$. Since α acts R-linearly on $\bigwedge_n(V)$ we conclude that (a) holds.

(b) Let $\alpha, \beta \in End_R(V)$ and $f \in \bigwedge_n(V)$. Then

$$f(\alpha\beta) = (f\alpha)\beta = (r_\alpha f)\beta = r_b eta(r_\alpha f) = (r_\beta r_\alpha)f = (r_\alpha r_\beta)f$$

Hence $r_{\alpha\beta} = r_\alpha r_\beta$ and (b) holds.

(c) We will compute $\det(\alpha)$. We have

$$\det(\alpha) = r_\alpha = r_\alpha 1_R = r_\alpha f^{b1}(b) = (r_\alpha f^{b1})(b) = (f^{b1}\alpha)(b) = f^{b1}(\alpha(b)) = f^{b1}\big((\alpha(b_j)_{j\in I}\big)$$

Let $A = (a_{ij})$ be the matrix for α with respect to $\mathcal{B}$. Then $\alpha(b_j) = \sum_{i\in I} a_{ij}b_i$. So

$$\det(\alpha) = f^{b1}\left(\Big(\sum_{i\in I} a_{ij}b_i\Big)_{j\in J}\right)$$

Since f^{bl} is alternating we see the function $\tau : \mathrm{M}_I(R) \to R, A \to \det(\alpha)$ is alternating and R-multilinear in the columns of A. Also if $A = \mathrm{id}_n$, then $\alpha = id_V$, $\det(\mathrm{id}_V) = 1_R$ and so $\tau(I_n) = 1_R$. 7.1.9 shows that τ is uniquely determined and so $\tau = \det$, that is $\det(\alpha) = \det(A)$. □

Lemma 7.1.14. *Let V be an R-module and I a finite set. Then V^I is an $\mathrm{M}_I(R)$-module via*

$$Av = \big(\sum_{i\in I} a_{ij}v_j\big)_{i\in I}$$

for all $A = (a_{ij})_{(i,j)\in I\times I} \in \mathrm{M}_I(R)$ and $v = (v_j)_{j\in I} \in V^I$.

Proof. Let $A = (a_{ij})$, $B = (b_{jk})$ and $C := AB$. Then $C = (c_{ik})$ with $c_{ik} = \sum_{j\in I} a_{ij}b_{jk}$. Let $v = (v_k)_{k\in I} \in V^I$. Then

$$\begin{aligned}
A(Bv) &= A\big(\textstyle\sum_{k\in I} b_{jk}v_k\big)_{j\in I} \\
&= \Big(\textstyle\sum_{j\in I} a_{ij}\big(\sum_{k\in I} b_{jk}v_k\big)\Big)_{i\in I} \\
&= \Big(\textstyle\sum_{k\in I} \big(\sum_{j\in I} a_{ij}b_{jk}\big)v_k\Big)_{i\in I} \\
&= \big(\textstyle\sum_{k\in I} c_{ik}v_k\big)_{i\in I} \\
&= Cv = (AB)v
\end{aligned}$$

□

Definition 7.1.15. *Let $n \in \mathbb{N}$ and $A \in \mathrm{M}_n(R)$. Note that $xI_n - A \in \mathrm{M}_n(R[x])$ and $R[x]$ is a commutative ring. So we can define*

$$\chi_A := \det(xI_n - A) \in R[x]$$

χ_A is called the characteristic polynomial of A.

Theorem 7.1.16 (Cayley Hamilton)**.** *Let $n \in \mathbb{N}$ and $A \in \mathrm{M}_n(R)$. Then $\chi_A(A) = O_n$, where $O_n = 0_{\mathrm{M}_n(R)}$ is the $n \times n$-zero matrix over R.*

Proof. By 3.2.5 the maps $R[x] \to \mathrm{M}_n(R), f \to f(A)$ is a ring homomorphism. Note that (for example by 7.1.14) R^n is an $\mathrm{M}_n(R)$ module via $Bv = (\sum_{j=1}^n b_{ij}v_j)_{i=1}^n$ for all $B = (b_{ij}) \in \mathrm{M}_n(R)$ and all $v = (v_j)_{j=1}^n \in R^n$. Thus $V := R^n$ is also an $R[x]$ module via $fv = f(A)v$ for all $f \in R[x]$ and $v \in V$. Note that $xv = Av$ for all $v \in R^n$. Since V is an $R[x]$-module we conclude from 7.1.14 that V^n is a $\mathrm{M}_n([R[x])$-module. Put $e_k = (\delta_{kj})_{j=1}^n \in V$. Then

$$xe_k = Ae_k = \left(\sum_{j=1}^n a_{ij}\delta_{kj}\right)_{i=1}^n = (a_{ik})_{i=1}^n = \sum_{i=1}^n a_{ik}e_i$$

Let $D = xI_n - A^{\mathrm{T}} \in \mathrm{M}_n(R[x])$ and $e = (e_j)_{j=1}^n \in V^n$. Then $D = (d_{ij})$ with $d_{ij} = \delta_{ij}x - a_{ji}$. Hence

$$De = \left(\sum_{j=1}^n d_{ij}e_j\right)_{i=1}^n = \left(\sum_{j=1}^n (\delta_{ij}x - a_{ji})e_j\right)_{i=1}^n = \left(xe_i - \sum_{j=1}^n a_{ji}e_j\right)_{i=1}^n = (xe_i - xe_i)_{i=1}^n = 0_{V^n}$$

Hence also $D^{\mathrm{ad}}(De) = 0_{V^n}$ and so $(D^{\mathrm{ad}}D)e = 0_{V^n}$. By 7.1.8 $D^{\mathrm{ad}}D = \det(D)I_n$ we have $(D^{\mathrm{ad}}D)e = (\det(D)e_i))_{i=1}^n$. Hence $\det(D)e_i = 0_V$ for all $1 \leq i \leq n$. By Homework 6#10, $\det(D) = \det(D^{\mathrm{tr}}) = \chi_A$ and so $\chi_A e_i = 0$ for all $v \in V$. But $\chi_A e_i = \chi_A(A)e_i$ and so the i-column of $\chi_A(A)$ is zero. Thus $\chi_A(A) = 0_n$. □

Lemma 7.1.17. *Let V and W be $R[x]$-modules and $\pi : W \to W$ a function. Then π is $R[x]$-linear if and only if π is R-linear and $\pi(xv) = x\pi(v)$ for all $v \in V$.*

Proof. The forward direction is obvious. So suppose π is R-linear and $\pi(xv) = x\pi(v)$ for all $v \in V$. Let $S = \{f \in R[x] \mid \pi(fv) = f\pi(v) \text{ for all } v \in V\}$. We will show that S is a subring of $R[x]$. Indeed let $f, g \in S$ and $v \in V$. Then

$$\pi((f+g)v(= \pi(fv + gv) = \pi(fv) + \pi(gv) = f\pi(v) = g\pi(v) = (f+g)\pi(v)$$

and

$$\pi((fg)(v) = \pi(f(gv)) = f\pi(gv)f(g\pi(v)) = (fg)\pi(v))$$

So $f + g, fg \in V$. Similarly 0_R and $-f \in S$. So S is a subring of $R[x]$. Since π is R-linear, $R \subseteq S$ and by assumption $x \in S$. Thus $S = R[x]$ and π is $R[x]$-linear. □

Theorem 7.1.18. *Let M be a finitely generated R-module and $\alpha \in \mathrm{End}_R(M)$. Then there exists a monic polynomial $f \in R[x]$ with $f(\alpha) = 0_{\mathrm{End}_R(M)}$.*

Proof. Let I be a finite subset of M with $M = \langle I \rangle_R = RI$. Then for each $j \in I$ there exist $a_{ij} \in R$, $i \in I$ with $\alpha(j) = \sum_{i \in I} a_{ij} i$. (Note here that the a_{ij} are not necessarily unique.) View R^I as an $R[x]$-module via $fv = f(A)v$ and view M as an $R[x]$ module via $fm = f(\alpha)(m)$. Define $\pi : R^I \to M, (r_i)_{i \in I} \to \sum_{i \in I} r_i i$. Then π is onto and R-linear. Let $e_i = (\delta_{ij})_{j \in I}$. By definition of π and A

$$x\pi(e_j) = xj = \alpha(j) \sum_{i=1}^{n} a_{ij} i$$

and

$$\pi(xe_j) = \pi(Ae_j) = \pi((a_{ij})_{i \in I}) = \sum_{i \in I} a_{ij} i$$

Thus $x\pi(e_j) = \pi(xe_j)$.. Since x acts R-linearly on R^I and M this implies $x\pi(v) = pi(xv)$ for all $v \in R^n$. Thus by 7.1.17, π is $R[x]$ linear. Put $f = \chi_A$. Then f is monic, $f \in R[x]$, $f(A) = 0$ and so for all $v \in R^n$,

$$f(\alpha)(\pi(v)) = f\pi(v) = \pi(f(v)) = \pi(f(A)v) = \pi(0v) = \pi(0) = 0.$$

Since π is onto we conclude that $f(\alpha) = 0$. □

7.2 Ring Extensions

Definition 7.2.1. *Let R and S be rings with $R \leq S$ and $1_S = 1_R$. Then S is called a* ring extension *of R. Such a ring extension is denoted by $R \leq S$.*

Definition 7.2.2. *Let $R \leq S$ be a ring extension.*

(a) Let $s \in S$. s is called integral *over R if $f(s) = 0$ for some monic polynomial $f \in R[x]$.*

(b) $R \leq S$ is called integral *if all $s \in S$ are integral over R.*

(c) $R \leq S$ is called finite if S is finitely generated as an R-module (by left multiplication)

Example 7.2.3. (1) Suppose $R \leq S$ is a ring extension with R a field and S an integral domain. Let $s \in S$. Then s is integral over R if and only if s is algebraic over R. $R \leq S$ is integral if and only if its algebraic. Note that then by 5.1.11 S is a field. $R \leq S$ is a finite ring extension if and only if its a finite field extension.

(2) Let $R = \mathbb{Z}$ and $S = \mathbb{C}$. Then $\sqrt{2}$ is integral over $\mathbb{Z}$. $\frac{1}{2}$ is not integral over $\mathbb{Z}$. Indeed suppose that $\frac{1}{2}$ is integral over $\mathbb{Z}$. Then there exists $a_0, a_1, \ldots, a_{n-1} \in \mathbb{Z}$ with

$$a_0 + a_1\frac{1}{2} + a_2\frac{1}{4} + \ldots a_{n-1}\frac{1}{2}^n + \frac{1}{2^n} = 0$$

Multiplication with 2^n shows that

$$1 = -(a_0 2^n + a_1 2^{n-1} + \ldots + a_{n-1}2)$$

since the left hand side of this equation is even, we derived a contradiction.

Theorem 7.2.4. *Let $R \leq S$ be a ring extension and $s \in S$. Then the following statements are equivalent:*

(a) s is integral over R.

(b) $R \leq R[s]$ is finite.

(c) There exists a subring T of S containing $R[s]$ such that $R \leq T$ is finite.

(d) There exists a faithful $R[s]$- module M, which is finitely generated as an R-module.

Proof. (a) $\Longrightarrow$ (b): Suppose $f(s) = 0$ for a monic $f \in R[x]$. Let $J = \{g \in R[x] \mid g(s) = 0\}$. Then $R[s] \cong R[x]/J$ and $R[x]f \leq J$.

We claim that $R[x]/R[x]f$ is finitely generated as an R-module. Indeed let $g \in R[x]$. Since f is monic we can apply the division algorithm and so $g = qf + r$, where $q, r \in R[x]$ with $\deg q < \deg f$. Let $n = \deg f$. We conclude that $g + R[x]$ is in the R span of $(x^i + R[x]f)_{i=0}^{n-1}$.

This proves the claim. Since $R[x]/J$ is isomorphic to a quotient of $R[x]/R[x]f$, also $R[X]/J$ and $R[s]$ are finitely generated as an R-module.

(b) $\Longrightarrow$ (c): Just choose $T = R[s]$.

(c) $\Longrightarrow$ (d): Put $B = T$. Since $1 \in T$, $aT \neq 0$ for all $0 \neq a \in R[s]$. Thus T is a faithful $R[s]$ module.

(d) $\Longrightarrow$ (a): By 7.1.18 there exists a monic $f \in R[x]$ with $f(s)M = 0$. Since M is faithful for $R[s]$, $f(s) = 0$. □

Corollary 7.2.5. *Let $R \leq S$ be a finite ring extension. Then $R \leq S$ is integral.*

Proof. This follows immediately from 7.2.4(c) applied with $T = S$. □

Lemma 7.2.6. *Let $R \leq E$ and $E \leq S$ be finite ring extensions. Then $R \leq S$ is a finite ring extension.*

Proof. Let I be a finite subset of E with $RI = E$ and J a finite subset of S with $S = EJ$. Then by 5.1.4(aa) $S = R\{ij \mid i \in I, J\}$. So also $R \leq S$ is finite □

Proposition 7.2.7. *Let $R \leq S$ be a ring extension and $I \subseteq S$ such that each $b \in I$ is integral over R.*

(a) If I is finite, $R \leq R[I]$ is finite and integral.

(b) $R \leq R[I]$ is integral.

(c) The set $Int(R,S)$ of the elements in S which are integral over R is a subring S. Moreover, $R \leq Int(R,S)$ is integral.

Proof. (a) By induction on $|I|$. If $|I| = 0$ there is nothing to prove. So suppose there exists $i \in I$ and let $J = I \setminus \{i\}$. Put $E = R[J]$. By induction $R \leq E$ is finite. Since i is integral over R, f is integral over E. Thus by 7.2.4(b), $E \leq E[i]$ is finite. Note that $E[i] = R[J][i] = R[I]$ and so (a) follows from 7.2.6.

(b) By 5.1.3(a) $R[I] = \bigcup\{R[J] \mid J \subseteq I, |J| < \infty\}$. By (a) each of the extensions $R \leq R[J]$ are integral. So(b) holds.

(c) Follows from (b) applied to $I = \mathrm{Int}(T,S)$. □

Definition 7.2.8. *Let $R \leq S$ be a ring extension and let $Int(R,S)$ the set of elements in S which are integral over R. Then $Int(R,S)$ is ca called to* integral closure *of R in S. If $R = Int(R,S)$, then R is called T*integrally closed *in S.*

If R is an integral domain and R is integrally closed in $\mathbb{F}_R$ (the field of fraction of R), then R is called integrally closed.

Example 7.2.9. Let $A = \mathrm{Int}(\mathbb{Z}, \mathbb{C})$. Then A is the set of complex numbers which are the roots of an integral monic polynomial. So A is the set of algebraic integers (see Homework 2#6). We now know from 7.2.7 that A is a subring of $\mathbb{C}$, which generalized Homework 2#6(c). By Homework 2#6(b), $A \cap \mathbb{Q} = \mathbb{Z}$. Thus $\mathrm{Int}(\mathbb{Z}, \mathbb{Q}) = \mathbb{Z}$ and so $\mathbb{Z}$ is integrally closed. But $\mathbb{Z}$ is not integrally closed in $\mathbb{C}$ since $\sqrt{2} \in A$.

Lemma 7.2.10. *Let $R \leq E$ and $E \leq S$ be integral ring extensions. Then $R \leq S$ is integral.*

Proof. Let $s \in S$ and let $f \in E[x]$ be the monic with $f(s) = 0$. Let I be the set of non-zero coefficients f. Then I is a finite subset of E and so by 7.2.7(a), $R \leq R[I]$ is finite. Since $f \in R[I][x]$, 7.2.4 implies that $R[I] \leq R[I][s]$ is finite. So by 7.2.6, $R \leq R[I][s]$ is finite. So by 7.2.4, s is integral over R. □

7.3 Ideals in Integral Extensions

Definition 7.3.1. *Let R be ring and I an ideal in R. Then*

$$\mathrm{rad}\, I = \mathrm{rad}_R I = \{r \in R \mid r^n \in I \text{ for some } n \in \mathbb{Z}^+\}.$$

$\mathrm{rad}_R I$ *is called the* radical *of I in R . If $I = \mathrm{rad}_R I$, I is called a radical ideal in R.*

Lemma 7.3.2. *Let R be a ring and P an ideal in R.*

(a) $\mathrm{rad}P$ *is an ideal in R and $P \leq \mathrm{rad}P$.*

(b) $\mathrm{rad}P$ *is a radical ideal.*

(c) All primes ideals in R are radical ideals.

Proof. (a) Note that $r \in \mathrm{rad}P$ if and only if $r + P$ is nilpotent in R/P. By Homework 6#5 in MTH 818, the nilpotent elements of R/P form an ideal in R/P. So (a) holds.

(b) Homework 6#5 in MTH 818 $R/\mathrm{rad}\, R/\mathrm{rad}\, R/P$ has no non-zero nilpotent elements.

(c) If P is a prime ideal, then R/P has no zero divisors and so also no non-zero nilpotent elements. □

Lemma 7.3.3. *Let $R \leq S$ be an integral ring extension.*

(a) Let P be an ideal in R and $p \in P$.

- *(a) $Sp \cap R \subseteq \mathrm{rad}_R P$.*
- *(b) If P is a prime ideal or a radical ideal in R, then $Sp \cap R \subseteq P$.*

(b) Suppose S is an integral domain

- *(a) Let $0 \neq b \in S$, then $Sb \cap R \neq 0$.*
- *(b) Let Q be a non-zero ideal in S, then $Q \cap R \neq 0$.*

Proof. (a) Let $s \in S$ such that $r := sp \in R$. Since $R \leq S$ is integral there exists $r_0, r_1 \ldots r_{n-1} \in R$ with

$$s^n = r_{n-1}s^{n-1} + \ldots + r_1 s + r_0$$

Multiplying this equation with p^n we obtain:

$$(sp)^n = (r_{n-1}p)(sp)^{n-1} + \ldots + r_1 p^{n-1}(sp) + r_0 p^n$$

Hence

$$r^n = (r_{n-1}p)r^{n-1} + \ldots + (r_1 p^{n-1})r + r_0 p^n$$

As P is an ideal and $r_i r^i \in R$ we have $r_i r^i p^{n-i} \in P$ for all $0 \leq i < n$. So the right side of the last equation lies in P. Thus $r^n \in P$ and $r \in \mathrm{rad}P$.

(a:b) In both cases 7.3.2 implies that $P = \mathrm{rad}_R P$. So (a:b) follows from(a:a).

(b:a) Let $f \in R[x]$ be a monic polynomial of minimal degree with $f(b) = 0$. Let $f = xg + r$ where $r \in R$ and $g \in R[x]$ is monic of degree one less than f. Then

$$0 = f(b) = bg(b) + r$$

and so $r = -g(b)b$

If $r = 0$, we get $g(b)b = 0$. Since $b \neq 0$ and S is an integral domain, $g(b) = 0$. But this contradicts the minimal choice of $\deg f$.

Hence $0 \neq r = -g(b)b \in R \cap Sb$.

(b:b) Let $0 \neq b \in Q$. Then by (b:a) $\{0\} \neq R \cap Sb \subseteq R \cap Q$. □

Theorem 7.3.4. *Let $R \leq S$ be an integral extension and P a prime ideal in R. Put*

$$\mathcal{M} := \{I \mid I \text{ is an ideal in } R, R \cap I \subseteq P\}$$

Order $\mathcal{M}$ by inclusion. Let $Q \in \mathcal{M}$.

(a) Q is contained in a maximal member of $\mathcal{M}$

(b) The following are equivalent:

(a) Q is maximal in $\mathcal{M}$.

(b) Q is a prime ideal and $R \cap Q = P$.

Proof. Put $\mathcal{M}_Q := \{I \in \mathcal{M} \mid Q \leq I\}$. Then a maximal element of $\mathcal{M}_Q$ is also a maximal element of $\mathcal{M}$.

(a) Since $Q \in \mathcal{M}_Q$, $\mathcal{M}_Q \neq \emptyset$. So by Zorn's Lemma A.6 it remains to show that every non-empty chain $\mathcal{D}$ in $\mathcal{M}_Q$ has an upper bound in $\mathcal{M}_Q$. Put $D = \bigcup \mathcal{D}$. By 3.3.15(a) D is an ideal in S. Let $E \in \mathcal{D}$. Then $Q \leq E \leq D$. Moreover,

$$R \cap D = \bigcup_{E \in \mathcal{D}} R \cap E \subseteq P$$

Thus $D \in \mathcal{M}_Q$ and D is an upper bound for $\mathcal{D}$.

(b) For $E \subseteq S$ put $\overline{E} = E + Q/Q \subseteq S/Q$. Since S is integral over R, $\overline{S}$ is integral over $\overline{R}$. (Indeed let $\overline{s} \in \overline{S}$. Then $\overline{s} = s + Q$ for some $s \in S$ and there exists a monic polynomial $f \in R[x]$ with $f(s) = 0$. The $\overline{f}$ is a monic polynomial in $\overline{S}[x]$ and $\overline{f}(\overline{s}) = 0$.

$$\begin{array}{rclcrcl} \overline{R}/\overline{P} & = & R + Q/Q/P + Q/Q & \cong & R + Q/P + Q & = & R + (P + Q)/P + Q \\ & \cong & R/R \cap (P + Q) & = & R/P + (R \cap Q) & = & R/P \end{array}$$

Since P is a prime ideal in R we conclude that $\overline{P}$ is a prime ideal in $\overline{R}$. Let I be an ideal in S with $Q \leq I$. We have

$$\begin{array}{rc} & \overline{R} \cap \overline{I} \leq \overline{P} \\ \iff & ((R + Q)/Q) \cap I/Q \leq P + Q/Q \\ \iff & (R + Q) \cap I \leq P + Q \\ \iff & Q + (R \cap I) \leq P + Q \end{array}$$

If $R \cap I \leq P$ we have $Q + (R \cap I) \leq P + Q$. If $Q + (R \cap I) \leq P + Q$, then $R \cap I \leq (P + Q) \cap R = P + (Q \cap R) \leq P$. So

$$\overline{R} \cap \overline{I} \leq \overline{P} \iff R \cap I \leq I$$

Therefore $\{\overline{I} \mid I \in \mathcal{M}_Q\} = \{J \leq \overline{S} \mid J \text{ is an ideal in } \overline{S}, \overline{R} \cap J \subseteq \overline{P}\}$.

If follows that (b) holds if and only if (b) holds for $(\overline{S}, \overline{R}, \mathrm{P}, \overline{Q}$ in place of (S, R, P, Q). Since $\overline{Q} = 0$ we thus may assume that $Q = 0$.

(b:a) $\Longrightarrow$ (b:b): Suppose that Q is not a prime ideal. As $Q = 0$, this means S is not an integral domain. Hence there exists $b_1, b_2 \in S^\sharp$ with $b_1 b_2 = 0$. Since $Q = 0$ is maximal in $\mathcal{M}$, $Sb_i \notin \mathcal{M}$ and so $R \cap Sb_i \not\leq P$. Hence there exist $s_i \in S$ with $0 \neq r_i := s_i b_i \in R \setminus P$. But then $r_1 r_2 = (s_1 b_1)(s_2 b_2) = (s_1 s_2)(b_1 b_2) = 0 \in P$. But this contradicts the fact that P is a prime ideal in R.

So Q is a prime ideal. Suppose that $P \neq R \cap Q$, that is $P \neq 0$. Let $0 \neq p \in P$. Then by 7.3.3(a) , $Sp \cap R \leq P$. Hence $Sp \in \mathcal{M}$, contradiction the maximality of $Q = 0$. So (b:a) implies (b:b).

(b:b) $\Longrightarrow$ (b:a): Suppose now that Q is a prime ideal and $P = R \cap Q$. Since $Q = 0$ this means that S is an integral domain and $P = 0$. Let I be any non-zero ideal in S. Then by 7.3.3(b:b) $R \cap I \neq 0$ and so $R \cap I \not\leq P$ and $I \notin \mathcal{M}$. Thus $\mathcal{M} = \{0\}$ and Q is maximal. □

Corollary 7.3.5. *Let $R \leq S$ be an integral extension.*

(a) Let P be a prime ideal in R and Q an ideal in S with $R \cap Q \leq P$. Then there exists a prime ideal M in S with $R \cap M = P$ and $Q \leq M$.

(b) Let P be a prime ideal in R. Then there exists a prime ideal M in S with $R \cap M = P$.

(c) Let Q_1 and Q_2 be prime ideals in S with $R \cap Q_1 = R \cap Q_2$ and $Q_1 \leq Q_2$. Then $Q_1 = Q_2$.

(d) Let Q be a maximal ideal in S. Then $Q \cap R$ is a maximal ideal in R.

(e) Let P be a maximal ideal in S. Then there exists a maximal ideal M of S with $R \cap M = P$.

Proof. (a) We apply 7.3.4. Let $\mathcal{M}$ be defined as there. By part (a) there exists a maximal element M of $\mathcal{M}$ containing Q. By part (b) M is a prime ideal and $R \cap M = P$.

(b) follows from (a) applied with $Q = 0$.

(c) By 7.3.4, applied with $P = R \cap Q_1$ and $Q = Q_1$ we get that Q_1 is maximal in $\mathcal{M}$. As $Q_2 \in \mathcal{M}$ and $Q_1 \leq Q_2$, $Q_1 = Q_2$.

(d) Since $1 \notin Q$, $R \cap Q \neq R$. So by 3.2.15, $Q \cap R$ is contained in a maximal ideal P of R. By (a) there exists an ideal M in S with $P = R \cap M$ and $Q \leq M$. Since Q is maximal, $M = Q$. Thus $R \cap Q = R \cap M = P$ and so $R \cap Q$ is a maximal ideal in R.

(e) By 3.2.16, P is a prime ideal in R. So by (b) there exists an ideal Q of S with $R \cap Q = P$. Let M be a maximal ideal in S with $Q \leq M$. Then $P = R \cap Q \leq R \cap M < R$ and since P is a maximal ideal in R, $P = R \cap M$. □

7.4 Noether's Normalization Lemma

Definition 7.4.1. *Let $\mathbb{K}$ be a field. A $\mathbb{K}$-algebra is a ring R with $\mathbb{K}$ as a subring. A $\mathbb{K}$-algebra R is called finitely generated if $R = \mathbb{K}[I]$ for some finite subset I of K*

Theorem 7.4.2. *Let $\mathbb{K}$ be a field and R a $\mathbb{K}$-algebra. Suppose that there exists a finite subset I of R such that $\mathbb{K}[I] \leq R$ is integral. Then there exists a finite subset J of R such that J is algebraically independent over $\mathbb{K}$ and $\mathbb{K}[J] \leq R$ is integral.*

Proof. Choose a finite subset I of R of minimal size such that$\mathbb{K}[I] \leq R$ is integral. Suppose that $u =: (i)_{i\in i}$ is not algebraic independent over $\mathbb{K}$ and pick $0 \neq f \in \mathbb{K}[x_i, i \in I]$ with $f(u) = 0$. Put $J = \bigoplus_{j\in J} \mathbb{N}$. Then $f = \sum_{\alpha\in J} k_\alpha x^\alpha$, where $k_\alpha \in \mathbb{K}$, almost all 0. Put $J^* = \{\alpha \in \mathbb{I} \mid k_\alpha \neq 0\}$. Then

$$\sum_{\beta\in J^*} k_\beta u^\beta = 0 \tag{1}$$

where $u^\beta = \prod_{i\in I} i^{\beta_i}$. Since J^* is finite, we can pick $c \in \mathbb{Z}^+$ with $\alpha_i < c$ for all $\alpha \in \mathbb{J}^*$ and $i \in I$. Fix $l \in I$ and let $(t_i) \in \mathbb{N}^I$ be a 1-1 function with $t_l = 0$. Define

$$\rho : \mathbb{J}^* \to \mathbb{Z}^+, \quad \alpha \to \sum_{i\in I} c^{t_i}\alpha_i$$

We claim that ρ is one to one. Indeed suppose that $\rho(\alpha) = \rho(\beta)$ for $\alpha \neq \beta \in \mathbb{J}^*$. Let $I^* = \{i \in I \mid \alpha_i \neq \beta_i$ and $k \in I^*$ with t_k is minimal.

$$0 = \rho(\alpha) - \rho(\beta) = c^{t_k}\left(\alpha_k - \beta_k + \sum_{i\in I^*\setminus\{k\}} c^{t_i - t_k}(\alpha_i - \beta_i)\right)$$

Since t is 1-1, $t_k < t_j$ for all $i \in I^* \setminus \{k\}$. So we conclude that c divides $\alpha_k - \beta_k$, a contradiction to $c > \alpha_j$ and $c > \beta_j$.

Since ρ is 1-1, we can choose $\alpha \in J^*$ with $\rho(\alpha) < \rho(\beta)$ for all $\beta \in J^* \setminus \{\alpha\}$.

For $i \in I$ define $v_i = i - l^{c^{t_i}}$. Put $S := \mathbb{K}[v_i, i \in I]$. Note that $v_l = l - l^{c^{t_l}} = l - l^{c^0} = l - l^1 = 0$. So

$$S = \mathbb{K}[V_i, \in I \setminus \{l\}] \tag{2}$$

We will show that l is integral over S. Let $\beta \in J^*$. Since $i = l^{c^{t_i}} + v_i$ we have

$$u^\beta = \prod_{i\in I} i^{\beta_i} = \prod_{i\in I}(l^{c^{t_i}} + v_i)^{\beta_i}.$$

Thus $u^\beta = g_\beta(l)$ where $g_\beta \in S[x]$ is a monic of degree $\rho(\beta)$. Put

$$g := \sum_{\beta\in J^*} k_\beta g_\beta \in S[x].$$

Then maximality of $\rho(\alpha)$ shows that g has degree $\rho(\alpha)$ and leading coefficient k_α. Moreover,

$$g(l) = \sum_{\beta \in J^*} k_\beta g_\beta(l) = \sum_{\beta \in J^*} k_\beta u^\beta = 0.$$

Thus $k_\alpha - 1g$ is a monic polynomial over S with l as a root and so l is integral over S. Note that $i = v_i + l^{c_i^t} \in S[l]$ and thus $\mathbb{K}[I] \leq S[l] \leq \mathbb{K}[I]$. So $\mathbb{K}[I] = S[l]$ and $S \leq \mathbb{K}[l]$ is integral. Since also $\mathbb{K}[I] \leq R$ is integral we conclude from 7.2.10 that $S \leq R$ is integral. But this contradicts (2) and the minimal choice of $|J|$. □

Proposition 7.4.3. *Let $R \leq S$ be an integral extension and suppose that that R and S are integral domains. Then S is a field if and only if R is a field.*

Proof. Suppose first that R is a field. Then $R \leq S$ is algebraic and so by 5.1.11(c) , S is a field.

Suppose next that S is a field and let $r \in R^\sharp$. Since S is a field, $1 \in Sr \cap R$. Hence by 7.3.3(a:b) applied with $P = Rr$, $1^n \in Rr$ for some $n \in \mathbb{Z}^+$. Thus $1 = tr$ for some $t \in T$. Hence r is invertible in R, and R is a field. □

Proposition 7.4.4. *(a) Let $\mathbb{K} \leq \mathbb{F}$ be a field extensions such that $\mathbb{F}$ is finitely generated over $\mathbb{K}$ as a ring. Then $\mathbb{K} \leq \mathbb{F}$ is finite. In particular, if $\mathbb{K}$ is algebraically closed then $\mathbb{F} = \mathbb{K}$.*

(b) Let $\mathbb{K}$ be an algebraically closed field, A a finitely generated $\mathbb{K}$-algebra and M a maximal ideal in A. Then $A = \mathbb{K} + M$.

Proof. (a) By 7.4.2 there exists a finite subset J of $\mathbb{K}$ such that $\mathbb{K}[J] \leq \mathbb{F}$ is integral and J is algebraically independent over $\mathbb{K}$. By 7.4.3, $\mathbb{K}[J]$ is a field. Since the units in $\mathbb{K}[J]$ are $\mathbb{K}$ we get $J = \emptyset$. Hence $\mathbb{K} \leq \mathbb{F}$ is integral and so algebraic. Thus by 5.1.11 $\mathbb{K} \leq \mathbb{F}$ is finite.

(b) Note that $\overline{A} := A/M$ is a field. Also $\overline{\mathbb{K}} = (\mathbb{K} + M)/M$ is a subfield of $\overline{A}$ isomorphic to $\mathbb{K}$ and $\overline{A}$ is a finitely generated $\overline{\mathbb{K}}$ algebra. So by (a) $\overline{A} = \overline{\mathbb{K}}$ and thus $A = \mathbb{K} + M$. □

7.5 Affine Varieties

Hypothesis 7.5.1. *Throughout this section $\mathbb{K} \leq \mathbb{F}$ is field extension with $\mathbb{F}$ algebraically closed. D is a finite set, $A = \mathbb{K}[x_d, d \in D]$ and $B = \mathbb{F}[x_d, d \in D]$, with A viewed as a subset of B.*

Definition 7.5.2. *Let $S \subseteq A$ and $U \subseteq \mathbb{F}^D$.*

(a) $V(S) = V_{\mathbb{F}^D}(S) = \{v \in \mathbb{F}^D \mid f(v) = 0$ *for all* $f \in S\}$.

$V(S)$ *is called an* affine variety *in* $\mathbb{F}^D$ *defined over* $\mathbb{K}$*, or a* $\mathbb{K}$*-variety in* $\mathbb{F}^D$.

(b) $U \subseteq \mathbb{F}^D$ *define* $J(U) := J_A(U) = \{f \in A \mid f(u) = 0$ *for all* $u \in U\}$.

(c) U *is called* closed *if* $U = V(J(U))$ *and* S *is called closed if* $S = J(V(S))$.

Lemma 7.5.3. *Let $U \subseteq \tilde{U} \subseteq \mathbb{F}^D$ and $S \subseteq \tilde{S} \subseteq A$.*

(a) $J(U)$ is an ideal in R.

(b) $J(\tilde{U}) \subseteq J(U)$.

(c) $V(\tilde{S}) \subseteq V(S)$.

(d) $U \subseteq V(J(U))$.

(e) $S \subseteq J(V(S))$.

(f) The following are equivalent:

(a) U is $\mathbb{K}$-variety in $\mathbb{F}^D$.

(b) $U = V(S)$ for some $S \subseteq A$.

(c) U is closed.

(d) $U = V(I))$ for some ideal I of A.

(g) S is closed if and only if $S = J(U)$ for some $U \subseteq F^D$.

(h) $V(S) = V(AS)$.

Proof. (a) Clearly $0 \in J(U)$. Let $f, g \in J(U)$, $h \in A$ and $u \in U$. Then $(f-g)(u) = f(g) - g(u) = 0$ and $(hf)(u) = h(u)f(u) = 0$. So $f - g \in J(U)$ and $hf \in J(U)$.

(b) and and (c) are obvious.

(d) Let $u \in U$. Then for all $f \in J(U)$, $f(u) = 0$. So (d) holds.

(e) Similar to (d).

(f) Suppose U is $\mathbb{K}$-variety in $\mathbb{F}^D$, then by definition $U = V(S)$ for some $S \subseteq A$. So (f:a) implies (f:b)

Suppose $U = V(S)$ for some $S \subseteq A$. Then by (d) $S \subseteq J(U)$ and so by (b) $V(J(U)) \subseteq J(S) = U$. By (d), $U \leq (J(U))$ and hence $U = V(J(U))$. So (f:b) implies (f:c)

Suppose U is closed. Then $U = U(J(U))$. By (a) $J(U)$ is an ideal in A and so (f:c) implies (f:d).

Clearly (f:d) implies (f:a). So (f) holds.

(g) If S is closed then $S = J(U)$ for $U = J(S)$. The other direction is similar to the implication (f:b) $\Longrightarrow$ (f:c).

(h) Since $S \subseteq AS$, $V_(AS) \leq V(S)$. By (e) $S \subseteq J(V(S))$ and by (a), $J(V(S))$ is an ideal. Thus $AS \subseteq J(V(S))$ and so $V(S) \subseteq V(AS)$. □

Example 7.5.4. (1) Suppose that $|D| = 1$ and so $A = \mathbb{K}[x]$ and $\mathbb{F}^D = \mathbb{F}$. Let U be a affine $\mathbb{K}$-variety in $\mathbb{F}$. Then by 7.5.3(f), $U = V(I)$ for some ideal I in $\mathbb{K}[x]$. By 3.4.6, $\mathbb{K}[x]$ is a PID and so there exists $f \in \mathbb{K}[x]$ with $I = \mathbb{K}[x]f$. Thus by 7.5.3(h), $U = V(I) = V(f)$. So U is the set of roots of f in $\mathbb{F}$. So either $f = 0$ and $U = \mathbb{F}$ or U is finite.

Now let U be any finite subsets of $\mathbb{K}$ and put $f = \prod_{u \in U} x - u$. Then $V(f) = U$ and so any finite subsets of $\mathbb{K}$ are is an affine $\mathbb{K}$-variety in $\mathbb{F}$.

If $\mathbb{K} = \mathbb{F}$ we see the affine $\mathbb{F}$-varieties in $\mathbb{F}$ are $\mathbb{F}$ itself and the finite subsets of $\mathbb{F}$.

(2) Let $\mathbb{K} = \mathbb{R}$, $\mathbb{F} = \mathbb{C}$ and $D = \{1, 2\}$. Then $A = \mathbb{K}[x_1, x_2]$. Let $f = x_1^2 + x_2^1 - 1$. Then $V(f) = \{(a, b) \in \mathbb{C}^2 \mid a^2 + b^2 = 1\}$.

(3) Let $n \in \mathbb{Z}^+$ and $D = \{(i, j) \mid 1 \leq i, j \leq n\}$. Then $\mathbb{F}^D = M_n(\mathbb{F})$ is the set of $n \times n$-matrices with coefficients in $\mathbb{F}$. Write x_{ij} for $x_{(i,j)} \in A$ and consider the matrix $X := (x_{ij}) \in M_n(A)$. Put $f = \det(X) \in A$. Let $u = (u_{ij}) \in \mathbb{F}^D = M_n(\mathbb{F})$. Then $f(u) = \det u$. Thus

$$V(f - 1) = \{u \in M_n(\mathbb{F}) \mid \det(u) = 1\} = \mathrm{SL}_n(\mathbb{K})$$

Lemma 7.5.5. *Let $u \in \mathbb{F}^D$.*

(a) $J(u)$ is the kernel of the evaluation map: $\Phi : A \to \mathbb{F}, f \to f(u)$.

(b) If $\mathbb{K} \leq \mathbb{F}$ is algebraic, $J(u)$ is a maximal ideal in A.

Proof. (a) is obvious. (b) Note that $\mathbb{K} \leq \Phi(\mathbb{K}) \leq \mathbb{F}$. Therefore $\Phi(\mathbb{K})$ is an integral domain which is algebraic over $\mathbb{K}$. So by 5.1.11 $\Phi(\mathbb{K})$ is an field. By (a) and the first isomorphism theorem for rings, $A/J(u)$ is a field and so by 3.2.19 $J(u)$ is a maximal ideal in A. □

Lemma 7.5.6. *Let M be a maximal ideal in B.*

(a) There exists $u = (u_d)_{d \in D} \in \mathbb{F}^D$ with $M = J_B(u)$.

(b) M is the ideal in B generated by $\{x_d - u_d \mid d \in D\}$.

(c) $V(M) = \{u\}$.

Proof. (a) and (b) By 7.4.4, $B = \mathbb{F} + M$. Hence for each $d \in D$ there exists $u_d \in \mathbb{F}$ with $x_d - u_d \in M$. Put $u = (u_d)_{d \in D}$ and let I be the ideal generated by $\{x_d - u_d \mid d \in D)$. Then $x_d \in \mathbb{F} + I$ and so $\mathbb{F} + I$ is a subring of B containing $\mathbb{F}$ and all x_d. Hence $B = \mathbb{F} + I$ and B/I is a field. So I is a maximal ideal. Since $I \leq M$ we get $I = M$ and since $I \leq J_B(u)$, $I = J_B(u)$. So $M = I = J_B(u)$.

(c) Let $a \in V(M)$. Since $x_d - u_d \in M$, $0 = (x_d - u_d)(a) = a_d - u_d$. Hence $a_d = u_d$ and $a = d$. □

Proposition 7.5.7. *Let I be an ideal in A with $I \neq A$. Then $V(I) \neq \emptyset$*

Proof. By 3.2.15 I is contained in a maximal ideal P of A. Let $\mathbb{A}$ be the set of elements in $\mathbb{F}$ algebraic over $\mathbb{K}$. Then

$$V_{\mathbb{A}^D}(P) \subseteq V(P) \subseteq V(I),$$

and so we may assume that $\mathbb{F} = \mathbb{A}$ and I is maximal in A. Then $\mathbb{K} \leq \mathbb{F}$ is algebraic and so each $b \in \mathbb{F} \subseteq B$ is integral over $\mathbb{K}$ and so also over A. Since $B = A[\mathbb{F}]$ we conclude from 7.2.7 that $A \leq B$ is integral. Hence by 7.3.5, there exists a maximal ideal M of B with $I = A \cap M$. By 7.5.6, $V(M) \neq \emptyset$. Since $V(M) \subseteq V(I)$ the proposition is proved. □

Theorem 7.5.8 (Hilberts' Nullstellensatz)**.** *Let I be an ideal in A. Then $J(V(I)) = \mathrm{rad}I$. In other words, I is closed if and only if I is a radical ideal.*

Proof. Let $f \in \mathrm{rad}I$ and $u \in V(I)$. Then $f^n \in I$ for some $n \in \mathbb{Z}$. Thus $(f(u))^n = f^n(u) = 0$ and since $\mathbb{F}$ is an integral domain, $f(u) = 0$. Thus $f \in J(V(I))$ and $\mathrm{rad}I \subseteq J(V(I))$.

Next let $0 \neq f \in J(V(I))$. We need to show that $f \in \mathrm{rad}I$. Put $E = D \cup \{f\}$ and put $y = x_f$. Then $\mathbb{K}[x_e, e \in E] = A[y]$. Let L be the ideal in $A[y]$ generated by I and $yf - 1$.

Suppose for a contradiction that $V_{\mathbb{F}^E}(L) = \emptyset$ and pick $c \in V_{\mathbb{F}^E}(L)$. Then $c = (a, b)$ with $a \in \mathbb{F}^D$ and $b \in \mathbb{F}$. Let $g \in I$. Then $g \in L$ and $0 = g(a, b) = g(a)$. Thus $a \in V(I)$. Since $f \in J(V(I))$ we get $f(a) = 0$. Hence $0 = (yf - 1)(a, b) = bf(a) - 1 = -1 \neq 0$, a contradiction.

Thus $V_{\mathbb{F}^E}(L) = \emptyset$. 7.5.7 implies $L = A[y]$. So there exist $g_s(y) \in A[y], 0 \leq s \leq m$ and $f_s \in I, 1 \leq s \leq m$, with

$$(*) \qquad 1 = g_0(y)(yf - 1) + \sum_{s=1}^{m} g_i(y) f_i.$$

Let $\mathbb{F}_A = \mathbb{K}(x_d, d \in D)$ be the field of fractions of A. Let $\phi : A[y] \to \mathbb{F}_A$ be the unique ring homomorphism with $\phi(a) = a$ for all $a \in A$ and $\phi(y) = f^{-1}$. (see 5.2.29. Applying ϕ to (*) we obtain:

$$(**) \qquad 1 = g_0(f^{-1})(f^{-1}f - 1) + \sum_{s=1}^{m} g_i(f^{-1}) f_i = \sum_{s=1}^{m} g_i(f^{-1}) f_i.$$

Let $k \in \mathbb{Z}^+$ with $k \geq \deg_y g_i(y)$ for all $1 \leq i \leq m$. Then $g_i(f^{-1})f^k \in A$ for all i and thus $g_i(f^{-1})f^k f_i \in AI = I$. So multiplying equation (**) with f^k we get $f^k \in I$ and $f \in \mathrm{rad}I$. □

Corollary 7.5.9. *Then map $U \to J(U)$ is a inclusion reversing bijection with inverse $I \to V(I)$ between the affine $\mathbb{K}$-varieties in $\mathbb{F}^D$ and the radical ideals in A.*

Proof. Let U be an affine $\mathbb{K}$-variety in $\mathbb{F}^D$. Then by definition, $U = V(S)$ for some ideal S if A. So by 7.5.3(f)(g), U and $I := J(U)$ are closed. Thus $V(J(U)) = U$ and by Hilbert's Nullstellensatz, $I = J(V(I)) = \mathrm{rad}I$. So I is a radical ideal.

Suppose next that I is a radical ideal in A. Then by definition $V(I)$ is a affine $\mathbb{K}$-variety in $\mathbb{F}^D$ and by Hilbert's Nullstellensatz, $I = \mathrm{rad}I = J(V(I))$.

Finally by 7.5.3(b), $U \to J(U)$ is inclusion reversing. □

We would like to show that every affine variety is of the form $V(S)$ for a finite subset S of A. For this we need a little excursion:

Definition 7.5.10. *A ring R is called* Noetherian *if every ideal in R is finitely generated as an ideal.*

Theorem 7.5.11 (Hilbert's Basis Theorem)**.** *Let R be a Noetherian ring. Then also $R[x_d, d \in D]$ is Noetherian.*

Proof. By induction on $|D|$ it suffices to show that $R[x]$ is Noetherian.

Let J be an ideal in $R[x]$. For $n \in \mathbb{N}$ let J_n be the set of all $r \in R$ such that $r = 0$ or $r = \text{lead}(f)$ for some $f \in J$ with $\deg f = n$. Observe that J_n is an ideal in R. Since $\text{lead}(xf) = \text{lead}(f)$, $J_n \subseteq J_{n+1}$. Let $0 \leq n \leq t$. By 3.3.17 $\{J_n \mid n \in \mathbb{N}\}$ has a maximal element say J_t, for some $t \in \mathbb{N}$. Then $J_m = J_t$ for all $m \geq t$. By assumption each J_n is finitely generated and so we can choose $r_{nj}, 1 \leq j \leq k_n$ with

$$(*) \qquad J_n = \sum_{j=1}^{k_n} R r_{nj}.$$

For $0 \leq n \leq t$ and $1 \leq j \leq k_n$ pick $f_{nj} \in J$ with

$$(**) \qquad \text{lead}(f_{nj}) = r_{nj}.$$

Let I be the ideal in $R[x]$ generated by the $f_{nj}, 0 \leq n \leq t$ and $1 \leq j \leq k_n$. Note that $I \subseteq J$. For $m > t$ put $k_n := k_t$, $r_{mj} := r_{nj}$ and $f_{mj} := x^{m-t}$. Since $J_m = J_t$ we conclude that (*) and (**) holds for all $n \in \mathbb{N}$. Moreover $f_{nj} \in I$ for all n, j.

We will now show that $J = I$. So let $f \in J$. If $f = 0$, $f \in I$. So suppose $f \neq 0$ and let $n = \deg f$ and $s = \text{lead}(f)$. By (*),

$$s = \sum_{j=1}^{k_n} s_j r_{nj},$$

for some $s_k \in R, 1 \leq j \leq k_n$. Put

$$g := \sum_{j=1}^{k_n} s_j f_{nj}.$$

Then $\text{lead}(g) = s$, $g \in I$ and $\deg g = n$. Thus $f - g \in J$ and $\deg(f - g) < n$. By induction on $\deg f$, $f - g \in I$ and so $f = (f - g) + g \in I$. This shows that $I = J$ and so J is a finitely generated ideal in $R[x]$. □

Corollary 7.5.12. *(a) A is a Noetherian ring.*

(b) Let U be an affine $\mathbb{K}$-variety. Then $U = V(S)$ for some finite subset S of A.

(c) Let $\mathcal{V}$ be a non-empty set of $\mathbb{K}$ varieties in $\mathbb{F}^D$. Then $\mathcal{V}$ has a minimal element.

Proof. (a) Clearly $\mathbb{K}$ is Noetherian, so (a) follows Hilbert's Basis Theorem.

(b) By (a) $J(U)$ is finitely generated as an ideal. So $J(U) = AS$ for some finite subset S of A. Thus by 7.5.3

$$U = V(J(U)) = V(AS) = V(S).$$

(c) Let $\mathcal{I} = \{J(U) \mid U \in \mathcal{V}\}$. Then by (a) and 3.3.17 $\mathcal{I}$ has a maximal element say $J(U_0)$ for some $U_0 \in \mathcal{V}$. Let U be in $\mathcal{V}$ with $U \subseteq U_0$. Then $J(U_0) \subseteq J(U)$ and by maximality of $J(U_0)$, $J(U_0) = J(U)$. Thus

$$U = V(J(U)) = V(J(U_0)) = U_0$$

and so U_0 is a minimal element of $\mathcal{U}$. □

Lemma 7.5.13. *Let S and T be ideals in A. Then*

$$V(S) \cup V(T) = V(S \cap T) == V(ST)$$

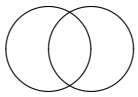
Proof. Clearly $V(S) \cap V(T) \subseteq V(S \cap T)$. Since S and T are ideals, $ST \subseteq S\alpha' T$ and so $V(S \cap T) \subseteq V(ST)$. So it remains to show that $V(ST) \subseteq V(S) \cup V(T)$. Let $u \in F^{\mathrm{D}}$ with $v \notin V(S) \cup V(T)$. Then there exists $s \in S$ and $t \in T$ with $s(u) \neq 0 \neq t(u)$. Then $(st)(u) = s(u)t(u) \neq 0$ and since $st \in ST$, $u \notin V(ST)$. So $V(ST) \subseteq V(S) \cup V(T)$. □

Definition 7.5.14. *An affine $\mathbb{K}$-variety U in $\mathbb{F}^D$ is called $\mathbb{K}$-irreducible provided that:*
Whenever U_1 and U_2 are $\mathbb{K}$-varieties in $\mathbb{F}^D$ with $U = U_1 \cup U_2$, then $U = U_1$ or $U = U_2$

Example 7.5.15. (1) Let $\mathbb{F} = \mathbb{C}$ and $U = V(x^2 - 2y^2)$. If $\mathbb{K} = \mathbb{R}$, then $(x^2 - 2y^2) = (x + \sqrt{2}y)(x - sqrt2y)$ and so by 7.5.13 $U = V(x + \sqrt{2}y) \cup V(x - sqrt2y)$ and so U is not an irreducible $\mathbb{R}$-variety.

But it can be shown that U is an irreducible $\mathbb{Q}$-variety.

(2) Let $\mathbb{F} = \mathbb{C}$ and $\mathbb{K} = \mathbb{Q}$. Let $U = V(x^2 + y^2 - 4)((x-1)^2 + y^2 - 4))$:



Then U is the union of two irreducible subvarieties namely the circles $V(x^2+y^2-4)$ and $V(x-1)^2+y^2-4)$. But U cannot be written has the disjoint union of two subvarieties.

Lemma 7.5.16. *Let U be a affine $\mathbb{K}$-variety in $\mathbb{F}^D$. Then U is $\mathbb{K}$- irreducible if and only if $J(U)$ is a prime ideal in A.*

Proof. Suppose first that $J(U)$ is a prime ideal in A and let U_1, U_2 be affine $\mathbb{K}$-varieties with $U = U_1 \cup U_2$. Then $U = U_1 \cup U_2 \subseteq V(J(U_1)J(U_2))$ and so $J(U_1)J(U_2) \subseteq J(U)$. Since $J(U)$ is a prime ideal we conclude $J(U_i) \subseteq J(U)$ for some $i \in \{1, 2\}$. Hence $U \subseteq V(J(U_i)) = U_i \subseteq U$ and so $U = U_i$.

Suppose next that U is irreducible and let J_1 and J_2 be ideal in A with $J_1 J_2 \subseteq J(U)$. We need to show that $J_k \subseteq J(U)$ for some i. Replacing J_i be $J_i + J(U)$ we may assume that $J(U) \subseteq J_i$ for $i = 1$ and 2. Then $V(J_i) \subseteq U_i$. By 7.5.13

$$V(J_1) \cup V(J_2) = V(J_1 \cap J_2) = V(J_1 J_2)$$

Since $J_1 J_2 \leq J(U) \leq J_1 \cap J_2$ we have

$$V(J_1 \cap J_2) \subseteq U = J(V(U)) \subseteq V(J_1 J_2)$$

Thus $U = V(J_1) \cup V(J_2)$ and since U is irreducible, $V(J_k) = U$ for some k. Thus $J_k \leq J(U)$ and $J(U)$ is a prime ideal in A. □

Chapter 8

Simple Rings and Simple Modules

8.1 Jacobson's Density Theorem

Definition 8.1.1. *let R be a ring and M an R-module. M is called a* simple R-module *if $RM \neq 0$ and M has no proper R-submodules.*

Example 8.1.2. 1. Let I is be left ideal in R, then R/I is simple if and only if I is a maximal left ideal.

2. Let D be a division ring and V is an D-module. We will show that V is a simple $\mathrm{End}_D(V)$ module. For this we first show that for each $u, v \in V$ with $u \neq 0_V$ there exists $\alpha \in \mathrm{End}_D(V)$ with $\alpha(u) = v$. For this let $\mathcal{B}$ be a basis for V with $u \in \mathcal{B}$. Then there exists a unique D-linear map $V \to W$ with $\alpha(w) = v$ for all $w \in \mathcal{B}$. In particular, $\alpha(u) = v$.

   Now let U be any non-zero $\mathrm{End}_D(V)$-submodule of B. Let $u \in U^\sharp$ and $v \in V$. Then by the above there exists $\alpha \in \mathrm{End}_D(V)$ with $\alpha(u) = v$. Thus $v \in U$ and $U = V$.

Lemma 8.1.3 (Schur's Lemma)**.** *Let M, N be simple R-modules and $f \in \mathrm{Hom}_R(M, N)$. If $f \neq 0$, then f is R-isomorphism. In particular,* $\mathrm{End}_R(M)$ *is a division ring.*

Proof. Since $f \neq 0$, $\ker f \neq M$. Also $\ker f$ is an R-submodule and so $\ker f = 0$ and f is 1-1. Similarly, $\mathrm{Im}\, f \neq 0$, $\mathrm{Im}\, f = N$ and so f is onto. So f is a bijection and has an inverse f^{-1}. An easy computation shows that $f^{-1} \in \mathrm{Hom}_R(N, M))$. Choosing $N = M$ we see that $\mathrm{End}_R(M)$ is a division ring. □

Definition 8.1.4. *Let R be a ring and M be an R-module.*

(a) Let $N \subseteq M$. N is called R-closed *in M if $N = \mathrm{Ann}_M(\mathrm{Ann}_R(N))$.*

(b) Let $I \subseteq R$. I is called M-closed *in R if $I = \mathrm{Ann}_R(\mathrm{Ann}_M(I))$.*

Lemma 8.1.5. *Let R be a ring and M an R module. Let $U \subseteq \tilde{U} \subseteq M$ and $S \subseteq \tilde{S} \subseteq R$.*

(a) $\mathrm{Ann}_R(\tilde{U}) \subseteq \mathrm{Ann}_R(U)$.

(b) $\mathrm{Ann}_M(\tilde{S}) \subseteq \mathrm{Ann}_M(S)$.

(c) $U \subseteq \mathrm{Ann}_M(\mathrm{Ann}_R(U))$.

(d) $S \subseteq \mathrm{Ann}_R(\mathrm{Ann}_M(S))$.

(e) U *is* R*-closed in* M *if and only if* $U = \mathrm{Ann}_M(S)$ *for some* $S \subseteq M$.

(f) S *is* M*-closed in* R *if and only if* $S = \mathrm{Ann}_R(U)$ *for some* $U \subseteq M$.

Proof. Readily verified. □

Lemma 8.1.6. . *Let* M *be a simple* R*-module,* V *a* R*-closed subset of* M *and* $w \in M \setminus V$. *Put* $I = \mathrm{Ann}_R(A)$. *Then* $M = Iw$ *and the map* $\beta : I/\mathrm{Ann}_I(w) \to M, i + \mathrm{Ann}R_J(w) \to jw$ *is a well defined* R*-isomorphism.*

Proof. Since V is closed, $V = \mathrm{Ann}_R(V)$ and so $Iw \neq 0$. By 4.1.11 I is a left ideal in R and so $R(Iw) = (RI) \subseteq Iw$. Thus Iw is an R-submodule of M. Since M is simple, $M = Iw$. Define $\phi : I \to M, i \to iw$. Then ϕ is R-linear, onto and $\ker \phi = \mathrm{Ann}_I(w)$ $J/\mathrm{A}_J(m) \cong M$. So the second statement follows from the First Isomorphism Theorem of R-modules. □

Lemma 8.1.7. *Let* M *be simple* R*-module and* $\mathbb{D} = \mathrm{End}_R(M)$. *Let* $V \leq W$ *be* $\mathbb{D}$ *submodules of* M *with* $\dim_{\mathbb{D}}(W/V)$ *finite. If* V *is closed in* M *so is* W. *In particular, all finite dimensional* $\mathbb{D}$ *subspaces of* M *are closed.*

Proof. By induction on $\dim_{\mathbb{D}} W/V$ we may assume that $\dim \mathbb{D}W/V = 1$. Let $w \in W \setminus V$. Then $W = V + \mathbb{D}w$. Put $I = \mathrm{Ann}_R(V)$ and $J = \mathrm{Ann}_I(w)$. We will show that $W = \mathrm{A}_R(J)$. So let $m \in \mathrm{A}_M(J)$. Then $J \subseteq \mathrm{A}_I(m)$ and hence the map $\alpha : I/J \to M, i + J \to im$ is well defined and R-linear. By 8.1.6 the map $\beta : I/J \to M, i + J \to iw$ is an R-isomorphism. Put $\delta = \alpha\beta^{-1}$. Then $\delta : M \to M$ is R-linear and $\delta(iw) = im$ for all $i \in I$. Hence $\delta \in \mathbb{D}$ and

$$i(m - \delta(w)) = im - i\delta(w)) = im - \delta(iw) = im - im = 0$$

for all $i \in I$. Since V is closed, $V = \mathrm{Ann}_M(I)$ and so $\delta(w) - m \in V$. Thus $m \in \delta(w) + V \leq W$ and $\mathrm{Ann}_M(J) \subseteq W$

Let $j \in J$, $\delta \in \mathrm{D}$ and $v \in V$. Then $i(v + \delta(w) = iv + \delta(iw) = 0 + \delta(0) = 0$ and so $W \subseteq \mathrm{Ann}_M(J)$. Thus $W = \mathrm{Ann}_W(J)$ and so by 8.1.5(e), W is R-closed in M.

Since M is a simple R-module, $RM \neq 0$, $\mathrm{Ann}_M(R) \neq M$ and $\mathrm{Ann}_M(R) = 0$. So 0 is a R-closed in M. Hence the first statement of the lemma implies the second. □

Definition 8.1.8. *Let* M *be an* R*-module and* $\mathbb{D} \leq \mathrm{End}_R(M)$ *a division ring. Then we say that* R *acts densely on* M *with respect to* $\mathbb{D}$ *if for each* $\mathbb{D}$*-linearly independent tuple* $(m_i)_{i=1}^n \in M^n$ *and each* $(w_i)_{i=1}^n \in M^n$, *there exists* $r \in R$ *with* $rm_i = w_i$ *for all* $1 \leq i \leq n$.

Theorem 8.1.9 (Jacobson's Density Theorem)**.** *Let* R *be a ring and* M *a simple* R*-module. Put* $\mathbb{D} := \mathrm{End}_R(M)$, *then* R *acts densely on* M *with respect to* $\mathbb{D}$.

Proof. Let $(m_i)_{i=1}^n \in M^n$ be $\mathbb{D}$-linear independent and $(w_i)_{i=1}^n \in M^n$. By induction on n we will show that there exists $r \in R$ with $rm_i = w_i$ for all $1 \leq i \leq n$. For $n = 0$, there is nothing to prove. By induction there exists $s \in R$ with $sm_i = w_i$ for all $1 \leq i < n$. Let $V = \sum_{i=1}^{n-1} \mathbb{D}m_i$. Then by 8.1.7 V is closed and so by 8.1.6 there exists $t \in \mathrm{A}_R(V)$ with $tm_n = w_n - sm_n$. Put $r = s + t$. For $1 \leq i < n$, $tm_i = 0 =$ and so $rm_i = sm_i = w_i$. Also $rm_i = sm_i + tm_i = sm_n + (w_n - sm_n) = w_n$ and the theorem is proved. □

Corollary 8.1.10. *Let M be a simple R-module, $\mathbb{D} = \mathrm{End}_R(M)$ and W a finite dimensional $\mathbb{D}$-submodule of M. Put $N_R(W) = \{r \in R \mid rW \subseteq W\}$. Then $\mathrm{N}_R(W)$ is a subring of W, W is a $N_R(W)$-submodule of M, $\mathrm{Ann}_R(W)$ is an ideal in $\mathrm{N}_R(W)$ and, if $R|_W$ denotes the image of $\mathrm{N}_R(W)$ in $\mathrm{End}(W)$, then*

$$\mathrm{N}_R(W)/\mathrm{Ann}_R(W) \cong R|_W = \mathrm{End}_{\mathrm{D}}(W).$$

Proof. Let $r, s \in \mathrm{N}_R(W)$ and $w \in W$. Then $(r + s)w = rw + sw \in W$ and $(rs)w = r(sw) \in W$. Thus $\mathrm{N}_R(W)$ is a subring of R. Note that W is an $N_R(W)$-module and $\mathrm{Ann}_R(W) = \mathrm{Ann}_{\mathrm{N}_R(W)}W)$. So by 4.1.11(b) $\mathrm{Ann}_R(W)$ is an ideal in $\mathrm{N}_R(W)$. Clearly $R|_W$ is contained in $\mathrm{End}_{\mathrm{D}}(W)$. Let $\phi \in \mathrm{End}_{\mathrm{D}}(W)$ and choose a basis $(v_i, 1 \leq i \leq n)$ for W over $\mathbb{D}$. By 8.1.9 there exists $r \in R$ with $rv_i = \phi v_i$ for all $1 \leq i \leq n$. Then $rW \leq W$ and so $r \in \mathrm{N}_R(W)$. The image of r in $\mathrm{End}(W)$ is ϕ. Thus $\phi \in R|_W$ and so $R|_W = \mathrm{End}_{\mathbb{D}}(W)$. □

8.2 Semisimple Modules

Definition 8.2.1. *Let R be a ring and M an R-module. M is a* semisimple *if M is the direct sum of simple R-submodules.*

Lemma 8.2.2. *Let G be a group and $(G_i, i \in I)$ a family of G. Let $(I_j)_{j \in J}$ be a partition of I and put $H_j = \langle i \in I_j \rangle$. Then*

$$G = \bigoplus_{i \in I} G_i$$

if and only if

$$G = \bigoplus_{j \in J} H_j \text{ and for all } j \in J, H_j = \bigoplus_{i \in I_j} G_i.$$

Proof. Note first that $\langle G_i \mid i \in I \rangle = \langle \langle G_i, i \in I_j \rangle \mid j \in I \rangle = \langle H_j, h \in J \rangle$. So in both cases $G = \langle G_i, i \in I \rangle = \langle H_j, j \in J \rangle$.

Suppose first that $G = \bigoplus_{i \in I} G_i$. Then $G_i \trianglelefteq G$ for all $i \in i$ and so $H_j \trianglelefteq G$ and $G_i \trianglelefteq H_j$ for all $j \in I$, $i \in I_j$.

$$G_i \cap \langle G_t \mid i \neq t \in I_j \rangle \leq G_i \cap \langle G_t \mid i \neq t \in I \rangle = \{e\}$$

and so $H_j = \bigoplus_{i\in I_j} G_i$. Let $h \in H_j \cap \langle H_k \mid j \neq k \in J\rangle$. Since $h \in H_j$ and $H_j = \bigoplus_{i\in I_j} G_i$ there exists $h_i \in G)i$ for $i \in I_j$ with $h = \prod_{i\in I_j} h_i$. Since $h \in \langle H_k \mid j \neq k \in J\rangle \leq \langle G_k \mid i \neq k \in I\rangle$ we conclude

Then

$$h_i = h \prod_{i\neq k\in I_j} h_k^{-1} \in G_i \cap \langle G_k \mid i \neq k \in I\rangle = \{e\}$$

Thus $h_i = e$ for all $i \in H_i$ and so $h = e$ and $G = \bigoplus_{j\in I} H_j$.

Suppose next that $G = \bigoplus_{j\in J} H_j$ and for all $j \in J, H_j = \bigoplus_{i\in I_j} G_i$. Then $G_i \trianglelefteq H_j$ for all $j \in I$, $i \in I_j$. Since $[H_j, H_k] = 1$ for $j \neq k \in I$ and $G = \langle H_k \mid k \in J\rangle$ we conclude that $G_i \trianglelefteq G$. Let $g \in G_i \cap \langle G_k \mid i \neq k \in I\rangle$. Put $X = \langle H_k \mid j \neq k \in J$ and $Y := \langle G_k \mid i \neq k \in I_j\rangle$. Then $Y \leq H_i$ and so $[X, Y] = 1$ and $\langle G_k \mid i \neq k \in I\rangle = XY$. Note that $g \in H_j \cap XY = X(H_j \cap Y) = X$ and so $g \in G_i \cap X = \{e\}$. Thus $G = \bigoplus_{i\in I} G_i$. □

Lemma 8.2.3. *Let $\mathcal{S}$ a set of simple R-submodules of the R-module M. Also let N be a R-submodule of M and suppose that $M = \sum \mathcal{S}$.*

(a) There exists a subset $\mathcal{M}$ of $\mathcal{S}$ with $M = N \oplus \bigoplus \mathcal{M}$.

(b) $M = \bigoplus \mathcal{T}$ for some $\mathcal{T} \subseteq \mathcal{S}$.

(c) $M/N \cong \bigoplus \mathcal{T}$ for some subset $\mathcal{T}$ of $\mathcal{S}$.

(d) M/N is semisimple.

(e) $N \cong \bigoplus \mathcal{T}$ for some subset $\mathcal{T}$ of $\mathcal{S}$.

(f) N is semisimple.

(g) If N is a simple submodule of M, then $N \cong S$ for some $S \in \mathcal{S}$.

(h) Suppose N is a maximal N-submodule of M, then then $M/X \cong S$ for some $S \in \mathcal{S}$.

Proof. (a) Let $\mathcal{M}$ consists of all subsets $\mathcal{T}$ of $\mathcal{S}$ with $N \cap \sum \mathcal{T} = 0$ and $\sum \mathcal{T} = \bigoplus \mathcal{T}$. Since $\emptyset \in \mathcal{M}$, $\mathcal{M} \neq \emptyset$. Order $\mathcal{M}$ by inclusion and let $\mathcal{C}$ be a chain in $\mathcal{M}$. Let $\mathcal{D} = \bigcup \mathcal{C}$. Let $m \in M \cap \sum \mathcal{D}$. Then there exists $D_i \in \mathcal{D}$, $1 \leq i \leq n$ and $d_i \in D_i$ with $m = \sum_{i=1}^n d_i$. For each D_i there exists $C_i \in \mathcal{C}$ with $D_i \in C_i$. As $\mathcal{C}$ is a chain we may assume that $C_1 \subseteq C_2 \subseteq \ldots C_n$. Then $D_i \in C_n$ for all $1 \leq i \leq n$ and so $m \in N \cap \sum C_n = 0$. A similar argument shows that $\sum \mathcal{D} = \bigoplus \mathcal{D}$.

Therefore $N \cap \sum \mathcal{D} = 0$ and $\mathcal{D} \in \mathcal{M}$. So we can apply Zorn's lemma to obtain a maximal element $\mathcal{T}$ in $\mathcal{M}$. Put $W = \sum \mathcal{T}$. Suppose that $M \neq N + W$. Then there exists $S \in \mathcal{S}$ with $S \not\leq N + W$. Since S is simple, $(N + W) \cap S = 0$. So $(N + W) \cap (S + W) = W + ((N + W) \cap S) = W$ and so $N \cap (S + W) \leq N \cap W = 0$. Also $W \cap S = 0$ implies that $\sum \mathcal{S} \cup \{M\} = W \oplus S = \bigoplus \mathcal{M} \cup \{S\}$. Thus $\mathcal{T} \cup \{S\}$ is linearly independent and so $\mathcal{T} \cup \{S\} \in \mathcal{M}$, a contradiction to the maximality of $\mathcal{M}$.

Thus $M = N \oplus W$. So (a) holds.

(b) follows from (a) applied with $N = 0$. (c) follows from (a).

(d) follows from (c).

Note that $N \cong M/W$. So (e) follows from (c) applied to W in place of N.

(f) follows from (e).

(e) Suppose N is simple. Then the set $\mathcal{T}$ from (e) only contains one element, say S. So $N \cong S$ and (g) is proved.

Suppose that N is a maximal R-submodule of M. Then the set $\mathcal{T}$ from (b) only contains one elements, sat S. Thus $M/N \cong S$. □

Corollary 8.2.4. *Let R be a ring, M a semisimple R-module and A and B R-submodules of M with $A \leq B$. Then A/B is semisimple.*

Proof. 8.2.3(f) implies that B is semisimple. Then 8.2.3(d) applied to (A, B) in place of (N, M) shows that B/A is semisimple. □

Lemma 8.2.5. *Let M a semisimple R-module and N an R-submodule of M with $N \neq M$. Let $\mathcal{M}$ be the set of maximal R-submodules of M containing N. Then $\bigcap \mathcal{M} = N$.*

Proof. By 8.2.4 M/N is a semisimple R-module. Thus replacing M by M/N we may assyme that $N = 0$. Let $\mathcal{S}$ be a set of simple R-submodules of M with $M = \bigoplus \mathcal{S}$. For $S \in \mathcal{S}$, put $S^* = \sum_{S \neq T \in \mathcal{S}} T$. Then $M/S^* \cong S$ and so S^* is a maximal R-submodule of S. Then $0 \leq \bigcap \mathcal{M} \subseteq \bigcap_{S \in \mathcal{S}} S^* = 0$ and so $\bigcap \mathcal{M} = 0 = N$. □

8.3 Simple Rings

Proposition 8.3.1. *Let R be a simple ring with identity. Then there exists a simple R-module M. Moreover, if M is a simple R-module and $\mathbb{D} = \mathrm{End}_R(M)$, then R is isomorphic to a dense subring of $\mathrm{End}_{\mathbb{D}}(M)$.*

Proof. let $\mathcal{C}$ be non-empty chain of proper left ideal in R. Then $1_R \notin \bigcup \mathcal{C}$ and so $\mathcal{C}$ is a proper left ideal in R. Hence by By Zorn's Lemma, R has a maximal left ideal I. We claim that $M := R/I$ is a simple R-module. Indeed since I is maximal, M has no proper R-submodules and since R is simple, $R^2 \neq 0$, $R^2 = R$ and $RM = M \neq \{0\}$.

Now let M be any simple R-module. Since $RM \neq 0$, $\mathrm{Ann}_R(M) \neq R$. Since M is simple and $\mathrm{Ann}_R(M)$ is an ideal in M, $\mathrm{Ann}_R(M) = 0$. Thus $R \cong R|_M$ and by 8.1.9, R and so also $R|_M$ is dense on M. □

Proposition 8.3.2. *Let M be faithful, simple R-module and put $\mathbb{D} = \mathrm{End}_R(M)$. Suppose that $n := \dim_{\mathbb{D}} M$ is finite.*

(a) $R \cong R|_M = \mathrm{End}_{\mathbb{D}}(M)$.

(b) $R \cong M^n$ *as a left R-module.*

(c) Let I be a maximal left ideal in R. Then $I = \mathrm{Ann}_R(m)$ for some $m \in M^\sharp$ and $R/I \cong M$

(d) Each left ideal in R is closed in R with respect to M.

(e) The map $I \to \mathrm{Ann}_R(I)$ is a bijection between the left ideals in R and the $\mathbb{D}$-subspaces in M. Its inverse is $M \to \mathrm{Ann}_M(I)$.

(f) Each simple R-module is isomorphic to M.

(g) R is a simple ring.

Proof. (a) Note that $\mathrm{N}_R(M) = R$ and so (a) follows from 8.1.10.

(b) Let $\mathcal{B}$ be a basis for M over $\mathbb{D}$. Define

$$\gamma : R \to M^{\mathcal{B}}, r \to (rb)_{b\in\mathcal{B}}$$

Then γ is R-linear and by the density theorem γ is onto. Let $r \in \ker\phi$. Then $rb = 0$ for all $b \in \mathcal{B}$. Thus $\mathrm{Ann}_M(r)$ is a $\mathbb{D}$-submodule containing $\mathcal{B}$ and so $\mathrm{Ann}_M(r) = M$. Since M is a faithful R-module, $r = 0$ and so γ is 1-1. Thus γ is an R-isomorphism.

(c) By (b) 8.2.3(h), $R/I \cong M$. Note that by (a), R has an identity. Let $\phi : R/I \to M$ be an R-isomorphism and put $m = \phi(1_R + I)$. Then

$$\mathrm{Ann}_R(m) = \mathrm{Ann}_R(1_R + I/I) = \{r \in R \mid r(1_R + I) = 0_{R/I}\} = \{r \in R \mid r_I = I\} = I.$$

(d) Let I be an left ideal in R and $\mathcal{M}$ the set of maximal ideals in R containing I. Since R is a semisimple R-module 8.2.5 implies that $\bigcap\mathcal{M} = I$. By (c), for each $J \in \mathcal{M}$ there exists $m_J \in M$ with $J = \mathrm{Ann}_R(m_J)$. Put $N = \{m_J \mid J \in \mathcal{M}\}$. Then

$$\mathrm{Ann}_R(N) = \bigcap_{J\in\mathcal{M}} \mathrm{Ann}_R(m_J) = \bigcap_{J\in\mathcal{M}} J = I.$$

So I is closed in R with respect to M.

(e) Let I be a left ideal in R. Then $\mathrm{Ann}_M(R)$ is a $\mathbb{D}$-submodule of M and by (d), $I = \mathrm{Ann}_R(\mathrm{Ann}_M(I))$.

Let N be a $\mathbb{D}$-submodule of M. Then $\mathrm{Ann}_R(N)$ is a left ideal in R. Since M and so N is finite dimensional over $\mathbb{D}$ we conclude from 8.1.7 that N is closed in M with respect to R and so $\mathrm{Ann}_M(\mathrm{Ann}_R(N)) = N$.

(f) Let W be a simple R-module and $w \in W^\sharp$. Then $R/\mathrm{Ann}_R(w) \cong Rw = W$. Hence $\mathrm{Ann}_R(w)$ is maximal left ideal in R and so and (c) $W \cong R/\mathrm{Ann}_R(w) \cong M$.

(g) Let I be an ideal in R. Then $\mathrm{Ann}_M(I)$ is an R-submodule of M. Since M is simple, $\mathrm{Ann}_M(I) = 0$ or $\mathrm{Ann}_M(I) = M$. By (e), $I = \mathrm{Ann}_R(\mathrm{Ann}_M(I)$ and so $I = \mathrm{Ann}_R(0) = R$ or $I = \mathrm{Ann}_R(M) = 0$. Since R has an identity, $R^2 \neq 0$ and so R is simple. □

Definition 8.3.3. *A ring R is called* Artinian *for every descending chain*

$$I_1 \geq I_2 \geq I_3 \geq \dots I_k \geq I_{k+1} \geq \dots$$

there exists $n \in \mathbb{Z}^+$ with $I_k = I_n$ for all $k \leq n$.
In other words, R does not have an infinite strictly descending chain of left ideals.

Lemma 8.3.4. *Let R be an Artinian ring and M a simple R-module. Then M is finite dimensional over $\mathbb{D} = \mathrm{End}_R(M)$.*

Proof. Suppose that $\dim_{\mathbb{D}} M = \infty$. Then there exists an infinite strictly ascending series

$$M_1 < M_2 < M_3 < \dots$$

of finite dimensional $\mathbb{D}$-subspaces. By 8.1.7 each M_i is closed. Thus

$$\mathrm{Ann}_R(M_1) > \mathrm{Ann}_R(M_2) > \mathrm{Ann}_R(M_3) > \dots$$

is a strictly descending chain of left ideals in R, contradicting the definition of an Artinian ring. □

Theorem 8.3.5. *Let R be a simple Artinian ring with identity. Then there exists a simple R-module M, M is unqiue up to isomorphism and if $\mathrm{D} := \mathrm{End}_R(M)$, then $\dim_{\mathrm{D}} M$ is finite dimensional and $R \cong \mathrm{End}_{\mathrm{D}}(M)$.*

Proof. By 8.3.1 R has a simple module M. By 8.3.4 $\dim_{\mathbb{D}}(M)$ is finite. Thus by 8.3.2(g), M is unique up to isomorphism and by 8.3.2(a), $R \cong \mathrm{End}_{\mathbb{D}}(M)$. □

Appendix A

Zorn's Lemma

This chapter is devoted to prove Zorn's lemma: Let M be a nonempty partially ordered set in which every chain has an upper bound. Then M has a maximal element.

To be able to do this we assume throughout this lecture notes that the *axiom of choice* holds. The axiom of choice states that if $(A_i, i \in I)$ is a nonempty family of nonempty sets then also $\prod_{i\in I} A_i$ is not empty. That is there exists a function $f : I \to \bigcup_{i\in I} A_i$ with $f(i) \in A_i$. Naively this just means that we can pick an element from each of the sets A_i.

Definition A.1. *A* partially ordered set *is a set M together with a reflexive, anti-symmetric and transitive relation " $\leq$ ". That is for all $a, b, c \in M$*

(a) $a \leq a$ (reflexive)

(b) $a \leq b$ *and* $b \leq a \Longrightarrow a = a$ (anti-symmetric)

(c) $a \leq b$ *and* $b \leq c \Longrightarrow a \leq c$ (transitive)

Definition A.2. *Let $(M, \leq)$ be a partially ordered set, $a, b \in M$ and $C \subseteq M$.*

(a) a are called comparable *if $a \leq b$ or $b \leq a$.*

(b) $(M, \leq)$ is called linearly ordered *if any two elements are comparable.*

(c) C is called a chain *if any two elements in C are comparable.*

(d) An upper bound *m for C is an element m in M so that $c \leq m$ for all $c \in C$.*

(e) An least upper bound *for C is an upper bound m so that $m \leq d$ for all upper bounds d of C.*

(f) An element $m \in C$ is called a maximal element *of C if $c = m$ for all $c \in C$ with $m \leq c$.*

(g) An element $m \in C$ is called a minimal element *of C if $c = m$ for all $c \in C$ with $c \leq m$.*

(h) A function $f : M \to M$ is called increasing if $a \leq f(a)$ for all $a \in M$.

As the main steps toward our proof of Zorn's lemma we show:

Lemma A.3. *Let M be a non-empty partially ordered set in which every non-empty chain has a least upper bound. Let $f : M \to M$ be an increasing function. Then $f(m_0) = m_0$ for some $m_0 \in M$.*

Proof. To use that M is not empty pick $a \in M$. Let $B := \{m \in M \mid a \leq m\}$. If $b \in B$, then $a \leq b \leq f(b)$ and so $f(B) \in B$. Note also that the least upper bound of any non-empty chain in B is contained in B. So replacing M by B we may assume that

1°. *$a \leq m$ for all $m \in M$.*

We aim is to find a subset of M which is a chain and whose upper bound necessarily a fixed-point for f. We will not be able to reach both these properties in one shot and we first focus on the second part. For this we define a subset A of M to be closed if:

(Cl i) $a \in A$

(Cl ii) $f(b) \in A$ for all $b \in A$.

(Cl iii) If C is a non-empty chain in A then its least upper bound is in A.

Since M is closed, there exists at least one closed subset of M.

2°. *Let D be closed chain and d its upper bound in M. Then $f(d) = d$.*

By (i), D is not empty and so has a least upper bound d. By (iii), $d \in D$ and by (ii), $f(d) \in D$. Since d is a upper bound for D, $f(d) \leq d$ and since f is increasing, $d \leq f(d)$. Since $\leq$ is antisymmetric $f(d) = d$.

In view of(2°) we just have to find a closed chain in M. There is an obvious candidate: It is immediate from the three conditions of closed that intersections of closed sets are closed. So we define A be the intersection of all the closed sets.

Call $e \in A$ to extreme if

$$(Ex) \qquad f(b) \leq e \text{ for all } b \in A \text{ with } b < e$$

Note that a is extreme, so the set E of extreme elements in A is not empty.

Here comes the main point of the proof:

3°. *Let e be extreme and $b \in A$. Then $b \leq e$ or $f(e) \leq b$. In particular, e and b are comparable.*

To prove (3°) put

$$A_e = \{b \in A \mid b \leq e \text{ or } f(e) \leq b\}$$

We need to show that $A_e = A$. Since A is the unique minimal closed set this amounts to proving that A_e is closed.

Clearly $a \in A_e$. Let $b \in A_e$. If $b < e$, then as e is extreme, $f(b) \leq e$ and so $f(b) \in A_e$. If $b = e$, then $f(e) = f(b) \leq f(b)$ and again $f(b) \in A_e$. If $f(e) \leq b$, then $f(e) \leq b \leq f(b)$ and $f(e) \leq f(b)$ by transitivity. So in all cases $f(b) \in A_e$.

let D be a non-empty chain in A_e and m its least upper bound. If $d \leq e$ for all d in D, then e is an upper bound for D and so $m \leq e$ and $m \in A_e$. So suppose that $d \not\leq e$ for some $d \in D$. As $d \in A_e$, $f(e) \leq d \leq m$ and again $m \in A_e$.

We proved that A_e is closed. Thus $A_e = A$ and (3°) holds.

4°. *E is closed*

Indeed, $a \in E$. Let $e \in E$. To show that $f(e)$ is extreme let $b \in A$ with $b < f(e)$. By (3°) $b \leq e$ or $f(e) \leq b$. The latter case is impossible by anti-symmetry. If $b < e$, then since e is extreme, $f(b) \leq e \leq f(e)$. If $e = b$, then $f(b) = f(e) \leq f(e)$. So $f(e)$ is extreme.

Let D be a non-empty chain in E and m its least upper bound. We need to show that m is extreme. Let $b \in A$ with $b < m$. As m is a least upper bound of D, b is not an upper bound and there exists $e \in D$ with $e \not\leq b$. By (3°), e and b are comparable and so $b < e$. As e is extreme, $f(b) \leq e \leq m$ and so m is extreme. Thus E is closed.

As E is closed and $E \subseteq A$, $A = E$. Hence by (4°) , any two elements in A are comparable. So A is a closed chain and by (2°), the lemma holds. □

As an immediate consequence we get:

Corollary A.4. *Let M be a non-empty partially ordered set in which every non-empty chain has a least upper bound. Then M has a maximal element.*

Proof. Suppose not. Then for each $m \in M$ there exists $f(m)$ with $m < f(m)$.(The axiom of choice is used here). But then f is a strictly increasing function, a contradiction to A.3. □

Lemma A.5. *Let M be any partial ordered set. Order the set of chains in M by inclusion. Then M has a maximal chain.*

Proof. Let $\mathcal{M}$ be the set of chains in M. The union of a chain in $\mathcal{M}$ is clearly a chain in M and is an least upper bound for the chain. Thus A.4 applied to $\mathcal{M}$ yields a maximal member of $\mathcal{M}$. That is a maximal chain in M. □

Theorem A.6 (Zorn's Lemma)**.** *Let M be a nonempty partially ordered set in which every chain has an upper bound. Then M has a maximal element.*

Proof. By A.5 there exists a maximal chain C in M. By assumption C has an upper bound m. Let $l \in M$ with t $m \leq l$. Then $C \cup \{m, l\}$ is a chain in M and the maximality of C implies $l \in C$. Thus $l \leq m$, $m = l$ and m is maximal element. □

Definition A.7 (Structures)**.** *Let S be a set. A structure $\mathcal{G}$ on S consists of sets I and J, a family of sets $(X_j, j \in J)$, a subset K of J with $X_k = S$ for all $k \in K$ and for each $i \in I$ a subset J_i of J and a function*

$$f_i : \bigtimes_{j \in J_i} X_j \to S.$$

A subset T of S is called $\mathcal{G}$ invariant if

$$f_i((x_j)_{j \in J_i}) \in T$$

for all $i \in I$ and all $(x_j)_{j \in I_j} \in \bigtimes_{j \in J_i} X_j$ with $x_k \in T$ for all $k \in I_j \cap K$.

Here are a few examples: Let G be a group. Let $\mathcal{G}$ be the structure consisting of $I = \{1, 2, 3\}$, $J = \{0, 1, 2\}$, $X_0 = \{e_G\}$, $X_1 = X_2 = G$, $K = \{1, 2\}$
for $i = 1$, $J_1 = \{1, 2\}$ and

$$f_1 : G \times G \to G, (a, b) \to ab$$

for $i = 2$, $J_2 = \{1\}$ and

$$f_2 : G \to G, a \to a^{-1}$$

For $i = 3$ $J_3 = \{0\}$ and

$$f_3 : \{e_G\} \to G, e_G \to e_G$$

Then $H \subseteq G$ is $\mathcal{G}$-invariant if and only if

$$\begin{aligned} ab = f_1(a, b) \in T \quad & \text{for all } a, b \in H \\ a^{-1} = f_2(a) \in T \quad & \text{for all } a \in H \\ e_G = f_3(e_G) \in T & \end{aligned}$$

So a $\mathcal{G}$ invariant subset of G is just a subgroup of G.

As a second example consider a group G acting on a set S. Let $\mathcal{G}$ be the structure on S given by $I = \{1\}$, $K = \{2\}$, $X_1 = G, X_2 = S$ and $f_1 : G \times S \to S, (g, s) \to gs$. Let $T \subseteq S$. Then T is $\mathcal{G}$-invariant if and only if

$$gt = f_1(g, t) \in T \text{ for all } g \in G, t \in T$$

So T is $\mathcal{G}$-invariant if and only if T is G-invariant.

As last example consider a ring R. Let $\mathcal{G}$ be the structure on R given by
$I = \{1, 2, 3, 4\}$, $J = \{0, 1, 2, 3\}$, $K = \{2, 3\}$, $X_0 = \{0_R\}$, $X_1 = X_2 = X_3 = R$,
for $i = 1$, $J_1 = \{2, 3\}$ and

$$f_1 : R \times R \to R, (a, b) \to a + b$$

for $i = 2$, $J_2 = \{2\}$ and

$$f_2 : G \to G, a \to -a$$

For $i = 3$ $J_3 = \{0\}$ and

$$f_3 : \{0_R\} \to R, 0_R \to 0_R$$

For $i = 4$, $J_4 = \{1, 2\}$ and

$$f_4 : R \times R \to R, (a, b) \to ab$$

Let $I \subseteq R$. Then I is $\mathcal{G}$-invariant if and only if

$$\begin{aligned} a + b = f_1(a, b) \in I \quad & \text{for all } a, b \in I \\ -a = f_2(a) \in I \quad & \text{for all } a \in I \\ e_G = f_3(e_G) \in I & \\ ri = f_4(r, i) \in I \quad & \text{for all } r \in T, i \in I \end{aligned}$$

So the $\mathcal{G}$-invariant subsets of R are just the left ideals in R.

We will now prove:

1°. *Let $\mathcal{G}$ be a structure on the set S and $(T_q, q \in Q)$ a non-empty family of $\mathcal{G}$-invariant subsets of S. Then $\bigcap_{q\in Q} T_q$ is $\mathcal{G}$-invariant.*

Put $T = \bigcap_{q\in Q} T_q$. Let $i \in I$ and $(x_j)_{j\in J_i} \in \bigtimes_{j\in J_i} X_j$ with $x_k \in T$ for all $k \in J_i \cap K$. Note that $x_k \in T_q$ for all $q \in Q$ and $k \in J_i \cap K$. Since T_q is $\mathcal{G}$-invariant we get $f((x_j)_{j\in J_i}) \in T_q$ for all $q \in Q$ and so also $f((x_j)_{j\in J_i}) \in T$. Thus T is $\mathcal{G}$-invariant.

2°. *Let $\mathcal{G}$ be a structure on the set S and $(T_q, q \in Q)$ a non-empty family of $\mathcal{G}$-invariant subsets of S. Suppose that $\{T_q, g \in Q\}$ is linearly order by inclusion and that $J_i \cap K$ is finite for all $i \in I$. Then $\bigcup_{q\in Q} T_q$ is $\mathcal{G}$-invariant.*

Put $T = \bigcap_{q\in Q} T_q$. Fix $i \in I$ and $(x_j)_{j\in J_i} \in \bigtimes_{j\in J_i} X_j$ with $x_k \in T$ for all $k \in J_i \cap K$. Let $k \in J_i \cap K$. Then $x_k \in nT$ and so $x_k \in T_{q_k}$ for some $q_k \in Q$. Since $\{T_{q_k} \mid k \in J_i \cap K\}$ is finite and linearly ordered and since Q is non-empty there exists $q \in Q$ with $T_{q_k} \subseteq T_q$ for all $k \in I_j \cap K$. Since $x_k \in T_{q_k}$ this implies $x_k \in T_q$. Since T_q is $\mathcal{G}$ invariant, we get $f((x_j)_{j\in J_i}) \in T_q$ and so also $f((x_j)_{j\in J_i}) \in T$. Thus T is $\mathcal{G}$-invariant.

As an immediate consequence of (2°) and the above example we have

3°. *Let G be a group and $(G_q, q \in Q)$ a non-empty chain of subgroups of G. Then $\bigcup_{q\in Q} G_q$ is a subgroup of G.*

4°. *Let R be a ring and $(I_q, q \in Q$ a non-empty chain of ideals in R. Then $\bigcup_{q\in Q} I_q$ is an ideal in R.*

As an application of Zorn's lemma we prove the Well-Ordering Principal.

Definition A.8. *(a) A linearly ordered set M is called* well-ordered *if every non-empty subset of M has a minimal element.*

(b) We say that a set T can be well-ordered *if there exists a relation " $\leq$ " on T such that $(T,$ " $\leq$ ") is well ordered set.*

Example A.9. Let J be a non-empty well-ordered set with minimal element m and let $(I_j)_{j\in J}$ a family of non-empty well -ordered sets. Let m_j be the minimal element of I_j. Define

$$K = \{a \in \bigtimes_{j\in J} I_j \mid a_j = m \text{ for almost all } j \in J\}.$$

Order K as follows $a < b$ if $a \neq b$ and $a_j < b_j$ where $j = j(a,b) \in J$ is maximal with $a_j \neq b_j$ (Note here that there are only finitely many $j \in J$ with $a_j \neq b_j$, so there does exists a maximal such j.) We claim that this is a well ordering:

Suppose $a < b$ and $b < c$ and let $j = j(a,b)$ and $k = j(b,c)$. If $j \leq k$, then $a_l = b_l = c_l$ for all $l > k$ and $a_k \leq b_k < c_k$ so $a < c$. And if $j > k$, then $a_l = b_l = c_l$ for all $l > j$ and $a_j < b_j = c_j$ and again $a < c$. So K is linearly ordered. Let S be a non-empty subset of K. For $j \in J$ define

$$S(j) = \{t \in S \mid t_k \leq s_k \text{ for all } k \in J \text{ with } j \leq k\}$$

Define $U = U_S = \{u \in J \mid s_u = t_u \text{ for all } s,t \in S\}$. Suppose $U \neq J$ (that is $|S| > 1$). We will show there exists $j \in J \setminus U$ with $S(j) \neq \emptyset$. Suppose first that $J \setminus U$ has maximal element j. Choose $s \in S$ with s_j-minimal. Then $s \in S(j)$. Suppose next that $J \setminus U$ has no maximal element. Let $s \in S$. Since $\{k \in J \setminus U \mid s_k \neq m_k\}$ is finite and $J \setminus U$ has no maximal element there exists $j \in J \setminus U$ with $s_k = m_k$ for all $k \in J \setminus U$ with $k \geq j$. Then $s \in S(j)$.

If $|S| = 1$ define $j_S = m$ and if $|S| \neq 1$, pick $j_S \in J \setminus U$ minimal with $S^* := S(j_S) \neq \emptyset$. Put $j = j_S$. Let $s \in S^*$ and $t \notin S \setminus S^*$. Then $s_k \leq t_k$ for all $k \geq j$. If $s_k = t_k$ for all $k \geq j$, then $t \in S^*$. Hence there exists $k \geq j$ with $s_k < t_k$ and so $s < t$.

Let $k \in J$ with $k \geq j$. Then $s_k \leq t_k$ for all $s,t \in S^*$ and so $k \in U_{S^*}$. If $|S^*| \neq 1$, then $j_{s^*} \notin U_{S^*}$ and so $j_{S^*} < j = j_S$. If $|S^*| = 1$, then $j_{S^*} = m \leq j_S$. Put $S_0 = S$, $S_{i+1} = S_i^*$ and $j_i = j_{S_i}$. Then $j_0 \geq j_1 \geq j_2 \geq \ldots$ and since H is well ordered there exists $k \in \mathbb{N}$ with $j_k = j_{k+1}$. Then $S_{k+1} = S_k^*$ contains a unique element say r. Also for all $u \leq k$ and $s \in S_u \setminus S_{u+1}$ we have $r \in S_{u+1}$ and $r \leq s$. So r is the minimal element of S. Thus K is well-ordered.

Theorem A.10 (Well-ordering principal)**.** *Every set M can be well ordered.*

Proof. W be the set of well orderings $\alpha = (M_\alpha, \leq_\alpha)$ with $M_\alpha \subseteq M$. As the empty set can be well ordered, W is not empty. For $\alpha, \beta \in W$ define $\alpha \leq \beta$ if

< 1 $M_\alpha \subseteq M_\beta$

< 2 $\leq_\beta|_{M_\alpha}=\leq_\alpha$.

< 3 $a \leq_\beta b$ for all $a \in M_\alpha, b \in M_\beta \setminus M_\alpha$

It is easy to see that $\leq$ is a partial ordering on W. We would like to apply Zorn's lemma to obtain a member in W. For this let $\mathcal{A}$ be a chain in W. Put $M_* = \bigcup_{\alpha\in\mathcal{A}} M_\alpha$ and for $a, b \in M_*$ define $a \leq_* b$ if there exits $\alpha \in \mathcal{A}$ with $a, b \in M_\alpha$ and $a \leq_\alpha b$. Again it is readily verified that $\leq_*$ is a well define partial ordering on M_*. Is it well ordered? Let I be any non-empty subset of M^* and pick $\alpha \in \mathcal{A}$ so that $I \cap M_\alpha \neq \emptyset$. Let m be the least element of $I \cap M_\alpha$ with respect to $\leq_\alpha$. We claim that m is also the least element of I with respect to $\leq_*$. Indeed let $i \in I$. If $i \in M_\alpha$, then $m \leq_\alpha i$ by choice of m. So also $m \leq_* i$. If $i \notin M_\alpha$, pick $\beta \in \mathcal{A}$ with $i \in M_\beta$. As $\mathcal{A}$ is a chain, α and β are comparable. As $i \in M_\beta \setminus M_\alpha$ we get $\alpha < \beta$ and (< 3) implies $m \leq_\beta i$. Again $m \leq_* i$ and we conclude that $(M_*, \leq_*)$ is well ordered. Clearly it is also an upper bound for $\mathcal{A}$.

So by Zorn's lemma there exists a maximal element $\alpha \in W$. Suppose that $M_\alpha \neq M$ and pick $m \in M \setminus M_\alpha$. Define the partially ordered set $(M_*, \leq_*)$ by $M_* = M_\alpha \cup \{m\}$, $\leq_*|_{M_\alpha\times M_\alpha}=\leq_\alpha$ and $i <_* m$ for all $i \in M_\alpha$. Then clearly $(M_*, \leq_*)$ is a well-ordered set and $\alpha < (M_*, \leq_*)$, a contradiction to the maximality of α.

Thus $M_\alpha = M$ and $\leq_\alpha$ is a well ordering on M. □

Remark A.11 (Induction)**.** *The well ordering principal allows to prove statement about the elements in an arbitrary set by induction.*

This works as follows. Suppose we like to show that a statement $P(m)$ is true for all elements m in a set M. Endow M with a well ordering $\leq$ and suppose that we can show

$$P(a) \text{ is true for all } a < m \Longleftrightarrow P(m)$$

then the statement is to true for all $m \in M$.

Indeed suppose not and put $I = \{i \in M \mid P(i) \text{ is false }\}$. Then I has a least element m. Put then $P(a)$ is true for all $a < i$ and so $P(i)$ is true by the induction conclusion.

The well-ordering principal can also be used to define objects by induction:

Lemma A.12. *Let I a well ordered set and S any set. For $a \in I$ let $I^a = \{i \in I \mid i \leq a$ and $I_a = \{i \in I \mid i < a\}$. Suppose that for each $a \in I$, $\mathcal{F}_a$ is a set of functions from $I^a \to S$.*

Also suppose that if $f : I_a \to S$ is a function with $f\,|_{I^b} \in \mathcal{F}_b$ for all $b \in I_a$, then there exists $\tilde{f} \in \mathcal{F}_a$ with $\tilde{f}\,|_{I_a} = f$.

Then there exists $f : I \to S$ with $f\,|_{I_a} \in \mathcal{F}_a$ for all $a \in A$.

Proof. Let $\mathcal{I}$ be the set of all subsets J of I so $a \leq b \in J$ implies $a \in J$. Note that either $J = R$ or $J = I_a$ where a is the least element of $R \setminus I$. Put

$$W = \{f : J_f \to S \mid J_f \in \mathcal{I}, f\,|_{I^a} \in \mathcal{F}_a, \forall a \in J\}$$

Order W by $f \leq g$ if $J_f \subset J_g$ and $g\,|_{J_f}$. Let $\mathcal{C}$ be a chain in W. Put $J = \bigcup_{f\in\mathcal{C}} J_f$. Clearly $J \in \mathcal{I}$. Define $f : J \to S$ by $f(j) = g(j)$ where $g \in \mathcal{C}$ with $j \in J_g$. Then also $f\,|_{I^j} = g\,|_{I^j}$

and so $f \in W$. Thus f is an upper bound for W. By Zorn's lemma, $\mathcal{M}$ has a maximal member f. If $J_f = R$ we are done. So suppose $J_f \neq R$. Then $J_f = I_a$ for some $a \in I$. Buy assumptions there exists $\tilde{f} \in \mathcal{F}_a$ with $\tilde{f} \mid_{I_a} = f$. But then $\tilde{F} \in W$ and $f < \tilde{f}$, a contradiction to the maximal choice of f. □

Appendix B

Categories

In this chapter we give a brief introduction to categories.

Definition B.1. *A* category *is a class of objects $\mathcal{C}$ together with*

- *for each pair A and B of objects a set*

$$\mathrm{Hom}(A, B),$$

  an element f of $\mathrm{Hom}(A, B)$ is called a morphism *from A to B and denoted by $f : A \to B$;*

- *for each triple A, B, C of objects a function*

$$\mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C),$$

  for $f : A \to B$ and $g : B \to C$ we denote the image of (g, f) under this function by $g \circ f$, $g \circ f : A \to C$ is called the composite *of f and g;*

such that the following rules hold:

- *[*Associative*] If $f : A \to B$, $g : B \to C$ and $h : C \to D$ are morphisms then*

$$h \circ (g \circ f) = (h \circ g) \circ f$$

- *[*Identity*] For each object A there exists a morphism $\mathrm{id}_A : A \to A$ such that for all $f : A \to B$ and $g : B \to A$*

$$f \circ \mathrm{id}_A = f \text{ and } \mathrm{id}_A \circ g = g$$

A morphism $f : A \to B$ in the category $\mathcal{C}$ is called an *equivalence* if there exists $g : B \to A$ with

$$f \circ g = \mathrm{id}_B \text{ and } g \circ f = id_A$$

Two objects A and B are called *equivalent* if there exists an equivalence $f : A \to B$. Note that associativity implies that the composite of two equivalences is again an equivalence.

Examples

Let $\mathcal{S}$ be the class of all sets. For $A, B \in \mathcal{S}$, let $\mathrm{Hom}(A, B)$ be the set of all functions from $A \to B$. Also let the composites be defined as usual. Note that a morphism is an equivalence if an only if it is a bijection.

The class of all groups with morphisms the group homomorphisms forms category $\mathcal{G}$.

Let $\mathcal{C}$ be a category with a single object A. Let $G = \mathrm{Hom}(A, A)$. The the composite

$$G \times G \to G$$

is a binary operation on G. (I) and (II) now just mean that G is a monoid. Conversely every monoid gives rise to a category with one object which we will denoted by $\mathcal{C}_G$. An object in $\mathcal{C}_G$ is equivalent to $e_G = \mathrm{id}_A$ if and only if it has an has inverse.

Let G be a monoid. For $a, b \in G$ define $\mathrm{Hom}(a, b) = \{x \mid xa = b\}$. If $x : a \to b$ and $y : b \to c$. Then $(yx)a = y(xa) = yb = c$ so $yx : a \to c$. So composition can be defined as multiplication. The resulting category is denoted by $\mathcal{C}(G)$.

The class of all partially ordered sets with morphisms the increasing functions is a category.

Let I be a partially ordered set. Let $a, b \in I$. If $a \leq b$ define $\mathrm{Hom}(a, b) = \emptyset$. If $a \leq b$ then $\mathrm{Hom}(a, b)$ has a single element, which we denote by "$a \to b$". Define the composite by

$$(b \to c) \circ (a \to b) = (a \to c)$$

this is well defined as partial orders are transitive. Associativity is obvious and $a \to a$ is an identity for A. We denote this category by $\mathcal{C}_I$

Let $\mathcal{C}$ be any category. Let $\mathcal{D}$ be the class of all morphisms in $\mathcal{C}$. Given morphisms $f : A \to B$ and $g : C \to D$ in $\mathcal{C}$ define $\mathrm{Hom}(f, g)$ to be the sets of all pairs (α, β) with $\alpha : A \to C$ and $\beta : B \to D$ so that $g \circ \alpha = \beta \circ f$, that is the diagram:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ C & \xrightarrow{g} & D \end{array}$$

commutes.

Let $\mathcal{C}$ be a category. The *opposite* category $\mathcal{C}^{\mathrm{op}}$ is defined as follows:
The objects of $\mathcal{C}^{\mathrm{op}}$ are the objects of $\mathcal{C}$.
$\mathrm{Hom}^{\mathrm{op}}(A, B) = \mathrm{Hom}(B, A)$ for all objects A, B. $f \in \mathrm{Hom}^{\mathrm{op}}(A, B)$ will be denoted by

$$f : A \overset{\mathrm{op}}{\to} B \quad \text{or} \quad f : A \leftarrow B$$

$$f \overset{\text{op}}{\circ} g = g \circ f$$

The opposite category is often also called the *dual* or *arrow reversing* category. Note that two objects are equivalent in $\mathcal{C}$ if and only if they are equivalent in $\mathcal{C}^{\text{op}}$.

Definition B.2. *(a) An object I in a category is called* universal *(or* initial*) if for each object C of $\mathcal{C}$ there exists a unique morphism $I \to C$.*

(b) An object I in a category is called couniversal *(or* terminal*) if for each object C of $\mathcal{C}$ there exists a unique morphism $C \to I$.*

Note that I is initial in $\mathcal{C}$ if and only if its terminal in $\mathcal{C}^{\text{op}}$.

The initial and the terminal objects in the category of groups are the trivial groups.

Let I be a partially ordered set. A object in $\mathcal{C}_I$ is initial if an only if its a least element. Its terminal if and only if its a greatest element.

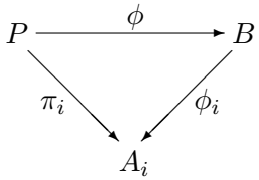
Let G be a monoid and consider the category $\mathcal{C}(G)$. Since $g : e \to g$ is the unique morphism form e to G, e is a initial object. e is a terminal object if and only if G is a group.

Theorem B.3. [**uniuni**] *Any two initial (resp. terminal) objects in a category I are equivalent.*

Proof. Let A and B be initial objects. In particular, there exists $f : A \to B$ and $g : B \to A$. Then id_A and $g \circ f$ both are morphisms $A \to A$. So by the uniqueness claim in the definition of an initial object, $\text{id}_A = g \circ f$, by symmetry $\text{id}_B = f \circ g$.

Let A and B be terminal objects. Then A and B are initial objects in $\mathcal{C}^{\text{op}}$ and so equivalent in $\mathcal{C}^{\text{op}}$. Hence also in $\mathcal{C}$. □

Definition B.4. *Let $\mathcal{C}$ be a category and $(A_i, i \in I)$ a family of objects in $\mathcal{C}$. A* product *for $(A_i, i \in I)$ is an object P in $\mathcal{C}$ together with a family of morphisms $\pi_i : P \to A_i$ such that any object B and family of homomorphisms $(\phi_i : B \to A_i, i \in I)$ there exists a unique morphism $\phi : B \to P$ so that $\pi_i \circ \phi = \phi_i$ for all $i \in I$. That is the diagram commutes:*



commutes for all $i \in I$.

Any two products of $(G_i, i \in I)$ are equivalent in $\mathcal{C}$. Indeed they are the terminal object in the following category $\mathcal{E}$

The objects in $\mathcal{E}$ are pairs $(B, (\phi_i, i \in I))$ there B is an object and $(\phi_i : B \to A_i, i \in I)$ is a family of morphism. A morphism in $\mathcal{E}$ from $(B, (\phi_i, i \in I))$ to $(D, (\psi_i, i \in I)$ is a morphism $\phi : B \to D$ with $\phi_i = \psi_i \circ \phi$ for all $i \in I$.

A *coproduct* of a family of objects $(G_i, i \in I)$ in a category $\mathcal{C}$ is its product in $\mathcal{C}^{\text{op}}$. So it is an initial object in the category $\mathcal{E}$. This spells out to:

Definition B.5. *Let $\mathcal{C}$ be a category and $(A_i, i \in I)$ a family of objects in $\mathcal{C}$. A* coproduct *for $(A_i, i \in I)$ is an object P in $\mathcal{C}$ together with a family of morphisms $\pi_i : A_i \to B$ such that for any object B and family of homomorphisms $(\phi_i : A_i \to B, i \in I)$ there exists a unique morphism $\phi : P \to B$ so that $\phi \circ \pi_i = \phi_i$ for all $i \in I$.*

Bibliography

[Gro] Larry C. Grove, *Algebra* Pure and Applied Mathematics 110, Academic Press, (1983) New Work.

[Hun] Thomas W. Hungerford, *Algebra* Graduate Text in Mathematics 73, Springer-Verlag (1974) New York.

[Lan] Serge Lang, *Algebra* Addison-Wesley Publishing Company, (1965) New York.

Index