Lectures on Abstract Algebra for Graduate Students

Alexander Kleshchev

Contents

Introduction

These are lecture notes for a year long graduate course in abstract algebra given at the University of Oregon in 2002-2003. The text is *Advanced Modern Algebra* by J. Rotman. *I will greatly appreciate if you will let me know of any misprints or errors you can find in these lecture notes.*

This is a difficult course to take and to teach: there is a *lot* of material to cover. As a result, it is difficult to get into things in a deep way, so the course might sometimes even seem boring.

The homework assignments will be given weekly on Mondays and collected also on Mondays. Only part of the problems will be graded. The assignment will usually include sections from the textbook or these lecture notes to read. This part of the assignment should never be ignored! Sometimes I might assign sections which were not explained in class.

The midterm will be on Wednesday of the 6th week of each term, from 6 to 8:15 p.m., if possible. Final during the finals week according to schedule.

I *will* assume that the material usually covered in 500 Algebra courses has been mastered.

Finally, never fear! I am there to help.

1
Groups

1.1 Things known

Throughout the course I will assume some maturity in linear algebra (vector spaces, bases, linear transformations, bilinear forms, duals, ...)

As far as groups are concerned we assume that sections 2.1-2.5 and parts of section 2.6 in Rotman are well understood, as these matters are covered in detail in 500 Algebra.

Among other things, these sections discuss the following important topics (which I will skip):

- 2.2 Symmetric group as an interesting example of a finite group. You need to understand symmetric and alternating groups very well, as these will occur throughout the course as main examples of finite groups.
- 2.3 Formal definition of a group and more examples: cyclic groups, dihedral groups, general linear groups, etc. All of them should be your friends, just like symmetric groups.
- 2.4 Subgroups, cosets, Lagrange's Theorem, Fermat's Theorem, Subgroups generated by subsets.
- 2.5 Homomorphisms and automorphisms, kernels, images, conjugation, normal subgroups, center.
- 2.6 Quotient groups.

1.2 Cyclic groups

We will review some useful properties of cyclic groups. Most of the proofs are omitted. The cyclic group of order n is denoted by C_n, and the infinite cyclic group is denoted by C_∞ (or $\mathbb{Z}$ if we use additive notation).

Lemma 1.2.1 *Let $C_n = \langle g \rangle$. For each divisor d of n, C_n has exactly one subgroup of order d, namely $\langle g^{n/d} \rangle = \{h \in C_n \mid h^d = 1\}$.*

Lemma 1.2.2 *Let $C_n = \langle g \rangle$. Every automorphism of C_n has form $\alpha(g^i) = g^{ki}$ for a fixed k with $(k, n) = 1$. Hence $\mathrm{Aut}(C_n) \cong (\mathbb{Z}/n\mathbb{Z})^\times$, the multiplicative group of the residue ring $\mathbb{Z}/n\mathbb{Z}$.*

Remark 1.2.3 Note that $\mathrm{Aut}(C_n)$ does not have to be cyclic. What is $\mathrm{Aut}(C_8)$?

Lemma 1.2.4 *If an abelian group G has elements of orders k and l, then it has an element of order* $\mathrm{LCM}(k, l)$

Proof Let g and h have orders k and l, respectively. If $(k, l) = 1$ it is easy to see that the order of gh is kl. Otherwise let $d = (k, l)$, and consider the elements g and $h^{l/d}$. □

The following nice result is used very often.

Lemma 1.2.5 *Let F be a field, and G be a finite subgroup of the multiplicative group $F^\times$. Then G is cyclic.*

Proof First observe that, for every $d \in \mathbb{Z}_{>0}$, there are at most d solutions of the equation $x^d = 1$ in $F^\times$—indeed, the polynomial $x^d - 1$ has at most d roots.

Now let g be an element of G having a maximal possible order n. We claim that $G = \langle g \rangle$. Otherwise, pick an element $h \in G \setminus \langle g \rangle$. By the choice of g, the order k of h is at most n. If $k = n$, then $1, g, g^2, \ldots, g^{n-1}, h$ are $n + 1$ solutions of $x^n = 1$, giving a contradiction. So we have $k < n$. Now, k divides n, for otherwise Lemma 1.2.4 yields an element of order $\mathrm{LCM}(k, n) > n$, giving a contradiction. Finally, we get a contradiction anyway, as now $1, g^{n/k}, g^{2n/k}, \ldots, g^{(k-1)n/k}, h$ are $k + 1$ solutions of $x^k = 1$. □

1.3 Simplicity of A_n

Recall that a group is called *simple* if it has exactly two normal subgroups (which then have to be $\{1\}$ and G itself). Just to start our course somewhere, in this section we will prove the classical result that the alternating group A_n is simple for $n \geq 5$. Simple groups are very

important in group theory, and A_n was historically the first example of a simple group.

Lemma 1.3.1 *Let $n \geq 5$. Then all 3-cycles are conjugate in A_n.*

Proof Let $i, j, k, l, m, , \ldots$ be arbitrary (distinct) numbers from

$$\{1, 2, \ldots, n\}.$$

We can conjugate $(1, 2, 3)$ to (i, j, k) using the permutation σ which maps 1 to i, 2 to j, and 3 to k, and leaves $4, 5, \ldots$ invariant. If σ happens to be odd, then use $(l, m)\sigma$ instead. □

Lemma 1.3.2 *Let $n \geq 3$. Then A_n is generated by the 3-cycles.*

Proof Any element of A_n is a product of an even number of transpositions. Consider a product of two transpositions $(i, j)(k, l)$. If all numbers i, j, k, l are distinct, then

$$(i, j)(k, l) = (i, j)(j, k)(j, k)(k, l) = (i, j, k)(j, k, l).$$

Otherwise the product looks like $(i, j)(j, k)$, which equals (i, j, k). The lemma follows. □

If g, h are two elements of a group G, we write $[g, h]$ for the element $ghg^{-1}h^{-1}$, called the *commutator* of g and h. This terminology comes from the fact that g and h commute if and only if their commutator is 1.

Theorem 1.3.3 *Let $n \geq 5$. Then A_n is simple.*

Proof Let $\{1\} \neq H \trianglelefteq A_n$. By the lemmas above, it suffices to show that H contains a 3-cycle. Take $\sigma \in H \setminus \{1\}$.

Let us suppose first that $n = 5$. Then either $\sigma = (i, j)(k, l)$ or $\sigma = (i, j, k, l, m)$. In the former case take $\tau = (i, j)(k, m)$. Then $(m, l, k) = [\tau, \sigma] \in H$. In the latter case, take $\tau = (i, j, k)$. Then $(i, j, l) = [\tau, \sigma] \in H$. We are done in both cases.

Now, let $n = 6$. If σ fixes some $i \in \{1, 2, \ldots, 6\}$, then σ belongs to a subgroup $A_5 < A_6$ permuting the remaining 5 symbols. Thus, $\sigma \in H \cap A_5$. Therefore $H \cap A_5$ is a non-trivial normal subgroup of A_5. As we already know that A_5 is simple, this implies $H \geq A_5$. In particular, H contains a 3-cycle, and we are done. So assume σ does not fix any element. Then either $\sigma = (i, j)(k, l, m, r)$ or $\sigma = (i, j, k)(l, m, r)$.

In the first case, $1 \neq \sigma^2 \in H$ fixes i (and j), and we are reduced to the case already considered above. In the second case take $\tau = (j, k, l)$. We have again $(k, i, m)(j, l, k) = [\sigma, \tau] \in H$ fixes r.

Finally, let $n \geq 7$. There exist $i \neq j$ with $\sigma(i) = j$. Chose a 3-cycle α which fixes i but moves j. Then $\alpha\sigma \neq \sigma\alpha$, as the two elements differ on i. Hence $\gamma := [\alpha, \sigma]$ is a non-trivial element of H. But $\sigma\alpha^{-1}\sigma^{-1}$ is a 3-cycle. So γ is a product of two 3-cycles. Hence it moves at most 6 elements, say, $i_1, \ldots, i_6$. Let $F \cong A_6$ be the alternating group on $\{i_1, \ldots, i_6\}$ considered as a subgroup of A_n. Then $\gamma \in H \cap F$, whence $H \geq F$ by simplicity of F, and so H contains a 3-cycle. □

Example 1.3.4 A permutation of the set $\{1, 2, \ldots\}$ is called *finitary* if it fixes all but finitely many points. Denote by A_∞ the finitary alternating group, i.e. the group of all even finitary permutations. Using the fact that $A_\infty = \cup_{n \geq 1} A_n$, it is easy to see that A_∞ is simple, giving us an example of an infinite simple group.

1.4 Isomorphism and Correspondence Theorems

Theorem 1.4.1 (First Isomorphism Theorem) *If $f : G \to H$ is a homomorphism of groups, then $K := \ker f \lhd G$, and the map*

$$\bar{f} : G/K \to \operatorname{im} f, \quad gK \mapsto f(g)$$

is an isomorphism.

Proof It is routine to check that K is normal, that $\bar{f}$ is a well-defined homomorphism (the most important part, so make sure you understand this), and that $\bar{f}$ is surjective and injective. □

Example 1.4.2 The cyclic group C_m of order m is isomorphic to $\mathbb{Z}/m\mathbb{Z}$.

Example 1.4.3 Let S^1 be the group of all complex numbers of absolute value 1. Then considering the map $\mathbb{R} \to S^1$, $x \mapsto e^{2\pi i x}$, gives an isomorphism $\mathbb{R}/\mathbb{Z} \cong S^1$.

Example 1.4.4 The determinant map $\det : GL_n(F) \to F^\times$ yields an isomorphism $GL_n(F)/SL_n(F) \cong F^\times$.

Example 1.4.5 The sign map $S_n \to \{\pm 1\}$ shows that $S_n/A_n \cong C_2$.

Example 1.4.6 If $d \mid n$ then $C_n/C_d \cong C_{n/d}$.

Example 1.4.7 For any group G, we have $G/Z(G) \cong \mathrm{Inn}(G)$, where $Z(G)$ is the center of G, and $\mathrm{Inn}(G)$ is the group of the inner automorphisms of G.

Example 1.4.8 Let p be a prime, and $C_{p^\infty} < \mathbb{C}^\times$ be the group of all p^nth roots of 1 for all $n \geq 0$. Considering the map $z \mapsto z^p$ yields an isomorphism $C_{p^\infty}/C_p \cong C_{p^\infty}$. It can be proved that actually any non-trivial quotient of C_{p^∞} is isomorphic to C_{p^∞}. Another curious property of this group is that any finitely generated subgroup of it is cyclic, even though it is not cyclic itself.

Example 1.4.9 Let p be a prime $\mathbb{Q}_{(p)}$ be a subgroup of $(\mathbb{Q}, +)$ which consists of all numbers of the form m/p^n for $m, n \in \mathbb{Z}$. Considering the map $\mathbb{Q}_{(p)} \to \mathbb{Z}_{p^\infty},\ m/p^n \mapsto e^{2\pi i m/p^n}$ yields an isomorphism $\mathbb{Q}_{(p)}/\mathbb{Z} \cong \mathbb{Z}_{p^\infty}$.

Theorem 1.4.10 (Second Isomorphism Theorem) *Let $H \trianglelefteq G$, $K \leq G$. Then $H \trianglelefteq HK \leq G$, $H \cap K \trianglelefteq K$, and the map*

$$\varphi : K/(H \cap K) \to HK/H, \quad k(H \cap K) \mapsto kH$$

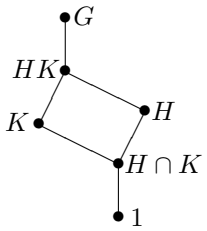
is an isomorphism of groups.

Proof Things to check: (1) $HK < G$, (2) $H \trianglelefteq HK$, (3) $H \cap K \trianglelefteq K$, (4) φ is a well-defined homomorphism, (5) φ is surjective and injective. All are routine. For example, for (1): given elements $h_1, h_2 \in H$ and $k_1, k_2 \in K$,

$$h_1k_1(h_2k_2)^{-1} = h_1k_1k_2^{-1}h_2^{-1} = h_1h_3k_1k_2^{-1} \in HK,$$

where $h_3 = k_1k_2^{-1}h_2^{-1}(k_1k_2^{-1})^{-1} \in H$, as H is normal. $\square$

Remark 1.4.11 The following picture might help you to remember what the Second Isomorphism Theorem is saying.



Example 1.4.12 Let V_1, V_2 be finite dimensional subspaces of a vector space V. Then the map in the Second Isomorphism Theorem gives an isomorphism (of groups, but actually vector spaces) $V_1/(V_1 \cap V_2) \cong (V_1 + V_2)/V_2$, which implies the well-known dimension formula

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

Theorem 1.4.13 (Third Isomorphism Theorem) *Let $H, K \trianglelefteq G$, $K \leq H$. Then $L := \{hK \mid h \in H\} \subseteq G/K$ is a normal subgroup isomorphic to H/K, and the map*

$$\frac{G/K}{L} \to G/H, \quad (gK)L \mapsto gH$$

is an isomorphism of groups.

Proof Apply the First Isomorphism Theorem to the map $f : G/K \to G/H,\ gK \mapsto gH$. □

Remark 1.4.14 Slightly informally, the Third Isomorphism Theorem claims that the following cancellation rule is true:

$$\frac{G/K}{H/K} \cong G/H.$$

The following theorem is used very often, and so you need to make sure that you understand what it says, and how the proof is 'obvious' or 'routine'.

Theorem 1.4.15 (Correspondence Theorem) *Let $K \trianglelefteq G$ and $\pi : G \to G/K$ be the natural projection. Then:*

(i) *For any subgroup S of G containing K, K is a normal subgroup of S.*

(ii) *$\pi(S)$ is a subgroup of G/K isomorphic to S/K.*

(iii) *The maps π and π^{-1} establish a bijection between the subgroups of G containing K and the subgroups of G/K.*

(iv) *The bijection respect inclusions and indexes, i.e. $T \leq S$ if and only if $\pi(T) \leq \pi(S)$, in which case $[S : T] = [\pi(S) : \pi(T)]$.*

(v) *The bijection respects normality and quotients, i.e. $T \trianglelefteq S$ if and only if $\pi(T) \trianglelefteq \pi(S)$, in which case $S/T \cong \pi(S)/\pi(T)$.*

Proof (i) is obvious. (ii) follows from the First Isomorphism Theorem. (iii) is clear because π and π^{-1} induce mutually inverse maps between our two sets of subgroups. The first part of (iv) is obvious and the second is a routine exercise with cosets representatives. Finally, (v) follows from the Third Isomorphism Theorem. □

Example 1.4.16 We have $SL_n(\mathbb{F}_5) \triangleleft GL_n(\mathbb{F}_5)$, and

$$GL_n(\mathbb{F}_5)/SL_n(\mathbb{F}_5) \cong \mathbb{F}_5^\times \cong C_4.$$

So there is only one subgroup strictly between $SL_n(\mathbb{F}_5)$ and $GL_n(\mathbb{F}_5)$: the group of matrices with determinant ± 1.

1.5 Group Actions and First Applications

Group actions are a very powerful tool for studying groups themselves. A special case of this is when a group acts with linear transformations on a vector space. Such actions are a subject of *group representation theory*, a very important and active area of mathematics. However, we start from groups acting on sets:

Definition 1.5.1 Let X be a set and G be a group. We say that G *acts on* X or X is a *G-set*, if there is a function

$$G \times X \to X, \quad (g, x) \mapsto g \cdot x$$

such that

(i) $(gh) \cdot x = g \cdot (h \cdot x)$ for all $g, h \in G$, $x \in X$;

(ii) $1 \cdot x = x$ for all $x \in X$.

Two G-sets X and Y are *isomorphic* if there is a bijection $f : X \to Y$ satisfying $f(g \cdot x) = g \cdot f(x)$ for all $g \in G$ and $x \in X$.

Remark 1.5.2 Having an action of G on X is equivalent to having a homomorphism $G \to S(X)$, where $S(X)$ is the group of bijective maps from X to itself.

Definition 1.5.3 Let X be a G-set. An *orbit* of $x \in X$, denoted $G \cdot x$ or $\mathcal{O}(x)$, is the set $\{g \cdot x \mid g \in G\}$. An *orbit of G on X* is an orbit of some $x \in X$. If G has only one orbit on X we say that G acts *transitively* on X. The *stabilizer of $x \in X$* is the subgroup $G_x := \{g \in G \mid g \cdot x = x\}$. The *kernel* of the action is the subgroup $\{g \in G \mid g \cdot x = x \text{ for all } x \in X\}$. The action is called *faithful* if the kernel is trivial. If the action of G on X is faithful we also say that G is a permutation group on X. In this case, if G acts transitively on X we also say that G is a *transitive permutation group on X*.

We first prove some simple general results on group actions and then consider examples.

Proposition 1.5.4 *Every G-set is a disjoint union of the G-orbits.*

Proof The relation $x \equiv y$ if and only if there exists $g \in G$ with $g \cdot x = y$ is an equivalence relation, and equivalence classes are orbits. □

Theorem 1.5.5 *Let X be a G-set, and $x \in X$. The map*

$$G/G_x \to G \cdot x, \quad gG_x \mapsto g \cdot x$$

is a well-defined bijection. In particular, if $[G : G_x]$ is finite, then $[G : G_x] = |G \cdot x|$.

Proof Routine. □

Lemma 1.5.6 *Let X be a G-set, $x \in X$, and $g \in G$. Then $G_{g \cdot x} = gG_xg^{-1}$.*

Proof Pretty obvious. □

The interesting exaples of permutation groups abound. Here are some illustrations.

Example 1.5.7 *Regular action* of G on itself, i.e. $g \cdot h = gh$. Note that the homomorphism $G \to S(G)$ corresponding to this action according to Remark 1.5.2, is injective. This proves that any finite group is a

subgroup of the symmetric group $S_{|G|}$, the fact usually referred to as Cayley's Theorem. The regular action is transitive, and stabilizers of all elements are trivial: $G_g = \{1\}$.

Example 1.5.8 If $H \leq G$ is a subgroup, we have the action of G on G/H given by $g \cdot (g'H) = gg'H$. This gives a homomorphism of G into $S_{[G:H]}$. The action is again transitive, and for the stabilizers we have: $G_{gH} = gHg^{-1}$. The kernel of the action is contained in H. In fact, the kernel is $\cap_{g \in G} gHg^{-1}$. Actions of groups on the left cosets are very important because every transitive action is isomorphic to such one. Indeed Theorem 1.5.5 shows that any transitive G-set X is isomorphic to the G-set G/G_x for any $x \in X$.

Example 1.5.9 Action of G on the elements of a normal subgroup $H \trianglelefteq G$ by *conjugation*, i.e. $g \cdot h = ghg^{-1}$. The action is transitive only if $H = \{1\}$, as $G \cdot 1 = \{1\}$. Let $H = G$ (which is the most important special case). Then the orbits are called the *conjugacy classes* of G. The conjugacy class of x is denoted by x^G. The stabilizer G_x is called the *centralizer* of x in G and denoted $C_G(x)$. Theorem 1.5.5 in this case reads:

$$|x^G| = [G : C_G(x)]. \tag{1.1}$$

The conjugacy class of x consists of only one element if and only if x belongs to the *center*

$$Z(G) := \{z \in G \mid zg = gz \text{ for any } g \in G\}.$$

Now (1.1) together with Proposition 1.5.4 give what is called the *class equation* of a finite group G:

$$|G| = |Z(G)| + \sum_i [G : C_G(x_i)], \tag{1.2}$$

where one x_i is selected from each conjugacy class having more than one element.

Example 1.5.10 Action of G on the set of its subgroups by conjugation, i.e. $g \cdot H = gHg^{-1}$ for $g \in G, H \leq G$. There is a special name and notation for the stabilizer G_H: it is called the *normalizer* of H in G and denoted $N_G(H)$.

Example 1.5.11 By definition, D_8 acts naturally on the vertices of the square. The action is transitive, the stabilizer of any point x is $\{1, r\}$, where r is the (unique) reflection in D_8, which leaves x invariant.

Example 1.5.12 By definition, the group G of rotations of cube acts on the four diagonals of the cube. It is easy to see that the action is faithful. Moreover, for every pair of diagonals there exists a rotation swapping them and leaving the other two diagonals invariant. This shows that $G \cong S_4$.

Example 1.5.13 By definition, $GL(V)$ acts naturally on the elements of V. There are two orbits: the origin $\{0\}$ and $V \setminus \{0\}$.

Example 1.5.14 The action of $GL_n(F)$ on all matrices $M_n(F)$ by conjugation. If the ground field F is algebraically closed then the orbits are parametrized by the *Jordan normal forms* of matrices.

Example 1.5.15 The natural action of $GL(V)$ on the set $\mathbb{P}(V)$ of the lines in V is transitive.

Example 1.5.16 The natural action of $GL(V)$ on the bases of V is transitive. The stabilizer of any element is trivial. Assume that $V = \mathbb{F}_q^n$, where $\mathbb{F}_q$ is a field with q elements. Theorem 1.5.5 implies that $|GL_n(\mathbb{F}_q)|$ equals the number of bases in V. This number is easy to calculate! Indeed, for the first element of a basis we can take any non-zero element, and so we have $(q^n - 1)$ options. After the first element has been chosen, we can choose from $(q^n - q)$ elements of V which do not belong to the line spanned by the first element. When the first two elements have been chosen, we may choose the third element from any of the $(q^n - q^2)$ elements of V, which do not belong to the plane spanned by the first two elements, and so on. Thus,

$$\begin{aligned} |GL_n(\mathbb{F}_q)| &= (q^n - 1)(q^n - q) \dots (q^n - q^{n-1}) \\ &= q^{n(n-1)/2}(q-1)(q^2-1)\dots(q^n-1). \end{aligned} \tag{1.3}$$

Therefore using Example 1.4.4, we get

$$|SL_n(\mathbb{F}_q)| = q^{n(n-1)/2}(q^2-1)(q^3-1)\dots(q^n-1). \tag{1.4}$$

Example 1.5.17 The natural action of the symmetric group S_n on the set $\{1, \dots, n\}$ is transitive with the stabilizer of any element isomorphic

to S_{n-1}. This action induces a transitive action on the ordered k-tuples of distinct elements from $\{1,\ldots,n\}$, whose stabilizers are isomorphic to S_{n-k} (here $k \leq n$).

Example 1.5.18 Let us use group actions to prove that S_3 is the only non-abelian group of order 6. Let G be such a group. Then there must be an element $c \in G$ of order 3 and an element $s \in G$ of order 2 (the fact that there is an element of order 2 follows by considering pairs $\{g, g^{-1}\}$, and the fact that not every element is of order 2 follows as G is not abelian). Now, $\langle c\rangle \triangleleft G$, as a subgroup of index 2 is always normal. Let $scs^{-1} = c^i$ for $i \in \{0,1,2\}$. Then i can only be 2, as s and c do not commute. Now consider the action of G on the cosets $G/\langle s\rangle$. The kernel of the action is contained in $\langle s\rangle$. But s is not in the kernel, as otherwise $sc\langle s\rangle = c\langle s\rangle$, hence $c^{-1}sc = s$, i.e. $sc = cs$, giving a contradiction. Thus the kernel is trivial, and so we have constructed an embedding of G into S_3. As both groups have order 6 they must be isomorphic.

Example 1.5.19 Let $H, K \leq G$ be two subgroups. We can restrict the action on left cosets G/H from G to K. Orbits of this K-action are called (K,H)-double cosets of G. A (K,H)-double cosets is thus the set

$$KgH = \{kgh \mid k \in K,\ h \in H\}.$$

By Proposition 1.5.4, G is a disjoint union of double cosets. There is one important difference between single and double cosets: all single cosets have the same cardinality, while double cosets might have different cardinalities (find an example!) Theorem 1.5.5 yields

$$|KgH| = \frac{|K||H|}{|gHg^{-1} \cap K|}.$$

Example 1.5.20 The *Klein* 4*-group* is defined to be

$$V_4 := \{1, (12)(34), (13)(24), (14)(23)\}. \tag{1.5}$$

This is a transitive subgroup of S_4 isomorphic to $C_2 \times C_2$. It can be shown that any transitive subgroup of S_4 is either (i) S_4, (ii) A_4, (iii) V_4, (iv) cyclic group of order 4, or (v) D_4 acting on the verices of the square.

We now obtain some more applications of group actions.

Let p be a prime. Recall that a finite group is called a *p-group* if the order of G is a power of p.

Theorem 1.5.21 *The center of a non-trivial p-group is non-trivial.*

Proof Consider the class equation (1.2) of our p-group G. By assumption $|G|$ and all $[G : C_G(x_i)]$ are divisible by p, so the same must be true for $|Z(G)|$. □

Corollary 1.5.22 *Let p be a prime. Every group of order p^2 is abelian.*

Proof In view of the previous theorem we may assume that $|Z(G)| = p$. Then $G/Z(G) \cong C_p$. Take any element $g \in G \setminus Z(G)$. Then any element of $G \setminus Z(G)$ looks like $g^k z$ for some k and some $z \in Z(G)$, so g commutes with any such element. As g also commutes with the elements of $Z(G)$, it follows that $g \in Z(G)$, giving a contradiction. □

Remark 1.5.23 The result above does not generalize to groups of order p^3. Indeed, the group of all strictly upper triangular 3 by 3 matrices over $\mathbb{F}_p$ is a non-abelian group of order p^3.

We now prove the important Cauchy's Theorem. First we deal with the abelian version.

Lemma 1.5.24 *If G is a finite abelian group whose order is divisible by a prime number p, then G contains an element of order p.*

Proof Apply induction on $|G|$, starting from $|G| = p$, when $G \cong C_p$, and so the result holds. Now let $|G| > p$. Let $g \neq 1$ be any element of G, and k be the order of g. If p divides k then the order of $g^{k/p}$ is p. Otherwise, consider $G/\langle g \rangle$. By inductive assumption there is an element $h\langle g \rangle \in G/\langle g \rangle$ of order p. Let h have order m. Then $(h\langle g \rangle)^m = \langle g \rangle$ in $G/\langle g \rangle$, so $p|m$, and we can take $h^{m/p}$. □

Theorem 1.5.25 (Cauchy's Theorem) *If G is a finite group whose order is divisible by a prime number p, then G contains an element of order p.*

Proof Let $|G| = pm$. We apply induction on m. If $m = 1$ then $G = C_p$, and the result is true. Let $m > 1$. Consider the class equation (1.2). If p divides some $|C_G(x_i)|$, we apply the inductive hypothesis to the subgroup $C_G(x_i)$. So we may assume that $p \not| |C_G(x_i)|$ for every i. Hence

$p \mid [G : C_G(x_i)]$ for every i, whence $p \mid |Z(G)|$. Therefore, we may assume that G is abelian, and apply Lemma 1.5.24. □

Corollary 1.5.26 *If $|G| = p^e$, then G has a normal subgroup of order p^k for every $k \leq e$.*

Proof Apply induction on $e \geq 1$, the induction base being clear. For the inductive step, we know by Theorem 1.5.21 and Lemma 1.5.24 that G has a normal subgroup of order p. Now the result follows from the inductive hypothesis and Correspondence Theorem. □

Corollary 1.5.27 *An abelian group is simple if and only if it has a prime order.*

Proof Follows immediately from Lemma 1.5.24. □

Example 1.5.28 Let G be a finite group with $g \sim g^2$ for every $g \in G$. Then $G = \{1\}$. Indeed, let p be the smallest prime dividing $|G|$, and take $g \in G$ of order p. As $g^2 = hgh^{-1}$ for some $h \in G$, we have $h\langle g\rangle h^{-1} = \langle g^2\rangle \leq \langle g\rangle$. Hence $h\langle g\rangle h^{-1} = \langle g\rangle$, i.e. $h \in N_G(\langle g\rangle)$. Moreover $h \notin C_G(\langle g\rangle)$. In particular, $N_G(\langle g\rangle)/C_G(g)$ is non-trivial. Note that $N_G(\langle g\rangle)/C_G(g)$ embeds into $\mathrm{Aut}(\langle g\rangle) \cong C_{p-1}$, so $[N_G(\langle g\rangle) : C_G(g)]$ divides $p-1$. Take a prime smaller than p dividing $[N_G(\langle g\rangle) : C_G(g)]$. This prime should also divide $|G|$, which gives a contradiction.

The following beautiful qualitative result is another application of group actions.

Theorem 1.5.29 (Landau) *For each positive integer k there exists a bound $B(k)$ such that a finite group having exactly k conjugacy classes satisfies $|G| \leq B(k)$.*

Proof Let G has k conjugacy classe of sizes $c_1 = 1, c_2, \ldots, c_k$. Write $|G|/c_i = n_i \in \mathbb{Z}_{>0}$. Then $1 = (1/n_1) + \cdots + (1/n_k)$. Note that $n_1 = |G|$, so it suffices to prove that the equation

$$1 = \frac{1}{x_1} + \cdots + \frac{1}{x_k}$$

has only finitely many solutions in positive integers, which is left as an exercise. □

Remark 1.5.30 For further reading on permutation groups we recommend [Ca] and [Wi].

1.6 Direct and Semidirect Products

(Semi)direct product is a building tool: it allows you to form new groups from old ones. This construction occurs very often in 'real life'. We start from direct products.

Definition 1.6.1 Let $G_1, G_2, \ldots, G_n$ be groups. The *direct product* $G_1 \times G_2 \times \cdots \times G_n$ is the cartesian product $G_1 \times G_2 \times \cdots \times G_n$ (whose elements are all n-tuples $(g_1, g_2, \ldots g_n)$ such that $g_i \in G_i$ for $1 \leq i \leq n$), with the component-wise multiplication.

Example 1.6.2 Let m, n be positive integers relatively prime to each other. Then, by Lemma 1.2.4, we have $C_{mn} \cong C_m \times C_n$. Conversely, $C_m \times C_n$ is cyclic only if $(m, n) = 1$.

Clearly, the subset

$$G_i' := \{(g_1, \ldots, g_n) \in G_1 \times \cdots \times G_n \mid g_1 = \cdots = g_{i-1} = g_{i+1} = \cdots = g_n = 1\}$$

is a normal subgroup of $G_1 \times \cdots \times G_n$ isomorphic to G_i. Observe that the elements of G_i' and G_j' commute whenever $i \neq j$, every element $g \in G_1 \times \cdots \times G_n$ can be written uniquely as a product $g = g_1 \ldots g_n$ in which $g_i \in G_i'$ for all i, and $G_1' G_2' \ldots G_i' \cap G_{i+1}' = \{1\}$ for all $i < n$. The following theorem shows how such properties actually characterize direct products.

Theorem 1.6.3 *Let $G_1, \ldots, G_n$ be subgroups of a group G.*

(i) *Assume that $g_i \in G_i$ and $g_j \in G_j$ commute whenever $i \neq j$, and that every $g \in G$ can be written uniquely as a product $g = g_1 \ldots g_n$ in which $g_i \in G_i$ for all i. Then the map*

$$G_1 \times \cdots \times G_n \to G, \quad (g_1, \ldots, g_n) \mapsto g_1 \ldots g_n$$

is an isomorphism of groups.

(ii) *Assume that $G_i \trianglelefteq G$ for all i, that $G_1 G_2 \ldots G_i \cap G_{i+1} = \{1\}$ for all $i < n$, and that $G = G_1 G_2 \ldots G_n$. Then $G \cong G_1 \times \cdots \times G_n$.*

Proof (i) is a simple check. We prove (ii) by showing that the subgroups G_i satisfy the assumptions of (i). Let $g_i \in G_i$, $g_j \in G_j$, and, say, $i < j$. Then, using the normality of G_i and G_j, we see that the commutator $[g_i, g_j]$ belongs to $G_i \cap G_j \subseteq (G_1 \dots G_{j-1}) \cap G_j = \{1\}$, so g_i and g_j commute. By assumption, every $g \in G$ can be written as a product $g_1 \dots g_n$ with $g_i \in G_i$, and it remains to prove that such presentation is unique. Well, assume there is another such presentation: $g = h_1 \dots h_n$. Then

$$g_n h_n^{-1} = (g_1 \dots g_{n-1})^{-1}(h_1 \dots h_{n-1}).$$

Now, using the commuting property which we just proved, we see that $g_n h_n^{-1} \in G_n \cap (G_1 \dots G_{n-1}) = \{1\}$. Thus $g_h = h_n$. Now, we use a similar argument to prove that $g_{n-1} = h_{n-1}$, etc. □

Now we study a generalization of a direct product called a semidirect product.

Definition 1.6.4 Let G and H be two groups, and $\varphi : H \to \mathrm{Aut}(G)$ be a (group) homomorphism. Denote $(\varphi(h))(g)$ by $h \cdot g$. The *semidirect product of G and H* (relative to φ) is the cartesian product of G and H with multiplication

$$(g, h)(g', h') = (g(h \cdot g'), hh') \qquad (g, g' \in G,\ h, h' \in H).$$

The semidirect product of G and H is denoted by $G \rtimes H$ or, if one wants to emphasize which φ is used, by $G \times_\varphi H$.

Having a homomorphism $\varphi : H \to \mathrm{Aut}(G)$ is equivalent to having an *action of H on G by group automorphisms*. This is why it is convenient to use the notation $h \cdot g$ for $(\varphi(h))(g)$. In particular,

$$h_1 \cdot (h_2 \cdot g) = (h_1 h_2) \cdot g \qquad (h_1, h_2 \in H,\ g \in G). \tag{1.6}$$

One needs to verify that $G \rtimes H$ is a group. It is easy to see that $(1, 1) \in G \rtimes H$ is the identity element, and $(g, h)^{-1} = (h^{-1} \cdot g^{-1}, h^{-1})$. The associativity boils down to the property (1.6).

Note that $H' := \{(1, h) \in G \rtimes H \mid h \in H\}$ is a subgroup of $G \rtimes H$, isomorphic to H, and $G' := \{(g, 1) \in G \rtimes H \mid g \in G\}$ is a *normal* subgroup of $G \rtimes H$ isomorphic to G. Moreover, $G' \cap H' = \{1\}$ and $G \rtimes H = G'H'$. These properties characterize $G \rtimes H$ uniquely:

Theorem 1.6.5 *Let K be a group, and G, H be two subgroups of K. Suppose that $G \trianglelefteq K$, $G \cap H = \{1\}$, and $K = GH$. Then $K \cong G \rtimes H$.*

Proof First of all, define the homomorphism $\varphi : H \to \mathrm{Aut}(G)$ via

$$(\varphi(h))(g) = hgh^{-1} \qquad (g \in G,\ h \in H).$$

We claim that G is isomorphic to the semidirect product $G \rtimes H$ taken relative to this homomorphism. Consider the map

$$\theta : G \rtimes H \to K, \quad (g, h) \mapsto gh \qquad (g \in G,\ h \in H).$$

It follows from the definitions that θ is a group homomorphism, and assumptions imply that θ is surjective and injective in the usual way. □

Example 1.6.6 Let D_n be the dihedral group of order $2n$, generated by the rotation x and by the reflection y. Thus $x^n = 1$, $y^2 = 1$, $yxy^{-1} = x^{-1}$. Let $C_n = \langle x \rangle$ and $C_2 = \langle y \rangle$. By the theorem above, $D_n \cong C_n \rtimes C_2$. It is also true that $D_8 \cong (C_2 \times C_2) \rtimes C_2$ (make sure you believe it!). Finally, $Z := \langle x^2 \rangle$ is normal subgroup of D_8, and $D_8/Z \cong C_2 \times C_2$, but it is not true that $D_8 \cong C_2 \rtimes (C_2 \times C_2)$.

Example 1.6.7 We have $C_4/C_2 \cong C_2$, but $C_4 \ncong C_2 \rtimes C_2$. In fact, abelian group is a semidirect product $H \rtimes K$ if and only if it is a direct product $H \times K$.

Example 1.6.8 Let Q_8 be the quaternion group, i.e.

$$Q_8 = \{\pm 1, \pm i, \pm j, \pm k\} < \mathbb{H}^\times.$$

Let $x = i$ and $y = j$. Then the following relations hold:

$$x^4 = 1, \quad x^2 = y^2, \quad yxy^{-1} = x^{-1}.$$

Observe that $C_4 := \langle c \rangle$ is a normal subgroup of Q_8, and $Q_8/C_4 \cong C_2$. However, $Q_8 \ncong C_4 \rtimes C_2$. Indeed, it suffices to notice that $-1 = x^2$ is the only element of Q_8 of order 2 and the subgroup $\langle -1 \rangle$ is contained in every subgroup of Q_8 isomorphic to C_4. In fact, Q_8 cannot be presented as a non-trivial semiderect product.

Example 1.6.9 $S_n \cong A_n \rtimes C_2$.

Example 1.6.10 Let V be a finite dimensional vector space over a field F. The *affine* group of V is defined to be the semidirect product

$$AGL(V) := V \rtimes GL(V),$$

where the action of $GL(V)$ on V is taken to be natural. Elements of $AGL(V)$ are called *affine transformations* of V.

Example 1.6.11 Let p and q be primes, and $p \equiv 1 \pmod{q}$. Then there exists a non-abelian group of the form $C_p \rtimes C_q$. Indeed, $\mathrm{Aut}(C_p) \cong C_{p-1}$, and, since $q|(p-1)$, there is an embedding $C_q \hookrightarrow \mathrm{Aut}(C_p)$. This allows us to define the semidirect product which is not commutative.

Remark 1.6.12 If $H \trianglelefteq G$, and $G/H \cong K$, we say that G is an *extension* of H by K. For example, D_8, Q_8, C_8, and $C_4 \times C_2$ are extensions of C_4 by C_2. Given H and K, the *extension problem* is to classify all extensions of H by K. This problem leads naturally to group *cohomology* to be studied later.

As for the extensions of C_4 by C_2, our list in the previous paragraph is complete and irredundant. This follows from from Lemma 1.11.17.

1.7 Sylow Theorems

The following results are important and surprising. If you don't think they are surprising, try to generalize them along the following lines: if k is a divisor of $|G|$, then there is a subgroup of G of order k. You will find out that Sylow goes about as far as one can go in this direction in the generality of *any finite group*. In fact, there aren't so many deep results valid for *any finite group*. Sylow theorems are perhaps the most important of these. Sylow p-subgroups provide us with a powerful tool for studying finite groups. Let us state the theorems.

Theorem 1.7.1 (First Sylow Theorem) *If G is a finite group of order $p^e m$ with $(m,p) = 1$, then every p-subgroup of G is contained in a subgroup of order p^e. In particular, subgroups of order p^e exist.*

Definition 1.7.2 Subgroups of order p^e in the First Sylow theorem are called *Sylow p-subgroups* of G. In view of the theorem, they can also be defined as maximal p-subgroups of G.

Theorem 1.7.3 (Second Sylow Theorem) *Sylow p-subgroups of G are conjugate to each other.*

Theorem 1.7.4 (Third Sylow Theorem) *The number r of Sylow p-subgroups of G is a divisor of $|G|$, and also satisfies $r \equiv 1 \pmod{p}$.*

We now prove Sylow theorems. Notice how group actions come into play in a crucial way again. When we speak of a maximal p-subgroup below we mean maximal with respect to inclusion, not size. In the end these will turn out to be equivalent though.

Lemma 1.7.5 *Let P be a maximal p-subgroup of a finite group G.*

(i) *Every conjugate of P is also a maximal p-subgroup.*
(ii) *$[N_G(P) : P]$ is prime to p.*
(iii) *If $g \in G$ is a p-element, and $gPg^{-1} = P$ then $g \in P$.*

Proof (i) is clear, and (iii) follows from (ii). Let us prove (ii). If p divides $[N_G(P) : P]$ then, by Cauchy's theorem, $N_G(P)/P$ contains an element gP of order P. By the Correspondence theorem, this yields a subgroup S of $N_G(P)$ with $S/P \cong C_p$, which contradicts the maximality of P. □

Theorem 1.7.6 *Let G be a finite group.*

(i) *All maximal p-subgroups are conjugate in G.*
(ii) *The number r of maximal p-subgroups divides $|G|$ and satisfies $r \equiv 1 \pmod{p}$.*

Proof Let P be a maximal p-subgroup of a finite group G, and let $X = \{P = P_1, P_2, \ldots, P_r\}$ be set of its conjugates. Now, P acts on X by conjugation. By Lemma 1.7.5(iii), there is only one orbit of size 1, namely $\{P\}$. All other orbits must have lengths divisible by p, and so $r \equiv 1 \pmod{p}$. Moreover, considering X as a G-set, we have $r = [G : N_G(P)]$, so r divides $|G|$.

Finally, let Q be any maximal p-subgroup in G. It acts on X by conjugation. As $r \equiv 1 \pmod{p}$, there must be a Q-orbit of size 1, say $\{P_j\}$. By Lemma 1.7.5(iii), $Q \leq P_j$, hence $Q = P_j$. □

Theorem 1.7.7 *Let G be a finite group of order $p^e m$ with $(m, p) = 1$. If P is a maximal subgroup of G, then $|P| = p^e$.*

Proof It suffices to prove that $p \not| [G : P]$. Well,

$$[G : P] = [G : N_G(P)][N_G(P) : P],$$

and $[G : N_G(P)] \equiv 1 \pmod{p}$ by Theorem 1.7.6, while $[N_G(P) : P]$ is prime to p thanks to Lemma 1.7.5(ii). □

Note that Theorems 1.7.6 and 1.7.7 imply the three Sylow theorems. In what follows we denote by $r_p(G)$ the number of Sylow p-subgroups of G.

Example 1.7.8 The order of the Sylow 2-subgroup of S_4 is 8. On the other hand, D_8 is a subgroup of S_4, because it acts faithfully on the four vertices of the square. It follows that Sylow 2-subgroups of S_4 are isomorphic to D_8.

Example 1.7.9 It follows from (1.3) that the group of upper unitriangular n by n matrices is a Sylow p-subgroup of $GL_n(\mathbb{F}_{p^n})$, where $\mathbb{F}_{p^n}$ is a field with p^n elements.

Example 1.7.10 Sylow 3-subgroup of S_9 is isomorphic to $(C_3 \times C_3 \times C_3) \rtimes C_3$, where the generator c of C_3 acts on $(C_3 \times C_3 \times C_3)$ as follows: $c \cdot (c_1, c_2, c_3) = (c_3, c_1, c_2)$. Think what a Sylow 3-subgroup of S_{27} might look like...

Example 1.7.11 There is no simple group of order 120. Assume for a contradiction that G is such a group. Then are exactly 6 Sylow 5-subgroups. The conjugation action on the Sylow 5-subgroups yields an embedding of G into S_6. By simplicity of G, we have $G \leq A_6$, and $[A_6 : G] = 3$. But A_6 does not have subgroups of index 3, as it is simple and therefore cannot act on a three element set. A contradiction.

Here are some easy consequences of Sylow theorems. These are the first example of how one might apply Sylow p-subgroups of G to studying a group structure of G.

Proposition 1.7.12 *Let G be a finite group.*

(i) *Let P be a Sylow p-subgroup of G. Then G has a unique Sylow p-subgroup if and only if $P \trianglelefteq G$.*

(ii) *All Sylow p-subgroups of G are normal if and only if G is a direct product of its Sylow p-subgroups. In particular, every finite abelian group is a direct product of its Sylow p-subgroups.*

Proof (i) follows immediately from Sylow theorems.

(ii) The 'if' part is clear. We prove the 'only if part'. Let $|G| = p_1^{e_1} p_2^{e_2} \dots p_t^{e_t}$, and let G_i be the unique Sylow p_i-subgroup of G, see (i). In view of Theorem 1.6.3(ii), it suffices to prove that (a) $G = G_1 G_2 \dots G_t$,

and (b) $G_i \cap G_1 G_2 \dots G_{i-1} = \{1\}$ for all $1 < i \leq t$. But (a) holds because every $p_i^{e_i}$ divides the order of the group $G_1 G_2 \dots G_t$. For (b), note first of all that the elements of G_j and G_k commute as long as $j \neq k$. Now pick an element $g \in G_i \cap G_1 G_2 \dots G_{i-1}$. On one hand, g is a p_i-element. On the other hand, $g = g_1 \dots g_{i-1}$ for $g_j \in G_j$. As the elements g_j commute, g is a p'-element, i.e. its order is prime to p. It follows that $g = 1$. □

Lemma 1.7.13 *Let p be a prime, G be a finite group, and $|G| = p^e m$ with $(p, m) = 1$, $p^e \not| (m-1)!$. Then G is not simple.*

Proof In view of Theorem 1.5.21, we may assume that $m > 1$. Let P be a Sylow p-subgroup of G. The action og G on G/P gives a homomorphism from G to S_m, with the kernel contained in P, see Example 1.5.8. If G is simple, this implies that the kernel is trivial, whence $p^e m$ divides $m!$, or p^e divides $(m-1)!$, giving a contradiction. □

Corollary 1.7.14 *A_5 is a non-abelian simple group of minimal possible order.*

Proof A_5 has order 60. Now, 30, 40 and 56 are the only numbers between 2 and 59, which are not primes and which cannot be rules out using Lemma 1.7.13.

Let $r_p = r_p(G)$. If G is simple, then $r_p > 1$ for every p, as otherwise the Sylow p-subgroup is normal, see Proposition 1.7.12(i). We will also use repeatedly that r_p divides G and $r_p \equiv 1 \pmod{p}$.

Assume $|G| = 30$. Then the order of the Sylow 5-subgroup is 5, and $r_5 = 6$. It follows that the union of Sylow 5-subgroups has 24 non-trivial elements. The order of the Sylow 3-subgroup is 3, and $r_3 = 10$. So the union of Sylow 3-subgroups has 20 non-trivial elements. Contradiction.

Let $|G| = 40$. Then $r_5 = 1$. Contradiction.

Finally, let $|G| = 56$. Then r_7=8. So the union of Sylow 7-subgroups has 48 non-trivial elements. Moreover, $r_2 \geq 1$, so we get at least 9 more elements. Contradiction. □

We now obtain further general properties of Sylow p-subgroups.

Proposition 1.7.15 *Let P be a Sylow p-subgroup of G, and H be a subgroup of G containing $N_G(P)$. Then $N_G(H) = H$.*

Proof Let $g \in N_G(H)$. Then P and gPg^{-1} are both Sylow p-subgroups of H. By the Second Sylow Theorem applied to H, there exists $h \in H$ with $hgPg^{-1}h^{-1} = P$. So $hg \in N_G(P) \leq H$, whence $g \in H$. □

The above proposition should be contrasted with:

Proposition 1.7.16 *Let H be a p-subgroup of a finite group G, which is not a Sylow p-subgroup. Then $N_G(H) \gneq H$.*

Proof By assumption p divides $[G : H]$. If p does not divide $[G : N_G(H)]$ then clearly $N_G(H) \gneq H$, and we are done. So we may assume that $p|[G : N_G(H)]$.

Let $\mathcal{O}$ be set of subgroups conjugate to H. Then $|\mathcal{O}| = [G : N_G(H)]$. The group H act on $\mathcal{O}$ by conjugation, and has a trivial orbit $\{H\}$. As all orbits of H have sizes powers of p, we now see that there must be at least p trivial orbits. Let $\{K\} \neq \{H\}$ be another trivial orbit. We have $H \leq N_G(K)$, so $N_G(K) \gneq K$. As H and K are conjugate, this implies $N_G(H) \gneq H$. □

Corollary 1.7.17 *In a p-group, maximal proper subgroups have index p, and every subgroup of index p is normal.*

Example 1.7.18 There is no simple group of order $2^3 \cdot 3^3 \cdot 11$.

Indeed, let G be such a simple group. By the Third Sylow Theorem, $r_{11}(G) = 12$. Let P_{11} be a Sylow 11-subgroup, $N = N_G(P_{11})$, and $C = C_G(P_{11})$. Then $|N| = 2 \cdot 3^2 \cdot 11$. Moreover, N/C embeds into $\mathrm{Aut}(P_{11}) \cong C_{10}$, see Lemma 1.2.2. As 5 is not a divisor of $|G|$, it follows that $|N/C| = 1$ or 2. In particular, 3^2 divides $|C|$.

Let P be a Sylow 3 subgroup of C, and $H := N_G(P)$. We have $|P| = 3^2$. Note that $P_{11} \leq C_G(P) \leq H$. So $|H|$ is divisible by 11. Next, let P_3 be a Sylow 3-subgroup of G containing P. In view of Corollary 1.7.17, we have $P \lhd P_3$, and so $P_3 \leq H$. Hence 3^3 divides $|H|$. It follows that the index of H in G is at most 8. Thus, G embeds into S_n for $n \leq 8$, which contradicts the fact that 11 is a divisor of $|G|$.

Example 1.7.19 Let us classify groups of order pq, where p and q are primes. If $p = q$, then the group is abelian in view of Corollary 1.5.22. Then the group is either C_{p^2} or $C_p \times C_p$ (to see this, consider two cases: when G has an element of order p^2, and when all non-trivial elements have order p). So let $p > q$, say. By the Third Sylow Theorem, the Sylow

p-subgroup of G is normal. Moreover, if $p \not\equiv 1 \pmod{q}$, the Sylow q-subgroup is also normal, and so $G \cong C_p \times C_q \cong C_{pq}$ by Proposition 1.7.12 and Example 1.6.2.

Finally, let $p \equiv 1 \pmod{q}$. Of course, it is still possible that $G \cong C_p \times C_q \cong C_{pq}$. However, we have already seen in Example 1.6.11 that there also exists a non-abelian group of order pq. In fact, the group described in that example is the only one up to isomorphism. Indeed, the First Sylow Theorem and Theorem 1.6.5 imply that $G \cong C_p \rtimes C_q$. But we need to be careful about the following problem: semidirect product $H \rtimes K$ in principle depends on the action of K on H. However, we claim that in our situation this dependence is not essential in the sense that any two choices of non-trivial actions lead to isomorphic semi-direct products (and the choice of the trivial action leads, of course, to the direct product, which is the abelian case already taken care of). As C_{p-1} has only one subgroup of order q, there are exactly $q-1$ non-trivial homomorphisms from C_q to $\mathrm{Aut}(C_p) \cong C_{p-1}$, all with the same image, the last fact being the key. Therefore, if $\varphi : C_q \to \mathrm{Aut}(C_p)$ is one such homomorphism, then any other looks like $\varphi_j : C_q \to \mathrm{Aut}(C_p),\ c \mapsto \varphi(c^j)$ for some $1 \leq j < q$. Now, the desired isomorphism $C_p \times_{\varphi_j} C_q \to C_p \times_{\varphi} C_q$ is given by the formula $(d, c) \mapsto (d, c^j)$.

1.8 Jordan-Hölder Theorem

Definition 1.8.1 A *normal series* of a group G is a finite sequence of subgroups

$$G = G_0 \geq G_1 \geq \cdots \geq G_n = \{1\}$$

such that $G_i \trianglelefteq G$ for all i. A *subnormal series* of a group G is a finite sequence of subgroups

$$G = G_0 \geq G_1 \geq \cdots \geq G_n = \{1\}$$

such that $G_{i+1} \trianglelefteq G_i$ for all i. The *factor groups* of the subnormal series above are the groups $G_0/G_1, G_1/G_2, \ldots, G_{n-1}/G_n$. The *length* of a subnormal series is the number of non-trivial factor groups. A *composition series* is a subnormal series, all of whose non-trivial factor groups are simple. The non-trivial factor groups of a composition series are called *composition factors* of G.

Remark 1.8.2 An infinite group might not have a composition series, as the example of the infinite cyclic group C_∞ shows. However:

Theorem 1.8.3 (Jordan-Hölder) *Every finite group G has a composition series. Moreover, any two composition series have the same composition factors up to the order of their appearance. In particular, the length of a composition series is an invariant of G.*

We will skip the proof of this theorem, as it is a little technical. Later in the course we will prove a similar result for modules.

Example 1.8.4 The Fundamental Theorem of Arithmetic follows from Jordan-Hölder theorem applied to $G = C_n$.

Example 1.8.5 Non-isomorphic groups can have the same composition factors: consider $C_2 \times C_2$ and C_4. This is really the extension problem again, cf. Remark 1.6.12.

Example 1.8.6 $S_4 > A_4 > V_4 > C_2 > \{1\}$ is a composition series of S_4 with composition factors C_2, C_3, C_2.

Example 1.8.7 The group $C_6 \cong C_2 \times C_3$ has two composition series: $C_6 > C_3 > \{1\}$ and $C_6 > C_2 > \{1\}$.

Example 1.8.8 It follows from Corollary 1.5.26 that all composition factors of a p-group are isomorphic to C_p.

Example 1.8.9 We claim that $GL_2(\mathbb{F}_4) > SL_2(\mathbb{F}_4) > \{1\}$ is a composition series of $GL_2(\mathbb{F}_4)$. We just have to show that $SL_2(\mathbb{F}_4)$ is simple. We claim that in fact

$$SL_2(\mathbb{F}_4) \cong A_5.$$

Well, in view of (1.4), the order of $SL_2(\mathbb{F}_4)$ is $60 = |A_5|$. So it suffices to embed $SL_2(\mathbb{F}_4)$ into A_5. This embedding comes from the action of $SL_2(\mathbb{F}_4)$ on the five lines of the 2-dimensional space $\mathbb{F}_4^2$. It is easy to see that the action is faithful, for to act trivially on all lines, the matrix must be scalar, but there are no non-trivial scalar matrices in $SL_2(\mathbb{F}_4)$. There are many ways to see that $SL_2(\mathbb{F}_4)$ acts on the lines with even permutations. The easiest one is to notice that A_5 is the only subgroup of S_5 having 60 elements.

1.9 Solvable and Nilpotent Subgroups

There is an awful lot of groups. It is impossible to imagine a classification of all groups and even all finite groups. Jordan-Hölder theorem shows that any finite group can be built out of simple ones using group extensions. In some sense this reduces the study of finite groups to, first, understanding finite simple groups, and, second, trying to understand the building process, or extension problem. Both are extremely hard tasks. Simple groups are only simple by name, but at least they can be classified, which we will discuss a little later. The extension problem in full generality really seems to be out of reach, as is indicated by the following circumstance. Let us consider only the groups whose composition factors are as easy as possible, namely cyclic groups of prime order. One might expect to get some benign class of groups. However, it turns out that the class of groups obtained in this way is very rich and complicated. For example, there seems to be no chance that such groups can be classified.

Definition 1.9.1 A finite group is called *solvable* if all its composition factors are cyclic (of prime order).

Lemma 1.9.2 *Every quotient group and every subgroup of a solvable group is solvable. Every extension of a solvable group by a solvable group is solvable.*

Proof The result for extensions follows from the Correspondence Theorem.

Let $N \trianglelefteq G$, and we have to prove that G/N is solvable. Choose a composition series for G/N, lift its terms to subgroups containing N, using Correspondence Theorem, then choose a composition series for N, and join the two series to get a composition series for G. We know that composition factors for G are cyclic. Now, by uniqueness of composition factors, composition factors of G/N are also cyclic.

Let H be a subgroup of G, and

$$G = G_0 > G_1 > \cdots > G_t = \{1\}$$

be a composition series of G. Consider the series

$$H = H \cap G_0 \geq H \cap G_1 \geq \cdots \geq H \cap G_t = \{1\} \tag{1.7}$$

Then $H \cap G_i \trianglelefteq H \cap G_{i-1}$, and, using the Second Isomorphism Theorem,

we get

$$\frac{H \cap G_{i-1}}{H \cap G_i} = \frac{H \cap G_{i-1}}{(H \cap G_{i-1} \cap G_i)} \cong \frac{G_i(H \cap G_{i-1})}{G_i} \subseteq \frac{G_{i-1}}{G_i}.$$

So the factors groups of the series (1.7) are either trivial or cyclic. □

Example 1.9.3 By Example 1.8.8, any p-group is solvable.

Example 1.9.4 Let B be the subgroup of all upper triangular matrices in $GL_n(\mathbb{F}_q)$ (here diagonal entries do not have to equal 1). Then B is solvable, as it is an extension of the p-group U of all upper unitriangular matrices by the abelian group $\mathbb{F}_q^\times \times \cdots \times \mathbb{F}_q^\times$.

Example 1.9.5 In view of Example 1.8.6, A_4 is solvable.

Definition 1.9.6 If H and K are subgroups of a group G, we denote by $[H, K]$ the subgroup of G generated by all commutators $[h, k] = hkh^{-1}k^{-1}$ for $h \in H$, $k \in K$. The *commutator subgroup* G' of a group G is defined to be $[G, G]$.

The following result show that G' can be characterized as the smallest normal subgroup of G such that G/G' is abelian.

Proposition 1.9.7 *Let G be a group.*

(i) *$G' \trianglelefteq G$, and G/G' is abelian.*
(ii) *If $H \trianglelefteq G$ and G/H is abelian, then $G' \leq H$.*

Proof (i) follows from the properties

$$[x, y]^{-1} = [y, x], \quad g[x, y]g^{-1} = [gxg^{-1}, gyg^{-1}],$$

in G and the property $[xG', yG'] = [x, y]G'$ in G/G'.
(ii) If H is as in the assumption, then $[x, y] \in H$ for any $x, y \in G$. □

Example 1.9.8 (i) If G is a non-abelian simple group then $G' = G$.
(ii) If G is abelian then $G' = \{1\}$.
(iii) $S_n' = A_n$.

Definition 1.9.9 The *derived series* of G is

$$G = G^{(0)} \geq G^{(1)} \geq G^{(2)} \geq \cdots \geq G^{(i)} \geq \ldots,$$

where $G^{(0)} := G$, and $G^{(i+1)} = (G^{(i)})'$ for $i \geq 0$.

Every $G^{(i)}$ is clearly a *characteristic subgroup*, i.e. a subgroup invariant with respect to every automorphism of G. So $G^{(i)} \trianglelefteq G$.

The following proposition may be thought of as the statement that solvable groups are in some sense close to abelian groups. In this respect the non-abelian simple groups show a completely opposite behavior, see Example 1.9.8(i).

Proposition 1.9.10 *A finite group G is solvable if and only if $G^{(n)} = \{1\}$ for some n.*

Proof If G is solvable, it is easy to see that $G^{(i)} \leq G_i$ for any composition series $G = G_0 \geq G_1 \geq \ldots$. Conversely a finite derived series ending on $\{1\}$ is a normal series with abelian factor groups. Such series can be refined to a composition series with cyclic factor groups using Correspondence Theorem. □

Remark 1.9.11 There are two remarkable theorems on solvable groups, which should be mentioned at this point:

(i) Burnside's $p^a q^b$ Theorem says that if p and q are two primes and G is a group of order $p^a q^b$, then G is solvable. As the example of A_5 shows the theorem cannot be improved to three primes. This theorem will be proved using group representation theory later in this course.

(ii) The Odd Order Theorem of Feit and Thompson says that a group of odd order is solvable. This is equivalent to the fact that a non-abelian simple group has even order (why?). The Odd Order Theorem is very hard to prove (the proof takes a whole volume of the *Pacific Journal of Mathematics*, and has not been dramatically simplified since it originally appeared in 1963). It has played a crucial role in the classification of finite simple groups.

Example 1.9.12 Here is a special case of Burnside's Theorem which we can treat with bare hands. Let p and q be primes. Then any group of order $p^2 q$ is solvable.

Well, if $p = q$ we are done by Example 1.9.3, and if $p > q$, then there exists only one Sylow p-subgroup by the Third Sylow theorem, which is normal. Let $p < q$. By the Third Sylow Theorem, $r_q = 1$ or p^2. If $r_q = 1$, we are done. If $r_q = p^2$, then the Sylow q-subgroups comprise $p^2(q-1)+1$ elements, whence the remaining $p^2 - 1$ elements are exactly the p-elements of our group. It follows that there is only one Sylow p-subgroup.

Example 1.9.13 A key fact established in the proof of the general Burnside's Theorem is that no conjugacy class in a finite simple group has order a prime power > 1. Given this fact, it is not difficult to finish the proof.

Indeed, apply induction on $a+b$ to prove that a group G of order $p^a q^b$ is solvable. If $a + b = 1$, we are fine. Let $a + b > 1$. We may assume that $a \geq 1$ and $b \geq 1$ in view of Example 1.9.3. If G has a non-trivial proper normal subgroup N, then apply inductive hypothesis to N, G/N and use Lemma 1.9.2.

So we may assume that G is simple. Consider a Sylow p-subgroup $P < G$. It has a non-trivial center. Let z be a non-trivial element of this center, and set $C := C_G(z)$. Clearly $P \leq C$. So $[G : C]$ is a power of q. But $[G : C]$ is the size of the conjugacy class of x, so by the key fact mentioned above, we have $C = G$. Tus $z \in Z(G)$, which contradicts the simplicity of G.

Now we work to introduce an important subclass of solvable groups, called nilpotent groups.

Definition 1.9.14 The *descending central series* of a group G is

$$G = \gamma_0(G) \geq \gamma_1(G) \geq \ldots,$$

where $\gamma_{i+1}(G) = [\gamma_i(G), G]$ for any $i \geq 1$. We say that the descending central series *terminates* if $\gamma_n(G) = \{1\}$ for some n.

A group is called *nilpotent* if its descending central ceries terminates.

Clearly, nilpotent groups are solvable, but there exist solvable groups which are not nilpotent.

Example 1.9.15 (i) Consider the group B of upper triangular matrices over $\mathbb{F}_q$ from Example 1.9.4. Assume $q > 2$. Then $\gamma_1(B) = [B, B] = U$, where U is the group of upper unitriangular matrices. On the other hand $\gamma_2(B) = [U, B] = U$. So $\gamma_n(U) = U$ for any $n \geq 1$, and so B is solvable but not nilpotent.

(ii) S_4 is solvable but not nilpotent, as $[S_4, A_4] = A_4$.

There exists another central series of G:

Definition 1.9.16 The *ascending central series* of a group G is

$$\{1\} = Z_0(G) \leq Z_1(G) \leq Z_2(G) \leq \ldots,$$

where Z_{i+1} is defined from $Z_{i+1}(G)/Z_i(G) = Z(G/Z_i(G))$. We say that the ascending central series *terminates* if $Z_n(G) = G$ for some n.

It is easy to see that descending and ascending central series are normal series. Terminating descending and ascending central series are special cases of a general terminating central series:

Definition 1.9.17 A finite normal series

$$\{1\} = G_0 \leq G_1 \leq \cdots \leq G_n = G,$$

such that $G_i/G_{i-1} \leq Z(G/G_{i-1})$ for all $1 \leq i \leq n$ is called a *terminating central series* of G.

It turns out that nilpotent groups can be defined in terms of various central series:

Lemma 1.9.18 *The following condition on a group G are equivalent:*

(i) *The descending central series of G terminates.*
(ii) *The ascending central series of G terminates.*
(iii) *G has a terminating central series.*

Proof In this proof we will write Z_i for $Z_i(G)$ and γ_i for $\gamma_i(G)$. It is easy to see that (i) implies (iii) and (ii) implies (iii). We now assume that

$$\{1\} = G_0 \leq G_1 \leq \cdots \leq G_n = G,$$

is a terminating central series of G.

(iii) $\Rightarrow$ (ii) It suffices to show that $G_i \subseteq Z_i$ for $i = 0, 1, \ldots, n$. Apply induction on i, the case $i = 0$ being clear. Assume $i > 0$ and $G_{i-1} \subseteq Z_{i-1}$. Then we have a natural surjection

$$\pi : G/G_{i-1} \to G/Z_{i-1}$$

with $\pi(G_i/G_{i-1}) = (G_i Z_{i-1})/Z_{i-1}$. Moreover, note that

$$G_i/G_{i-1} \subseteq Z(G/G_{i-1}),$$

and so its image under the surjective homomorphism π is central:

$$(G_i Z_{i-1})/Z_{i-1} \leq Z(G/Z_{i-1}) = Z_i/Z_{i-1}.$$

It follows that $G_i Z_{i-1} \leq Z_i$, whence $G_i \leq Z_i$.

(iii) $\Rightarrow$ (i) It suffices to show that $\gamma_i \subseteq G_{n-i}$ for $i = 0, 1, \ldots, n$. Apply induction on i, the case $i = 0$ being clear. Assume $i > 0$ and

$\gamma_{i-1} \subseteq G_{n-i+1}$. Let $x \in G$ and $y \in \gamma_{i-1} \subseteq G_{n-i+1}$. Then $[x, y] \in G_{n-i}$, since in G/G_{n-i} we have $[xG_{n-i}, yG_{n-i}] = 1$. As γ_i is generated by all $[x, y]$ as above, we have $\gamma_i \subseteq G_{n-i}$. □

Corollary 1.9.19 *Every non-trivial nilpotent group has a non-trivial center.*

Proposition 1.9.20 *Every subgroup and quotient group of a nilpotent group is nilpotent. Direct product of nilpotent groups is nilpotent. Moreover, if $N \trianglelefteq Z(G)$ and G/N is nilpotent then G is nilpotent.*

Proof Let $H \leq G$. It is clear that $\gamma_i(H) \leq \gamma_i(G)$. So the result on subgroups follows. For the quotient G/K, it suffices to observe that $\gamma_i(G/K) = (\gamma_i(G)K)/K$. For direct products note that $\gamma_i(G \times H) = \gamma_i(G) \times \gamma_i(H)$.

Finally, let $N \trianglelefteq Z(G)$ and G/N be nilpotent. Define $Z_i \geq N$ from $Z_i(G/N) = Z_i/N$. By assumption, $Z_n(G/N) = G/N$ for some n. Now,

$$G = Z_n \geq Z_{n-1} \geq \cdots \geq Z_0 = N \geq \{1\}$$

is a terminating central series of G. □

A rich source of nilpotent groups is provided by the following:

Proposition 1.9.21 *Every finite p-group is nilpotent.*

Proof Induction on $|G|$, using Theorem 1.5.21 and Proposition 1.9.20. □

Example 1.9.22 If n is a power of 2, then D_n is a 2-group, and hence nilpotent. However, in all other cases D_n is not nilpotent. This follows from the fact that the center of D_n is trivial when $n/2$ is odd, while $D_n/Z(D_n) \cong D_{n/2}$ when $n/2$ is even.

Using Proposition 1.9.20 we can construct lots of nilpotent groups by forming direct products of p-groups. It turns out that all finite nilpotent groups are obtained in this way:

Theorem 1.9.23 *For a finite group G the following conditions are equivalent:*

(i) *G is nilpotent.*
(ii) *Every Sylow p-subgroup of G is normal.*

(iii) *G is a direct product of p-groups.*

Proof (ii) and (iii) are equivalent thanks to Proposition 1.7.12(ii).

(i) $\Rightarrow$ (ii) Let G be nilpotent. First we show that $Z_i(G) \leq H \leq G$ implies $Z_{i+1}(G) \leq N_G(H)$. Let $z \in Z_{i+1}(G)$ and $h \in H$. As $zZ_i \in Z(G/Z_i)$, we have $hzZ_i = zhZ_i$ or $h^{-1}z^{-1}hz \in Z_i \leq H$, whence $z^{-1}hz \in H$. Therefore $z \in N_G(H)$.

Now let P be a Sylow p-subgroup of G. Then $N_G(P) = N_G(N_G(P))$ by Proposition 1.7.15. We have $\{1\} = Z_0(G) \leq N_G(P)$. By the previous paragraph, $Z_i(G) \leq N_G(P)$ for all i. Taking i sufficiently large, we get $N_G(P) = G$.

(iii) $\Rightarrow$ (i) by Propositions 1.9.21 and 1.9.20. □

We finish this section with several agreeable properties of nilpotent groups. None of this holds even for solvable groups.

Corollary 1.9.24 *If G is a finite nilpotent group and d is a divisor of $|G|$, then G has a normal subgroup of order d.*

Proof Follows from Theorem 1.9.23 and Corollary 1.5.26. □

Proposition 1.9.25 *If N is a non-trivial normal subgroup of a nilpotent group G, then $N \cap C(G) \neq \{1\}$.*

Proof As $Z_m(G) = G$ for some m, there exists $m \geq 1$ with $Z_m(G) \cap N \neq \{1\}$ and $Z_{m-1}(G) \cap N = \{1\}$. Take some non-trivial $z \in Z_m(G) \cap N$. Then for any $g \in G$ we have $[g, z] \in Z_{m-1}(G)$, $[g, z] \in N$ (as N is normal). Hence $[g, z] = 1$, i.e. $z \in Z(G)$. □

Here is another characterization of nilpotent groups:

Proposition 1.9.26 *A finite group is nilpotent if and only if every maximal proper subgroup of G is normal.*

Proof Let M be a maximal subgroup of G. By Theorem 1.9.23, $G = S_1 \times \cdots \times S_l$, a direct product of p-groups. Let $|S_i| = p_i^{e_i}$. If P_i is the Sylow p_i-subgroup of M, then $P_i \leq S_i$. It follows that $M = P_1 \times \cdots \times P_l$. Now the 'only-if' part follows from the analogous property for p-groups proved in Corollary 1.7.17.

Conversely, let every maximal proper subgroup of G be normal. Consider a Sylow p-subgroup $P \leq G$. In view of Theorem 1.9.23, it suffices

to prove that $N := N_G(P) = G$. If $N \lneq G$, then there exists a maximal proper subgroup $M < G$ containing N. By assumption we have $N_G(M) = G$, which contradicts Proposition 1.7.15. □

1.10 More on simple groups: projective unimodular groups

The only simple groups we have seen so far are cyclic groups of prime order and the alternating groups A_n for $n \geq 5$. We will now show that certain groups of matrices are also simple.

Let F be an arbitrary field. Consider the group $SL_n(F)$ of all n by n matrices over F with determinant 1. The most important case will be when $F = \mathbb{F}_q$ is a finite field.

Definition 1.10.1 The (i, j) *matrix unit* $E_{i,j}$ is the matrix with 1 in the position (i, j) and 0's elsewhere. We write I_n for the identity $n \times n$ matrix. A transvection is a matrix of the form

$$t_{i,j}(a) := I_n + aE_{i,j} \qquad (a \in F,\ 1 \leq i \neq j \leq n).$$

Proposition 1.10.2 *The group $SL_n(F)$ is generated by transvections.*

Proof This is essentially equivalent to Gaussian elimination. One just needs to observe that multiplying an $n \times n$ matrix A on the left by $t_{ij}(a)$ leads to the elementary row transformation which adds to the ith row of A the jth row of A multiplied by a, and multiplying A by $t_{ij}(a)$ on the right leads to a similar elementary column transformation. But be careful, as the operation of multiplying rows or columns with scalars is not available! □

Lemma 1.10.3 *The center of $SL_n(F)$ consists of the scalar matrices in $SL_n(F)$.*

Proof The scalar matrices are clearly central. Conversely, let $A \in Z(SL_n(F))$. Then $At_{ij}(1) = t_{ij}(1)A$ or, equivalently, $AE_{ij} = E_{ij}A$. However, $E_{ij}A$ is the matrix whose ith row coincides with the jth row of A and other rows are zero, while Ae_{ij} is the matrix whose jth column coincides with the ith column of A and other columns are zero. This implies that $a_{ii} = a_{jj}$ and that all other elements in the ith row and jth column of A are zero. □

Definition 1.10.4 The group

$$PSL_n(F) := SL_n(F)/Z(SL_n(F))$$

is called a *projectuve unimodular group.*

Lemma 1.10.5 *We have* $|PSL_n(\mathbb{F}_q)| = q^{n(n-1)/2}\prod_{i=2}^{n}(q^i-1)/(n,q-1)$.

Proof In view of (1.4) and Lemma 1.10.3, we just have to observe using Lemmas 1.2.5 and 1.2.1 that the number of solutions of equation $x^n = 1$ in $\mathbb{F}_q^\times$ is $(n, q-1)$. □

Example 1.10.6 We have $|PSL_2(\mathbb{F}_2)| = 6$, so $PSL_2(\mathbb{F}_2)$ is solvable (why?). Also, $|PSL_2(\mathbb{F}_3)| = 12$, so $PSL_2(\mathbb{F}_3)$ is solvable (why?). However, we will soon prove that all other projective unimodular groups are simple.

The following property will play an important role in proving simplicity of $PSL_n(\mathbb{F}_q)$.

Proposition 1.10.7 *Let* $G = PSL_n(\mathbb{F}_q)$. *Then* $G' = G$, *unless* $n = 2$ *and* $q \leq 3$.

Proof It will suffice if we prove the result with SL in place of PSL. Denote by $\mathrm{diag}(b_1, \ldots, b_n)$ the diagonal matrix with entries $b_1, \ldots, b_n$ down the diagonal. It is easy to check that

$$[t_{ik}(a), t_{kj}(b)] = t_{ij}(ab) \qquad (i, j, k \text{ all distinct}), \quad (1.8)$$

$$[t_{ij}(a), \mathrm{diag}(b_1, \ldots, b_n)] = t_{ij}(ab_i/b_j - a) \qquad (i \neq j) \quad (1.9)$$

for all $a, b \in F$, $b_1, \ldots, b_n \in F^\times$. By (1.8), G' contains all transvections for $n > 2$. Now from (1.9) we see that G' also contains all transvections if $n = 2$ and $q > 3$. It remains to use Lemma 1.10.2. □

In order to establish the simplicity of $PSL_n(\mathbb{F}_q)$ (in most cases), we are going to use somewhat more advanced properties of permutation groups.

Definition 1.10.8 Let X be a set and G be a transitive permutation group on X. The group G also acts on the subsets of X. A partition

$$X = \sqcup_{i\in I} X_i$$

of X is called a *partition into imprimitivity blocks* if G permutes the subsets X_i, i.e. for every $g \in G$ and $i \in I$ there exists $j \in I$ with $gX_i = X_j$. Of course, there always exist two trivial partitions into imprimitivity blocks: the partition with just one block X and the partition with the one-element blocks. If there are no non-trivial partitions into imprimitivity blocks, the permutation group is called *primitive*. Otherwise it is called *imprimitive*.

Remark 1.10.9 The idea behind the notion of primitivity is that many questions about permutation groups can be reduced to primitive groups, so primitive groups are in some sense building blocks for arbitrary permutation groups. Indeed, let us start from an arbitrary permutation group G on a finite set X. First of all 'we may assume' that G is transitive because otherwise we just split X into the G-orbits and then study each of them separately. Further we often 'may assume' that G is primitive because otherwise the action of G can be understood from its action on the set of imprimitivity blocks on one hand, and the action of the stabilizer G_{X_i} of any block X_i on the elements of X_i on the other hand. Both actions are on smaller sets, so we can repeat this argument until we reach primitive permutation groups.

Example 1.10.10 (i) The action of S_n on $\{1, 2, \dots, n\}$ is primitive.

(ii) D_8 acts imprimitively on the vertices of a square: the blocks of imprimitivity are pairs of opposite vertices.

(iii) Let $m, n \in \mathbb{Z}_{>1}$, and $X = \{1, 2, \dots, mn\}$. Consider the partition $X = \sqcup_{i=1}^{n} X_i$, where $X_i = \{mi + 1, mi + 2, \dots, m(i + 1)\}$. Let G be the subgroup in S_{mn} which consists of all permutations g preserving this partition—this means that $gX_i = X_j$ for every i. By definition, G is imprimitive with imprimitivity blocks $X_1, \dots, X_n$. This groups is called the *wreath product* of S_m with S_n and denoted $S_m \wr S_n$. We have

$$S_m \wr S_n \cong \underbrace{(S_m \times \cdots \times S_m)}_{n \text{ copies}} \rtimes S_n$$

with S_n acting on the n-tuples in $S_m \times \cdots \times S_m$ by place permutations.

The following fact is fundamental:

Theorem 1.10.11 *Let G be a transitive permutation group on X. Then G is primitive if and only if G_x is a maximal subgroup for any $x \in X$.*

Proof Assume that G is imprimitive. Let H be the stabilizer of an imprimitivity block X_i. As G is transitive, H is a proper subgroup of G. Take $x \in X_i$. Clearly, $G_x \leq H$. Moreover, G_x must be strictly contained in H—otherwise we would never be able to move x to other elements of X_i, contrary to the transitivity again.

Conversely, assume that G_x is not maximal. Let $G_x \lneq H \lneq G$. Write $G = g_1H \sqcup \cdots \sqcup g_lH$, and define $X_i := \{g_ih \cdot x \mid h \in H\}$. We claim that $X = \sqcup_{i=1}^{l} X_i$. Indeed, $X = \cup_{i=1}^{l} X_i$ by transitivity. On the other hand, if $g_ih \cdot x = g_jh' \cdot x$, it follows that $h^{-1}g_i^{-1}g_jh' \in G_x < H$, whence $g_i^{-1}g_j \in H$, and so $i = j$. Now, G obviously permutes X_i's, and so it just remains to notice that the blocks are non-trivial. □

We will also need the following easy observation:

Lemma 1.10.12 *Every non-trivial normal subgroup N of a primitive group G is transitive.*

Proof Note that the N-orbits are permuted by G. □

Definition 1.10.13 Let G be a permutation group on X. The group is called *k-transitive* if for every pair of k-tuples of distinct elements $(x_1, \ldots, x_k)$ and $(y_1, \ldots, y_k)$ in X^k there exists $g \in G$ with $g \cdot x_i = y_i$ for all $1 \leq i \leq k$.

Of course 1-transitive is the same as transitive. On the other hand, the properties of being 2-transitive, 3-transitive, and so on, show how 'rich' the permutation group is in a certain sense.

Lemma 1.10.14 *If G is 2-transitive on X then it is primitive on X.*

Proof If G is imprimitive, take a pair (x_1, x_2) with x_1, x_2 being in the same imprimitivity block, and a pair (y_1, y_2) with x_1, x_2 being in different imprimitivity blocks. Now G cannot 'move' (x_1, x_2) into (y_1, y_2), so it is not 2-transitive. □

Example 1.10.15 (i) The group $SL_n(\mathbb{F}_q)$ acts naturally on the lines of the vector space $\mathbb{F}_q^n$. It follows from linear algebra that this action is 2-transitive, and so it is primitive by the above lemma.

Finally, note that under our action the center of $SL_n(\mathbb{F}_q)$ acts trivially, see Lemma 1.10.3. So the action factors through to a primitive action of $PSL_n(\mathbb{F}_q)$.

(ii) Let $1 < m < n-1$, and G be the symmetric group S_n acting on the m-element subsets of $\{1,\dots,n\}$. Then G is not 2-transitive. Indeed, take a pair of subsets (X_1, X_2) such that $|X_1 \cap X_2| = m-1$ and a pair of subsets (Y_1, Y_2) such that $|Y_1 \cap Y_2| = m-2$. Then G cannot move (X_1, X_2) into (Y_1, Y_2). On the other hand, the action is primitive, unless $m = n/2$, as the point stabilizer is the maximal subgroup $S_m \times S_{n-m} < S_n$.

The following technical proposition will be used to verify simplicity of $PSL_n(\mathbb{F}_q)$.

Proposition 1.10.16 *Let G be a permutation group with the following properties:*

(i) *G is primitive;*
(ii) *$G = G'$;*
(iii) *the stabilizer G_x of some element $x \in X$ contains an abelian normal subgroup A such that its conjugates generate G.*

Then G is simple.

Proof Let $N \neq \{1\}$ be a normal subgroup of G. We have to show that $G = N$. Note first that $G = NG_x$, as N is transitive in view of Lemma 1.10.12.

Now, we claim that $G = NA$. Indeed, by the assumption (iii), every $g \in G$ can be written as

$$g = (g_1 a_1 g_1^{-1}) \dots (g_l a_l g_l^{-1}) \qquad (g_i \in G,\ a_i \in A).$$

As $G = NG_x = G_x N$ and $A \triangleleft G_x$, we may assume that every g_i belongs to N. Now the claim follows, as $AN = NA$.

Finally, let $a_1, a_2 \in A$ and $n_1, n_2 \in N$. Then using the normality of N and commutativity of A,

$$\begin{aligned} [n_1 a_1, n_2 a_2] &= n_1 a_1 n_2 a_2 a_1^{-1} n_1^{-1} a_2^{-1} n_2^{-1} \\ &= (n_1)(a_1 n_2 a_1^{-1})(a_2 n_1^{-1} a_2^{-1})(n_2^{-1}) \in N. \end{aligned}$$

So $[NA, NA] \leq N$, which completes the proof, since $G = [G,G] = [NA, NA]$ by (ii) and the previous paragraph. □

Now we can achieve our goal:

Theorem 1.10.17 *The group $G = PSL_n(\mathbb{F}_q)$ is simple unless $n = 2$ and $q \leq 3$.*

Proof We verify the conditions (i)-(iii) of Proposition 1.10.16. According to Example 1.10.15(i), G acts primitively on the lines in $\mathbb{F}_q^n$. This gives us (i), and (ii) comes from Proposition 1.10.7.

For (iii), take x to be the line in $\mathbb{F}_q^n$ spanned by the vector $(1, 0, \ldots, 0)$. The stabilizer G_x consists of all the cosets $gZ(G)$ where $g \in SL_n(\mathbb{F}_q)$ is a matrix with entries $(g_{ij})_{1 \le i,j \le n}$ such that $g_{i1} = 0$ for any $i \ge 2$. These matrices look like

$$\begin{pmatrix} m & * \\ 0 & M \end{pmatrix}, \tag{1.10}$$

where M is an arbitrary non-degenerate $(n-1) \times (n-1)$ matrix, 0 stands for a column of zero entries, $m = (\det M)^{-1}$, and $*$ stands for a row of arbitrary entries. Now, take A to be the subgroup of G_x, which consists of (the cosets of) all the matrices of the form (1.10) such that M is the identity matrix. It is easy to see that A is abelian and normal in G_x. It remains to prove that the conjugates of A generate G.

In view of Lemma 1.10.2, it suffices to show that any $t_{ij}(a)$ is conjugate to some $t_{1k}(a)$ in $SL_n(\mathbb{F}_q)$. If $i = 1$, we are done. Let $i > 1$, and $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{F}_q^n$. Note that $t_{ij}(a)$ maps e_j to $e_j + ae_i$ and leaves e_k invariant for $k \ne j$. Now consider the new basis $\{e'_1, \ldots, e'_n\}$, where $e'_1 = e_i$, $e'_i = -e_1$, and $e'_j = e_j$ for $j \ne 1, i$. Then, $t_{ij}(a)$ maps e'_j to $e'_j + ae'_1$ and leaves e'_k invariant for $k \ne j$. So the matrix of our transvection in the new basis looks like $t_{1j}(a)$. By linear algebra, this means that in the group $GL_n(\mathbb{F}_q)$ we have $gt_{ij}(a)g^{-1} = t_{1j}(a)$, where g is the change of basis matrix. But g has determinant 1, and so we are done. □

Apart from $PSL_n(q)$, one can construct roughly three more families of finite simple groups which are labelled by two parameters n and q. In fact, just like $PSL_n(\mathbb{F}_q)$, these are described roughtly as some groups of $n \times n$ matrices over the field $\mathbb{F}_q$. Moreover, there are five more families of simple groups of matrices labelled only by a parameter q, which means that the sizes of the matrices is fixed in these cases. All such simple groups of matrices belong to the class of finite Chevalley groups.

Are there any finite simple groups other than alternating groups and finite Chevalley groups? Yes, five more examples were known since 19th century—these are Mathieu grous $M_{11}, M_{12}, M_{22}, M_{23}$, and M_{24}. The index k in M_k indicates that M_k is constructed as a permutation group on k symbols. (Note, by the way, the typical situation that the groups do not come from thin air, but appear either as groups of linear transformations/matrices or as permutation groups). Mathieu groups are quite

small, for example $|M_{11}| = 7920$. It took 100 years to discover more finite simple groups. Between 1965 (Zvonimir Janko) and 1982 (Robert Griess) various people discovered 21 new simple groups. Thus we have 26 'exceptional' finite simple groups, known as sporadic simple groups. The largest one discovered by Griess is enormous: it has order

$$808,017,424,794,512,875,886,459,904,$$
$$961,710,757,005,754,368,000,000,000$$

and is aptly called the *Monster* (there also exists a *Baby Monster*). What is even much more exciting, people were able to prove the Classification Theorem, which claims that the groups described above are actually *all* finite simple groups! This magnificent result is of great importance for mathematics. But is extremely difficult to prove. In fact the proof itself is *monstrous*: it takes 500 articles and about $15,000$ journal pages. At the moment, a *revision program* is under way. This will eventually result in having the proof checked and written neatly in 10 or so book volumes.

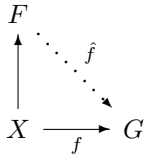
1.11 Generators and Relations

How can we describe a group. So far we mainly had groups appearing either as groups of permutations (symmetric and alternating groups) or as groups of matrices (general linear groups, special linear groups). Dihedral groups might look different, but they too have an explicit geometric realization. So we can say that most groups we have seen so far come as symmetry groups of certain known objects or systems. For example, permutation groups are symmetry groups of sets (on which they operate), linear groups are symmetry groups of vector spaces, dihedral groups are symmetry groups of geometric objects (n-gons). Perhaps only cyclic groups are different. Of course, they too can be described as groups acting somewhere as symmetries, but originally we just defined them as 'abstract groups' by 'multiplication properties' we want them to satisfy (one generator g, other elements look like g^i, and $g^n = 1$, or something like this). In this section we will make this philosophy precise by studying groups described in an abstract fashion using generators and relations.

The notion of a *free group* turns out to be the key. The free group on the set X may be thought of as the group given by generators X and no relations.

Definition 1.11.1 If X is a subset of a group F, then F is a *free group*

on X if for every group G and every map $f : X \to G$, there exists a unique group homomorphism $\hat{f} : F \to G$ with $\hat{f}(x) = f(x)$ for all $x \in X$, see the diagram below.



Remark 1.11.2 It is not so obvious that a free group on X exists. However, it is clear from the definition that if a free group on X exists, then it is *unique* ('up to unique isomorphism'): if F and F' are free groups on X, then there exists a unique isomorphism ψ from F to F' which is identity on X. Slightly more generally, let X and Y be two sets and $\varphi : X \to Y$ be a bijection. If $F(X)$ is a free group on X and $F(Y)$ is a free group on Y then there exists an isomorphism $\hat{\varphi} : F(X) \to F(Y)$, extending φ.

Now we work to construct a free group on X. We assume that X is non-empty (interpreting the free group on $\varnothing$ as the trivial group $\{1\}$ for convenience). Let X^{-1} be a set bijective to X but disjoint from it. Pick a bijection between X and X^{-1}, and denote the image of $x \in X$ under this bijection by x^{-1}. Thus,

$$X^{-1} = \{x^{-1} \mid x \in X\}.$$

Definition 1.11.3 If $n \in \mathbb{Z}_{>0}$, we define a *word* on X of length n to be a function $w : \{1, 2, \ldots, n\} \to X \sqcup X^{-1}$. If $w(i) = x_i^{\varepsilon_i}$, where $x_i \in X$, $\varepsilon_i = \pm 1$ $(1 \leq i \leq n)$, we write

$$w = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n},$$

and we denote the length n by $|w|$. We allow the empty word, denoted 1, whose length is 0 by convention. A *subword* of a word $w = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n}$ is either the empty word or a word of the form $x_r^{\varepsilon_r} x_{r+1}^{\varepsilon_{r+1}} \ldots x_s^{\varepsilon_s}$ for $1 \leq r \leq s \leq n$. The *inverse word* of $w = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n}$ is $w^{-1} = x_n^{-\varepsilon_n} \ldots x_1^{-\varepsilon_1}$. A word w is *reduced* if $w = 1$ or w has no subwords of the form xx^{-1} or $x^{-1}x$ for $x \in X$. Any two words can be *multiplied* by juxtaposing them (writing one after another, if you prefer normal English): if $w = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n}$ and $u = y_1^{\delta_1} \ldots y_n^{\delta_n}$, then $wu = x_1^{\varepsilon_1} \ldots x_n^{\varepsilon_n} y_1^{\delta_1} \ldots y_n^{\delta_n}$. We interpret $1w = w1 = w$.

Example 1.11.4 Let $X = \{x, y\}$. Then $w = x^{-1}xy$ is a word of length 3 on X (even if you believe it is more logical to think that its length is 1). The word w is not reduced. For the inverse word we have $w^{-1} = y^{-1}x^{-1}x$. The subwords of length 2 are $x^{-1}x$ and xy. Note that $x^{-1}y$ is not a subword.

Now, the free group on X should be something like the group generated by X with as few relations as possible ('free' is for 'free from relations'!). So the elements of the group will be words on X with multiplication given by juxtaposition and the empty word 1 being the identity element. Of course, relations like $xx^{-1} = 1$ hold in every group, so we will have to tolerate those as necessary evil. Thus, we must identify the words which differ by erasing or inserting subwords of the form xx^{-1} and $x^{-1}x$.

Definition 1.11.5 Let u and v be words on X (possibly empty), and $w = uv$. An *insertion* is changing w to $uxx^{-1}v$ or $ux^{-1}xv$. A *deletion* is changing $uxx^{-1}v$ or $ux^{-1}xv$ to w. By an *elementary operation* we mean insertion or delition. We write $w \to w'$ if w' is obtained from w by an elementary operation. Two words u and w on X are *equivalent*, written $u \sim w$, if we can transform one into another using finitely many elementary operations. The equivalence class of a word w is denoted $[w]$.

Proposition 1.11.6 *Let $F(X)$ be the set of equivalence classes of words with respect to the equivalence relation $\sim$. The multiplication on $F(X)$ given by $[u][w] := [uw]$ is well-defined. It gives $F(X)$ the structure of a group with idenity element $[1]$ and inverse $[w]^{-1} = [w^{-1}]$.*

Proof Note that $u' \sim u$ and $w' \sim w$ imply that $u'v' \sim uv$, which takes care of 'well defined'. The axioms of group are now obvious. □

Proposition 1.11.7 *Let $i : X \to F(X)$, $x \mapsto [x]$. For every group G and map $f : X \to G$ there exists a unique group homomorphism $\hat{f} : F(X) \to G$ such that $\hat{f}(i(x)) = f(x)$.*

Proof Uniqueness of $\hat{f}$ is clear. For existence just set $\hat{f}([x_1^{\varepsilon_1} \dots x_n^{\varepsilon_n}]) := f(x_1)^{\varepsilon_1} \dots f(x_n)^{\varepsilon_n}$. By definition of $F(X)$, the map $\hat{f}$ is a well-defined group homomorphism, which clearly has the desired property. □

The two propositions above got us very close to proving that $F(X)$ is the free group on X. It just remains to check that the map i is an

embedding. As for every $x \in X$, $[x]$ is a reduced word, the required fact follows from the following natural

Proposition 1.11.8 *Every equivalence class of words on X contains a unique reduced word.*

Proof Let w be a word on X. Let us apply consecutive deletions of the *rightmost* subword of the form xx^{-1} or $x^{-1}x$. After finitely many steps we will reach a reduced word $\rho(w) \sim w$. We make a number of observations:

(1) $\rho(uv) = \rho(u\rho(v))$ (obvious).

(2) $\rho(x^{\varepsilon}x^{-\varepsilon}v) = \rho(v)$ (follows from (1)).

(3) $\rho(ux^{\varepsilon}x^{-\varepsilon}v) = \rho(uv)$. Indeed, using (1) and (2), we have

$$\rho(ux^{\varepsilon}x^{-\varepsilon}v) = \rho(u\rho(x^{\varepsilon}x^{-\varepsilon}v)) = \rho(u\rho(v)) = \rho(uv).$$

Now let $u \sim v$ be two reduced words. By definition, there exists a sequence of words $u = u_1, u_2, \ldots, u_l = v$, in which u_{i+1} is obtained from u_i by an elementary operation. In view of (3), $\rho(u_i) = \rho(u_{i+1})$. As $\rho(u) = u$ and $\rho(v) = v$, this implies that $u = v$. □

Remark 1.11.9 The properties (1)-(3) proved in Proposition 1.11.8 imply that

$$\rho(uv) = \rho(\rho(u)\rho(v)). \tag{1.11}$$

Indeed, apply the induction on $|u|$, the induction base $|u| = 0$ being quite clear. Let $|u| > 0$. If u is reduced, then everything is again clear. Otherwise, $u = u_1x^{\varepsilon}x^{-\varepsilon}u_2$. Now, using (2), (3) and inductive hypothesis, we obtain

$$\begin{aligned}
&\rho(uv) = \rho(u_1x^{\varepsilon}x^{-\varepsilon}u_2v) = \rho(\rho(u_1)\rho(x^{\varepsilon}x^{-\varepsilon}u_2v))\\
&= \rho(\rho(u_1)\rho(u_2v)) = \rho(u_1u_2v) = \rho(\rho(u_1u_2)\rho(v))\\
&= \rho(\rho(u_1x^{\varepsilon}x^{-\varepsilon}u_2)\rho(v)) = \rho(\rho(u)\rho(v)).
\end{aligned}$$

Now Proposition 1.11.8 and (1.11) yield another definition of the free group $F(X)$: its elements are reduced words on X and the multiplication is given by $uv = \rho(uv)$. Of course we could have tried to use this construction of the free group instead of our approach with equivalence classes. But then it would be hard to prove associativity of multiplication.

Corollary 1.11.10 *Let G be a group and $X \subseteq G$ be a subset. Then X is free on X if and only if each element $g \in G$ can be uniquely written in the form*

$$g = x_1^{e_1} \dots x_l^{e_l}, \tag{1.12}$$

where $x_i \in X$, $e_i \neq 0$ and $x_i \neq x_{i+1}$.

Proof If G is free on X, we may assume that $G = F(X)$, where $F(X)$ has been explicitly constructed above. Now instead of each $x_i^{e_i}$ we write the word $x_i \dots x_i$ (e_i times) if $e_i > 0$, and $x_i^{-1} \dots x_i^{-1}$ ($|e_i|$ times) if $e_i < 0$. The resulting word of length $|e_1| + \dots + |e_l|$ is reduced, and so the uniqueness of the presentation (1.12) follows from Proposition 1.11.8.

Conversely, assume that every element of G can be written uniquely in the form (1.12). Let $F(X)$ be the free group on X. By the universal property of $F(X)$, the embedding f of X into G induces a surjective homomorphism $\hat{f} : F(X) \to G$, and our assumption guarantees that $\hat{f}$ is injective, thanks to Proposition 1.11.8 again. □

Now we can proceed to describe groups given by generators and relations. Let G be a group generated by some set X of its elements. By the universal property of free groups, there is a natural surjection $\pi : F(X) \to G$, $[x] \mapsto x$. Let S be the kernel of π. Elements of S are called *relations* of the presentation π. If R is a subset of S such that S is the minimal normal subgroup containing R, we say that G is given by (the set of) generators X and (the set of) relations R and write

$$G = \langle X \mid R \rangle.$$

Of course this entirely describes G as the quotient of $F(X)$ by the minimal normal subgroup containing R. If $R = \{[r_1], [r_2], \dots, [r_l]\}$ is finite we also write

$$G = \langle X \mid r_1 = 1, r_2 = 1, \dots, r_l = 1 \rangle.$$

A group G is called *finitely generated* (resp. *finitely presented*) if it has a presentation $G = \langle X \mid R \rangle$ with X finite (resp. both X and R finite).

The following result is a universal property of $G = \langle X \mid R \rangle$. Note that if $G = \langle X \mid R \rangle$ then we have a natural map $i : X \to G$ (which does not have to be an embedding).

Theorem 1.11.11 (von Dyck's Theorem) *Let $G = \langle X \mid R \rangle$. Let $f : X \to H$ be a map of X into another group H. Assume that for every $r = [x_1^{\varepsilon_1} \dots x_l^{\varepsilon_l}] \in R$, we have $f(x_1)^{\varepsilon_1} \dots f(x_l)^{\varepsilon_l} = 1$ in H. Then there*

exists a unique group homomorphism $\hat{f} : G \to H$ *such that* $\hat{f}(i(x)) = f(x)$ *for all* $x \in X$.

Proof By the universal property of the free group $F(X)$, there exists a surjective homomorphism $g : F(X) \to H$ with $g(x) = f(x)$ for every $x \in X$. It follows from the assumption that the kernel K of this homomorphism is at least as large as the minimal normal subgroup S generated by R. Now, we can construct $\hat{f}$ as the natural map $F(X)/S \to F(X)/K = \operatorname{im} g \subseteq H$. Finally, uniqueness of $\hat{f}$ is clear as the elements of $i(X)$ generate G. □

Example 1.11.12 The dihedral group D_{2n} is given by generators and relations as follows:

$$D_{2n} \cong \langle a, b \mid a^n = 1, b^2 = 1, bab^{-1} = a^{-1} \rangle. \tag{1.13}$$

Note first of all that, strictly speaking, instead of $bab^{-1} = a^{-1}$ we should have written $bab^{-1}a = 1$, but this meaning is clear. You should be aware that it is typical to see such rewritten relations if people think that in the rewritten form the reation looks nicer.

Now we prove (1.13). Let x be a rotation by $360/n$ degrees and y be a reflection in D_{2n}. Then it is easy to see that $x^n = 1$, $y^2 = 1$, $yxy^{-1} = x^{-1}$. By von Dyck's Theorem, there exists a surjective homomorphism from the group $G := \langle a, b \mid a^n = 1, b^2 = 1, bab^{-1} = a^{-1} \rangle$ onto D_{2n}, which maps a to x and b to y. So it remains to prove that $|G| = |D_{2n}| = 2n$. This will follow if we can show that every element of G can be written in the form $a^m b^n$, which is clear from defining relations (which allow us to pull all a's past b's to the left).

We leave it as an exercise that D_{2n} also has another useful presentation

$$D_{2n} \cong \langle s_1, s_2 \mid s_1^2 = 1, s_2^2 = 1, (s_1 s_2)^n = 1 \rangle. \tag{1.14}$$

Example 1.11.13 Define the *generalized quaternion* group

$$Q_{4m} := \langle a, b \mid a^{2m} = 1, a^m = b^2, bab^{-1} = a^{-1} \rangle.$$

We claim that the order of Q_{4m} is $4m$. First of all $|Q_{4m}| \leq 4m$ because every element of it can be written in the form $a^i b^j$ for $0 \leq i < 2m$ and $j \in \{0, 1\}$ (use the relations). How to prove that the order is exactly $4m$? This is usually the hard part, and a typical trick is to find a group of order at least $4m$ generated by two elements which satisfy the same relations as defining relations of Q_{4m}. Application of von Dyck's

Theorem then completes the proof. Well, let us construct the required group. Let $\mathbb{H} = \mathbb{R} \cdot 1 \oplus \mathbb{R} \cdot i \oplus \mathbb{R} \cdot j \oplus \mathbb{R} \cdot k$ be the algebra of quaternions. Let $x := e^{\pi i/m}$ and $y := j$. It is easy to check that $x^{2m} = 1$, $x^m = y^2$, and $yxy^{-1} = x^{-1}$. Let G be the subgroup of $\mathbb{H}^\times$ generated by x and y. Now it is easy to see working explicitly in $\mathbb{H}$ that the elements $x^i y^j$ with $0 \leq i < 2m$ and $j \in \{0, 1\}$ are all distinct which completes the proof.

It is easy to check that $Z(Q_{4m}) = \{1, a^m\}$, and $Q_{4m}/Z(Q_{4m}) \cong D_{2m}$. So Q_{4m} is solvable. It is nilpotent if and only if m is a power of 2.

Example 1.11.14 The symmetric group S_n has the following presentation:

$$\begin{aligned} S_n \cong \langle s_1, s_2, \ldots, s_{n-1} \mid \\ s_i^2 = 1 \ (1 \leq i < n), \\ s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} (1 \leq i < n-1), \\ s_i s_j = s_j s_i \ (1 \leq i, j < n, \ |i-j| > 1) \rangle. \end{aligned} \tag{1.15}$$

Let G be the group given by generators and relations as in the right hand side of (1.15). Note that the elements $(1,2), (2,3), \ldots, (n-1, n)$ generate S_n and satisfy the same relations as $s_1, s_2, \ldots, s_{n-1}$. Therefore, by von Dicks's Theorem, there exists a surjection from G onto S_n mapping s_j to $(j, j+1)$ for $1 \leq j < n$, and it suffices to show that $|G| \leq n!$. For this we apply induction on n, the cases $n = 1, 2$ being clear. For inductive step, consider $H < G$ generated by $s_1, \ldots, s_{n-2}$. It follows from the inductive hypothesis that $|H| \leq (n-1)!$, so it is enough to prove that $[G : H] \leq n$. This will follow if we can show that $G = K$, where

$$K := H \cup s_{n-1} H \cup s_{n-2} s_{n-1} H \cup \cdots \cup s_1 s_2 \ldots s_{n-1} H.$$

The equality $G = K$ is equivalent to the property $gK = K$ for every $g \in G$, which is in turn equivalent to $s_i K = K$ for every $1 \leq i < n$. Consider $s_i(s_j \ldots s_{n-1} H)$. If $i = j - 1$ or j, then clearly $s_i(s_j \ldots s_{n-1} H) \subset K$. If $i < j - 1$, we get

$$s_i(s_j \ldots s_{n-1} H) = s_j \ldots s_{n-1} s_i H = s_j \ldots s_{n-1} H \subset K.$$

Finally, if $i > j$, we have

$$\begin{aligned} & s_i(s_j \ldots s_{n-1} H) = s_j \ldots s_{i-2}(s_i s_{i-1} s_i) s_{i+1} \ldots s_{n-1} H \\ & = s_j \ldots s_{i-2}(s_{i-1} s_i s_{i-1}) s_{i+1} \ldots s_{n-1} H = s_j \ldots s_{n-1} s_{i-1} H \\ & = s_j \ldots s_{n-1} H \subset H. \end{aligned}$$

Remark 1.11.15 The presentations (1.14) and (1.15) can be generalized as follows. Let r be a positive integer. For every pair $1 \leq i, j \leq r$ fix a parameter $m_{ij} \in \mathbb{Z}_{>0} \cup \{\infty\}$ such that $m_{ij} = m_{ji} \geq 2$ for $1 \leq i \neq j \leq r$ and $m_{ii} = 1$ for $1 \leq i \leq r$. Consider the group generated by $s_1, \ldots, s_r$ subject to the relations

$$(s_i s_j)^{m_{ij}} = 1 \qquad (1 \leq i, j \leq r).$$

This group is called a Coxeter group. The parameters m_{ij} can be visualized conveniently using the corresponding *Coxeter graph*. The vertices of the graph are labelled by $1, 2, \ldots, r$. There is no edge between the vertices i and j if either $i = j$ or $m_{ij} = 2$, there is an edge without label between i and j if $m_{ij} = 3$, and there is an edge with label $m_{i,j}$ between i and j otherwise. For example, for the presentation (1.15) of S_4 the Coxeter graph is

The most remarkable is the following classification theorem: a Coxeter group is finite if and only if its Coxeter graph is a finite disjoint union of graphs appearing in Figure 1.1. If you are stunned (and I bet you are!) by the beauty of this result, read more about it, for example in [Hu]. This result is a beginning of Lie theory.

Remark 1.11.16 Besides Lie theory, we have touched another large and beautiful area of mathematics: combinatorial group theory. This theory investigates what can be said about a group given by generators and relations, and turns out to be closely related with topology and geometry. One of the prominent results is the unsolvability of the word problem. A group G has a solvable word problem if it has a presentation $G = \langle X \mid R \rangle$ for which there exists an algorithm to determine whether an arbitrary word on X is equal to the identity in G. We would also like to know when two different presentations determine the same group (we saw that it is indeed possible to have different presentations for the same group). Another important problem is to determine whether a group determined by a presentation is finite or infinite. If you want to read about these matters, we recommend, for example, [Ol].

To conclude the section, we classify groups of order < 16. In view of Example 1.7.19, we understand groups of order pq, so the only interesting cases are groups of order 8 and 12. Abelian groups can be easily understood using the Fundamental Theorem on Finitely Gener-

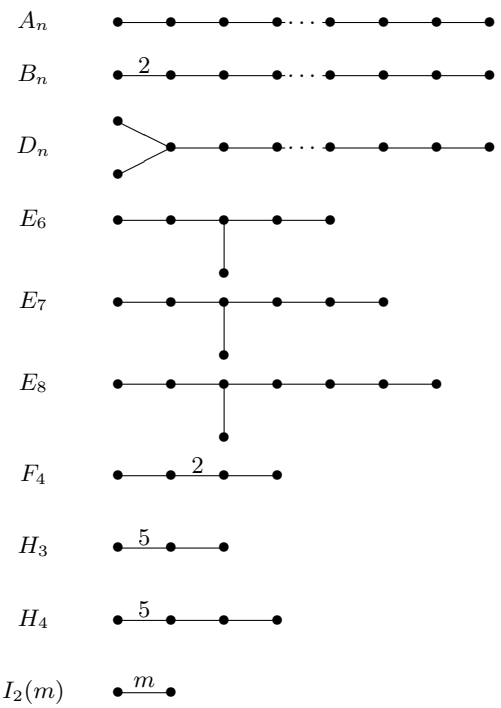


Fig. 1.1. Coxeter graphs of finite Coxeter groups

ated Abelian Groups, to be studied later. So here we concentrate on non-abelian groups.

Lemma 1.11.17 *Q_8 and D_8 are the only non-abelian groups of order 8.*

Proof Let G be a non-abelian group of order 8. Not every element of G has order 2, so there must be an element $a \in G$ of order 4. The subgroup $A := \langle a \rangle$ in G is of index 2, hence normal. Pick $b \in G \setminus A$. Then $G = \langle a, b \rangle$, and $b^2 \in A$. Hence, either $b^2 = 1$ or $b^2 = a^2$—otherwise b has order 8. Also, $bab^{-1} = a^3$, as it must be an element of order 4 and cannot be a—otherwise G is abelian. Now von Dick's theorem implies that G is a homomorphic image of D_8 or Q_8, and the result follows by comparing orders. □

Lemma 1.11.18 *Up to isomorphism, there are three non-abelian groups of order 12: A_4, D_{12}, and Q_{12}.*

Proof The three groups have, respectively, 3, 7, and 1 elements of order 2, so they are non-isomorphic. Now let G be a non-abelian group of order 12, and let $P = \langle c \rangle$ be a Sylow 3-subgroup of G. The action on G/P provides a homomorphism $G \to S_4$. Let K be the kernel of this homomorphism.

If $K = \{1\}$, then G is isomorphic to a subgroup of S_4 of order 12. Such subgroup must be contained in $S_4' = A_4$, so $G \cong A_4$.

If $K \neq \{1\}$, then $K = P$, and $P \triangleleft G$. By the Sylow Theorems, P is the only subgroup of order 3. Hence c and c^2 are the only elements of order 3 in G. Hence c has one or two conjugates, and so $|C_G(c)| = 6$ or 12. So $C_G(c)$ contains an involution d. Let $a = cd$. Then a has order 6, and so $A := \langle a \rangle$ is a normal subgroup of order 6.

Take any $b \notin A$. Then $G = \langle a, b \rangle$. Observe that $bab^{-1} \in A$ is of order 6, and $bab^{-1} \neq a$, as G is not abelian. Hence $bab^{-1} = a^{-1}$. Finally, b^2 belongs to A and commutes with b, whence either $b^2 = 1$ or $b^2 = a^3$. Now application of von Dick's Theorem completes the proof. □

1.12 Problems on Groups

Problem 1.12.1 True or false? Let G be a group such that any finitely generated subgroup of G is cyclic. Then G is cyclic.

Problem 1.12.2 Let G be a finite group with $|G| > n$. Then G is isomorphic to a transitive subgroup of S_n if and only if G contains a subgroup H of index n such that neither H nor any proper subgroup of H is normal in G.

Problem 1.12.3 Let G be a finite group. We choose an element $g \in G$ randomly. Then replace it and make another random choice of an element $h \in G$. Prove that the probability that g and h commute equals to $k/|G|$, where k is the number of conjugacy classes in G.

Problem 1.12.4 Any infinite group has infinitely many subgroups.

Problem 1.12.5 Let G be a finite group. The *Frattini subgroup* $\Phi(G)$ is the intersection of all maximal subgroups of G. An element $g \in G$ is called a *non-generator* if whenever $\langle X, g \rangle = G$, we have $\langle X \rangle = G$ for subsets $X \subseteq G$. Show that $\Phi(G)$ is the set of non-generators of G.

Problem 1.12.6 True or false? $\mathrm{Aut}(C_8) \cong C_4$.

Problem 1.12.7 Let G be a finite group and p be a prime. Show that there exists a normal p-subgroup $O_p(G)$ such that $H \leq O_p(G)$ for any normal p-subgroup H in G. Show that there exists a normal subgroup $O_{p'}(G)$ of order prime to p such that $H \leq O_{p'}(G)$ for any normal subgroup H in G whose order is prime to p.

Problem 1.12.8 Let G be a finite group and p be a prime. Show that there exists a normal subgroup $O^p(G)$ such that $G/O^p(G)$ is a p-group, and $H \geq O_p(G)$ for any normal p-subgroup H in G such that G/H is a p-group. Show that there exists a normal subgroup $O^{p'}(G)$ such that $|G/O^{p'}(G)|$ is prime to p, and $H \leq O_{p'}(G)$ for any normal subgroup H in G with $|G/H|$ prime to p.

Problem 1.12.9 True or false? If G is a finite nilpotent group, and m is a positive integer dividing $|G|$. Then there exists a subgroup of G of order m.

Problem 1.12.10 A non-trivial group G has a proper subgroup H which contains every proper subgroup of G. What can you say about G? (Another version: Let G be a finite group such that for all subgroups $H, K \leq G$ we have $H \subseteq K$ or $K \subseteq H$. What can you say about G?)

Problem 1.12.11 Let G be a finite group. For each prime p dividing $|G|$, let $\mathcal{S}_p(G)$ denote the set of all p-subgroups of G. Suppose for each p dividing $|G|$, that $\mathcal{S}_p(G)$ is totally ordered by inclusion (i.e. we have $H \subseteq K$ or $K \subseteq H$ for any $H, K \in \mathcal{S}_p(G)$). Prove that G is cyclic.

Problem 1.12.12 True or false? If G is a group with even number of elements, then the number of elements in G of order 2 is odd.

Problem 1.12.13 True or false? If N is a normal subgroup of G and N and G/N are nilpotent, then G is nilpotent.

Problem 1.12.14 True or false? If X is a normal subgroup of Y and Y is a normal subgroup of Z, then X is a normal subgroup of Z.

Problem 1.12.15 Let G be a finite group and $N \triangleleft G$. If $(|N|, [G : N]) = 1$, prove that N is the unique subgroup of G having order $|N|$.

Problem 1.12.16 Let G be a group and $H, K \leq G$ be subgroups. Then HK is a subgroup of G if and only if $HK = KH$.

Problem 1.12.17 If H, K are subgroups of G, then $[H : H \cap K] \leq [G : K]$. If $[G : K]$ is finite, then $[H : H \cap K] = [G : K]$ if and only if $G = HK$.

Problem 1.12.18 If H, K are subgroups of finite index of a group G such that $[G : H]$ and $[G : K]$ are relatively prime, then $G = HK$.

Problem 1.12.19 True or false? All subgroups of Q_8 are normal.

Problem 1.12.20 The center of S_n is trivial for $n \geq 3$.

Problem 1.12.21 If $N \triangleleft G$, $|N|$ is finite, $H \leq G$, $[G : H]$ is finite, and $([G : H], |N|) = 1$, then $N \leq H$.

Problem 1.12.22 $(\mathbb{Q}, +)$ does not have subgroups of finite index.

Problem 1.12.23 If H and K are finite index subgroups in G, then so is $H \cap K$.

Problem 1.12.24 If H is a proper subgroup of a finite group G, then the union $\cup_{g \in G} gHg^{-1}$ is not the whole G. (See 1.12.25 for a more general problem)

Problem 1.12.25 If $H < G$ is a proper subgroup of finite index, then the union $\cup_{g \in G} gHg^{-1}$ is not the whole G.

Problem 1.12.26 Let G be a finite group G, and $g_1, \ldots, g_k$ be representatives of the conjugacy classe of G. Then $G = \langle g_1, \ldots, g_l \rangle$.

Problem 1.12.27 The group $GL_n(\mathbb{F}_q)$ has an element of order $q^n - 1$.

Problem 1.12.28 Let H be a subgroup of a group G. Let

$$C := \{g \in G \mid H \cap gHg^{-1} \text{ has finite index in both } H \text{ and } gHg^{-1}\}.$$

Show that C is a subgroup of G.

Problem 1.12.29 Suppose that a finite group G has exactly three Sylow

2-subgroups. Show that every permutation of these Sylow subgroups can be obtained by conjugation by some suitable element of G.

Problem 1.12.30 Prove that any infinite simple group G has no subgroup of finite index.

Problem 1.12.31 Prove that there is no simple group of order 120.

Problem 1.12.32 The group of order $2^4 \cdot 7^2$ is not simple.

Problem 1.12.33 Prove that there is no simple group of order 150.

Problem 1.12.34 Show that a group of order 80 must have a non-trivial normal subgroup of order a power of 2.

Problem 1.12.35 Prove that the group of upperunitriangular 3×3 matrices over $\mathbb{F}_2$ is isomorphic to D_8.

Problem 1.12.36 Let G be a finite group with $g \sim g^2$ for every $g \in G$. Prove that $G = \{1\}$.

Problem 1.12.37 Suppose that a finite group G has exactly two conjugacy classes. Determine G up to isomorphism.

Problem 1.12.38 True or false: $S_4/V_4 \cong S_3$.

Problem 1.12.39 True or false: every subgroup of order 5 of S_5 is transitive.

Problem 1.12.40 True or false: if p and q are primes, then a group of order pq is nilpotent.

Problem 1.12.41 A group G is called *metabelian* if there exists a normal subgroup N of G with N and G/N both abelian. Prove that every subgroup of a metabelian group is metabelian. Prove that every quotient of a metabelian group is metabelian.

Problem 1.12.42 True or false: If P is a Sylow p-subgroup of the finite group G, then $N_G(P)$ contains just one Sylow p-subgroup of G.

Problem 1.12.43 If $H \leq G$ are finite groups, then $r_p(H) \leq r_p(G)$.

Problem 1.12.44 Let $H \leq G$ be finite groups, P be a Sylow p-subgroup of H, and $N_G(P) \subseteq H$. Then P is a Sylow p-subgroup of G.

Problem 1.12.45 If $|G| = pqr$, show that G is not simple.

Problem 1.12.46 Let $|G| = p(p+1)$, where p is prime. Show G has either a normal subgroup of order p or a normal subgroup of order $p+1$.

Problem 1.12.47 Let a finite group G have a cyclic Sylow 2-subgroup. Show that G has a subgroup of index 2.

Problem 1.12.48 Let $n \neq 6$. Then every automorphism of S_6 is inner.

Problem 1.12.49 The group S_6 has an automorphism mapping the stabilizer of a point in S_6 to a transitive subgroup. No such automorphism can be inner.

Problem 1.12.50 Let G be a finite group. Then G is nilpotent if and only if $N_G(H) \gneq H$ whenever $H \lneq G$.

Problem 1.12.51 True or false: D_n is nilpotent.

Problem 1.12.52 Let $\varphi : G \to H$ be a surjective homomorphism of finite groups. If P is a Sylow p-subgroup of G, then $\varphi(P)$ is a Sylow p-subgroup of H. Conversely, every Sylow p-subgroup of H is the image of a certain Sylow p-subgroup of G.

Problem 1.12.53 Let G be a finite group, $N \triangleleft G$, and P a Sylow p-subgroup of G for some prime p. Show that PN/N is a Sylow p-subgroup of G/N and $P \cap N$ is a Sylow p-subgroup of N.

Problem 1.12.54 If a group G contains an element having exactly two conjugates, then G has a non-trivial proper normal subgroup.

Problem 1.12.55 Any finite group is isomorphic to a subgroup of A_n for some n.

Problem 1.12.56 True or false? A_5 contains no subgroup of order 15.

Problem 1.12.57 Find the smallest n such that A_n contains a subgroup of order 15.

Problem 1.12.58 True or false? Let G be a finite group, and let P be its Sylow p-subgroup. If $P \triangleleft N \triangleleft G$, then $P \triangleleft G$.

Problem 1.12.59 Let G be a finite p-group, and $N \triangleleft G$ be a normal subgroup of order p. Then N is in the center of G.

Problem 1.12.60 If G is a finite p-group, $N \triangleleft G$, and $N \neq \{1\}$, then $N \cap Z(G) \neq \{1\}$.

Problem 1.12.61 Let H, K, N be non-trivial normal subgroups of a group G, and suppose that $G = H \times K$. Prove that N is in the center of G or N intersects one of H, K non-trivially. Give an example where N is in the center and does not intersect either H or K non-trivially. Give an example where N is not in the center but intersects both H and K non-trivially. Give an example when N is in the center and intersects both H and K non-trivially.

Problem 1.12.62 Let $N_1 \triangleleft G_1$, $N_2 \triangleleft G_2$. True or false?

(a) If $G_1 \cong G_2$, and $N_1 \cong N_2$ then $G_1/N_1 \cong G_2/N_2$.
(b) If $G_1 \cong G_2$ and $G_1/N_1 \cong G_2/N_2$ then $N_1 \cong N_2$.
(c) If $N_1 \cong N_2$ and $G_1/N_1 \cong G_2/N_2$ then $G_1 \cong G_2$.

Problem 1.12.63 Let G be a finite group. If G is solvable, then G contains a non-trivial normal abelian subgroup. If G is not solvable then it contains a normal subgroup H such that $H' = H$.

Problem 1.12.64 Let p be a prime. Find the number of subgroups of $C_p \times C_p$ (counting $\{1\}$ and itself).

Problem 1.12.65 Let G be a group, and suppose $|\mathrm{Aut}(G)| = 1$. Prove that G has at most two elements.

Problem 1.12.66 True or false: every group of order 18 is nilpotent.

Problem 1.12.67 Let G_1 and G_2 be finite groups. Is it true that every subgroup of $G_1 \times G_2$ is of the form $H_1 \times H_2$, where $H_1 \leq G_1$ and $H_2 \leq G_2$. What if we assume, additionally, that $(|G_1|, |G_2|) = 1$?

Problem 1.12.68 True or false: Any element of order p in a finite p-group is central.

Problem 1.12.69 Let p be a prime and G be a finite simple group having a subgroup H of index p. Find the isomorphism type of a Sylow p-subgroup of G.

Problem 1.12.70 Classify the groups of order 175.

Problem 1.12.71 Let F be a free group with basis $X = \{x_1, \ldots, x_n\}$. Then $F/F' \cong C_\infty \times \cdots \times C_\infty$ (n copies).

Problem 1.12.72 Let $X \subseteq Y$. Show that $F(X) \leq F(Y)$.

Problem 1.12.73 Find the group of rotations and the full group of symmetries of a regular tetrahedron.

Problem 1.12.74 Describe the conjugacy classes of A_5 and S_5.

Problem 1.12.75 True or false? $D_{12} \cong S_3 \times C_2$.

Problem 1.12.76 Let p be prime g be any p-cycle in S_p and h be any transposition in S_p. Prove that $\langle g, h\rangle = S_p$.

Problem 1.12.77 Let D_{2n} be the dihedral group of order $2n$. For which value of n:

(a) The center of D_{2n} is trivial?

(b) All involutions in D_{2n} are conjugate to each other?

(c) D_{2n} is a direct product of two proper subgroups?

2
Fields

2.1 Things known

We assume that material reviewed in sections 3.1-3.5 and parts of section 3.8 of Rotman is well understood. Among other things, these sections discuss the following topics:

3.2. Definition of a commutative ring (with 1), basic examples; domains, zero divisors, divisibility theory; definition of a field, field of fractions of a domain, the field $\mathbb{F}_p$.
3.3. Polynomials; rational functions.
3.4. Division algorithm for polynomials; roots of polynomials; GCD and LCM for polynomials; irreducible polynomials; Euclidean algorithm; Unique Factorization Theorem for polynomials; algebraic integers;
3.5. Ring homomorphism and isomorphism; ideals; PID's; UFD's; GCD and LCM in UFD's.
3.8. Quotient rings, First Isomorphism Theorem for rings.

We want to emphasize the following fundamental universal properties:

Theorem 2.1.1 (Universal Property of Polynomials) *Let R and S be commutative rings, $\varphi : R \to S$ be a ring homomorphism, and $s \in S$ be an arbitrary element. Then there exists a unique homomorphism $\hat{\varphi} : R[x] \to S$, which maps x to s and $r \cdot 1$ to $\varphi(r)$ for any $r \in R$. Moreover, for general $r_0 + r_1 x + \cdots + r_n x^n \in k[x]$ we have*

$$\hat{\varphi}(r_0 + r_1 x + \cdots + r_n x^n) = \varphi(r_0) + \varphi(r_1)s + \cdots + \varphi(r_n)s^n.$$

Theorem 2.1.2 (Universal Property of Fraction Fields) *Let R and S be commutative rings and $\varphi : R \to S$ be a ring homomorphism.*

Assume that R is a domain and that $\varphi(r)$ is invertible in S for any $r \neq 0$. Let $Q(R)$ be the quotient field of R. Then there exists a unique homomorphism $\hat{\varphi} : Q(R) \to S$, which maps r/s to $\varphi(r)\varphi(s)^{-1}$.

Make sure you know how to prove the theorems above.

2.2 More on irreducible polynomials

Apart from what is explained in sections 3.3 and 3.4 of the textbook, we will need some more facts on irreducible polynomials.

Lemma 2.2.1 *Let $f \in \mathbb{Z}[x]$ be a polynomial which cannot be written as a product of two polynomial in $\mathbb{Z}[x]$ of positive degrees. Then f is irreducible when considered as an element of $\mathbb{Q}[x]$.*

Proof Assume that f is reducible in $\mathbb{Q}[x]$, i.e. $f = gh$ for polynomials $g, h \in \mathbb{Q}[x]$ of degrees smaller than $\deg f$. Multiplying by the product of denominators of coefficients of g and h, we get $nf = g'h'$, where $n \in \mathbb{Z}$ and $g'h' \in \mathbb{Z}[x]$. We now show that one can cancel prime factors of n one by one without going outside $\mathbb{Z}[x]$.

Suppose p is a prime factor of n. We claim that either p divides all coefficients of g' or else p divides all coefficients of h'. Otherwise let i and j be the minimal possible such that p does not divide the ith coefficient of g' and jth coefficient of h'. It follows that p does not divide the $(i+j)$th coefficient of $g'h'$, giving a contradiction. Now cancel by p and continue this way until $n = 1$. □

Theorem 2.2.2 (Eisenstein's Criterion) *Let $(x) = a_0 + a_1 x + \cdots + a_n t^n \in \mathbb{Z}[x]$. Suppose that there is a prime p such that*

(1) *p does not divide a_n,*
(2) *p divides a_i for $0 \leq i < n$,*
(3) *p^2 does not divide a_0.*

Then f is irreducible as an element of $\mathbb{Q}[x]$.

Proof By Lemma 2.2.1, it suffices to show that f is irreducible in $\mathbb{Z}[x]$. Suppose for a contradiction that $f = gh$, where $g = b_0 + b_1 x + \cdots + b_r x^r$ and $h = c_0 + c_1 x + \cdots + c_s x^s$ are integral polynomials of degrees less than n. Then $a_0 = b_0 c_0$, so by (2), $p|b_0$ or $p|c_0$. By (3), we may assume that $p|b_0$ and $p \not| c_0$. If all coefficients b_i were divisible by p then a_n would be divisible by p, contrary to (1). Let b_i be the first coefficient of g not

divisible by p. Note that $i < n$. Then $a_i = b_i c_0 + \cdots + b_0 c_i$ is not divisible by p, giving a contradiction. □

Example 2.2.3 $f(x) = \frac{2}{9}x^5 + \frac{5}{3}x^4 + x^3 + \frac{1}{3} \in \mathbb{Q}[x]$ is irreducible if and only if $9f(x) = 2x^5 + 15x^4 + 9x^3 + 3$ is irreducible, which is the case by the Eisenstein's criterion with $p = 3$.

Example 2.2.4 Eisenstein's criterion does not apply to the polynomial $f(x) = x^3 - 3x - 1 = 0$. But $f(x)$ is irreducible if and only if $f(x+1)$ is irreducible. Now $f(x+1) = x^3 + 3x^2 - 3$ is irreducible by Eisenstein's criterion.

Example 2.2.5 Let p be a prime. Then the polynomial $f(x) = 1 + x + \cdots + x^{p-1}$ is irreducible. Indeed, consider $f(x+1)$ instead. As $f(x) = (x^p - 1)/(x - 1)$, we have $f(x+1) = ((x+1)^p - 1)/x$, and Eisenstein's criterion applies.

There is another method to verify irreducibility. By the Universal Property of Polynomials, the natural homomorphism $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ extends to a homomorphism $\mathbb{Z}[x] \to (\mathbb{Z}/n\mathbb{Z})[x], f \mapsto \bar{f}$. Now, assume that the top coefficient of f is not 0 modulo n. Then clearly f is irreducible if $\bar{f}$ is irreducible (the converse is not necessarily true). Now, verifying if $\bar{f}$ is irreducible is a 'finite problem', as there are only finitely many polynomials of bounded degree over $\mathbb{Z}/n\mathbb{Z}$.

Example 2.2.6 We claim that $f(x) = x^4 + 15x^3 + 7$ is irreducible. Modulo 5 we have $\bar{f} = x^4 + 2$. If $\bar{f}$ has a factor of degree 1 then it has a root in $\mathbb{Z}/5\mathbb{Z}$, which is not the case. Let f have two factors of degree 2:

$$x^4 + 2 = (x^2 + ax + b)(x^2 + cx + d).$$

Then $a + c = 0$, $ac + b + d = 0$, $bd = 2$. So $b + d = a^2$, which can take only the values 0, 1, 4, since these are the only squares in $\mathbb{Z}/5\mathbb{Z}$. Hence, either $-b^2 = 2$ or $b(1-b) = 2$ or $b(4-b) = 2$. Trying all possible values for b we see that none of these equations can hold.

Finally, one more useful trick:

Lemma 2.2.7 *Let $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{Z}[x]$. If $\frac{r}{s} \in \mathbb{Q}$ with $(r, s) = 1$ is a root of f then $r \mid a_0$ and $s \mid a_n$.*

Proof As $f(r/s) = 0$, we have

$$a_0 s^n = r(\sum_{i=1}^{n}(-a_i)r^{i-1}s^{n-i}) \quad \text{and} \quad -a_n r^n = s\sum_{i=0}^{n-1} r^i s^{n-i-1},$$

which implies the required result. □

Example 2.2.8 Lemma 2.2.7 implies that the polynomial $x^3 - 3x + 1 \in \mathbb{Q}[x]$ does not have rational roots and hence is irreducible over $\mathbb{Q}$ as it has degree 3.

2.3 First steps in fields

Let F be a field. If $k_i \subseteq F,\ i \in I$ is a family of subfields, then $\cap_{i\in I} k_i$ is also a subfield. This implies that every field F has the smallest subfield, which is just the intersection of all its subfields.

Definition 2.3.1 The smallest subfield of a field F is called the *prime subfield* of F and denoted F_0.

It turns out that for an arbitrary field F, we have $F_0 \cong \mathbb{Q}$ or $\mathbb{F}_p$ for some prime p. Indeed, consider the ring homomorphism $\chi : \mathbb{Z} \to F,\ m \mapsto m \cdot 1_F$. We have two cases:

(1) χ is injective. Then $\mathbb{Z} \cong \operatorname{im}\chi$ is a subring of F_0. By the universal property of fraction fields, the field of fractions of $\mathbb{Z}$, which is $\mathbb{Q}$, is a subfield of F_0, so $F_0 \cong \mathbb{Q}$. In this case we say that *characteristic* of F is 0, written $\operatorname{char} F = 0$.
(2) $\ker\chi = (p)$ for a prime number p. Then $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z} \cong \operatorname{im}\chi$ is a subfield of F_0. So $F_0 \cong \mathbb{F}_p$. In this case we say that *characteristic* of F is p, written $\operatorname{char} F = p$.

Remark 2.3.2 We have $\operatorname{char} F = p > 0 \ \Leftrightarrow\ p \cdot a = 0$ for any $a \in F$ $\Leftrightarrow\ p \cdot 1_F = 0$, where $p \cdot a$ means $a + \cdots + a$ (p summands). This implies the following nice property which holds in any field F of characteristic p and is sometimes called 'Freshman's Dream':

$$(a + b)^p = a^p + b^p \qquad (a, b \in F).$$

Definition 2.3.3 If k is a subfield of another field K, we say that K is a field extension over k, written K/k.

For any field extension K/k, we may consider K as a k-vector space in a natural way.

Proposition 2.3.4 *If F is a finite field then $|F| = p^n$ for some prime p and some $n \in \mathbb{Z}_{>0}$.*

Proof Consider F as a vector space over F_0. □

We will later prove that for any prime p and any $n \in \mathbb{Z}_{>0}$ there does exist a field with p^n elements, and, moreover, such field is unique (up to isomorphism).

The following is a very useful result on degrees of finite extensions.

Proposition 2.3.5 (Tower Law) *Let $K/E/k$ be field extensions with K/E and E/k finite. Then K/k is finite, and $[K : k] = [K : E][E : k]$.*

Proof If $\{a_1, \ldots, a_m\}$ is a basis of E over k and $\{b_1, \ldots, b_n\}$ is a basis of K over E, it is easy to see that $\{a_i b_j \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$ is a basis of K over k. □

Definition 2.3.6 (i) Let K/k be a field extension. The extension is called *finite* if $\dim_k K < \infty$. The dimension $\dim_k K$ is called the *degree* of K/k and denoted $[K : k]$.

(ii) An element $\alpha \in K$ is called *algebraic* over k if it is a root of some non-zero polynomial $f(x) \in k[x]$. Otherwise α is called *transcendental* over k. The extension K/k is called *algebraic* if all elements of K are algebraic over k. Otherwise the extension is called *transcendental*.

Lemma 2.3.7 *If K/k is a finite field extension, then it is algebraic.*

Proof Let $\dim_k K = n$. Take any $\alpha \in K$, and consider $n + 1$ elements $1, \alpha, \alpha^2, \ldots, \alpha^n$. They must be linearly dependent over k, so there exist $a_0, a_1, \ldots, a_n \in k$, not all zero, with $a_0 + a_1\alpha + \cdots + a_n\alpha^n = 0$. This shows that α is algebraic. □

The following simple theorem is of fundamental importance. It is the key tool in constructing field extensions.

Theorem 2.3.8 *If k is a field, $p \in k[x]$, and $I = (p) \lhd k[x]$, then $k[x]/I$ is a field if and only if the polynomial p is irreducible. Moreover, $\dim k[x]/I = \deg p$.*

Proof Let $n = \deg p$. It is easy to see that $\{1, x, \dots, x^{n-1}\}$ is a basis of $k[x]/I$, whence $\dim k[x]/I = n$. If $p = fg$ is not irreducible, then the quotient $k[x]/I$ has zero divisors: $(f + I)(g + I) = 0 + I$; in particular, $k[x]/I$ is not a field.

Now, let p be irreducible. Then $1 + I \neq 0 + I$ is clear. Moreover, take $f + I \neq 0$. As $f \notin I$, we have $(f, p) = 1$, whence $1 = af + bp$ for some $a, b \in k[x]$. Hence $a + I$ is the inverse of $f + I$ in $k[x]/I$. Thus, $k[x]/I$ is a field. □

Remark 2.3.9 Let k is a field, $p \in k[x]$ be an irreducible polynomial, and set $K := k[x]/(p)$. Then $K \supseteq k$ is a field extension, and $x + (p) \in K$ is a root of p in K. So sometimes we refer to the passage from k to $k[x]/(p)$ as 'adjoining a root of an irreducible polynomial p to k'.

Definition 2.3.10 Let K/k be a field extension and $\alpha \in K$. The smallest subfield of K containing k and α is called the subfield of K obtained by *adjoining α to K*. This subfield is denoted $k(\alpha)$. If $K = K(\alpha)$ for some $\alpha \in K$ we say that K is a *simple extension* of k.

More generally, if $X \subset K$ be a subset, then $k(X)$ denotes the smallest subfield of K containing k and α. If $X = \{\alpha_1, \dots, \alpha_n\}$ we also write $k(\alpha_1, \dots, \alpha_n)$ for $k(X)$.

Remark 2.3.11 An extension may be simple without appearing to be. Consider $K := \mathbb{Q}(i, \sqrt{5})$. As written, it appears to require adjunction of two elements. But in fact $K = \mathbb{Q}(i + \sqrt{5})$. To prove this it is enough to show that i and $\sqrt{5}$ belong to $\mathbb{Q}(i + \sqrt{5})$. Well, $\mathbb{Q}(i + \sqrt{5})$ contains

$$(i + \sqrt{5})^2 = 4 + 2i\sqrt{5}.$$

So it also contains

$$(i + \sqrt{5})(4 + 2i\sqrt{5}) = 16i,$$

whence it contains i and $\sqrt{5}$.

Theorem 2.3.12 *Let K/k be a field extension, and $\alpha \in K$ be an element algebraic over k. Then there exists a unique monic irreducible polynomial $p \in k[x]$ having α as a root. This polynomial can be characterized as the monic polynomial of minimal possible degree having α as a root. Finally, there is an isomorphism $k(\alpha) \cong k[x]/(p)$ which is identity on k.*

Proof Consider the ring homomorphism $\varphi : k[x] \to K,\ f(x) \mapsto f(\alpha)$. As $k[x]$ is a PID, we have $\ker \varphi = (p)$ for some unique monic polynomial $p(x)$. Moreover, p is irreducible, as otherwise $k[x]/(p) \subseteq K$ has zero divisors. Moreover, p is the only monic irreducible polynomial of (p) and the unique monic polynomial in (p) of minimal possible degree. Now observe that the ideal (p) consists precisely of all the polynomials in $k[x]$ which have α as a root. For the last part of the theorem, note that $k(\alpha) = \operatorname{im} \varphi$. □

Definition 2.3.13 The irreducible monic polynomial $p \in k[x]$ introduced in Theorem 2.3.12, is called the *minimal polynomial* of α over k and denoted $\operatorname{irr}(\alpha; k)$.

Remark 2.3.14 Theorem 2.3.12 shows that adjoining elements, as in Definition 2.3.10, and adjoining roots of irreducible polynomials, as in Remark 2.3.9, are closely related. The difference is that when adjoining elements, we already have a larger field K given, and we construct $k(\alpha)$ as a subfield of K, while when we adjoin roots of irreducible polynomials, we are building a larger field ourselves. However, once the new field $K = k[x]/(p) \supseteq k$ is constructed, we can write it in the form $K = k(\alpha)$ where $\alpha = x + (p)$. At any rate the field we get is the same up to isomorphism identical on k.

If $\varphi : k \to k'$ is a field homomorphism and $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial in $k[x]$, we write $\varphi(f)$ for the polynomial $\varphi(a_0) + \varphi(a_1)x + \cdots + \varphi(a_n)x^n \in k'[x]$.

Theorem 2.3.15 *Let K/k and K'/k' be field extensions, and $\varphi : k \to k'$ be a field isomorphism. Let $\alpha \in K$ and $\alpha' \in K'$ be elements algebraic over k and k', respectively. Then there exists a field isomorphism $\hat{\varphi} : k(\alpha) \to k'(\alpha')$ extending φ and sending α to α' if and only if* $\operatorname{irr}(\alpha'; k') = \varphi\big(\operatorname{irr}(\alpha; k)\big)$.

Proof If $\hat{\varphi} : k(\alpha) \to k'(\alpha')$ is an isomorphism extending φ and mapping α to α', then the monic polynomial $\varphi(\operatorname{irr}(\alpha; k))$ clearly annihilates α'. To see that $\varphi(\operatorname{irr}(\alpha; k))$ is of minimal possible degree among polynomials annihilating α' use $\hat{\varphi}^{-1}$.

Conversely, by Theorem 2.3.12, we have $k(\alpha) \cong k[x]/(\operatorname{irr}(\alpha; k))$ and $k'(\alpha') \cong k'[x]/(\operatorname{irr}(\alpha'; k'))$, with α and α' corresponding to the cosets $x + (\operatorname{irr}(\alpha; k))$ and $x + (\operatorname{irr}(\alpha'; k'))$, respectively. Now use the universal

property of polynomials to see that there is a ring isomorphism

$$k[x] \to k'[x]/,$$

extending the map $\varphi : k \to k'$ and mapping x to x. It induces an isomorphism $\hat{\varphi}$ of fields

$$k[x]/(\mathrm{irr}\,(\alpha;k)) \to k'[x]/\big(\varphi(\mathrm{irr}\,(\alpha;k))\big),$$

mapping $x + (\mathrm{irr}\,(\alpha;k))$ to $x + \big(\varphi(\mathrm{irr}\,(\alpha;k))\big)$ and extending φ. □

Let $K = k(\alpha)$ be a simple extension. If α is algebraic, we saw in Theorem 2.3.12 that $K \cong k[x]/(\mathrm{irr}\,(\alpha;k))$. The following result investigates the case where α is transcendental.

Theorem 2.3.16 *Let $K = k(\alpha)$ with α transcendental over k, and $k(x)$ be the field of rational functions in x. Then there exists an isomorphism $k(x) \tilde{\to} k(\alpha)$, which is identity on k and maps x to α.*

Proof By the universal properties, there exists a homomorphism $\varphi : k(x) \to k(\alpha)$, which maps $f(x)/g(x)$ to $f(\alpha)/g(\alpha)$. It is non-zero, so injective. It is surjective, as all elements of the form $f(\alpha)/g(\alpha)$ (with non-zero g) form a subfield of K containing k and α. □

Combining Theorems 2.3.12 and 2.3.16 we get

Corollary 2.3.17 *Let K/k be a field extension and $\alpha \in K$. Then α is algebraic over k if and only if $[k(\alpha) : k]$ is finite.*

Theorem 2.3.18 *Let $F/K/k$ be field extensions. If F/K and K/k are algebraic then F/k is algebraic.*

Proof Let $\alpha \in F$ and $\mathrm{irr}\,(\alpha;K) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Then α is algebraic over $k(a_0,\ldots,a_{n-1})$. Now using Corollary 2.3.17 we see that $[k(\alpha,a_0,\ldots,a_{n-1}) : k(a_0,\ldots,a_{n-1})] < \infty$ and $[k(a_0,\ldots,a_{n-1}) : k] < \infty$. By the tower law $[k(\alpha,a_0,\ldots,a_{n-1}) : k] < \infty$, whence α is algebraic. □

Example 2.3.19 There exist algebraic extensions which are not finite. For example, let $\mathbb{A}$ be the set of all complex numbers which are algebraic over $\mathbb{Q}$. We claim that $\mathbb{A}$ is an algebraic extension of $\mathbb{Q}$ of infinite degree.

First of all we need to see that $\mathbb{A}$ is a field. Note by Corollary 2.3.17 that $\alpha \in \mathbb{C}$ belongs to $\mathbb{A}$ if and only if $[\mathbb{Q}(\alpha) : \mathbb{Q}] < \infty$. Let $\alpha, \beta \in \mathbb{A}$.

Note that $\mathbb{Q}(\alpha,\beta) = \mathbb{Q}(\alpha)(\beta)$, and β is algebraic over $\mathbb{Q}$, hence over $\mathbb{Q}(\alpha)$. So $[\mathbb{Q}(\alpha,\beta) : \mathbb{Q}(\alpha)] < \infty$. Therefore $[\mathbb{Q}(\alpha,\beta) : \mathbb{Q}] = [\mathbb{Q}(\alpha,\beta) : \mathbb{Q}(\alpha)][\mathbb{Q}(\alpha) : \mathbb{Q}] < \infty$. As $\mathbb{Q}(\alpha\beta^{-1}), \mathbb{Q}(\alpha-\beta) \subseteq \mathbb{Q}(\alpha,\beta)$, these fields also have finite degrees over $\mathbb{Q}$, and so $\alpha\beta^{-1}, \alpha-\beta \in \mathbb{A}$, which shows that $\mathbb{A}$ is a field. By the way, the argument just given works in larger generality—it shows that given a field extension K/k, the set L of all elements of K which are algebraic over k form a subfield of K.

To show that $[\mathbb{A} : \mathbb{Q}] = \infty$, it suffices to prove that there exists an irreducible polynomial in $\mathbb{Q}[x]$ of arbitrarily large degree. Now take

$$p + px + \cdots + px^{n-1} + x^n$$

and use Eisenstein's criterion.

Example 2.3.20 If the extension K/k is algebraic and k is countable then K is also countable. This follows from the fact that there are only countably many polynomials in $k[x]$. It follows that $\mathbb{R}/\mathbb{Q}$ and $\mathbb{C}/\mathbb{Q}$ are not algebraic, and $\mathbb{A} \subsetneq \mathbb{C}$.

2.4 Ruler and compass

First we formalize the intuitive idea of a ruler and compass construction. Given a set P_0 of points in $\mathbb{R}^2$, consider operations of the following two kinds:

(1) (Ruler) Through any two points of P_0 draw a straight line.
(2) (Compass) Draw a circle whose center is a point in P_0, and whose radius is equal to the distance between some pair of points in P_0.

Definition 2.4.1 The points of intersection of any two distinct lines or circles, drawn using operations (1) or (2), are called *constructible in one step* from P_0. A point $r \in \mathbb{R}^2$ is *constructible* from P_0 if there is a finite sequence $r_1, \ldots, r_n = r$ such that for each $i = 1, \ldots, n$, the point r_i is constructible in one step from the set $P_0 \cup \{r_1, \ldots, r_{i-1}\}$.

Fields enter in the following natural way. Let K_0 be the subfield of $\mathbb{R}$ generated by the x- and y-coordinates of the points in P_0. If $r_=(x_i, y_i)$ then inductively define $K_i = K_{i-1}(x_i, y_i)$.

The following observation is crucial.

Lemma 2.4.2 *With the above notation, x_i and y_i are zeros in K_i of quadratic polynomials over K_{i-1}.*

Proof Write $F := K_{i-1}$. There are three cases to consider: line meets line, line meets circle, and circle meets circle. Let us consider the x-coordinates, the y-coordinates being similar.

Case 1. Let (x_i, y_i) be the intersection point of the lines AB and CD for points A, B, C, D in F^2. It is clear that (x_i, y_i) is a solution of a linear system whose coefficients belong to F, so x_i and y_i also belong to F. Now x_i is of course a solution of the quadratic equation $(x - x_i)^2 = 0$.

Case 2. Let (x_i, y_i) be an intersection of the line through $A = (a_1, a_2), B = (b_1, b_2) \in F^2$ and the circle with the center $C = (c_1, c_2) \in F^2$ of radius s. As s is the distance between the two points with coordinates in F, we have $s^2 \in F$. The equation of the line AB is

$$\frac{x - a_1}{b_1 - a_1} = \frac{y - a_2}{b_2 - a_2},$$

and the equation of the circle is

$$(x - c_1)^2 + (y - c_2)^2 = s^2.$$

Solving these two equations, we obtain

$$(x - c_1)^2 + \left(\frac{(b_2 - a_2)(x - a_1)}{b_1 - a_1} + a_2 - c_2 \right)^2 = s^2,$$

which is a quadratic equation for x.

Case 1. Let (x_i, y_i) be an intersection point of the circles with equations

$$(x - c_1)^2 + (y - c_2)^2 = s^2, \qquad (x - d_1)^2 + (y - d_2)^2 = t^2.$$

We will use greek letters to denote constants in F. From the first equation we have

$$\begin{aligned} (y - d_2)^2 &= (y - c_2 + (c_2 - d_2))^2 \\ &= (y - c_2)^2 + \alpha(y - c_2) + \beta \\ &= s^2 - (x - c_1)^2 + \alpha\sqrt{s^2 - (x - c_1)^2} + \beta \\ &= -(x - c_1)^2 + \alpha\sqrt{\delta - (x - c_1)^2} + \gamma. \end{aligned}$$

Substitute this to the second equation to get

$$(x - d_1)^2 - (x - c_1)^2 + \alpha\sqrt{\delta - (x - c_1)^2} + \gamma - t^2 = 0,$$

which can be rewritten in the form

$$\varepsilon x + \alpha\sqrt{\delta - (x - c_1)^2} + \kappa = 0,$$

and so we again get a quadratic equation on x. □

Theorem 2.4.3 *If $r = (x, y)$ is constructible from a subset P_0 of $\mathbb{R}^2$, and if K_0 is the field generated by $\mathbb{Q}$ and the coordinates of the points in P_0, then the degrees $[K_0(x) : K_0]$ and $[K_0(y) : K_0]$ are powers of 2.*

Proof We will use notation introduced earlier in this section. By Lemma 2.4.2 we have $[K_{i-1}(x_i) : K_{i-1}], [K_{i-1}(y_i) : K_{i-1}] \in \{1, 2\}$. By the Tower Law, $[K_{i-1}(x_i, y_i) : K_{i-1}] \in \{1, 2, 4\}$. By induction, $[K_n : K_0]$ is a power of 2. As $[K_0(x) : K_0]$ and $[K_0(y) : K_0]$ are divisors of $[K_n : K_0]$, they must also be powers of two. □

Corollary 2.4.4 *The cube twice the volume of a given cube cannot constructed using rules-and-compass constructions.*

Proof We may assume that $P_0 = \{(0, 0), (1, 0)\}$ so that $K_0 = \mathbb{Q}$. If we could duplicate the cube then we could construct the point $(0, \alpha)$, where $\alpha^3 = 2$. Therefore, by Theorem 2.4.3, $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ would be a power of 2. But $t^3 - 2$ is irreducible over $\mathbb{Q}$, so $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$. □

Corollary 2.4.5 *The angle $\pi/3$ cannot be trisected using ruler-and-compass constructions.*

Proof To construct the angle trisecting $\pi/3$ is equivalent to constructing the point $(\cos(\pi/9), 0)$ given $(0, 0)$ and $(1, 0)$. From this we could construct $(2\cos(\pi/9), 0)$. Denote $\beta = 2\cos(\pi/9)$. Now the formula $\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta)$ with $\theta = \pi/9$ implies that $\beta^3 - 3\beta - 1 = 0$. So $[\mathbb{Q}(\beta) : \mathbb{Q}] = 3$, as the polynomial $f(x) := x^3 - 3x - 1 = 0$ is irreducible by Example 2.2.4. It remains to apply Theorem 2.4.3. □

In the following result we use the known fact that π is not algebraic over $\mathbb{Q}$.

Corollary 2.4.6 *The circle cannot be squared using ruler-and-compass constructions.*

Proof Such a construction is equivalent to one of the point $(0, \sqrt{\pi})$ from $\{(0, 0), (1, 0)\}$. So $\sqrt{\pi}$ is algebraic, whence π is algebraic, giving a contradiction. □

2.5 What is Galois theory?

We begin with a definition:

Definition 2.5.1 Let K/k be a field extension. An automorphism of K is called *k-automorphism* if it fixes the subfield k element-wise.

It is clear that the set of all k-automorphisms forms a group.

Definition 2.5.2 The *Galois group* $\mathrm{Gal}(K/k)$ is the group of all k-automorphisms of K.

Example 2.5.3 (i) $\mathrm{Gal}(\mathbb{C}/\mathbb{R}) \cong C_2$ with the only non-trivial element given by complex conjugation. Indeed, let σ be an $\mathbb{R}$-automorphism of $\mathbb{C}$. Set $j = \sigma(i)$. Then $j^2 = -1$. Hence $j = i$ or $-i$. It is easy to see that in the first case $\sigma = \mathrm{id}$, and in the second case σ is complex conjugation. Finally, it is easy to check that complex conjugation is indeed an $\mathbb{R}$-automorphism of $\mathbb{C}$.

(ii) $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}) = \{1\}$. Indeed, any element σ of the Galois group must clearly move $\sqrt[3]{2}$ to a primitive third root of 1 in $\mathbb{C}$. However, $\sqrt[3]{2}$ is the only such root in $\mathbb{Q}(\sqrt[3]{2})$, so $\sigma(\sqrt[3]{2}) = \sqrt[3]{2}$. It follows that $\sigma = \mathrm{id}$.

The main idea of Galois theory is that under certain natural assumptions there is a one-to-one correspondence between subgroups of $\mathrm{Gal}(K/k)$ and *intermediate subfields* of the extension K/k, i.e. subfields L of K containing k. This correspondence has many nice properties and is constructed as follows. To each intermediate subfield L we associate the subgroup

$$L^* := \{\sigma \in \mathrm{Gal}(K/k) \mid \sigma(a) = a \text{ for every } a \in L\} \leq \mathrm{Gal}(K/k).$$

Clearly, L^* is nothing but $\mathrm{Gal}(K/L)$, and $L \subseteq M$ implies $L^* \geq M^*$. Conversely, to each subgroup H of $\mathrm{Gal}(K/k)$ one associates the subfield

$$H^* := \{a \in K \mid \sigma(a) = a \text{ for every } \sigma \in H\} \subseteq K.$$

It is easy to check that H^* is an intermediate subfield of the extension K/k, and that $H \leq E$ implies $H^* \supseteq E^*$. Finally, the following inclusions are also clear from definitions:

$$H \leq (H^*)^*, \quad L \subseteq (L^*)^*.$$

The extra conditions needed for Galois theory to work are *separability* and *normality*. They will be studied in the next section. Already

now you can see that something may go wrong if you look at Example 2.5.3(ii).

2.6 Normality and separability

We start from the notion of a splitting field which is closely related to the idea of normality.

Definition 2.6.1 Let K/k be a field extension and $f \in k[x]$ be a polynomial. We say that f *splits over* K if over K it can be written as a product of linear factors:

$$f(x) = a(x - \alpha_1) \dots (c - \alpha_n) \qquad (a, \alpha_1, \dots, \alpha_n \in K).$$

Definition 2.6.2 Let k be a field and $f \in k[x]$. A field $K \supseteq k$ is called a *splitting field* for f over k if f splits over K and $K = k(\alpha_1, \dots, \alpha_n)$ where $\alpha_1, \dots, \alpha_n$ are the roots of f in K.

Example 2.6.3 (i) The polynomial $x^3 - 1 \in \mathbb{Q}[x]$ splits over $\mathbb{C}$:

$$x^3 - 1 = (x - 1)(x - e^{2\pi i/3})(x - e^{4\pi i/3}).$$

(ii) The field $\mathbb{Q}(i, \sqrt{5})$ is a splitting field for $f(x) = (x^2 + 1)(x^2 - 5)$. Note using the Tower Law that $[\mathbb{Q}(i, \sqrt{5}) : \mathbb{Q}] = 4$.

Example 2.6.4 Let p be a prime. The polynomial $f(x) = x^p - 2 \in \mathbb{Q}[x]$ is irreducible by Eisenstein's criterion. It has one real root $\sqrt[p]{2}$. We have $[\mathbb{Q}(\sqrt[p]{2}) : \mathbb{Q}] = p$. Let ω be a primitive pth root of 1. It is easy to see that $\{\omega^i \sqrt[p]{2} \mid 0 \le i < p\}$ are exactly all the roots of f. Thus, $\mathbb{Q}(\sqrt[p]{2}, \omega)$ is a splitting field for f. We want to find the degree $[\mathbb{Q}(\sqrt[p]{2}, \omega) : \mathbb{Q}]$. By the tower law, it is divisible by both $p = [\mathbb{Q}(\sqrt[p]{2} : \mathbb{Q}]$ and $p - 1 = [\mathbb{Q}(\omega) : \mathbb{Q}]$ (see Example 2.2.5). So $[\mathbb{Q}(\sqrt[p]{2}, \omega) : \mathbb{Q}] \ge p(p - 1)$. On the other hand, $[\mathbb{Q}(\omega)(\sqrt[p]{2}) : \mathbb{Q}(\omega)] \le p$. This proves that $[\mathbb{Q}(\sqrt[p]{2}, \omega) : \mathbb{Q}] = p(p - 1)$.

Example 2.6.5 $f(x) = x^6 - 1$ factorizes as

$$f(x) = (x - 1)(x^2 + x + 1)(x + 1)(x^2 - x + 1).$$

If ω is a root of $x^2 + x + 1$ then

$$f(x) = (x - 1)(x - \omega)(x - \omega^2)(x + 1)(x + \omega)(x + \omega^2).$$

It follows that $\mathbb{Q}(\omega)$ is a splitting field for f and $[\mathbb{Q}(\omega) : \mathbb{Q}] = 2$.

Our next goal is to prove that for any polynomial splitting field exists and is in some sense unique. The proof of existence is obtained by successive adjoining of roots of irreducible polynomials in the sense of Remark 2.3.9:

Theorem 2.6.6 *If k is any field and $f \in k[x]$ is any polynomial then there exists a splitting field for f over k.*

Proof We use induction on $\deg f$. If $\deg f \leq 1$, there is nothing to prove for f splits over k. Let $\deg f > 1$. If f does not split over k, it has an irreducible factor f_1 of degree > 1. Adjoin a root of f_1 to k to get a field $k(\alpha_1) \cong k[x]/(f_1)$, see Remark 2.3.14. Then in $k(\alpha_1)[x]$ we have $f = (x - \alpha_1)g(x)$. By induction, there is a splitting field F for g over $k(\alpha_1)$. Now note that F is a splitting field for f over k. □

We need the following lemma to prove uniqueness of splitting fields.

Lemma 2.6.7 *Let $\varphi : k \to k'$ be an isomorphism of fields, $f \in k[x]$, and let $K \supseteq k$ be a splitting field for f over k. If $K' \supseteq k'$ is a field extension of k' such that $\varphi(f)$ splits over K' then there exists a monomorphism $\hat{\varphi} : K \to K'$, which extends φ.*

Proof Induction on $n := \deg f$. Over K we have

$$f(x) = a(x - \alpha_1) \dots (x - \alpha_n).$$

Set $g := \operatorname{irr}(\alpha_1; k)$. Then g is an irreducible factor of f. It follows that $\varphi(g)$ is an irreducible factor of $\varphi(f)$. Moreover, $\varphi(f)$ can be written over K' as a product of linear factors: $\varphi(f) = b(x-\beta_1) \dots (x-\beta_n)$. It follows that $\varphi(g)$ is a product of some $(x - \beta_i)$. Without loss of generality we may assume that $(x-\beta_1)$ is a factor of $\varphi(g)$. Then $\varphi(g) = \operatorname{irr}(\beta_1; k')$. By Theorem 2.3.15, there is an isomorphism $\tilde{\varphi} : k(\alpha_1) \to k'(\beta_1)$, extending φ. Now, K is a splitting field of the polynomial $f(x)/(x-\alpha_1) \in k(\alpha_1)[x]$. By induction, there exists a homomorphism $\hat{\varphi} : K \to K'$ extending $\tilde{\varphi}$. □

Now we are ready to prove the desired uniqueness theorem for splitting fields.

Theorem 2.6.8 *Let $\varphi : k \to k'$ be a field isomorphism. Let K be a splitting field for $f \in k[x]$ over k and K' be a splitting field for $\varphi(f) \in k'[x]$ over k'. Then there exists an isomorphism $\hat{\varphi} : K \to K'$, extending φ.*

Proof By Lemma 2.6.7, there exists a monomorphism $\hat{\varphi} : K \to K'$, extending φ. Moreover $\varphi(K) \subseteq K'$ are both splitting fields for $\varphi(f)$, so $\varphi(K) = K'$. □

Example 2.6.9 The same field can be a splitting field for two different polynomials. For example, $\mathbb{Q}(\sqrt{3}) \subset \mathbb{C}$ is a splitting field for $x^2 - 3$ and $x^2 - 2x - 2$ over $\mathbb{Q}$.

We now arrive to the notion of normality.

Definition 2.6.10 An extension K/k is *normal* if every irreducible polynomial over k which has at least one zero in K splits over K.

Example 2.6.11 (i) $\mathbb{C}/\mathbb{R}$ is normal as every polynomial in $\mathbb{R}[x]$ splits over $\mathbb{C}$ (Fundamental Theorem of Algebra).

(ii) $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ is not normal, as it contains the root $\sqrt[3]{2}$ of the irreducible polynomial $x^3 - 2 \in \mathbb{Q}[x]$, but does not contain the other two roots of the same polynomial.

Compare Example 2.6.11 with Example 2.5.3. The 'bad' behavior of the Galois group $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q})$ occurs because the extension is not normal.

The following result explains the close connection between normal extensions and splitting fields and provides us with a wide range of normal extensions.

Theorem 2.6.12 *An extension K/k is normal and finite if and only if K is a splitting field for some polynomial over k.*

Proof Suppose K/k is normal and finite. Write $K = k(\alpha_1, \ldots, \alpha_n)$ for algebraic elements α_i over k. Let $f_i = \mathrm{irr}\,(\alpha_i; k)$, and $f = f_1 \ldots f_n$. By normality, f splits over K, so K is a splitting field for f.

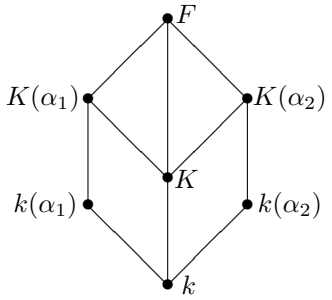
Conversely, let K be a splitting field for $g \in k[x]$. The extension K/k is then obviously finite. To show that it is normal, take an irreducible polynomial $f \in k[x]$ with a root in K. Let $F \supseteq K$ be a splitting field for f over K.

Suppose that α_1 and α_2 are two roots of f in F. We need to prove that $\alpha_1 \in K$ implies $\alpha_2 \in K$. This will follow if we can show that

$$[K(\alpha_1) : K] = [K(\alpha_2) : K]. \tag{2.1}$$

Indeed, if $\alpha_1 \in K$, then $[K(\alpha_1) : K] = 1$, so by (2.1), $[K(\alpha_2) : K] = 1$, and hence $\alpha_2 \in K$.

For the proof of (2.1), consider several subfields of F whose inclusions are illustrated by the following diagram:



For $i = 1, 2$, we have

$$[K(\alpha_i) : K][K : k] = [K(\alpha_i) : k] = [K(\alpha_i) : k(\alpha_i)][k(\alpha_i) : k]. \qquad (2.2)$$

As $\operatorname{irr}(\alpha_1; k) = \operatorname{irr}(\alpha_2; k) = f$, there is an isomorphism $\varphi : k(\alpha_1) \to k(\alpha_2)$. Moreover, $K(\alpha_i)$ is a splitting for g over $k(\alpha_i)$, so by Theorem 2.6.8, φ extends to an isomorphism $\hat{\varphi} : K(\alpha_1) \to K(\alpha_2)$. In particular, $[K(\alpha_1) : k] = [K(\alpha_2) : k]$. Now (2.1) follows from (2.2). □

Corollary 2.6.13 *Let $F/K/k$ be finite field extensions, K/k be normal, and F/K be a splitting field over K of a polynomial $f(x) \in K[x]$ all of whose coefficients belong to k. Then F/k is normal.*

Proof We just need to apply Theorem 2.6.12 twice. Indeed, as K/k is normal it is a splitting field of a polynomial $g \in k[x]$. Now F is a splitting field over k of the polynomial $fg \in k[x]$. □

Example 2.6.14 It is not true in general that being normal is transitive; that is if $F/K/k$ and F/K, K/k are normal, then F/k need not be normal. For a counterexample consider $\mathbb{Q}(\sqrt[4]{2})/\mathbb{Q}(\sqrt{2})/\mathbb{Q}$.

Now we define the second important property needed for the Galois theory to work well.

Definition 2.6.15 An irreducible polynomial f over a field k is called *separable* over k if it has no multiple roots in a splitting field. Otherwise f is called *inseparable*.

An arbitrary polynomial over a field k is called *separable* over k if all its irreducible factors are separable over k.

If K/k is a field extension, then an algebraic element $\alpha \in K$ is called *separable* over k if its minimal polynomial over k is separable over k.

An algebraic extension K/k is called *separable* if every $\alpha \in K$ is separable over k.

In view of the uniqueness of splitting fields, the notion just defined does not depend on which splitting field is used.

Example 2.6.16 We give an example of an inseparable polynomial. Let p be a prime, and $k = \mathbb{F}_p(u)$, the field of fractions of $\mathbb{F}_p[u]$, where u is an indeterminate. Consider the polynomial $f(x) := x^p - u \in k[x]$. Let K be a splitting field of f over k, and $\alpha \in K$ be a root of f. Then $\alpha^p = u$, hence $(x-\alpha)^p = x^p - \alpha^p = f(x)$, using the 'freshman's dream'. This shows that all roots of f are equal.

We now prove that f is irreducible over k. Indeed, assume that $f = gh$, where degrees of g and h are lower that the degree of f. By the previous paragraph, we must have $g(x) = (x-\alpha)^d$ for $0 < d < p$. Hence the constant coefficient $\alpha^d \in k$. As d is prime to p, we can write $1 = rd + sp$, and it follows that $\alpha \in k$, i.e. $\alpha = a(u)/b(u)$ for some $a, b \in \mathbb{F}_p[u]$. Therefore $a(u)^p/b(u)^p = u$ or $a(u)^p = b(u)^p u$, which is impossible.

Let k be an arbitrary field and $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \in k[x]$. The *formal derivative* of f is the polynomial

$$f' := a_1 + 2a_2 x + \cdots + n a_n x^{n-1}.$$

It is easy to see that the formal derivative has the following familiar properties:

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

Lemma 2.6.17 *Let k be any field. A non-zero polynomial $f \in k[x]$ has a multiple root in a splitting field if and only if f and f' have a common factor of degree ≥ 1 as polynomials over k.*

Proof Suppose that f has multiple roots in a splitting field K, so that over K we have $f(x) = (x-\alpha)^2 g(x)$. Then

$$f'(x) = (x-\alpha)(2g(x) + (x-\alpha)g'(x)).$$

So f and f' have a common factor $(x-\alpha)$ over K. Now, the minimal polynomial $\operatorname{irr}(\alpha; k)$ must divide both f and f' over k.

Conversely, suppose f has no repeated roots. We show by induction on the degree of f that f and f' are coprime over K, hence also over k. If $\deg f = 1$ this is obvious. Otherwise $f(x) = (x-\alpha)g(x)$ where $(x-\alpha) \not| g(x)$. Then

$$f'(x) = (x-\alpha)g'(x) + g(x). \tag{2.3}$$

Now, $(x-\alpha)$ does not divide f'. So, if h is a common factor of f and f', h must divide g. By (2.3), h also divides g'. By inductive assumption, g is a scalar. □

Proposition 2.6.18 *Let k be a field and f be an irreducible polynomial over k. Then f is inseparable over k if and only if k has a positive characteristic p and $f(x)$ has the form $a_0 + a_1x^p + \cdots + a_lx^{lp}$.*

Proof Using irreducibility of f, Lemma 2.6.17 implies that f is inseparable over k if and only $f' = 0$. □

The following result shows that separability 'carries over to intermediate fields'.

Lemma 2.6.19 *Let $K/L/k$ be field extensions with K/k being (algebraic) separable. Then L/k and K/L are (algebraic) separable.*

Proof For L/k this is clear. Let $\alpha \in K$. Then $\mathrm{irr}\,(\alpha; L) \mid \mathrm{irr}\,(\alpha; k)$, and the result follows. □

The converse of this theorem is also true. We prove it later in a special case of finite extensions, see Theorem 2.7.20 below.

The Fundamental theorem of Galois theory applies to finite extensions which are normal and separable. It is convenient to give the following

Definition 2.6.20 We say that a field extension is *Galois* if it is finite normal and separable.

2.7 The Fundamental Theorem

The following result is quite amazing.

Lemma 2.7.1 (Dedekind's Lemma) *If K and L are fields then every set of distinct monomorphisms from K to L is linearly independent over L.*

Proof Let n be the smallest number such that n distinct monomorphisms $\lambda_1, \ldots, \lambda_n : K \to L$ are linarly dependent, i.e. there exist $a_1, \ldots, a_n \in L$, all non-zero, such that

$$a_1\lambda_1(x) + \cdots + a_n\lambda_n(x) = 0 \qquad (x \in K). \tag{2.4}$$

As $\lambda_1 \neq \lambda_n$, there exists $y \in K$ with $\lambda_1(y) \neq \lambda_n(y)$. Such y is clearly non-zero. Now, equation (2.4) holds with yx in place of x, so

$$a_1\lambda_1(y)\lambda_1(x) + \cdots + a_n\lambda_n(y)\lambda_n(x) = 0 \qquad (x \in K). \tag{2.5}$$

Multiply equation (2.4) by $\lambda_1(y)$ and subtract equation (2.5) to get an equation of the form (2.4) with fewer terms. This contradicts the minimality of n. □

Corollary 2.7.2 *Let K/k be a finite extension. Then* $\mathrm{Gal}(K/k)$ *is finite.*

Proof Let $d = [K : k]$. Then the dimension of the vector space of k-linear maps from K to itself is d^2. So any $n^2 + 1$ elements of $\mathrm{Gal}(K/k)$ are k-lineary dependent, hence K-linearly dependent. But this contradicts Dedekind's Lemma. □

Theorem 2.7.3 *Let G be a finite subgroup of the group of automorphisms of a field K, and let k be the fixed field of G. Then $[K : k] = |G|$.*

Proof Let $n := |G|$, and let $1 = \sigma_1, \sigma_2, \ldots, \sigma_n$ are the elements of G.

Suppose first that $[K : k] =: m < n$, and let $\{\alpha_1, \ldots, \alpha_m\}$ be a basis for K over k. The system

$$\sum_{i=1}^{n} \sigma_i(\alpha_j)x_i = 0 \qquad (1 \leq j \leq m) \tag{2.6}$$

of m linear equations with unknowns $x_1, \ldots, x_n$ has a non-trivial solution $(\beta_1, \ldots, \beta_n)$. Let α be any element of K. Then $\alpha = a_1\alpha_1 + \cdots + a_m\alpha_m$ with $a_j \in k$. Using (2.6), we obtain

$$\sum_{i=1}^{n} \sigma_i(\alpha)\beta_i = \sum_{i=1}^{n} \sigma_i(\sum_{j=1}^{m} a_j\alpha_j)\beta_i = \sum_{j=1}^{m} a_j \sum_{i=1}^{n} \sigma_i(\alpha_j)\beta_i = 0.$$

Hence distinct monomorphisms $\sigma_1, \ldots, \sigma_n$ are linearly dependent, contrary to Dedekind's Lemma. Therefore $m \geq n$.

Next suppose that $m > n$. Then there exists a set of $n + 1$ elements

$\{\alpha_1, \ldots, \alpha_{n+1}\}$ of K linearly independent over k. The system

$$\sum_{i=1}^{n+1} \sigma_j(\alpha_i) x_i = 0 \qquad (1 \leq j \leq n) \tag{2.7}$$

of n linear equations with unknowns $x_1, \ldots, x_{n+1}$ has non-trivial solutions. Let $(\beta_1, \ldots, \beta_{n+1})$ be a non-trivial solution with minimal possible amount of non-zeros. By renumbering β's and α's, we may assume that $\beta_1, \ldots, \beta_r \neq 0$ and $\beta_{r+1} = \cdots = \beta_{n+1} = 0$. Then (2.7) gives

$$\sum_{i=1}^{r} \sigma_j(\alpha_i) \beta_i = 0 \qquad (1 \leq j \leq n). \tag{2.8}$$

Operating on (2.8) with $\sigma \in G$ yields

$$\sum_{i=1}^{r} \sigma\sigma_j(\alpha_i) \sigma(\beta_i) = 0 \qquad (1 \leq j \leq n).$$

As $\{\sigma\sigma_j \mid 1 \leq j \leq n\} = \{\sigma_j \mid 1 \leq j \leq n\}$, this system is equivalent to

$$\sum_{i=1}^{r} \sigma_j(\alpha_i) \sigma(\beta_i) = 0 \qquad (1 \leq j \leq n). \tag{2.9}$$

Multiply the equation (2.8) by $\sigma(\beta_1)$ and subtract the equation (2.9) multiplied by β_1 to get

$$\sum_{i=2}^{r} \sigma_j(\alpha_i)(\sigma(\beta_1)\beta_i - \sigma(\beta_i)\beta_1) = 0 \qquad (1 \leq j \leq n).$$

This is a system of linear equation like (2.7) with $r-1$ terms. This gives a contradiction with the choice of r unless all solutions $\sigma(\beta_1)\beta_i - \sigma(\beta_i)\beta_1$ are zero. If this happens then $\beta_i\beta_1^{-1} = \sigma(\beta_i\beta_1^{-1})$ for all $\sigma \in G$, i.e. $b_i := \beta_i\beta_1^{-1} \in k$. Thus (2.8) with $j = 1$ is $\sum_{i=1}^{r} \alpha_i b_i \beta_1 = 0$ or $\sum_{i=1}^{r} b_i \alpha_i = 0$ As all b_i are non-zero, this contradicts the linear independence of the α_i. Therefore $m \leq n$. □

Corollary 2.7.4 *Let K/k be a finite extension and $G = \mathrm{Gal}(K/k)$. If $H < G$ is a subgroup then $|H|[H^* : k] = [K : k]$.*

Proof By Theorem 2.7.3, $|H|[H^* : k] = [K : H^*][H^* : k] = [K : k]$. □

Definition 2.7.5 Let K/k and L/k be two field extensions. Then a k-monomorphism of K into L is a field monomorphism $K \to L$ which is identity on k.

Example 2.7.6 Let $k = \mathbb{Q}$, $K = \mathbb{Q}(\sqrt[3]{2})$, and $L = \mathbb{C}$. Note that any k-monomorphism φ of K into L must map $\sqrt[3]{2}$ to a cubic root of 2. Moreover, φ is uniquely determined by the choice of $\varphi(\sqrt[3]{2})$. So there are at most three different k-monomorphisms of K into L. In fact, there are exactly three, as all cubic roots of 2 have the same minimal polynomial, and so Theorem 2.3.15 applies.

In general if $L/K/k$ are field extensions then any k-automorphism of L restricts to a k-monomorphism from K to L. We are interested in when the process can be reversed.

Theorem 2.7.7 *Let $N/K/k$ be field extensions, and suppose that N/k is finite and normal. Then any k-monomorphism τ of K into N can be extended to a k-automorphism of N.*

Proof By Theorem 2.6.12, N is a splitting field over k of some polynomial f over k. Hence N is a splitting field of f over both K and $\tau(K)$, and $\tau(f) = f$. It remains to apply Theorem 2.6.8. □

Proposition 2.7.8 *Let N/k be a finite normal extension, and $\alpha, \beta \in N$ be roots of the same irreducible polynomial over k. Then there exists $\sigma \in \mathrm{Gal}(N/k)$ such that $\sigma(\alpha) = \beta$.*

Proof Follows from Theorems 2.3.15 and 2.7.7. □

Definition 2.7.9 Let K/k be an algebraic extension. A *normal closure* of K/k is an extension N/K such that (1) N/k is normal, and (2) if $N/L/K$ is an intermediate extension with L/k normal then $L = N$.

Informally speaking, the normal closure is obtained by adjoining all the 'missing' roots.

Theorem 2.7.10 *If K/k is a finite extension then there exists a normal closure N of K/k with N/k finite. If M is another normal closure of K/k then the extensions N/k and M/k are isomorphic.*

Proof Let $\alpha_1, \ldots, \alpha_d$ be a basis of K over k and let $f_i = \mathrm{irr}\,(\alpha_i; k)$ for every i. Let N be a splitting field for $f = f_1 f_2 \ldots f_d$. Then N/k is normal and finite by Theorem 2.6.12. Suppose that $N/L/K$ and L/k is normal. Each polynomial f_i has a root $\alpha_i \in L$. So by normality f splits in L. Since N is a splitting field for f, we have $N = L$.

Now suppose N and M are both normal closures. The above polynomial splits in both M and N, so each M and N contains a splitting field for f over k. These splitting fields contain K and are normal over k, so must equal to M and N respectively. The uniqueness of splitting fields now says that M/k and N/k are isomorphic. □

Example 2.7.11 The extension $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ is not normal. Consider the field $N := \mathbb{Q}(\sqrt[3]{2}, \omega\sqrt[3]{2}, \omega\sqrt[3]{2})$ where $\omega = e^{2\pi i/3}$. Then N is a normal closure of $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$. Indeed, N is normal, because it is a splitting field for the polynomial $x^3 - 2$, and clearly no intermediate field extension is normal.

Lemma 2.7.12 *Suppose that $F/N/K/k$ are field extensions, with K/k finite and N a normal closure of K/k. Let τ be any k-monomorphism from K into F. Then $\tau(K) \subseteq N$.*

Proof Let $\alpha \in K$. Applying τ to $\operatorname{irr}(\alpha; k)(\alpha) = 0$, we see that $\tau(\alpha)$ is a root of $\operatorname{irr}(\alpha; k)$. So $\tau(\alpha)$ lies in N. □

Theorem 2.7.13 *For a finite extension K/k the following are equivalent:*

(i) *K/k is normal;*

(ii) *there exists a normal extension N of k containing K such that every k-monomorphism τ from K into N is a k-automorphism of K.*

(iii) *for every extension F of k containing K every k-monomorphism τ from K into F is a k-automorphism of K.*

Proof (i) ⇒ (iii) If K/k is normal then K is a normal closure of K/k so by Lemma 2.7.12 we have $\tau(K) \subseteq K$, whence $\tau(K) = K$ by considering dimensions.

(iii) ⇒ (ii) Take N to be a normal closure of K/k.

(ii) ⇒ (i) Let $f \in k[x]$ be an irreducible polynomial with a root in K. Then f splits over N by normality, and if β is any root of f in N there exists $\sigma \in \operatorname{Gal}(N/k)$ such that $\sigma(\alpha) = \beta$, see Proposition 2.7.8. By hypothesis, $\sigma|_K \in \operatorname{Gal}(K/k)$, so $\beta = \sigma(\alpha) \in K$, and hence f splits over K. Therefore K is normal. □

Theorem 2.7.14 *Let K/k be a finite separable extension of degree d. Then there are precisely d distinct k-monomorphisms of K into a normal closure N and hence into any given normal extension M containing K.*

Proof Induction on d, the case $d = 1$ being clear. Let $d > 1$. Let $\alpha \in K \setminus k$ and $f := \operatorname{irr}(\alpha; k)$. If $r := \deg f$ then $[k(\alpha) : k] = r$. Now f is a separable irreducible polynomial in the normal extension N, so f splits in N and its roots $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_r$ are distinct.

Let $s := [K : k(\alpha)] = d/r$. By induction, there are precisely s distinct $k(\alpha)$-monomorphisms $\rho_1, \ldots, \rho_s : K \to N$. By Proposition 2.7.8, there are r distinct k-automorphisms $\tau_1, \ldots, \tau_r$ of N, such that $\tau(\alpha) = \alpha_i$. The maps $\varphi_{ij} = \tau_i \rho_j$ give rs distinct k-monomorphisms $K \to N$.

Now let $\varphi : K \to N$ be a k-monomorphism. Then $\varphi(\alpha)$ is a root of f in N, so that $\varphi(\alpha) = \alpha_i$ for some i. The map $\psi = \tau_i^{-1}\varphi$ is a $k(\alpha)$-monomorphism $K \to N$, so by induction $\psi = \rho_j$ for some j. Hence $\varphi = \varphi_{ij}$. □

Corollary 2.7.15 *If K/k is a separable normal extension of degree $d < \infty$ then $|\mathrm{Gal}(K/k)| = d$.*

Proof Use Theorem 2.7.14 and 2.7.13. □

Theorem 2.7.16 *If K/k is a Galois extension with Galois group G then k is the fixed field of G.*

Proof Let K_0 be the fixed field of G and $[K : k] = d$. By Corollary 2.7.15, $|G| = d$. By Theorem 2.7.3, $[K : K_0] = d$. It remains to notice that $k \subseteq K_0$. □

Theorem 2.7.17 *Let $K/L/k$ be field extensions with $[K : k] < \infty$ and $[L : k] = n$. Then there are at most n k-monomorphisms of L into K. Moreover, if L/k is not separable there are strictly less than n k-monomorphisms of L into K.*

Proof Let N be a normal closure of K/k. Then N/k is finite and every k-monomorphism of L into K is also a k-monomorphism of L into N. We now argue by induction on $[L : k]$ as in the proof of Theorem 2.7.14, except that we can now only deduce that there are s' $k(\alpha)$-monomorphisms $L \to N$, where $s' \leq s$ (by induction) and there are r'

distinct k-automorphisms of N where $r' \leq r$ (since the roots of f in N need not be distinct). The rest of the argument goes through.

Finally, if L/k is not separable then $r' < r$, and the second claim of the theorem follows. □

The following result is converse to Theorem 2.7.16. It shows that we *have* to consider normal and separable extensions if we want Galois theory to 'work well'.

Theorem 2.7.18 *If K/k is a finite extension with Galois group G such that k is the fixed field of G then K/k is Galois.*

Proof By Theorem 2.7.3, $[K : k] = |G| = n$, say. There are exactly n distinct k-monomorphisms $K \to K$, namely the elements of the Galois group. By Theorem 2.7.17, K/k is separable.

Let N be a normal extension of K containing k. By Theorem 2.7.17, there are at most n k-monomorphisms from K to N, and so they are just the k-automorphisms of K (followed by the embedding of K into N). By Theorem 2.7.13, K/k is normal. □

The results obtained above can be obtained to deduce some useful facts about separable extensions.

Lemma 2.7.19 *Suppose we have a chain $k = k_0 \subseteq k_1 \subseteq \ldots$ of field extensions where $k_i = k_{i-1}(\alpha_i)$ for $\alpha_i \in k_i$ that is algebraic over k_{i-1} and whose minimal polynomial over k_{i-1} is separable. Then k_i/k is separable for all i.*

Proof Proceed by induction on i, the case $i = 0$ being trivial. Consider k_i/k assuming inductively that k_{i-1}/k is separable. Let $f_j = \operatorname{irr}(\alpha_j; k)$, $j = 1, 2, \ldots$, and $f = \operatorname{irr}(\alpha_i; k_{i-1})$. Let $K \supseteq k_i$ be a splitting field for $f_1 \ldots f_i$ over k. By inductive hypothesis and Theorem 2.7.14, there are exactly $[k_{i-1} : k]$ k-monomorphisms $k_{i-1} \to K$. Take any of these, τ, say. Note that $f \mid f_i$ and f_i splits in K, so f and $\tau(f)$ split in K. By separability of f we know that the number of roots of $\tau(f)$ equals $\deg f = [k_i : k_{i-1}]$. So by Theorem 2.3.15 there are $[k_i : k_{i-1}]$ extensions of τ to a monomorphism $k_i \to K$. Hence in total there are $[k_i : k] = [k_i : k_{i-1}][k_{i-1} : k]$ k-monomorphisms $k_i \to k$. This implies by Theorem 2.7.17 that k_i/k is separable. □

Theorem 2.7.20 (Transitivity of Separable Extensions) *If F/K and K/k are finite separable field extensions then so is F/k.*

Proof Write $F = K(\alpha_1, \ldots, \alpha_n)$ and $K = k(\alpha_{n+1}, \ldots, \alpha_m)$. Set $k_i = k(\alpha_1, \ldots, \alpha_i)$. Then each $\operatorname{irr}(\alpha_i; k_{i-1})$ is separable. Indeed, for $i \leq n$ we use the fact that $\operatorname{irr}(\alpha_i; k_{i-1}) \mid \operatorname{irr}(\alpha_i; k)$ and for $i > n$ we use the fact that $\operatorname{irr}(\alpha_i; k_{i-1}) \mid \operatorname{irr}(\alpha_i; K)$. Now we can apply Lemma 2.7.19. □

We are now in a position to establish the fundamental properties of the Galois correspondence. Let K/k be a field extension with Galois group G. Let $\mathcal{F}$ be the set of intermediate fields F in K/k, and $\mathcal{G}$ be the set of all subgroups $H \leq G$. In §2.5 we have defined two maps

$$^* : \mathcal{F} \to \mathcal{G},\ F \mapsto F^*, \qquad ^* : \mathcal{G} \to \mathcal{F},\ H \mapsto H^*. \tag{2.10}$$

Theorem 2.7.21 (Fundamental Theorem of Galois Theory) *Let K/k be a Galois extension of degree n with Galois group G. Then*

(i) $|G| = n$;
(ii) *The maps* (2.10) *are mutual inverses; they set up an order-reversing one-to-one correspondence between $\mathcal{F}$ and $\mathcal{G}$.*
(iii) *If $F \in \mathcal{F}$ then $[K : F] = |F^*|$ and $[F : k] = [G : F^*]$.*
(iv) *$F \in \mathcal{F}$ is a normal extension of k if and only if $F^* \lhd G$.*
(v) *If $F \in \mathcal{F}$ is a normal extension of k then $\operatorname{Gal}(F/k) \cong G/F^*$.*

Proof (i) is proved in Corollary 2.7.15.

(ii) Let $F \in \mathcal{F}$. By Lemma 2.6.19, the extension K/F is separable and from Theorem 2.6.12 it is normal. Therefore by Theorem 2.7.16, F is the fixed field of F^*, i.e.

$$(F^*)^* = F. \tag{2.11}$$

Now, let $H \leq G$. It is clear that $H \leq (H^*)^*$. Moreover $((H^*)^*)^* = H^*$ by (2.11). So applying Theorem 2.7.3 twice we have

$$|H| = [K : H^*] = [K : ((H^*)^*)^*] = |(H^*)^*|,$$

hence $H = (H^*)^*$.

(iii) We have proved above that K/F is Galois. By Corollary 2.7.15, $[K : F] = |F^*|$, and the other equality follows immediately.

(iv) First observe that for any $\sigma \in G$ and $F \in \mathcal{F}$ we have

$$\sigma(F)^* = \sigma(F^*)\sigma^{-1}. \tag{2.12}$$

Now, if F/k is normal then $\sigma(F) = F$ for any $\sigma \in G$ thanks to

Theorem 2.7.13. So F^* is normal by (2.12). Conversely, suppose $F^* \triangleleft G$. Let τ be a monomorphism $F \to K$. By Theorem 2.7.7, there is $\sigma \in G$ such that $\sigma|_F = \tau$. By (2.12), $\sigma(F^*)\sigma^{-1} = F^*$ implies $\sigma(F)^* = F^*$. By (ii), $\sigma(F) = F$, hence $\tau(F) = F$. Now, F/k is normal by Theorem 2.7.13.

(v) We define a group homomorphism

$$\varphi : G \to \mathrm{Gal}(F/k),\ \sigma \mapsto \sigma|_F,$$

using Theorem 2.7.13. By Theorem 2.7.7, φ is surjective, and its kernel is obviously F^*. □

2.8 Galois group of a polynomial

Definition 2.8.1 Let $f \in k[x]$ be a polynomial and K/k be its splitting field. The *Galois group* $\mathrm{Gal}(f;k)$ of f over k is defined to be $\mathrm{Gal}(K/k)$.

By 'uniqueness' of splitting fields, $\mathrm{Gal}(f;k)$ is well defined up to isomorphism. Note also that $\mathrm{Gal}(f;k)$ is naturally a permutation group on the set of the roots of f in the splitting field K. Indeed, it is clear that elements of $\mathrm{Gal}(f;k)$ permute the roots and that an element $\sigma \in \mathrm{Gal}(f;k)$ gives a trivial permutation if and only if $\sigma = 1$. In fact, if f is irreducible then $\mathrm{Gal}(f;k)$ is a transitive permutation group, as follows from Theorems 2.3.15 and 2.7.7.

In general it is difficult to calculate explicitly Galois groups of polynomials and field extensions. Here is a 'baby' example of such calculation.

Example 2.8.2 Let $f(x) = x^4 - 2 \in \mathbb{Q}[x]$. Choose a splitting field K of f so that $K \subseteq \mathbb{C}$. Let $G = \mathrm{Gal}(K/\mathbb{Q})$. Denote $\alpha = \sqrt[4]{2}$. Then over $\mathbb{C}$

$$f(x) = (x-\alpha)(x+\alpha)(x-i\alpha)(x+i\alpha).$$

So $K = \mathbb{Q}(\alpha, i)$. As characteristic is 0, K is separable, and K is normal because it is a splitting field. Finally, $K/\mathbb{Q}$ is finite. So Galois theory applies.

First we determine $[K:\mathbb{Q}]$. We have $[K:\mathbb{Q}] = [K:\mathbb{Q}(\alpha)][\mathbb{Q}(\alpha):\mathbb{Q}]$. As f is irreducible be Eisenstein's criterion, $[\mathbb{Q}(\alpha):\mathbb{Q}] = 4$. Moreover, $[K:\mathbb{Q}(\alpha)] > 1$ and i satisfies the equation $x^2 + 1 = 0$ of degree 2, so $[K:\mathbb{Q}(\alpha)] = 2$. Thus, $[K:\mathbb{Q}] = 8$.

We claim that there are $\sigma, \tau \in G$ such that

$$\sigma(i) = i,\ \sigma(\alpha) = i\alpha, \qquad \tau(i) = -i,\ \tau(\alpha) = \alpha.$$

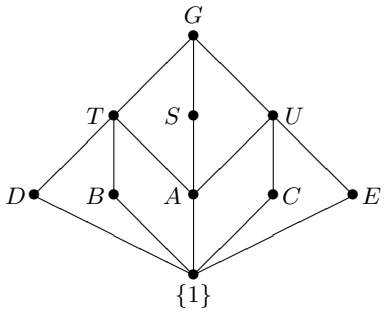
For example, let us show that there exists σ. By Theorem 2.3.15,

there exists a $\mathbb{Q}$-isomorphism $\bar{\sigma} : \mathbb{Q}(\alpha) \to \mathbb{Q}(i\alpha)$ with $\bar{\sigma}(\alpha) = i\alpha$. As $\operatorname{irr}(\mathbb{Q}(\alpha); i) = \operatorname{irr}(\mathbb{Q}(i\alpha); i)$, the same theorem now implies that there exists an automorphism σ of K extending $\bar{\sigma}$ and mapping i to i.

By looking at the images of the elements α and i we see that

$$\{\sigma^i\tau^j \mid 0 \le i \le 3,\ 0 \le j \le 1\}$$

are distinct elements of G. It follows that $G \cong D_8$. The subgroup structure of D_8 is represented by the following picture:



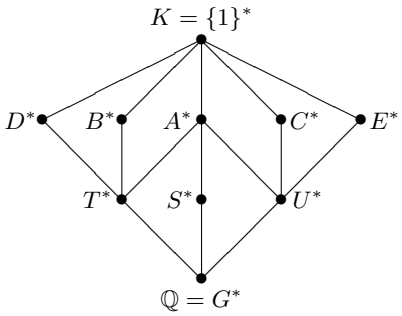
Here the subgroups of order 2 are

$$D = \langle\sigma^2\tau\rangle,\ B = \langle\tau\rangle,\ A = \langle\sigma^2\rangle,\ C = \langle\sigma\tau\rangle,\ E = \langle\sigma^3\tau\rangle,$$

and subgroups of order 4 are

$$T = \langle\sigma^2, \tau\rangle,\ S = \langle\sigma\rangle,\ U = \langle\sigma^2, \sigma\tau\rangle$$

Under the Galois correspondence we obtain the intermediate fields



We now describe these intermediate fields explicitly. It is clear that

$$S^* = \mathbb{Q}(i),\ T^* = \mathbb{Q}(\sqrt{2}),\ U^* = \mathbb{Q}(i\sqrt{2}).$$

Indeed, the fields have degree 2 over 2, and are clearly fixed by the corresponding subgroups.

Now let us find C^*. Any element of K can be uniquely expressed in the form

$$x = a_0 + a_1\alpha + a_2\alpha^2 + a_3\alpha^3 + a_4 i + a_5 i\alpha + a_6 i\alpha^2 + a_7 i\alpha^3 \quad (a_j \in \mathbb{Q}).$$

Then

$$\begin{aligned} \sigma\tau(x) &= a_0 + a_1 i\alpha - a_2\alpha^2 - a_3 i\alpha^3 - a_4 i + a_5\alpha + a_6 i\alpha^2 - a_7\alpha^3 \\ &= a_0 + a_5\alpha - a_2\alpha^2 - a_7\alpha^3 - a_4 i + a_1 i\alpha + a_6 i\alpha^2 - a_3 i\alpha^3. \end{aligned}$$

Therefore x is fixed by $\sigma\tau$ if and only if $a_2 = a_0 = 0$, $a_1 = a_5$, and $a_3 = -a_7$, i.e.

$$\begin{aligned} \sigma\tau(x) &= a_0 + a_1(1+i)\alpha + a_6 i\alpha^2 + a_3(1-i)\alpha^3 \\ &= a_0 + a_1\big((1+i)\alpha\big) + \frac{a_6}{2}\big((1+i)\alpha\big)^2 - \frac{a_3}{2}\big((1+i)\alpha\big)^3, \end{aligned}$$

which means that $C^* = \mathbb{Q}((1+i)\alpha)$.

Similarly we get

$$A^* = \mathbb{Q}(i, \sqrt{2}),\ B^* = \mathbb{Q}(\alpha),\ D^* = \mathbb{Q}(i\alpha),\ E^* = \mathbb{Q}((1-i)\alpha).$$

The normal subgroups of G are G, S, T, U, A, $\{1\}$. The corresponding fields are normal by the Fundamental Theorem. It is easy to see it directly, as the corresponding fields are splitting fields for the polynomials x, x^2+1, x^2-2, x^2+2, x^4-x^2-2, x^4-2, respectively. On the other hand, for example B^* is not normal, as α is a root of x^4-2 in B^*, but does not split in B^*.

As another illustration, note that $G/A \cong C_2 \times C_2$. By the Fundamental Theorem we must have $\mathrm{Gal}(A^*/\mathbb{Q}) = C_2 \times C_2$. This is indeed the case, as $A^* = \mathbb{Q}(i, \sqrt{2})$, and the four automorphisms send $(i, \sqrt{2})$ to $(\pm i, \pm\sqrt{2})$.

Remark 2.8.3 The *inverse Galois problem* asks for existence and an explicit construction of the extension $K/\mathbb{Q}$ whose Galois group is a given group G. For example, how to construct the extension $K/\mathbb{Q}$ with $\mathrm{Gal}(K/\mathbb{Q}) \cong GL_n(\mathbb{F}_q)$ (or does it even exist)? The positive solution of the inverse Galois problem for solvable groups was obtained by Shafarevich (1954). If you want to learn about this difficult area of mathematics, we recommend [MM] or [Vo]

2.9 Discriminant

Suppose that f is a separable irreducible monic cubic polynomial in $k[x]$. Then $\mathrm{Gal}(f)$ is transitive on the three roots of f, so it must be either A_3 or S_3. How can we determine which it is? This is a special case of a more general problem.

Let f be an arbitrary polynomial over a field k, and $\alpha_1, \ldots, \alpha_n$ be roots of f in a splitting field K for f (repeated according to multiplicity). Define the *disriminant* of f to be

$$\Delta = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j)^2.$$

A priori, Δ is the element of K. However, it turns out that $\Delta \in k$. Indeed, if f has multiple roots then $\Delta = 0$. Otherwise f is separable, and the result follows from Galois theory since Δ is clearly $\mathrm{Gal}(K/k)$-invariant.

Theorem 2.9.1 *Let k be a field of characteristic $\neq 2$, $f \in k[x]$ be a polynomial without multiple roots, K be a splitting field for f over k, and Δ be the discriminant of f. Consider the Galois group $G = \mathrm{Gal}(K/k)$ as a permutation group on n roots of f. Then $G \le A_n$ if and only if Δ has a square root in k. Moreover, if Δ has no square root in k, it has a square root δ in K, and $k(\delta)$ is the fixed field of $G \cap A_n$.*

Proof Let $\alpha_1, \ldots, \alpha_n \in K$ be the roots of f. Set

$$\delta = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j) \in K. \tag{2.13}$$

It is clear that δ is a square root of Δ in K, and Δ has a square root in k if and only if $\delta \in k$. By Galois theory, $\delta \in k$ if and only if δ is G-invariant, which in turn is equivalent to $G \le A_n$. Finally, if $\delta \notin k$ then it is of course fixed by $G \cap A_n$, which has index 2 in G. As $[k(\delta) : k] = 2$, we must have $k(\delta) = (G \cap A_n)^*$ by Galois theory. □

The following trick is useful for calculating discriminants.

Lemma 2.9.2 *Let $\alpha_1, \ldots, \alpha_n$ be roots of a polynomial f. Set $p_j =$*

$\sum_{i=1}^n \alpha_i^j$ *for* $j \geq 1$. *Then*

$$\Delta = \begin{vmatrix} n & p_1 & \dots & p_{n-1} \\ p_1 & p_2 & \dots & p_n \\ p_2 & p_3 & \dots & p_{n+1} \\ & & \vdots & \\ p_{n-1} & p_n & \dots & p_{2n-2} \end{vmatrix}.$$

Proof Note that δ defined in (2.13) is nothing but the Vandermonde determinant: $\delta = \det A$, where

$$A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha_1 & \alpha_2 & \dots & \alpha_n \\ & & \vdots & \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \dots & \alpha_n^{n-1} \end{pmatrix}.$$

Now it remains to note that $\Delta = \delta^2 = \det AA^t$. □

This result is useful because the quantities p_j can be expressed in terms of the coefficients of f, and so we can calculate Δ. This will be explained in general in §2.15. Right now we will only consider some examples.

Example 2.9.3 (i) Let $f(x) = x^2 + a_1x + a_0$ be a monic polynomial of degree 2. Then in terms of the roots α_1, α_2, we have $a_0 = \alpha_1\alpha_2$ and $a_1 = -\alpha_1 - \alpha_2$. So $p_1 = -a_1$ and $p_2 = a_1^2 - 2a_0$. By Lemma 2.9.2,

$$\Delta = a_1^2 - 4a_0.$$

(ii) Let $f(x) = x^3 + a_2x^2 + a_1x + a_0$ be a monic polynomial of degree 3. A tedious exercise using Lemma 2.9.2 gives

$$\Delta = -4a_0a_2^3 + a_1^2a_2^2 + 18a_0a_1a_2 - 4a_1^3 - 27a_0^2. \tag{2.14}$$

Example 2.9.4 (i) The polynomial $x^3 - 3x + 1 \in \mathbb{Q}[x]$ is irreducible by Example 2.2.8. Its discriminant equals 81, which is a square in $\mathbb{Q}$, so the Galois group is A_3.

(ii) Consider the polynomial $x^3 + 3x^2 - x - 1 \in \mathbb{Q}[x]$. Consider $g(x) = f(x-1) = x^3 - 4x + 2$. It is irreducible by Eisenstein and the discriminant is 148, which is not a square in $\mathbb{Q}$. So the Galois group is S_3.

Solving Qubic Equations

Assume that $g(y) = y^3 + a_2y^2 + a_1y + a_0$ is an *irreducible* monic polynomial of degree 3. Assume for simplicity that the characteristic of the ground field k is not 2 or 3, so that f is separable. We can simplify the expression for f by setting $x = y + a_2/3$. Then

$$g(y) = f(x) = x^3 + px + q.$$

Let K/k be a splitting field for f over k. By (2.14), f has discriminant

$$\Delta = -4p^3 - 27q^2.$$

Let $\delta \in K$ be a square root of Δ.

Let us try to solve f by radicals. For this it is convenient to work in a splitting field $L \supseteq K$ for the plolynomial $x^3 - 1$ over K. The field L contains a cube root $\omega \neq 1$ of 1, and actually $L = K(\omega)$. In L we set

$$\beta := \alpha_1 + \omega\alpha_2 + \omega^2\alpha_3, \quad \gamma := \alpha_1 + \omega^2\alpha_2 + \omega\alpha_3.$$

Then

$$\begin{aligned}\beta\gamma &= \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + (\omega + \omega^2)(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)\\ &= (\alpha_1 + \alpha_2 + \alpha_3)^2 - 3(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)\\ &= -3p.\end{aligned}$$

Then $\beta^3\gamma^3 = -27p^3$. Moreover,

$$\begin{aligned}\beta^3 + \gamma^3 &= (\alpha_1 + \omega\alpha_2 + \omega^2\alpha_3)^3 + (\alpha_1 + \omega^2\alpha_2 + \omega\alpha_3)^3\\ &\quad +(\alpha_1 + \alpha_2 + \alpha_3)^3\\ &= 3(\alpha_1^3 + \alpha_2^3 + \alpha_3^3) + 18\alpha_1\alpha_2\alpha_3\\ &= -27q,\end{aligned}$$

using that $\alpha_1 + \alpha_2 + \alpha_3 = 0$, which in turn implies that $\alpha_1^3 + \alpha_2^3 + \alpha_3^3 = 3\alpha_1\alpha_2\alpha_3$. Thus,

$$(x - \beta^3)(x - \gamma^3) = x^2 + 27qx - 27p^3.$$

The discriminant of this quadratic equation is

$$(27q)^2 + 4 \cdot 27p^3 = -27\Delta.$$

So

$$\{\beta^3, \gamma^3\} = -\frac{27}{2}q \pm \frac{3}{2}\sqrt{-3\Delta} = -\frac{27}{2}q \pm \frac{3}{2}(2\omega + 1)\delta.$$

Now, we can obtain β by adjoining a cube root of say $\frac{27}{2}q + \frac{3}{2}(2\omega+1)\delta$ to L, and then $\gamma = -3p/\beta$. Finally,

$$\alpha_1 = (\beta+\gamma)/3, \quad \alpha_2 = (\omega^2\beta + \omega\gamma)/3, \quad \alpha_3 = (\omega\beta + \omega^2\gamma)/3.$$

Solving Quartic Equations

Suppose that $g(y) = y^4 + a_3y^3 + a_2y^2 + a_1y + a_0$ be an irreducible monic quartic in $k[x]$. We continue to suppose that $\operatorname{char} k \neq 2, 3$. The substitution $x = y + a_3/4$ reduces g to the form

$$f(x) = x^4 + px^2 + qx + r.$$

Let K/k be a splitting field for f over k, and let $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be the roots of f in K. Let $G = \operatorname{Gal}(K/k)$ considered as a subgroup of S_4 permuting the roots. If $V_4 \triangleleft S_4$ is the Klein 4-group then $H := V_4 \cap G \triangleleft G$. Let $M = H^*$. By the Fundamental Theorem, $\operatorname{Gal}(K/M) = H$ and $\operatorname{Gal}(M/k) \cong G/H$. Now, in view of Example 1.5.20, $H \cong C_2$ or $H \cong V_4 \cong C_2 \times C_2$. We first determine the intermediate field M. Let

$$\beta := \alpha_1 + \alpha_2, \quad \gamma := \alpha_1 + \alpha_3, \quad \delta := \alpha_1 + \alpha_4.$$

Then

$$\begin{aligned}
\beta^2 &= -(\alpha_1+\alpha_2)(\alpha_3+\alpha_4),\\
\gamma^2 &= -(\alpha_1+\alpha_3)(\alpha_2+\alpha_4),\\
\delta^2 &= -(\alpha_1+\alpha_4)(\alpha_2+\alpha_3).
\end{aligned}$$

Consequently $\beta^2, \gamma^2, \delta^2 \in M$. So $k(\beta^2, \gamma^2, \delta^2) \subseteq M$. On the other hand, it is easy to check that if $\sigma \in S_4$ fixes $\beta^2, \gamma^2, \delta^2$ then $\sigma \in V_4$. Hence

$$\operatorname{Gal}(K/k(\beta^2, \gamma^2, \delta^2)) \leq H = \operatorname{Gal}(K/M),$$

and so $k(\beta^2, \gamma^2, \delta^2) \supseteq M$. Thus $k(\beta^2, \gamma^2, \delta^2) = M$. Tedious calculations show that

$$\begin{aligned}
&\beta^2 + \gamma^2 + \delta^2 = -2p,\\
&\beta^2\gamma^2 + \beta^2\delta^2 + \gamma^2\delta^2 = p^2 - 4r\\
&\beta\gamma\delta = -q.
\end{aligned}$$

Thus $k(\beta^2, \gamma^2, \delta^2)$ is a splitting field for

$$x^3 + 2px^2 + (p^2 - 4r)x - q^2.$$

This cubic is called the *cubic resolvent* for f. By the previous results on the cubics we can construct $\beta^2, \gamma^2, \delta^2$ by adjoining square roots and cube

roots. We can then construct β, γ, δ by adjoining square roots (choosing those so that $\beta\gamma\delta = -q$). Then

$$\begin{aligned}\alpha_1 &= (\beta + \gamma + \delta)/2,\\ \alpha_2 &= (\beta - \gamma - \delta)/2,\\ \alpha_3 &= (-\beta + \gamma - \delta)/2,\\ \alpha_4 &= (-\beta - \gamma + \delta)/2.\end{aligned}$$

Notice now that $K = k(\beta, \gamma, \delta)$.

Which are possible Galois groups of an irreducible quartic

$$f = x^4 + px^2 + qx + r \in k[x]?$$

Let K be a splitting field for f over k. The answer is given by the following table, where

$$g = x^3 + 2px^2 + (p^2 - 4r)x - q^2$$

is the cubic resolvent for f, and M is a splitting field for g in K.

Discriminant	g	f	$\mathrm{Gal}(f;k)$
No square root in k	Irreducible over k		S_4
Has square root in k	Irreducible over k		A_4
Has square root in k	Factorizes in $k[x]$		V_4
No square root in k	Factorizes in $k[x]$	Factorizes in $M[x]$	C_4
No square root in k	Factorizes in $k[x]$	Irreducible over M	D_8

Indeed, if g is irreducible over k, G/H is A_3 or S_3 depending on whether G is contained in A_4 or not. This explains the first two lines. Now, let g factorize in $k[x]$. If the disciminant has a square root in k the Galois group must be V_4, as this is the only remaining group contained in A_4. Otherwise we have to distinguish between C_4 and D_8. Note that if f is irreducible over M then $[K : M] = 4$. So in this case we have $|H| = 4$, and so G must be D_8. Conversely, if $G = D_8$, then $\mathrm{Gal}(K/M) = H = V_4$ is transitive, so the polynomial f must be irreducible over M, see Problem (2.17.13).

Note that in the five cases above the Galois group of g over k is S_3, C_3, $\{1\}$, C_2, and C_2, respectively. So everything is uniquely determined in terms of the cubic resolvent except for one ambiguity.

Example 2.9.5 The polynomials in $\mathbb{Q}[x]$

(i) $x^4 + 4x + 2$;
(ii) $x^4 + 8x + 12$;
(iii) $x^4 + 1$;
(iv) $x^4 + x^3 + x^2 + x + 1$;
(v) $x^4 - 2$

provide examples of the cases described in the tables above.

2.10 Finite fields

In this section we give a complete classification of finite fields. It turns out that a finite field is determined up to isomorphism by its order, that the order is a prime power and that all prime powers occur.

Lemma 2.10.1 *Let k be a field of characteristic $p > 0$. Then the map*

$$\mathrm{Fr} : k \to k,\ a \mapsto a^p$$

is a field monomorphism. If k is finite, Fr *is an automorphism.*

Proof Follows from the Freshman's Dream. □

Definition 2.10.2 The map Fr defined in the previous lemma is called the *Frobenius morphism.*

Theorem 2.10.3 *Let p be a prime and $q = p^n$ for some $n \in \mathbb{Z}_{>0}$. A field F has q elements if and only if it is a splitting field for $f(x) = x^q - x$ over the prime subfield $F_0 \cong \mathbb{F}_p$.*

Proof Suppose that $|F| = q$. Then $F^\times \cong C_{q-1}$, see Lemma 1.2.5. It follows that all elements of F are roots of f. Therefore f splits in F, and its roots generate F over F_0.

Conversely, let K be a splitting field for f over $\mathbb{F}_p$. As $f' = -1$ is prime to f, f has exactly q roots. Using Lemma 2.10.1 one easily checks that this set of roots is a subfield of K, hence the set equals K. Therefore $|K| = q$. □

As splitting fields are unique up to isomorphism, we deduce using Lemma 2.3.4:

Theorem 2.10.4 *A finite field must have $q = p^n$ elements where p is a prime number and n is a positive integer. For each such q there exists precisely one field with q elements up to isomorphism. This field can be constructed as a splitting field for $x^p - x$ over $\mathbb{F}_p$.*

Let $\mathbb{F}_p$ be the prime subfield of $\mathbb{F}_q$. It is clear that $\mathbb{F}_q/\mathbb{F}_p$ is normal. It follows from Corollary 2.10.5 below that it is also separable.

Corollary 2.10.5 *All polynomials in $\mathbb{F}_q[x]$ are separable.*

Proof As Frobenius morphism on $\mathbb{F}_q$ is an automorphism, every element of $\mathbb{F}_q$ is a pth power. Let $f \in \mathbb{F}_q[x]$ be irreducible. By Proposition 2.6.18, f is not separable only if $f(x) = g(x^p) = a_0 + a_1x^p + \cdots + a_{n-1}x^{(n-1)p} + x^{np}$. Now every a_i is a pth power, i.e. $a_i = b_i^p$ for some $b_i \in k$. So $f(x) = (b_0 + b_1x + \cdots + x^n)^p$, a contradiction. □

Corollary 2.10.6 *Any extension of finite fields is Galois.*

Proof Separability follows from Corollary 2.10.5 and normality from Theorem 2.10.4. □

Example 2.10.7 We have seen that we can construct fields of all prime power orders p^n by constructing splitting fields for $x^{p^n} - x$ over $\mathbb{F}_p$. The polynomial $x^{p^n} - x$ is of course not irreducible. In certain circumstances we obtain more information by considering splitting fields of irreducible polynomials. Let us illustrate this by considering fields of order p^p.

Consider the polynomial $g(x) := x^p - x - 1 \in \mathbb{F}_p[x]$. Note that $g(a) = -1$ for every $a \in \mathbb{F}_p$. In particular g has no roots in $\mathbb{F}_p$. Let K be a splitting field for g over $\mathbb{F}_p$, and α be a root of g in K. It is easily checked using 'Freshman's Dream' that $\alpha + a$ is also a root of g for any $a \in \mathbb{F}_p$. So the roots of g are precisely $\{\alpha + a \mid a \in \mathbb{F}_p\}$.

Next we show that g is irreducible over $\mathbb{F}_p$. Suppose that $g = g_1g_2$, where g_1, g_2 are monic with $1 \leq d := \deg g_1 < p$. Let

$$A = \{a \in \mathbb{F}_p \mid \alpha + a \text{ is a root of } g_1\}.$$

Then the coefficient of x^{d-1} in g_1 is $-\sum_{a\in A}(\alpha + a) = -d\alpha + b$ for some $b \in \mathbb{F}_p$. It follows that $\alpha \in \mathbb{F}_p$, giving a contradiction.

This means that $[K : \mathbb{F}_p] = p$, and so $|K| = p^p$. So we can think of the field $\mathbb{F}_{p^p}$ as the field obtained from $\mathbb{F}_p$ by adjoining a root of the irreducible polynomial $x^p - x - 1$, which is of course much more convenient than splitting field for $x^{p^p} - x$. Note by the way that as a by-product we obtain that $x^p - x - 1$ divides $x^{p^p} - x$. Finally note that a similar argument applies to any polynomial $x^p - x - a$ with $a \in \mathbb{F}_p^\times$ in place of $x^p - x - 1$. This easily implies that $x^{p^p} - x = \prod_{a \in \mathbb{F}_p}(x^p - x - a)$.

Theorem 2.10.8 *Let $q = p^n$ for a prime p. Then*

$$\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) = \{\mathrm{id} = \mathrm{Fr}^0, \mathrm{Fr}, \mathrm{Fr}^2, \ldots, \mathrm{Fr}^{n-1}\} \cong C_n.$$

Proof In view of Corollary 2.10.6, we can apply the Fundamental Theorem of Galois Theory. It tells us that the Galois group has order n. By Lemma 2.10.1, powers of the Frobenius morphism are elements of the Galois group. It just remains to show that the first n powers are distinct automorphisms of $\mathbb{F}_q$. To see that, apply these powers to a generator of the cyclic group $\mathbb{F}^\times \cong C_{q-1}$ (cf. Lemma 1.2.5). □

Now the Fundamental Theory of Galois Theory and the subgroup structure of cyclic groups described in Lemma 1.2.1 imply

Corollary 2.10.9 *Let p be a prime. Then the field $\mathbb{F}_{p^n}$ contains exactly one subfield isomorphic to $\mathbb{F}_{p^d}$ for each $d|n$, and there are no other subfields in $\mathbb{F}_{p^n}$. Moreover,*

$$\mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_{p^d}) = \{\mathrm{id} = \mathrm{Fr}^0, \mathrm{Fr}^d, \mathrm{Fr}^{2d}, \ldots, \mathrm{Fr}^{(n/d-1)d}\} \cong C_{n/d}.$$

Corollary 2.10.10 *Let q be a prime power and $m, n \geq 1$ be integer. Then $m \mid n$ if and only if $q^m - 1$ divides $q^n - 1$.*

Proof If m divides n then we clearly have that the polynomial $x^m - 1$ divides the polynomial $x^n - 1$. Conversely, if $q^m - 1$ divides $q^n - 1$, then the cyclic group $C_{p^n-1} \cong \mathbb{F}_{q^n}^\times$ contains a subgroup H of order $q^m - 1$. Now, each of the q^m elements of $H \cup \{0\}$ is a solution of $x^{q^m} - x = 0$. So this polynomial splits over $\mathbb{F}_{q^n}$. The corresponding splitting field is isomorphic to $\mathbb{F}_{q^m}$ and is contained in $\mathbb{F}_{q^n}$, whence $m \mid n$. □

2.11 Cyclotomic Polynomials

In this section we consider splitting fields and Galois groups of polynomials of the form $x^m - 1$ over a field k. We comment right away on a

technical point. Assume that $\operatorname{char} k = p | m$. Write $m = p^r q$ where $p \not| q$. Then in $k[x]$ we have $x^m - 1 = (x^q - 1)^{p^r}$. Thus a splitting field for $x^q - 1$ is a splitting field for $x^m - 1$. For this reason *in this section we assume that* $\operatorname{char} k$ *does not divide* m. In this case $(x^m - 1)' = m x^{m-1} \neq 0$, and so $x^m - 1$ has m distinct roots in a splitting field.

Suppose that K/k is a splitting field for $x^m - 1$ over k. As $x^m - 1$ has m distinct roots, the extension K/k is Galois. The set of roots $R \subset K$ clearly forms a groop under multiplication, and so R is a cyclic group of order m, see Lemma 1.2.5. An element $\varepsilon \in R$ is called a *primitive* m*th root of* 1 if it generates R. For example in $\mathbb{C}$ only i and $-i$ are primitive 4th roots of 1. Note that if ε is a primitive mth root of 1 then $K = k(\varepsilon)$.

We now define the m*th cyclotomic polynomial* Φ_m to be

$$\Phi_m(x) := \prod_{\varepsilon} (x - \varepsilon), \tag{2.15}$$

where the product is over all *primitive* mth roots of 1. Observe that

$$x^m - 1 = \prod_{d|m} \Phi_d(x). \tag{2.16}$$

Also note that $\deg \Phi_m = \varphi(m)$, where $\varphi(m)$ is the *Euler function* defined to be the number of natural numbers which are less than m and are prime to m.

A priori, Φ_m are polynomials in $K(x)$, but in fact much more is true.

Theorem 2.11.1 *We have* $\Phi_m \in k_0[x]$*, where* k_0 *is the prime subfield of* k*. Moreover, if* $\operatorname{char} k = 0$*, then* $\Phi_m \in \mathbb{Z}[x]$.

Proof As $x^m - 1 = \prod_{d|m} \Phi_d(x)$, the theorem follows by induction from the following lemma. □

Lemma 2.11.2 (i) *If* K/k *is a field extension,* $q \in K[x]$*, and there exist non-zero* $f, g \in k[x]$ *such that* $f = qg$*, then* $q \in k[x]$.

(ii) *Suppose that* K *is the field of fractions of an integral domain* R*,* $q \in K[x]$*, and there exist monic* $f, g \in R[x]$ *such that* $f = qg$*. Then* $q \in R[x]$.

Proof (i) Let $q = a_0 + a_1 x + \cdots + a_m x^m$, $g = b_0 + b_1 x + \cdots + b_n x^n$, $f = c_0 + c_1 x + \cdots + c_n x^{m+n}$, where $a_m, b_n, c_{m+n} \neq 0$. As $a_m b_n = c_{m+n}$, we have $a_m \in k$. Now, $a_m b_{n-1} + a_{m-1} b_n = c_{m+n-1}$, whence $a_{m-1} = (c_{m+n-1} - a_m b_{n-1})/b_n \in k$, and so on.

(ii) In this case $b_n = 1$, and the same induction goes through. □

Theorem 2.11.3 *For each m the polynomial Φ_m is irreducible over $\mathbb{Q}$.*

Proof Suppose Φ_m is not irreducible. By Lemma 2.2.1, we can write $\Phi_m = fg$, where $f, g \in \mathbb{Z}[x]$ and f is irreducible monic with $1 \leq \deg f < \deg \Phi_m$. Let $K/\mathbb{Q}$ be a splitting field for Φ_m over $\mathbb{Q}$.

We first show that if ε is a root of f in K and p is a prime which does not divide m, then ε^p is also a root of f. Suppose not. Note that ε is a primitive mth root of 1, and hence so is ε^p. Thus, ε^p is a root of Φ_m, and so $g(\varepsilon^p) = 0$. Let $h(x) = g(x^p) \in \mathbb{Z}[x]$. As $h(\varepsilon) = 0$ and $f = \operatorname{irr}(\varepsilon; \mathbb{Q})$, it follows that $f|h$ in $\mathbb{Q}[x]$. By Lemma 2.11.2(ii), we can write $h = fe$ for $e \in \mathbb{Z}[x]$. Now we consider the quotient map $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z},\ n \mapsto \bar{n}$, and the induced map $\mathbb{Z}[x] \to \mathbb{Z}/p\mathbb{Z}[x],\ d \mapsto \bar{d}$. We have using Freshman's Dream

$$\bar{f}(x)\bar{e}(x) = \bar{h}(x) = \bar{g}(x^p) = (\bar{g}(x))^p.$$

Let $\bar{q}$ be any irreducible factor of $\bar{f}$ in $\mathbb{F}_p[x]$. Then $\bar{q}$ divides $\bar{g}^p$, and so $\bar{q}$ divides $\bar{g}$. Therefore $\bar{q}^2$ divides $\bar{f}\bar{g} = \bar{\Phi}_m$. So $\bar{\Phi}_m$ has multiple roots in a splitting field extension of $\mathbb{F}_p$. But this is not so since p does not divide m.

Now let η be a root of f, and let θ be a root of g. As η and θ are both primitive roots of 1 there exists r such that $\theta = \eta^r$, where r and m are relatively prime. Write r as a product of primes and apply the previous paragraph several times to see that θ is a root of f. As Φ_m has no multiple roots, $f = \Phi_m$. Contradiction. □

Remark 2.11.4 The statement of the theorem is wrong in positive characteristic. For example $\Phi_3(x) = x^2 + x + 1$ is reducible over $\mathbb{F}_7$ (but not over $\mathbb{F}_5$!).

Theorem 2.11.5 *Let Φ_m be a cyclotomic polynomial over k. Then $\operatorname{Gal}(\Phi_m; k)$ is a subgroup of $(\mathbb{Z}/m\mathbb{Z})^\times$. Moreover, Φ_m is irreducible over k if and only if $\operatorname{Gal}(\Phi_m; k) \cong (\mathbb{Z}/m\mathbb{Z})^\times$.*

Proof Let $K = k(\varepsilon)$ be a splitting field for Φ_m over k where ε is a primitive mth root of 1. We can write the primitive mth roots of 1 as $\varepsilon^{n_1}, \varepsilon^{n_2}, \ldots, \varepsilon^{n_\varphi}$, where $n_1, n_2, \ldots, n_\varphi$ are the positive integers less than m which are prime to m. Now, for any i, there exist integers a and b such that $an_i + bm = 1$. So in the ring $\mathbb{Z}/m\mathbb{Z}$ we have that $\bar{n}_i$ is a unit. Conversely, if $\bar{n}$ is a unit in $\mathbb{Z}/m\mathbb{Z}$ then n and m are relatively prime. Thus $\{\bar{n}_1, \ldots, \bar{n}_\varphi\} = (\mathbb{Z}/m\mathbb{Z})^\times$.

Now, let $\sigma \in \mathrm{Gal}(K/k)$. As $K = k(\varepsilon)$, σ is determined by its action on ε. As $\sigma(\varepsilon)$ is also a primitive mth root of 1, we have $\sigma(\varepsilon) = \sigma^{n_{i(\sigma)}}$ for some $1 \leq i(\sigma) \leq \varphi$. If τ is another element of the Galois group, it is easy to check that $\bar{n}_{i(\tau\sigma)} = \bar{n}_{i(\tau)}\bar{n}_{i(\sigma)}$, so the mapping $\sigma \mapsto \bar{n}_{i(\sigma)}$ is a group embedding of $\mathrm{Gal}(K/k)$ into $(\mathbb{Z}/m\mathbb{Z})^\times$.

Further, $|\mathrm{Gal}(K/k)| = \varphi$ if and only if the group acts transitively on the roots of Φ_m, and this happens if and only if Φ_m is irreducible, see Problem 2.17.13. □

We complete this section with proving the theorem of Wedderburn which says that the only finite division rings are commutative.

Lemma 2.11.6 *Let $m > 1$. Then $|\Phi_m(x)| > x - 1$ for all real $x \geq 2$.*

Proof Since $x > 1$, the point on the unit circle closest to x is 1. It follows that $|x-\varepsilon| > |x-1|$ for every non-trivial complex roots of unity ε. Since $\Phi_m(x)$ is a product of $\varphi(m)$ factors of the form $x - \varepsilon$, it follows that $|\Phi_m(x)| > |x-1|^{\varphi(m)} \geq x - 1$, where the last inequality holds because $x - 1 \geq 1$. □

Definition 2.11.7 A ring R is called a *division ring* if every non-zero element in R is invertible.

Lemma 2.11.8 *Let D be a division ring and $a \in D$. Then the centralizer $C(a) := \{x \in D \mid xa = ax\}$ is a subring of D and the center $Z(D) = \{x \in D \mid xy = yx \text{ for all } y \in D\}$ is a subfield.*

Proof Easy exercise. □

Theorem 2.11.9 (Wedderburn) *Let D be a finite division ring. Then D is commutative (and thus is a finite field).*

Proof Let $Z = Z(D)$, so that $Z \cong \mathbb{F}_q$ for some prime power q. For each $a \in D$, we have $Z \subseteq C(a)$, which makes $C(a)$ into a Z-vector space. Writing $d(a)$ for $\dim_Z C(a)$ we deduce that $|C(a)| = q^{d(a)}$. In particular, $|D| = q^n$, where $n = d(1)$.

Now, $D^\times$ is a finite group. Let S be the set of representatives of non-central conjugacy classes in $D^\times$. Of course we want to show that $S = \varnothing$. Let us assume that $S \neq \varnothing$. If $a \in S$, then the conjugacy class K_a of a contains exactly $[D^\times : C_{D^\times}(a)] = (q^n - 1)/(q^{d(a)} - 1)$ elements,

and we have

$$q^n - 1 = |D^\times| = |Z^\times| + \sum_{a \in S} |K_a| = (q-1) + \sum_{a \in S} \frac{q^n - 1}{q^{d(a)} - 1}.$$

As $(q^n-1)/(q^{d(a)}-1)$ is an integer, $d(a)$ divides n by Corollary 2.10.10. Because $d(a) < n$, we deduce from the irreducibility of cyclotomic polynomials that Φ_n divides $(x^n - 1)/(x^{d(a)} - 1)$, and so each $\frac{q^n-1}{q^{d(a)}-1}$ is a multiple of $\Phi_n(q)$. Also $q^n - 1$ is a multiple of $\Phi_n(q)$. So it follows that $\Phi_n(q)$ divides $q - 1$. Hence $|\Phi_n(q)| \leq q - 1$. By Lemma 2.11.6, $n = 1$, and thus $D = Z$. □

2.12 The theorem of the primitive element

In this section we give an application of Galois theory to the following problem: if K/k is an algebraic extension, under what circumstances is K a simple extension of k.

Theorem 2.12.1 (Steinitz) *An algebraic extension K/k is simple if and only if there are only finitely many intermediate fields.*

Proof First suppose there are only finitely many intermediate fields. In this case K must be finitely generated over k, say $K = k(\alpha_1, \alpha_2, \ldots, \alpha_r)$. Thus K/k is finite. Moreover, we may assume that $|k| = \infty$, as extensions of finite fields are clearly simple. Let r be the minimal amount of generators of K over k. We want to prove that $r = 1$. If $r \geq 2$, let $L = k(\alpha_1, \alpha_2)$. For each $a \in k$, let $F_a = k(\alpha_1 + a\alpha_2)$. As k is infinite and there are only finitely many intermediate fields there exist $a \neq b$ in k with $F_a = F_b$. But then $(\alpha_1 + b\alpha_2) - (\alpha_1 + a\alpha_2) = (b - a)\alpha_2 \in F_b$, and so $\alpha_2 \in F_b$. Also $\alpha_1 = (\alpha_1 + b\alpha_2) - b\alpha_2 \in F_b$, so that $k(\alpha_1, \alpha_2) \subseteq k(\alpha_1 + b\alpha_2)$. Consequently $K = k(\alpha_1 + b\alpha_2, \alpha_3, \ldots, \alpha_r)$, contradicting the minimality of r.

Conversely, suppose that $K = k(\alpha)$. Let $f = \operatorname{irr}(\alpha; k)$. Note that f has only finitely many monic divisors $g_1, g_2, \ldots, g_m$, say, in $K[x]$. Let L be an intermediate field and $g_L = \operatorname{irr}(\alpha; L)$. Then $g_L | f$ in $L[x]$, and hence in $K[x]$. But this means that $g_L = g_i$ for some i. The proof will therefore be complete if we can show that g_L determines L. Let $g_L = a_0 + a_1 x + \cdots + x^r$, and let $L_0 = k(a_0, a_1, \ldots, a_{r-1})$. Then $L_0 \subseteq L$, and so g_L is irreducible over L_0. Thus $g_L = \operatorname{irr}(\alpha; L_0)$. As $K = L_0(\alpha)$, we have $[K : L_0] = r$. Similarly $[K : L] = r$, so $L = L_0$. In particular g_L determines L. □

The theorem is of course false for transcendental extensions as the example of $k(x)/k$ shows.

Theorem 2.12.2 (Theorem of the primitive element) *Suppose that K/k is finite and separable. Then K/k is simple.*

Proof Let $K = k(\alpha_1, \ldots, \alpha_n)$. Let $f = \prod_{i=1}^{n} \operatorname{irr}(\alpha_i; k)$. Then f is separable over k. Let N be a splitting field of f over K. Then N is also a splitting field for f over k. Thus N/k is Galois. By the Fundamental Theorem, there are only finitely many intermediate subfields, and it remains to apply Theorem 2.12.1. □

Corollary 2.12.3 *If K/k is a finite normal separable field extension, then there exists an irreducible polynomial $f \in k[x]$ such that K is a splitting field for f over k.*

Example 2.12.4 We demonstrate that the assumption of separability in Theorem 2.12.2 is necessary. Let u, v be indeterminates, $k = \mathbb{F}_p(u^p, v^p)$, and $K = \mathbb{F}_p(u, v)$. It is easy to check that $[K : k] = p^2$ and $\alpha^p \in k$ for any $\alpha \in K$. It follows that K/k is not simple.

2.13 Solution of equations by radicals

In this section we give a condition which must be satisfied by an equation soluble by radicals, namely the associated Galois group must be a solvable group. We then construct a quintic polynomial whose Galois group is not solvable, which shows that the quintic equation is not solvable by radicals. Solvability of the Galois group is also a sufficient condition for an equation to be solvable by radicals. We give the necessary definitions.

Definition 2.13.1 Let K/k be a field extension. An element $\alpha \in K$ is called *radical* over k if $\alpha^d \in k$ for some positive integer d.

The extension K/k is called *radical* if $K = k(\alpha_1, \ldots, \alpha_m)$ where α_i is radical over $k(\alpha_1, \ldots, \alpha_{i-1})$ for all $1 \leq i \leq m$.

A polynomial $f \in k[x]$ is called *solvable by radicals* if there exists a radical extension K/k in which f splits into linear factors.

We want to characterize radical extensions in terms of Galois groups.

Lemma 2.13.2 *Let K/k be a radical extension and N be a normal closure of K/k. Then N/k is radical.*

Proof Let α_i be as in Definition 2.13.1. Then N is a splitting field of the polynomial $\prod_{i=1}^{m} \operatorname{irr}(\alpha_i; k)$. For every root β_{ij} of $\operatorname{irr}(\alpha_i; k)$ in N there exists a k-isomorphism $\sigma : k(\alpha_i) \to k(\beta_{ij})$ by Theorem 2.3.15. Since α_i is radical over k, so is β_{ij}. □

The next two lemmas show that certain Galois groups are abelian.

Lemma 2.13.3 *Let k be a field of characteristic 0, and K be a splitting field of the polynomial $x^m - 1$ over k. Then $\operatorname{Gal}(K/k)$ is abelian.*

Proof As $(x^m - 1)' = mx^{m-1}$, the polynomial $x^m - 1$ has no multiple roots in K. Clearly the roots form a group under multiplication. By Lemma 1.2.5, this group is cyclic. Let ε be its generator. Then $K = k(\varepsilon)$. So any $\sigma \in \operatorname{Gal}(K/k)$ is determined by its effect on ε, and $\sigma(\varepsilon)$ must be of the form ε^i. The result follows. □

Lemma 2.13.4 *Let k be a field of characteristic 0 in which $x^n - 1$ splits into linear factors. Let $a \in k$ and K be a splitting field of $x^n - a$ over k. Then $\operatorname{Gal}(K/k)$ is abelian.*

Proof Let $\alpha \in K$ be a root of $x^n - a$. By assumption, any root of $x^n - a$ in K looks like $\varepsilon\alpha$, where ε is root of $x^n - 1$ in k. It follows that $K = k(\alpha)$, so any $\sigma \in \operatorname{Gal}(K/k)$ is determined by its effect on α, and $\sigma(\alpha)$ must be of the form $\varepsilon\alpha$. The result follows. □

Lemma 2.13.5 *If K/k is a normal and radical field extension and k has characteristic 0, then $\operatorname{Gal}(K/k)$ is solvable.*

Proof Suppose that $K = k(\alpha_1, \ldots, \alpha_m)$ with $\alpha_i^{d(i)} \in k(\alpha_1, \ldots, \alpha_{i-1})$ for all i. Inserting extra elements α_i if necessary we may assume that all $d(i)$ are prime. Set $p := d(1)$. We prove the result be induction on m, starting from the trivial case $m = 0$.

We may assume that $\alpha_1 \notin k$, as otherwise $K = k(\alpha_2, \ldots, \alpha_m)$ and induction applies. Hence $f := \operatorname{irr}(\alpha_1; k)$ has degree at least 2. As K/k is normal, f splits over K. Since $\operatorname{char} k = 0$, K/k is separable, and so f does not have multiple roots. Let $\beta \neq \alpha_1$ be another root of f. Then $\varepsilon := \alpha_1\beta^{-1}$ satisfies $\varepsilon^p = 1$. As $\varepsilon \neq 1$, the elements $1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{p-1}$ are distinct roots of $x^p - 1$ in K. Therefore $x^p - 1$ splits in K.

Let $M = k(\varepsilon)$. Then we have field extensions $K/M(\alpha_1)/M/k$. Observe that K/k is Galois, hence so is L/M. As $x^p - 1$ splits over M and

$\alpha_1^p \in M$, $M(\alpha_1)$ is a splitting field of $x^p - \alpha_1^p$ over M. Thus $M(\alpha_1)/M$ is normal. By Lemma 2.13.4, $\mathrm{Gal}(M(\alpha_1)/M)$ is abelian. By the Fundamental Theorem of Galois Theory applied to the extension K/M, we have

$$\mathrm{Gal}(M(\alpha_1)/M) = \mathrm{Gal}(K/M)/\mathrm{Gal}(K/M(\alpha_1)).$$

Now $K = M(\alpha_1)(\alpha_2, \ldots, \alpha_m)$ and $K/M(\alpha_1)$ is normal, so by induction $\mathrm{Gal}(K/M(\alpha_1)$ is solvable. Finally, $\mathrm{Gal}(M/k)$ is abelian in view of Lemma 2.13.3, and M/k is normal, so the Fundamental Theorem gives

$$\mathrm{Gal}(M/k) = \mathrm{Gal}(K/k)/\mathrm{Gal}(K/M).$$

By Lemma 1.9.2, $\mathrm{Gal}(K/k)$ is solvable. □

Theorem 2.13.6 *If $F/K/k$ are field extensions,* $\mathrm{char}\, k = 0$*, and F/k is radical, then* $\mathrm{Gal}(K/k)$ *is solvable.*

Proof Let L be the fixed field of $\mathrm{Gal}(K/k)$, and N/L be a normal closure of F/L. We have field extensions $N/F/K/L/k$. By Lemma 2.13.2, N/L is radical, and so $\mathrm{Gal}(N/L)$ is solvable thanks to Lemma 2.13.5. By Theorem 2.7.18, K/L is normal. So the Fundamental Theorem yields

$$\mathrm{Gal}(K/L) \cong \mathrm{Gal}(N/L)/\mathrm{Gal}(N/K).$$

Now $\mathrm{Gal}(K/L)$ is solvable by Theorem 1.9.2, and it remains to observe that $\mathrm{Gal}(K/k) = \mathrm{Gal}(K/L)$. □

Corollary 2.13.7 *Let f be a polynomial over a field k of characteristic 0. If f is solvable by radicals then the Galois group of f over k is solvable.*

We want to exhibit a polynomial whose Galois group is not solvable.

Lemma 2.13.8 *Let p be a prime and $f \in \mathbb{Q}[x]$ be an irreducible polynomial of degree p. Suppose that f has precisely two non-real roots in $\mathbb{C}$. Then* $\mathrm{Gal}(f; \mathbb{Q}) \cong S_p$.

Proof By the Fundamental Theorem of Algebra, $\mathbb{C}$ contains a splitting field K for f. Let $G = \mathrm{Gal}(f; \mathbb{Q})$ considered as a permutation group on the roots of f in K. If α is a root of f, then $[\mathbb{Q}(\alpha) : \mathbb{Q}] = p$ so $[K : \mathbb{Q}]$ is divisible by p. By the Fundamental Theorem of Galois Theory, p divides $|G|$, so G contains a p-cycle.

Complex conjugation induces an element of G which leaves the $p - 2$ real roots of f invariant while transposing the two non-real roots.

Therefore G contains a 2-cycle. But it is easy to see that a 2-cycle and a p-cycle generate S_p, see problem (1.12.76) from §1.12. □

Example 2.13.9 The polynomial $f(x) = x^5 - 6x + 3 \in \mathbb{Q}[x]$ is not solvable by radicals. Indeed, by Eisenstein's Criterion f is irreducible over $\mathbb{Q}$. In view of Lemma 2.13.8, it suffices to show that f has exactly three real roots, each with multiplicity 1. But this is an easy exercise from Calculus (sketch the graph!).

In §2.14 for any n we will demonstrate a polynomial whose Galois group is S_n. The only drawback of that construction will be that the polynomial will have coefficients not in $\mathbb{Q}$ but in a rather 'large' field. Now we want to obtain a converse to Corollary 2.13.7.

Definition 2.13.10 Let K/k be a finite normal extension with Galois group G. The *norm* of an element $\alpha \in K$ is defined to

$$N(\alpha) = \sigma_1(\alpha)\sigma_2(\alpha)\dots\sigma_n(\alpha)$$

Where $G = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$.

It is easy to see that $N(\alpha)$ lies in the fixed subfield of G, so if the extension is also separable then $N(\alpha) \in k$, see Theorem 2.7.16.

Theorem 2.13.11 (Hilbert's Theorem 90) *Let K/k be a finite normal extension with cyclic Galois group $G = \langle\sigma\rangle$. Then $\alpha \in K$ has norm $N(\alpha) = 1$ if and only if $\alpha = \beta/\sigma(\beta)$ for some $\beta \in K^\times$.*

Proof We have $N(\frac{\beta}{\sigma(\beta)}) = \frac{\beta}{\sigma(\beta)}\frac{\sigma(\beta)}{\sigma^2(\beta)}\dots\frac{\sigma^{n-1}(\beta)}{\sigma^n(\beta)} = 1$. Conversely, suppose that $N(\alpha) = 1$. Let $\gamma \in K$ and define $\delta_j = \sigma^j(\gamma)\prod_{j=0}^{i}\sigma^j(\alpha)$ for $0 \le i < n$. Note that $\delta_{i+1} = \alpha\sigma(\delta_i)$ for $0 \le i \le n-2$, and $\delta_{n-1} = \sigma^{n-1}(\gamma)N(\alpha) = \sigma^{n-1}(\gamma)$. Set $\beta = \sum_{i=0}^{n-1}\delta_i$. Note that

$$\sigma(\beta) = \sum_{i=0}^{n-1}\sigma(\delta_i) = \sigma^n(\gamma) + \alpha^{-1}\sum_{i=1}^{n-1}\delta_i = \alpha^{-1}\sum_{i=0}^{n-1}\delta_i = \beta/\alpha.$$

So the result will follow if we can choose γ in such a way that $\beta \neq 0$. Suppose on the contrary that $\beta = 0$ for all choices of γ, i.e.

$$\sum_{i=0}^{n-1}\Big(\prod_{j=1}^{i}\sigma^j(\alpha)\Big)\sigma^i(\gamma) = 0 \qquad (\gamma \in K).$$

But this means that the distinct automorphisms σ^i are linearly dependent over K, contrary to Dedekind's Lemma. □

Theorem 2.13.12 *Suppose that K/k is a finite separable normal extension with Galois group G isomorphic to C_p for a prime p. Assume that* $\operatorname{char} k$ *is 0 or prime to p, and that $x^p - 1$ splits in k. Then $K = k(\alpha)$ where α is a root of an irreducible polynomial of the form $x^p - a$ over k.*

Proof Let $G = \langle\sigma\rangle$. The roots of $x^p - 1$ form a cyclic subgroup of $k^\times$. Let ε be a generator of this group. As $\varepsilon \in k$, we have $\sigma^i(\varepsilon) = \varepsilon$ for any i, and so $N(\varepsilon) = \varepsilon^p = 1$. By Hilbert's Theorem 90, $\varepsilon = \alpha/\sigma(\alpha)$ for some $\alpha \in K$. It follows that $\sigma(\alpha) = \varepsilon^{-1}\alpha$, $\sigma^2(\alpha) = \varepsilon^{-2}\alpha$, etc. So $a := \alpha^p$ is fixed by G, hence lies in k.

The roots of the polynomial $x^p - a$ are of the form $\alpha\varepsilon^i$, so $k(\alpha)$ is a splitting field of $x^p - a$ over k. The k-automorphisms $1, \sigma, \ldots, \sigma^{p-1}$ map α to distinct elements, so they give p distinct elements of $\operatorname{Gal}(k(\alpha)/k)$. By the Fundamental Theorem, $[k(\alpha) : k] \geq p$. But $[K : k] = |G| = p$, so $K = k(\alpha)$. Moreover $x^p - a$ is irreducible, for otherwise the minimal polynomial of α over k has degree less than p and $[k(\alpha) : k] < p$. □

Theorem 2.13.13 *Let K/k be a finite normal extension and* $\operatorname{char} k = 0$. *If* $\operatorname{Gal}(K/k)$ *is solvable then there exists an extension F/K such that F/k is radical.*

Proof Note that all extensions are separable as we are in the characteristic 0 case. Set $G := \operatorname{Gal}(K/k)$. We use induction on $|G|$, the case $|G| = 1$ being clear. Let $|G| > 1$. As G is solvable, it has a normal subgroup H with $G/H \cong C_p$ for a prime p. Let N be a splitting field over K of the polynomial $x^p - 1$. Then N/k is normal by Corollary 2.6.13. The group $\operatorname{Gal}(N/K)$ is abelian by Lemma 2.13.3, and by the Fundamental Theorem, $\operatorname{Gal}(K/k) \cong \operatorname{Gal}(N/k)/\operatorname{Gal}(N/K)$. It follows that $\operatorname{Gal}(N/k)$ is solvable, see Lemma 1.9.2.

Let M be the subfield of N generated by k and the roots of $x^p - 1$. Clearly M/k is radical, so the desired result will follow if we can find an extension F of N such that F/M is radical.

We claim that $\operatorname{Gal}(N/M)$ is isomorphic to a subgroup of G. Let us map any $\sigma \in \operatorname{Gal}(N/M)$ to $\sigma|_K$. As K/k is normal, $\sigma|_K \in G$. So we have a well-defined group homomorphism $\operatorname{Gal}(N/M) \to G$, which is injective as K and M generate N. Let $J \leq G$ be the image of this

embedding. If J is a proper subgroup of G then by induction there is a required extension F/N.

Now assume $J = G$. Then $H \lhd J$ yields a normal subgroup $I \lhd \mathrm{Gal}(N/M)$ of index p. By the Fundamental Theorem, $[I^* : M] = p$, and I^*/M is normal. By Theorem 2.13.12, $I^* = M(\alpha)$ where $\alpha^p = a \in M$. But N/I^* is a normal extension with solvable Galois group of order less than $|G|$, so by induction there exists an extension F/N such that F/I^* is radical. But then F/M is radical and we are done. □

Corollary 2.13.14 *Over a field of characteristic* 0 *a polynomial is solvable by radicals if and only if it has solvable Galois group.*

As S_4 and all its subgroups are solvable, Corollary 2.13.14 implies that polynomials of degree ≤ 4 can be solved by radicals. We can use our insight into the structure of the symmetric group to find out how.

2.14 Transcendental extensions

Definition 2.14.1 Let K/k be a field extension and $A = \{\alpha_1, \ldots, \alpha_n\}$ be a finite subset of K. We say that A is *algebraically dependent* over k if there exists a non-zero polynomial $f \in k[x_1, \ldots, x_n]$ such that $f(\alpha_1, \ldots, \alpha_n) = 0$. Otherwise A is called *algebraically independent*.

An infinite set is called *algebraically dependent* if it has a finite subset which is algebraically dependent and *algebraically independent* otherwise.

Denote by $k(x_1, \ldots, x_n)$ the field of rational functions in variables $x_1, \ldots, x_n$, which is the quotient field of the ring $k[x_1, \ldots, x_n]$. The following result is an analogue of Theorem 2.3.16 and is proved similarly.

Theorem 2.14.2 *Let K/k be a field extension and $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a subset of K algebraically independent over k. Then there exists a unique isomorphism $k(x_1, \ldots, x_n) \tilde{\to} k(\alpha_1, \ldots, \alpha_n)$, which is identity on k and maps x_i to α_i for every i.*

The next result gives a useful practical criterion for a set to be algebraically independent.

Theorem 2.14.3 (The Main Criterion) *Let K/k be a field extension and $\alpha_1, \alpha_2, \ldots, \alpha_n$ be distinct elements of K. Let $k_0 = k$, $k_i =$*

$k(\alpha_1, \ldots, \alpha_i)$ for $1 \leq i \leq n$. Then $A = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is algebraically independent over k if and only if α_i is transcendental over k_{i-1} for every $1 \leq i \leq n$.

Proof Suppose that α_i is algebraic over k_{i-1}, i.e. there is a non-zero polynomial

$$f(x) = a_0 + a_1 x + \cdots + a_r x^r \in k_{i-1}[x]$$

with $f(\alpha_i) = 0$. We can write each a_j as

$$a_j = p_j(\alpha_1, \ldots, \alpha_{i-1})/q_j(\alpha_1, \ldots, \alpha_{i-1})$$

for $p_j, q_j \in k[x_1, \ldots, x_{i-1}]$. We clear denominators. Let

$$l_j = p_j \prod_{m \neq j} q_m \qquad (0 \leq j \leq r).$$

Then each $l_j \in k[x_1, \ldots, x_{i-1}]$ and $g = l_0 + l_1 x_i + \cdots + l_r x_i^r$ is a non-zero element of $k[x_1, \ldots, x_{i-1}, x_i]$ with $g(\alpha_1, \ldots, \alpha_i) = 0$, whence A is algebraically dependent.

Conversely, assume that A is algebraically dependent. There is j such that $\{\alpha_1, \ldots, \alpha_{j-1}\}$ is algebraically independent and $\{\alpha_1, \ldots, \alpha_j\}$ is not. Thus there is a non-zero $f \in k[x_1, \ldots, x_j]$ with $f(\alpha_1, \ldots, \alpha_j) = 0$. We can write $f = a_0 + a_1 x_j + \cdots + a_r x_j^r$ where $a_i \in k[x_1, \ldots, x_{j-1}]$. Let

$$g = a_0(\alpha_1, \ldots, \alpha_{j-1}) + a_1(\alpha_1, \ldots, \alpha_{j-1})x + \cdots + a_r(\alpha_1, \ldots, \alpha_{j-1})x^r.$$

Then g is a non-zero polynomial of $k_{j-1}[x]$, as $\alpha_1, \ldots, \alpha_{j-1}$ are algebraically independent. As $g(\alpha_j) = 0$, α_j is algebraic over k_{j-1}. □

You may have noticed an analogy between algebraic and linear dependence. We now push this analogy further.

Definition 2.14.4 Let K/k be a field extension. A *transcendence basis* of K over k is a subset $A \subseteq K$ which is algebraically independent over k and is maximal (with respect to inclusion) in the set of all algebraically independent subsets of K.

Lemma 2.14.5 *For any field extension K/k a transcendence basis exists. Moreover, if $C \subset K$ is an algebraically independent subset then there exists a transcendence basis B of K over k containing C.*

Proof Follows from Zorn's Lemma. □

Lemma 2.14.6 *Let K/k be a field extension and B be a subset of K. Then B is a transcendence basis of K over k if and only if B is algebraically independent over k and $K/k(B)$ is algebraic.*

Proof Suppose that B is a transcendence basis of K over k. If $\alpha \in K \setminus k(B)$ then by the maximality of B, $B \cup \{\alpha\}$ is algebraically dependent. So there exist distinct $b_1, \ldots, b_n \in B$ such that $\{b_1, \ldots, b_n, \alpha\}$ are algebraically dependent. Now α is algebraic over $k(b_1, \ldots, b_n)$ and hence over $k(B)$ thanks to the Main Criterion.

Conversely, suppose that B is algebraically independent and $K/k(B)$ is algebraic. If $\alpha \in K \setminus B$, then α is algebraic over $k(B)$, so there exists a non-zero $g = a_0 + a_1 x + \cdots + a_j x^j \in k(B)[x]$ annihilating α. Each coefficient of g involves only finitely many elements of B, so there exists a finite subset $\{b_1, \ldots, b_n\} \subseteq B$ such that $a_i \in k(b_1, \ldots, b_n)$. Thus α is algebraic over $k(b_1, \ldots, b_n)$, and so $\{b_1, \ldots, b_n, \alpha\}$ is algebraically dependent by the Main Criterion. Hence $B \cup \{\alpha\}$ is algebraically dependent, and B is maximal. □

Corollary 2.14.7 *Let K/k be a field extension, A be a subset of K with $K/k(A)$ algebraic, and C be a subset of A which is algebraically independent over k. Then there exists a transcendence basis of K over k satisfying $C \subseteq B \subseteq A$.*

Proof Use Zorn lemma to choose a maximal element among the algebraically independent subsets B satisfying $C \subseteq B \subseteq A$. Then by the Main Criterion, every element of A is algebraic over $k(B)$, hence $k(A)/k(B)$ is algebraic, and so $K/k(B)$ is algebraic by transitivity of algebraic extensions. Now Lemma 2.14.6 implies that B is a transcendence basis. □

Corollary 2.14.8 *If $K = k(A)$ then there exists a subset $B \subseteq A$ which is a transcendence basis of K over k.*

Proof Take $C = \varnothing$ in Corollary 2.14.7. □

Theorem 2.14.9 *Let K/k be a field extension, $C = \{c_1, \ldots, c_r\}$ be a subset of K (with r distinct elements) which is algebraically independent over k, and $A = \{a_1, \ldots, a_s\}$ be a subset of K (with s distinct elements) such that $K/k(A)$ is algebraic. Then $r \leq s$ and there exists a set D with $C \subseteq D \subseteq A \cup C$ and such that $|D| = s$ and $K/k(D)$ is algebraic.*

Proof Induction on r. If $r = 0$ take $D = A$. Suppose the result is true for $r - 1$. As the set $C_0 := \{c_1, \ldots, c_{r-1}\}$ is algebraically independent, there exists a set D_0 with $C_0 \subseteq D_0 \subseteq A \cup C_0$ and such that $|D_0| = s$ and $K/k(D_0)$ is algebraic. By relabeling A if necessary we may assume that

$$D_0 = \{c_1, \ldots, c_{r-1}, a_r, a_{r+1}, \ldots, a_s\}.$$

As $K/k(D_0)$ is algebraic, c_r is algebraic over $k(D_0)$. As $\{c_1, \ldots, c_r\}$ is algebraically independent, c_r is transcendental over $k(c_1, \ldots, c_{r-1})$ see the Main Criterion. Thus $s \geq r$. By the same theorem,

$$E := \{c_1, \ldots, c_{r-1}, c_r, a_r, a_{r+1}, \ldots, a_s\}$$

is algebraically dependent. Using the Main Criterion once more we conclude that there exists t with $r \leq t \leq s$ such that a_t is algebraic over $k(c_1, \ldots, c_r, a_r, \ldots, a_{t-1})$. Set

$$D := \{c_1, \ldots, c_r, a_r, \ldots, a_{t-1}, a_{t+1}, \ldots, a_s\}.$$

Then a_t is algebraic over $k(D)$, and so $K(E)/K(D)$ is algebraic. As $E \supseteq D_0$, $K/k(E)$ is algebraic, and so $K/k(D)$ is algebraic. This completes the proof. □

The theorem easily implies

Corollary 2.14.10 *Let K/k be a field extension. If B and C are two transcendence bases of K over k then either B and C are both infinite or B and C have the same number of elements.*

This result shows that the following notions are well-defined.

Definition 2.14.11 If K/k is a field extension we define its *transcendence degree*, denoted $\mathrm{tr.deg}(K/k)$, to be infinity if there exists a transcendence basis of K over k with infinitely many elements. If there exists a transcendence basis B with finitely many elements, we define $\mathrm{tr.deg}(K/k) = |B|$.

Our next goal is to prove the analogue of tower law for transcendental extensions.

Proposition 2.14.12 *Let $F/K/k$ be field extensions. If $A \subset K$ is algebraically independent over k and $B \subseteq F$ is algebraically independent over K then $A \cup B$ is algebraically independent over k.*

Proof Let C be a finite subset of $A \cup B$. We can write

$$C = \{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s\}$$

where $\alpha_i \in A$, $\beta_j \in B$. By the Main Criterion, every α_i is transcendental over $k(\alpha_1, \ldots, \alpha_{i-1})$ and every β_j is transcendental over $K(\beta_1, \ldots, \beta_{j-1})$. So β_j is transcendental over $k(\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_{j-1})$. Thus C is algebraically independent by the Main Criterion. □

Theorem 2.14.13 *Let $F/K/k$ be field extensions. If A is a transcendence basis for K over k and B is a transcendence basis for F over K then $A \cup B$ is a transcendence basis for F over k.*

Proof By Lemma 2.14.6 and Proposition 2.14.12, it suffices to prove that $F/k(A \cup B)$ is algebraic. We know that $K/k(A)$ is algebraic, whence $K(B)/k(A \cup B)$ is algebraic, see Problem 2.17.43 from §2.17. Moreover, $F/K(B)$ is algebraic, so $F/k(A \cup B)$ is algebraic by Theorem 2.3.18. □

Corollary 2.14.14 (Tower Law For Transcendence degree) *If $F/K/k$ are field extensions then the transcendence degree of F/k is the sum of the transcendence degrees of F/K and K/k.*

2.15 Symmetric functions and generic polynomials

Let K be a field, $x_1, \ldots, x_n \in K$ and consider polynomial

$$f(x) = (x - x_1)(x - x_2)\ldots(x - x_n) \in K[x].$$

If we write

$$f(x) = x^n - e_1 x^{n-1} + \cdots + (-1)^n e_n$$

then

$$\begin{aligned} e_1 &= x_1 + \cdots + x_n, \\ e_2 &= \sum_{1 \le i < j \le n} x_i x_j, \\ &\vdots \\ e_n &= x_1 x_2 \ldots x_n. \end{aligned}$$

The expressions e_n considered as elements of $\mathbb{Z}[x_1, \ldots, x_n]$ are the *elementary symmetric functions in n variables*. When necessary, they can also be considered as elements of $k[x_1, \ldots, x_n]$ for any field k.

Now let $K = k(x_1, \ldots, x_n)$ where $x_1, \ldots, x_n$ are algebraically independent. Then $e_1, \ldots, e_n$ can be considered as polynomials in $k[x_1, \ldots, x_n]$ and as elements of K. Note that the symmetric group S_n acts on K by k-automorphisms permuting the variables: if $\sigma \in S_n$ then

$$\begin{aligned} &\sigma\big(f(x_1, \ldots, x_n)/g(x_1, \ldots, x_n)\big) \\ &= f(x_{\sigma(1)}, \ldots, x_{\sigma(n)})/g(x_{\sigma(1)}, \ldots, x_{\sigma(n)}). \end{aligned}$$

Let L be the fixed field

$$L := S_n^* = K^{S_n}.$$

By Theorems 2.7.3 and 2.7.18, and the Fundamental Theorem, K/L is a Galois extension with Galois group S_n.

Theorem 2.15.1 *In the above notation, the elementary symmetric polynomials $e_1, \ldots, e_n \in K$ are algebraically independent over k, and $L = k(e_1, \ldots, e_n)$.*

Proof First of all, it is clear that $k(e_1, \ldots, e_n) \subseteq L$, and $[K : L] = n!$. Therefore to establish the equality $L = k(e_1, \ldots, e_n)$, it suffices to prove that $[K : k(e_1, \ldots, e_n)] \leq n!$. But this follows from the fact that $K/k(e_1, \ldots, e_n)$ is a splitting field extension for f, when f is considered as an element of $k(e_1, \ldots, e_n)[x]$.

By Corollary 2.14.8, $e_1, \ldots, e_n$ contains a transcendence basis for $k(e_1, \ldots, e_n)$ over k. Moreover, $K/k(e_1, \ldots, e_n)$ is algebraic, so this basis will be a basis for K over k (use Lemma 2.14.6 twice). As $\mathrm{tr.deg}(K/k) = n$, this basis must be the whole of $\{e_1, \ldots, e_n\}$. □

Theorem 2.15.2 *Let K/k be a field extension with $K = k(a_1, \ldots, a_n)$, and $a_1, \ldots, a_n$ be algebraically independent over k. Let*

$$g(x) = x^n - a_1 x^{n-1} + \cdots + (-1)^n a_n \in K[x].$$

Then g is irreducible and separable. Moreover, $\mathrm{Gal}(g; K) \cong S_n$.

Proof In view of Theorem 2.15.1, there is a k-isomorphism

$$k(a_1, \ldots, a_n) \tilde{\to} k(e_1, \ldots, e_n),$$

which maps a_i to e_i for every i. This isomorphism transforms our polynomial g to the polynomial f from Theorem 2.15.1. Now everuthing follows from the remarks preceding Theorem 2.15.1. □

2.16 The algebraic closure of a field

Definition 2.16.1 A field K is called *algebraically closed* if every $f \in K[x]$ splits over K. An extension field $K \subseteq k$ is called an *algebraic closure* of k if K/k is algebraic and K is algebraically closed.

Thus the fundamental theorem of algebra states that $\mathbb{C}$ is algebraically closed. Note that the only possible algebraic closure of an algebraically closed field K is K itself.

Example 2.16.2 $\mathbb{C}$ is an algebraic closure of $\mathbb{R}$. However, $\mathbb{C}$ is not an algebraic closure of $\mathbb{Q}$ as $\mathbb{C}/\mathbb{Q}$ is not algebraic, see Example 2.3.20.

The next theorem gives two useful characterizations of an algebraic closure:

Theorem 2.16.3 *Suppose that K/k is a field extension. Then the following are equivalent:*

(i) *K is an algebraic closure of k.*
(ii) *K/k is algebraic and every $f \in k[x]$ splits over K.*
(iii) *K/k is algebraic and if L/K is algebraic then $L = K$.*

Proof (i) $\Rightarrow$ (ii) is clear.

(ii) $\Rightarrow$ (iii) By Theorem 2.3.18, L/k is algebraic. Suppose $\alpha \in L$. By hypothesis, and $\operatorname{irr}(\alpha; k)$ splits over K, whence $\alpha \in K$.

(iii) $\Rightarrow$ (i) Let $f \in K[x]$. Let L be a splitting field of f over K. The extension L/K is algebraic, so, by hypothesis, $L = K$. Thus f splits over K, and so K is algebraically closed. □

Corollary 2.16.4 *Suppose K/k is a field extension and K be algebraically closed. Let $\bar{k}$ be the field of all elements of K which are algebraic over k. Then $\bar{k}$ is an algebraic closure of k.*

In particular, $\mathbb{A}$ is an algebraic closure of $\mathbb{Q}$, see Example 2.3.19.

Proposition 2.16.5 *Let K/k be an algebraic extension. Then every homomorphism φ of k into an algebraically closed field F can be extended to K.*

Proof Let $\mathcal{S}$ be the set of all pairs (L, ψ), where $k \subseteq L \subseteq K$ and $\psi : L \to F$ is a field homomorphism which extends φ. Order $\mathcal{S}$ by

$(L_1, \psi_1) \leq (L_2, \psi_2)$ if and only if $L_1 \subseteq L_2$ and $\psi_1 = \psi_2|_{L_1}$. By Zorn's Lemma, $\mathcal{S}$ has a maximal element (M, μ). Assume $M \neq K$. Let $\alpha \in K \setminus M$ and $f = \operatorname{irr}(\alpha; M)$. As F is algebraically closed, $\mu(f)$ has a root in F, so μ can be extended to $M(\alpha)$ by Theorem 2.3.15, which contradicts the maximality of M. □

Theorem 2.16.6 *Let k be any field. Then there exists an algebraic closure of k. Moreover any two algebraic closures of k are k-isomorphic.*

We will not prove this theorem in these notes. If you are interested, please read the proof in section 6.4 of Rotman (Theorems 6.58 and 6.62).

2.17 Problems on Fields

Problem 2.17.1 True or false? The Galois group of $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$ has order 3.

Problem 2.17.2 If the extension K/k is algebraic and k is countable then K is also countable.

Problem 2.17.3 The field $\mathbb{A}$ is countable.

Problem 2.17.4 True or false? If $F/K/k$ are field extensions with F/K and K/k algebraic then F/k is also algebraic.

Problem 2.17.5 True or false? The extension $\mathbb{Q}(i, \sqrt{5})$ is simple.

Problem 2.17.6 Let p be prime. Then the polynomial $1 + x + \cdots + x^{p-1} \in \mathbb{Q}[x]$ is irreducible.

Problem 2.17.7 True or false? If $F/K/k$ are field extensions with F/K and K/k normal then F/k is also normal.

Problem 2.17.8 True or false? $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt{2})$ are isomorphic fields.

Problem 2.17.9 If β is algebraic over $k(\alpha)$ and β is transcendental over k then α is algebraic over $k(\beta)$.

Problem 2.17.10 True or false? If K/k is algebraic and D is an integral domain such that $k \subseteq D \subseteq K$ then D is a field.

Problem 2.17.11 Let K/k be a field extension. This extension is algebraic if and only if for every intermediate field F every k-monomorphism of F is in fact an automorphism of F.

Problem 2.17.12 A field F is called *perfect* if every irreducible polynomial over F is separable. Show that a field F of characteristic p is perfect if and only if every element of F has a pth root in F. Show that every finite field is perfect.

Problem 2.17.13 Let $f \in k[x]$, K/k be a splitting field for f over k, and $G = \mathrm{Gal}(K/k)$. Show that G acts on the set of the roots of f. Show that G acts transitively if f is irreducible. Conversely, if f has no multiple roots and G acts transitively then f is irreducible.

Problem 2.17.14 Prove that if F is an infinite field, then its multiplicative group $F^\times$ is never cyclic.

Problem 2.17.15 Let K/k be a normal field extension and f be an irreducible polynomial over k. Show that all irreducible factors of f in $K[x]$ all have the same degree.

Problem 2.17.16 True or false: $\mathrm{Gal}(k(x)/k) = \{1\}$.

Problem 2.17.17 Construct subfields of $\mathbb{C}$ which are splitting fields over $\mathbb{Q}$ for the polynomials $x^3 - 1$, $x^4 - 5x^2 + 6$, $x^6 - 8$. Find the degrees of those fields as extensions over $\mathbb{Q}$.

Problem 2.17.18 Which of the following extensions are normal?

(a) $\mathbb{Q}(x)/\mathbb{Q}$;
(b) $\mathbb{Q}(\sqrt{-5})/\mathbb{Q}$;
(c) $\mathbb{Q}(\sqrt[7]{5})/\mathbb{Q}$.
(d) $\mathbb{Q}(\sqrt{5}, \sqrt[7]{5})/\mathbb{Q}(\sqrt[7]{5})$;
(e) $\mathbb{R}(\sqrt{-7})/\mathbb{R}$.

Problem 2.17.19 Let K/k be a splitting field for a polynomial $f \in k[x]$ of degree n. Show that $[K : k]$ divides $n!$.

Problem 2.17.20 True or false: if K/k is a field extension then every k-monomorphism $K \to K$ is an automorphism.

Problem 2.17.21 Determine Galois groups of the following extensions:
(a) $\mathbb{Q}(\sqrt{2},\sqrt{3})/\mathbb{Q}$;
(b) $\mathbb{Q}(\sqrt[5]{3},e^{2\pi i/5})/\mathbb{Q}$;
(c) $\mathbb{Q}(\sqrt[3]{2},\sqrt{2})/\mathbb{Q}$;
(d) $\mathbb{Q}(\sqrt[3]{2},\sqrt{2},e^{2\pi i/3})/\mathbb{Q}$.

Solution. (a) $C_2 \times C_2$, see more difficult Problem 2.17.48.

(b) Denote $\alpha := \sqrt[5]{3}$, $\varepsilon := e^{2\pi i/5}$. Note that $\mathbb{Q}(\sqrt[5]{3},e^{2\pi i/5})$ is a splitting field of the irreducible polynomial x^5-3, whose roots are $\alpha,\alpha\varepsilon,\alpha\varepsilon^2,\alpha\varepsilon^3,\alpha\varepsilon^4$, so the Galois group G is a transitive permutation group on the roots. Moreover,

$$|G| = [\mathbb{Q}(\sqrt[5]{3},e^{2\pi i/5}) : \mathbb{Q}(\sqrt[5]{3})][\mathbb{Q}(\sqrt[5]{3}) : \mathbb{Q}] \leq 4 \cdot 5 = 20.$$

(one can see that the index $[\mathbb{Q}(\sqrt[5]{3},e^{2\pi i/5}) : \mathbb{Q}]$ is exactly 20 as both 5 and 4 divide it, but this will follow anyway). Now, as in Example 2.8.2, one checks that there are elements $\sigma,\tau \in G$ with $\sigma(\alpha) = \alpha\varepsilon, \sigma(\varepsilon) = \varepsilon$ and $\tau(\alpha) = \alpha, \sigma(\varepsilon) = \varepsilon^2$. Note that in terms of permutations on five roots, σ is the 5-cycle $(1,2,3,4,5)$ and τ is the 4-cycle $(2,3,5,4)$. These two cycles generate a subgroup of S_5 of order 20, which must be the Galois group. This group can be described by generators and relations as follows:

$$G = \langle \sigma,\tau \mid \sigma^5 = \tau^4 = 1, \tau^{-1}\sigma\tau = \sigma^3 \rangle.$$

(Alternatively, one may deduce as noted above that G is a transitive subgroup of S_5 of order 20, and then show that $C_5 \rtimes C_4$ is the only such (up to a conjugation).)

(c) $\sqrt[3]{2}$ is the only root of $x^3-2 \in \mathbb{Q}[x]$ contained in $\mathbb{Q}(\sqrt[3]{2},\sqrt{2})$, so any element of the Galois group G fixes the subfield $\mathbb{Q}(\sqrt[3]{2})$, whence $G \cong \mathrm{Gal}(\mathbb{Q}(\sqrt[3]{2},\sqrt{2})/\mathbb{Q}(\sqrt[3]{2})) \cong C_2$.

(d) $C_2 \times S_3$ (using the Problem 4.17 from Rotman).

Problem 2.17.22 Prove that the quotient ring $R := \mathbb{F}_3[x]/(x^2+1)$ is a field of order 9. Exhibit an explicit generator for $R^\times$ (which should be the cyclic group of order 8).

Problem 2.17.23 True or false: if F is a field of characteristic p, $\alpha \in F$, then F contains at most one p^kth root of α.

Problem 2.17.24 True or false: every finite normal extension of $\mathbb{C}$ is normal over $\mathbb{R}$.

Problem 2.17.25 Let F be a field, and $F(x)$ be the field of rational functions over F. Prove that $F(x)/F(\frac{x^3}{x+1})$ is a simple extension. Find its degree and the minimal polynomial $\operatorname{irr}(x; F(\frac{x^3}{x+1}))$.

Solution. Let $K := F(\frac{x^3}{x+1})$. It is clear that $F(x) = K(x)$. Moreover, the polynomial

$$f(y) := y^3 - \frac{x^3}{x+1}y - \frac{x^3}{x+1} \in K[y]$$

clearly annihilates x. We claim that this polynomial is irreducible over K, which implies that $[F(x) : K] = 3$ and $f = \operatorname{irr}(x; K)$.

To prove the claim, note that $\deg f = 3$, and so it is reducible over K if and only if it has a root in K. Over $F(x)$ we can decompose

$$f(y) = (y - x)(y^2 + xy + \frac{x^2}{x+1}).$$

So if f has a root in K, this root is either x or is a root of

$$g(y) := y^2 + xy + \frac{x^2}{x+1}.$$

Note that non-zero rational functions in x has a well-defined notion of degree: if $r(x) = p(x)/q(x)$ is such a function we define $\deg r = \deg p - \deg q$. Next, observe that all elements of K has even degree. Now, $x \notin K$ as it has degree 1. Moreover, let $r \in K$ be a root of g. Then r^2 has even degree, xr has odd degree, and $\frac{x^2}{x+1}$ also has odd degree. It follows that $xr + \frac{x^2}{x+1} = 0$ and $r^2 = 0$, whence $r = 0$, which leads to a contradiction.

The problem is a special case of claim (*) proved in Problem 2.17.42.

Problem 2.17.26 True or false: if $[K : k] = 2$ the K/k is normal.

Problem 2.17.27 For extensions $F/K/k$, if F/K and K/k are separable then F/k is separable.

Problem 2.17.28 Let p be a prime. Then there are exactly $(q^p - q)/p$ monic irreducible polynomials of degree p in $\mathbb{F}_q[x]$ (q is not necessarily a power of p).

Solution. Consider the field extension $\mathbb{F}_{q^p}/\mathbb{F}_q$. Let $\alpha \in \mathbb{F}_{q^p} \setminus \mathbb{F}_q$. The extension $\mathbb{F}_{q^p}/\mathbb{F}_q$ has no non-trivial intermediate fields (as p is prime), so $\mathbb{F}_q(\alpha) = \mathbb{F}_{q^p}$, whence $\deg \operatorname{irr}(\alpha; \mathbb{F}_q) = p$. Moreover, $\operatorname{irr}(\alpha; \mathbb{F}_q)$ has p

distinct roots in $\mathbb{F}_{q^p}$, as the extension $\mathbb{F}_{q^p}/\mathbb{F}_q$ is normal and separable. Therefore considering the minimal polynomials of the elements of $\mathbb{F}_{q^p}\setminus\mathbb{F}_q$ yields $(q^p-q)/p$ distinct irreducible monic polynomials over $\mathbb{F}_q$.

It remains to show that any irreducible monic polynomial $f\in\mathbb{F}_q[x]$ of degree p has form $\operatorname{irr}(\alpha;\mathbb{F}_q)$ for some element $\alpha\in\mathbb{F}_{q^p}\setminus\mathbb{F}_q$. Note that the field F obtained from $\mathbb{F}_q$ by adjoining a root α of f has degree p over $\mathbb{F}_q$, so it has q^p elements. By uniqueness of finite fields, we know that F is $\mathbb{F}_q$-isomorphic to $\mathbb{F}_{q^p}$, and the result follows.

Problem 2.17.29 Let $q=p^n$ and $d|n$. Then $\mathbb{F}_q$ contains exactly one subfield with p^d elements. Conversely, if $\mathbb{F}_{p^d}$ is a subfield of $\mathbb{F}_{p^n}$ then $d|n$.

Problem 2.17.30 If K/k is a finite normal separable field extension, then there exists an irreducible polynomial $f\in k[x]$ such that K is a splitting field for f over k.

Problem 2.17.31 True or false? Every finite field extension is simple

Problem 2.17.32 If F is a finite field and $a,b\in F$ are not squares then ab is a square.

Problem 2.17.33 Let $F/K/k$ with K/k finite separable and F/K simple. Then F/k is simple.

Problem 2.17.34 Suppose that $f\in k[x]$ with roots $\alpha_1,\dots,\alpha_n$ in a splitting field for f. Show that $\Delta=\varepsilon_n\prod_{i=1}^n f'(\alpha_i)$, where $\varepsilon_n=1$ if $n=0$ or $1\pmod 4$, and $\varepsilon_n=-1$ otherwise.

Problem 2.17.35 True or false? There is an irreducible polynomial of degree 4 over $\mathbb{Q}$ whose splitting field has degree 6 over $\mathbb{Q}$.

Problem 2.17.36 Let φ be the Euler function.

(i) Prove that $\varphi(ab)=\varphi(a)\varphi(b)$ providing $(a,b)=1$.

(ii) Prove that $\varphi(m)=m\prod_{p|n}\frac{p-1}{p}$, where the product is over all prime divisors of m.

(iii) Prove that $m=\sum_{d|m}\varphi(d)$.

Problem 2.17.37 Let $m\geq 3$, φ be the Euler function, and $\varepsilon\in\mathbb{C}$ be a primitive mth root of 1. Prove that $[\mathbb{Q}(\varepsilon+\varepsilon^{-1}):\mathbb{Q}]=\varphi(m)/2$.

Solution. It suffices to prove that $[\mathbb{Q}(\varepsilon) : \mathbb{Q}(\varepsilon+\varepsilon^{-1})] = 2$. As $\varepsilon+\varepsilon^{-1} \in \mathbb{R}$ and $\mathbb{Q}(\varepsilon) \not\subseteq \mathbb{R}$, we have $[\mathbb{Q}(\varepsilon) : \mathbb{Q}(\varepsilon+\varepsilon^{-1})] \geq 2$. The converse inequality follows from the fact that the quadratic polynomial $x^2-(\varepsilon+\varepsilon^{-1})x+1 \in \mathbb{Q}(\varepsilon+\varepsilon^{-1})[x]$ annihilates ε.

Problem 2.17.38 Let $m > 1$ be an odd integer. Show that $\Phi_{2m}(x) = \Phi_m(-x)$.

Solution. Note that if ξ is a primitive mth root of 1 then $-\xi$ is a primitive $2m$th root of 1. Therefore $\Phi_m(-x)$ divides $\Phi_{2m}(x)$ over $\mathbb{Q}$, see Lemma 2.11.2(i). As both polynomials are irreducible and monic, it now follows that they are equal.

Problem 2.17.39 If p is prime then $\Phi_{p^n}(x) = 1+x^{p^{n-1}}+x^{2p^{n-1}}+\cdots+x^{(p-1)p^{n-1}}$.

Problem 2.17.40 True or false? For $f \in k[x]$, if $\mathrm{Gal}(f;k)$ acts transitively on the roots of f then f is irreducible over k.

Problem 2.17.41 True or false? Let K/k be a field extension. If $\alpha_1, \ldots, \alpha_n \in K$ are algebraically independent over k, and $\alpha \notin k$ is the element of $k(\alpha_1, \ldots, \alpha_n)$, then α is transcendental over k.

Solution. True

Problem 2.17.42 Let k be a field and x be transcendental over k. Describe the group $\mathrm{Gal}(k(x)/k)$.

Solution. Let $f/g \in k(x)$ with $f/g \notin k$ and f, g relatively prime in $k[x]$. Denote $z := f/g$. We first prove the following claim:

(*) The element x is algebraic over $k(z)$ and

$$[k(x) : k(z)] = \max(\deg f, \deg g).$$

Note that x is a root of

$$\varphi(y) := zg(y) - f(y) \in k(z)[y]$$

of degree $\max(\deg f, \deg g)$. So it remains to prove that φ is irreducible. Note that z is transcendental over k, as otherwise $k(x)/k(z)/k$ would be algebraic. In order to prove that $\varphi(y)$ is irreducible over $k(z)$ it suffices to check that it cannot be decomposed over $k[z]$ (this is a version of

Lemma 2.2.1 generalized to $k[z]$ instead of $\mathbb{Z}$ and $k(z)$ instead of $\mathbb{Q}$ but the proof is the 'same'). Well, the degree of z in φ is 1, so if $\varphi(y) = \varphi_1(y)\varphi_2(y)$ in $k[z][y]$, then z appears in the coefficients of only one of φ_1, φ_2, say of φ_1. Then we can write

$$\varphi_1 = (a_0 + zb_0) + \cdots + (a_n + b_n z)y^n, \quad \varphi_2 = \alpha_0 + \cdots + \alpha_m y^m,$$

whence

$$\begin{aligned} zg(y) - f(y) = (b_0 + \cdots + b_n y^n)(\alpha_0 + \cdots + \alpha_m y^m)z \\ +(a_0 + \cdots + a_n y^n)(\alpha_0 + \cdots + \alpha_m y^m). \end{aligned}$$

Therefore $g(y) = (b_0 + \cdots + b_n y^n)(\alpha_0 + \cdots + \alpha_m y^m)$ and $-f(y) = (a_0 + \cdots + a_n y^n)(\alpha_0 + \cdots + \alpha_m y^m)$, which contradicts the fact that $(f, g) = 1$. The claim (*) is proved.

Now, let σ be an k-automorphism of $k(x)$. If $\sigma(x) = f/g$, it follows from (*) that $\max(\deg f, \deg g) = 1$. So we can write $f(x) = (ax + b)/(cx + d)$. Remember that we have $(f, g) = 1$, so the determinant of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is non-zero. Note however, that the matrix is uniquely defined only up to a scalar. So to each $\sigma \in \mathrm{Gal}(k(x)/k)$ we have associated an element of $g(\sigma) \in PGL_2(k)$. Moreover, it is easy to see that $g(\sigma\tau) = g(\sigma)g(\tau)$. It follows that $g : \mathrm{Gal}(k(x)/k) \to PGL_2(k)$ is a group homomorphism. It is now easy to check that g is an isomorphism.

Problem 2.17.43 Let $F/K/k$ be a field extensions and $A \subseteq F$ be a subset. If K/k is algebraic then $K(A)/k(A)$ is also algebraic.

Problem 2.17.44 True or false? If G is a finite group of automorphisms of a field K and $k = K^G$ is the fixed field of G, then K/k is Galois and $\mathrm{Gal}(K/k) = G$.

Problem 2.17.45 True or false? If k is a field, $f \in k[x]$, and K/k is a splitting field for f over k then $[K : k] \leq (\deg f)!$.

Problem 2.17.46 True or false? If k is a field, $f \in k[x]$, K/k is a splitting field for f over k, and $[K : k] = (\deg f)!$, then f is separable and irreducible.

Problem 2.17.47 Calculate $\mathrm{Gal}(\mathbb{R}/\mathbb{Q})$.

Problem 2.17.48 Calculate $\mathrm{Gal}(\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})/\mathbb{Q})$.

Solution. Consider

$$K := \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})/\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}(\sqrt{2})/\mathbb{Q}.$$

We claim that the containments are strict and of dimension 2. We will prove even more: all extensions

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{p_1}) \subset \mathbb{Q}(\sqrt{p_1}, \sqrt{p_2}) \subset \dots$$

are strict for $p_1 < p_2 < \dots$ being all the primes. Induction on n starting from $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$. If $\sqrt{p_{n+1}} \in \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_n})$, then $\sqrt{p_{n+1}} = a + b\sqrt{p_n}$, where $a, b \in \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_{n-1}})$. Then $p_{n+1} = a^2 + 2ab\sqrt{p_n} + b^2 p_n$, whence $\sqrt{p_n} \in \mathbb{Q}(\sqrt{p_1}, \dots, \sqrt{p_{n-1}})$, a contradiction.

Now $1, \sqrt{2}, \sqrt{3}, \sqrt{6}, \sqrt{5}, \sqrt{10}, \sqrt{15}, \sqrt{30}$ is a basis for K over $\mathbb{Q}$. Define three automorphisms from $\mathrm{Gal}(K/\mathbb{Q})$ by requiring that they negate exactly one of $\sqrt{2}, \sqrt{3}, \sqrt{5}$ and fix the other two. Then $\langle \sigma_1, \sigma_2, \sigma_3 \rangle \cong C_2 \times C_2 \times C_2$.

Problem 2.17.49 Let K/k be a Galois extension, and L, M be intermediate fields. Denote by LM the minimal subfield of K containing L and M.

(i) Prove that $(L \cap M)^* = \langle L^*, M^* \rangle$.

(ii) Prove that $(LM)^* = L^* \cap M^*$.

(iii) Assume that L/k is normal. Prove that

$$\mathrm{Gal}(LM/M) \cong \mathrm{Gal}(L/(L \cap M)).$$

Solution. (i) $\langle L^*, M^* \rangle$ is the smalles subgroup containing both L^* and M^*. By the Fundamental theorem, $\langle L^*, M^* \rangle^*$ is the largest subfield contained in both L and M, i.e. $\langle L^*, M^* \rangle^* = L \cap M$. Hence $(L \cap M)^* = (\langle L^*, M^* \rangle^*)^* = \langle L^*, M^* \rangle$. The proof of (ii) is entirely similar.

(iii) Let $\sigma \in \mathrm{Gal}(LM/M)$. Then $\sigma(L) \subseteq L$, as by assumption L is generated over k by the roots of a certain polynomial over k. So restriction yields a well-defined homomorphism $\varphi : \mathrm{Gal}(LM/M) \to \mathrm{Gal}(L/(L \cap M))$. The kernel of φ is clearly trivial. To show that φ is surjective, let $\tau \in \mathrm{Gal}(L/(L \cap M))$. As LM/M is normal, so it suffices to show that the second group has the order not exceeding the order of the first group. Note that both LM/M and $L/(L \cap M)$ are Galois, so we just need to show that $[LM : M] \geq [L : L \cap M]$. Well, using the fundamental theorem, (i), (ii), and the second isomorphism theorem, we

have

$$[LM : M] = [M^* : (LM)^*] = [M^*/(L^* \cap M^*)]$$
$$= [L^*M^*/L^*] = [L : L \cap M].$$

Problem 2.17.50 Let k be a subfield of $\mathbb{R}$ and $f \in k[x]$ is irreducible cubic with discrininant D. Then

(i) $D > 0$ if and only if f has three real roots.

(ii) $D < 0$ if and only if f has precisely one real root.

Problem 2.17.51 Let $\operatorname{char} k \neq 2$ and $f \in k[x]$ is a cubic whose discriminant has a square root in k, then f is either irreducible or splits in k.

Problem 2.17.52 Let f be an irreducible separable quartic over a field k and α be a root of f. There is no field properly between k and $k(\alpha)$ if and only if the Galois group of f is either A_4 or S_4.

Problem 2.17.53 Every element in a finite field can be written as a sum of two squares.

Solution. Let our field F have p^n elements. If $p = 2$, the map $x \mapsto x^2$ is a bijection, and so every element is a square ($+0^2$). Let $p > 2$. Then the multiplicative group $F^\times \cong C_{p^n-1}$ has even order, and so exactly half of its elements are squares. As 0 is also a square, there are $(p^n + 1)/2$ squares in F. Now, let $\alpha \in F$ be any element. Consider the sets $X = \{\alpha - s \mid s \in F \text{ is a square}\}$ and $S = \{s \in F \mid s \text{ is a square}\}$. As $|X| + |S| > p^n$, we have $X \cap S \neq \varnothing$. So α minus a square is a square, and the result follows.

Problem 2.17.54 Determine the Galois group of $x^3 + 11$ over $\mathbb{Q}$, determine all subfields of its splitting field, and decide whether these subfields are normal over $\mathbb{Q}$. Describe at least two of the subfields by their generators.

Solution. The discriminant $D = -27 \cdot 11^2$ does not have square root in $\mathbb{Q}$, so the Galois group G is S_3 acting on the three roots

$$\{-\sqrt[3]{11}, -\varepsilon\sqrt[3]{11}, -\varepsilon^2\sqrt[3]{11}\},$$

where $\varepsilon = e^{2\pi i/3}$. There are four non-trivial proper subgroups in G: $G_1 = \langle(12)\rangle$, $G_2 = \langle(13)\rangle$, $G_3 = \langle(23)\rangle$, $G_4 = A_3$, only the last of

which is normal. The corresponding fixed fields are the only proper non-trivial intermediate subfields, of which only G_4^* is normal. Moreover, G_4^* is generated by $\sqrt{D}$, so $G_4^* = \mathbb{Q}(i\sqrt{3})$. Finally, (23) is the complex conjugation, $G_3^* = \mathbb{Q}(\sqrt[3]{11})$.

Problem 2.17.55 Let K be a splitting field for x^4-2 over $\mathbb{Q}$. Determine $\mathrm{Gal}(K/\mathbb{Q})$. How many subfields does K have of degree 2 over $\mathbb{Q}$.

Problem 2.17.56 Find all subfields of the splitting field of $x^3 - 7$ over $\mathbb{Q}$. Be sure to justify your assertions. Which of the subfields are normal over $\mathbb{Q}$?

Solution. The splitting field $K = \mathbb{Q}(\sqrt[3]{7}, \omega)$ where ω is a primitive cube root of 1. Then $[K : \mathbb{Q}] = [K : \mathbb{Q}(\sqrt[3]{7})][\mathbb{Q}(\sqrt[3]{7}) : \mathbb{Q}] = 6$. So the Galois group is S_3. The proper subfields correspond to the proper subgroups $\langle(123)\rangle$, $\langle(12)\rangle$, $\langle(13)\rangle$, $\langle(23)\rangle$. The corresponding fixed fields are the fields obtained by adjoining ω, and the different roots of $x^3 - 7$.

Problem 2.17.57 Let K be a splitting field for $x^4 + 6x^2 + 5$ over $\mathbb{Q}$. Find all subfields of K.

Solution. We have $x^4+6x^2+5 = (x^2+1)(x^2+5)$. And the Galois group is $C_2 \times C_2$ by the Problem 4.17 in Rotman (or directly).

Problem 2.17.58 Let K be a splitting field for $x^4 - 3$ over $\mathbb{Q}(i)$ (where $i = \sqrt{-1} \in \mathbb{C}$). Find the Galois group of K over $\mathbb{Q}(i)$.

Solution. We claim that the group is C_4. Indeed, it suffices to note that $[K : \mathbb{Q}(i)] = 4$ and the Galois group is transitive. That the degree is 4 follows by considering the tower $K/\mathbb{Q}(\sqrt[4]{3})/\mathbb{Q}$ (see Example 2.8.2 for more details). That the Galois groups is transitive follows from the fact that $K = \mathbb{Q}(i)(\sqrt[4]{3})$.

Problem 2.17.59 Let $\varepsilon \in \mathbb{C}$ be a primitive 7th root of 1. Determine the minimal polynomial of ε, the structure of the Galois group of $\mathbb{Q}(\varepsilon)$ over $\mathbb{Q}$, and the proper subfields of $\mathbb{Q}(\varepsilon)$ by giving field generators.

Problem 2.17.60 Let $K = \mathbb{Q}(i, e^{2\pi i/3})$, where $i = \sqrt{-1} \in \mathbb{C}$. Find $[K : \mathbb{Q}]$ and determine $\mathrm{Gal}(K/\mathbb{Q})$.

Problem 2.17.61 Let K be a splitting field over $\mathbb{Q}$ of the polynomial

$(x^3 - 5)(x^2 + 1)$. Describe the Galois group $\mathrm{Gal}(K/\mathbb{Q})$. How many subfields does K have with extension degree 2 over $\mathbb{Q}$.

Problem 2.17.62 Find the Galois groups of $x^4 + 1$ and $x^5 + 1$ over $\mathbb{Q}$.

Problem 2.17.63 Let K be a splitting field for $x^5 - 2$ over $\mathbb{Q}$ and let $G = \mathrm{Gal}(K/\mathbb{Q})$. Compute the order of G and prove that G is not abelian.

Problem 2.17.64 Show that the Galois group of $x^5 - 2$ over $\mathbb{Q}$ has a normal Sylow 5-subgroup.

Problem 2.17.65 True or false? The Galois group of the polynomial $x^3 - 5$ over $\mathbb{Q}$ is abelian.

Problem 2.17.66 Let K be a finite field with a prime field k.

(a) Prove that $\mathrm{Gal}(K/k)$ is cyclic.

(b) Prove that if L is an arbitrary subfield of K then $\mathrm{Gal}(K/L)$ is cyclic.

Problem 2.17.67 Let $K = \mathbb{F}_2[x]/(x^6+x+1)$ and $\alpha = x+(x^6+x+1) \in K$.

(a) Prove that the polynomial $x^6 + x + 1$ is irreducible over $\mathbb{F}_2$.

(b) Prove that $\alpha^i \neq 1$ for $i = 1, 2, 3, \ldots, 22$.

(c) Find $\mathrm{irr}\,(\alpha; \mathbb{F}_2)$.

(d) Show that the mapping $a \mapsto a^4$ is an automorphism of K.

(e) Find $\mathrm{irr}\,(\alpha^4; \mathbb{F}_2)$.

(f) Show that $\{0, \alpha^9, \alpha^{18}, \alpha^{27}, \alpha^{36}, \alpha^{45}, \alpha^{54}, \alpha^{63}\}$ is a subfield of K.

Problem 2.17.68 True or false? Let $\mathrm{char}\, k = p$. Then k contains exactly one p^mth root of 1 for any natural number m.

Problem 2.17.69 Prove that the quotient ring $R = \mathbb{F}_3[x]/(x^2 + 1)$ is a field of order 9. Exhibit an explicit generator for the multiplicative group $R^\times$.

Problem 2.17.70 True or false? A field of order 27 is a Galois extension of a field of order 9.

Problem 2.17.71 True or false? If p is a prime and n is a positive

integer, then a field of order p^n contains a subfield of order p^m for each integer $1 \leq m \leq n$.

Problem 2.17.72 True or false? If F is a finite field and $f \in F[x]$ is irreducible then $f' \neq 0$.

Problem 2.17.73 True or false? If F and K are finite fields whose additive groups are isomorphic then F and K are isomorphic as rings.

Problem 2.17.74 True or false? If F and K are finite fields whose multiplicative groups are isomorphic then F and K are isomorphic as rings.

Problem 2.17.75 True or false? $\mathbb{F}_{81}$ has exactly three subfields counting itself.

Problem 2.17.76 True or false? There is a Galois extension of $\mathbb{F}_8$ with Galois group $C_2 \times C_2$.

Problem 2.17.77 True or false? There is a Galois extension of $\mathbb{F}_{125}$ with the Galois group C_6.

Problem 2.17.78 True or false? Every Galois extension of $\mathbb{C}$ is Galois over $\mathbb{R}$.

Problem 2.17.79 True or false? Let k be a finite field. Then for every prime p, k has a finite Galois extension with Galois group isomorphic to the symmetric group S_p.

Problem 2.17.80 Let k be a subfield of an algebraically closed field K such that the transcendence degree of K over k is finite. Prove that if $\varphi : K \to K$ is a ring homomorphism, which is identity on k, then φ is an automorphism of K.

Solution. First of all φ is injective. Now, let B be a (necessarily finite) transcendence basis. Then $\varphi(B)$ is algebraically independent over k and has the same order as B. So $\varphi(B)$ is also a transcendence basis for K and for $\varphi(K)$. Thus K is algebraic over $k(\varphi(B))$, and hence over $\varphi(K)$. But $\varphi(K)$ is algebraically closed, so $\varphi(K) = K$.

Problem 2.17.81 True or false? A field with exactly 16 elements has a unique subfield with exactly 8 elements.

Problem 2.17.82 True or false? A field with exactly 16 elements has a unique subfield with exactly 4 elements.

Problem 2.17.83 True or false? Let $f \in \mathbb{F}_p[x]$ such that $f' = 0$. Then the splitting field for f over $\mathbb{F}_p$ is not separable over $\mathbb{F}_p$.

Problem 2.17.84 True or false? A field of order 243 contains exctly one proper subfield.

Problem 2.17.85 True or false? Let $K/\mathbb{F}_q$ be a finite extension, and L, M be two intermediate subfields. Then either $L \subseteq M$ or $M \subseteq L$.

Problem 2.17.86 True or false? In a finite extension of a finite field every intermediate field is stable (with respect to the Galois group).

Problem 2.17.87 Let K/k be a finite Galois extension and L be an intermediate field. Prove that L/k is Galois implies $GL \subseteq L$.

Problem 2.17.88 True or false? The field extension $\mathbb{Q}(x)/\mathbb{Q}(x^6)$ is a Galois extension.

Problem 2.17.89 Let K/k be a finite field extension and L, M be intermediate fields. Prove $[LM : k] \leq [L : k][M : k]$ and show by example that a strict inequality is possible even if $L \cap M = k$.

Solution. We know that $[LM : k] = [LM : L][L : k]$, so it suffices to show that $[LM : L] \leq [M : K]$. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis of M over k. Then, as $LM = L(\alpha_1, \ldots, \alpha_n) = L[\alpha_1, \ldots, \alpha_n]$, whence any element of LM is an L-linear combination of elements in M, and so the α_i span LM over L. An example is provided by K, the splitting field of $x^3 - 2$ over $k = \mathbb{Q}$. Let ω and τ be distinct complex roots of $x^3 - 2$. Let $\mathbb{Q}(\omega) = L$ and $\mathbb{Q}(\tau) = M$. Then $[L : k] = [M : k] = 3$ and so $L \cap M = k$ but $LM = K$, and $[K : k] = 6$.

Problem 2.17.90 True or false? The splitting field K for $x^3 - 5$ over $\mathbb{Q}$ has exactly three automorphisms.

Problem 2.17.91 Let k be a field, $p(x)$ be an irreducible polynomial

in $k[x]$ of degree n, and let K be a Galois extension of k containing a root α of $p(x)$. Let $G = \mathrm{Gal}(K/k)$, and G_α be the set of all $\sigma \in G$ with $\sigma(\alpha) = \alpha$.

(a) Show that G_α has index n in G.

(b) Show that if G_α is normal in G, then $p(x)$ splits in the fixed field of G_α.

Solution. Note that G acts transitively on the roots of p. Moreover, p must be separable as α belongs to the separable extension K of k, and so p has n roots. This proves (a). Note further that $G_\alpha^* = k(\alpha)$. If G_α is normal it follows from the fundamental theorem that $k(\alpha)$ is normal, which implies (b).

Problem 2.17.92 Let $k(\alpha)/k$ be an field extension obtained by adjoining a root α of an irreducible separable polynomial $f \in k[x]$. Then there exists an intermediate field $k \subsetneq F \subsetneq k(\alpha)$ if and only if the Galois group $\mathrm{Gal}(f; k)$ is imprimitive. If the group is imprimitive then the subfield F can be chosen so that $[F : K]$ is equal to the number of imprimitivity blocks.

Solution. We proved in Problem 2.17.91 that $k(\alpha) = G_\alpha^*$. Moreover a transitive permutation group $G = \mathrm{Gal}(f; k)$ on the roots is primitive if and only if the point stabilizer G_α is maximal. This proves the first claim, thanks to the Galois correspondence. If G is imprimitive, we can choose an overgroup H which is maximal and $[G : H]$ equals the number of the imprimitivity blocks (choose H to be a stabilizer of the block B containing α). Now take $F = H^*$.

Problem 2.17.93 Let K/k be a Galois extension and p be a prime number.

(a) Prove that K has an intermediate subfield L such that $[K : L]$ is a prime power.

(b) Prove that if L_1 and L_2 are intermediate subfields with $[K : L_1]$, $[K : L_2]$ both p-powers, and $[L_1 : k]$, $[L_2 : k]$ both prime to p, then L_1 is k-isomorphic to L_2.

Problem 2.17.94 True or false? Let K/k be a field extension with $1 < [K : k] < \infty$. Then $|\mathrm{Gal}(K/k)| > 1$.

Problem 2.17.95 True or false? The Galois group $\mathrm{Gal}(\mathbb{Q}(\sqrt[3]{5})/\mathbb{Q})$ is cyclic of order 3.

Problem 2.17.96 True or false? $\mathrm{tr.deg}(\mathbb{C}/\mathbb{Q})$ is infinite.

Problem 2.17.97 Show that F is algebraically closed if and only if F has no proper algebraic extensions.

Problem 2.17.98 There exist field extensions E/k that do not have an intermediate field K with K/k algebraic and E/K purely transcendental.

Solution. An example of such is given by $\mathbb{C}/\mathbb{Q}$. Indeed, assume that K is an intermediate field such that $K/\mathbb{Q}$ is algebraic. Then $K \subsetneq \mathbb{C}$. As $\mathbb{C}$ is algebraically closed, the result now follows from the following claim: purely transcendental extension is never algebraically closed. To prove the claim, consider a purely transcendental extension $F(B)/F$ for any field F. Here B is a transcendence basis of $F(B)$ over F. Let $x \in B$. Then the polynomial $y^2 - x \in F(B)[y]$ does not split in $F(B)$: otherwise there is an element $p/q = p(x, x_1, \ldots, x_n)/q(x, x_1, \ldots, x_n) \in F(B)$ with $(p/q)^2 = x$ or $p^2 = xq^2$, which is a contradiction because the degree of the left hand side with respect to x is even and the degree of the right hand side with respect to x is odd.

3
Modules

3.1 Definition and the first properties

The idea of a module is similar to that of a group acting on a set, except that there is more structure. Roughly speaking, a module is a 'ring acting on an abelian group'. As usual we assume that all rings have 1.

Definition 3.1.1 Let R be a ring.

A *left R-module* is an abelian group V together with a map

$$R \times V \to V, (r, v) \mapsto rv$$

(called the left action of R on V or left R-module structure on V) such that

(i) $r_1(r_2v) = (r_1r_2)v$ for all $r_1, r_2 \in R$, $v \in V$;
(ii) $r(v_1 + v_2) = rv_1 + rv_2$ for all $r \in R$, $v_1, v_2 \in V$;
(iii) $(r_1 + r_2)v = r_1v + r_2v$ for all $r_1, r_2 \in R$, $v \in V$;
(iv) $1v = v$ for all $v \in V$.

A *right R-module* is an abelian group V together with a map

$$V \times R \to V, (r, v) \mapsto vr$$

(called the right action of R on V or right R-module structure on V) such that

(i) $(vr_1)r_2 = v(r_1r_2)$ for all $r_1, r_2 \in R$, $v \in V$;
(ii) $(v_1 + v_2)r = v_1r + v_2r$ for all $r \in R$, $v_1, v_2 \in V$;
(iii) $v(r_1 + r_2) = vr_1 + vr_2$ for all $r_1, r_2 \in R$, $v \in V$;
(iv) $v1 = v$ for all $v \in V$.

If S is another ring, and V is both left R-module and right S-module, such that $r(vs) = (rv)s$ for all $r \in R$, $s \in S$, and $v \in V$, then we say that V is an (R, S)-*bimodule*. If $R = S$, we speak of an R-*bimodule*.

If we want to emphasize that V is a left R-module (resp. right R-module, resp. (R,S)-bimodule) we write ${}_RV$ (resp. V_R, resp. ${}_RV_S$).

Even before we consider the first examples, some informal remarks are in order. First of all, informally speaking there is not much difference between left and right modules. To be more precise, the following construction construction 'reduces' right modules to left modules. If R is a ring, the *opposite* ring R^{op} has the same underlying abelian group as R and the opposite multiplication $r * s = sr$. Now, given a right R-module V we can define the left R^{op}-module V^{op} via $rv = vr$. It is not hard to believe that V and V^{op} have the 'same properties' (this will be made precise later when we understand what kind of properties are of interest). Note that if R is commutative, $R = R^{\mathrm{op}}$, and 'there is no difference' between left and right R-modules. So in this case we will just speak of *R-modules*.

Here are the first examples.

Example 3.1.2 (i) If F is a field, the F-module is the same as F-vector space.

(ii) The notion of an abelian group is equivalent to the notion of a $\mathbb{Z}$-module (make sure you agree).

(iii) If R is a ring, the ring multiplication defines on R the structure of a left R-module (resp. right R-module, resp. R-bimodule) called the *regular (bi)module*. The notation is ${}_RR$, R_R or ${}_RR_R$, respectively. More generally, if $R \subseteq S$ is a subring, then ring multiplication defines on R the structure of a left S-module (resp. right S-module, resp. S-bimodule).

(iv) If I is a left ideal in R then R/I is naturally a left R-module.

(v) If V is an F-vector space and θ is a linear transformation of V, V becomes an $F[x]$-module where x acts as θ.

The following properties follow immediately from the definition.

Lemma 3.1.3 *Let V be a left R-module. Then*

$$r0 = 0,\ 0v = 0,\ r(v-w) = rv - rw,\ (r-s)v = rv - sv$$

for all $r, s \in R,\ v, w \in V$.

We note that like with group actions on sets, there is another way to think about left R-modules: having an R-module structure on an abelian group V is equivalent to having a ring homomorphism $R \to \mathrm{End}(V)$, where $\mathrm{End}(V)$ is the ring of all group homomorphisms from V to itself.

The kernel of this homomorphism is called the *annihilator* of V, denoted $\mathrm{Ann}(V)$ or $\mathrm{Ann}_R(V)$. In other words,

$$\mathrm{Ann}(V) = \{r \in R \mid rv = 0 \text{ for any } v \in V\}.$$

The module V is called *faithful* if $\mathrm{Ann}(V) = 0$. If S is any subset of V we denote by $\mathrm{Ann}(S)$ the set of all $r \in R$ which annihilate every element in S. Note that $\mathrm{Ann}(S)$ is a left ideal in R. We need more definitions.

Definition 3.1.4 Let V, W be two left R-modules. A *homomorphism* from V to W (also called R-homomorphism) is a homomorphism

$$\varphi : V \to W$$

of abelian groups such that $\varphi(rv) = r\varphi(v)$ for all $r \in R$ and $v \in V$. The set of all homomorphisms from V to W is denoted by $\mathrm{Hom}_R(V, W)$. It has a natural structure of an abelian group.

Surjective (resp. injective, resp. bijective) homomorphism is called *monomorphism* (resp. *epimorphism*, resp. *isomorphism*).

A homomorphism from V to V is also called an *endomorphism* of V. The set of all endomorphisms of V is denoted by $\mathrm{End}_R(V, W)$. It has a natural structure of ring.

Definition 3.1.5 Let V be a left R-module. A subgroup $W \subseteq V$ is called a *submodule* (or R-submodule) if $rw \in W$ for all $r \in R$ and $w \in W$.

Note that it follows from this definition that a submodule contains $0 \in V$ and so it is a non-empty subset.

Example 3.1.6 Note that submodules of ${}_RR$ (resp. ${}_RR_R$, resp R_R) are precisely left (resp. right, resp. two-sided) ideals of R.

The following is easy to check:

Lemma 3.1.7

(i) *If $\varphi : V \to W$ is a homomorphism of left R-modules, then* $\ker \varphi \subseteq V$ *and* $\mathrm{im}\, \varphi \subseteq W$ *are R-submodules.*

(ii) *Intersection of submodules is a submodule, union of a non-empty ascending chain of submodules is a submodule.*

Let X be a subset of an R-module V. The minimal submodule of V containing X is called the submodule generated by X and denoted $\langle X \rangle$.

It is clear that $\langle S\rangle$ consists of all finite linear combinations of the form $r_1x_1 + \cdots + r_nx_n$ with $r_i \in R, x_i \in X$. An R-module V is called *cyclic* if it is generated by a single element: $V = \langle v\rangle$ for some $v \in V$. In this case we also write $R = Rv$. An R module is called *finitely generated* if $V = \langle v_1, \ldots, v_n\rangle$ for some $v_1, \ldots, v_n \in V$.

If $\{V_i\}_{i\in I}$ is a family of R-submodules in V, their sum is defined as

$$\sum_{i\in I} V_i := \langle \cup_{i\in I} V_i\rangle.$$

It is easy to see that $\sum_{i\in I} V_i$ consists of all sums $\sum_{i\in I} v_i$, where $v_i \in V_i$ and all but finitely many v_i's are zero.

The notion of a quotient module is defined in an obvious way and we leave it as an exercise. Another standard (and highly recommended!) exercise is to state and prove the three isomorphism theorems and the correspondence theorem.

Lemma 3.1.8 *A left R-module is cyclic if and only if it is isomorphic to R/I for some left ideal I in R. If $V = Rv$ then $V \cong R/\mathrm{Ann}(v)$. If R is commutative then* $\mathrm{Ann}(Rv) = \mathrm{Ann}(R)$.

Proof The module R/I is cyclic because it is generated by the coset $1+I$. On the other hand, if $V = Rv$, then the map ${}_RR \to V, r \mapsto rv$ is a surjective R-module homomorphism, whose kernel is $\mathrm{Ann}(v)$, so by the First Isomorphism Theorem, we have $V \cong R/\mathrm{Ann}(v)$. The rest is clear. □

3.2 Direct sums and products

Definition 3.2.1 Let $\{V_i\}_{i\in I}$ be a family of left R-modules.

The *direct product* of the this family of modules is defined to be the cartesian product of the corresponding sets, i.e. the set $\prod_{i\in I} V_i$ of all families $(v_i)_{i\in I}$ such that $v_i \in V_i$ for all $i \in I$, with the pointwise operations:

$$(v_i)_{i\in I} + (w_i)_{i\in I} = (v_i + w_i)_{i\in I}, \quad r(v_i)_{i\in I} = (rv_i)_{i\in I}.$$

If $I = \varnothing$ then the product is interpreted as (0). If $I = \{1, 2, \ldots, n\}$ we also write $V_1 \times \cdots \times V_n$.

The direct sum of the family $\{V_i\}_{i\in I}$ is the submodule

$$\bigoplus_{i\in I} V_i := \{(v_i)_{i\in I} \in \prod_{i\in I} V_i \mid v_i = 0 \text{ for almost all } i\} \subseteq \prod_{i\in I} V_i.$$

If $I = \{1, 2, \ldots, n\}$ we also write $V_1 \oplus \cdots \oplus V_n$.

Of course, there is no difference between finite direct products and direct sums: $V_1 \oplus \cdots \oplus V_n = V_1 \times \cdots \times V_n$. This is not the case for the infinite families.

The direct product $\prod_{i\in I} V_i$ comes with *natural projections*

$$\pi_i : \prod_{i\in I} V_i \to V_i,\ (v_i)_{i\in I} \mapsto v_i \qquad (i \in I),$$

and the direct sum $\sum_{i\in I} V_i$ comes with *natural injections*

$$\iota_i : V_i \to \bigoplus_{i\in I} V_i, \qquad (i \in I),$$

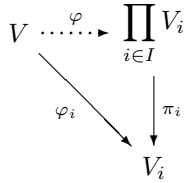
where $\iota_i(v_i)_j = v_i$ if $j = i$ and 0 otherwise. The following follows easily

Lemma 3.2.2 *Natural injections $\iota_i : V_i \to \bigoplus_{i\in I} V_i, \qquad (i \in I)$ are indeed injective. Moreover every $v \in \bigoplus_{i\in I} V_i, \qquad (i \in I)$ can be written uniquely in the form $v = \sum_{i\in I} \iota_i(v_i)$, with almost all $v_i = 0$.*

We have the following important universal properties:

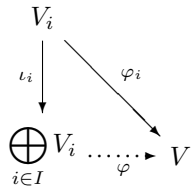
Theorem 3.2.3

(i) *Let V be a left R-module and $\{\varphi_i : V \to V_i\}_{i\in I}$ be a family of R-module homomorphisms. Then there exists a unique R-module homomorphism $\varphi : V \to \prod_{i\in I} V_i$ such that the following diagram commutes for every $i \in I$:*



(ii) *Let V be a left R-module and $\{\varphi_i : V_i \to V\}_{i\in I}$ be a family of R-module homomorphisms. Then there exists a unique R-module*

homomorphism $\varphi : \sum_{i\in I} V_i \to V$ such that the following diagram commutes for every $i \in I$:



Proof We prove (ii). Define

$$\varphi : \bigoplus_{i\in I} V_i \to V,\ (v_i)_{i\in I} \mapsto \sum_{i\in I} \varphi_i(v_i).$$

Note that this definition makes sense and the map has the desired properties. On the other hand if $\varphi \circ \iota_i = \varphi_i$ for all i then φ must be given by this formula. □

Corollary 3.2.4 *Let $\{V_i\}_{i\in I}$ be a family of left R-modules, and V be another left R-module. Then there exists natural isomorphisms of abelian groups*

$$\mathrm{Hom}_R(\bigoplus_{i\in I} V_i, V) \cong \prod_{i\in I} \mathrm{Hom}_R(V_i, V)$$

and

$$\mathrm{Hom}_R(V, \prod_{i\in I} V_i) \cong \prod_{i\in I} \mathrm{Hom}_R(V, V_i).$$

In particular, if $V_1, \ldots, V_n$ and $W_1, \ldots, W_m$ are left R-modules, we have a natural isomorphism

$$\mathrm{Hom}_R(\bigoplus_{i=1}^{n} V_i, \bigoplus_{j=1}^{m} W_j) \cong \bigoplus_{1\le i\le n,\ 1\le j\le m} \mathrm{Hom}_R(V_i, W_j).$$

Note that a family of module homomorphisms $\{\varphi_i : V_i \to W_i\}_{i\in I}$ induces natural module homomorhisms

$$\prod_{i\in I} \varphi_i : \prod_{i\in I} V_i \to \prod_{i\in I} W_i, \quad \text{and} \quad \bigoplus_{i\in I} \varphi_i : \bigoplus_{i\in I} V_i \to \bigoplus_{i\in I} V_i.$$

The map $\bigoplus_{i\in I} \varphi_i$ should not be confused with the map $\sum_{i\in I} \varphi_i$ which appears in the following circumstances. Given a family of module homomorphisms $\{\varphi_i : V \to W\}_{i\in I}$, such that for every $v \in V$ we have

$\varphi_i(v) = 0$ for almost all i, we have the map

$$\sum_{i\in I} \varphi_i : V \to W, \quad v \mapsto \sum_{i\in I} \varphi_i(v).$$

We give another characterization of direct sums.

Lemma 3.2.5 *Let $V, \{V_i\}_{i\in I}$ be left R-modules. Then $V \cong \bigoplus_{i\in I} V_i$ if and only if there exist homomorphisms $V_i \xrightarrow{\mu_i} V \xrightarrow{\rho_i} V_i$ such that*

(i) $\rho_i \circ \mu_i = \mathrm{id}_{V_i}$;
(ii) $\rho_i \circ \mu_j = 0$ *whenever* $i \neq j$;
(iii) *for each* $v \in V$, $\rho_i(v) = 0$ *for almost all* i;
(iv) $\sum_{i\in I} \mu_i \circ \rho_i = \mathrm{id}_V$.

Proof If $V = \bigoplus_{i\in I} V_i$ then we can take μ_i to be the standard injections and ρ_i to be the standard projections.

Conversely, by Lemma 3.2.3(ii), there exists a homomorphism $\theta : \bigoplus_{i\in I} V_i \to V$ such that $\theta \circ \iota_i = \mu_i$. Moreover, $\theta((v_i)_{i\in I}) = \sum_{i\in I} \mu_i(v_i)$. If $\theta((v_i)_{i\in I}) = 0$, then for every j, using (i) and (ii), we have

$$0 = \rho_j(\sum_{i\in I} \mu_i(v_i)) = v_j,$$

whence θ is injective. If $w \in V$, then $v := ((\rho_i(w))_{i\in I}) \in \bigoplus_{i\in I} V_i$ by (iii), and (iv) yields $w = \sum_{i\in I} \mu_i(\rho_i(w)) = \theta(v)$. So θ is surjective. □

The most important special case of the lemma is given by

Corollary 3.2.6 *A left R module X is isomorphic to a direct sum of left R-modules $V \oplus W$ if and only if there exist homomorphisms*

$$V \underset{\pi}{\overset{\mu}{\rightleftarrows}} X \underset{\nu}{\overset{\rho}{\rightleftarrows}} W$$

such that $\pi \circ \mu = \mathrm{id}_V$, $\rho \circ \nu = \mathrm{id}_W$, $\pi \circ \nu = 0$, $\rho \circ \mu = 0$, and $\mu \circ \pi + \nu \circ \rho = \mathrm{id}_X$.

Direct sums can also be characterized in terms of submodules rather than homomorphisms. The following result is a converse to Lemma 3.2.2.

Lemma 3.2.7 *Assume that $\{V_i\}_{i\in I}$ is a family of R-submodules in a left module V. If every element $v \in V$ can be written uniquely in the form $v = \sum_i v_i$ with $v_i \in V_i$ for all i and $v_i = 0$ for almost all i, then $V \cong \bigoplus_{i\in I} V_i$.*

Proof The inclusion homomorphisms $V_i \to V$ induce a homomorphism $\theta : \bigoplus_{i\in I} V_i \to V$ such that $\theta((v_i)_{i\in I}) = \sum_{i\in I} v_i$ for every $(v_i)_{i\in I} \in \bigoplus_{i\in I} V_i$. The hypothesis shows that θ is bijective. □

Definition 3.2.8 A left R-module V is the *internal direct sum* of submodules $\{V_i\}_{i\in I}$ if every element $v \in V$ can be written uniquely in the form $v = \sum_i v_i$ with $v_i \in V_i$ for all i and $v_i = 0$ for almost all i, then $V \cong \bigoplus_{i\in I} V_i$.

The notation $\bigoplus_{i\in I} V_i$ is also used to denote internal direct sums, which is justified by Lemma 3.2.7.

Proposition 3.2.9 *A left R-module V is the internal direct sum of submodules $\{V_i\}_{i\in I}$ if and only if*

(i) $V_i \cap (\sum_{j\neq i} V_j) = 0$ *for all* $i \in I$;
(ii) $V = \sum_{i\in I} V_i$.

Proof " $\Rightarrow$ " It is clear that $V = \sum_{i\in I} V_i$. If $v_i \in \cap(\sum_{j\neq i} V_j)$ then $v_i = \sum_{j\neq i} v_j$ can be written in this form in two ways, unless $v_i = 0$.

" $\Leftarrow$ " By (ii) every element $v \in V$ can be written in the form $v = \sum_{i\in I} v_i$ with $v_i \in V_i$ for all i and $v_i = 0$ for almost all i. If $\sum_{i\in I} v_i = \sum_{i\in I} w_i$ then for each i we have $v_i - w_i = \sum_{j\neq i}(w_j - v_j) \in V_i \cap (\sum_{j\neq i} V_j)$, which is 0 by (i). □

Corollary 3.2.10 *A left R-module X is the internal direct sum of its submodules V and W if and only if $V \cap W = 0$ and $V + W = X$.*

3.3 Simple and Semisimple modules

The following notion of a simple module is very important. It is similar to the notion of a simple group in the sense that simple modules are building blocks for more complicated modules, just like simple groups are building blocks for all finite groups.

Definition 3.3.1 A left R-module V is called *simple* (or *irreducible*) if $V \neq 0$ and V has no submodule different from 0 and V.

Example 3.3.2 (i) If F is a field, we saw that an F-module is the same as F-vector space. Such module is simple if and only if the corresponding vector space has dimension 1.

(ii) We saw a $\mathbb{Z}$-module is the same as an abelian group G. It is irreducible if and only if the group is non-trivial and has no subgroups except $\{1\}$ and itself. So the module is irreducible if and only if $G \cong C_p$, a cyclic group of a prime order.

The easiest (but not the only!) way to build more general modules out of simple modules is by taking direct sums. The modules obtained by this method are called semisimple.

Definition 3.3.3 A left R-module V is called *semisimple* (or *completely reducible*) if for any submodule $W \subseteq V$ there is a submodule $X \subseteq V$ such that $V = W \oplus X$.

Lemma 3.3.4 *Any submodule and factor-module of a semisimple module is semisimple.*

Proof Let V be a semisimple module and $W \subseteq V$. Now, let Y be a submodule of W. As Y is also a submodule of V there is a submodule $Z \subseteq V$ such that $V = Y \oplus Z$. Then it is easy to see that $W = Y \oplus (Z \cap W)$. Moreover, $V = W \oplus X$ for some submodule $X \subseteq V$, and $V/W \cong X$, so the result for quotient follows from the result on submodule. □

Theorem 3.3.5 *Let V be a left R-module. Then the following are equivalent:*

(i) *V is semisimple;*
(ii) *V is a direct sum of simple submodules;*
(iii) *V is a sum (not necessarily direct) of simple submodules.*

Proof We may assume that $V \neq 0$—otherwise the theorem is obvious.

(i) $\Rightarrow$ (ii) We first show that any nonzero submodule $W \subseteq V$ contains a simple submodule. For a fixed element $w \in W \setminus \{0\}$ consider the set of all submodules $W' \subseteq W$ such that $w \notin W'$. This set is non-empty, as it contains the zero submodule. By Zorn's Lemma, there exists a submodule W_0 which is a maximal element of this set. By Lemma 3.3.4, W is semisimple, so $W = W_0 \oplus W_1$ for some submodule $W_1 \subseteq W$. We claim that W_1 is irreducible. Indeed, if W_2 is a proper submodule of W_1, then $W_1 = W_2 \oplus W_3$, and $W = W_0 \oplus W_2 \oplus W_3$. Moreover, $(W_0 + W_2) \cap (W_0 + W_3) = W_0$, so either $w \notin W_0 + W_2$ or $w \notin W_0 + W_3$, which contradicts the maximality of W_0.

Now, let $\{X_\alpha\}_{\alpha \in A}$ be the set of all simple submodules of V. By

the previous paragraph this set is non-empty. Let $\mathcal{B}$ be the set of all subsets $B \subseteq A$ such that the sum $\sum_{\alpha\in B} X_\alpha$ is direct. It is easy to check the conditions of Zorn lemma to deduce that there is a maximal element $B_0 \in \mathcal{B}$, and we just have to verify that $V = W' := \sum_{\alpha\in B_0} X_\alpha$. Well, otherwise write $V = W' \oplus W''$ and pick an irreducible submodule $X_\beta \subseteq W''$. Then the sum $X_\beta + \sum_{\alpha\in B_0} X_\alpha$ is direct which contradicts the maximality of B_0.

(ii) $\Rightarrow$ (iii) is obvious.

(iii) $\Rightarrow$ (i) Let $W \subseteq V$ be a submodule. Choose $W' \subseteq V$ to be a maximal among submodules $Y \subseteq V$ such that $Y \cap W = 0$. We just need to prove that $V = W + W'$. Well, otherwise there is an element $v \in V \setminus (W + W')$. By assumption (iii), we may write $v = v_1 + \cdots + v_n$, where v_i is an element of a simple submodule $V_i \subseteq V$. As $v \notin W + W'$, we must have $v_i \notin W + W'$ for some v_i. Hence $V_i \cap (W + W') = 0$. Hence $W' \subsetneq W' + V_i$ and $(W' + V_i) \cap W = 0$. This contradicts the maximality of W'. □

3.4 Finiteness conditions

To develop a reasonable theory of modules we often have to assume that they are not too large in such or another sense. There are different ways to make this formal, and these are studied in this section.

Definition 3.4.1 Let V be a left R-module.

(i) We say that V satisfies *A.C.C.* (or V is *noetherian*) if every ascending sequence

$$V_1 \subseteq V_2 \subseteq \cdots \subseteq V_i \subseteq \ldots$$

terminates, i.e. there exists n such that $V_i = V_n$ for all $i \geq n$.

(ii) We say that V satisfies *D.C.C.* (or V is *artinian*) if every descending sequence

$$V_1 \supseteq V_2 \supseteq \cdots \supseteq V_i \supseteq \ldots$$

terminates, i.e. there exists n such that $V_i = V_n$ for all $i \geq n$.

Definition 3.4.2 Let R be a ring. We say that R is *left noetherian* (resp. *left artinian*) if so is the left regular module ${}_RR$.

Remark 3.4.3 (i) One can also define the obvious notions of right noetherian and right artinian. If the rings are commutative, there is

of course no difference, and we just speak of noetherian and artinian commutative rings. In the non-commutative case however, there are examples of rings which are left artinian and left noetherian but not right artinian or right noetherian, see Problem 3.17.1 from §3.17.

(ii) In case of the rings one should not consider the conditions of being artinian and noetherian as somehow opposite to each other. Indeed, it will turn out that every artinian ring is also noetherian (quite amazing, isn't it?).

Lemma 3.4.4 *Let V be a left R-module.*

(i) *V satisfies A.C.C. if and only if every non-empty subset of submodules in V has a maximal element (by inclusion).*
(ii) *V satisfies D.C.C. if and only if every non-empty subset of submodules in V has a minimal element (by inclusion).*

Proof Obvious. □

Lemma 3.4.5 *Let V be a left R-module and $W \subseteq V$ be a submodule. Then*

(i) *V satisfies A.C.C. if and only if W and V/W do.*
(ii) *V satisfies D.C.C. if and only if W and V/W do.*

Proof We prove (i) and leave (ii) as an exercise. If V satisfies *A.C.C.* then it is clear that W does. To see this for V/W use the correspondence theorem. Conversely, assume that both W and V/W satisfy A.C.C. Let $V_1 \subseteq V_2 \subseteq \dots$ be an ascending chain of submodules in V. Then $W \cap V_1 \subseteq W \cap V_2 \subseteq \dots$ is an ascending chain of submodules in W, and $(V_1 + W)/W \subseteq (V_2 + W)/W \subseteq \dots$ is an ascending chain of submodules in V/W. Since both sequences terminate, there exists n such that $V_n \cap W = V_m \cap W$ and $(V_n + W)/W = (V_m + W)/W$ for all $m \geq n$. Then $V_n + W = V_m + W$ for all $m \geq n$. This implies that $V_n = V_m$ for all $m \geq n$: if $v \in V_m$ then $v \in V_m + W = V_n + W$, so $v = x + w$ for $x \in V_n$ and $w \in W$. Now, $w = v - x \in V_m \cap W = V_n \cap W \subseteq V_n$. Thus $v \in V_n$. □

We prove two more facts about noetherian modules.

Proposition 3.4.6 *A left R-module V satisfies A.C.C. if and only if every submodule of V is finitely generated.*

Proof If the module does not satisfy A.C.C. there is an infinite strictly increasing chain of submodules:

$$V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_i \subsetneq \ldots.$$

Let $W = \cup_{i=1}^{\infty} V_i$. This is a submodule of V. Assume $W = \langle v_1, \ldots, v_n \rangle$. Then there is N such that all $v_j \in V_N$, and so $\langle v_1, \ldots, v_n \rangle \subseteq V_N \subsetneq W$, giving a contradiction. The argument is easily reversed. □

Theorem 3.4.7 *Let R be a left noetherian ring and V be a left R-module. Then V is noetherian if and only if it is finitely generated.*

Proof By Proposition 3.4.6 we only have to prove the 'if'-part. Now, if $V = \langle v_1, \ldots, v_n \rangle$, let $V_i = \langle v_1, \ldots, v_i \rangle$. Then for every i, V_i/V_{i-1} is cyclic, so is a quotient of ${}_RR$, and so noetherian by Lemma 3.4.5. Now V is noetherian using induction and Lemma 3.4.5 again. □

3.5 Jordan-Hölder and Krull-Schmidt

Definition 3.5.1 A *composition series* of a left R-module is a finite chain of submodules

$$V = V_0 \supset V_1 \supset V_2 \supset \cdots \supset V_m = 0$$

such that the quotients V_i/V_{i+1} are simple for all $i = 0, 1, \ldots, m-1$. The simple modules V_i/V_{i+1} are called its *composition factors*.

Of course a composition series for a module V does not have to exist (consider ${}_{\mathbb{Z}}\mathbb{Z}$). If it does, we say that V has *finite length*.

Theorem 3.5.2 *A left R-module V has a finite length if and only if it satisfies A.C.C. and D.C.C.*

Proof Let V satisfy A.C.C. and D.C.C. Then by Lemma 3.4.4, in the set of all proper submodules of V there exists a maximal element V_1. Now, either $V_1 = 0$ or in the set of all proper submodules of V_1 there exists a maximal element V_2. As V satisfies D.C.C, this process must stop after finitely many steps, providing us with a composition series.

The converse follows using Lemma 3.4.5 and induction. □

Let

$$V = X_1 \supset X_2 \supset \cdots \supset X_{m+1} = 0,$$
$$V = Y_1 \supset Y_2 \supset \cdots \supset Y_{n+1} = 0$$

be two composition series of V. We say that they are *equivalent* if $m = n$ and there is a permutation $\sigma \in S_n$ such that $Y_i/Y_{i+1} \cong X_{\sigma(i)}/X_{\sigma(i)+1}$ for all $1 \leq i \leq n$.

Theorem 3.5.3 (Jordan-Hölder) *Let V be a left R-module of finite length. Then any two composition series of V are equivalent.*

Proof Chose two composition series:

$$V = X_0 \supset X_1 \supset X_2 \supset \cdots \supset X_m = 0, \tag{3.1}$$
$$V = Y_0 \supset Y_1 \supset Y_2 \supset \cdots \supset Y_n = 0, \tag{3.2}$$

and apply induction on n. If $n = 1$ then V is simple, and the result is clear. Let $n > 1$. If $X_1 = Y_1$, then just apply the inductive assumption. Otherwise we have $X_1 + Y_1 = V$, and

$$\begin{aligned} V/X_1 = (X_1 + Y_1)/X_1 \cong Y_1/(X_1 \cap Y_1), \\ V/Y_1 = (X_1 + Y_1)/Y_1 \cong X_1/(X_1 \cap Y_1). \end{aligned} \tag{3.3}$$

Choose a composition series of $X_1 \cap Y_1$:

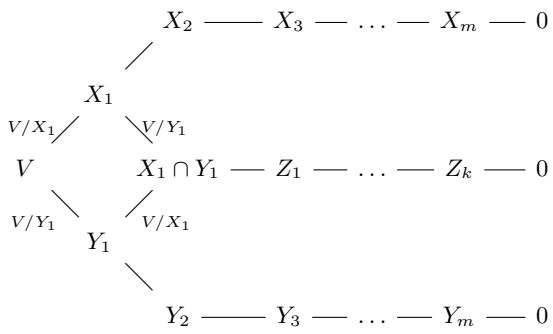
$$X_1 \cap Y_1 = Z_0 \supset Z_1 \supset Z_2 \supset \cdots \supset Z_k = 0.$$

Then we have two different composition series for V:

$$V = X_0 \supset X_1 \supset X_1 \cap Y_1 = Z_0 \supset Z_1 \supset Z_2 \supset \cdots \supset Z_k = 0, \tag{3.4}$$
$$V = Y_0 \supset Y_1 \supset X_1 \cap Y_1 = Z_0 \supset Z_1 \supset Z_2 \supset \cdots \supset Z_k = 0, \tag{3.5}$$

which clearly have the same set of composition factors up to permutation. Now, by induction, (3.1) and (3.4) are equivalent, in particular $k = n - 2$. So we can apply induction to deduce that (3.2) and (3.5) are equivalent. Hence (3.1) and (3.2) are equivalent, as required.

The following picture illustrates the main idea of the proof:



□

There is another way to approach a notion of a building block for an R-module.

Definition 3.5.4 A left R-module V is called *indecmposable* if $V \neq 0$ and it cannot be presentaed as a direct sum of its two non-trivial submodules.

It is clear that irreducible modules are indecomposable, but the converse does not have to be true. For example $\mathbb{Z}/4\mathbb{Z}$ is indecomposable but not irreducible $\mathbb{Z}$-module.

The role of indecomposable module is explained by the fact that any finite length module can be decomposed as a direct sum of finitely many indecomposable modules, and the latter are defined uniquely up to an isomorphism:

Theorem 3.5.5 (Krull-Schmidt) *Let V be a left R-module of finite length. Then V can be decomposed as a finite direct sum of its indecomposable submodules. Moreover, if $V = X_1 \oplus \cdots \oplus X_m$ and $V = Y_1 \oplus \cdots \oplus Y_n$ are two such decompositions, then $m = n$ and there is a permutation $\sigma \in S_n$ such that $X_i \cong Y_{\sigma(i)}$ for all $1 \leq i \leq n$.*

We skip the proof of this theorem (sorry!)

3.6 Free modules

Of all modules, free modules are most like vector spaces. On the other hand, they have universal properties like those enjoyed by free groups.

Definition 3.6.1 Let V be a left R-module and X be a subset of V.

(i) X is *linearly independent* (over R) in case $\sum_{x\in X} r_x x = 0$, with $r_x \in R$ and almost all $r_x = 0$, implies $r_x = 0$ for all x.

(ii) A *basis* of V is a linearly independent subset which generates V.

(iii) We say that V is *free* on X if X is a basis of V. We say that V is *free* if it has a basis.

It is clear that V is free on X if and only if every element $v \in V$ can be written uniquely as a linear combination $v = \sum_{x\in X} r_x x$ where $r_x \in R$ and almost all $r_x = 0$.

The following proposition classifies free R-modules.

Proposition 3.6.2 *Let R be a ring.*

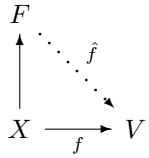
(i) *A left R-module is free if and only if it is isomorphic to a direct sum $\bigoplus_{i\in I} {}_R R$ (for some I) of regular modules.*

(ii) *A left R-module has basis X if and only if it is isomorphic to $\bigoplus_{x\in X} {}_R R$. Conversely, for any set I, the module $\bigoplus_{i\in I} {}_R R$ is free on the set $\{e_i\}_{i\in I}$ where e_i has components $(e_i)_i = 1$ and $(e_i)_j = 0$ for $j \neq i$.*

Proof Of course (i) follows from (ii). It is clear that $\bigoplus_{i\in I} {}_R R$ is free on $\{e_i\}_{i\in I}$. For the converse, let V have basis X. The universal property of a direct sum yiels a map $\theta : \bigoplus_{x\in X} {}_R R \to V$, $(r_x)_{x\in X} \mapsto \sum_{x\in X} r_x x$, which is an isomorphism by the properties of bases. □

Corollary 3.6.3 *For every set X there exists a free left R-module with basis X, and it is unique up to isomorphism.*

Theorem 3.6.4 (Universal property of Free Modules) *Let F be a left R-module free on X. Then for every left R-module V and every map $f : X \to V$, there exists a unique R-module homomorphism $\hat{f} : F \to V$ with $\hat{f}(x) = f(x)$ for all $x \in X$, see the diagram below.*



Proof The homomorphism $\hat{f}$ must preserve linear combinations, and so send $\sum r_x x$ to $r_x f(x)$. Since X is a basis for F, there exists a unique such map $\hat{f}$, which clearly is a module homomorphism. □

Corollary 3.6.5 *A left R-module which generated by a subset X is a homomorphic image of the free left R-module on X.*

Just like for groups, we can use free modules to describe arbitrary modules by generators and relations. We do not pursue this here.

Instead, let us note that, unlike vector spaces, the notion of 'dimension' of a free module is not well-defined in general. For example the regular module ${}_R R$ has basis, which consists of just one element $\{1\}$. On the other hand, it might happen for some rings that ${}_R R$ also has a basis with two elements, see Problem 3.17.5. However, this does not happen if the free module has an infinite basis (use a cardinality argument). Another important case when the 'dimension' happens to be well-defined is that of a commutative ring:

Theorem 3.6.6 *If R is commutative, then all bases of a free R-module have the same number of elements.*

Proof Let F be a free R-module with a basis X. Let I be a maximal ideal of R, and $k := R/I$. Then k is a field. Note that IF consists of all elements of the form $\sum_{x \in X} r_x x$ such that all $r_x \in I$. Consider F/IF as a k-vector space. It is easy to check that $\{x + IF \mid x \in B\}$ is a basis of this vector space. As we know that dimensions of vector spaces are well-defined, the theorem follows. □

Yet another important case where all bases have the same cardinality is when R is a division ring. This is proved in the same manner as for the vector spaces over fields.

When all bases of a free left R-module have the same number of elements, that number is called the *rank* of the module. For example, every free abelian group has a rank.

3.7 Modules over PID's

In this section we recall some basic facts about PID's and then study finitely generated modules over PID's. Throughout the section we assume that R is a PID. Here is a list of facts which I assume you know:

- The definition of a PID.
- Examples of PIDs: Euclidean domains, such as integers $\mathbb{Z}$, polynomials over a field $F[x]$, Gaussian integers $\mathbb{Z}[i]$ (There are PID's which are not euclidean, but examples are rather technical). Examples of domains which are not PID's: $\mathbb{Z}[x]$, $F[x, y]$.
- Any PID is noetherian.
- An element r of a PID is irreducible if and only if it is prime if and only if (r) is a non-zero maximal ideal (a non-zero, non-unit element of a domain is called irreducible if it can not be written as a product of two non-units; a non-zero, non-unit element p of a domain is called prime if $p \mid ab$ implies $p \mid a$ or $p \mid b$ or, equivalently, if (p) is a prime ideal).
- PID is a UFD. This means two things: (a) in a PID every non-zero non-unit is a product of irreducible elements; (b) if two products of irreducible elements $p_1 p_2 \dots p_m = q_1 q_2 \dots q_n$ are equal, then $m = n$, and up to a permutation we have $(p_i) = (q_i)$.
- The notion of the LCM and GCD in PID's. If $d = (a, b)$ is the GCD of a and b (defined up to a unit) then $(a) + (b) = (d)$. If $a = p_1^{a_1} \dots p_n^{a_n}$ and $b = p_1^{b_1} \dots p_n^{b_n}$ are decompositions into irreducible factors, then $(a, b) = p_1^{c_1} \dots p_n^{c_n}$ where $c_i = \min(a_i, b_i)$ for all i (note the equalities are up to a unit).

Lemma 3.7.1 *Let R be a PID, and V be a free R-module with finite basis $\{v_1, \dots, v_n\}$. Every submodule W of F is free and* $\operatorname{rank} W \le n$.

Proof Induction on n, starting from $n = 0$. Let $n > 0$. Every element $w \in W$ can be written uniquely in the form $w = r_1 v_1 + \dots + r_n v_n$ where $r_i \in R$. The set of all coefficients r_1 when w runs over W is an ideal $I \lhd R$. As R is a PID, there exists $s \in R$ such that $I = (s)$. Then $w_1 = s v_1 + \dots + r_n v_n$ for some $w_1 \in W$. Set $V' = R v_2 \oplus \dots \oplus R v_n$. By induction, the submodule $V' \cap W$ has a basis $w_2, \dots, w_m$ for some $m \le n$. Moreover, it is clear that $\{w_1, w_2, \dots, w_m\}$ is a basis of W (except for the case $s = 0$ when the element w_1 is absent). □

For a converse to Lemma 3.7.1, see Problem 3.17.14.

The study of finitely generated R-modules is closely related with the theory of invariant factors of rectangular matrices over R. The matrices appear as follows. Let $\{v_1, \dots, v_n\}$ be a basis of V. For a matrix $X = (x_{ij}) \in M_n(R)$ (the ring of n by n matrices with entries in R), define

the elements $v'_1, \ldots, v'_n \in V$ via

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)X,$$

i.e. $v'_i = \sum_{j=1}^{n} x_{ji} v_j$. It is easy to see that $\{v'_1, \ldots, v'_n\}$ is a basis of V if and only if the matrix X is *unimodular*, i.e. there exists $Y \in M_n(R)$ such that $XY = YX = I_n$, which is equivalent to the fact that $\det X \in R^\times$ ($R^\times$ denotes the units in R).

Now let $W \subseteq V$ be a submodule, and $\{w_1, \ldots, w_m\}$ be a basis of W. We know that $m \leq n$. We can write

$$(w_1, \ldots, w_m) = (v_1, \ldots, v_n)A,$$

for some $A \in M_{n,m}(R)$ ($n \times m$ matrices over R). Assume that we have new bases in V and W:

$$(v'_1, \ldots, v'_n) = (v_1, \ldots, v_n)X, \quad (w'_1, \ldots, w'_m) = (w_1, \ldots, w_m)Y$$

for unimodular matrices $X \in M_n(R)$ and $Y \in M_m(R)$. Then

$$(w'_1, \ldots, w'_n) = (v'_1, \ldots, v'_n)X^{-1}AY.$$

Our first goal is to show that the matrices X and Y can be chosen so that $X^{-1}AY = \operatorname{diag}(r_1, \ldots, r_m)$, the matrix in $M_{n,m}(R)$ with $r_1, \ldots, r_m$ on the main diagonal, and zeros elsewhere. In fact, this can be done in a careful way so that certain division properties hold and then we can even say something about uniqueness of the resulting diagonal matrix, see Theorems 3.7.3 and 3.7.7 below.

Definition 3.7.2 Two matrices A and B in $M_{n,m}(R)$ are called *equivalent*, written $A \sim B$, if there exist unimodular matrices $X_1 \in M_n(R)$ and $X_2 \in M_m(R)$ such that $B = X_1 A X_2$.

Theorem 3.7.3 *Let R be a PID. Every matrix $A \in M_{n,m}(R)$ is equivalent to a matrix of the form* $\operatorname{diag}(\delta_1, \ldots, \delta_k, 0, \ldots, 0)$*, where $\delta_i \in R \setminus \{0\}$, δ_i divides δ_{i+1} for all $1 \leq i < k$.*

Proof We will demonstrate an effective method of reducing any matrix to the required form using elementary transformations with rows and columns. Let X_{ij} be the unimodular matrix obtained from the identity matrix I (of appropriate size) by permuting rows i and j. Note that for $B \in M_{n,m}(R)$, the matrix $X_{ij}B$ is obtained from B by permuting rows i and J, while BX_{kl} is obtained from B by permuting columns k and l. Now for $i \neq j$, let $Y_{ij}(r) = I + rE_{ij}$ be a transvection matrix (of an

appropriate size). Again $Y_{ij}(r)$ is unimodular, and $Y_{ij}(r)B$ is obtained from B by adding the jth row multiplied by r to the ith row, while $BY_{kl}(r)$ is obtained from B by adding the kth column multiplied by r to the lth column.

We assume by induction that the theorem is proved for all $(n-1) \times (m-1)$ matrices over R and that $A \neq 0$. Consider the following cases.

Case 1. Some element a_{ij} of A divides all other elements. Then multiplying A on the right and on the left by the matrices of the form X_{kl}, we will get a matrix $B \sim A$ with b_{11} dividing all other entries. Now we can use multiplication on the left and on the right with matrices of the form $Y_{i1}(r)$ and $Y_{1j}(r)$ to get a new matrix $C \sim A$ of the form

$$\begin{pmatrix} b_{11} & \mathbf{0} \\ \mathbf{0} & C_1 \end{pmatrix}$$

and all elements of C_1 are divisible by b_{11}. By induction, there are unimodular matrices X_1 and Y_1 such that $Y_1C_1X_1$ has the required form. Let

$$X = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & X_1 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Y_1 \end{pmatrix}.$$

Then

$$C \sim YCX = \begin{pmatrix} b_{11} & \mathbf{0} \\ \mathbf{0} & Y_1C_1X_1 \end{pmatrix}.$$

Moreover, it is easy to see that b_{11} divides the first diagonal element of the diagonal matrix $Y_1C_1X_1$. Thus, YCX has the required form.

Case 2. None of the elements a_{ij} of A divides all other elements. Let $a \in R$ be an entry of the matrix A with the minimal number k of irreducible factors. As a is not a unit, $k > 0$. It suffices to show that $A \sim B$ where B contains an element which has less than k irreducible factors. We may assume that $a = a_{11}$. If a_{11} divides all elements of the first row and the first column, then, as in the case 1,

$$A \sim B = \begin{pmatrix} a_{11} & \mathbf{0} \\ \mathbf{0} & B_1 \end{pmatrix},$$

and $a_{11} \not| b$ for some element b of the matrix B_1. Then adding to the first column of B the column of B containing b, we get a matrix equivalent to A in which a_{11} does not divide some element of the first column. So we may assume from the beginning that a_{11} does not divide some element of the first row or column. Assume for definiteness that a_{11} does not divide an element of the first column. We may assume, passing

if necessary to an equivalent matrix, that $a_{11} \not| a_{21}$. Let $d = (a_{11}, a_{21})$. Then $(d) = (a_{11}) + (a_{21})$, so $d = c_1 a_{11} + c_2 a_{21}$ for some $c_1, c_2 \in R$ with $(c_1) + (c_2) = R$. So there are elements $d_1, d_2 \in R$ such that $d_1 c_1 - d_2 c_2 = 1$. Then the matrix

$$X := \begin{pmatrix} c_1 & c_2 & \mathbf{0} \\ d_2 & d_1 & \\ \mathbf{0} & & I \end{pmatrix}$$

is unimodular, and the matrix XA has an element d in the position $(1,1)$. It remains to note that d has less than k irreducible factors. □

The elements δ_i in the theorem above are called invariant factors of the matrix A. It will follow from Theorem 3.7.7 that they are defined uniquely up to units in R.

We make some remarks on torsion in R-modules. Let V be an R-module and $v \in V$. We have $\mathrm{Ann}(v) = (r)$ for some element $r \in R$ defined up to a unit. We say that r is the *order* of v or (r) is the *order ideal* of v. If the order ideal of v is non-zero we say that v is a *periodic* element; otherwise v is *aperiodic*. The module V is called *torsion* module if all of its elements are periodic; V is *torsion-free* if all non-zero elements of V are aperiodic. The following lemma follows easily from the definitions.

Lemma 3.7.4 *Let R be a PID, and V be an R-module. The set T of all periodic elements of V is a torsion submodule of V, and V/T is torsion-free.*

The submodule T consisting of all periodic elements of V is called the *torsion* of V.

Now we make an important step to proving uniqueness of invariant factors.

Lemma 3.7.5 *Let R be a PID, δ be a non-zero non-unit element of R, and $\delta = \pi_1 \dots \pi_n$, where the elements π_i are powers of irreducible factors of δ (up to units). Then we have an isomorphism of R-modules*

$$R/(\delta) \cong R/(\pi_1) \oplus \dots \oplus R/(\pi_n),$$

and the modules $R/(\pi_i)$ are indecomposable.

Proof We first prove that $R/(\pi_i)$ is indecomposable. Let $\pi_i = p^m$ for an irreducible element $p \in R$. Note that submodules $X \subseteq R/(p^n)$

are in one-to-one correspondence with the ideals of R containing (p^n). Hence they form a chain. In particular $R/(p^n)$ has a unique minimal submodule, whence $R/(\pi_i)$ is indecomposable.

Let $\zeta_i = \prod_{j\neq i} \pi_j$, $1 \leq i \leq n$, and denote

$$\bar{\zeta_i} = \zeta_i + (\delta) \in R/(\delta).$$

As $\mathrm{GCD}(\zeta_1, \ldots, \zeta_n) = 1$, there exist elements ξ_i such that $1 = \sum_{i=1}^n \xi_i \zeta_i$. It follows that

$$R/(\delta) = R\bar{\zeta_1} + \cdots + R\bar{\zeta_n}.$$

Moreover, $R\bar{\zeta_i} \cong R/(\pi_i)$, and so it just remains to show that the sum is direct. Assume $\eta_i\bar{\zeta_i} = \sum_{j\neq i} \eta_j\bar{\zeta_j}$. There are $\alpha, \beta \in R$ with $\alpha\pi_i + \beta\zeta_i = 1$. As $\pi_i\bar{\zeta_i} = 0$, we have

$$\eta_i\bar{\zeta_i} = (1 - \alpha\pi_i)\eta_i\bar{\zeta_i} = \beta\zeta_i \sum_{j\neq i} \eta_j\bar{\zeta_j} = 0.$$

□

Lemma 3.7.6 *Let R be a PID, and $\delta_1 \mid \delta_2 \ldots \mid \delta_k$, $\delta_1' \mid \delta_2' \ldots \mid \delta_{k'}'$ be non-zero non-unit elements of R such that*

$$R/(\delta_1) \oplus \cdots \oplus R/(\delta_k) \cong V/W \cong R/(\delta_1') \oplus \cdots \oplus R/(\delta_k').$$

Then $k = k'$ and $\delta_i = \delta_i'$ up to a unit for all $1 \leq i \leq k$.

Proof Decompose each δ_i and δ_i' into the products of powers of distinct irreducible elements:

$$\delta_i = \pi_{i,1} \ldots \pi_{i,j_i}, \quad \delta_i' = \pi_{i,1}' \ldots \pi_{i,j_i'}'.$$

Then by Lemma 3.7.5,

$$\bigoplus_{1\leq i\leq k,\ 1\leq m\leq j_i} R/(\pi_{i,m}) \cong \bigoplus_{1\leq i\leq k',\ 1\leq m\leq j_i'} R/(\pi_{i,m}'),$$

with all the summands being indecomposable. Let $X := \{\pi_{i,m} \mid 1 \leq i \leq k,\ 1 \leq m \leq j_i\}$ and $\{\pi_{i,m}' \mid 1 \leq i \leq k',\ 1 \leq m \leq j_i'\} =: X'$. Then Krull-Schmidt's Theorem implies that $X = X'$ (some elements might appear in both sets several times, and we do have that the corresponding multiplicities are the same).

We claim that the elements δ_i (resp. δ_i') can be recovered from X (resp. X'). Indeed, note that δ_k is the LCM of the elements of X, δ_{k-1} is the LCM of the elements of X, which do not appear in the decomposition

of δ_k, etc. The same algorithm works for X'. As $X = X'$, we are done. □

Theorem 3.7.7 *Let R be a PID, V be a free left R-module of finite rank, and $W \subseteq V$ be a non-zero submodule. Then there exists a basis $\{v_1, \ldots, v_n\}$ of V and non-zero elements $\delta_1, \ldots, \delta_k \in R$ $(1 \leq k \leq n)$ such that $\delta_i \,|\, \delta_{i+1}$ $(1 \leq i < k)$ and $\{\delta_1 v_1, \ldots, \delta_k v_k\}$ is a basis of W. The elements δ_i are called invariant factors of the pair (V, W) and are defined uniquely up to units in R.*

Proof The existence of the basis $\{v_1, \ldots, v_n\}$ and the elements $\delta_1, \ldots, \delta_k$ follows from Theorem 3.7.3. Assume that $\{v'_1, \ldots, v'_n\}$ and $\delta'_1, \ldots, \delta'_k$ also have the desired properties (k is the same as the rank of W is well-defined). Note that

$$R/(\delta_1) \oplus \cdots \oplus R/(\delta_k) \cong V/W \cong R/(\delta'_1) \oplus \cdots \oplus R/(\delta'_k).$$

Now, apply Lemma 3.7.6. □

Theorem 3.7.8 *Let R be a PID, and V be a finitely generated R-module. Then*

$$V \cong ({}_R R)^{\oplus m} \oplus R/(\delta_1) \oplus \cdots \oplus R/(\delta_k)$$

where $\delta_1 \,|\, \delta_2 \,|\, \ldots \,|\, \delta_k$ are non-zero non-unit elements of R. Moreover, the number m is defined uniquely and the elements δ_i are defined uniquely up to units.

Proof Let $V = \sum_{i=1}^n Rv_i$, and let $F = \oplus_{i=1}^n Rf_i$ be a free module with basis $\{f_1, \ldots, f_n\}$. Then there exists a surjective homomorphism $\theta : F \to V$, $\sum_{i=1}^n r_i f_i \mapsto \sum_{i=1}^n r_i v_i$. Let $W = \ker\theta$. Then $V \cong F/W$. If $W = 0$, then V is free. Let $W \neq 0$. By Theorem 3.7.7, there is a basis $\{e_1, \ldots, e_n\}$ of F and non-zero non-unit elements $\delta_1 \,|\, \delta_2 \,|\, \ldots \,|\, \delta_k$ of R such that $W = R\delta_1 e_1 \oplus \cdots \oplus R\delta_k e_k$ for some $k \leq n$. Now it is easy to see that $F/W \cong R/(\delta_1) \oplus \cdots \oplus R/(\delta_k) \oplus ({}_R R)^{\oplus n-k}$.

For uniqueness, note that $T := R/(\delta_1) \oplus \cdots \oplus R/(\delta_k)$ is the torsion of V, and that $V/T \cong ({}_R R)^{\oplus m}$. Now the uniqueness statement follows from Lemma 3.7.6 and the fact that the rank of free R-modules is well-defined. □

We can decompose the torsion part of the module further:

Theorem 3.7.9 *Let R be a PID and V be a finitely generated torsion R-module. Then*

$$V \cong \oplus_{i=1}^{n} R/(\pi_i),$$

where the π_i are prime powers in R. Moreover, the π_i are defined uniquely up to units.

Proof Everything follows from Lemma 3.7.5, Theorem 3.7.8, and Krull-Schmidt. □

Applying the above results to the case $R = \mathbb{Z}$ gives

Theorem 3.7.10 (Fundamental Theorem of Abelian Groups) *Let G be a finitely generated abelian group. Then*

(i)

$$G \cong (C_\infty)^{\times m} \times C_{\delta_1} \times \cdots \times C_{\delta_k}$$

where $\delta_1 \mid \delta_2 \mid \ldots \mid \delta_k$ are positive integers greater than 1. Moreover, the numbers m and the δ_i are defined uniquely.

(ii)

$$G \cong (C_\infty)^{\times m} \times \prod_{i=1}^{n} C_{\pi_i},$$

where the π_i are prime powers. Moreover, m and the π_i are defined uniquely.

3.8 Normal forms of a matrix over a field

In this section we apply the theory developed for modules over PIDs to obtain canonical forms of matrices over a field F. Let V be a finite dimensional vector space over F, and $\varphi : V \to V$ be a non-zero linear operator. We want to find a basis of V in which the matrix of φ has a particularly nice form. The link with the theory of modules comes from the following easy observation: the vector space V is a module over the polynomial ring $F[x]$ via the action $fv := f(\varphi)v$ for $v \in V$, $f \in F[x]$.

As the dimension of the space $\mathrm{End}(V)$ is finite, the elements $\mathrm{id} = \varphi^0, \varphi, \varphi^2, \ldots, \varphi^N$ are linearly dependent for sufficiently large N. This means that there is a non-zero polynomial $g \in F[x]$ such that $g(\varphi) = 0$. So the annihilator of the module V is a non-trivial proper ideal in $F[x]$. This annihilator is the principal ideal generated by some polynomial $m(x)$ defined uniquely up to a unit. Hence the additional requirement

that this generator $m(x)$ is monic defines $m(x)$ uniquely. Clearly, another way to characterize m is to say that this is the monic polynomial of minimal possible degree annihilating φ. Such polynomial is called the *minimal polynomial* of φ.

The discussion above shows that the $F[x]$-module V is torsion. The discussion above shows that the $F[x]$-module V is torsion. By Theorem 3.7.8, we can decompose V as a direct sum

$$V = V_1 \oplus \cdots \oplus V_k$$

where each V_i is isomorphic to $F[x]/(\delta_i)$, where $\delta_1 \mid \delta_2 \mid \ldots \mid \delta_k$ are monic polynomials of positive degree. Moreover, these polynomials are unique.

For any monic polynomial $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0$ define its *companion matrix* $C(f)$ to be the matrix of the left multiplication by x in the vector space $F[x]/(f)$ with respect to the basis $\{1, x, \ldots, x^{d-1}\}$:

$$C(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \vdots & & \\ 0 & 0 & \cdots & 0 & -a_{d-2} \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix}.$$

Note that we have identified the action of φ on V_i with action of x in $F[x]/(\delta_i)$. So we have

Theorem 3.8.1 (The First Canonical Form) *If φ is a linear transformation in a finite dimensional vector space V over an arbitrary field F, then there exists a basis of V with respect to which the matrix of φ has a block diagonal form* $\operatorname{diag}(C(\delta_1), \ldots, C(\delta_k)$*, where $\delta_1 \mid \delta_2 \mid \ldots \mid \delta_k$ are monic polynomials over F of positive degree. Moreover, such presentation is unique.*

Similarly, Theorem 3.7.9 implies

Theorem 3.8.2 (The Second Canonical Form) *If φ is a linear transformation in a finite dimensional vector space V over an arbitrary field F, then there exists a basis of V with respect to which the matrix of φ has a block diagonal form* $\operatorname{diag}(C(p_1^{a_1}), \ldots, C(p_m^{a_m}))$*, where $p_1^{a_1}, \ldots, p_m^{a_m}$ are positive powers of irreducible polynomials over F. Such presentation is unique up to a permutation of blocks.*

If $p \in F[x]$ is of degree d, consider the following basis of $F[x]/(p^a)$:

$$\{p^{a-1}, xp^{a-1}, \ldots, x^{d-1}p^{a-1}, p^{a-2}, xp^{a-2}, \ldots, x^{d-1}p^{a-2}, \ldots, \\ 1, x, \ldots, x^{d-1}\}.$$

In this matrix the multiplication by x has the following matrix

$$J(p,a) := \begin{pmatrix} C(p) & M & 0 & \ldots & 0 & 0 \\ 0 & C(p) & M & \ldots & 0 & 0 \\ & & & \vdots & & \\ 0 & 0 & 0 & \ldots & C(p) & M \\ 0 & 0 & 0 & \ldots & 0 & C(p) \end{pmatrix},$$

where M is the $m \times m$ matrix with 1 in the $(1,m)$ position and zeros elsewhere. We will call this matrix a *generalized Jordan block*.

Theorem 3.8.3 (Generalized Jordan Normal Form) *If φ is a linear transformation in a finite dimensional vector space V over an arbitrary field F, then there exists a basis of V with respect to which the matrix of φ has a block diagonal form* $\operatorname{diag}(J(p_1,a_1), \ldots, J(p_m,a_m))$, *where $p_1, \ldots, p_m$ are irreducible polynomials over F, and $a_1, \ldots, a_m$ are positive integers. Such presentation is unique up to a permutation of generalized Jordan blocks.*

Of course, if the field F is algebraically closed then the only irreducible polynomials are linear, and so the theorem above boils down to the usual Jordan normal form result.

Finally we are interested in an algorithm for finding the canonical forms of linear operators. Let A be the matrix of our linear operator φ in some basis. The matrix $A - xI$ can be considered as a matrix over the polynomial ring $F[x]$. The proof of Theorem 3.7.3 provides us with the algorithm for finding invariant factors $\delta_1, \ldots, \delta_k$ and hence elementary divisors $p_1^{a_1}, \ldots, p_m^{a_m}$ of $A - xI$. It is not hard to see that these are the ones appearing in the canonical forms of φ.

3.9 Algebras and Modules over Algebras

Many of our favorite rings are actually more than just rings. For example the ring $F[x]$ of polynomials over a field F has the additional structure of a vector space over F. This leads us to the idea of an algebra.

Definition 3.9.1 Let R be a commutative ring. An *R-algebra* (or an *associative R-algebra* or an *associative algebra over R*) is a ring A (as usual with identity) with an additional structure of an R-module such that $a(\alpha b) = (\alpha a)b = \alpha(ab)$ for all $a, b \in A$, $\alpha \in R$.

Remark 3.9.2 Every ring is automatically a $\mathbb{Z}$-algebra. The most 'popular' algebras are algebras over fields. These are rings with additional structure of a vector space over a field F such that the natural axiom of Definition 3.9.1 is satisfied.

Remark 3.9.3 In this course we will only consider associative algebras, i.e. those whose multiplication is associative, and so we will just call them algebras. However, not only associative algebras are interesting. A very important class of non-associative algebras is that of Lie algebras. A typical example of a Lie algebra is the vector space of all $n \times n$ matrices over F with multiplication given by $[A, B] = AB - BA$.

Note that the field F is itself an F-algebra, and every F-algebra has F as a subalgebra: consider the set of all scalar multiples of $1 \in A$. Note that this subalgebra is central.

If A is an F-algebra, we can speak of A-modules because A is a ring to start with. Note, however, that any A-module is now naturally an F-vector space, because the subalgebra $F \subseteq A$ acts on it. This allows us to speak of the dimension of an A-module for example.

3.10 Endomorphism Ring of a Module

Recall that if V is a left R-module then $\mathrm{End}_R(V)$ has a structure of a ring: if $\theta, \eta \in \mathrm{End}_R(V)$ then $(\theta \pm \eta)(v) = \theta(v) \pm \eta(v)$ and $\theta\eta = \theta \circ \eta$. If R is an F-algebra then $\mathrm{End}_R(V)$ is also an F-algebra.

The following is an easy exercise:

Lemma 3.10.1 *There is an isomorphism of rings* $\mathrm{End}_R({}_RR)^{\mathrm{op}} \tilde{\to} R$ *which assigns to each $x \in R$ the endomorphism $r \mapsto rx$. If R is an F-algebra then the map above is an isomorphism of algebras.*

Lemma 3.10.2 (Schur's Lemma) *Let V and W be simple left R-modules.*

(i) *A module homomorphism from V to W is either* 0 *or isomorphism. In particular,* $\mathrm{End}_R(V)$ *is a division ring.*

(ii) *If R is an F-algebra and F is algebraically closed then* $\mathrm{End}_R(V) = \{c\,\mathrm{id}_V \mid c \in F\} \cong F$.

Proof (i) If $\theta \in \mathrm{Hom}_R(V,W) \setminus \{0\}$ then $\ker\theta \neq V$ and $\mathrm{im}\,\theta \neq 0$. Hence $\ker\theta = 0$ and $\mathrm{im}\,\theta = W$ by simplicity of V and W.

(ii) Now, let R be an F-algebra and $F = \bar{F}$. Then any $\theta \in \mathrm{End}_R(V)$ is a linear operator on V. Let $c \in F$ be its eigenvalue. Then $\theta - c\,\mathrm{id}_V \in \mathrm{End}_R(V)$, and $\theta - c\,\mathrm{id}_V$ has a non-trivial kernel (the c-eigenspace of θ). It follows from (i) that $\theta - c\,\mathrm{id}_V = 0$, i.e. $\theta = c\,\mathrm{id}_V$. □

Let V be a left R-module and $E := \mathrm{End}_R(V)^{\mathrm{op}}$. Then V becomes a right E-module via evaluation: $v\theta = \theta(v)$ for $v \in V$, $\theta \in E$. Now we want to address the natural question: what do we know about $\mathrm{End}_E(V)$? Of course we always have a ring homomorphism

$$R \to \mathrm{End}_E(V), r \mapsto \theta_r, \qquad \text{where } \theta_r(v) = rv \quad (v \in V). \tag{3.6}$$

If R is an algebra the map (3.6) is a homomorphism of algebras.

In the special case where $V = {}_RR$ Lemma 3.10.1 implies

Lemma 3.10.3 *If $V = {}_RR$ then the map (3.6) is an isomorhism.*

For a semisimple V the image of the map (3.6) is "large" or *"dense"* in $\mathrm{End}_E(V)$ in the sense of Theorem 3.10.6 below. To prove Jacobson Density Theorem we need to do some preliminary work. Let V be a left R-module. We denote by $V^{\oplus n}$ an (outer) direct sum of n copies of V. If $f : V \to V$ is a map we denote via $f^{\oplus n}$ the map $V^{\oplus n} \to V^{\oplus n}$, which maps $(v_1, \ldots, v_n)$ to $(f(v_1), \ldots, f(v_n))$.

For any ring E, let $M_n(E)$ be the ring of n by n matrices over E. Then, if V is a right E-module, $V^{\oplus n}$ is a right $M_n(E)$-module in a natural way: if $A = (a_{ij})_{1 \leq i,j \leq n} \in M_n(E)$, then

$$(v_1, \ldots, v_n) \cdot A = (w_1, \ldots, w_n),$$

where $w_i = \sum_{j=1}^n v_i a_{ij}$ for all $i = 1, 2, \ldots, n$.

Lemma 3.10.4 *Let E be a ring, $D = M_n(E)$, and V be a right E-module. Then*

$$\alpha : \mathrm{End}_E(V) \to \mathrm{End}_D(V^{\oplus n}) : \ f \mapsto f^{\oplus n}$$

is a ring isomorphism.

Proof First of all, if $f \in \mathrm{End}_E(V)$, it is clear that $f^{\oplus n} \in \mathrm{End}_D(V^{\oplus n})$. Now it is also easy to see that α is an injective ring homomorphism. To see that α is surjective, let $g \in \mathrm{End}_D(V^{\oplus n})$. Denote by $\iota_i : V \to V^{\oplus n}$ the ith natural injection and by $\pi_j : V^{\oplus n} \to V$ the jth natural projection. Then

$$g_{ij} := \pi_j \circ g \circ \iota_i : V \to V \in \mathrm{End}_{\mathbb{Z}}(V).$$

Denote $e_{jk} \in D$ the (j,k)-matrix unit. Then e_{jk} acts on $V^{\oplus n}$ as $\iota_j \circ \pi_k$. So

$$g((v_1, \dots, v_n)\zeta_{jk}) = g((v_1, \dots, v_n))\zeta_{jk},$$

implies $g_{jk} = 0$ for $j \neq k$. Moreover, for every element $\sigma \in S_n$ we have an element $\rho_\sigma \in D$, which maps $(v_1, \dots, v_n)$ to $(v_{\sigma(1)}, \dots, v_{\sigma(n)})$. The fact that g commutes with ρ_σ now implies that $g_{jj} = g_{kk}$ for all j, k. Thus $g = f^{\oplus n}$, where $f = g_{ii}$. To see that $f \in \mathrm{End}_E(V)$, pick any $\eta \in E$ and use the fact that $f^{\oplus n} = g$ commutes with $\mathrm{diag}(\eta, \dots, \eta) \in D$. □

Corollary 3.10.5 *Let V be a left R-module, $E = \mathrm{End}_R(V)^{\mathrm{op}}$, and $D = \mathrm{End}_R(V^{\oplus n})^{\mathrm{op}}$. Then $f \mapsto f^{\oplus n}$ is a ring isomorphism*

$$\mathrm{End}_E(V) \tilde{\to} \mathrm{End}_D(V^{\oplus n}).$$

Proof We can identify D with $M_n(E)$ using Corollary 3.2.4. So everything follows from Lemma 3.10.4. □

Theorem 3.10.6 (Jacobson Density Theorem) *Let V be a semisimple left R-module and $E = \mathrm{End}_R(V)^{\mathrm{op}}$. For every $f \in End_E(V)$ and $v_1, \dots, v_n \in V$ there exists $r \in R$ such that $f(v_i) = rv_i$ for all i.*

Proof Let $f \in \mathrm{End}_E(V)$ and $v \in V$. As V is semisimple, we have $V = Rv \oplus W$ for some submodule W. The projection $\pi : V \to Rv$ along W is an element of E. So $f(v) = f(v\pi) = f(v)\pi$, whence $f(v) \in Rv$. This proves the theorem in the case $n = 1$.

In general, let $D = \mathrm{End}_R(V^{\oplus n})^{\mathrm{op}}$. Then $V^{\oplus n}$ is a semisimple left R-module. By Corollary 3.10.5, $f^{\oplus n} \in \mathrm{End}_D(V^{\oplus n})$. By the case $n = 1$ for each $(v_1, \dots, v_n) \in V^{\oplus n}$ there exists $r \in R$ such that $f^{\oplus n}((v_1, \dots, v_n)) = r(v_1, \dots, v_n)$. □

Example 3.10.7 (i) Assume that R is an F-algebra, F is algebraically closed, and V is a finite dimensional simple R-module. By Schur's

Lemma, $E = F$, and the Density Theorem implies that for any $f \in \mathrm{End}_F(V)$ there is $r \in R$ with $f(v) = rv$ for all $v \in V$.

(ii) Assume that R is a division ring and V is a finite dimensional R-vector space with basis $\{v_1, \ldots, v_n\}$. Then V is semisimple R-module and $E \cong M_n(R^{\mathrm{op}}) \cong M_n(R)^{\mathrm{op}}$. Now the Density Theorem implies that the center of $M_n(R)$ consists of the scalar matrices $\{cI_n \mid c \in Z(R)\}$. Of course, it is easy to see this directly.

3.11 The Wedderburn-Artin Theorem

The main result of this section is a fundamental structure theorem for semisimple rings. The class of semisimple rings is very important: it includes for example group algebras of finite groups over fields of characteristic 0.

Definition 3.11.1 A ring R is called *left semisimple* if every left R-module is semisimple.

Remark 3.11.2 More standard terminology here is *semisimple artinian*. To gain right to use this terminology we will need to prove that any left semisimple ring is right semisimple (i.e. all right R-modules are semisimple) and also is both left and right artinian. Once these facts are established we will switch to the standard terminology.

We will have lots of equivalent reformulations of the notion of left semisimple. The following lemma provides us with the first two. Recall that an element $e \in R$ is called an *idempotent* if $e^2 = e$.

Lemma 3.11.3 *For a ring R the following properties are equivalent:*

(i) *R is left semisimple.*
(ii) *The left regular module ${}_RR$ is semisimple.*
(iii) *Every left ideal of R is generated by an idempotent.*

Proof Obviously (i) implies (ii). Moreover (ii) implies (i) as every module is a quotient of a free module, and a quotient of a semisimple module is semisimple.

(ii) $\Rightarrow$ (iii) Let I be a left ideal of R. By assumption, $R = I \oplus J$ for some left ideal J. Then $1 = e + f$ for $e \in I$ and $f \in J$. We claim that e is an idempotent and $I = Re$. Clearly, $Re \subseteq I$. Conversely, if $x \in I$,

then $x = x \cdot 1 = xe + xf$, which implies $x = xe$ (and $xf = 0$). It follows that e is an idempotent and $I \subseteq Re$.

(iii) $\Rightarrow$ (ii) Let $I = Re$ be a left ideal. We claim that ${}_RR = I \oplus J$ where $J = R(1-e)$. Indeed, each $r \in R$ can be written as $r = re + r(1-e)$, whence ${}_RR = I + J$. To see that the sum is direct, let $r \in Re \cap R(1-e)$. Then $r = xe = y(1-e)$ implies $r = xe = xee = y(1-e)e = 0$. □

Note that a left ideal of a ring R is minimal (non-zero) if and only if it is simple as a left R-module. Thus a ring is left semisimple if and only if it is a direct sum of minimal left ideals.

The following result gives an important family of left-semisimple rings:

Theorem 3.11.4 (Maschke's Theorem) *Let G be a finite group and F be a field of characteristic $p \geq 0$. Then the group algebra FG is semisimple if and only if p does not divide $|G|$.*

Proof If $p \mid |G|$ it follows from Problem 3.17.7 that FG is not semisimple. Now, let $p \nmid |G|$, and $W \subseteq V$ be left FG-modules. We need to show that there is a submodule $X \subseteq V$ with $V = W \oplus X$.

Let Y be an F-subspace of V with $V = W \oplus Y$. The projection $\varphi : V \to W$ is a linear transformation. We define a map $\varphi : V \to V$ by the formula

$$\varphi(v) = \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(gv) \qquad (v \in V).$$

It is clear that φ is linear. We claim that φ is actually an FG-module homomorphism. To check this fact it suffices to verify that $\varphi(hv) = h\varphi(v)$ for all $h \in G, v \in V$. Well, we have

$$\begin{aligned}
\varphi(hv) &= \frac{1}{|G|} \sum_{g \in G} g^{-1} \pi(ghv) \\
&= \frac{1}{|G|} \sum_{g \in G} hh^{-1} g^{-1} \pi(ghv) \\
&= \frac{1}{|G|} h \sum_{g \in G} h^{-1} g^{-1} \pi(ghv) \\
&= \frac{1}{|G|} h \sum_{g \in G} g^{-1} \pi(gv) = h\varphi(v).
\end{aligned}$$

As W is a submodule we have $\operatorname{im} \varphi \subseteq W$. We claim that actually $\operatorname{im} \varphi = W$. This follows from the fact that $\varphi(w) = w$ for any $w \in W$.

Indeed, note that $\pi(w') = w'$ for any $w' \in W$, so

$$\begin{aligned}\varphi(w) &= \frac{1}{|G|}\sum_{g\in G} g^{-1}\pi(gw)\\ &= \frac{1}{|G|}\sum_{g\in G} g^{-1}gw\\ &= \frac{1}{|G|}\sum_{g\in G} w = w.\end{aligned}$$

By a standard argument we now have a module decomposition $V = W \oplus \ker\varphi$. □

Proposition 3.11.5 *Let R be a left semisimple ring.*

(i) *Let V be a simple left R-module and I be a minimal left ideal of R. If $V \ncong I$ then $IV = 0$.*

(ii) *Every simple left R-module is isomorphic to a minimal left ideal of R.*

Proof (i) Assume $IV \neq 0$. Then $Iv \neq 0$ for some $v \in V$. As V is simple, $V = Iv$. Thus we have a non-zero homomorphism $x \mapsto xv$ from I to V. By Schur's Lemma, it is an isomorphism.

(ii) Let V be a simple R-module. As R is a sum of minimal left ideals, there is a minimal left ideal I of R with $IV \neq 0$. Now $V \cong I$ by (i). □

We now produce examples of left semisimple rings, which will turn out to be 'general enough'...

Proposition 3.11.6 *Let D be a division ring. Then $R = M_n(D)$ is a direct sum of n minimal left ideals. All simple left R-modules are faithful, isomorphic to each other, and have dimension n over D.*

Proof Let I_j be the set of all matrices in R, in which all entries are zero outside of jth column. It is readily checked that I_j has dimension n over D, is a minimal left ideal of R, and is faithful when considered as a left R-module. As $R = I_1 \oplus \cdots \oplus I_n$, R is left semisimple. It remains to observe that $I_j \cong I_k$ for any j, k and use Proposition 3.11.5. □

The following result follows from the definition of left semisimple rings.

Proposition 3.11.7 *If rings $R_1, \ldots, R_m$ are semisimple, then so is $R_1 \oplus \cdots \oplus R_m$.*

It follows from the previous two propositions that any ring of the form $M_{n_1}(D_1)\oplus\cdots\oplus M_{n_m}(D_m)$ is semisimple. The fundamental Wedderburn-Artin Theorem which we will prove by the end of this section claims that all left semisimple rings look like this.

Lemma 3.11.8 *Let $R = R_1\oplus\cdots\oplus R_m$ be a direct sum of rings. The left (resp. right, resp. two-sided) ideals of R_i coincide with the left (resp. right, resp. two-sided) ideals of R contained in R_i. The minimal left (resp. right, resp. two-sided) ideals of R_i coincide with the minimal left (resp. right, resp. two-sided) ideals of R contained in R_i.*

Proof A left ideal of R which is contained in R_i is obviously a left ideal of R_i. Conversely, a left ideal of R_i is a left ideal of R. The lemma for the left ideals follows. For other ideals the argument is similar. □

Definition 3.11.9 A ring R is *simple* if R has no two-sided ideals but 0 and R.

Remark 3.11.10 In spite of what the terminology might suggest it is not true that simple ring is necessarily left semisimple, although this is true under the assumption that the ring is left artinian.

Let R be an arbitrary ring. We will say that two left ideals I, J of R are *isomorphic* if they are isomorphic as left R-modules.

Theorem 3.11.11 *Let R be a left semisimple ring and $(C_a)_{a\in A}$ be the family of isomorphy classes of minimal left ideals of R. For each class C_a let R_a be the sum of all minimal left ideals $I \in C_a$.*

(i) *A is finite. In particular, there are only finitely many non-isomorphic simple R-modules.*
(ii) *Any two-sided ideal of R equals $\oplus_{b\in B}R_b$ for some $B \subseteq A$. In particular, each R_a is a minimal two-sided ideal of R.*
(iii) *R is isomorphic to the direct sum of finitely many rings R_a.*
(iv) *Each R_a is a simple ring and a left semisimple ring. Moreover, all simple R_a-modules are isomorphic.*

Proof Since R is left semisimple, we have ${}_RR = \sum_{a\in A} R_a$, see Theorem 3.3.5. Each R_a is a left ideal of R. By Proposition 3.11.5,

$$R_aR_b = 0 \quad \text{whenever} \quad a \neq b. \tag{3.7}$$

Hence $R_aR = R_a \sum_{a\in A} R_a = R_aR_a \subseteq R_a$, and R_a is a two-sided ideal of R.

Now, let J be any non-zero two-sided ideal of R. Then J is a semisimple left R-module and contains a simple submodule, that is a minimal left ideal I of R. Say, $I \in C_a$, and let $I' \in C_a$. Since I is a direct summand of R, composing the projection $R \to I$, the isomorphism $I \xrightarrow{\sim} I'$, and the embedding $I' \to R$ yields an endomorphism η of ${}_RR$. By Lemma 3.10.3, there exists $r \in R$ such that $I' = \eta(I) = Ir \subseteq J$. It follows that $R_a \subseteq J$. Now, let J' be the sum of all R_a contained in J. As R is left semisimple, we can write $J = J' \oplus J''$ where J'' is a left ideal of R. If $J'' \neq 0$ then it contains a minimal left ideal I', and as above $R_b \subseteq J$ if $I' \in C_b$. So $I' \subseteq R_b \subseteq J'$ and so $J' \cap J'' = 0$ leads to a contradiction. We have proved that any two-sided ideal J looks like a direct sum of some R_a. This gives (ii).

Since $R = \sum_{a\in A} R_a$, we have $1 = \sum_{a\in A} e_a$, where $e_a \in R_a$ and $e_a = 0$ for almost all a. If $r \in R_a$, then $r = r \cdot 1 = \sum_{b\in A} re_b = re_a$, in view of (3.7). Similarly $r = e_ar$. Thus e_a is the identity of R_a, and R_a is a ring. Moreover, since $C_a \neq \varnothing$, we have $R_a \neq 0$. Therefore $e_a \neq 0$ for all $a \in A$ proving (i).

We have $R_a \cap (\sum_{b\neq a} R_b) = 0$, as $r \in R_a$ implies $r = e_ar$ and $x \in \sum_{b\neq a} R_b$ implies $e_ar = 0$. Moreover, using (3.7) we deduce that

$$(\sum_{a\in A} r_a)(\sum_{b\in A} s_b) = \sum_{a\in A} r_as_a \qquad (r_a, s_a \in R_a).$$

This proves (iii).

By Lemma 3.11.8 and (ii), R_a is a simple ring. Also, the minimal left ideals of R_a coincide with the minimal left ideals of R contained in R_a. So R_a is a sum of minimal left ideals of R_a, hence left semisimple. Moreover, the minimal left ideals of R_a are all in C_a, hence they are all isomorphic as R-modules, and so they are also all isomorphic as R_a-modules. □

Corollary 3.11.12 *A left semisimple ring R is simple if and only if all simple left R-modules are isomorphic. In particlular, $M_n(D)$ is simple and left semisimple for every division ring D.*

Theorem 3.11.13 *A ring is both simple and left semisimple if and only if it is isomorphic to a ring of $n \times n$ matrices over a division ring.*

Proof The 'if'-part is contained in Corollary 3.11.12. Conversely, let R be simple and left semisimple. Then R is a direct sum $R = \oplus_{a\in A} I_a$ of minimal left ideals. Now $1 = \sum_{a\in A} e_a$, where $e_a \in I_a$ and almost all $e_a = 0$. If $r \in I_b$, then $r = r \cdot 1 = \sum_{a\in A} re_a \in \oplus_{a\in A} I_a$ implies $r = re_b$. Therefore $e_b \neq 0$. It follows that A is finite. By Corollary 3.11.12, ${}_R R \cong V^{\oplus m}$ for some irreducible R-module V and some $m > 0$.

By Schur's Lemma, $D := \mathrm{End}_R(V)$ is a division ring. Moreover, it is easy to see that $\mathrm{End}_R(V^{\oplus m}) \cong M_m(D)$. Finally,

$$R \cong \mathrm{End}_R({}_R R)^{\mathrm{op}} \cong \mathrm{End}_R(V^{\oplus m}) \cong M_m(D),$$

whence $R \cong M_m(D)^{\mathrm{op}} \cong M_m(D^{\mathrm{op}})$. □

Theorems 3.11.11 and 3.11.13 now yield:

Theorem 3.11.14 (Wedderburn-Artin) *A ring is left semisimple if and only if it is isomorphic to a direct sum $M_{n_1}(D_1) \oplus \cdots \oplus M_{n_m}(D_m)$ of finitely many finite dimensional matrix rings over division rings.*

Corollary 3.11.15

(i) *A ring is left semisimple if and only if it is right semisimple.*
(ii) *A left semisimple ring is left and right artinian.*

Proof (i) By definition R is right semisimple if and only if R^{op} is left semisimple. Now observe that $M_n(D)^{\mathrm{op}} \cong M_n(D^{\mathrm{op}})$, using the transpose map. So the Wedderbur-Artin theorem implies that R is left semisimple if and only if R^{op} is left semisimple if and only if R is right semismiple.

(ii) By Proposition 3.11.6 and Wedderburn-Artin, a left semisimple ring R is a direct sum of finitely many minimal left ideals, so it is left artinian for example by Lemma 3.4.5. Now using (i) we see that R is also right artinian. □

As agreed in the beginning in the section, from now on we will speak of *semisimple artinian* rings instead of *left semisimple* rings.

Remark 3.11.16 Let R be a semisimple artinian ring which is also an algebra. Then we know that the division rings $D_1, \ldots, D_m$ appearing in the Wedderburn-Artin theorem are actually division algebras over F: just recall from the proof that the D_i's arised as endomorphism rings of simple R-modules. If F is algebraically closed then Schur's Lemma implies that every D_i is just F, and we get the following version of

the Wedderburn Artin theorem for algebras: every semisimple artinian algebra R over an algebraically closed field F is a direct sum of finite dimensional matrix algebras over F; in particular, R is finite dimensional; moreover, R is simple if and only if it is of the form $M_n(F)$. If F is not algebraically closed then of course it is not true that the $D_i's$ have to equal F. However, if R is a finite dimensional F-algebra, the D_i's are finite dimensional division algebras over F. For example, in this case, if F is a finite field, then the D_i's will have to be finite fields too, thanks to Wedderburn's Theorem 2.11.9.

We finish this section with a uniqueness statement for Wedderburn-Artin theorem.

Proposition 3.11.17 *Let*

$$M_{n_1}(D_1) \oplus \cdots \oplus M_{n_m}(D_m) \cong M_{n'_1}(D'_1) \oplus \cdots \oplus M_{n'_m}(D'_m).$$

Then $m = m'$, and up to a permutation,

$$n_1 = n'_1, D_1 \cong D'_1, \ \ldots, \ n_m = n'_m, D_m \cong D'_m.$$

The proposition follows from the following two lemmas.

Lemma 3.11.18 *Let $R = \oplus_{i=1}^m J_i$ be a semisimple artinian ring written as an inner direct direct sum of simple subrings J_i. Then every two-sided ideal of R is of the form $J_{i_1} \oplus \cdots \oplus J_{i_k}$ for some $1 \leq i_1 < \cdots < i_k \leq m$. In particular, $J_1, \ldots, J_m$ can be characterized as the minimal non-zero two-sided ideals of R.*

Proof Follows from Theorem 3.11.11. □

Proposition 3.11.19 *Let C and D be division rings. If $M_m(C) \cong M_n(D)$ then $m = n$ and $C \cong D$.*

Proof Let $V = C^m$ be the natural $M_m(C)$-module and $W = D^n$ be the natural $M_n(D)$-module. We know that R has only one simple module up to an isomorphism, see Proposition 3.11.6. As isomorphic modules have isomorphic endomorphism rings, we have $C \cong \mathrm{End}_R(V)^{\mathrm{op}} \cong \mathrm{End}_R(W)^{\mathrm{op}} \cong D$.

It remains to prove that $M_m(C) \cong M_n(C)$ implies $m = n$. For that observe that the right action of $C = \mathrm{End}_R(V)^{\mathrm{op}}$ on an irreducible R-module $V = C^m$ comes from the left action of C on C^m. So the R-

isomorphism between the irreducible modules C^m and C^n should be an isomorphism of C-vector spaces, which implies $m = n$. □

3.12 The Jacobson Radical

In the previous section we studied (left) semisimple rings, that is the rings over which every (left) module is semisimple. Of course, not every ring is semisimple and we now set to study Jacobson's radical of a ring which in some sense 'measures' how far a ring is from being semisimple.

Definition 3.12.1 The *Jacobson radical* $J(R)$ of a ring R is the intersection of all maximal left ideals of R.

Remark 3.12.2 Strictly speaking we should have called $J(R)$ the *left* Jacobson radical. However it will turn out later that $J(R)$ is also the intersection of the maximal right ideals, and so there is no need in this 'one-sided' terminology.

The following result is one of many equivalent descriptions of Jacobson radical. Informally speaking, it says that $J(R)$ is the part of the ring which 'does not see' simple (and hence semisimple) modules.

Proposition 3.12.3 *$J(R)$ is the intersection of the annihilators of all simple left R-modules. Hence $J(R)$ is a two-sided ideal.*

Proof Let I be a maximal left ideal of R. Then $V = R/I$ is a simple left R-module, and $\mathrm{Ann}(V) \subseteq I$. Hence the intersection of all $\mathrm{Ann}(V)$ is contained in $J(R)$. Conversely, let $r \in J(R)$ and V be a simple left R-module. Take any $v \in V \setminus \{0\}$. We need to prove that $rv = 0$. We have $Rv = V$, hence ${}_RR/\mathrm{Ann}(v) \cong Rv$ is simple, and so $\mathrm{Ann}(v)$ is a maximal left ideal of R. Therefore $r \in \mathrm{Ann}(v)$. □

Here is another description of $J(R)$:

Proposition 3.12.4 *$x \in J(R)$ if and only if $1 + rx$ has a left inverse for every $r \in R$.*

Proof If $x \in J(R)$ and $r \in R$, then rx belongs to every maximal left ideal of R and $1+rx$ belongs to no maximal left ideal of R, as otherwise 1 would belong to a maximal left ideal. It follows that $R(1 + rx) = R$ for otherwise we would find a maximal left ideal containing the left ideal

$R(1+rx)$ using Zorn's lemma, and $1+rx$ would lie in this maximal ideal.

Conversely, if $x \notin J(R)$, then $x \notin I$ for some maximal left ideal I, whence $I + Rx = R$, and $1 = y + rx$ for some $y \in I, r \in R$. Now, $1 - rx$ does not have a left inverse. □

And one more description:

Proposition 3.12.5 *$J(R)$ is the largest two-sided ideal J of R such that $1+x$ is a unit in R for every $x \in J$.*

Proof Let $x \in J(R)$. Then $1+x$ has a left inverse y by Proposition 3.12.4. Then $y = 1 - yx$ has a left inverse z. But y already has a right inverse $1 + x$. Hence $z = (1 + x)$ has a right inverse y.

Conversely, let J be a two-sided ideal of R such that $1 + x$ is a unit for every $x \in J$. If $x \in J$, then $1 + rx$ is a unit for every $r \in R$, and $x \in J(R)$ by Proposition 3.12.4. Thus $J \subseteq J(R)$. □

As the previous proposition describes $J(R)$ in a 'side-independent' manner, we have the following pleasant fact, which means that all the left-handed descriptions of $J(R)$ give the same object as the right-handed ones.

Corollary 3.12.6 $J(R) = J(R^{\mathrm{op}})$.

Proposition 3.12.7 $J(R/J(R)) = 0$.

Proof Follows from the definition of Jacobson radical and the correspondence theorem for ideals. □

Definition 3.12.8 An element $r \in R$ is called *nilpotent* if $r^n = 0$ for some positive integer n. A left, right, or two-sided ideal $I \in R$ is called *nilpotent* if there exists a positive integer n such that $I^n = 0$, i.e. $x_1 \dots x_n = 0$ for all $x_1, \dots, x_n \in I$.

Proposition 3.12.9 *$J(R)$ contains all nilpotent ideals of R.*

Proof In view of Corollary 3.12.6, it suffices to prove the result for left ideals. Let N be a nilpotent left ideal of R, say $N^m = 0$. We just need to show that $NV = 0$ for any simple R-module V. Well, if $NV \neq 0$ then $V = NV = N^2V = \cdots = N^mV = 0$, giving a contradiction. □

Corollary 3.12.10 *If R is commutative then $J(R)$ contains all nilpotent elements of R.*

Proof Let $r \in R$ be nilpotent. Then the principal ideal (r) is nilpotent, and so $r \in (r) \subseteq J(R)$ by Proposition 3.12.9. □

Lemma 3.12.11 (Nakayama's Lemma) *Let V be a finitely generated left R-module. If W is a submodule of V and $W + J(R)V = V$, then $W = V$.*

Proof First consider the case $W = 0$.Assume that $J(R)V = V$. Since V is finitely generated, it has a minimal generated subset X, which is finite. Then $V = \sum_{x \in X} Rx$ implies $V = J(R)V = \sum_{x \in X} J(R)x$. So, if $y \in X$, then

$$y = \sum_{x \in X} r_x x$$

for some $r_x \in J(R)$, or

$$(1 - r_y)y = \sum_{x \in X,\ x \neq y} r_x x.$$

By Proposition 3.12.5, $(1 - r_y)$ is invertible, so we have

$$y = \sum_{x \in X,\ x \neq y} (1 - r_y)^{-1} r_x x.$$

Therefore V is generated by $X \setminus \{y\}$. This contradiction shows that $X = \varnothing$, i.e. $V = 0$.

In the general case apply what has just been proved to the module V/W. □

3.13 Artinian Rings

Artinian rings form an important class of rings containing for example all finite dimensional algebras.

Lemma 3.13.1 *If R is left artinian then $J(R)$ is nilpotent. In particular, $J(R)$ is the largest nilpotent two-sided ideal of R.*

Proof Let $J := J(R)$. The sequence $J \supseteq J^2 \supseteq J^3 \supseteq \ldots$ stabilizes at some J^n. Assume $J^n \neq 0$. As $J^n J = J^n \neq 0$ and the ring is artinian, Lemma 3.4.4(ii) shows that there exists a left ideal L minimal

with respect to the property that $J^nL \neq 0$. Then $J^nx \neq 0$ for some $x \in L \setminus \{0\}$, so J^nx is a nonzero left ideal contained in L. Moreover, $J^n(J^nx) = J^nx \neq 0$, so by the minimality of L, $J^nx = L$. Hence $rx = x$ or $(1-r)x = 0$ for some $r \in J^n \subseteq J$. By Proposition 3.12.5, $x = 0$, giving a contradiction.

The second statement now follows from Proposition 3.12.9. □

Lemma 3.13.2 *If R is left artinian then $J(R)$ is the intersection of finitely many maximal left ideals of R.*

Proof Let $\mathcal{S}$ be the set of all intersections of finitely many maximal left ideals of R. By Lemma 3.4.4(ii), $\mathcal{S}$ has a minimal element J. Then J is contained in every maximal left ideal L as $J \cap L \subsetneq J$ is not possible. It follows that $J = J(R)$. □

Theorem 3.13.3 *A ring R is left semisimple if and only if R is left artinian and $J(R) = 0$.*

Proof Let R be left artinian and $J(R) = 0$. By Lemma 3.13.2, there exist maximal ideals $L_1, \ldots, L_n$ with $L_1 \cap \cdots \cap L_n = 0$. Then the map $R \to R/L_1 \oplus \cdots \oplus R/L_n$ induced by the projections $R \to R/L_i$ is injective. Thus ${}_RR$ is a submodule of a semisimple module $R/L_1 \oplus \cdots \oplus R/L_n$, whence R is left semisimple.

Conversely, let R be left semisimple. By Corollary 3.11.15(ii), R is left artinian. Moreover, in view of Wedderburn-Artin and R is isomorphic to $M_{n_1}(D_1) \oplus \cdots \oplus M_{n_m}(D_m)$. But it is easy to see that in such ring the intersection of maximal left ideals is zero. □

The rings R with $J(R) = 0$ are usually called *semiprimitive*. The theorem above shows that under the assumption that R is artinian this property is equivalent to being semisimple artinian, see §3.11. So the terminology 'semiprimitive' is usually reserved for non-artinian rings.

Lemma 3.13.4 *Let R be left artinian and V be a left R-module. Then V is semisimple if and only if $J(R)V = 0$.*

Proof Let $J := J(R)$. By Proposition 3.12.3, $JV = 0$ for a semisimple module V. Conversely, assume that $JV = 0$. Then the structure of the left R-module on V factors through to give a structure of the left R/J-

module. As R/J is semisimple by Proposition 3.12.7, this R/J-module is semisimple. Hence the original module V is also semisimple. □

Theorem 3.13.5 *Every left artinian ring is left noetherian.*

Proof By Lemma 3.13.1, ${}_RR$ has a descending sequence

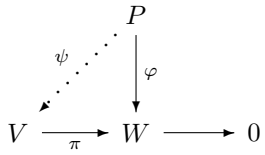
$$R = J^0 \supseteq J \supseteq J^2 \supseteq \cdots \supseteq J^n = 0.$$

Every $J^k \subseteq {}_RR$ is an artinian R-submodule, hence every J^k/J^{k+1} is an artinian R-module. Also, J^k/J^{k+1} is semisimple by Lemma 3.13.4. It follows that J^k/J^{k+1} is a direct sum of finitely many simple modules and hence noetherian. It remains to apply Lemma 3.4.5. □

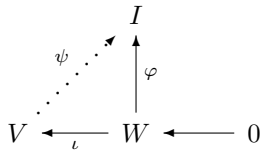
3.14 Projective and Injective Modules

Projective and injective modules are classes of modules which possess nice (categorical) properties.

Definition 3.14.1 A left R-module P is *projective* if for any two left R-modules V and W, any homomorphism $\varphi : P \to W$ and any surjective homomorphism $\pi : V \to W$ there exists a homomorphism $\psi : P \to V$ such that $\varphi = \pi \circ \psi$:



A left R-module I is *injective* if for any two left R-modules V and W, any homomorphism $\varphi : W \to I$ and any injective homomorphism $\iota : W \to V$ there exists a homomorphism $\psi : V \to I$ such that $\varphi = \psi \circ \iota$:



Remark 3.14.2 (i) The 'weird' arrows involving 0 indicate that π is surjective and ι is injective. The origin of this notation will become clear later.

(ii) Observe that the notions of projective and injective are dual to each other in the sense that one is obtained from another by 'inverting the arrows'.

The proof of the following proposition is left as an exercise:

Proposition 3.14.3

(i) *A direct sum $\oplus_{i\in I} V_i$ of left R-modules is projective if and only if each V_i is projective.*

(ii) *A direct product $\prod_{i\in I} V_i$ of left R-modules is injective if and only if each V_i is injective.*

The following theorem describes some pleasant properties of projective and injective modules.

Theorem 3.14.4

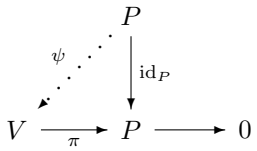
(i) *A left R-module P is projective if and only if any R-module epimorphism $\pi : V \to P$ splits, i.e. there exists a homomorphism $\psi : P \to V$ such that $\pi \circ \psi = \mathrm{id}_P$:*

$$V \underset{\psi}{\overset{\pi}{\rightleftarrows}} P \longrightarrow 0$$

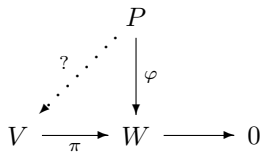
(ii) *A left R-module I is injective if and only if any R-module monomorphism $\iota : I \to V$ splits, i.e. there exists a homomorphism $\psi : V \to I$ such that $\psi \circ \iota = \mathrm{id}_I$:*

$$0 \longrightarrow I \underset{\psi}{\overset{\iota}{\rightleftarrows}} V$$

Proof We prove (i), the proof of (ii) is dual. Assume that P is projective and consider the diagram



to see that π splits. Conversely, consider the diagram



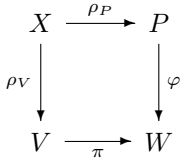
Define a submodule $X \subseteq V \oplus P$ by setting

$$X = \{(v, p) \in V \oplus P \mid \pi(v) = \varphi(p)\},$$

and the maps

$$\rho_V : X \to V,\ (v, p) \mapsto v, \qquad \rho_P : X \to P,\ (v, p) \mapsto p.$$

Note that the diagram



is commutative. Moreover, ρ_P is surjective, because π is surjective. So ρ_P splits: there exists $\alpha : P \to X$ with $\rho_P \circ \alpha = \mathrm{id}_P$. If we take $\psi = \rho_V \circ \alpha$, then

$$\pi \circ \psi = \pi \circ \rho_V \circ \alpha = \varphi \circ \rho_P \circ \alpha = \varphi \circ \mathrm{id}_P = \varphi,$$

as required. □

Remark 3.14.5 (i) Theorem 3.14.4(i) can be restated as follows: a module P is projective if and only if $P \cong V/W$ implies that W is a direct summand of V, i.e. there exists a submodule $P' \subseteq V$ with $V = P' \oplus W$. In this case we necessarily have $P' \cong P$.

Indeed, if $V = P' \oplus W$ then we can take ψ to be the isomorphism from P to P'. Conversely, if we know that the natural epimorphism $\pi : V \to V/W = P$ splits with the splitting map $\psi : P \to V$, then we can decompose $V = W \oplus \operatorname{im} \psi$. To check the last equality first pick $v \in V$ and write it as

$$v = (v - \psi(\pi(v))) + \psi(\pi(v)).$$

As $\pi(v - \psi(\pi(v))) = 0$, it follows that $v - \psi(\pi(v)) \in W$, and so $V = W + \operatorname{im} \psi$. On the other hand, if $v \in W \cap \operatorname{im} \psi$, then $v = \psi(p)$, for some $p \in P$, and $0 = \pi(v) = \pi(\psi(p)) = p$, whence $v = 0$.

(ii) Similarly, Theorem 3.14.4(ii) can be restated as follows: a module I is injective if and only if $I \subseteq V$ implies that I is a direct summand of V, i.e. there exists a submodule $J \subseteq V$ with $V = I \oplus J$.

We now explain yet another way to characterize projective and injective modules.

Definition 3.14.6 A sequence of R-maps and R-modules

$$\ldots \longrightarrow V_i \xrightarrow{\delta_i} V_{i+1} \xrightarrow{\delta_{i+1}} \ldots$$

is called *exact* if $\ker \delta_{i+1} = \operatorname{im} \delta_i$ for all i. An exact sequence of the form

$$0 \longrightarrow X \xrightarrow{\alpha} Y \xrightarrow{\beta} Z \longrightarrow 0 \tag{3.8}$$

is called a *short exact sequence*.

Thus the sequence (3.8) is eact if and only if the following three conditions are satisfied: (1) α is injective; (2) $\operatorname{im} \alpha = \ker \beta$; (3) β is surjective.

If X and Z are R-modules, it is easy to construct the stupid exact sequence of the form (3.8): namely, take Y to be $X \oplus Z$ with $\alpha = \iota_X$ the natural embedding and $\beta = \pi_Z$ the natural projection. Such 'stupid' short exact sequences are called *split*:

Definition 3.14.7 A short exact sequence (3.8) is called *split* if there exists a homomorphism $\varphi : Y \to X \oplus Z$ which makes the following diagram commutative:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & X & \xrightarrow{\alpha} & Y & \xrightarrow{\beta} & Z & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \mathrm{id}_X} & & \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \mathrm{id}_Z} & & \\
0 & \longrightarrow & X & \xrightarrow{\iota_X} & X \oplus Z & \xrightarrow{\pi_Z} & Z & \longrightarrow & 0
\end{array}$$

Note that in the definition above φ will automatically have to be an isomorphism. The following is easy to check:

Lemma 3.14.8 *The following conditions on a short exact sequence (3.8) are equivalent:*

(i) *The sequence is split.*
(ii) *The monomorphism α splits.*
(iii) *The epimorphism β splits.*

Now, it is clear that Theorem 3.14.4 can be restated as follows:

Theorem 3.14.9

(i) *An R-module P is projective if and only if any short exact sequence of the form*

$$0 \to X \to Y \to P \to 0$$

splits.

(ii) *An R-module I is projective if and only if any short exact sequence of the form*

$$0 \to I \to Y \to Z \to 0$$

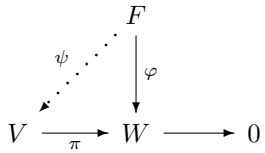
splits.

The following characterization of projective modules does not have an analogue for injective modules.

Theorem 3.14.10 *A module is projective if and only if it is isomorphic to a direct summand of a free module. In particular every free module is projective.*

Proof Let P be projective. By Corollary 3.6.5, P is a quotient of a free module F, and by Remark 3.14.5(i), P is a direct summand of F.

Next, we prove that every free module is projective. Let F be a free module with basis $(f_i)_{i\in I}$, and consider the diagram



As π is surjective, there is for each $i \in I$ some $v_i \in V$ such that $\pi(v_i) = \varphi(f_i)$. By the universal property of free modules, there is a homomorphism $\psi : F \to V$ with $\psi(f_i) = v_i$ for all i. Then $\pi(\psi(f_i)) = \varphi(f_i)$ for all i, whence $\pi \circ \psi = \varphi$.

Finally, if P is a direct summand of a free module F, then P is projective by Proposition 3.14.3.

□

We know that all modules over a division ring are free, hence projective. On the other hand, we know that a submodule of a finitely generated free module over a PID is again free. It can be proved that this result is true even without assuming that the free module is finitely

generated. It follows that every projective module over a PID is free. The converse is of course not true in general. For example consider an irreducible module module over $M_n(D)$ and use the following

Proposition 3.14.11 *For a ring R the following conditions are equivalent:*

(i) *R is semisimple artinian;*
(ii) *Every left R-module is projective;*
(iii) *Every left R-module is injective*

Proof Assume that R is semisimple artinian. Then every left R-module is semisimple. In particular, every short exact sequence of R-modules splits. So every R-module is projective and injective by Lemma 3.14.9.

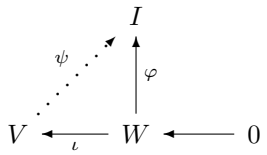
Assume that every left R-module is projective or that every every left R-module is injective. By Lemma 3.14.9, every short exact sequence of R-modules is split, which implies that every R-module is semisimple, whence R is semisimple artinian. □

We now turn to some properties specific for injective modules.

Proposition 3.14.12 (Baer's Criterion) *For a left R-module I the following conditions are equivalent:*

(i) *I is injective;*
(ii) *Every R-module homomorphism of a left ideal of R into I can be extended to ${}_RR$.*

Proof (i) implies (ii) by definition of an injective module. To see that (ii) implies (i), consider the diagram



Let $\mathcal{F}$ be the family of all R-module homomorphisms $\theta : X \to I$ where $\operatorname{im}\iota \subseteq X \subseteq V$ and $\theta \circ \iota = \varphi$. Note $\mathcal{F}$ is non-empty, as it contains $\varphi \circ \iota^{-1} : \operatorname{im}\iota \to I$. Partially order $\mathcal{F}$ by $\theta \leq \theta'$ if and only if the domain of θ is contained in the domain of θ' and θ is the restriction of θ' to the domain of θ. By Zorn's lemma, there exists a maximal element $\psi : X \to I$ in $\mathcal{F}$. We just need to prove that $X = V$. Otherwise, take

$v \in V \setminus X$ and consider the left ideal $L = \{r \in R \mid rv \in X\}$. The map $L \to I$ given by $r \mapsto \psi(rv)$ is an R-module homomorphism. By hypothesis, there exists an R-homomorphism $\alpha : {}_RR \to I$ such that $\alpha(r) = \psi(rv)$ for any $r \in L$. Let $w = \alpha(1_R)$ and consider the map

$$\hat{\psi} : X + Rv \to I,\ x + rv \mapsto \psi(x) + rw.$$

We claim that $\hat{\psi}$ is well-defined. Indeed, if $x_1 + r_1v = x_2 + r_2v \in X + Rv$, then $x_1 - x_2 = (r_2 - r_1)v \in X \cap Rv$. Hence $r_2 - r_1 \in L$ and

$$\begin{aligned}\psi(x_1) - \psi(x_2) = \psi(x_1 - x_2) = \psi((r_2 - r_1)v) = \alpha(r_2 - r_1)\\ = (r_2 - r_1)\alpha(1_R) = (r_2 - r_1)w = r_2w - r_1w.\end{aligned}$$

Therefore

$$\hat{\psi}(x_1 + r_1v) = \psi(x_1) + r_1w = \psi(x_2) + r_2w = \hat{\psi}(x_2 + r_2v).$$

Finally, it is easy to see that $\hat{\psi}$ is an element of $\mathcal{F}$. This contradicts the maximality of ψ. Therefore $X = V$. □

Definition 3.14.13 A left R-module V is called *divisible* if $rV = V$ for all $r \in R \setminus \{0\}$.

Corollary 3.14.14 *If R is a PID, an R-module is injective if and only if it is divisible.*

Proof Let I be injective, $v \in I$, and $r \in R \setminus \{0\}$. There is a module homomorphism $\varphi : Rr \to I,\ sr \mapsto sv$. As I is injective, φ may be extended to a module homomorphism $\hat{\varphi} : {}_RR \to I$. Then there exists $w \in I$ such that $\hat{\varphi}(s) = sw$ for all $s \in R$. Then

$$v = \varphi(r) = \hat{\varphi}(r) = rw,$$

hence I is divisible.

Conversely, let I be divisible. Any ideal of R looks like Rr. Let $\varphi : Rr \to I$ be a module homomorphism. In view of Baer's criterion, it suffices to prove that φ extends to a homomorphism ${}_RR \to I$. If $r = 0$ just extend to the zero homomorphism. Otherwise, let $v = \varphi(r) \in I$. As I is divisible, we have $v = rw$ for some $w \in I$. Now, define the desired extension to be $\hat{\varphi}(s) = sw$. □

Example 3.14.15 In view of the previous corollary, $\mathbb{Q}$ is an injective $\mathbb{Z}$-module. By the way it is not projective, because it is clearly not free,

and we noted above that all projective modules over a PID are free (also see Problem 3.17.102).

Another example of a divisible $\mathbb{Z}$-module is C_{p^∞}, see Example 1.4.8 (check that it is indeed divisible!). In fact it can be shown that a $\mathbb{Z}$-modue is divisible (equivalently injective) if and only if it is a direct sum of copies of $\mathbb{Q}$ and $\mathbb{C}_{p^\infty}$.

On the other hand, the regular module ${}_{\mathbb{Z}}\mathbb{Z}$ is not divisible and hence not injective.

We know that every module is a quotient of a projective (and even free) module. We now want to prove the dual statement that every module is a submodule of an injective module. This turns out to be much harder to prove. First we prove this for $\mathbb{Z}$-modules:

Lemma 3.14.16 *Every $\mathbb{Z}$-module may be embedded in a divisible $\mathbb{Z}$-module.*

Proof Let V be a Z-module. We know that V is an epimorphic image of a free $\mathbb{Z}$-module F with kernel K. Since F is a direct sum of copies of $\mathbb{Z}$ and $\mathbb{Z} \subset \mathbb{Q}$, F may be embedded in a direct sum D of copies of $\mathbb{Q}$. It is clear that D is divisible. Now, the embedding $\iota : F \to D$ induces an embedding $F/K \to D/\iota(K)$. It remains to notice that a quotient of a divisible module is again divisible, and so by Baer's criterion we have embedded $V \cong F/K$ into an injective module $D/\iota(K)$. □

If V is a $\mathbb{Z}$-module, the abelian group $\mathrm{Hom}_{\mathbb{Z}}(R, V)$ can be considered as a left R-module via

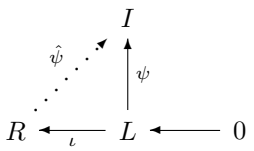
$$(r\theta)(s) = \theta(sr) \qquad (\theta \in \mathrm{Hom}_{\mathbb{Z}}(R, V),\ r, s \in R).$$

Lemma 3.14.17 *If I is an injective $\mathbb{Z}$-module then $\mathrm{Hom}_{\mathbb{Z}}(R, I)$ is an injective left R-module.*

Proof We will use Baer's criterion. So, let L be a left ideal of R and $\varphi : L \to \mathrm{Hom}_{\mathbb{Z}}(R, I)$ be an R-homomorphism. The map

$$\psi : L \to I, r \mapsto [\varphi(r)](1_R)$$

is a $\mathbb{Z}$-homomorphism. As I is an injective $\mathbb{Z}$-module, the diagram



where ι is an embedding $L \subseteq R$, yields an extension $\hat{\psi} : R \to I$ of ψ. Define $\hat{\varphi} : R \to \mathrm{Hom}_{\mathbb{Z}}(R, I),\ r \mapsto \hat{\varphi}(r)$, where $\hat{\varphi}(r) : R \to I$ is the map given by $[\hat{\varphi}(r)](s) = \hat{\psi}(sr)$ $(s \in R)$. It is routine to check that each $\hat{\varphi}(r) \in \mathrm{Hom}_{\mathbb{Z}}(R, I)$, and that $\hat{\varphi}$ is $\mathbb{Z}$-linear. We next check that $\hat{\varphi}$ is R-linear. Indeed,

$$[\hat{\varphi}(tr)](s) = \hat{\psi}(str) = [\hat{\varphi}(r)](st) = [t\hat{\varphi}(r)](s) \qquad (r, s, t \in R),$$

as required.

Finally, we check that $\hat{\varphi}$ is an extension of φ. Suppose $r \in L$ and $s \in R$. Then $sr \in L$ and

$$\begin{aligned}\hat{\varphi}(r)(s) &= \hat{\psi}(sr) = \psi(sr) = [\varphi(sr)](1_R) \\ &= [s\varphi(r)](1_R) = [\varphi(r)](1_R s) = [\varphi(r)](s),\end{aligned}$$

as required. □

Theorem 3.14.18 *Every R-module V may be embedded in an injective R-module.*

Proof As a $\mathbb{Z}$-module we may embed V into a free $\mathbb{Z}$-module I, see Lemma 3.14.16. This induces an R-module embedding $\mathrm{Hom}_{\mathbb{Z}}(R, V) \to \mathrm{Hom}_{\mathbb{Z}}(R, I)$ (see more about it in §3.15). Moreover, $V \cong \mathrm{Hom}_R({}_R R, V)$ embeds into $\mathrm{Hom}_{\mathbb{Z}}(R, V)$ as an R-module (again, if this is confusing, read §3.15 and in particular Lemma 3.15.8). Thus we embedded V into the R-module $\mathrm{Hom}_{\mathbb{Z}}(R, I)$, which is injective by Lemma 3.14.17. □

Remark 3.14.19 The result about embedding any R-module into an injective module can be made more precise. Let us say that a submodule V of a module W is *essential* if $V \cap X \neq 0$ for every submodule $X \subseteq W$. It can be proved that every module V can be embedded in an essential way into an injective module I_V, and this injective module I_V is defined uniquely up to isomorphism. I_V is called the *injective hull* of V.

The dual object, called the *projective cover* of V does not exist in general, but its existence for every module is guaranteed under very reasonable assumptions on the ring R.

3.15 Hom and Duality

In this section we will give perhaps the most important characterization of projective and injective modules in terms of Hom-functors.

Let V, W, X be R-modules, and $f : V \to W$ be an R-module homomorphism. It induces homomorphisms of abelian groups

$$f_* = f_*^X : \mathrm{Hom}_R(X, V) \to \mathrm{Hom}_R(X, W),\ \theta \mapsto f \circ \theta, \tag{3.9}$$

$$f^* = f_X^* : \mathrm{Hom}_R(W, X) \to \mathrm{Hom}_R(V, X),\ \theta \mapsto \theta \circ f. \tag{3.10}$$

If $g : Y \to V$ is another R-module map, then we have the following obvious properties:

$$(f \circ g)_* = f_* \circ g_*, \quad \text{and} \quad (f \circ g)^* = g^* \circ f^*. \tag{3.11}$$

Recall from Definition 3.14.6 that the sequence of R-modules and R-module homomorphisms

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W$$

is called exact if f is injective and $\ker g = \mathrm{im}\, f$.

Theorem 3.15.1 *The sequence of R-modules and R-module homomorphisms*

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \tag{3.12}$$

is exact if and only if the corresponding sequence

$$0 \longrightarrow \mathrm{Hom}_R(X, U) \xrightarrow{f_*} \mathrm{Hom}_R(X, V) \xrightarrow{g_*} \mathrm{Hom}_R(X, W) \tag{3.13}$$

of abelian groups is exact for every R-module X.

Proof Assume that (3.12) is exact. We need to show that f_* is injective and $\mathrm{im}\, f_* = \ker g_*$.

To see that f_* is injective, assume that $f_*(\theta) = 0$ for some $\theta \in \mathrm{Hom}_R(X, U)$. Well, that means $f \circ \theta = 0$. But f is injective, so $f \circ \theta = 0$ is only possible if $\theta = 0$.

To see that $\operatorname{im} f_* = \ker g_*$, note first of all that $\operatorname{im} f_* \subseteq \ker g_*$. Indeed, $\operatorname{im} f_* \subseteq \ker g_*$ is equivalent to $g_* \circ f_* = 0$, which is true in view of (3.11):

$$g_* \circ f_* = (g \circ f)_* = 0_* = 0.$$

Finally, we prove that $\operatorname{im} f_* \supseteq \ker g_*$. Let $\theta \in \ker g_*$, i.e. $g \circ \theta = 0$. Hence $\operatorname{im} \theta \subseteq \ker g = \operatorname{im} f$. As f is a monomorphism, the map $\tilde{f} : U \to \operatorname{im} f$ is an isomorphism. Define the map $\psi \in \operatorname{Hom}_R(X, U)$ as the composition

$$X \xrightarrow{\theta} \operatorname{im} \theta \subseteq \ker g = \operatorname{im} f \xrightarrow{\tilde{f}^{-1}} U.$$

Note that $f_*(\psi) = \theta$, so that $\theta \in \operatorname{im} f_*$, as required.

Conversely, assume that (3.13) is exact. To see that f is injective take $X = \ker f$ and apply f_* to the embedding $\ker f \to U$. Next take $X = U$. As $g_*(f_*(\mathrm{id}_U)) = 0$, we have $g \circ f = 0$, i.e. $\operatorname{im} f \subseteq \ker g$. Finally, let $X = \ker g$ and $\theta : X \to V$ be the inclusion map. Note that $g_*(\theta) = g \circ \theta = 0$, i.e. $\theta \in \ker g_* = \operatorname{im} f_*$, i.e. there exists $\varphi \in \operatorname{Hom}_R(X, U)$ with $\theta = f \circ \varphi$. Therefore for every $v \in \ker g$, we have $v = \theta(v) = f(\varphi(v)) \in \operatorname{im} f$, as required. □

The following dual statement to Theorem 3.15.1 is left as an (extremely useful) exercise.

Theorem 3.15.2 *The sequence of R-modules and R-module homomorphisms*

$$U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0 \tag{3.14}$$

is exact if and only if the corresponding sequence

$$0 \longrightarrow \operatorname{Hom}_R(W, X) \xrightarrow{g^*} \operatorname{Hom}_R(V, X) \xrightarrow{f^*} \operatorname{Hom}_R(U, X) \tag{3.15}$$

of abelian groups is exact for every R-module X.

It is not true in general that a short exact sequence of modules induces a short exact sequence of Hom's. However, the following three theorems show that this is true in some important situations.

Theorem 3.15.3 *The following conditions are equivalent:*

(i) $0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$ *is a split exact sequence of R-modules.*

(ii) $0 \longrightarrow \operatorname{Hom}_R(X, U) \xrightarrow{f_*} \operatorname{Hom}_R(X, V) \xrightarrow{g_*} \operatorname{Hom}_R(X, W) \longrightarrow 0$ *is a split exact sequence of abelian groups for every R-module X.*

(iii) $0 \longrightarrow \mathrm{Hom}_R(W,X) \xrightarrow{g^*} \mathrm{Hom}_R(V,X) \xrightarrow{f^*} \mathrm{Hom}_R(U,X) \longrightarrow 0$ *is a split exact sequence of abelian groups for every R-module X.*

Proof We prove (i) $\Leftrightarrow$ (iii). The proof of (i) $\Leftrightarrow$ (ii) is similar. By Lemma 3.14.8, if (i) holds, there exists a homomorphism $h : V \to U$ such that $hf = \mathrm{id}_U$. Then $f^* \circ h^* = (f \circ h)^* = (\mathrm{id}_U)^* = \mathrm{id}_{\mathrm{Hom}_R(U,X)}$. Hence f^* is an epimorphism. Now, Theorem 3.15.2 and Lemma 3.14.8 imply that the sequence in (iii) is split exact.

Conversely, take $X = U$. Then by surjectivity of f^* there exists $\theta \in \mathrm{Hom}_R(V,X)$ such that $\theta \circ f = f^*(\theta) = \mathrm{id}_U$. Hence f is injective, and the sequence in (i) is split exact thanks to Theorem 3.15.2 and Lemma 3.14.8. □

Theorem 3.15.4 *The following conditions on an R-module P are equivalent:*

(i) *P is projective;*

(ii) *For any short exact sequence*

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

of R-modules the corresponding sequence

$$0 \longrightarrow \mathrm{Hom}_R(P,U) \xrightarrow{f_*} \mathrm{Hom}_R(P,V) \xrightarrow{g_*} \mathrm{Hom}_R(P,W) \longrightarrow 0$$

of abelian groups is exact.

(iii) *For any epimorphism $g : V \to W$ of R-modules the map*

$$g_* : \mathrm{Hom}_R(P,V) \to \mathrm{Hom}_R(P,W)$$

is an epimorphism.

Proof (ii) $\Leftrightarrow$ (iii) is clear by Theorem 3.15.1. As for (i) $\Leftrightarrow$ (iii), note that the surjectivity of g_* is a restatement of Definition 3.14.1. □

The proof of the following theorem is dual:

Theorem 3.15.5 *The following conditions on an R-module I are equivalent:*

(i) *I is injective;*

(ii) *For any short exact sequence*

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

of R-modules the corresponding sequence

$$0 \longrightarrow \mathrm{Hom}_R(W, I) \xrightarrow{g^*} \mathrm{Hom}_R(V, I) \xrightarrow{f^*} \mathrm{Hom}_R(U, I) \longrightarrow 0$$

of abelian groups is exact.

(iii) *For any monomorphism $f : U \to V$ of R-modules the map*

$$f^* : \mathrm{Hom}_R(V, I) \to \mathrm{Hom}_R(U, I)$$

is an epimorphism.

Remark 3.15.6

(i) If X is not projective, it is not in general true that for a surjective homomorphism $g : V \to W$, the corresponding homomorphism $g_* : \mathrm{Hom}_R(X, V) \to \mathrm{Hom}_R(X, W)$ is also surjective. Indeed consider the natural projection of $\mathbb{Z}$-modules

$$g : \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$$

and take $X = \mathbb{Z}/2\mathbb{Z}$. Then

$$\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}) = 0 \neq \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/2\mathbb{Z}) \neq 0,$$

so g_* cannot be surjective.

(ii) If X is not injective, it is not in general true that for a monomorphism $f : U \to V$, the corresponding homomorphism

$$f^* : \mathrm{Hom}_R(V, I) \to \mathrm{Hom}_R(U, I)$$

is surjective. Indeed consider the monomorohism of $\mathbb{Z}$-modules

$$f : \mathbb{Z} \to \mathbb{Z},\ m \mapsto 2m$$

and take $X = \mathbb{Z}/2\mathbb{Z}$. Check that f^* is the zero map between two modules isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

We now explain an important construction concerning Hom spaces. If V is a left (R, T)-bimodule and W is an (R, S)-bimodule, then the Hom-space $\mathrm{Hom}_R(V, W)$ inherits a structure of (T, S)-bimodule:

$$(\theta s)(v) = \theta(v)s, \quad (t\theta)(v) = \theta(vt)$$

for all $\theta \in \mathrm{Hom}_R(V, W),\ v \in V,\ s \in S$, and $t \in T$. A special case of this construction yields a structure of a right R-module on the Hom-space in the following definition.

Definition 3.15.7 Let V be a left R-module. The right module

$$V^* := \mathrm{Hom}_R(V, {}_R R_R)$$

is called the *dual* module of V.

Similarly, if V is a right R-module the corresponding left module V^* can be defined.

As expected, there is a natural (R-module) homomorphism $V \to V^{**}$. A module V is called *reflexive* if this homomorphism is an isomorphism. An example of a reflexive module is given by a free module with a finite basis and a finitely generated projective R-module, see Problems 3.17.116 and 3.17.117.

We finish the section with the following useful lemma:

Lemma 3.15.8 *Let V be a left R-module. Then there is an isomorphism of left R-modules*

$$\mathrm{Hom}_R({}_R R_R, V) \tilde{\to} V, \ \theta \mapsto \theta(1_R).$$

Proof Exercise. □

3.16 Tensor Products

Tensor products (just like Hom) is a very useful basic construction involving modules. Recall that to a pair of left R-modules ${}_R V$ and ${}_R W$ we have associated an abelian group $\mathrm{Hom}_R({}_R V, {}_R W)$. Moreover, if, additionally V was an (R, S)-bimodule and W was an (R, T)-bimodule, then $\mathrm{Hom}_R({}_R V_S, {}_R W_T)$ has a natural structure of an (S, T)-bimodule. This construction is fundamental for module theory. In this section we describe another fundamental construction, called the *tensor product*, which associates to a right R-module V_R and the left R-module ${}_R W$ the abelian group

$$V_R \otimes_R {}_R W.$$

Moreover,

$${}_S V_R \otimes_R {}_R W_T$$

will turn out to have a natural structure of an (S, T)-bimodule. Of course, most of the time we will drop the indices from the module notation and just write

$$V \otimes_R W.$$

We are usually *not* going to drop the index from the symbol $\otimes_R$ though, just like for the Hom-notation.

The abelian group $V \otimes_R W$ may be defined by a universal property or by explicit construction. It is important to understand both. We start from the universal property.

Definition 3.16.1 Let V be a right R-module, W be a left R-module, and A be an abelian group. A map $\varphi : V \times W \to A$ is called *R-biadditive* if the following three conditions are satisfied:

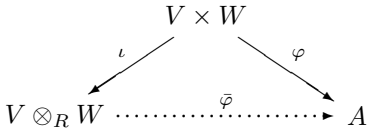
(i) $\varphi(v_1 + v_2, w) = \varphi(v_1, w) + \varphi(v_2, w)$ for all $v_1, v_2 \in V$, $w \in W$;

(ii) $\varphi(v, w_1 + w_2) = \varphi(v, w_1) + \varphi(v, w_2)$ for all $v \in V$, $w_1, w_2 \in W$;

(iii) $\varphi(vr, w) = \varphi(v, rw)$ for all $v \in V$, $w \in W$, $r \in R$.

We now define $V \otimes_R W$ via the universal property.

Definition 3.16.2 Let V be a right R-module and W be a left R-module. The *tensor product of V and W* is a pair $(V \otimes_R W, \iota)$ where $V \otimes_R W$ is an abelian group, $\iota : V \times W \to V \otimes_R W$ is an R-biadditive map, which is universal among all R-biadditive maps from $V \times W$ to abelian groups in the following sense: for any R-biadditive map $V \times W \to A$ there exists a unique map $\bar{\varphi} : V \otimes_R W \to A$ such that $\varphi = \bar{\varphi} \circ \iota$. The situation is described by the following diagram:



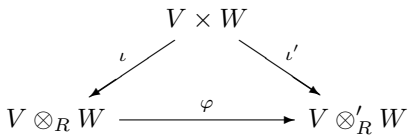
(3.16)

Remark 3.16.3 You need to get used to people being sloppy and just referring to the *abelian group* $V \otimes_R W$ as the *tensor product* of V and W. That is, the map ι is not mentioned, but it is *there!*

The first lemma is standard for things defined by universal properties: it claims that if the universal object exists then it is unique.

Lemma 3.16.4 *Assume that the pairs $(V \otimes_R W, \iota)$ and $(V \otimes'_R W, \iota')$ both satisfy the universal property of Definition 3.16.2. Then there exists a*

unique isomorphism $\varphi : V \otimes_R W \xrightarrow{\sim} V \otimes'_R W$ *such that* $\varphi \circ \iota = \iota'$*:*



Proof A standard exercise, which by now should be a piece of cake for you. □

We now deal with the existence of tensor product. Thus we give an explicit construction of the pair $(V \otimes_R W, \iota)$ which satisfies the required universal property. Let F be a free abelian group (equivalently, free $\mathbb{Z}$-module) on the set $V \times W$. Let K be the subgroup of F generated by all elements of the form:

(i) $(v_1 + v_2, w) - (v_1, w) - (v_2, w)$ for all $v_1, v_2 \in V$, $w \in W$;
(ii) $(v, w_1 + w_2) - (v, w_1) - (v, w_2)$ for all $v \in V$, $w_1, w_2 \in W$;
(iii) $(vr, w) - (v, rw)$ for all $v \in V$, $w \in W$, $r \in R$.

Denote $V \otimes_R W := F/K$, $v \otimes w := (v, w) + K$ for all $v \in V, w \in W$, and set

$$\iota : V \times W \to V \otimes_R W,\ (v, w) \mapsto v \otimes w.$$

Theorem 3.16.5 *The pair* $(V \otimes_R W, \iota)$ *is the tensor product of* V *and* W *in the sense of Definition 3.16.2.*

Proof First of all observe that the map ι is indeed R-biadditive. Moreover, note that every element of $V \otimes_R W$ is a finite sum of elements of the form $v \otimes w$.

Now, consider the diagram (3.16). In view of the previous remark, if $\bar{\varphi}$ exists, it is unique, because it must send $v \otimes w$ to $\varphi((v, w))$.

To show the existence of $\bar{\varphi}$, first define $\hat{\varphi} : F \to A$ as the homomorphism of abelian groups with $\hat{\varphi}((v, w)) = \varphi((v, w))$. Such $\hat{\varphi}$ exists by the universal property of free abelian groups (equivalently, free $\mathbb{Z}$-modules). Now, it is not hard to see that $\hat{\varphi}$ annihilates K. So it factors through to give the map $\bar{\varphi} : V \otimes_R W = F/K \to A$, which is the one we are looking for. □

Remark 3.16.6 Informally, people think of the tensor product $V \otimes_R W$ as sums of *pure tensors* $v \otimes w$, which are objects having the following properties:

(i) $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$ for all $v_1, v_2 \in V$, $w \in W$;
(ii) $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$ for all $v \in V$, $w_1, w_2 \in W$;
(iii) $vr \otimes w = v \otimes rw$ for all $v \in V$, $w \in W$, $r \in R$.

Observe there is no mentioning of ι althought it is of course there: it maps (v, w) to $v \otimes w$. Also, in the notation for pure tensors $v \otimes w$ it is customary to drop the index R.

We now consider what happens if V and W have more structure than descrihan they did in Definition 3.16.2.

Proposition 3.16.7 *Assume that V is an (S, R)-bimodule and W is an (R, T)-bimodule. Then $V \otimes_R W$ is an (S, T)-bimodule with the action on elementary tensors given by*

$$s(v \otimes w)t = (sv) \otimes (wt)$$

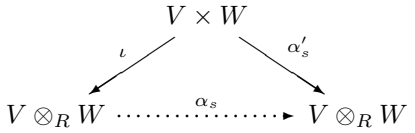
for all $s \in S, t \in T, v \in V, w \in W$.

Proof Read the following proof carefully, especially if you think that everything is clear and there is *nothing to prove* here.

We prove for example that there is an S-action on $V \otimes_R W$ given by $s(v \otimes w) = (sv) \otimes w$. Everything is indeed more or less clear, *except* the fact that for every $s \in S$ there exists a homomorphism of abelian groups

$$\alpha_s : V \otimes_R W \to V \otimes_R W,\ v \otimes w \mapsto (sv) \otimes w,$$

the problem being that there are *relations* between pure tensors, such as $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$, and we cannot just define a map on pure tensors by sending them wherever we like. Well, having said that, it's not a big deal to show that α_s exists—just use the universal property for the following diagram



where the map $\alpha'_s : (v, w) \mapsto sv \otimes w$ needs to be checked to be R-biadditive.

Now, that we've got α_s, it is easy to check that it defines the structure of the left S-module on $V \otimes_R W$ via $sx = \alpha_s(x)$ for any $x \in V \otimes_R W$ (verify all the axioms!). The structure of a right T-module on $V \otimes_R W$ is

defined similarly. Finally, one checks that the two structures commute. □

Note that if R is *commutative* then any R-module can be considered as an an (R, R)-bimodule. So in this case the tensor product of two R-modules is again an R-module in a natural way: we have

$$r(v \otimes w) = (rv) \otimes w = v \otimes (rw) \qquad (v \in V, w \in W, r \in R),$$

In particular, tensor product of two F-vector spaces is again an F-vector space.

The following pleasant property of tensor products is reminiscent of Lemma 3.15.8.

Lemma 3.16.8 *Let V be a right R-module and W be a left R-module. Then there exist isomorphisms of modules*

$$V \otimes_R R_R \tilde{\to} V, \ v \otimes r \mapsto vr,$$
$$_R R \otimes_R W \tilde{\to} W, \ r \otimes w \mapsto rw.$$

Proof We sketch the proof for the first isomorphism. First, use the universal property of tensor products to convince yourself that there exists a map $\alpha : V \otimes_R R \to V$ of right R-modules, which on pure tensors works as follows:

$$\alpha(v \otimes r) = vr.$$

Now, define the map $\beta : V \to V \otimes_R R$ via

$$\beta(v) := v \otimes 1.$$

Note (the pleasant fact) that we do not need to check that β is well-defined! What one needs to check (and this is very easy) is that β is a homomorphism of right R-modules, inverse to α. So α is an isomorphism. □

Another important fundamental fact on tensor products:

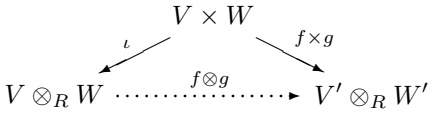
Lemma 3.16.9 *Let V, V' be right R-modules and W, W' be left R-modules. Suppose $f : V \to V'$ and $g : W \to W'$ are module homomorphisms. Then there exists a homomorphism of abelian groups*

$$f \otimes g : V \otimes_R W \to V' \otimes_R W', \ v \otimes w \mapsto f(v) \otimes g(w).$$

Moreover, if V is an (S, R)-bimodule, W is an (R, T)-bimodule, and f, g

respect these additional structures, then $f \otimes g$ is a homomorphism of (S,T)-bimodules.

Proof As usual, everything is routine, except for the existence of $f \otimes g$, which, also as usual, follows from the universal property of tensor products:



□

Compare the following result with Theorems 3.15.1 and 3.15.2.

Theorem 3.16.10 (Right Exactness of Tensor Product) *The sequence of left R-modules and R-module homomorphisms*

$$U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0 \tag{3.17}$$

is exact if and only if the corresponding sequence

$$X \otimes_R U \xrightarrow{\mathrm{id}_X \otimes f} X \otimes_R V \xrightarrow{\mathrm{id}_X \otimes g} X \otimes_R W \longrightarrow 0 \tag{3.18}$$

of abelian groups is exact for every right R-module X.

Proof In view of Lemma 3.16.8, the sequence (3.18) with $X = R_R$ is isomorphic to the sequence (3.17), so the 'if' part follows (check the details—everything is not as obvious as I am trying to present here). For the 'only-if' part, assume (3.17) is exact. As g is surjective, every pure tensor $x \otimes w$ belongs to the image of $\mathrm{id}_X \otimes g$, whence $\mathrm{id}_X \otimes g$ is surjective. Moreover,

$$(\mathrm{id}_X \otimes g) \circ (\mathrm{id}_X \otimes f) = \mathrm{id}_X \otimes (g \circ f) = 0,$$

i.e. $\mathrm{im}\ \mathrm{id}_X \otimes f \subseteq \ker \mathrm{id}_X \otimes g$.

The difficult part of the proof is to establish the converse inclusion. For that it suffices to show that the natural map

$$\beta : X \otimes_R V/(\mathrm{im}\ \mathrm{id}_X \otimes f) \to X \otimes_R V/(\ker \mathrm{id}_X \otimes g)$$

is an injection. But β is injective if and only if $\alpha := \gamma \circ \beta$ is, where

$$\gamma : X \otimes_R V/(\ker \mathrm{id}_X \otimes g) \to X \otimes_R W$$

is the natural isomorphism which maps the coset of $x \otimes v$ to $x \otimes g(v)$. Thus, we are after proving that the map

$$\alpha : X \otimes_R V/(\text{im id}_X \otimes f) \to X \otimes_R \otimes W,\ \pi(x \otimes v) \mapsto x \otimes g(v)$$

is injective, where

$$\pi : X \otimes_R V \to X \otimes_R V/(\text{im id}_X \otimes f)$$

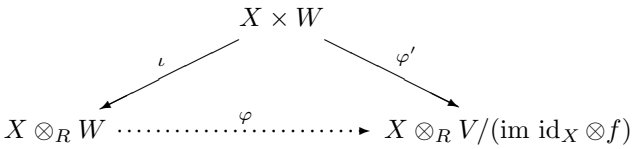
is the natural projection. Define a map

$$\varphi' : X \times W \to X \otimes_R V/(\text{im id}_X \otimes f), (x, w) \mapsto \pi(x \otimes v),$$

where $v \in g^{-1}(w)$. Note that φ' is well-defined: if $g(v_1) = g(v_2) = w$, then $g(v_1 - v_2) = 0$, and so $v_1 - v_2 = f(u)$ for some $u \in U$, hence

$$\begin{aligned}\pi(x \otimes v_1) - \pi(x \otimes v_2) &= \pi(x \otimes v_1 - x \otimes v_2) = \pi(x \otimes (v_1 - v_2)) \\ &= \pi(x \otimes f(u)) = \pi((\text{id}_X \otimes f)(x \otimes u)) = 0.\end{aligned}$$

Also, it is clear that φ' is R-biadditive. Considering the diagram



we get a homomorphism φ, which maps (x, w) to $\pi(x \otimes v)$, where $v \in g^{-1}(w)$. Now,

$$\varphi(\alpha(\pi(x \otimes v))) = \varphi(x \otimes g(v)) = \pi(x \otimes v),$$

i.e. φ is a left inverse to α, whence α is injective, as required. □

The obvious version of Theorem 3.16.10 for right modules is left as an exercise.

Compare the following remark with Remark 3.15.6.

Remark 3.16.11 In general, it is not true that applying $X \otimes_R -$ to an exact sequence yields an exact sequences, the problem being that even if $f : U \to V$ is injective, the map $\text{id}_X \otimes f : X \otimes_R U \to X \otimes_R V$ does not have to be injective. For example, consider the injective map of $\mathbb{Z}$-modules

$$f : \mathbb{Z} \to \mathbb{Z}, m \mapsto 2m.$$

Tensoring this with the $\mathbb{Z}$-module $\mathbb{Z}/2\mathbb{Z}$ gives the map

$$\text{id} \otimes f : \mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z},$$

which is actually the zero map between two modules isomorphic to $\mathbb{Z}/2\mathbb{Z}$. Check this.

The next definition introduces an important class of modules for tensor products play a role similar to that played by projective and injective modules for Hom's:

Definition 3.16.12 A right R-module X is called *flat* if whenever

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

is an exact sequence of left R-modules,

$$0 \longrightarrow X \otimes_R U \xrightarrow{\mathrm{id}_X \otimes f} X \otimes_R V \xrightarrow{\mathrm{id}_X \otimes g} X \otimes_R W \longrightarrow 0$$

is an exact sequence of abelian groups.

By Theorem 3.16.10, in order to check flatness we only have to worry about injective maps going to injective maps, everything else being automatic.

Theorem 3.16.13 (Associativity of tensor product) *Given U_R, ${}_RV_S$, and ${}_SW$, there is an isomorphism*

$$(U \otimes_R V) \otimes_S W \xrightarrow{\sim} U \otimes_R (V \otimes_S W),\ (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$$

Proof By now you should be good in this. So, try this one on your own. Well, just in case, for every $w \in W$, consider the diagram

$$\begin{array}{ccc} & U \times V & \\ {\scriptstyle \iota}\swarrow & & \searrow{\scriptstyle \alpha'_w} \\ U \otimes_R V & \xrightarrow{\ \alpha_w\ } & U \otimes_R (V \otimes_S W) \end{array}$$

to get the automorphism of abelian groups

$$\alpha_w : U \otimes_R V \to U \otimes_R (V \otimes_S W),\ u \otimes v \mapsto u \otimes (v \otimes w).$$

Next consider the diagram

$$\begin{array}{ccc} & (U \otimes_R V) \times W & \\ {\scriptstyle \iota}\swarrow & & \searrow{\scriptstyle \alpha'} \\ (U \otimes_R V) \otimes_S W & \xrightarrow{\ \alpha\ } & U \otimes_R (V \otimes_S W) \end{array}$$

where

$$\alpha'((\sum_i u_i \otimes v_i), w) = \alpha_w(\sum_i u_i \otimes v_i) = \sum_i u_i \otimes (v_i \otimes w).$$

This gives a map

$$\alpha : (U \otimes_R V) \otimes_S W \to U \otimes_R (V \otimes_S W),\ (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w).$$

The inverse map

$$\beta : U \otimes_R (V \otimes_S W) \to (U \otimes_R V) \otimes_S W,\ u \otimes (v \otimes w) \mapsto (u \otimes v) \otimes w$$

is constructed in a similar way. □

Theorem 3.16.14 (Additivity of tensor product)

(i) *Given U_R, ${}_RV$, and ${}_RW$, there is an isomorphism of abelian groups*

$$U \otimes_R (V \oplus W) \tilde{\to} (U \otimes_R V) \oplus (U \otimes_R W),\ u \otimes (v, w) \mapsto (u \otimes v, u \otimes w).$$

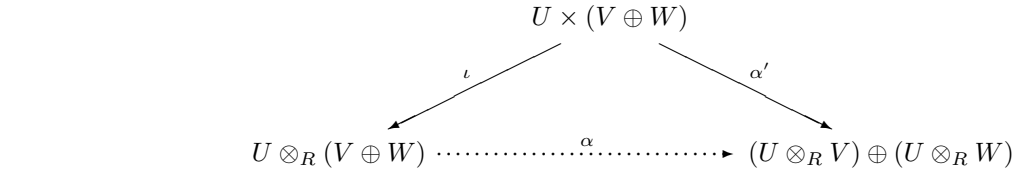
(ii) *Given U_R, V_R, and ${}_RW$, there is an isomorphism of abelian groups*

$$(U \oplus V) \otimes_R W \tilde{\to} (U \otimes_R W) \oplus (V \otimes_R W),\ (u, v) \otimes w \mapsto (u \otimes w, v \otimes w).$$

Proof We prove (i). Construct a homomorphism of abelian groups

$$\begin{aligned} \alpha : & U \otimes_R (V \oplus W) \to (U \otimes_R V) \oplus (U \otimes_R W), \\ & u \otimes (v, w) \mapsto (u \otimes v, u \otimes w) \end{aligned}$$

using the diagram



The inverse map

$$\begin{aligned} \beta : & (U \otimes_R V) \oplus (U \otimes_R W) \to U \otimes_R (V \oplus W), \\ & (\sum_i u_i \otimes v_i, \sum_j u_j \otimes w_j) \mapsto \sum_i u_i \otimes (v_i, 0) + \sum_j u_j \otimes (0, w_j) \end{aligned}$$

comes as a direct sum of two maps

$$\beta_1 : U \otimes_R V \to U \otimes_R (V \oplus W),\ \sum_i u_i \otimes v_i \mapsto \sum_i u_i \otimes (v_i, 0)$$

and

$$\beta_2 : U \otimes_R W \to U \otimes_R (V \oplus W),\ \sum_j u_j \otimes w_j \mapsto \sum_j u_j \otimes (0, w_j),$$

which, in turn, exist by the universal property of tensor products. □

Remark 3.16.15 The theorem above is actually true for a direct sum of an arbitrary family of modules and not only a direct sum of two modules. The 'same' proof works.

Theorem 3.16.16 (Commutativity of Tensor Product) *If R is a commutative ring and V, W are R-modules, then there is an R-isomorphism*

$$V \otimes_R W \tilde{\to} W \otimes_R V,\ v \otimes w \mapsto w \otimes v.$$

Proof (By now) a trivial exercise. □

We now treat an important special case.

Theorem 3.16.17 *Assume that R is commutative and V and W are free R-modules with bases $(v_i)_{i\in I}$ and $(w_j)_{j\in J}$, respectively. Then $V \otimes_R W$ is a free R-module with basis $(v_i \otimes w_j)_{i\in I, j\in J}$.*

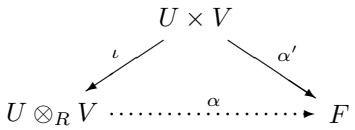
Proof Let F be a free module on the symbols $(v_i \otimes w_j)_{i\in I, j\in J}$. The map

$$\beta : F \to V \otimes_R W,\ v_i \otimes w_j \mapsto v_i \otimes w_j$$

exists by the universal property of free R-modules. The inverse map

$$\alpha : V \otimes_R W \to F,\ v_i \otimes w_j \mapsto v_i \otimes w_j$$

comes from the diagram



where the map α' is defined via

$$\alpha'((\sum_i r_i v_i, \sum_j s_j w_j)) = \sum_{i,j} r_i s_j v_i \otimes w_j.$$

□

Remark 3.16.18 We can now show that in general not every element of a tensor product can be written as a pure tensor. For example let V and W be 2-dimensional vector spaces over a field F with bases $\{v_1, v_2\}$ and $\{w_1, w_2\}$. We claim that $v_1 \otimes w_1 + v_2 \otimes w_2$ can not be written as a pure tensor. If you think this is obvious, be careful! For example, $v_1 \otimes w_1 + v_1 \otimes w_2$ does not look like a pure tensor but it is equal to one: $v_1 \otimes w_1 + v_1 \otimes w_2 = v_1 \otimes (w_1 + w_2)$. So, assume that $v_1 \otimes w_1 + v_2 \otimes w_2$ can be written as a pure tensor:

$$v_1 \otimes w_1 + v_2 \otimes w_2 = v \otimes w. \tag{3.19}$$

Write $v = c_1 v_1 + c_2 v_2$ and $w = d_1 w_1 + d_2 w_2$ for $c_i, d_i \in F$. Then

$$v \otimes w = c_1 d_1 v_1 \otimes w_1 + c_1 d_2 v_1 \otimes w_2 + c_2 d_1 v_2 \otimes w_1 + c_2 d_2 v_2 \otimes w_2.$$

By the theorem above, $\{v_1 \otimes w_1, v_1 \otimes w_2, v_2 \otimes w_1, v_2 \otimes w_2\}$ is a basis of $V \otimes_F W$. So (3.19) leads to a contradiction.

We now consider some useful examples of tensor products.

Example 3.16.19 Let V be a $\mathbb{Z}$-module and m be a positive integer. Then

$$V \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z} \cong V/mV. \tag{3.20}$$

Indeed, one can use the universal property to write down a homomorphism

$$\alpha : V \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z} \cong V/mV,\ v \otimes (n + m\mathbb{Z}) \mapsto nv + mV,$$

and the inverse map is given by $v + mV \mapsto v \otimes (1 + m\mathbb{Z})$. It is easy to deduce from (3.20) that

$$\mathbb{Z}/n\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}/\mathrm{GCD}(m, n)\mathbb{Z}. \tag{3.21}$$

For example, $\mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z} = 0$. If this is hard to believe, I highly recommend to see directly why for example a pure tensor

$$\bar{1} \otimes \bar{1} \in \mathbb{Z}/2\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/3\mathbb{Z}$$

is zero. Another useful exercise is to convince yourself that

$$\mathbb{Z}/m\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Q} = 0.$$

We conclude this section with another important tensor product construction. If A and B are two R-algebras for a commutative ring R (see Definition 3.9.1), then the tensor product $A \otimes_R B$ also has a structure of an R-algebra with multiplication satisfying

$$(a_1 \otimes b_1)(a_2 \otimes b_2) = (a_1 a_2) \otimes (b_1 b_2) \qquad (a_1, a_2 \in A,\ b_1, b_2 \in B).$$

Details are left as an exercise.

Remark 3.16.20 Explain Hopf algebra idea here...

3.17 Problems on Modules

Problem 3.17.1 Let K be an infinite extension of a field k (e.g. $K = k(x)$). Let R be the ring of all 2×2 upper triangular matrices

$$\begin{pmatrix} \alpha & \beta \\ 0 & c \end{pmatrix}$$

with $\alpha, \beta \in K$ and $c \in k$. Show that R is left artinian and left noetherian but is neither right artinian or right noetherian.

Problem 3.17.2 Let V be a left R-module and $W \subseteq V$ be a submodule. Then V satisfies D.C.C. if and only if W and V/W do.

Problem 3.17.3 If R is left artinian and V is a finitely generated left R-module then V satisfies D.C.C.

Problem 3.17.4 Assume that a left R-module V is written as a finite sum of its submodules:

$$V = \sum_{i=1}^{n} V_i.$$

Show that V is noetherian (resp. artinian) if and only if so is every V_i.

Problem 3.17.5 Let k be a field and V be a vector space over k with an infinite basis $\{e_0, e_1, e_2, \dots\}$, and $R = \mathrm{End}_k(V)$. Let $r, s \in R$ be the elements defined by $r(e_{2n}) = e_n, r(e_{2n+1}) = 0$ and $s(e_{2n+1}) = e_n, s(e_{2n}) = 0$ for all n. Prove that $\{r, s\}$ is a basis of ${}_R R$.

Problem 3.17.6 True or false? Let $V = W \oplus X$ and $V = W \oplus Y$

be decompositions of a left R-module V as direct sums of submodules. Then $X = Y$.

Problem 3.17.7 Let G be a finite group, F a field of characteristic p dividing $|G|$, and FG the group algebra. Consider the 1-dimensional submodule of the left regular module ${}_{FG}FG$ spanned by the element $\sum_{g\in G} g$. Show that this submodule is not a direct summand of the regular module.

Solution. Denote $V := F\sum_{g\in G} g$ and assume that X is a complement to V in FG. Let $\sum_{h\in G} a_h h \in X$. As X is a submodule, we have $(\sum_{g\in G} g)\cdot(\sum_{h\in G} a_h h) \in X$. Note however that

$$(\sum_{g\in G} g)\cdot(\sum_{h\in G} a_h h) = (\sum_{h\in G} a_h)(\sum_{g\in G} g) \in V.$$

As $V \cap X = 0$ it follows that $\sum_{h\in G} a_h = 0$. Now by a dimension argument X must consist of all elements of the form $\sum_{h\in G} a_h h$ with $\sum_{h\in G} a_h = 0$. However, if p divides the order of G, then $\sum_{g\in G} g \in X$, giving a contradiction.

Problem 3.17.8 True or false?

(a) Let R be a commutative ring and $I, J \lhd R$ be two ideals of R. If the modules R/I and R/J are isomorphic then $I = J$.

(b) Let R be a ring and I, J be two left ideals in R. If the modules R/I and R/J are isomorphic then $I = J$.

Solution. (a) This is **true!** Note that $\mathrm{Ann}(R/I) = I$, $\mathrm{Ann}(R/J) = J$ and annihilators of isomorphic modules coincide. (If you gave a counterexample, please return to it and find out what was wrong).

(b) This is false. For example, the matrix algebra $R = M_2(\mathbb{C})$ is a direct sum of two left ideals $I_1 \oplus I_2$, where I_1 consists of the matrices with zeros in the second column and I_2 consists of the matrices with zeros in the first column. Now, $R/I_1 \cong I_2$ and $R/I_2 \cong I_1$. It remains to observe that $I_1 \cong I_2$.

Problem 3.17.9 Let $R = \mathbb{C}[[x]]$, the ring of formal power series over $\mathbb{C}$. Consider the submodule W of the free module $V = Rv_1 \oplus Rv_2$ generated by

$$(1-x)^{-1}v_1 + (1-x^2)^{-1}v_2 \quad \text{and} \quad (1+x)^{-1}v_1 + (1+x^2)^{-1}v_2.$$

Find a basis $\{v_1', v_2'\}$ of V and elements $\delta_1 \mid \delta_2 \in R$ such that W is generated by $\delta_1 v_1'$ and $\delta_2 v_2'$. Describe V/W.

Solution. Note that R is a P.I.D. Applying (row 2) - (row 1)$\times(1+x)^{-1}$ to the matrix

$$\begin{pmatrix} (1-x)^{-1} & (1+x)^{-1} \\ (1-x^2)^{-1} & (1+x^2)^{-1} \end{pmatrix}$$

yields the matrix

$$\begin{pmatrix} (1-x)^{-1} & (1+x)^{-1} \\ 0 & (1+x^2)^{-1} - (1+x)^{-2} \end{pmatrix}.$$

Applying (column 2) - (column 1)$\times(1-x)(1+x)^{-1}$ yields

$$\begin{pmatrix} (1-x)^{-1} & 0 \\ 0 & (1+x^2)^{-1} - (1+x)^{-2} \end{pmatrix}.$$

Thus the elementary divisors are $\delta_1 = (1+x)^{-1}$ and $\delta_2 = (1+x^2)^{-1} - (1+x)^{-2}$ (note that $\delta_1 \mid \delta_2$ as δ_1 is a unit. Moreover,

$$\begin{pmatrix} (1-x)^{-1} & 0 \\ 0 & (1+x^2)^{-1} - (1+x)^{-2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -(1+x)^{-1} & 1 \end{pmatrix} \times \begin{pmatrix} (1-x)^{-1} & (1+x)^{-1} \\ (1-x^2)^{-1} & (1+x^2)^{-1} \end{pmatrix} \begin{pmatrix} 1 & -(1-x)(1+x)^{-1} \\ 0 & 1 \end{pmatrix}.$$

Hence

$$(v_1', v_2') = (v_1, v_2) \begin{pmatrix} 1 & 0 \\ -(1+x)^{-1} & 1 \end{pmatrix} = (v_1 - (1+x)^{-1} v_2, v_2).$$

Finally

$$\begin{aligned} V/W &\cong R/((1-x)^{-1}) \oplus R/((1+x^2)^{-1} - (1+x)^{-2}) \\ &\cong (0) \oplus R/(x) \cong R/(x) \cong \mathbb{C}. \end{aligned}$$

Problem 3.17.10 True or False? Every finitely generated module over a commutative noetherian ring has a composition series.

Solution. False: take $R = \mathbb{Z}$.

Problem 3.17.11 True or False? If R is a commutative ring then any submodule of a free module is free.

Solution. False: consider $\mathbb{Z}/3\mathbb{Z} \subset {}_{\mathbb{Z}/6\mathbb{Z}}\mathbb{Z}/6\mathbb{Z}$.

Problem 3.17.12 True or False? If R is a domain then any submodule of a free module is free.

Solution. False: consider $(x, y) \subset F[x, y]$ and use the fact that there is a non-trivial relation $y \cdot x - x \cdot y = 0$.

Problem 3.17.13 True or False? If R is a PID then any submodule of a finitely generated free R-module is free.

Problem 3.17.14 True or False? If R is commutative and every submodule of a free R-module is free then R is a PID.

Solution. True. Clearly R is a domain. Assume that there is an ideal I which is free R-module of rank > 1. Let $\{v_1, v_2, \dots\}$ be a basis. Then $v_1 v_2 - v_2 v_1 = 0$ is a relation, giving a contradiction.

Problem 3.17.15 True or False? If R is a PID and I is a proper ideal in R then R/I is a PID.

Solution. False: does not have to be a domain.

Problem 3.17.16 True or False? If R is a PID then any subring of R is a PID.

Solution. False: $\mathbb{Z}[x] \subset \mathbb{Q}[x]$.

Problem 3.17.17 True or False? If R is a PID then $R[x]$ is a PID.

Solution. False: take $R = \mathbb{Z}$.

Problem 3.17.18 True or False? If R is artinian then $R[x]$ is artinian.

Solution. False: take $R = \mathbb{Q}$.

Problem 3.17.19 True or False? Any finitely generated torsion free module over a PID is free.

Solution. True.

Problem 3.17.20 Let R be a PID. Calculate $\mathrm{Hom}_R((R/(a), R/(b))$.

Solution. The homomorphism space is isomorphic to $R/(GCD(a, b))$.

This can be seen by decomposing $R/(a)$ and $R/(b)$ into primary components and showing that $\mathrm{Hom}_R(R/(p^k), R/(p^l)) \cong R/(p^{\min(k,l)})$ for an irreducible element $p \in R$.

Problem 3.17.21 Let R be a PID, and V be an R-module. A submodule $W \subseteq V$ is called *pure* if $W \cap rV = rW$ for all $r \in R$. If V is a finitely generated R-module, prove that a submodule $W \subseteq V$ is a pure submodule if and only if W is a direct summand of V.

Solution. First of all it is clear that a direct summand of V is pure. Conversely, let $W \subseteq V$ be a pure submodule. Then V/W is torsion free. So W contains the torsion $T(V)$ of V, and it suffices to show that the pure submodule $W/T(V)$ of the free module $V/T(V)$ is a direct summand. Thus we may assume that V is free. Then by a theorem there exist a basis $\{v_1, \ldots, v_n\}$ of V such that $\delta_1 v_1, \ldots, \delta_k v_k$ is a basis of W. As V/W is torsion-free, it follows that all $\delta_i = 1$, and so W is a direct summand.

Problem 3.17.22 Calculate the invariant factors of the following matrices, working over the ring $\mathbb{Z}[i]$ of Gaussian integers:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 2+i \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 2i & i & 2+i \\ i-1 & 1+i & 0 \\ 0 & 0 & 2+i \\ 1+i & -1 & 2+i \end{pmatrix}.$$

Problem 3.17.23 Let

$$A = \begin{pmatrix} -4 & -6 & 7 \\ 2 & 2 & 4 \\ 6 & 6 & 15 \end{pmatrix}.$$

Find unimodular matrices X and Y over $\mathbb{Z}$ such that ZAY has the form $\mathrm{diag}(\delta_1, \delta_2, \delta_3)$ with $\delta_1 \mid \delta_2 \mid \delta_3$.

Problem 3.17.24 True or false? Let R be a PID and V be a finitely generated R-module with invariant factors $\delta_1 \mid \delta_2 \mid \ldots \mid \delta_k$. Then V cannot be generated by less than k elements.

Solution. True. If the module is generated by $m < k$ elements then it is a quotient of the free module of rank m. By Theorem 3.7.7, V will have less than k invariant factors, which contradicts the fact that the invariant factors of a module are well-defined.

Problem 3.17.25 Let R be a commutative ring. Is it true that the left regular R-module is indecomposable? What if R is a PID?

Problem 3.17.26 True or false? Let R be a PID and A be an $n \times n$ matrix over R. Then A is invertible if and only if A is equivalent to the identity matrix.

Problem 3.17.27 Find the invariant factors and the primary decomposition for the $\mathbb{Z}$-module $\mathbb{Z}/2000\mathbb{Z}$.

Problem 3.17.28 Let p be a prime and V be a finitely generated $\mathbb{Z}$-module with $p^a V = 0$. Suppose that $v \in V$ has order exactly p^a. Show that $V = Rv \oplus W$ for some submodule $W \subseteq V$.

Solution. Decompose V according to Theorem 3.7.9:

$$V = \oplus_{i=1}^{n} \mathbb{Z}/(p^{a_i}).$$

It follows from the assumption $p^a V = 0$ that all prime powers appearing in this decomposition are indeed powers of p. Moreover, at least one of a_i must equal to a, as otherwise the module would not have elements of order p^a. Now, let $v = v_1 + v_2 + \cdots + v_n$ with v_i in the component $\mathbb{Z}/(p^{a_i})$. As v has order exactly p^a, at least one of the non-zer components v_i must correspond to $a_i = a$. We may assume without loss of generality that it is V_1. Thus, $v_1 \neq 0$ and $a_1 = a$. Now

$$W := \oplus_{i=2}^{n} \mathbb{Z}/(p^{a_i})$$

will do the job.

Problem 3.17.29 True or false? Let V be a finitely generated torsion-free $\mathbb{Q}[x]$-module, and $\theta \in \mathrm{End}_{\mathbb{Q}[x]}(V)$ be surjective. Then θ is injective.

Solution. True for any PID R. Indeed, the assumption implies that V is free of finite rank m. Now consider the kernel W of θ and use Theorem 3.7.7 to see that either V/W has torsion or it is free of rank smaller than m. In both cases we cannot have $V/W \cong V$.

Problem 3.17.30 True or false? Let V be a finitely generated torsion-free $\mathbb{Q}[x]$-module, and $\theta \in \mathrm{End}_{\mathbb{Q}[x]}(V)$ be injective. Then θ is surjective.

Solution. False. Multiplication by t in a free module of rank 1 is injective but not surjective.

Problem 3.17.31 True or false? A (not necessarily finitely generated) torsion-free module over a PID is free.

Problem 3.17.32 True or false? Let $V \subsetneq W$ be a proper containment of free $\mathbb{Z}$-modules. Then $\text{rank}_{\mathbb{Z}}(V) < \text{rank}_{\mathbb{Z}}(W)$.

Solution. False: $2\mathbb{Z} \subset \mathbb{Z}$.

Problem 3.17.33 True or false? If $V \subseteq W$ are free $\mathbb{Z}$-modules of equal finite rank, then W/V is a finite group.

Problem 3.17.34 True or false? The $\mathbb{C}[x,y]$-modules $\mathbb{C}[x,y]/(x,y)$ and $\mathbb{C}[x,y]/(x-1,y-1)$ are isomorphic.

Problem 3.17.35 Classify, up to isomorphism, all $\mathbb{Q}[t]$-modules V which are annihilated by $(t^3-2)(t-2)^3$ and satisfy $\dim_{\mathbb{Q}} V = 5$.

Solution. Let $R = \mathbb{Q}[t]$, $X = R/(t^3-2)$, and $W_i = R/((t-2)^i)$ for $1 \leq i \leq 3$. By the Fundamental Theorem on finitely generated modules over a PID, the only possibilities for V are: $X \oplus W_2$, $X \oplus W_1 \oplus W_1$, $W_3 \oplus W_2$, $W_3 \oplus W_1 \oplus W_1$, $W_2 \oplus W_2 \oplus W_1$, $W_2 \oplus W_1 \oplus W_1 \oplus W_1$, $W_1^{\oplus 5}$.

Problem 3.17.36 How many abelian groups are there of order $5^6 \cdot 7^5$ (up to isomorphism)?

Problem 3.17.37 Find the isomorphism classes of abelian groups of order 108 having exactly 4 subgroups of order 6.

Solution. The groups are as follows: $C_{27} \times C_4$, $C_{27} \times C_2 \times C_2$, $C_9 \times C_3 \times C_4$, $C_9 \times C_3 \times C_2 \times C_2$, $C_3 \times C_3 \times C_3 \times C_4$, $C_3 \times C_3 \times C_3 \times C_2 \times C_2$. A subgroup of order 6 is $C_3 \times C_2$, and C_3, C_2 must be in the appropriate Sylow subgroups. Now, C_{27} has one subgroup of order 3, while $C_9 \times C_3$ has four, and $C_3 \times C_3 \times C_3$ has thirteen. Similarly, C_4 has just one subgroup of order 2, and $C_2 \times C_2$ has three. Thus, the only possibility is $C_9 \times C_3 \times C_4$.

Problem 3.17.38 True or false? If V is a noetherian module over a ring then any surjective endomorphism of V is bijective.

Solution. True. Let K be the kernel of our surjective homomorphism. Then $V/K \cong V$. If $K \neq 0$ we construct an infinite strictly ascending chain of submodules, which gives a contradiction. To get the chain, we

start with $W_0 = K$. Then use the correspondence theorem for submodules to get a submodule W_1 of V containing W_0 and such that $W_1/W_0 \cong K$. As $V/W_1 \cong V/K \cong V$, we can use the correspondence theorem again to get $W_2 \supset W_1$ with $W_2/W_1 \cong K$. Continuing like this we get the desired chain.

Problem 3.17.39 Let V and W be simple left R-modules. Suppose there exist non-zero elements $v \in V$ and $w \in W$ such that (v, w) generates a proper submodule of $V \oplus W$. Then $V \cong W$.

Problem 3.17.40 If V_1 and V_2 are non-isomorphic simple R-modules, then the R-module $V_1 \oplus V_2$ is cyclic.

Problem 3.17.41 If V_1 and V_2 are non-isomorphic simple R-modules, then $V_1 \oplus V_2$ has exactly four submodules: 0, $0 \oplus V_2$, $V_1 \oplus 0$, and $V_1 \oplus V_2$.

Solution. Let W be a non-zero submodule of $V_1 \oplus V_2$. If W is not one of the four modules mentioned above, then W contains a vector (v_1, v_2) with both v_1 and v_2 non-zero. But the module generated by (v_1, v_2) has both V_1 and V_2 as its composition factors. (To see this, project to V_1 and V_2 and observe that both projections are of course non-zero.) Hence both V_1 and V_2 are composition factors of W. Now, W must equal $V_1 \oplus V_2$ as otherwise $V_1 \oplus V_2$ will have more than two composition factors.

Problem 3.17.42 True or false? Let V be a left R-module and $V_1 \neq V_2$ be maximal submodules. Then $V/(V_1 \cap V_2) \cong V/V_1 \oplus V/V_2$.

Problem 3.17.43 Let R be a ring and V be an R-module. Set $E := \mathrm{End}_R(V)$, the endomorphism ring of V.

(a) If V is the direct sum of two non-trivial R-submodules, show that E contains an idempotent $e \neq 0, 1$.

(b) Suppose that all zero divisors of E lie in a proper ideal J of E. Show that V is indecomposable.

Problem 3.17.44 True or false? Let R be a subring of $M_n(\mathbb{Q})$ which is finitely generated as a $\mathbb{Z}$-module. Then R is free as a $\mathbb{Z}$-module.

Problem 3.17.45 True or false? If V is a left R-module and $\mathrm{End}_R(V)$ is a division ring, then V is simple.

Solution. False. For example prove that $\mathrm{End}_{\mathbb{Z}}(\mathbb{Q}) \cong \mathbb{Q}$.

Problem 3.17.46 True or false? Every finitely generated module over $\mathbb{R}[x, y, z]/(x^2 - y^3, y^2 + z^2)$ is noetherian.

Problem 3.17.47 True or false? If G be a finite abelian group then the group algebra $\mathbb{Q}G$ is a domain.

Solution. False: if g is an element of order n, then

$$(1 - g)(1 + g + \cdots + g^{n-1}) = 0.$$

Problem 3.17.48 True or false: A commutative ring is left semisimple if and only if it is isomorphic to a direct sum of finitely many fields.

Solution. True by Wedderburn-Artin.

Problem 3.17.49 Prove that $J(R)$ contains no non-zero idempotent.

Problem 3.17.50 True or false? In a ring R the set of nilpotent elements is an ideal.

Problem 3.17.51 True or false? If R is a commutative semisimple ring and $r \in R$ with $r^2 = 0$ then $r = 0$.

Solution. True. By Weddreburn-Artin (or Problem 3.17.48) R is a direct sum of fields, and the result is clear.

Problem 3.17.52 R is a semisimple finite ring with $|R| = 4$. What are the possibilities for R.

Problem 3.17.53 True or false? For every left noetherian ring R, the Jacobson radical $J(R)$ is the largest nilpotent ideal of R.

Problem 3.17.54 True or false? If R is a noetherian commutative ring then $R/J(R)$ is a semisimple ring (i.e. every R-module is semisimple).

Problem 3.17.55 Calculate the Jacobson radical of the ring $\mathbb{Z}/m\mathbb{Z}$.

Solution. Decompose $m = p_1^{a_1} \dots p_k^{a_k}$ where p_i's are distinct primes and all $a_i > 0$. I claim that $J(\mathbb{Z}/m\mathbb{Z})$ is the principal ideal I generated by the element $p_1 p_2 \dots p_k + m\mathbb{Z}$. To prove this, observe that the

ideal I coincides with the set of nilpotent elements in $\mathbb{Z}/m\mathbb{Z}$ and use Lemma 3.13.1.

Problem 3.17.56 Let $e \in R$ be a non-zero idempotent. Then eRe is a ring (with identity), and $J(eRe) = J(R) \cap eRe$

Solution. Hint: first prove that when V is a simple R-module, eV is either zero or simple R-module. Then use Proposition 3.12.3.

Problem 3.17.57 Let R be a ring. Show that $J(M_n(R)) = M_n(J(R))$.

Solution. Hint: Observe that if V is a simple R-module, then $V^{\oplus n}$ is a simple $M_n(R)$-module in a natural way. Then use Proposition 3.12.3. You may also want to use Problem 3.17.56: observe that $R = eM_n(R)e$ where $e = E_{1,1}$.

Problem 3.17.58 Let V be a finitely generated left R-module, and $\pi : V \to V/J(R)V$ be the projection. If $\pi(v_1), \ldots, \pi(v_n)$ generate $V/J(R)V$, then $v_1, \ldots, v_n$ generate V.

Problem 3.17.59 True or false? $J(R_1 \oplus \cdots \oplus R_n) = J(R_1) \oplus \cdots \oplus J(R_n)$.

Problem 3.17.60 Let F be a field and $R = F[x]/I$ where I is the ideal of $F[x]$ generated by $x^2 + 2x + 1$. What is the Jacobson radical of R.

Problem 3.17.61 Prove the following generalization of Nakayama's lemma: if R is an arbitrary ring, I a two-sided ideal of R contained in the Jacobson radical of R, and V a left finitely generated R-module such that $IV = V$, then $V = 0$.

Problem 3.17.62 True or false? If V is an irreducible R-module, the center of $\mathrm{End}_R(V)$ is a field.

Problem 3.17.63 Say all you can about an artinian ring containing no non-zero nilpotent elements.

Problem 3.17.64 True or false? If R is a left artinian ring with no zero divisors, then R is a division ring.

Problem 3.17.65 True or false? Let F be a field. A finite dimensional F-algebra without zero divisors is a division algebra.

Problem 3.17.66 True or false? If a ring R is simple artinian then the ring of all 2×2 matrices over R is simple artinian.

Problem 3.17.67 Let G be a finite group and F be an algebraically closed field of characteristic p. Prove the following:

(i) Up to isomorphism, that there are only finitely many irreducible FG-modules $L_1, \ldots, L_k$.

(ii) Let $d_i = \dim L_i$, $1 \leq i \leq k$. Then $\sum_{i=1}^{k} d_i^2 \leq |G|$, and the equality holds if and only if $p \not| |G|$.

(iii) Is it true that the inequality $\sum_{i=1}^{k} d_i^2 \leq |G|$ holds even if F is not algebraically closed?

Solution. The irreducible modules for FG are pullbacks of the irreducible modules over $FG/J(FG)$, which is a finite dimensional semisimple algebra. In particular it has only finitely many irreducibles for example by Wedderburn-Artin. This gives (i). Also by Wedderburn-Artin, we have $\dim(FG/J(FG)) = \sum_{i=1}^{k} d_i^2$, which gives (ii) using Maschke's Theorem. . Finally, (iii) is false: for example C_3 has a simple module of dimension 2 over $\mathbb{Q}$.

Problem 3.17.68 Let C_n be the cyclic group of order n. Decompose the group algebra $\mathbb{C}C_n$ as a direct sum of simple ideals. Do the same for $\mathbb{Q}C_n$.

Problem 3.17.69 True or false? Let A be a finite dimensional abelian algebra over an algebraically closed field F. Then all simple A-modules are 1-dimensional.

Problem 3.17.70 True or false? Let A be a finite dimensional abelian algebra over a field F. Then all simple A-modules are 1-dimensional.

Problem 3.17.71 Let A be a finite dimensional algebra over an algebraically closed field F, and V be a simple A-module.

(i) Show that V is finite dimensional.

(ii) Let $\{v_1, \ldots, v_n\}$ be a basis of v, and $\rho(a)$ be the matrix of a linear transformation $V \to V$, $v \mapsto av$ with respect to this basis. Show that for any matrix $A \in M_n(F)$ there exists $a \in A$ with $A = \rho(a)$.

(iii) Show that (ii) may fail if F is not algebraically closed.

Solution. (i) follows from the fact that every simple A-module is a quotient of A.

(ii) Either use Schur Lemma and Jacobson's Density Theorem or Wedderburn-Artin for algebras over algebraically closed fields (see Remark 3.11.16).

(iii) Consider $\mathbb{C}$ as an $\mathbb{R}$-algebra and the regular $\mathbb{C}$-module ${}_{\mathbb{C}}\mathbb{C}$.

Problem 3.17.72 A finite ring with no nilpotent elements is a direct product of fields.

Problem 3.17.73 Say all you can about a finite simple ring.

Solution. $M_n(\mathbb{F}_q)$ (by Wedderburn-Artin and Wedderburn's theorem that finite division rings are fields).

Problem 3.17.74 True or false? If A is a finite dimensional simple algebra over $\mathbb{C}$, then $\dim_{\mathbb{C}} A$ is a perfect square.

Problem 3.17.75 Let R be a ring and J be its Jacobson radical. Then R/J is a division ring if and only if R has a unique maximal left ideal.

Problem 3.17.76 True or false? If R is a semisimple artinian ring with $r^3 = r$ for all $r \in R$, then R is a division ring.

Problem 3.17.77 True or false? If A and B are semisimple complex algebras of dimension 4 then $A \cong B$.

Problem 3.17.78 True or false? If A and B are semisimple complex algebras of dimension 3 then $A \cong B$.

Problem 3.17.79 If R is a left semisimple ring with $r^2 = r$ for all $r \in R$ then R is isomorphic to a direct product of copies of $\mathbb{F}_2$.

Problem 3.17.80 True or false? If R is an artinian ring having no non-zero nilpotent elements then R is a direct sum of fields.

Problem 3.17.81 Classify all 2-dimensional $\mathbb{R}$-algebras.

Problem 3.17.82 Determine whether or not the matrices

$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & -1 \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & -6 \\ 1 & 0 & 1 \\ 0 & 1 & 4 \end{pmatrix}$$

are similar over rationals.

Problem 3.17.83 Determine whether or not the matrices

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & 1 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

are similar over rationals.

Problem 3.17.84 An $n \times n$ nilpotent matrix with entries in a field has characteristic polynomial x^n.

Problem 3.17.85 True or false? There are exactly 3 similarity classes of 4×4 matrices A over $\mathbb{F}_2$ satisfying $A^2 = 1$.

Problem 3.17.86 Give a list of 2×2 matrices over $\mathbb{F}_2$ such that every 2×2 matrix over $\mathbb{F}_2$ is similar to exactly one on your list.

Problem 3.17.87 Let V be a finite dimensional vector space and let φ, ψ be commuting diagonalizable linear transformations from V to V. Show that φ and ψ can be simultaneously diagonalized.

Problem 3.17.88 Let V be a finite dimensional vector space and let $(\varphi_i)_{i \in I}$ be a family of commuting diagonalizable linear transformations from V to V. Show that φ_i can be simultaneously diagonalized.

Solution. We use induction on $\dim V$, the induction base being $\dim V = 1$. Let $\dim V > 1$. If all linear transformations are scalar, we are done. Otherwise, pick φ_j which is not. Then V decomposes as a direct sum of the corresponding eigenspaces whose dimensions are less than $\dim V$. Each φ_i leaves those eigenspaces invariant (as φ_i commutes with φ_j), and so we can apply inductive assumption.

Problem 3.17.89 Let V be a 7-dimensional vector space over $\mathbb{Q}$.

(a) How many similarity classes of linear transformations on V have characteristic polynomial $(x-1)^4(x-2)^3$?

(b) Of the similarity classe in (a), how many have minimal polynomial $(x-1)^2(x-2)^2$?

(c) Let φ be a linear transformation from V to V having characteristic polynomial $(x-1)^4(x-2)^3$ and minimal polynomial $(x-1)^2(x-2)^2$. Find $\dim \ker(T - 2\,\mathrm{id})$.

Problem 3.17.90 Exhibit a 4×4 matrix A with integer coefficients such that $A^5 = 1 \neq A$.

Problem 3.17.91 Let $A, B \in M_5(\mathbb{Q})$ be non-zero 5×5 matrices over $\mathbb{Q}$ such that $AB = BA = 0$. Prove that if $A^4 = A$ and $B^4 = B^2 - B$, then $A + B$ is invertible.

Problem 3.17.92 Let V be a finite dimensional vector space over a field F. Let θ be a linear transformation on V. Assume that there do *not* exist proper, θ-invariant, subspaces V_1, V_2 of V such that $V = V_1 \oplus V_2$. Show that for some basis of V the matrix of θ is the companion matrix of $p(x)^e$, where $p(x)$ is some irreducible polynomial in $F[x]$.

Problem 3.17.93 Let V be a finite dimensional vector space over $\mathbb{Q}$. Let θ be a linear transformation on V having characteristic polynomial $(x-2)^4$.

(a) Describe the possible Jordan normal forms for θ, and for each of these give the minimal polynomial of θ.

(b) For each of the possibilities in (a) give the dimension of the 2-eigenspace of θ.

(c) Assume that θ leaves invariant only finitely many subspaces of V. What can be said about the Jordan normal form of θ.

Problem 3.17.94 Let R be an artinian ring. Show that $J(R)$ is a semisimple left R-module if and only if $J(R)^2 = 0$.

Problem 3.17.95 Suppose that R is a finite dimensional simple F-algebra, for F a field. Show that $\dim_F R = n^2 e$, where $ne = \dim_F V$ for some irreducible R-module V.

Problem 3.17.96 Let F be a field and V be a finite dimensional F-vector space. Let $\theta \in \mathrm{End}_F(V)$ have minimal polynomial $f \in F[x]$. Let R be the subalgebra of $F[\theta]$ of $\mathrm{End}_F(V)$. Prove that R is semisimple if and only if each prime factor of f in $F[x]$ has multiplicity 1.

Problem 3.17.97 True or false? For a short exact sequence

$$0 \to V \to W \to X \to 0,$$

if V and W are indecomposable then so is X.

Problem 3.17.98 Show that if $M_1, M_2, \ldots, M_n$ are distinct maximal ideals of the commutative ring R, then each R-module R/M_i, $1 \leq i \leq n$, is isomorphic to exactly one factor of the chain

$$R \supseteq M_1 \supseteq M_1 \cap M_2 \supseteq \cdots \supseteq M_1 \cap M_2 \cap \cdots \cap M_n.$$

Problem 3.17.99 True or false? If R is a commutative artinian ring then $R[x]$ is noetherian.

Problem 3.17.100 True or false? $\mathbb{Q}$ is an injective $\mathbb{Z}$-module.

Problem 3.17.101 Give three different definitions of projective module and show that your definitions are equivalent.

Problem 3.17.102 Prove that $\mathbb{Q}$ is not projective as a $\mathbb{Z}$-module by showing that it is not a direct summand of a free module.

Problem 3.17.103 Prove that a field F of characteristic 0 is not a projective $\mathbb{Z}$-module under the natural action.

Solution. For every element $\alpha \in F^\times$ and every non-zero integer n there exists an element $\beta \in F^\times$ with $n\beta = \alpha$. This shows that F cannot be a submodule of a free $\mathbb{Z}$-module, and so, of course, it is not a direct summand.

Problem 3.17.104 Let R be a PID. Which cyclic R-modules are projective?

Problem 3.17.105 Let R be commutative. True or false:

(i) Every submodule of a projective R-module is projective.
(ii) Every submodule of an injective R-module is injective.
(iii) Every quotient of a projective R-module is projective.
(iv) Every quotient of an injective R-module is injective.

Problem 3.17.106 Let $n \geq 1$. Then $\mathbb{Z}/n\mathbb{Z}$ is injective $\mathbb{Z}/n\mathbb{Z}$-module.

Solution. We need only to show that if I is an ideal of $\mathbb{Z}/n\mathbb{Z}$ and $\varphi : I \to \mathbb{Z}/n\mathbb{Z}$ is linear, then φ extends to a linear map $\psi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Now, $I = d\mathbb{Z}/n\mathbb{Z}$ for some $d|n$. Let $k \in \mathbb{Z}$ be such that $\varphi(d + n\mathbb{Z}) = k + n\mathbb{Z}$. Since $(n/d)(d + n\mathbb{Z}) = 0$ in I, we have $(n/d)k + n\mathbb{Z} = 0$. Therefore n divides nk/d. It follows that d divides k. Let $e = k/d$. We can define $\psi(x + n\mathbb{Z}) = xe + n\mathbb{Z}$.

Problem 3.17.107 True or false? Every abelian group monomorphism $C_{p^\infty} \to A$ splits.

Problem 3.17.108 True or false? $\mathbb{Q}[x, x^{-1}]$ is a projective $\mathbb{Q}[x]$-module.

Problem 3.17.109 Let R be a ring, and $V \subset W$, $V' \subset W'$ be R-modules such that $W/V \cong W'/V' \cong {}_RR$ and $V \cong V'$, then $W \cong W'$.

Solution. True: since ${}_RR$ is projective, we have $W \cong V \oplus R \cong V' \oplus R \cong W'$.

Problem 3.17.110 Let R be a domain and F be its field of fractions. Prove that F is an injective R-module.

Solution. Let L be an ideal in R, and $\varphi : L \to F$ be an R-module homomorphism. We may assume that $L \neq 0$. Pick any $r \in L \setminus \{0\}$ and extend φ to R by sending $s \in R$ to $s\varphi(r)/r \in F$. To see that the map is well-defined, note that $f(r_1r_2) = r_1f(r_2) = r_2f(r_1)$ implies that $f(r_1)/r_1 = f(r_2)/r_2$. It is easy to see that the map is an R-module map extending φ.

Problem 3.17.111 True or false? $\mathbb{Q}$ is the injective hull of $\mathbb{Z}$.

Problem 3.17.112 Every short exact sequence of $\mathbb{C}[x]/(x^2-1)$-modules is split.

Solution. True: the ring is isomorphic to $\mathbb{C} \times \mathbb{C}$, so we can apply Proposition 3.14.11.

Problem 3.17.113 True or false? Every short exact sequence of $\mathbb{Z}_{15}$-modules is split.

Problem 3.17.114 True or false? Every $\mathbb{R}[x]$-module is projective.

Solution. No, torsion module over a PID is not projective.

Problem 3.17.115 True or false? Every projective module over a commutative ring is free.

Solution. False: $\mathbb{C}$ over $\mathbb{C} \times \mathbb{C}$.

Problem 3.17.116 Let F be a free left R-module with finite basis.

Prove that F^* is a free right R-module with finite basis dual to the basis of F. Prove that F is reflexive.

Problem 3.17.117 Let P be a finitely generated projective left R-module. Prove that P^* is a finitely generated projective right R-module. Prove that P is reflexive. Demonstrate that both statements are false if we drop the assumption of P being finitely generated.

Solution. Since P is finitely generated, there is a surjective map $F \twoheadrightarrow P$, where F is a free module with finite basis. Then $F \cong P \oplus P'$, as P is projective. Note that $F^* \cong P^* \oplus P'^*$, whence P^* is projective and finitely generated, as F^* is free with finite basis by Problem 3.17.116. Moreover, let $\theta : F \to F^{**} = P^{**} \oplus P'^{**}$ be a canonical homomorphism, which is an isomorphism by Problem 3.17.116 again. Then $\theta(P) \subseteq P^{**}$ and $\theta(P') \subseteq P'^{**}$. It follows that the restriction $\theta\,|_P \colon P \to P^{**}$ is an isomorphism.

As for counterexamples, take an $\mathbb{F}_2$-vector space V on an infinite countable basis. Then $\mathbb{F}_2^{**}$ does not have a countable basis. Finally, consider a free $\mathbb{Z}$-module $F := \oplus_{i=1}^{\infty}\mathbb{Z}$ of infinite countable rank. We claim that $F^* = \prod_{i=1}^{\infty}\mathbb{Z}$ is not projective. Well, let $\varphi : F^* \to \mathbb{Z}_2$ be a map which sends a tuple $(n_1, n_2, \dots) \in \prod_{i=1}^{\infty}\mathbb{Z}$ to $\bar{1}$ if all n_i are odd, and to $\bar{0}$ otherwise. Now, if $\pi : \mathbb{Z} \to \mathbb{Z}_2$ is the natural surjection, then there no map $\psi : F^* \to \mathbb{Z}$ with $\pi \circ \psi = \varphi$. Indeed, if such ψ existed, then $\psi((1,1,1,\dots))$ must be odd and $\psi((1,0,0,\dots))$ is even, but $\psi((2,1,1,\dots))$ is even, giving a contradiction.

Problem 3.17.118 Let V be a finite dimensional vector space over a division ring D. Prove that V is reflexive.

Problem 3.17.119 Prove that a domain R is a field if and only if every R-module is projective.

Solution. If R is a field, we know that every R-module (i.e. a vector space over R) has a basis. This means that every R-module is free, and hence projective. Conversely, assume that every R-module is projective. By Proposition 3.14.11, R is semisimple artinian. By Wedderburn-Artin R is a field. (Another proof: take any non-zero $r \in R$. To show that r is invertible consider an epimomorphism $\theta : {}_RR \to Rr,\ s \mapsto sr$. As Rr is projective, the epimomorphism splits, i.e. there exists a homomorphism $\varphi : Rr \to {}_RR$ such that $\theta \circ \varphi = \mathrm{id}_{Rr}$. Using the fact that R is a domain, it

follows that $\varphi(r) = 1$. Now, if $xr = 0$, we have $0 = \varphi(xr) = x\varphi(r) = x$, as required.)

Problem 3.17.120 Let R be a subring of the ring S such that S is free with basis $(s_i)_{i\in I}$ when considered as a right R-module. If V is a left R-module, then, as abelian groups,

$$S \otimes_R V = \bigoplus_{i\in I} s_i \otimes V,$$

where $s_i \otimes V$ denotes the subspace of $S \otimes_R V$ generated by all pure tensors of the form $s_i \otimes v$.

Problem 3.17.121 Let V and W be $\mathbb{Z}$-modules (equivalently, abelian groups).

(i) $V \otimes_{\mathbb{Z}} \mathbb{Z}/m\mathbb{Z} \cong V/mV$.
(ii) $\mathbb{Z}/m\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}/k\mathbb{Z}$, where $k = \mathrm{GCD}(m, n)$.
(iii) Describe $V \otimes_{\mathbb{Z}} W$ if V and W are finitely generated.

Problem 3.17.122 True or false? If V is a torsion module over $\mathbb{Z}$ then $V \otimes_{\mathbb{Z}} \mathbb{Q} = 0$.

Problem 3.17.123 True or false? $\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}$.

Problem 3.17.124 Let V be a right R-module and W be a left R-module. True or false?

(i) There is an isomorphism of abelian groups $V \otimes_R W \cong V \otimes_{\mathbb{Z}} W$.
(ii) If $v \otimes w = v' \otimes w'$ in $V \otimes_R W$, then $v = v'$ and $w = w'$.

Problem 3.17.125 If V' is a submodule of the right R-module V and W' is a submodule of the left R-module W, then $(V/V') \otimes_R (W/W') \cong (V \otimes_R W)/U$, where U is the subgroup of $V \otimes_R W$ generated by all elements oof the form $v' \otimes w$ and $v \otimes w'$ where $v' \in V', v \in V, w' \in W', w \in W$.

Problem 3.17.126 If

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

is a split exact sequence of left R-modules, then

$$0 \longrightarrow X \otimes_R U \xrightarrow{\mathrm{id}_X \otimes f} X \otimes_R V \xrightarrow{\mathrm{id}_X \otimes g} X \otimes_R W \longrightarrow 0$$

is an exact sequence of abelian groups for any right R-module X.

Problem 3.17.127 Prove that a projective module is flat.

Problem 3.17.128

(i) If I is a right ideal of a ring R and V is a left R-module, then there is an isomorphism of abelian groups

$$R/I \otimes_R V \cong V/IV,$$

where IV is the subgroup of V generated by all elements xv with $x \in I$, $v \in V$.

(ii) If R is commutative and I, J are ideals in R, then $R/I \otimes_R R/J \cong R/(I + J)$.

Problem 3.17.129 True or false? If $f : V \to V'$ and $g : W \to W'$ are surjective maps of right and left R-modules respectively, then

$$f \otimes g : V \otimes_R W \to V' \otimes_R W'$$

is surjective.

Problem 3.17.130 Let R be a commutative ring and V, W be R-modules. Show that there exists a homomorphism of R-algebras

$$\theta : \mathrm{End}_R(V) \otimes_R \mathrm{End}_R(W) \to \mathrm{End}_R(V \otimes_R W)$$

such that

$$(\theta(f \otimes g))(v \otimes w) = f(v) \otimes g(w)$$

for all $v \in V, w \in W$, $f \in \mathrm{End}_R(V)$, $g \in \mathrm{End}_R(W)$.

Problem 3.17.131 Suppose $f : R \to S$ is a surjective ring homomorphism. Let V and W be a right and a left S-modules, respectively. Describe how to give V and W a right and a left R-module structures, and prove

$$V \otimes_R W \cong V \otimes_S W.$$

If f is not a surjection, is $V \otimes_R W \cong V \otimes_S W$?

Problem 3.17.132 True or false? If V and W are respectively right and left modules over a division ring D such that $V \otimes_D W = 0$ then either $V = 0$ or $W = 0$.

Problem 3.17.133 True or false? Let R be a commutative ring and V be an R-module. If $L \otimes_R V = 0$ for every simple R-module L then $V = 0$.

Problem 3.17.134 True or false? Let V be a left R-module. If

$$X \otimes_R V = 0$$

for every right R-module X then $V = 0$.

Problem 3.17.135 True or false? If V, V_1, V_2 are non-zero free modules of finite rank over a commutative ring R with $V \otimes_R V_1 \cong V \otimes_R V_2$, then $V_1 \cong V_2$.

Problem 3.17.136 True or false? $\mathbb{Z}_3 \otimes_{\mathbb{Z}} (\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}_2) \cong \mathbb{Z}_6$.

Problem 3.17.137 Find $(\mathbb{Q} \oplus \mathbb{Z}_7) \otimes_{\mathbb{Z}} \mathbb{Z}_5$.

Problem 3.17.138 True or false? $\mathbb{Z}_{35} \otimes_{\mathbb{Z}} \mathbb{Z}_5 \cong \mathbb{Z}_7$.

Problem 3.17.139 Find $(\mathbb{C} \oplus \mathbb{Z}_6) \otimes_{\mathbb{Z}} \mathbb{Z}_3$.

Problem 3.17.140 True or false? Let R be a ring, V be a right R-module and W be a left R-module. Then the additive group $V \otimes_R W$ is a quotient of the abelian group $V \otimes_{\mathbb{Z}} W$.

Problem 3.17.141 Give an example of a ring R, a right R-module V and a left R-module W such that $V \otimes_R W \ncong V \otimes_{\mathbb{Z}} W$ as abelian groups.

Problem 3.17.142 True or false? If $K/\mathbb{Q}$ and $L/\mathbb{Q}$ are finite field extensions then $K \otimes_{\mathbb{Q}} L$ is a semisimple ring.

Problem 3.17.143 Let K/k be a field extension, f be an irreducible polynomial over k, and α be a root of f in some extension field of k. Show that $k(\alpha) \otimes_k K$ is isomorphic to $K[x]/(f)$ as a k-algebra. Deduce that if α is separable over k, then $k(\alpha) \otimes_k K$ is semisimple.

Problem 3.17.144 Let A and B be finite dimensional semisimple algebras over $\mathbb{C}$. Prove that $A \otimes_{\mathbb{C}} B$ is semisimple.

Problem 3.17.145 Let R be a ring for which every simple left R-module is projective. Prove that R is semisimple artinian.

4
Categories and Functors

Category theory unifies concepts from many parts of mathematics. At the first glance it sounds like a boring piece of general nonsense or some kind of naive way of thinking about mathematics in terms of dots and arrows. In reality category theory is a very powerful tool. One of its principal ideas is that of naturality. Everything should be a functor, isomorphisms should be natural (or functorial), etc.

4.1 Categories

The idea of a category is that its object is characterized not as a set, not by describing of which elements it consists, but by its relations with other objects.

Definition 4.1.1 A *category* $\mathcal{C}$ consists of three ingredients:

- A class $\mathrm{Ob}\,\mathcal{C}$ of *objects* of the category.
- A set $\mathrm{Hom}(A,B) = \mathrm{Hom}_{\mathcal{C}}(A,B)$ of *morphisms* from A to B, for every ordered pair (A,B) of objects. If $f \in \mathrm{Hom}(A,B)$ we often write $f : A \to B$ or $A \xrightarrow{f} B$ and call A the *domain* and B the *target* of f.
- A *composition* map

$$\mathrm{Hom}(A,B) \times \mathrm{Hom}(B,C) \to \mathrm{Hom}(A,C),\ (f,g) \mapsto gf,$$

  for every ordered triple A, B, C of objects.

These ingredients should satisfy the following axioms:

(i) The Hom sets are pairwise disjoint; that is each morphism has a unique domain and a unique target.

(ii) For each object A, there is an identity morphism $\mathrm{id}_A \in \mathrm{Hom}(A, A)$ such that $f\,\mathrm{id}_A = f$ and $\mathrm{id}_B\, f = f$ for all $f : A \to B$.

(iii) Composition is associative, that is whenever

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D,$$

then $h(gh) = (hg)f$.

Remark 4.1.2 If you are confused by the word *class* in the definition above substitute for it the word *set* and read remarks in the beginning of section 7.2 of Rotman.

Example 4.1.3 (i) The category **Sets** of sets has all sets as its objects, maps between sets as morphisms and usual composition of maps as composition of morphisms.

(ii) The category **Groups**. Here objects are groups, morphisms are group homomorphisms.

(iii) The category **Ab**. Here objects are *abelian* groups, morphisms are group homomorphisms.

(iv) **Top** is the category of topological spaces and continuous maps.

(v) Let R be a fixed ring. The category R-Mod is the category of all left R-modules and R-module homomorphisms. Similarly Mod-R is the category of all right R-modules.

(vi) Let X be a partially ordered set, for example the set of all subgroups of a fixed group G ordered by inclusion. The category $\mathbf{PO}(X)$ has all elements of X as its objects. As for morphisms, we set $\mathrm{Hom}(x, y) = \varnothing$ if $x \not\leq y$. On the other hand, if $x \leq y$, then $\mathrm{Hom}(x, y) = \{f_{xy}\}$, that is there is exactly one morphism from x to y for each pair $x \leq y$. Finally, the composition is defined via $f_{yz} f_{xy} := f_{xz}$ for $x \leq y \leq z$.

(vii) Let G be a monoid (a 'group where we do not insist that every element has inverse element'). Define the category $\mathcal{C}(G)$ as follows: $\mathcal{C}(G)$ has only one object, call it $*$. Now, $\mathrm{Hom}_{\mathcal{C}(G)}(*, *) := G$ with the composition being the group multiplication.

It is clear how to translate the notion of isomorphism into a categorical language.

Definition 4.1.4 A morphism $f : A \to B$ in a category $\mathcal{C}$ is called an *isomorphism* if there exists a morphism $g : B \to A$ in $\mathcal{C}$ such that $gf = \mathrm{id}_A$ and $fg = \mathrm{id}_B$. The morphism g is called the *inverse* of f and denoted f^{-1}.

Remark 4.1.5 (i) The inverse morphism is defined uniquely: if g and g' are both inverse to f, then multiplyng the equality $fg = \mathrm{id}_B$ on the left by g' and using equalities $\mathrm{id}_A\, g = g$ and $g'\,\mathrm{id}_B = g'$, we get $g = g'$.

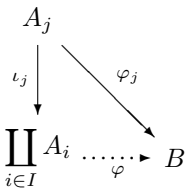
(ii) Clearly id_A is an isomorphism for any object A.

(iii) We describe isomorphisms in the categories from Example 4.1.3: (i) set bijections; (ii) and (iii) group isomorphisms, (iv) homeomorphisms; (v) module isomorphisms; (vi) id_x for each $x \in X$; (vii) invertible elements of G (in particular, if G is a group then all morphisms are isomorphisms).

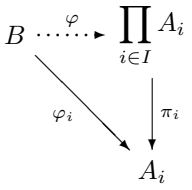
You will recognize the following definitions as generalizations of direct sum and direct product of modules.

Definition 4.1.6 Let $(A_i)_{i\in I}$ be a family of objects of a category $\mathcal{C}$ indexed by a set I.

(i) A *coproduct* of the family consists of an object $\coprod_{i\in I} A_i$ and a family of morphisms $(\iota_j : A_j \to \coprod_{i\in I} A_i)_{j\in I}$ such that for any object B and a family of morphisms $(\varphi_j : A_j \to B)_{j\in I}$ there exists a unique morphism $\varphi : \coprod_{i\in I} A_i \to B$ making the following diagram commutative for each $j \in I$:



(ii) A *coproduct* of the family consists of an object $\prod_{i\in I} A_i$ and a family of morphisms $(\pi_j : \prod_{i\in I} A_i \to A_j)_{j\in I}$ such that for any object B and a family of morphisms $(\varphi_j : B \to A_j)_{j\in I}$ there exists a unique morphism $\varphi : B \to \prod_{i\in I} A_i$ making the following diagram commutative for each $j \in I$:



First an expected uniqueness result:

Lemma 4.1.7 *Let $(A_i)_{i\in I}$ be a family of objects of a category $\mathcal{C}$ indexed by a set I.*

(i) *Assume that*

$$(\coprod_{i\in I} A_i, (\iota_j : A_j \to \coprod_{i\in I} A_i)_{j\in I})$$

and

$$(\coprod_{i\in I}{}' A_i, (\iota'_j : A_j \to \coprod_{i\in I}{}' A_i)_{j\in I})$$

are two coproducts. Then there exists an isomorphism

$$\varphi : \coprod_{i\in I} A_i \xrightarrow{\sim} \coprod_{i\in I}{}' A_i$$

such that the following diagram is commutative for every $j \in I$:

$$\begin{array}{ccc} & A_j & \\ {\scriptstyle \iota_j}\swarrow & & \searrow{\scriptstyle \iota'_j} \\ \coprod_{i\in I} A_i & \xrightarrow{\quad\varphi\quad} & \coprod_{i\in I}{}' A_i \end{array}$$

(ii) *Dual statement for products which I am lazy to type in.*

Proof Usual stuff. If you are still not comfortable with that please come to the office hours. □

What might be surprising, is that products and coproducts might not exist in certain categories.

Example 4.1.8

(i) In the category of sets coproduct is given by disjoint union and product by the cartesian product.

(ii) In the category of R-modules products are direct products and coproducts are direct sums. Similar constructions work for the category of abelian groups. As for the category of all groups, direct product works as a categorical product, and the categorical coproduct is given by the *free product*, not a direct sum! Also note that if you try infinite families of objects in the category of *finite groups* for example, you will be in trouble.

(iii) In the category $\mathbf{PO}(X)$ we have $x \coprod y$ is the least upper bound for x and y, and $x \prod y$ is the greatest lower bound of x and y, if they exist.

(iv) In the category of fields products and coproducts do not exist: think about what kind of field would need to exist if we wanted it to be a product of $\mathbb{F}_2$ and $\mathbb{F}_3$.

Definition 4.1.9 We say that a category $\mathcal{B}$ is a *subcategory* of a category $\mathcal{C}$, written $\mathcal{B} \subseteq \mathcal{C}$, if $\mathrm{Ob}\,\mathcal{B} \subseteq \mathrm{Ob}\,\mathcal{C}$ and for every pair A, B of objects in $\mathcal{B}$, we have $\mathrm{Hom}_{\mathcal{B}}(A, B) \subseteq \mathrm{Hom}_{\mathcal{C}}(A, B)$. A subcategory $\mathcal{B} \subseteq \mathcal{C}$ is called a *full subcategory* if in fact $\mathrm{Hom}_{\mathcal{B}}(A, B) \subseteq \mathrm{Hom}_{\mathcal{C}}(A, B)$ for every $A, B \in \mathrm{Ob}\,\mathcal{B}$.

Note that to specify a full subcategory of $\mathcal{C}$ it is sufficient to specify a subclass of objects.

Example 4.1.10 Finite groups and group homomorphisms is a full subcategory of **Groups**. The category of fields is a full subcategory of the category of rings.

4.2 Functors

If we want to 'compare' different categories we want to consider some kind of maps between categories:

Definition 4.2.1 Let $\mathcal{A}, \mathcal{B}$ be two categories. A *covariant functor* $F : \mathcal{A} \to \mathcal{B}$ is

(i) a rule that assigns to every object $A \in \mathcal{A}$ and object $F(A) \in \mathcal{B}$;

(ii) a rule that assigns to each morphism $f : A_1 \to A_2$ in $\mathcal{A}$ a morphism $F(f) : F(A_1) \to F(A_2)$ in $\mathcal{B}$.

These rules should satisfy the following axioms:

(a) $F(f \circ g) = F(f) \circ F(g)$ for all morphisms f, g in $\mathcal{A}$ for which the composition $f \circ g$ makes sense.

(b) $F(\mathrm{id}_A) = \mathrm{id}_{F(A)}$ for every object A of $\mathcal{A}$.

A *contravariant functor* $F : \mathcal{A} \to \mathcal{B}$ is the same as a covariant functor except that $F(f)$ is now a morphism from $F(A_2)$ to $F(A_1)$ (i.e. F 'inverts the arrows') and axiom (a) becomes

(a') $F(f \circ g) = F(g) \circ F(f)$ for all morphisms f, g in $\mathcal{A}$ for which the composition $f \circ g$ makes sense.

Often I will just write 'functor' for 'covariant functor'. By the way, the difference between covariant and contravariant functors is similar to the difference between left and right modules. We can even reduce the study of contravariant functors to covariant functors by introducing the notion of the *opposite category* $\mathcal{C}^{\mathrm{op}}$ to the category $\mathcal{C}$: it is the same as category $\mathcal{C}$, except that all arrows are turned around. Then a contravariant functor $\mathcal{A} \to B$ is the same as a covariant functor $\mathcal{A} \to \mathcal{B}^{\mathrm{op}}$. (Boring) details are left to the reader.

Example 4.2.2 (i) The *identity functor* $\mathrm{id}_{\mathcal{C}}$ on a category $\mathcal{C}$—guess what it is and check the axioms.

(ii) *Forgetful functor*: **Groups** $\to$ **Sets** associates to every group the underlying set and to every group homomorphism the same map between the corresponding underlying sets. This is a covariant functor (a rather stupid one) and you can come up with a lot of examples like that.

(iii) The *duality functor* D on the category F-**Vect** of vector spaces over a fixed field F assigns to every vector space V its dual V^*, and to every linear map $f : V \to W$ of vector spaces its dual map $f^* : W^* \to V^*$. Thus $D(V) = V^*$ and $D(f) = f^*$. This is a very nice example of a contravariant functor.

(iv) Let R be a ring and X be a fixed left R-module. Then we have a covariant functor

$$\mathrm{Hom}_R(X, -) : R\text{-}\mathbf{Mod} \to \mathbf{Ab}$$

and a contravariant functor

$$\mathrm{Hom}_R(-, X) : R\text{-}\mathbf{Mod} \to \mathbf{Ab}.$$

For example, the first functor maps an R-module Y to the abelian group $\mathrm{Hom}_R(X, Y)$ and a morphism $f : Y_1 \to Y_2$ to the morphism $f_* : \mathrm{Hom}_R(X, Y_1) \to \mathrm{Hom}_R(X, Y_2)$, see §3.15.

(v) Let R be a ring and X be a fixed right R-module. Then we have a covariant functor

$$X \otimes - : R\text{-}\mathbf{Mod} \to \mathbf{Ab}$$

which maps a left R-module Y to the abelian group $X \otimes_R Y$ and a morphism $f : Y_1 \to Y_2$ of left R-modules to a morphism $\mathrm{id}_X \otimes f$, see §3.16.

Lemma 4.2.3 *If $F : \mathcal{A} \to \mathcal{B}$ is a functor and $f : A_1 \to A_2$ is an isomorphism in $\mathcal{A}$ then $F(f) : F(A_1) \to F(A_2)$ is also an isomorphism.*

Proof If $g = f^{-1}$ apply F to $fg = \mathrm{id}$ and $gf = \mathrm{id}$. □

You might think this is crazy but we not only want 'maps' between categories—we also want 'maps' between functors! This will allow us to consider things like a category of functors from a category $\mathcal{A}$ to a category $\mathcal{B}$.

Definition 4.2.4 Let $F, G : \mathcal{A} \to \mathcal{B}$ be two functors. A *natural transformation* $\alpha : F \to G$ is a rule that assigns to every object $A \in \mathcal{A}$ a morphism $\alpha_A : F(A) \to G(A)$ in $\mathcal{B}$ such that for every morphism $f : A_1 \to A_2$ in $\mathcal{A}$ the following diagram is commutative:

$$\begin{array}{ccc} F(A_1) & \xrightarrow{F(f)} & F(A_2) \\ \Big\downarrow{\scriptstyle \alpha_{A_1}} & & \Big\downarrow{\scriptstyle \alpha_{A_2}} \\ G(A_1) & \xrightarrow{G(f)} & G(A_2) \end{array}$$

A natural transformation $\alpha : F \to G$ is called a *natural isomorphism* of functors if each α_A is an isomorphism in the category $\mathcal{B}$.

Example 4.2.5 (i) Let DD be the *double duality* functor F-**Vect** and id be the identity functor on F-**Vect**. For every vector space V there is a natural homomorphism $\alpha_V : V \to V^{**}$. It is easy to see that the $\alpha'_V s$ define a natural transformation $\alpha : \mathrm{id} \to DD$. Moreover, if we restrict ourselves to the full subcategory of *finite dimensional* vector spaces, α will become an isomorphism of functors.

(ii) Let id be the identity functor on R-**Mod**. We claim that id is naturally isomorphic to $\mathrm{Hom}_R(R, -)$. To check this we need to go back to Lemma 3.15.8 and check the 'naturality' (make sure you understand what this means). People would usually state the isomorphism of functors using the following words: the isomorphism of left R-modules

$$\mathrm{Hom}_R({}_R R_R, V) \tilde{\to} V,\ \theta \mapsto \theta(1_R).$$

is *natural* or *functorial*.

(iii) Using the terminology explained in (ii), we can now state stronger versions of many results above: for example for left R-modules V, W, X, the isomorphisms of abelian groups

$$\mathrm{Hom}_R(V \oplus W, X) \cong \mathrm{Hom}_R(V, X) \oplus \mathrm{Hom}_R(W, X)$$

coming from Corollary 3.2.4 is natural; the isomorphisms in 3.16.8 are natural, the isomorphisms in the associativity, additivity, and commutativity of tensor products are all natural.

Now we will try to come with a notion of an 'isomorphism of categories'. Let $\mathcal{A}$ and $\mathcal{B}$ be two categories. They are called *isomorphic* if there exist functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ such that $F \circ G = \mathrm{id}_{\mathcal{B}}$ and $G \circ F = \mathrm{id}_{\mathcal{A}}$. I should disappoint you right away: the isomorphism of categories is useless notion, because it almost never occurs in 'real life' (I should not dare to speak about real life after the definition of a natural transformation of functors above). For example, let us consider the category $\mathcal{A}$ of finite dimensional vector spaces over a fixed field $\mathbb{F}$. Now, if there is any justice in the world, the duality functor D from Example 4.2.2(iii) should be an isomorphism from $\mathcal{A}$ to $\mathcal{A}$, and its inverse should be D itself because V and V^{**} are naturally isomorphic. However, observe that $D \circ D$ is not $\mathrm{id}_{\mathcal{A}}$, because V is not the *same* vector space as V^{**}. This leads us to the following 'correct' definition

Definition 4.2.6 Two categories $\mathcal{A}$ and $\mathcal{B}$ are called *equivalent* if there exist functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ such that $F \circ G \cong \mathrm{id}_{\mathcal{B}}$ and $G \circ F \cong \mathrm{id}_{\mathcal{A}}$. In this case we say that F and G are *quasi-inverse* to each other.

Now it is easy to see that the duality functor D on finite dimensional F-vector spaces establishes a self-equivalence of categories with the inverse equivalence being D itself.

We now address the following question: given a functor $F : \mathcal{A} \to \mathcal{B}$, when does it define an equivalence of categories. In other words, what are the conditions on F which guarantee the existence of the quasi-inverse functor G? The answer is given in the useful Theorem 4.2.8.

Definition 4.2.7 A functor $F : \mathcal{A} \to \mathcal{B}$ is called *faithful* if the map

$$\mathrm{Hom}_{\mathcal{A}}(A_1, A_2) \to \mathrm{Hom}_{\mathcal{B}}(F(A_1), F(A_2)),\ \theta \mapsto F(\theta) \tag{4.1}$$

is injective, and F is called *full* if the map (4.1) is surjective.

Theorem 4.2.8 *A functor $F : \mathcal{A} \to \mathcal{B}$ is an equivalence of categories if and only if the following two conditions hold:*

(i) *F is full and faithful;*
(ii) *every object of $\mathcal{B}$ is isomorphic to an object of the form $F(A)$ for some $A \in \mathrm{Ob}\,\mathcal{A}$.*

Proof ($\Rightarrow$) Let F be an equivalence of categories and $G : \mathcal{B} \to \mathcal{A}$ be the quasi-inverse functor. Let $\alpha : GF \to \mathrm{id}_{\mathcal{A}}$ and $\beta : FG \to \mathrm{id}_{\mathcal{B}}$ be isomorphisms of functors. First of all, for any object B of $\mathcal{B}$ $\beta_B : F(G(B)) \to B$ is an isomorphism, which gives (ii). Next, for each $\varphi \in \mathrm{Hom}_{\mathcal{A}}(A_1, A_2)$ we have the commutative diagram

$$\begin{array}{ccc} GF(A_1) & \xrightarrow{\alpha_{A_1}} & A_1 \\ {\scriptstyle GF(\varphi)}\downarrow & & \downarrow{\scriptstyle \varphi} \\ GF(A_2) & \xrightarrow{\alpha_{A_2}} & A_2 \end{array}$$

Hence φ can be recovered from $F(\varphi)$ by the formula

$$\varphi = \alpha_{A_2} \circ GF(\varphi) \circ (\alpha_{A_1})^{-1}. \tag{4.2}$$

This shows that F is faithful. Similarly, G is faithful. To prove that F is full, consider an arbitrary morphism $\psi \in \mathrm{Hom}_{\mathcal{B}}(F(A_1), F(A_2))$, and set

$$\varphi := \alpha_{A_2} \circ G(\psi) \circ (\alpha_{A_1})^{-1} \in \mathrm{Hom}_{\mathcal{A}}(A_1, A_2).$$

Comparing this with (4.2) and taking into account that α_{A_1} and α_{A_2} are isomorphisms, we deduce that $G(\psi) = GF(\varphi)$. As G is faithful, this implies that $\psi = F(\varphi)$, which completes the proof that F is a full functor.

($\Leftarrow$) Assume that (i) and (ii) hold. In view of (i), we can (and will) identify the set $\mathrm{Hom}_{\mathcal{B}}(F(A_1), F(A_2))$ with the set $\mathrm{Hom}_{\mathcal{A}}(A_1, A_2)$ for any $A_1, A_2 \in \mathrm{Ob}\,\mathcal{A}$. Using (ii), for each object B in $\mathcal{B}$ we can pick an object A_B in $\mathcal{A}$ and an isomorphism $\beta_B : F(A_B) \to B$. We define a functor $G : \mathcal{B} \to \mathcal{A}$ which will turn out to be a quasi-inverse functor to F. on the objects we set $G(B) = A_B$ for any $B \in \mathrm{Ob}\,\mathcal{B}$. To define G on the morphisms, let $\psi \in \mathrm{Hom}_{\mathcal{B}}(B_1, B_2)$.

$$\begin{aligned} G(\psi) := \beta_{B_2}^{-1} \circ \psi \circ \beta_{B_1} \in & \mathrm{Hom}_{\mathcal{B}}(FG(B_1), FG(B_2)) \\ & = \mathrm{Hom}_{\mathcal{A}}(G(B_1), G(B_2)). \end{aligned}$$

It is easy to see that G is a functor, and $\beta = \{\beta_B\} : FG \to \mathrm{id}_{\mathcal{B}}$ is an isomorphism of functors. Further, $\beta_{F(A)} = F(\alpha_A)$ for the unique morphism $\alpha_A : GF(A) \to A$. Finally, it is not hard to see that $\alpha = \{\alpha_A\} : GF \to \mathrm{id}_{\mathcal{A}}$ is an isomorphism of functors. $\square$

4.3 Adjoint functors

Definition 4.3.1 Given functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$, we say that F is *left adjoint* to G (or G is *right adjoint* to F), if for each pair of objects $A \in \mathcal{A}$ and $B \in \mathcal{B}$ there are bijections

$$\alpha_{A,B} : \mathrm{Hom}_{\mathcal{B}}(FA, B) \tilde{\to} \mathrm{Hom}_{\mathcal{A}}(A, GB)$$

that are natural transformations in A and B.

The naturality in A in Definition 4.3.1 means that for every morphism $f \in \mathrm{Hom}_{\mathcal{A}}(A_1, A_2)$ and every object B in $\mathcal{B}$ the following diagram is commutative:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{B}}(FA_2, B) & \xrightarrow{(Ff)^*} & \mathrm{Hom}_{\mathcal{B}}(FA_1, B) \\ \Big\downarrow{\scriptstyle \alpha_{A_2,B}} & & \Big\downarrow{\scriptstyle \alpha_{A_1,B}} \\ \mathrm{Hom}_{\mathcal{A}}(A_2, GB) & \xrightarrow{f^*} & \mathrm{Hom}_{\mathcal{A}}(A_1, GB) \end{array}$$

The naturality in B in Definition 4.3.1 means that for every morphism $g \in \mathrm{Hom}_{\mathcal{B}}(B_1, B_2)$ and every object A in $\mathcal{A}$ the following diagram is commutative:

$$\begin{array}{ccc} \mathrm{Hom}_{\mathcal{B}}(FA, B_1) & \xrightarrow{g_*} & \mathrm{Hom}_{\mathcal{B}}(FA, B_2) \\ \Big\downarrow{\scriptstyle \alpha_{A,B_1}} & & \Big\downarrow{\scriptstyle \alpha_{A,B_2}} \\ \mathrm{Hom}_{\mathcal{A}}(A, GB_1) & \xrightarrow{(Gg)_*} & \mathrm{Hom}_{\mathcal{A}}(A, GB_2) \end{array}$$

Let V be an (R, S)-bimodule. The following theorem is (a slightly stronger version of) the statement that the functor

$$V \otimes_S - : S\textbf{-Mod} \to R\textbf{-Mod} \tag{4.3}$$

is left adjoint to the functor

$$\mathrm{Hom}_R(V, -) : R\textbf{-Mod} \to S\textbf{-Mod}. \tag{4.4}$$

Theorem 4.3.2 (Adjointness of $\otimes$ and Hom) *Let R and S be rings. Given an (R, S)-bimodule V, a left S-module U, and a left R-module W, there exists an isomorphism of abelian groups*

$$\mathrm{Hom}_R(V \otimes_S U, W) \cong \mathrm{Hom}_S(U, \mathrm{Hom}_R(V, W))$$

natural in U and W.

Proof This is one of the most instructive proofs in the whole course (and, boy, does it use lots of parenthesis)! We proceed in several steps.

1. Define a homomorphism of abelian groups

$$\alpha = \alpha_{U,W} : \mathrm{Hom}_R(V \otimes_S U, W) \to \mathrm{Hom}_S(U, \mathrm{Hom}_R(V, W))$$

via

$$\big[\alpha(f)(u)\big](v) := f(v \otimes u) \qquad (u \in U, v \in V, f \in \mathrm{Hom}_R(V \otimes_S U, W)).$$

2. We check that $\alpha(f)(u) \in \mathrm{Hom}_R(V, W)$:

$$\begin{aligned}\big[\alpha(f)(u)\big](rv) &= f(rv \otimes u) = f(r(v \otimes u)) \\ &= rf(v \otimes u) = r\big[\alpha(f)(u)\big](v).\end{aligned}$$

3. We check that $\alpha(f) \in \mathrm{Hom}_S(U, \mathrm{Hom}_R(V, W))$:

$$\begin{aligned}\big[\alpha(f)(su)\big](v) &= f(v \otimes su) = f(vs \otimes u) \\ &= \big[\alpha(f)(u)\big](vs) = \big(s\big[\alpha(f)(u)\big]\big)(v).\end{aligned}$$

4. We check that α is a homomorphism of abelian groups: note that

$$\alpha(f_1 - f_2) = \alpha(f_1) - \alpha(f_2)$$

if and only if

$$\alpha(f_1 - f_2)(u) = \alpha(f_1)(u) - \alpha(f_2)(u)$$

for every $u \in U$, which in turn holds if and only if

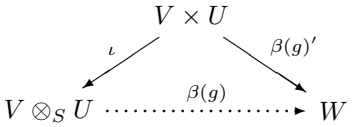
$$\big[\alpha(f_1 - f_2)(u)\big](v) = \big[\alpha(f_1)(u)\big](v) - \big[\alpha(f_2)(u)\big](v)$$

for every $u \in U$ and every $v \in V$. Now apply the definition of α.

5. To show that α is an isomorphism we want to define the inverse map β. Let $g \in \mathrm{Hom}_S(U, \mathrm{Hom}_R(V, W))$. The map

$$\beta(g)' : V \times U \to W,\ (v, u) \mapsto g(u)(v)$$

is checked to be S-biadditive (check!). Considering the diagram



yields the homomorphism of abelian groups

$$\beta(g) : V \otimes_S U \to W,\ v \otimes u \mapsto g(u)(v) \qquad (v \in V, u \in W).$$

6. We check that $\beta(g)$ is an R-homomorphism:

$$\begin{aligned}\beta(g)(r(v\otimes u)) &= \beta(g)(rv\otimes u) = g(u)(rv) = r(g(u)(v))\\ &= r\beta(g)(v\otimes u)\end{aligned}$$

7. Now we define the map

$$\beta = \beta_{U,V} : \mathrm{Hom}_S(U, \mathrm{Hom}_R(V,W)) \to \mathrm{Hom}_R(V\otimes_S U, W),\ g\mapsto \beta(g).$$

It is easy to see that β is a homomorphism of abelian groups.

8. We check that $\beta\circ\alpha = \mathrm{id}$. Let $f\in \mathrm{Hom}_R(V\otimes_S U, W)$. Then

$$\beta(\alpha(f))(v\otimes u) = \big[\alpha(f)(u)\big](v) = f(v\otimes u).$$

9. We check that $\alpha\circ\beta = \mathrm{id}$. Let $g\in \mathrm{Hom}_S(U, \mathrm{Hom}_R(V,W))$. Then

$$\big[\alpha(\beta(g))(u)\big](v) = \beta(g)(v\otimes u) = g(u)(v).$$

10. We check the naturality of α in W. Let $\theta : W_1 \to W_2$ be a homomorphism of left R-modules. We need to show that the following diagram commutes:

$$\begin{array}{ccc}
\mathrm{Hom}_R(V\otimes_S U, W_1) & \xrightarrow{\theta_*} & \mathrm{Hom}_R(V\otimes_S U, W_2)\\
\Big\downarrow{\scriptstyle \alpha_{U,W_1}} & & \Big\downarrow{\scriptstyle \alpha_{U,W_2}}\\
\mathrm{Hom}_S(U, \mathrm{Hom}_R(V,W_1)) & \xrightarrow{(G\theta)_*} & \mathrm{Hom}_S(U, \mathrm{Hom}_R(V,W_2))
\end{array}$$

where $G = \mathrm{Hom}_R(V,-)$. Well, let us take $f\in \mathrm{Hom}_R(V\otimes_S U, W_1)$. Then $\theta_*(f) = \theta\circ f$. So

$$\big[\alpha_{U,W_2}(\theta_*(f))\big](u)(v) = \theta(f(v\otimes u)).$$

On the other hand

$$\big[(G\theta)_*(\alpha_{U,W_1}(f))\big](u)(v) = \theta\big(\big[\alpha_{U,W_1}(f)\big](u)(v)\big) = \theta(f(v\otimes u)).$$

11. We check the naturality of α in W. Well, why don't *you* check it? (I guarantee a lot of excitement). □

There is one special case of Theorem 4.3.2 which is used particularly often. Assume that S is a subring of R, U is a left S-module and W is a left R-module. We have the *induction* and *restriction* functors

$$\mathrm{ind}_S^R : S\textbf{-Mod} \to R\textbf{-Mod},\ U\mapsto R\otimes_S U,$$

which is a special case of (4.3) and

$$\mathrm{res}\,{}^{R}_{S} : R\textbf{-Mod} \to S\textbf{-Mod}.$$

The restriction functor gives the S-structure to an R-module W by 'restricting' the action from R to S. Now Theorem 4.3.2 gives

Corollary 4.3.3 (Adjointness of ind **and** res **or Frobenius Reciprocity)** *Let S be a subring of a ring R. Given a left S-module U, and a left R-module W, there exists an isomorphism of abelian groups*

$$\mathrm{Hom}_R(\mathrm{ind}\,{}^{R}_{S}U, W) \cong \mathrm{Hom}_S(U, \mathrm{res}\,{}^{R}_{S}W)$$

natural in U and W.

4.4 Problems on Categories and Functors

Problem 4.4.1 Work in the category of abelian groups. Prove that the cartesian product $G \times H$ is both categorical product and categorical coproduct.

Problem 4.4.2 In the category $\mathbf{PO}(X)$, $x \coprod y$ is the least upper bound for x and y, and $x \prod y$ is the greatest lower bound of x and y, if they exist.

Problem 4.4.3 True or false? The functor $\mathbb{Z}_5 \otimes_{\mathbb{Z}} - : \mathbb{Z}\text{-Mod} \to \mathbb{Z}\text{-Mod}$ is exact.

Problem 4.4.4 True or false? The functor

$$\mathrm{Hom}_{\mathbb{Z}}(-, \mathbb{Q}) : \mathbb{Z}\text{-Mod} \to \mathbb{Z}\text{-Mod}$$

is exact.

Problem 4.4.5 If the functors $F : \mathcal{A} \to \mathcal{B}$ and $G : \mathcal{B} \to \mathcal{A}$ establish an equivalence of categories between $\mathcal{A}$ and $\mathcal{B}$ then F is both left and right adjoint to G.

Problem 4.4.6 In this problem I start the description of a functor. You need to finish it by saying what the functor is on the morphisms and check the axioms:

(a) A covariant functor from **Groups** to **Sets** that assigns to each group the set of all its subgroups.

(b) A covariant functor from **Rings** to **Rings** that assigns to each ring R the polynomial ring $R[x]$.

(c) A covariant functor from **Groups** to **Groups** that assigns to each group G its commutator subgroup G'.

Problem 4.4.7 Let $F, G : \mathcal{A} \to \mathcal{B}$ be covariant functors, and $\alpha : F \to G$ be a natural trasformation. Then α is a natural isomorphism if and only if there exists a natural transformation $\beta : G \to F$ such that $\beta\alpha = \mathrm{id}_F$ and $\alpha\beta = \mathrm{id}_G$.

5
Commutative algebra

From now on, all rings are commutative unless otherwise stated (and unital, as usual).

5.1 Noetherian rings

Theorem 5.1.1 (Hilbert Basis Theorem) *If the ring R is noetherian, so is the polynomial $R[x]$.*

Proof Let I be an ideal of $R[x]$. We have to show that it is finitely generated.

For each $n \geq 0$, let J_n be the set of all $r \in R$ for which there exists a polynomial $f(x) \in I$ of degree $\leq n$ in which the coefficient of x^n is r. It is easy to see that J_n is an ideal of R, as I is an ideal in $R[x]$. Moreover, $J_n \subseteq J_{n+1}$, as $xf \in I$. As R is noetherian, the ascending sequence $J_1 \subseteq J_2 \subseteq \ldots$ of ideals in R terminates, i.e. there is m such that $J_i = J_m$ for all $i \geq m$.

Moreover, each ideal $J_1, \ldots, J_m$ is finitely generated. Let $S_1, \ldots, S_m$ be the corresponding generating sets. For each $s \in S_i$, let $g_s(x) \in I$ be a polynomial of degree $\leq i$ in which the coefficient of x^i is s. We show that I coincides with the ideal K generated by the finite set

$$\{g_s \mid s \in S_1 \cup \cdots \cup S_m\}.$$

Note that $K \subseteq I$, as all g_s belong to I.

Conversely, we show by induction of the degree that every polynomial $f \in I$ is in K. Induction starts from $f = 0$. Let $f(x) = a_n x^n + \cdots + a_0$ have degree $n \geq 0$. Then $a_n \in J_n$.

If $n \leq m$, then $a_n = r_1 s_1 + \ldots r_k s_k$ for some $r_1, \ldots, r_k \in R$ and

$s_1 \ldots, s_k \in S_n$. Hence $g(x) := r_1 g_{s_1}(x) + \cdots + r_k g_{s_k}(x) \in K$ has degree n and leading coeffivient a_n. Then $\deg(f-g) < n$ and $f - g \in I$. By induction $f - g \in K$, whence $f \in K$.

If $n > m$, then $J_n = J_m$, and so $a_n = r_1 s_1 + \ldots r_k s_k \in R$ for some $r_1, \ldots, r_k$ and $s_1 \ldots, s_k \in S_m$. Hence $g(x) := r_1 g_{s_1}(x) + \cdots + r_k g_{s_k}(x) \in K$ has degree m and leading coeffivient a_n. As above, $f - x^{n-m} g(x) \in K$ be inductive hypothesis, and $f \in K$. □

Corollary 5.1.2 *If R is a noetherian ring, so is the ring of polynomials $R[x_1, x_2, \ldots, x_n]$. In particular, if F is a field, the ring $F[x_1, \ldots, x_n]$ is noetherian.*

Proof Note that $R[x_1, \ldots, x_n][x_n] \cong R[x_1, x_2, \ldots, x_n]$ and apply induction on n. □

Theorem 5.1.3 *If the ring R is noetherian, so is the ring $R[[x]]$ of formal power series.*

Proof The proof is similar to the one of Hilbert Basis Theorem, but uses order instead of degree (the order of the series $\sum_{i=0}^{\infty} r_i x^i$ is the minimal i with $i \neq 0$). We skip the details. □

Corollary 5.1.4 *If R is a noetherian ring, so is the ring of formal power series $R[[x_1, x_2, \ldots, x_n]]$. In particular, if F is a field, the ring $F[[x_1, \ldots, x_n]]$ is noetherian.*

Now we prove some results which are valid for any (commutative) ring.

Theorem 5.1.5 (Chinese Remainder Theorem) *Let $I_1, \ldots, I_n$ be ideals of a ring R. If $I_j + I_k = R$ for every $j \neq k$, then the homomorphism $\varphi : R \to R/I_1 \times \cdots \times R/I_n$ induced by projections $\pi_j : R \to R/I_j$ is surjective. In particular,*

$$R/(I_1 \cap \cdots \cap I_n) \cong R/I_1 \times \cdots \times R/I_n.$$

Proof Let $J_k := \prod_{l \neq k} I_l \lhd R$. We claim that $I_k + J_k = R$. Otherwise $I_k + J_k \subseteq M$ for some prime ideal P (for example a maximal ideal). As $J_k \subseteq P$ and P is prime, we have $I_l \subseteq P$ for some $l \neq k$. Now $I_l + I_k \subseteq P$ contradicts the assumption that $I_l + I_k = R$, proving the

claim. It follows that $\pi_k(J_k) = R/I_k$. On the other hand $\pi_l(J_k) = 0$ if $l \neq k$. It follows that $\varphi(J_1 + \dots J_n)$ is all of $R/I_1 \times \cdots \times R/I_n$. □

Remark 5.1.6 A special case of the theorem states that if $a_1, \dots a_n$ are integers with $(a_j, a_k) = 1$ for every $j \neq k$, then for each n-tuple of integers $b_1, \dots, b_n$ there exists an integer b such that $b \equiv b_i \pmod{a_i}$.

Definition 5.1.7 A subset S of a commutative ring R is called *multiplicative* if $1 \in S$ and $s_1 s_2 \in S$ whenever $s_1, s_2 \in S$. A multiplicative subset is called *proper* if $0 \notin S$.

Lemma 5.1.8 *Let $S \subset R$ be a proper multiplicative set. Let I be an ideal of R satisfying $I \cap S = \varnothing$. The set T of ideals $J \supseteq I$ such that $J \cap S = \varnothing$ has maximal elements, and each maximal element in T is a prime ideal.*

Proof That the set T has maximal elements follows from Zorn Lemma. Let M be such an element. Assume that $x, y \in R \setminus M$. By the choice of M, $M + Rx$ contains some $s_1 \in S$ and $M + Ry$ contains some $s_2 \in S$, i.e. $s_1 = m_1 + r_1 x$ and $s_2 = m_2 + r_2 y$. Hence

$$s_1 s_2 = (m_1 + r_1 x)(m_2 + r_2 y) \in M + Rxy.$$

It follws that $M + Rxy \neq M$, i.e. $xy \notin M$. □

Lemma 5.1.9 *Let $I \lhd R$ be a proper ideal. The intersection of prime ideals containing I is $\{r \in R \mid r^n \in I \text{ for some } n \geq 0\}$. In particular, the intersection of all prime ideals in R coincides with the set of nilpotent elements in R.*

Proof Let J be the intersection of all prime ideals containing I. If $r \in R$ does not belong to J then $r \notin P$ for a prime ideal $P \supseteq I$. But then $x^n \notin P$ for all $n \geq 0$, and so $x^n \notin I$.

Conversely, assume that $r^n \notin I$ for all $n \geq 0$. By Lemma 5.1.8 applied to the multiplicative set $S = \{1 = r^0, r, r^2, \dots\}$, there is a prime ideal P of R, which contains I but none of r^n. □

Definition 5.1.10 The *radical* of an ideal $I \lhd R$ in a commutative ring R is the set

$$\sqrt{I} = \{r \in R \mid r^n \in I \text{ for some } n \geq 0\}.$$

The set $\sqrt{0}$ of all nilpotent elements in R is also sometimes call the *radical* of R and denoted $\operatorname{Rad} R$.

We note that the radical and Jacobson radical of a commutative ring do not have to coincide, although we always have $J(R) \supseteq \operatorname{Rad} R$, see Corollary 3.12.10. This fact also follows from Lemma 5.1.9 and the definition of the Jacobson radical, because maximal ideals are of course prime. On the other hand, Lemma 3.13.1 shows that the two radicals coincide providing the ring is artinian. For a noetherian ring this is already not true, as the example of $R[[x]]$ shows, see Problem 5.6.2.

Lemma 5.1.11 *If $\sqrt{I}$ is finitely generated, then $(\sqrt{I})^n \subseteq I$ for some n.*

Proof Let $\sqrt{I} = Rx_1 + \cdots + Rx_k$. Then $x_i^{n_i} \in I$ for some n_i, $i = 1, \ldots, k$. Now take $n = (n_1 - 1) + \cdots + (n_k - 1) + 1$. □

We now work to obtain an analogue of a prime decomposition in $\mathbb{Z}$ for arbitrary noetherian rings, called *primary decomposition*.

Definition 5.1.12 A proper ideal $Q \triangleleft R$ is called *primary* if in the ring R/Q every zero divisor is nilpotent. Equivalently: $rs \in Q$ and $r \notin Q \Rightarrow s \in \sqrt{Q}$.

The primary ideals should play a role of powers of prime numbers in our primary decomposition as the following example suggests.

Example 5.1.13 An ideal in $\mathbb{Z}$ is primary if and only if it is of the form (p^n) where p is prime and n is a positive integer. Note that $\sqrt{(p^n)} = (p)$. This illustrates the following result.

Lemma 5.1.14 *If $Q \triangleleft R$ is primary then $P = \sqrt{Q}$ is prime.*

Proof Let $rs \in P$ and $r \notin P$. We have $r^n s^n = (rs)^n \in Q$ for some n. Note that $r^n \notin Q$, hence $s^n \in \sqrt{Q}$, i.e. $(s^n)^m \in Q$. Therefore $s \in P$. □

Definition 5.1.15 Let P be a prime ideal. An ideal Q is called *P-primary* (or associated with P) if it is primary and $\sqrt{Q} = P$.

Lemma 5.1.16 *The intersection of finitely many P-primary ideals is P-primary.*

Proof It suffices to prove that the intersection of two P-primary ideals Q_1 and Q_2 is P-primary. First of all, it is clear that

$$\sqrt{Q_1 \cap Q_2} \subseteq \sqrt{Q_1} = P.$$

Conversely, if $r \in P$, then $r^m \in Q_1$ and $r^n \in Q_2$, so $r^{m+n} \in Q_1 \cap Q_2$. Thus $\sqrt{Q_1 \cap Q_2} = P$.

To show that $Q_1 \cap Q_2$ is primary, let $rs \in Q_1 \cap Q_2$ and $r \notin Q_1 \cap Q_2$, then $s \in P = \sqrt{Q_1 \cap Q_2}$, since Q_1 and Q_2 are P-primary. □

Definition 5.1.17 A proper ideal $Q \lhd R$ is called *irreducible* if I cannot be written as the intersection $I = J \cap K$ of ideals $J, K \supsetneq I$.

Example 5.1.18 In $\mathbb{Z}$ the irreducible ideals coincide with primary ideals. Problem 5.6.3 gives an example of an ideal which is primary but not irreducible. On the other hand:

Lemma 5.1.19 *In a noetherian ring every irreducible ideal is primary.*

Proof Let I be irreducible. Assume that $rs \in I$, $s \notin \sqrt{I}$. We need to prove that $r \in I$. Let

$$I_n = \{x \in R \mid xs^n \in I\}.$$

Then $r \in I_1$, and we have an ascending chain of ideals

$$I \subseteq I_1 \subseteq I_2 \subseteq \dots.$$

As R is noetherian, there exists n such that $I_{2n} = I_n$. Let $J = I + Rs^n$. We claim that $J \cap I_n = I$. Indeed, clearly $I \subseteq J \cap I_n$. Conversely, let $x \in J \cap I_n$. We can write $x = t + ys^n$ for some $t \in I$, $y \in R$. We know that $xs^n = ts^n + ys^{2n} \in I$, which implies that $ys^{2n} \in I$, i.e. $y \in I_{2n} = I_n$. But then $ys^n \in I$, and so $x \in I$. Now $J \neq I$, as $s \notin \sqrt{I}$. As I is irreducible, we have $I_n = I$, whence $I_1 = I$, and so $r \in I$. □

Lemma 5.1.20 *In a noetherian ring every proper ideal is the intersection of finitely many irreducible ideals.*

Proof Assume this is false. Let $\mathcal{S}$ be the set of all proper ideals which are not intersections of finitely many irreducible ideals. Since R is noetherian, $\mathcal{S}$ has a maximal element M. Now M is not irreducible, so $M = J \cap K$ for some proper ideals $J \supsetneq M$ and $K \supsetneq M$. By the

choice of M, J and K are not in $\mathcal{S}$, so they are both intersections of finitely many irreducible ideals, and so is M. Contradiction. □

It follows from the last two lemmas that every ideal I of a noetherian ring R is the intersection $I = Q_1 \cap \cdots \cap Q_r$ of finitely many primary ideals. The primary ideals $Q_1, \ldots, Q_r$ are called a *primary decomposition* of I. By throwing redundant elements out of this decomposition we can always achieve the situation when $\cap_{i \neq j} Q_i \subsetneq \cap_i Q_i$, in which case we say that our primary decomposition is *irredundant*. It is now clear that every proper ideal I in a noetherian ring has an irredundant primary decomposition. However, we can do a little better. Note by Lemma 5.1.16 the intersection of all primary ideals in our decomposition which are associated with the same prime ideal P is again P-primary. So we can take this intersection as one new primary ideal in a new, 'shorter' decomposition. By doing this we may achieve that all primary ideals in our primary decomposition are associated with different primes. An irredundant primary decomposition $Q_1 \cap \cdots \cap Q_k$ is called *reduced* if for every $i \neq j$ we have $\sqrt{Q_i} \neq \sqrt{Q_j}$.

Theorem 5.1.21 (Primary Decomposition) *Every ideal of a commutative noetherian ring has a reduced primary decomposition. Moreover, if*

$$I = Q_1 \cap \cdots \cap Q_k = Q'_1 \cap \cdots \cap Q'_l$$

are two reduced primary decompositions of I, then $k = l$ and, up to a permutation, $\sqrt{Q_i} = \sqrt{Q'_i}$.

Proof We only need to prove the uniqueness. We show that in a reduced primary decomposition

$$I = Q_1 \cap \cdots \cap Q_k$$

the distinct prime ideals $P_1 = \sqrt{Q_1}, \ldots, P_k = \sqrt{Q_k}$ coincide with the set of prime ideals which have form

$$I(x) := \{y \in R \mid xy \in I\} \qquad (x \in R \setminus I).$$

Such prime ideals are called the *associated prime ideals* of I.

For any $1 \leq i \leq k$, let $J = \cap_{j \neq i} Q_j$. Then $I = J \cap Q_i \subsetneq J$. By Lemma 5.1.11, $P_i^n \subseteq Q_i$ for some n. Then $JP_i^n \subseteq J \cap Q_i = I$. Let n be minimal such that $JP_i^n \subseteq I$. Then we can take $x \in JP_i^{n-1} \setminus I$. Note

that $x \in J$ and $x \notin Q_i$, as otherwise $x \in J \cap Q_i = I$. Also,

$$I(x) = \{y \in R \mid xy \in I\} = \{y \in R \mid xy \in Q_i\} \subseteq P_i,$$

as Q_i is P_i-primary. On the other hand, $xP_i \subseteq JP_i^n \subseteq I$, whence $P_i \subseteq I(x)$. Thus $P_i = I(x)$ for some $x \notin I$.

Conversely, assume that $P = I(x)$ is a prime ideal where $x \notin I$. Then $x \notin Q_i$ for some i. Let $J = \prod_{x \notin Q_i} Q_i$. Then $xJ \subseteq Q_i$ for all i, as either $x \in Q_i$ or $J \subseteq Q_i$. Therefore $xJ \subseteq I$, i.e. $J \subseteq P$. As P is prime, this implies that $Q_i \subseteq P$ for some i such that $x \notin Q_i$. Then $P_i \subseteq P$. Conversely,

$$I(x) \subseteq \{y \in R \mid xy \in Q_i\} \subseteq P_i,$$

as $x \notin Q_i$ and Q_i is P_i-primary. Thus $P = P_i$. □

5.2 Rings of Quotients and Localization

The quotient field of a domain R can be thought of as a process of inverting all non-zero elements of R. More generally, if R is any commutative ring, we may try to invert only part, say $S \subset R$, of its elements to get a ring of quotients denoted $R[S^{-1}]$ or $S^{-1}R$. This simple idea is actually a powerful tool which reduces many problems in commutative algebra to problems about local rings.

We now give the formal construction. Let S be a multiplicative set in R (see Definition 5.1.7). As a set, the ring $S^{-1}R$ consists of the equivalence classes $[\frac{a}{s}]$ of all 'fractions' $\frac{a}{s}$ with $a \in R$ and $s \in S$. Two 'fractions' $\frac{a}{s}$ and $\frac{b}{t}$ are equivalent is there exists $u \in S$ such that

$$u(at - bs) = 0.$$

We now check that this is indeed an equivalence relation. The reflexivity and symmetricity are obvious, so let $\frac{a_1}{s_1} \sim \frac{a_2}{s_2}$ and $\frac{a_2}{s_2} \sim \frac{a_3}{s_3}$. So there are $u_1, u_2 \in S$ with

$$u_1(a_1s_2 - a_2s_1) = 0 \quad \text{and} \quad u_2(a_2s_3 - a_3s_2) = 0$$

or

$$u_1a_1s_2 = u_1a_2s_1 \quad \text{and} \quad u_2a_2s_3 = u_2a_3s_2.$$

Multiplying the first equality by u_2s_3 and the second equality by u_1s_1, we get

$$u_2s_3u_1a_1s_2 = u_2s_3u_1a_2s_1 \quad \text{and} \quad u_1s_1u_2a_2s_3 = u_1s_1u_2a_3s_2.$$

The right hand side of the first equality and the left hand side of the second equality coincide, so

$$u_1u_2s_2(a_1s_3 - a_3s_1) = u_1u_2s_2a_1s_3 - u_1u_2s_2a_3s_1 = 0,$$

proving that $\frac{a_1}{s_1} \sim \frac{a_3}{s_3}$, as S is multiplicative.

Now define the operations via

$$[\frac{a}{s}][\frac{b}{t}] = [\frac{ab}{st}], \quad [\frac{a}{s}] + [\frac{b}{t}] = [\frac{at+bs}{st}].$$

We need to check that the operations are well-defined. I will explain this for the sum, leaving the product as an exercise. So let $\frac{a_1}{s_1} \sim \frac{a_2}{s_2}$, i.e.

$$u(a_1s_2 - a_2s_1) = 0$$

for some $u \in S$. We need to prove that for any $b \in R$ and $t \in S$ we have

$$\frac{a_1t + bs_1}{s_1t} \sim \frac{a_2t + bs_2}{s_2t}.$$

Well,

$$u((a_1t + bs_1)s_2t - (a_2t + bs_2)s_1t) = ut^2(a_1s_2 - a_2s_1) = 0.$$

Now, it is clear that $S^{-1}R$ is a ring with unit $[\frac{1}{1}]$. Moreover, it is clear that

$$\varphi : R \to S^{-1}R,\ r \mapsto [\frac{r}{1}]$$

is a ring homomorphism. Note also that $\varphi(s)$ is invertible for any $s \in S$. By the way, this shows that $S^{-1}R = 0$ if $0 \in S$. Otherwise, the two properties above can be used to characterize $S^{-1}R$ as a universal object:

Lemma 5.2.1 *Let S be a proper multiplicative subset in a ring R, and $\varphi : R \to S^{-1}R$ be the canonical homomorphism. For any ring homomorphism $f : R \to R'$ such that $f(s)$ is invertible for every $s \in S$ there exists a unique ring homomorhism $\hat{f} : S^{-1}R \to R'$ such that $f = \hat{f} \circ \varphi$. Moreover, this property characterizes $S^{-1}R$ up to an isomorphism.*

Proof Define $\hat{f}([\frac{a}{s}]) = f(a)f(s)^{-1}$. To check that $\hat{f}$ is well-defined, assume that $u(a_1s_2 - a_2s_1) = 0$. Then $f(u)(f(a_1)f(s_2) - f(a_2)f(s_1)) = 0$. As $f(u)$ is invertible by assumption, we have $f(a_1)f(s_2) - f(a_2)f(s_1) = 0$, whence $f(a_1)f(s_1)^{-1} = f(a_2)f(s_2)^{-1}$.

The rest is standard. □

Example 5.2.2

(i) If R is a domain and $S = R \setminus \{0\}$, then $S^{-1}R$ is just the fraction field $F(R)$ of R. For general multiplicative $S \subseteq \setminus\{0\}$, $S^{-1}R$ is the subring of $F(R)$, which consists of all fractions $\frac{r}{s} \in F(R)$ such that $s \in S$.

(ii) Let $f \in R$, and $S = \{1 = f^0, f, f^2, f^3, \dots\}$. Then $S^{-1}R$ is usually denoted by R_f. Note for example that $\mathbb{C}[x]_x \cong \mathbb{C}[x, x^{-1}]$, the ring of Laurent polynomials.

(iii) Let P be a prime ideal in R. Then $S := R \setminus P$ is multiplicative. In this case we usually write R_P for $S^{-1}R$ and call if *localization* of R at P. Let us consider for example the prime ideal (x) in $\mathbb{C}[x]$. By part (i), $\mathbb{C}[x]_{(x)}$ consists of all rational functions in $\mathbb{C}(x)$ which can be written in the form $\frac{f}{g}$ with $g(0) \neq 0$. Such formal rational 'functions' can be considered as genuine functions from $\mathbb{C}$ to $\mathbb{C}$ defined almost everywhere (because g has only finitely many zeros). Moreover, two such functions are considered as the same one, if they coincide wherever they are defined.

(iv) Let $R = C(\mathbb{R})$ be the ring of all real-valued continuous functions on $\mathbb{R}$, and let $P \lhd R$ be the maximal (and therefore prime) ideal in R which consists of all functions such that $f(5) = 0$. Then R_P can be described as the algebra of stocks of continuous functions at the point 5. A stock at 5 is an equivalence class of continuous functions with respect to the equivalence relation $f \sim g$ if $f \equiv g$ in a neighborhood of 5. The ring operations on stocks are inherited from those on functions, e.g. $[f][g] = [fg]$.

Some easy general properties are listed below.

Lemma 5.2.3 *Let R be a commutative ring, S be a multiplicative subset of R, and $\varphi : R \to S^{-1}R$ be a canonical homomorphism.*

(i) $\ker \varphi = \{a \in R \mid rs = 0 \text{ for some } s \in S\}$.

(ii) *φ is an isomorphism if and only if S consists of units.*

(iii) *$S^{-1}R = 0$ if and only if $0 \in S$ if and only if S contains a nilpotent element.*

(iv) *If R is a domain and S is proper then $S^{-1}R$ is a domain and φ is injective.*

Proof (i) is clear. For (ii), $\varphi(s)$ is always a unit, and so if φ is an isomorphism, s should also be a unit. Conversely, if S consists of units, then φ is injective by (i), and surjective, as any $[\frac{r}{s}] = \varphi(rs^{-1})$. For (iii),

if $S^{-1}R = 0$ then $[\frac{1}{1}] = 0$, which implies that $u \cdot 1 = 0$ for some $u \in S$, and so $0 \in S$. The converse in (iii) is obvious, as is (iv). □

The notion of the ring of fractions extends to modules. Let V be an R-module and S be a multiplicative subset of R. The module of fractions $S^{-1}V$ as a set consists of equivalence classes of fractions of the form $\frac{v}{s}$ where $v \in V$ and $s \in S$. The equivalence relation is given by

$$\frac{v}{s} \sim \frac{w}{t} \quad \Leftrightarrow \quad u(tv - sw) = 0 \text{ for some } u \in S.$$

Moreover, $S^{-1}V$ is an $S^{-1}R$-module with respect to the obvious operations, e.g. $[\frac{a}{s}][\frac{v}{t}] = [\frac{av}{st}]$. We leave it as an exercise to check that everything makes sense. The following is a universal property characterization of $S^{-1}V$. Note that whenever we are given an $S^{-1}R$-module, it can be considered as an R-module via a homomorphism $\varphi : R \to S^{-1}R$.

Proposition 5.2.4 *Let V be an R-module and S be a multiplicative subset of R. Then the map*

$$\psi : V \to S^{-1}V, \ v \mapsto [\frac{v}{1}]$$

is a homomorphism of R-modules and every R-homomorphism from V to an $S^{-1}R$-module factors uniquely through ψ. Moreover, this properties characterize $S^{-1}V$ uniquely up to isomorphism of $S^{-1}R$-modules.

Proof Exercise. □

Proposition 5.2.5 *Let V be an R-module and S be a multiplicative subset of R. Then there is an isomorphism of $S^{-1}R$-modules*

$$\alpha : S^{-1}R \otimes_R V \tilde{\to} S^{-1}V, \ [\frac{a}{s}] \otimes v \mapsto [\frac{av}{s}].$$

Proof The existence of the map α follows from the universal property of tensor products. The inverse map is constructed using universal property of fraction modules, see Proposition 5.2.4. □

The following result shows that if $V \subseteq W$ is an R-submodule then $S^{-1}V$ can be considered as an $S^{-1}R$-submodule of $S^{-1}W$ which consists of all fractions of the form $[\frac{v}{s}] \in S^{-1}W$ with $v \in V$.

Theorem 5.2.6 *Let S be a multiplicative subset of R. Then $S^{-1}R$ is a flat R-module.*

Proof We just need to check that for every injective map $f : V \to W$ of R-modules, the corresponding map

$$\mathrm{id} \otimes F : S^{-1}R \otimes_R V \to S^{-1}R \otimes_R W$$

is also injective. Identifying $S^{-1}R \otimes_R V$ with $S^{-1}V$ and $S^{-1}R \otimes_R W$ with $S^{-1}W$ by Proposition 5.2.5, we see that we need to prove that the map

$$S^{-1}f : S^{-1}V \to S^{-1}W,\ [\frac{v}{s}] \mapsto [\frac{f(v)}{s}]$$

is injective. Well, assume that $[\frac{f(v)}{s}] = 0$. This means that $f(uv) = uf(v) = 0$ for some $u \in S$. But f is injective, and so $uv = 0$, whence $[\frac{v}{s}] = 0$. □

Our next goal is to study relations between ideals in R and $S^{-1}R$. We have two obvious operations (which are of course available in much larger generality, namely, whenever there is a ring homomorphism $\varphi : R_1 \to R_2$). If I is an ideal of R, its *expansion* $S^{-1}I$ is defined to be the ideal of $S^{-1}R$ generated by $\varphi(I)$. Clearly we have

$$S^{-1}I = \{[\frac{x}{s}] \in S^{-1}R \mid x \in I\}.$$

If J is an ideal of $S^{-1}R$, its *contraction* J^c is defined to be just $\varphi^{-1}(J)$.

Lemma 5.2.7 *Every ideal J of $S^{-1}R$ is the expansion of some ideal of R, namely of J^c.*

Proof We prove that $S^{-1}J^c = J$. If $[\frac{x}{s}] \in S^{-1}J^c$, we may assume that $x \in J^c$, which means that $[\frac{x}{1}] \in J$, hence $[\frac{x}{s}] = [\frac{1}{s}][\frac{x}{1}] \in J$. Conversely, if $[\frac{x}{s}] \in J$ then $\varphi(x) = [\frac{x}{1}] = [\frac{s}{1}][\frac{x}{s}] \in J$, so $x \in J^c$ and $[\frac{x}{s}] \in S^{-1}J^c$. □

Corollary 5.2.8 *If R is noetherian (resp. artinian), then so is $S^{-1}R$.*

Proof We prove the noetherian part. If $J_1 \subseteq J_2 \subseteq \ldots$ is an ascending chain of ideals in $S^{-1}R$, then $J_k = S^{-1}I_k$, where $I_k = J_k^c$, and the chain stabilizes because so does $I_1 \subseteq I_2 \subseteq \ldots$. □

We record the following nice properties.

Lemma 5.2.9 *Let I, J be ideals in R.*

(i) *$S^{-1}I = S^{-1}R$ if and only if $I \cap S \neq \varnothing$.*

(ii) $S^{-1}(I+J) = S^{-1}I + S^{-1}J$, $S^{-1}(I \cap J) = S^{-1}I \cap S^{-1}J$, *and* $S^{-1}(IJ) = (S^{-1}I)(S^{-1}J)$.

Proof This is routine. For example, let us prove that $S^{-1}(I \cap J) = S^{-1}I \cap S^{-1}J$. The inclusion $S^{-1}(I \cap J) = S^{-1}I \cap S^{-1}J$ is clear. Conversely, assume that $[\frac{x}{s}] \in S^{-1}I \cap S^{-1}J$. Then $[\frac{x}{s}] = [\frac{i}{t}] = [\frac{j}{u}]$ for some $i \in I$ and $j \in J$. Then there is $v \in S$ with $v(iu - jt) = 0$. It follows that $uiv \in I \cap J$. Hence $[\frac{x}{s}] = [\frac{i}{t}] = [\frac{iuv}{tuv}] \in S^{-1}(I \cap J)$. □

Definition 5.2.10 A commutative ring R is called *local* if it has only one maximal ideal.

The following result is a key trick which allows to reduce many questions in commutative algebra to local rings. It also explains the role of prime ideals and the term *localization at a prime* introduced in Example 5.2.2(iii).

Lemma 5.2.11 *Let S be a proper multiplicative subset of R and P be a prime ideal avoiding S. Then $[\frac{a}{s}] \in S^{-1}P$ if and only if $a \in P$. In particular P is the contraction of $S^{-1}P$.*

Proof If $a \in P$ then of course $[\frac{a}{s}] \in S^{-1}P$. Conversely assume that $[\frac{a}{s}] \in S^{-1}P$. This means that $[\frac{a}{s}] = [\frac{p}{t}]$ for some $p \in P$. Then $uta = usp \in P$ for some $u \in S$. Now $ut \notin P$ implies $a \in P$. □

Proposition 5.2.12 *Let S be a proper multiplicative subset of R. The expansion and contraction define a one-to-one correspondence between the prime ideals of $S^{-1}R$ and the prime ideals of R avoiding S.*

Proof If Q is a prime ideal of $S^{-1}R$, then by Lemma 5.2.7, Q is the expansion of Q^c, and it is clear that Q^c is a prime avoiding S.

Let P be a prime ideal of R such that $P \cap S = \varnothing$. In view of Lemma 5.2.11, it suffices to prove that $S^{-1}P$ is prime in $S^{-1}R$. Assume that $[\frac{a}{s}][\frac{b}{t}] = [\frac{ab}{st}] \in S^{-1}P$. Then by Lemma 5.2.11, $ab \in P$, so either a or b is in P, and so either $[\frac{a}{s}]$ or $[\frac{b}{t}]$ is in $S^{-1}P$. □

Corollary 5.2.13 *Let P be a prime ideal of R and $S = R \setminus P$. Then $P_P := S^{-1}P$ is the only maximal ideal of R_P.*

5.3 Ring extensions

Definition 5.3.1 A *ring extension* of a ring R is a ring A of which R is a subring.

If A is a ring extension of R, A is a fathful R-module in a natural way. Let A be a ring extension of R and S be a subset of A. The subring of A generated by R and S is denoted $R[S]$. It is quite clear that $R[S]$ consists of all R-linear combinations of products of elements of S.

Definition 5.3.2 A ring extension A of R is called *finitely generated* if $A = R[s_1, \ldots, s_n]$ for some finitely many elements $s_1, \ldots, s_n \in A$.

The following notion resembles that of an algebraic element for field extensions.

Definition 5.3.3 Let A be a ring extension of R. An element $\alpha \in A$ is called *integral* over R if $f(\alpha) = 0$ for some monic polynomial $f(x) \in R[x]$. A ring extension $R \subseteq A$ is called *integral* if every element of A is integral over R.

In Proposition 5.3.5 we give two equivalent reformulations of the integrality condition. For the proof we will need the following technical

Lemma 5.3.4 *Let V be an R-module. Assume that $v_1, \ldots, v_n \in V$ and $a_{ij} \in R$, $1 \leq i, j \leq n$ satisfy $\sum_{j=1}^n a_{kj} v_j = 0$ for all $1 \leq k \leq n$. Then $D := \det(a_{ij})$ satisfies $Dv_i = 0$ for all $1 \leq i \leq n$.*

Proof We expand D by the ith column to get $D = \sum_{k=1}^n a_{ki} C_{ki}$, where C_{ki} is the (k, i) cofactor. We then also have $\sum_{k=1}^n a_{kj} C_{ki} = 0$ for $i \neq j$. So

$$\begin{aligned} Dv_i &= \sum_{k=1}^n a_{ki} C_{ki} v_i = \sum_{k=1}^n a_{ki} C_{ki} v_i + \sum_{j \neq i} (\sum_{k=1}^n a_{kj} C_{ki}) v_j \\ &= \sum_{j=1}^n \sum_{k=1}^n a_{kj} C_{ki} v_j = \sum_{k=1}^n C_{ki} \sum_{j=1}^n a_{kj} v_j = 0. \end{aligned}$$

□

Proposition 5.3.5 *Let A be a ring extension of R and $\alpha \in A$. The following conditions are equivalent:*

(i) *α is integral over R.*

(ii) *$R[\alpha]$ is a finitely generated R-module.*

(iii) *There exists a faithful $R[\alpha]$-module which is finitely generated as an R-module.*

Proof (i) $\Rightarrow$ (ii) Assume $f(\alpha) = 0$, where $f(x) \in R[x]$ is monic of degree n. Let $\beta \in R[\alpha]$. Then $\beta = g(\alpha)$ for some $g \in R[x]$. As f is monic, we can write $g = fq+r$, where $\deg r < n$. Then $\beta = g(\alpha) = r(\alpha)$. Thus $R[\alpha]$ is generated by $1, \alpha, \ldots, \alpha^{n-1}$ as an R-module.

(ii) $\Rightarrow$ (iii) is clear.

(iii) $\Rightarrow$ (i) Let V be a faithful $R[\alpha]$-module which is generated as an R-module by finitely many elements $v_1, \ldots, v_n$. Write

$$\alpha v_i = a_{i1}v_1 + \cdots + a_{in}v_n \quad (1 \leq i \leq n).$$

Then

$$-a_{i1}v_1 - \cdots - a_{i,i-1}v_{i-1} + (\alpha - a_{ii})v_i - a_{i,i+1}v_{i+1} - \cdots - a_{in}v_n = 0$$

for all $1 \leq i \leq n$. By Lemma 5.3.4, we have $Dv_i = 0$ for all i, where

$$D = \begin{vmatrix} \alpha - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \alpha - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & a_{n2} & \cdots & \alpha - a_{nn} \end{vmatrix}.$$

As $v_1, \ldots, v_n$ generate V, this implies that D annihilates V. As V is faithful, $D = 0$. Expanding D shows that $D = f(\alpha)$ for some monic polynomial $f(x) \in R[x]$. □

Lemma 5.3.6 *Let $R \subseteq A \subseteq B$ be ring extensions. If A is finitely generated as an R-module and B is finitely generated as an A-module, then B is finitely generated as an R-module.*

Proof If $a_1, \ldots, a_m$ are generators of the R-module A and $b_1, \ldots, b_n$ are generators of the A-module B, then it is easy to see that $\{a_i b_j\}$ are generators of the R-module B. □

Proposition 5.3.7 *Let A be a ring extension of R.*

(i) *If A is finitely generated as an R-module, then A is integral over R.*

(ii) *If $A = R[\alpha_1, \ldots, \alpha_n]$ and $\alpha_1, \ldots, \alpha_n$ are integral over R, then A is finitely generated as an R-module and hence integral over R.*

(iii) *If $A = R[S]$ and every $s \in S$ is integral over R, then A is integral over R.*

Proof (i) Let $\alpha \in A$. Then A is a faithful $R[\alpha]$-module, and we can apply Proposition 5.3.5.

(ii) Note that $R[\alpha_1, \ldots, \alpha_i] = R[\alpha_1, \ldots, \alpha_{i-1}][\alpha_i]$. Now apply induction, Proposition 5.3.5 and Lemma 5.3.6.

(iii) Follows from (ii). □

Corollary 5.3.8 *Let A be a ring extension of R. The elements of A which are integral over R form a subring of A.*

Proof If $\alpha_1, \alpha_2 \in A$ are integral, then $\alpha_1 - \alpha_2$ and $\alpha_1\alpha_2$ belong to $R[\alpha_1, \alpha_2]$. So we can apply Proposition 5.3.7(ii). □

This result allows us to give the following definition

Definition 5.3.9 The *integral closure* of R in $A \supseteq R$ is the ring $\bar{R}$ of all elements of A that are integral over R. The ring R is *integrally closed* in $A \supseteq R$ in case $\bar{R} = R$. A domain R is called *integrally closed* if it is integrally closed in its field of fractions.

Example 5.3.10 The elements of the integral closure of $\mathbb{Z}$ in $\mathbb{C}$ are called *algebraic integers*. They form a subring of $\mathbb{C}$. In fact the field of algebraic numbers $\mathbb{A}$ is the quotient field of this ring.

We record some further nice properties of integral extensions.

Proposition 5.3.11 *Let R, A, B be rings.*

(i) *If $R \subseteq A \subseteq B$ then B is integral over R if and only if B is integral over A and A is integral over R.*

(ii) *If B is integral over A and $R[B]$ makes sense then $R[B]$ is integral over $R[A]$.*

(iii) *If A is integral over R and $\varphi : A \to B$ is a ring homomorphism then $\varphi(A)$ is integral over $\varphi(R)$.*

(iv) *If A is integral over R, then $S^{-1}A$ is integral over $S^{-1}R$ for every proper multiplicative subset S of R.*

Proof (i)-(iii) is an exercise.

(iv) First of all, it follows from definitions that $S^{-1}R$ is indeed a subring of $S^{-1}A$. Now, let $[\frac{a}{s}] \in S^{-1}A$. As $[\frac{a}{s}] = [\frac{a}{1}][\frac{1}{s}]$, it suffices to

show that both $[\frac{a}{1}]$ and $[\frac{1}{s}]$ are integral over $S^{-1}R$. But $\frac{1}{s} \in S^{-1}R$ and for $[\frac{a}{1}]$ we can use the monic polynomial which annihilates a. □

It follows from Proposition 5.3.11(i) that the closure of $\bar{R}$ in $A \supseteq R$ is again $\bar{R}$. In particular, if D is any domain and F is its field of fractions, then the closure $\bar{D}$ in F is an integrally closed domain (since the quotient field of $\bar{D}$ is also F).

We recall that a domain R is called a *unique factorization domain* or *UFD* if every non-zero non-unit element of R can be written as a product of irreducible elements, which is unique up to a permutation and units.

Proposition 5.3.12 *Every UFD is integrally closed.*

Proof Let R be a UFD and F be its field of fractions. Let $\frac{a}{b} \in F$ be integral over R. We may assume that no irreducible element of R divides both a and b. There is a monic polynomial $f(x) = x^n + r_{n-1}x^{n-1} + \cdots + r_0 \in R[x]$ with $f(\frac{a}{b}) = 0$, which implies $a^n + r_{n-1}a^{n-1}b + \cdots + r_0 b^n = 0$. So, if $p \in R$ is an irreducible element dividing b then p divides a^n, and hence p divides a, a contradiction. Therefore b is a unit and $\frac{a}{b} \in R$. □

Proposition 5.3.13 *If a domain R is integrally closed, then so is $S^{-1}R$ for any proper multiplicative subset S of R.*

Proof Exercise. □

Example 5.3.14 The ring $\mathbb{Z}[i]$ of Gaussian integers is Euclidean (the degree function is $\partial(a + bi) = a^2 + b^2$, hence it is a UFD, and so it is integrally closed by Proposition 5.3.13. On the other hand consider the ring $\mathbb{Z}[2i]$. The quotient field of both $\mathbb{Z}[i]$ and $\mathbb{Z}[2i]$ is $\mathbb{Q}(i)$, and we have $\mathbb{Z}[2i] \subset \mathbb{Z}[i] \subset \mathbb{Q}[i]$. Clearly $\mathbb{Z}[2i]$ is not integrally closed, as $i \notin \mathbb{Z}[2i]$ is integral over it. It is easy to see that $\overline{\mathbb{Z}[2i]} = \mathbb{Z}[i]$.

Next we are going to address the question of how prime ideals of R and A are related if $A \supseteq R$ is an integral extension.

Definition 5.3.15 Let $R \subseteq A$ be a ring extension. We say that a prime ideal P of A *lies over* a prime ideal $\mathfrak{p}$ of R if $P \cap R = \mathfrak{p}$.

The following lemma is a key technical trick.

Lemma 5.3.16 *Let $A \supseteq R$ be an integral ring extension, $\mathfrak{p}$ be a prime ideal of R, and $S := R \setminus \mathfrak{p}$.*

(i) *Let I be an ideal of A avoiding S, and P be an ideal of A maximal among the ideals of A which contain I and avoid S. Then P is a prime ideal of A lying over $\mathfrak{p}$.*

(ii) *If P is a prime ideal of A which lies over $\mathfrak{p}$, then P is maximal in the set T of all ideals in A which avoid S.*

Proof (i) Clearly, S is a proper multiplicative subset of A. So P is prime in view of Lemma 5.1.8. We claim that $P \cap R = \mathfrak{p}$. That $P \cap R \subseteq \mathfrak{p}$ is clear as $P \cap S = \varnothing$.

Assume that $P \cap R \subsetneq \mathfrak{p}$. Let $c \in \mathfrak{p} \setminus P$. By the maximality of P, $p + \alpha c = s \in S$ for some $p \in P$ and $\alpha \in A$. As A is integral over R, we have

$$0 = \alpha^n + r_{n-1}\alpha^{n-1} + \cdots + r_0$$

for some $r_0, \ldots, r_{n-1} \in R$. Multiplying by c^n yields

$$\begin{aligned} 0 &= c^n\alpha^n + cr_{n-1}c^{n-1}\alpha^{n-1} + \cdots + c^n r_0 \\ &= (s-p)^n + cr_{n-1}(s-p)^{n-1} + \cdots + c^n r_0. \end{aligned}$$

If we decompose the last expression as the sum of monomials, then the part which does not involve any positive powers of p looks like

$$x := s^n + cr_{n-1}s^{n-1} + \cdots + c^n r_0.$$

It follows that $x \in P$. On the other hand, $x \in R$, so $x \in R \cap P \subseteq \mathfrak{p}$. Now $c \in \mathfrak{p}$ implies $s^n \in \mathfrak{p}$. As $\mathfrak{p}$ is prime, $s \in \mathfrak{p}$, a contradiction.

(ii) If P is not maximal in T, then there exists an ideal I in T which properly contains P. As I still avoids S, it also lies over $\mathfrak{p}$. Take $u \in I \setminus P$. Then $u \notin R$ and u is integral over R. So the set of all polynomials $f \in R[x]$ such that $\deg f \geq 1$ and $f(u) \in P$ is non-empty. Take such $f(x) = \sum_{i=0}^{n} r_i x^i$ of minimal possible degree. We have

$$u^n + r_{n-1}u^{n-1} + \cdots + r_0 \in P \subseteq I,$$

whence $r_0 \in R \cap I = \mathfrak{p} = R \cap P \subseteq P$. Therefore

$$u^n + r_{n-1}u^{n-1} + \cdots + r_1 u = u(u^{n-1} + r_{n-1}u^{n-2} + \cdots + r_1) \in P.$$

By the choice of u and minimality of $\deg f$, $u \notin P$ and $u^{n-1} + r_{n-1}u^{n-2} + \cdots + r_1 \notin P$. We have contradiction because P is prime. □

Corollary 5.3.17 (Lying Over Theorem) *If A is integral over P then for every prime ideal $\mathfrak{p}$ of R there exists a prime ideal P of A which lies*

over $\mathfrak{p}$. *More generally, for every ideal* I *of* A *such that* $I \cap R \subseteq \mathfrak{p}$ *there exists a prime ideal* P *of* A *which contains* I *and lies over* $\mathfrak{p}$.

Corollary 5.3.18 (Going Up Theorem) *Let* $A \supseteq R$ *be an integral ring extension, and* $\mathfrak{p}_1 \subseteq \mathfrak{p}_2$ *be prime ideals in* R. *If* P_1 *is a prime ideal of* A *lying over* $\mathfrak{p}_1$, *then there exists a prime ideal* P_2 *of* A *such that* $P_1 \subseteq P_2$ *and* P_2 *lies over* $\mathfrak{p}_2$.

Proof Take $\mathfrak{p} = \mathfrak{p}_2$ and $I = P_1$ in Lemma 5.3.16(i). □

Corollary 5.3.19 (Incomparability) *Let* $A \supseteq R$ *be an integral ring extension, and* P_1, P_2 *be prime ideals of* A *lying over a prime ideal* $\mathfrak{p}$ *of* R. *Then* $P_1 \subseteq P_2$ *implies* $P_1 = P_2$.

Proof Use Lemma 5.3.16(ii). □

The relation between prime ideals established above has further nice properties.

Theorem 5.3.20 (Maximality) *Let* $A \supseteq R$ *be an integral ring extension, and* P *be a prime ideal of* A *lying over a prime ideal* $\mathfrak{p}$ *of* R. *Then* P *is maximal if and only if* $\mathfrak{p}$ *is maximal.*

Proof If $\mathfrak{p}$ is not maximal, we can find a maximal ideal $\mathfrak{m} \supsetneq \mathfrak{p}$. By the Going Up Theorem, there is an ideal M of A lying over $\mathfrak{m}$ and containing P. It is clear that M actually containg P properly, and so P is not maximal.

Conversely, let $\mathfrak{p}$ be maximal in R. Let M be a maximal ideal containing P. Then $M \cap R \supseteq P \cap R = \mathfrak{p}$ and we cannot have $M \cap R = R$, as $1_R = 1_S \notin M$. It follows that $M \cap R = \mathfrak{p}$. Now $M = P$ by Incomparability Theorem. □

The previous results can be used to prove some useful properties concerning extensions of homomorphisms.

Lemma 5.3.21 *Let* $A \supseteq R$ *be an integral ring extension. If* R *is a field then* $A \supseteq R$ *is an algebraic field extension.*

Proof Let $\alpha \in A$ be a non-zero element. Then α is algebraic over R, hence $R[\alpha] \subseteq A$ is a field, and α is invertible. Hence A is a field. □

Proposition 5.3.22 *Let A be integral over R. Every homomorphism φ of R to an algebraically closed field F can be extended to A.*

Proof If R is a field, then A is an algebraic field extension of R by Lemma 5.3.21. Now the result follows from Proposition 2.16.5.

If R is local, then $\ker\varphi$ is the maximal ideal $\mathfrak{m}$ of R. By Lying Over and Maximality Theorems, there is an ideal M of A lying over $\mathfrak{m}$. The inclusion $R \to A$ then induces an embedding of fields $R/\mathfrak{m} \to A/M$, which we use to identify $R/\mathfrak{m}$ with a subfield of A/M. Note that the field extension $A/M \supseteq R/\mathfrak{m}$ is algebraic. Since $\ker\varphi = \mathfrak{m}$, φ factors through the projection $R \to R/\mathfrak{m}$. The resulting homomorphism $\varphi : R/\mathfrak{m} \to F$ can be extended to $\psi : A/M \to F$ by Proposition 2.16.5. Now if $\pi : A \to A/M$ is the natural projection, then $\psi \circ \pi$ is the desired extension of φ.

Now we consider the general case. Let $\mathfrak{p} := \ker\varphi$, a prime ideal in R, and $S = R \setminus \mathfrak{p}$. Then $S^{-1}A$ is integral over $S^{-1}R$ by Proposition 5.3.11(iv). Now $S^{-1}R = R_{\mathfrak{p}}$ is local. By the universal property of localizations, φ extends to a ring homomorphism $\hat{\varphi} : S^{-1}R \to F$. By the local case, $\hat{\varphi}$ extends to $\hat{\psi} : S^{-1}A \to F$, and the desired extension $\psi : A \to F$ is obtained by composing ψ with the natural homomorphism $A \to S^{-1}A$. □

Proposition 5.3.23 *Every homomorphism of a field k into an algebraically closed field can be extended to every finitely generated ring extension of k.*

Proof Let $\varphi : k \to F$ be a homomorphism to an algebraically closed field F and R be a finitely generated ring extension of k, so that $R = k[\alpha_1, \ldots, \alpha_n]$ for some $\alpha_1, \ldots, \alpha_n \in R$.

First assume that R is a field. By Proposition 5.3.22, we may assume that R is not algebraic over k. Let $\{\beta_1, \ldots, \beta_t\}$ be a (necessarily finite) transcendence base of R over k. Each $\alpha \in R$ is algebraic over $k(\beta_1, \ldots, \beta_t)$, i.e. satisfies a polynomial $a_k\alpha^k + \cdots + a_1\alpha + a_0 = 0$ with coefficients $a_k, \ldots, a_0 \in k(\beta_1, \ldots, \beta_t)$, $a_k \neq 0$. Multiplying by a common denominator yields a polynomial equation

$$b_k\alpha^k + \cdots + b_1\alpha + b_0 = 0$$

with coefficients $b_k, \ldots, b_0 \in k[\beta_1, \ldots, \beta_t]$, $b_k \neq 0$. Hence α is integral over $k[\beta_1, \ldots, \beta_t, \frac{1}{b_k}]$. Applying this to $\alpha_1, \ldots, \alpha_n$ yields nonzero $c_1, \ldots, c_n \in k[\beta_1, \ldots, \beta_t]$ such that $\alpha_1, \ldots, \alpha_n$ are integral over

$k[\beta_1, \ldots, \beta_t, \frac{1}{c_1}, \ldots, \frac{1}{c_n}]$. Set $c = c_1 \ldots c_n$. Then $\alpha_1, \ldots, \alpha_n$ are integral over $k[\beta_1, \ldots, \beta_t, \frac{1}{c}]$, and hence R is integral over $k[\beta_1, \ldots, \beta_t, \frac{1}{c}]$, see Proposition 5.3.7(ii). Let c^φ be the image of c under the homomorphism

$$k[\beta_1, \ldots, \beta_t] \cong k[x_1, \ldots, x_t] \to F[x_1, \ldots, x_t]$$

induced by φ. As F is infinite there exist $\gamma_1, \ldots, \gamma_t \in F$ such that $c^\varphi(\gamma_1, \ldots, \gamma_t) \neq 0$. By the universal property of polynomial rings, there exists a homomorphism $\psi : k[\beta_1, \ldots, \beta_t] \to F$ which extends φ and sends $\beta_1, \ldots, \beta_t$ to $\gamma_1, \ldots, \gamma_t$, respectively. The universal property of localizations yields an extension of ψ to ring $k[\beta_1, \ldots, \beta_t, \frac{1}{c}] = k[\beta_1, \ldots, \beta_t]_c$. Now Proposition 5.3.22 extends φ to R, which completes the case where R is a field.

Now, let $R = k[\alpha_1, \ldots, \alpha_n]$ be any finitely generated ring extension of k. Let $\mathfrak{m}$ be a maximal ideal of R and $\pi : R \to R/\mathfrak{m}$ be the natural projection. Then $R/\mathfrak{m}$ is a field extension of $\pi(k) \cong k$ generated by $\pi(\alpha_1), \ldots, \pi(\alpha_n)$. By the first part of the proof, every homomorphism of $\pi(k)$ into F extends to $R/\mathfrak{m}$. Therefore every homomorphism of $k \cong \pi(k)$ extends to R. □

Example 5.3.24 We describe prime ideals of $\mathbb{Z}[i]$ and illustrate the Lying Over Theorem on integral extension $\mathbb{Z} \subset \mathbb{Z}[i]$. As $\mathbb{Z}[i]$ is a PID, the ideal (a) is prime if and only if a is irreducible. So we want to understand what are irreducible elements in $\mathbb{Z}[i]$. Recall that the degree function for the Euclidean ring $\mathbb{Z}[i]$ is $\partial(a + ib) = a^2 + b^2$. It has the property $\partial(xy) = \partial(x)\partial(y)$, whence ± 1 and $\pm i$ are the only units in $\mathbb{Z}[i]$. Moreover, it is clear that if $\partial(x) = p$, a prime number, then x is irreducible. This produces interesting irreducible elements when $p \equiv 1 \pmod 4$ in view of the following fact:

Fermat's Two-Squares Theorem An odd prime p is a sum of two squares if and only if $p \equiv 1 \pmod 4$. (*Proof.* Let $p = a^2 + b^2$. As p is odd, a and b have different parities, say $a = 2m$, $b = 2n + 1$. Then $a^2 + b^2 = 4m^2 + 4n^2 + 4n + 1 \equiv 1 \pmod 4$. Conversely, let $p \equiv 1 \pmod 4$. Then C_4 is a subgroup of C_{p-1}. Let $\bar{m}$ be a generator of C_4 inside $(\mathbb{Z}/pZ)^\times \cong C_{p-1}$. Then $m^2 \equiv -1 \pmod p$, and so we found an integer m such that $p \mid (m^2 + 1)$. Then in $\mathbb{Z}[i]$, $p \mid (m^2 + 1) = (m + i)(m - i)$. On the other hand, it is easy to see that p does not divide $m \pm i$. It follows that p is not an irreducible element of $\mathbb{Z}[i]$, i.e. $p = (a + ib)(c + id)$ with $(a + ib), (c + id) \in \mathbb{Z}[i]$ non-units. Then $p^2 = (a^2 + b^2)(c^2 + d^2)$, whence $p = a^2 + b^2 = c^2 + d^2$.

Now, let $\alpha = a + bi \in \mathbb{Z}[i]$ be non-zero non-unit. We claim that α is irreducible if and only if $\pm\alpha$ or $\pm i\alpha$ is one of the following:

(i) a prime $p \in \mathbb{Z}$ of the form $p = 4m + 3$;
(ii) $1 \pm i$;
(ii) $a \pm bi$, where $q = a^2 + b^2$ is a prime in $\mathbb{Z}$ of the form $4m + 1$.

First of all observe that the elements α described in (i)-(iii) are irreducible since for them we have $\partial(\alpha)$ is a prime in $\mathbb{Z}$. Conversely, let α be an arbitrary irreducible element in $\mathbb{Z}[i]$.

We prove that there is a unique prime p in $\mathbb{Z}$ with $\alpha \mid p$ in $\mathbb{Z}[i]$. Indeed, $\alpha\bar{\alpha} = \partial(\alpha)$ shows that $\alpha \mid \partial(\alpha)$. Decomposing $\partial(\alpha)$ as a product of prime numbers in $\mathbb{Z}$, we see that α divides some prime p. If α also divides another prime p', then α also divides $1 = (p.p')$, a contradiction.

Now, $\alpha \mid p$ implies $\partial(\alpha) \mid \partial(p) = p^2$, so that $\partial(\alpha) = p$ or p^2. If $p \equiv 3 \pmod 4$, then $\partial(\alpha) = p$ cannot occur, by Fermat's Two-Squares Theorem. Now, $\alpha \mid p$ implies $p = \alpha\beta$ for some β. Then $\partial(\alpha)\partial(\beta) = p^2$, and hence $\partial(\beta) = 1$, whence β is a unit, and so, up to a unit $\alpha = p$. If $p = 2$, then clealry $\alpha = 1 \pm i$ up to multiplication by units. Finally, let $p \equiv 1 \pmod 4$. We need to rule out the case $\partial(\alpha) = p^2$. By Fermat's Two-Squares Theorem, we can write $p = c^2 + d^2$. Let $\beta = c + di$, an irreducible element of $\mathbb{Z}[i]$. Then α divides $p = \beta\bar{\beta}$. As α is irreducible, *al* divides β or $\bar{\beta}$. But β and $\bar{\beta}$ are also irreducible, so, up to unit, α equals β or $\bar{\beta}$.

Now, that we know what prime ideals in $\mathbb{Z}[i]$ look like, it is easy to see that for a prime $p \in \mathbb{Z}$ of the form $p = 4m + 3$, exactly one prime lies over (p). Exactly two prime ideals $(1+i)$ and $(1-i)$ lie over (2). As for primes $p \in \mathbb{Z}$ of the form $p = 4m + 3$, the prime ideals of the form $(a + bi)$ lie over (p), for each $a, b \in \mathbb{Z}$ such that $a^2 + b^2 = p$.

5.4 Krull Theorems on Noetherian Rings

Lemma 5.4.1 *Let V be a finitely generated R-module. If $IV = V$ for some ideal I of R, then $(1 - x)V = 0$ for some $x \in I$.*

Proof We have $V = Rv_1 + \ldots Rv_n$ for some $v_1, \ldots, v_n \in V$. Then $IV = Iv_1 + \cdots + Iv_n$. If $V = IV$, then

$$v_i = \sum_{j=1}^{n} a_{ij} v_j \qquad (1 \leq i \leq n)$$

for some $a_{ij} \in I$. So the matrix $I-A$ satisfies assumptions of Lemma 5.3.4. So $d = \det(I - A)$ annihilates V. On the other hand, since all entries of A belong to I, we have $d = 1 - x$ for some $x \in I$. □

Theorem 5.4.2 (Krull Intersection Theorem) *Let R be a noetherian ring and I be an ideal of R. Let $J = \cap_{n>0} I^n$. Then $IJ = J$ and $(1 - x)J = 0$ for some $x \in I$. If $I \neq R$ and either R is a domain or R is local then $J = 0$.*

Proof First of all, $IJ \subseteq J$ because J is an ideal. For the converse inclusion, let Q be a primary ideal of R which contains IJ. As IJ is an intersection of primary ideals, it suffices to prove that Q contains J. Let $P = \sqrt{Q}$. By Lemma 5.1.11, $P^n \subseteq Q$ for some n. If $J \not\subseteq Q$, then $IJ \subseteq Q$ implies $I \subseteq P$, since Q is primary, and $J \subseteq I^n \subseteq P^n \subseteq Q$.

Applying Lemma 5.4.1 to the finitely generated R-module J yields $(1-x)J = 0$ for some $x \in I$. If R is a domain and $I \neq R$, then $1-x \neq 0$ and $J = 0$. Finally, if R is local then and $I \neq R$, then $1 - x$ is a unit and $J = 0$ again. □

Lemma 5.4.3 *Let I be an ideal of a commutative ring R. Every prime ideal P which contains I contains a prime ideal which contains I and is minimal with this property.*

Proof Consider the set $\mathcal{S}$ of prime ideals which contain I and are contained in P ordered by inverse inclusion: $P_1 \leq P_2$ if and only if $P_1 \supseteq P_2$. It suffices to show that this set satisfies the assumptions of Zorn's Lemma. Well, assume that we have a linearly ordered subset $\{P_i\}_{i \in I} \subseteq \mathcal{S}$. It suffices to prove that that $\cap_{i \in I} P_i$ is a prime ideal, which is easy to see. □

Definition 5.4.4 Let I be an ideal of a commutative ring R. A prime ideal P is *minimal over I* (or is *isolated prime ideal of I*) in case P is minimal among all prime ideals of R which contain I.

Proposition 5.4.5 *Let I be an ideal of a noetherian ring R. There are only finitely many prime ideals of R that are minimal over I.*

Proof Assume the result is false. As R is noetherian, there exists an ideal J which is maximal with the property that there are infinitely many prime ideals of R that are minimal over J. Then J is not prime,

i.e. there are ideals $J_1, J_2 \not\supseteq J$ such that $J_1 J_2 \subseteq J$. By the maximality of J, there are only finitely many prime ideals of R that are minimal over J_1 and there are only finitely many prime ideals of R that are minimal over J_2. Now, if a prime ideal P is minimal over J, then $J_1 J_2 \subseteq J \subseteq P$ whence $J_1 \subseteq P$ or $J_2 \subseteq P$, and P is minimal over J_1 or over J_2. Hence there are only finitely many primes minimal over J, a contradiction. □

Lemma 5.4.6 *If M is a maximal ideal of a noetherian ring, then R/M^n is an artinian R-module for every $n > 0$.*

Proof We show that $V = M^{n-1}/M^n$ is an artinian R-module for all $n > 0$. Since $(R/M^n)/(M^{n-1}/M^n) \cong R/M^{n-1}$, the result will follow by induction using Lemma 3.4.5. We have $MV = 0$, hence V is an R/M-module in a natural way. Now, M^{n-1} is a finitely generated R-module, hence V is a finitely generated R/M-module. Since R/M is a field, it follows that V is artinian R/M-module, hence an artinian R-module. □

Definition 5.4.7 Let P be a prime ideal of R. The nth *symbolic power* of P is defined to be the ideal

$$P^{(n)} := \{r \in R \mid sr \in P^n \text{ for some } s \in R \setminus P\}.$$

It is clear that $P^{(n)} \supseteq P^{(n+1)}$ and $P^{(n)} \supseteq P^n$ for all n. Another description of $P^{(n)}$ is as follows:

$$P^{(n)} = (P_P^n)^c. \tag{5.1}$$

Lemma 5.4.8 *Let P be a prime ideal of R. Then $P^{(n)}$ is a P-primary ideal of R for any $n > 0$.*

Proof We have $\sqrt{P^{(n)}} \supseteq \sqrt{P^n} = P$. Assume that $xy \in P^{(n)}$ and $x \notin \sqrt{P^{(n)}}$. Then $x \notin P$. By definition, $sxy \in P^n$ for some $s \notin P$. As $x \notin P$, it follows that $y \in P^{(n)}$. Thus $P^{(n)}$ is primary. It remains to show that $\sqrt{P^{(n)}} \subseteq P$. Well, otherwise there exists $r \in R \setminus P$ with $r^n \in P^{(n)} \subseteq P$, which is impossible as P is prime. □

Lemma 5.4.9 *Let R be a noetherian local domain, M be its maximal ideal, and $a \in M$. Assume that $R/(a)$ is an artinian R-module. Then the only prime ideal which does not contain a is* 0.

Proof Assume P is prime and $a \notin P$. By Lemma 5.4.8, we have a descending series $P = P^{(1)} \supseteq P^{(2)} \supseteq \dots$ of P-primary ideals. Let $I = P \cap (a)$. Then $P/I \cong (P + (a))/(a) \subseteq R/(a)$ is an artinian R-module. Hence a descending series

$$P/I = P^{(1)}/I \supseteq \cdots \supseteq (P^{(n)} + I)/I \supseteq (P^{(n+1)} + I)/I \supseteq \dots$$

stabilizes, i.e. there is some $m > 0$ such that $P^{(n)} + I = P^{(m)} + I$ for all $n \geq m$.

We now show that $P^{(n)} = P^{(m)}$ for all $n \geq m$. Let $x \in P^{(m)}$. As $P^{(n)} + I = P^{(m)} + I$, we may write $x = y + ra$ for some $y \in P^{(n)}$. Now $a \notin P$ and $ra = x - y \in P^{(m)}$ implies $r \in P^{(m)}$, as $P^{(m)}$ is P-primary. Thus $x \in P^{(n)} + P^{(m)}a$. Hence the finitely generated R-module $V = P^{(m)}/P^{(n)}$ satisfies $V = (a)V$, whence $V = MV$, and $V = 0$ by Nakayama's Lemma (Lemma 3.12.11).

By Krull's Intersection Theorem, applied to the local ring R_P, we have $\cap_{n \geq m}(P_P)^n = 0$. Applying contraction to the last equality and using (5.1), we get $P^{(m)} = \cap_{n \geq m} P^{(n)} = 0$. Hence $P = \sqrt{P^{(m)}} = \sqrt{0} = 0$. □

Definition 5.4.10 A prime ideal P *has finite length* if there is an integer $n > 0$ such that every strictly decrasing sequence

$$P = P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_m$$

of prime ideals has $m \leq n$. Then the *height* of P is the smallest such integer n.

Theorem 5.4.11 (Krull's Hauptidealsatz) *In a noetherian ring, a prime ideal which is minimal over a principal ideal has height at most* 1.

Proof First we consider the case where R is local and P is the maximal ideal of R. Assume that P is minimal over a principal ideal (a). By Krull's Intersection Theorem, $P^n \subseteq (a)$ for some $n > 0$. By Lemma 5.4.6, R/P^n is an artinian R-module, hence its quotient $R/(a)$ is also artinian. If $P'' \subsetneq P' \subsetneq P$ are prime ideals, then $a \notin P'$ by the minimality of P. By Lemma 5.4.9 applied to the domain R/P'', $P' = P''$, a contradiction.

Now, let R be any noetherian ring and P be a prime ideal of R which is minimal over a principal ideal (a). Then R_P is a local noetherian ring with maximal ideal P_P, and P_P is minimal over the principal ideal $([\frac{a}{1}]) = (a)_P$, by Proposition 5.2.12. Hence P_P has height at most 1 in

R_P, and it follows from Proposition 5.2.12 again that P has height at most 1 in R. □

Lemma 5.4.12 *Let $P_1, \ldots, P_r$ be prime ideals of a commutative ring. An ideal I which is contained in $P_1 \cup \cdots \cup P_r$ is contained in some P_i.*

Proof By induction we may assume that I is not contained in the union of any proper subset of $\{P_1, \ldots, P_r\}$. Then $P_i \not\subseteq P_j$ for any $i \neq j$. If $I \cap (\cap_{i \neq j} P_i) \not\subseteq P_j$ for every j, then pick $x_j \in (I \cap (\cap_{i \neq j} P_i)) \setminus P_j$, and observe that $x_1 + \cdots + x_r \in I \setminus (P_1 \cup \cdots \cup P_r)$, a contradiction. Therefore $I \cap (\cap_{i \neq j} P_i) \subseteq P_j$ for some j. Then $I \prod_{i \neq j} P_i \subseteq P_j$, whence $I \subseteq P_j$, as P_j is prime. □

Lemma 5.4.13 *Let R be a noetherian ring, $Q_1, \ldots, Q_r$ be prime ideals of R, and*

$$P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_m$$

be a chain of prime ideals of R. If P_0 is contained in no Q_j, then there exists a chain

$$P_0 \supsetneq P_1' \supsetneq \cdots \supsetneq P_{m-1}' \supsetneq P_m$$

of prime ideals of R such that $P_1', \ldots, P_{m-1}'$ are contained in no Q_j.

Proof Induction on m, the result being trivial for $m \leq 1$. For $m > 1$ the induction hypothesis yields a chain $P_0 \supsetneq P_1' \supsetneq \cdots \supsetneq P_{m-2}' \supsetneq P_{m-1}$ of prime ideals of R such that $P_1', \ldots, P_{m-2}'$ are contained in no Q_j. By Lemma 5.4.12, $P_{m-2}' \not\subseteq P_m \cup Q_1 \cup \cdots \cup Q_r$. Take

$$x \in P_{m-2}' \setminus (P_m \cup Q_1 \cup \cdots \cup Q_r).$$

By Lemma 5.4.3, P_{m-2}' contains a prime ideal P_{m-1}' which is minimal over $(x) + P_m$. By the choice of x, P_{m-1}' properly contains P_m and is contained in no Q_j.

It remains to prove that P_{m-2}' contains P_{m-1}' properly. Well, in the noetherian ring R/P_m, the ideal P_{m-1}'/P_m is minimal over the principal ideal $(a + P_m)$, so it has height at most 1 in view of Krull's Hauptidealsatz. On the other hand, P_{m-2}'/P_m has height at least 2: consider the chain $P_{m-2}'/P_m \supsetneq P_{m-1}/P_m \supsetneq P_m/P_m$. □

Theorem 5.4.14 (Krull's Finite Height Theorem) *In a noetherian ring, every prime ideal has finite height. Moreover, if P is minimal over an ideal with r generators, then P has height at most r.*

Proof Let P be minimal over an ideal $I = Rx_1 + \cdots + Rx_r$ (for the first part of the theorem take $I = P$). We apply induction on r to prove that P has height at most r. If $r = 0$, $I = 0$ and P is a minimal prime ideal, so has height 0. Let $r > 0$ and $J = Rx_1 + \cdots + Rx_{r-1} \subseteq I$. If P is minimal over J, we can apply inductive hypothesis. Otherwise there are only finitely many prime ideals $Q_1, \ldots, Q_s$ that are minimal over J, see Proposition 5.4.5. Clearly, $P \supseteq J$ is contained in no Q_j. Assume that the height of P is m, and let

$$P = P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_m$$

be a chain of prime ideals. Then the height of P/P_{m-1} is $m - 1$, and the ideal $(P_{m-1} + J)/P_{m-1}$ of R/P_{m-1} is generated by $r - 1$ elements, so it suffices to show that P/P_{m-1} is minimal over $(P_{m-1} + J)/P_{m-1}$ and apply the inductive hypothesis.

By Lemma 5.4.13 we may assume that $P_1, \ldots, P_{m-1}$ are contained in no Q_j. In the noetherian ring R/J, the prime ideals $Q_1/J, \ldots, Q_s/J$ are minimal over 0. Moreover, P is minimal over $J + (x_r)$ and so P/J is minimal over the principal ideal $(x_r + J)$ in R/J. On the other hand P/J, is not minimal over 0, and so it must have height 1, thanks to Krull's Hauptidealsatz. Thus the only prime ideals of R/J properly contained in P/J are Q_j/J and 0. Next, $(P_{m-1} + J)/J$ is contained in P/J but is contained in no Q_j/J. Therefore, P/J is minimal over $(P_{m-1} + J)/J$. Then P is minimal over $P_{m-1} + J$. Hence P/P_{m-1} is minimal over $(P_{m-1} + J)/P_{m-1}$. □

We complete this section with the notion of Krull dimension.

Definition 5.4.15 Let R be a noetherian ring. The *Krull dimension* (or *dimension*) of R, denoted $\dim R$, is the maximum of the heights of prime ideals in R. (If the heights are not bounded, we write $\dim R = \infty$).

Example 5.4.16

(i) R has dimension 0 if and only if R is a field.
(ii) A PID has dimension 1, since all its non-zero prime ideals are maximal.

The following result is an easy exercise.

Lemma 5.4.17 *Let P be a prime ideal of a noetherian ring R.*

(i) *The height of P is the dimension of R_P.*

(ii) *If* $P \neq 0$, *then* $\dim R \geq 1 + \dim R/P$.

Lemma 5.4.18 *Let* R *be a domain and* P *be a prime ideal of* $R[x]$. *If* $P \cap R = 0$, *then* P *is a minimal prime ideal.*

Proof Let F be the field of fractions of R. Then $F = S^{-1}R$, where $S = R \setminus \{0\}$. Note that $Q[x] \cong S^{-1}R[x]$, and we will identify the two rings. Let $0 \neq Q \subseteq P$ be a prime ideal of $R[x]$. Then by assumption $P \cap S = Q \cap S = \varnothing$, and so $S^{-1}Q \subseteq S^{-1}P$ are non-zero proper prime ideals of $Q[x]$, see Proposition 5.2.12. As $Q[x]$ is a PID, we must have $S^{-1}Q = S^{-1}P$, whence $P = Q$ by Proposition 5.2.12 again. □

Theorem 5.4.19 *Let* R *be a noetherian domain of dimension* n. *Then* $R[x]$ *has dimension* $n + 1$.

Proof By Lemma 5.4.17(ii), $\dim R[x] \geq n + 1$, as $R \cong R[x]/(x)$. We prove by induction on n that $\dim R[x] \leq n+1$. If $n = 0$ then R is a field, $R[x]$ is a PID, and so $\dim R[x] = 1$. Let $n > 0$, and $P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_m$ be a chain of prime ideals in $R[x]$. We need to show that $m \leq n + 1$. Since $n \geq 1$ we may assume that $m \geq 2$. We identify R with the subring of constant polynomials in $R[x]$.

If $P_{m-1} \cap R = 0$ then $P_{m-2} \cap R \neq 0$ by Lemma 5.4.18. Pick a non-zero $a \in P_{m-2} \cap R$. Now P_{m-2} has height at least 2 and is not minimal over (a) by Krull's Haupidealsatz. By Lemma 5.4.3, P_{m-2} contains a prime ideal P'_{m-1}, which is minimal prime over (a). Then

$$P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_{m-2} \supsetneq P'_{m-1} \supsetneq 0$$

is a chain of prime ideals of $R[x]$ in which $P'_{m-1} \cap R \supseteq (a) \neq 0$. Therefore, we may assume that $P_{m-1} \cap R \neq 0$.

Then $\mathfrak{p} = P_{m-1} \cap R$ is a non-zero prime ideal of R. By Lemma 5.4.17(ii), $\dim R/\mathfrak{p} \leq \dim R - 1 = n - 1$. By the inductive hypothesis, we have $\dim((R/\mathfrak{p})[x]) \leq n$. The projection $R \to R/\mathfrak{p}$ induces a homomorphism $R[x] \to (R/\mathfrak{p})[x]$ whose kernel $P = \mathfrak{p}[x] \subseteq P_{m-1}$ is a non-zero prime ideal of $R[x]$. Then the chain

$$P_0/P \supsetneq P_1/P \supsetneq \cdots \supsetneq P_{m-1}/P$$

of prime ideals of $R[x]/P \cong (R/\mathfrak{p})[x]$ shows that $m - 1 \leq n$, and $m \leq n + 1$. □

Corollary 5.4.20 *Let* F *be a field. Then* $\dim F[x_1, \ldots, x_n] = n$.

5.5 Introduction to Algebraic Geometry

Algebraic geometry is the subject which studies algebraic varieties. Naively, algebraic varieties are just algebraic sets.

Throughout this section we fix an algebraically closed ground field F. (It is much harder to develop algebraic geometry over non-algebraically closed fields and we will not try to do this).

Definition 5.5.1 Let $S \subseteq F[x_1, \ldots, x_n]$. A *zero* of the set S is an element $(x_1, \ldots, x_n)$ of F^n such that $f(x_1, \ldots, x_n) = 0$ for all $f \in S$. The *zero set* of S is the set $Z(S)$ of all zeros of S. An *algebraic set* in F^n (or *affine algebraic set*) is the zero set of some set $S \subseteq F[x_1, \ldots, x_n]$, in which case S is called a set of *equations* of the algebraic set.

Example 5.5.2 The straight line $x+y-1=0$ and the 'circle' $x^2+y^2-1=0$ are examples of algebraic sets in $\mathbb{C}^2$. More generally, algebraic sets in $\mathbb{C}^2$ with a single equation are called complex algebraic curves. Note that the curve given by the equation $(x+y-1)(x^2+y^2-1)=0$ is the union of the line and the 'circle' above. On the other hand, the zero set of $\{x+y-4, x^2+y^2-1\}$ consists of two points $(1,0)$ and $(0,1)$. Finally, two more examples: $\varnothing = Z(1)$, and $\mathbb{C}^2 = Z(0)$.

Note that $Z(S) = Z((S))$, where (S) is the ideal of $F[x_1, \ldots, x_n]$ generated by S. Therefore every algebraic set is the zero set of some ideal. Since $F[x_1, \ldots, x_n]$ is noetherian by Hilbert's Basis Theorem, every algebraic set is the sero set of a finite set of polynomials.

Proposition 5.5.3

(i) *Every intersection of algebraic sets is an algebraic set; the union of finitely many algebraic sets is an algebraic set.*
(ii) *F^n and $\varnothing$ are algebraic sets in F^n.*

Proof (i) Let $(V_j = Z(I_j))_{j \in J}$ be a family of algebraic sets, given as zero sets of certain ideals I_j. To see that their intersection is again an algebraic set, it is enough to note that $\cap_{j \in J} Z(I_j) = Z(\sum_{j \in J} I_j)$. For the union, let $Z(I)$ and $Z(J)$ be algebraic sets corresponding to ideals I and J, and note that $Z(I) \cup Z(J) = Z(I \cap J)$ (why?).
(ii) $F^n = Z(0)$ and $\varnothing = Z(1)$. □

The proposition above shows that algebraic sets in F^n are closed sets of some topology. This topology is called the *Zariski topology.* Zariski

toplology on F^n also induces Zariski topology on any algebraic set in F^n. This topology is very weird and it takes time to get used to it. The main unintuitive thing here is that the topology is 'highly non-Hausdorf'—its open sets are huge. For example, on $\mathbb{C}$ closed sets are exactly the finite sets, and so any two non-empty open sets intersect non-trivially.

The most important theorem of algebraic geometry is called Hilbert's Nullstellensatz (or theorem on zeros). It has many equivalent reformulations and many corollaries. The idea of the theorem is to relate algebraic sets in F^n (geometry) and ideals in $F[x_1, \dots, x_n]$ (commutative algebra). We have two obvious maps

$$Z : \{\text{ideals in } F[x_1, \dots, x_n]\} \to \{\text{algebraic sets in } F^n\}$$

and

$$I : \{\text{algebraic sets in } F^n\} \to \{\text{ideals in } F[x_1, \dots, x_n]\}.$$

We have already defined $Z(J)$ for an ideal J in $F[x_1, \dots, x_n]$. As for I, let V be any subset of F^n. Then the ideal $I(V)$ is defined to be

$$I(V) := \{f \in F[x_1, \dots, x_n] \mid f(z_1, \dots, z_n) = 0 \text{ for all } (z_1, \dots, z_n) \in Z\}.$$

Lemma 5.5.4 *Let V be any subset of F^n. Then $Z(I(V)) = \bar{V}$, the closure of V in Zariski topology. In particular, if V is an algebraic set, then $Z(I(V)) = V$.*

Proof We have to show that for any algebraic set $Z(J)$ containing V we actually have $Z(I(V)) \subseteq Z(J)$. Well, as $V \subseteq Z(J)$, we have $I(V) \supseteq J$, which in turn implies $Z(I(V)) \subseteq Z(J)$. □

Note, however, that Z and I do not give us a one-to one correspondence. For example, in F^1 we have $Z((x)) = Z((x^2)) = \{0\}$, that is the different ideals (x) and (x^2) give the same algebraic set. Also, note that $I(\{0\}) = (x) \neq (x^2)$. Nullstellensatz sorts out problems like this in a very satisfactory way.

The first formulation of the Nulltellensatz is as follows (don't forget that F is assumed to be algebraically closed throughout the section):

Theorem 5.5.5 (Hilbert's Nullstellensatz) *Let J be an ideal of $F[x_1, \dots, x_n]$. Then $I(Z(J)) = \sqrt{J}$.*

Proof First of all, it is easy to see that $\sqrt{J} \subseteq I(Z(J))$. Indeed, let

$f \in \sqrt{J}$. Then $f^n \in J$. Then f^n is zero at every point of $Z(J)$. But this implies that f is zero at every point of $Z(J)$, i.e. $f \in I(Z(J))$.

The converse is much deeper. Let $f \in I(Z(J))$ and assume that no power of f belongs to J. Applying Lemma 5.1.8 to the multiplicative set $\{1, f, f^2, \dots\}$ yields a prime ideal P containing J but not f. Let $R = F[x_1, \dots, x_n]/P$ and $\pi : F[x_1, \dots, x_n] \to R$ be the natural projection. Then R is a domain which is generated over $\pi(F) \cong F$ by $\alpha_1 := \pi(x_1), \dots, \alpha_n := \pi(x_n)$. We identify F and $\pi(F)$, and so π can be considered as a homomorphism of F-algebras. Under this agreement, $y := f(\alpha_1, \dots, \alpha_n) = \pi(f) \neq 0$, non-zero element of R, as $f \notin P$.

By Proposition 5.3.23, the identity isomorphism $F \to F$ can be extended to a homomorphism ψ from the subring $F[\alpha_1, \dots, \alpha_n, \frac{1}{y}]$ of the fraction field of R to F. Then $\psi(y) \neq 0$. So

$$f(\psi(\alpha_1), \dots, \psi(\alpha_n)) = \psi(f(\alpha_1, \dots, \alpha_n)) = \psi(y) \neq 0.$$

On the other hand, for any $g \in J \subseteq P$ we have

$$\begin{aligned} g(\psi(\alpha_1), \dots, \psi(\alpha_n)) &= \psi(g(\alpha_1, \dots, \alpha_n)) = \psi(g(\pi(x_1), \dots, \pi(x_n))) \\ &= \psi(\pi(g(x_1, \dots, x_n))) = \psi(\pi(g)) = \psi(0) = 0. \end{aligned}$$

Thus $(\psi(\alpha_1), \dots, \psi(\alpha_n))$ is a zero of J but not of f, i.e. $f \notin I(Z(J))$, a contradiction. □

Definition 5.5.6 We say that an ideal I of a commutative ring R is *radical* (or *semiprime*) if $\sqrt{I} = I$.

The following corollary is also often called Nullstellensatz.

Corollary 5.5.7 *The maps I and Z induce an order-reversing bijection between algebraic sets in F^n and radical ideals in $F[x_1, \dots, x_n]$.*

Proof Note that $I(V)$ is always a radical ideal for any subset $V \subseteq F^n$. Now the result follows from Theorem 5.5.5 and Lemma 5.5.4. □

Corollary 5.5.8 *Let J_1 and J_2 be two ideals of $F[x_1, \dots, x_n]$. Then $Z(J_1) = Z(J_2)$ if and only if $\sqrt{J_1} = \sqrt{J_2}$.*

Proof It is clear that $Z(J) = Z(\sqrt{J})$ for any ideal J, which gives the 'if'-part. The converse follows from Theorem 5.5.5. □

Corollary 5.5.9 *Every proper ideal of $F[x_1, \ldots, x_n]$ has at least one zero in F^n.*

Proof If $\sqrt{I} = F[x_1, \ldots, x_n]$, then $I = F[x_1, \ldots, x_n]$. Now the result follows from above. □

Let $a = (a_1, \ldots, a_n) \in F^n$. Denote $I(\{a\})$ by M_a, i.e.

$$M_a = \{f \in F[x_1, \ldots, x_n] \mid f(a_1, \ldots, a_n) = 0\}.$$

Corollary 5.5.10 *The mapping $a \to M_a$ is a one-to-one correspondence between F^n and the maximal ideals of $F[x_1, \ldots, x_n]$.*

Proof Note that the maximal ideals are radical and apply Nullstellensatz. □

The following gives another strange property of the Zariski topology.

Corollary 5.5.11 *The Zariski topology in F^n is noetherian, i.e. its open sets satisfy the ascending chain condition.*

Proof An ascending chain of open sets corresponds to a descending chain of closed sets, which, by the Nullstellensatz, corresponds to an ascending chain of radical ideals of $F[x_1, \ldots, x_n]$, which stabilizes since $F[x_1, \ldots, x_n]$ is noetherian. □

In the noetherian ring $F[x_1, \ldots, x_n]$ every ideal I is a reduced intersection of primary ideals with distinct radicals $P_1, \ldots, P_r$, which are determined uniquely up to a permutation. If I is radical, then

$$I = \sqrt{I} = \sqrt{Q_1 \cap \cdots \cap Q_r} = \sqrt{Q_1} \cap \cdots \cap \sqrt{Q_r} = P_1 \cap \cdots \cap P_r, \quad (5.2)$$

and the presentation is unique up to a permutation of the prime ideals P_i.

Definition 5.5.12 An algebraic set V is called *irreducible* if $V = Z(P)$ for some prime ideal P.

The following result reduces the study of algebraic sets to that of irreducible algebraic sets or, equivalently, to the study of prime ideals in $F[x_1, \ldots, x_n]$.

Corollary 5.5.13 *Every algebraic set is uniquely an irredundant finite union of irreducible algebraic sets.*

Proof Let V be an algebraic set. By the Nullstellensatz, $V = Z(I)$ for some radical ideal I. Write I as the irredundant intersection of primes, see (5.2). Then $V = Z(P_1 \cap \cdots \cap P_r) = Z(P_1) \cup \cdots \cup Z(P_r)$, and this decomposition is irredundant because of the Nulstellensatz. Moreover, if $V = V_1 \cup \cdots \cup V_s$ is an irredundant decomposition of V as a union of irreducible algebraic sets, then $I = I(V) = I(V_1 \cup \cdots \cup V_n) = I(V_1) \cap \cdots \cap I(V_n)$ is a presentation of I as the intersection of primes, and by uniqueness of such presentation $I(V_i) = P_i$ for every i (up to a permutation). Now, by Nullstellensatz, $V_i = Z(P_i)$ for every i. □

Another description of irreducible algebraic sets, which explains the term better, is as follows:

Lemma 5.5.14 *An algebraic set is irreducible if and only if it is not the union of two smaller algebraic sets.*

Proof The 'if' part follows from Corollary 5.5.13. Conversely, if V is an irreducible algebraic set and $V = V_1 \cup V_2$, a union of smaller algebraic sets, then $I(V) = I(V_1) \cap I(V_2)$, and the $I(V_i)$ contain $I(V)$ properly by the Nullstellensatz. Then $I(V_1)I(V_2) \subseteq I(V_1) \cap I(V_2) \subseteq I(V)$ contradicts the fact that $I(V)$ is prime. □

Corollary 5.5.15 *An algebraic set V is irreducible if and only if any open subset of V is dense in V.*

Proof If $V = Z_1 \cup Z_2$, the union of two smaller algebraic sets, then the open sets $U_1 = V \setminus Z_1$ and $U_2 = V \setminus Z_2$ do not intersect, whence $\bar{U}_1$ is not dense. Conversely, if an open set U is not dense then there is an open set W with with $U \cap W = \varnothing$, and $V = (V \setminus U) \cup (V \setminus W)$ is a union of two smaller algebraic sets. □

Now, we will try to define the dimension of an algebraic set V. We should be able to do it in terms of the ideal $I(V)$. If there is justice in the world our definition should be such that $\dim(F^n) = n$ and $\dim(\text{point}) = 0$. Also, in view of Corollary 5.5.13, we may work with irreducible algebraic sets. The following definition will do the job (see Definition 5.4.15, Example 5.4.16(i), and Corollary 5.4.20):

Definition 5.5.16 The *dimension* of an irreducible algebraic set V, written $\dim V$, is defined to be the Krull dimension of the domain $F[x_1, \ldots, x_n]/I(V)$.

The following lemma gives equivalent descriptions of dimension.

Lemma 5.5.17 *Let V be an irreducible algebraic set, $d = \dim V$, and $P = I(V)$. Then*

(i) *d is the length of the longest chain $P_0 \supsetneq P_1 \supsetneq \cdots \supsetneq P_d = P$ of prime ideals of $F[x_1, \ldots, x_n]$ containing P.*
(ii) *d is the length of the longest chain $V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_d = V$ of irreducible algebraic sets contained in V.*

Proof (i) follows from the definition of Krull dimension, as the quotient $F[x_1, \ldots, x_n]/I(V)$ is a domain. (ii) follows from (i) and the Nullstellensatz. □

Let $V \subseteq F^n$ be an algebraic set. Every polynomial $f \in F[x_1, \ldots, x_n]$ defined an F-valued function on F^n and hence on V via restriction. Such functions are called *regular* functions on V. The regular functions form a ring (even an F-algebra) with respect to the point-wise operations. The ring is called the *coordinate algebra* (or *coordinate ring*) of V (or simply the *ring of regular functions* on V) and denoted $F[V]$. Clearly,

$$F[V] \cong F[x_1, \ldots, x_n]/I(V).$$

If I is an ideal of $F[V]$ then we write $Z(I)$ for the set of all points $a \in V$ such that $f(a) = 0$ for every $f \in I$, and if Z is a subset of V we denote by $I(Z)$ the ideal of $F[V]$ which consists of all functions $f \in F[V]$ such that $f(z) = 0$ for every $z \in Z$. Note that closed subsets of V all look like $Z(I)$.

Now the Nullstellensatz and the correspondence theorem for the ideals imply:

Theorem 5.5.18 (Hilbert's Nullstellensatz) *Let V be an algebraic set.*

(i) *If J be an ideal of $F[V]$, then $I(Z(J)) = \sqrt{J}$.*
(ii) *The maps I and Z induce an order-reversing bijection between closed sets in V and radical ideals in $F[V]$.*
(iii) *Every proper ideal of $F[V]$ has at least one zero in V.*
(iv) *The mapping $a \to \mathfrak{m}_a = \{f \in F[V] \mid f(a) = 0\}$ is a one-to-one correspondence between V and the maximal ideals of $F[V]$.*

Proposition 5.5.19

(i) *Let V be an algebraic set. Then $F[V]$ is a commutative finitely generated F-algebra with* $\operatorname{Rad} F[V] = 0$. *If V is irreducible, then $F[V]$ is a domain.*

(ii) *Every commutative finitely generated F-algebra A with* $\operatorname{Rad} A = 0$ *is isomorphic to $F[V]$ for some algebraic set V.*

Proof (i) clear. For (ii), if $A = F[\alpha_1, \ldots, \alpha_n]$ is an F-algebra generated by $\alpha_1, \ldots, \alpha_n$, then by the universal property of polynomial rings, $A \cong F[x_1, \ldots, x_n]/I$ for some ideal I. As $\operatorname{Rad} A = 0$, the ideal I is radical, and so $I = I(V)$ for some algebraic set V by the Nulltellensatz. □

If V is irreducible, then $F[V]$ is a domain. Then the quotient field of $F[V]$ is called the *field of rational functions* on V and denoted $F(V)$. In a natural way, $F(V)$ is a field extension of F. As F is algebraically closed, $F(V)$ is purely transcendental. It can be shown that $\operatorname{tr.deg}(F(V)/F) = \dim V$.

We now define morphisms between algebraic sets. Let $V \subseteq F^n$, $W \subseteq F^m$ be two algebraic sets and $f : V \to W$ be a mapLet $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be the coordinate functions on F^n and F^m, respectively. Denote $y_i \circ f$ by f_i for all $1 \leq i \leq n$. So that we can think of f as the m-tuple of functions $f = (f_1, \ldots, f_m)$, where $f_i : V \to F$, and $f(a) = (f_1(a), \ldots, f_m(a)) \in F^m$. The map $f : V \to W$ is called a morphism of algebraic sets (or a *regular map* from V to W) if each function $f_i : V \to F$, $1 \leq i \leq n$ is a regular function on V. It is easy to see that algebraic sets and regular maps form a category, in particular a composition of regular maps is a regular map again.

Now, let $f : V \to W$ be a morphism of algebraic sets as above. This morphism defines the 'dual' morphism $f^\sharp : F[W] \to F[V]$ of coordinate algebras, as follows:

$$f^\sharp : F[W] \to F[V] : \ \theta \mapsto \theta \circ f.$$

It is clear that $f^\sharp$ is a homomorphism of F-algebras. Moreover, $(f \circ g)^\sharp = g^\sharp \circ f^\sharp$ and $\mathrm{id}^\sharp = \mathrm{id}$, i.e. we have a contravariant functor $\mathcal{F}$ from the category of algebraic sets to the category of finitely generated f-algebras with $\operatorname{Rad} A = 0$. To reiterate: $\mathcal{F}(V) = F[V]$ and $\mathcal{F}(f) = f^\sharp$. If we restrict our attention to irreducible algebraic sets, then the functor is an equivalence categories:

Theorem 5.5.20 *The functor $\mathcal{F}$ from the category of irreducible alge-*

braic sets to the category of finitely generated commutative F-algebras without zero divisors is an equivalence of categories.

Proof In view of Theorem 4.2.8 and Proposition 5.5.19(ii) we just need to show that $f \mapsto f^\sharp$ establishes a one-to one correspondence between regular mappings $f : V \to W$ and algebra homomorphisms $F[W] \to F[V]$, for arbitrary fixed irreducible algebraic sets $V \subseteq F^n$ and $W \subseteq F^m$. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be the coordinate functions on F^n and F^m, respectively.

Let $\varphi : F[W] \to F[V]$ be an F-algebra homomorphism. Set $q_j := y_j|_W \in F[w]$, $1 \leq j \leq m$. Then $\varphi(q_j)$ are regular functions on V. Define the regular map $\varphi_\sharp : V \to F^m$ as follows:

$$\varphi_\sharp := (\varphi(q_1), \ldots, \varphi(q_m)).$$

We claim that in fact $\varphi_\sharp(V) \subseteq W$. Indeed, let $a \in V$ and $f = \sum_{\mathbf{k}} c_{\mathbf{k}} y_1^{k_1} \ldots y_m^{k_m} \in I(W)$, where $\mathbf{k}$ stands for the m-tuple $(k_1, \ldots, k_m)$. Then $f(q_1, \ldots, q_m) = 0$ in $F[W]$, and, since φ is a homomorphism, we have

$$\begin{aligned} f(\varphi_\sharp(a)) &= f(\varphi(q_1)(a), \ldots, \varphi(q_m)(a)) \\ &= \sum_{\mathbf{k}} c_{\mathbf{k}} (\varphi(q_1)(a))^{k_1} \ldots (\varphi(q_m)(a))^{k_m} \\ &= \varphi(\sum_{\mathbf{k}} c_{\mathbf{k}} y_1^{k_1} \ldots y_m^{k_m})(a) \\ &= \varphi(f(q_1, \ldots, q_m))(a) = 0. \end{aligned}$$

Thus $\varphi_\sharp(V) \subseteq W$.

Now, to complete the proof of the theorem, it suffices to check that $(f^\sharp)_\sharp = f$ and $(\varphi_\sharp)^\sharp = \varphi$ for any regular map $f : V \to W$ and any F-algebra homomorphism $\varphi : F[W] \to F[V]$. Well, indeed,

$$(f^\sharp)_\sharp = (f^\sharp(q_1), \ldots, F^\sharp(q_m)) = (f_1, \ldots, f_m) = f.$$

On the other hand,

$$((\varphi_\sharp)^\sharp)(q_i) = q_i \circ \varphi_\sharp = \varphi(q_i)$$

for any $1 \leq i \leq m$. As the q_i generate $F[W]$, this implies $(\varphi_\sharp)^\sharp = \varphi$. □

Corollary 5.5.21 *Two irreducible algebraic sets are isomorphic if and only if their coordinate algebras are isomorphic.*

Lemma 5.5.22 *Regular maps are continuous in the Zariski topology.*

Proof Let $f : V \to W \subseteq F^m$ be a regular map. As the topology on W is induced by that on F^m, it suffices to prove that any regular map $f : V \to F^m$ is continuous. Let $Z = Z(I)$ be a closed subset of F^m. We claim that $f^{-1}(Z) = Z(J)$ where J is the ideal of $F[V]$ generated by $f^\sharp(I)$. Well, if $a \in Z(J)$, then $\varphi(f(a)) = f^\sharp(\varphi)(a) = 0$ for any $\varphi \in I$, so $f(a) \in Z(I)$, i.e. $a \in f^{-1}(Z)$. The argument is easily reversed. □

Remark 5.5.23 Note that regular maps from V to W usually do not exhaust all continuous maps from V to W, so the category of algebraic sets is not a full subcategory of the category of topological spaces. For example, if $V = W = \mathbb{C}$, the closed subsets in V and W are exactly the finite subsets, and there are lots of non-polynomial maps from $\mathbb{C}$ to $\mathbb{C}$ such that inverse image of a finite subset is finite (describe one!).

5.6 Problems on Commutative Algebra

Problem 5.6.1 True or false? $J(\mathbb{R}[x]) = 0$.

Solution. True: for every $a \in \mathbb{R}$ the ideal I_a is maximal and $\cap_{a\in\mathbb{R}} I_a = 0$.

Problem 5.6.2 Let $R = R[[x]]$. Calculate $J(R)$ and $\operatorname{Rad} R$.

Solution. $J(R) = (x)$, as this is the only maximal ideal in R. On the other hand $\operatorname{Rad} R = 0$, as R has no nilpotent elements.

Problem 5.6.3 The ideal $(4, 2x, x^2)$ in the ring $\mathbb{Z}[x]$ is primary but not irreducible.

Solution. The elements of $I := (4, 2x, x^2)$ are all polynomials $f(x) = a_0 + a_1x + \cdots \in \mathbb{Z}[x]$ with a_0 divisible by 4 and a_1 divisible by 2. It follows that $\sqrt{I} = (2)$. Now, let $(a_0 + a_1x + \dots)(b_0 + b_1x + \dots) \in I$. Assume that $(a_0+a_1x+\dots) \notin I$. Then either a_0 is not divible by 4 or a_1 is odd. In both cases it follows that b_0 is even, i.e. $(b_0 + b_1x + \dots) \in I$. We have proved that I is primary. On the other hand, $I = (4, x)\cap(2, x^2)$.

Problem 5.6.4 The ideal $I = (x^2, 2x)$ in $\mathbb{Z}[x]$ is not primary, but $(x)^2 \subset I \subset (x)$ and the ideal (x) is prime.

Solution. Obvious.

Problem 5.6.5 Represent the ideal $(9, 3x+3)$ in $\mathbb{Z}[x]$ as the intersection of primary ideals.

Solution. $(9, 3x + 3) = (3) \cap (9, x + 1)$. Indeed, it is easy to see that (3) and $(9, x + 1)$ are primary containing $(9, 3x + 3)$. On the other hand, if $r \in (3) \cap (9, x + 1)$, then r can be written as $9(a_0 + a_1x + \dots) + (x + 1)(b_0 + b_1x + \dots)$. The first summand clearly belongs to $(9, 3x+3)$, and the second one belongs to it too, because $r \in (3)$ implies that all b_i are divisible by 3.

Problem 5.6.6 Let I be an ideal of a commutative ring R such that $\sqrt{I}$ is a maximal ideal in R. Prove that I is primary.

Solution. $J := \sqrt{I}$ is the intersection of primes containing I. Since J is maximal, it is the only prime containing I. So, if $x + I$ is a zero divisor of the local ring R/I, then $x \in J$, whence $(x + I)$ is nilpotent.

Problem 5.6.7 In a noetherian ring, prove that $\sqrt{I}$ is the intersection of the associated prime ideals of I.

Solution. $\sqrt{I} = \sqrt{Q_1 \cap \dots \cap Q_r} = \sqrt{Q_1} \cap \dots \sqrt{Q_r}$, using the easily checked fact that $\sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$.

Problem 5.6.8 Let S be the set of all non-zero elements of $\mathbb{Z}_n$ which are not zero divisors. Determine $S^{-1}\mathbb{Z}_n$.

Solution. All non-zero divisors in $\mathbb{Z}_n$ happen to be units, so $S^{-1}\mathbb{Z}_n \cong \mathbb{Z}_n$.

Problem 5.6.9 Let S be a multiplicative subset of R and T be a multiplicative subset of $S^{-1}R$. Let $S_* = \{r \in R \mid [\frac{r}{s}] \in T \text{ for some } s \in S\}$. Then S_* is a multiplicative subset of R and there is a ring isomorphism $S_*^{-1}R \cong T^{-1}(S^{-1}R)$.

Solution. That S^* is multiplicative follows from the fact that so is T. It is also easy to see that $S^* \supseteq S$. Now, the universal property of localizations gives homomorphisms

$$\alpha : S_*^{-1}R \to T^{-1}(S^{-1}R),\ \left[\frac{r}{s}\right] \mapsto \left[\frac{[\frac{r}{s'}]}{[\frac{s}{s'}]}\right],$$

where $s' \in S$ is such that $[\frac{s}{s'}] \in T$, and

$$\beta : T^{-1}(S^{-1}R) \to S_*^{-1}R,\ \left[\frac{[\frac{r}{s}]}{[\frac{t}{s'}]}\right] \mapsto \left[\frac{rs'}{st}\right]$$

(note that the last formula makes sense, as $t \in S^*$, $s \in S \subseteq S^*$ and S^* is multiplicatively closed). Finally, it is clear that $\alpha\beta = \mathrm{id}$ and $\beta\alpha = \mathrm{id}$.

Problem 5.6.10 True or false? If R is a local ring, then there is a commutative ring R' and a prime ideal P of R' such that $R \cong R'_P$.

Solution. We can take $R' = R$ and P to be the maximal ideal of R.

Problem 5.6.11 Let P be a prime ideal of R. Show that R_P/P_P is isomorphic to the field of quotients of R/P.

Solution. Use the universal property of localizations to construct the homomorphism $f : R_P \to Q(R/P)$, which maps $[\frac{r}{s}]$ to $\frac{r+P}{s+P}$. It is easy to see that f is surjective and has P_P as its kernel.

Problem 5.6.12 True or false? The ideal $(x^2, 4)$ is $(x, 2)$-primary in $\mathbb{Z}[x]$.

Problem 5.6.13 Let R be a noetherian local ring with maximal ideal M and let $x_1, \ldots, x_n \in M$. Suppose that $\{x_1 + M^2, \ldots, x_n + M^2\}$ is a basis of the R/M-vector space M/M^2. Show that $M = Rx_1 + \cdots + Rx_n$.

Solution. Let N be an R-submodule of M generated by $x_1, \ldots, x_n$. Then $M = N + MM$. Since A is noetherian, M is finitely generated. We can therefore apply Nakayama's Lemma to deduce that $N = M$.

Problem 5.6.14 Let R be a commutative ring with a unique prime ideal P and let $r \in R$. Prove that x is nilpotent if and only if $x \in P$.

Problem 5.6.15 True or false? $\mathbb{C}[x, y]$ is a PID.

Problem 5.6.16 True or false? $\mathbb{C}[x, y]$ is a noetherian ring.

Problem 5.6.17 True or false? Every subring of an artinian ring is artinian.

Problem 5.6.18 True or false? $\mathbb{Z}[x]$ is a UFD.

Problem 5.6.19 True or false? $\mathbb{Z}[x, y]/(x - 2y)$ is a noetherian ring.

Problem 5.6.20 True or false? If R is a PID then R is noetherian.

Problem 5.6.21 True or false? If R is a noetherian ring then every non-zero prime ideal of R is maximal.

Problem 5.6.22 True or false? If R is a UFD then every non-zero prime ideal of R is maximal.

Problem 5.6.23 True or false? (a) Every quotient ring of a UFD is a UFD. (b) Every subring of a UFD is a UFD.

Problem 5.6.24 True or false? The intersection of two prime ideals of a commutative ring is prime.

Problem 5.6.25 True or false? If R is a local ring with maximal ideal M and V is a finitely generated R-module with $(R/M) \otimes_R V = 0$ then $V = 0$.

Solution. True: use $(R/M) \otimes_R V \cong V/MV$ and Nakayama's lemma.

Problem 5.6.26 True or false? If R is an integral domain and $r \in R$ is an irreducible element of R, then (r) is a prime ideal of R.

Problem 5.6.27 Show: if Q is a primary ideal in a commutative ring R, then $\sqrt{Q}$ is a prime ideal. Show that the converse holds if R is a PID.

Problem 5.6.28 Let R be a domain and F be its field of fractions. Then $r \in R$ is a unit if and only if $\frac{1}{r} \in F$ is integral over R.

Problem 5.6.29 True or false? Let S be a proper multiplicative subset of a commutative ring R and $I \neq J$ be ideals of R. Then $S^{-1}I \neq S^{-1}J$.

Solution. False. Take $R = \mathbb{Z}$, $S = 2\mathbb{Z}$, $I = 3\mathbb{Z}$, and $J = 5\mathbb{Z}$. Then $S^{-1}I = S^{-1}J = \mathbb{Z}_{(2)}$.

Problem 5.6.30 True or false? Let S be a proper multiplicative subset of a commutative ring R and I be an ideal of R. Then $\sqrt{S^{-1}I} = S^{-1}\sqrt{I}$.

Solution. True. If $a^n = [\frac{x}{s}]^n \in S^{-1}I$ then $s_1 x^n \in I$ for some $s_1 \in S$, whence $s_1 x \in \sqrt{I}$ and $a = [\frac{s_1 x}{s_1 s}] \in S^{-1}\sqrt{I}$. Conversely if $x^n \in I$, then $[\frac{x}{s}]^n \in S^{-1}I$ for every $s \in S$.

Problem 5.6.31 The ring $R = \mathbb{R}[x, y]$ is localized at the multiplicative set $S = \{f \in R \mid f(0,0) \neq 0\}$. Find all maximal ideals of $S^{-1}R$ and its Jacobson radical.

Problem 5.6.32 Let R be an integral domain with a quotient field F. Prove that for any maximal ideal M of R, R_M can be canonically embedded into F, and $\cap R_M = R$.

Solution. Clearly, $R \subseteq \cap R_M$. Conversely, take $x \in \cap R_M$. If $x \notin R$ set $I := \{y \in R \mid yx \in R\}$. Then I is a proper ideal, and so I is contained in a maximal ideal M. Now, $x \in R_M$ means $x = \frac{z}{w}$ for $z \in R$ and $w \in R \setminus M$. So $xw = z \in R$, a contradiction, since $w \notin I$.

Problem 5.6.33 Let R be a commutative ring and V be an R-module. Show that $V = 0$ if and only if $V_M = 0$ for every maximal ideal M of R.

Solution. 'Only if' is clear. Conversely, let $v \in V$ be a non-zero element. Pick a maximal ideal M containing the annihilator of v in R. Then it follows that $[\frac{v}{1}] \neq 0$ in V_M.

Problem 5.6.34 Let R be a domain which is integral over a ring R. Prove that A is a field if and only R is a field.

Solution. By Lemma 5.3.21, if R is a field, so is A. The converse follows from Problem 5.6.71(i).

Problem 5.6.35 True or false? The ring $\mathbb{Q}[x, y]$ is integrally closed.

Problem 5.6.36 True or false? The ring $\mathbb{Q}(x)[y]$ is integrally closed.

Solution. True, as $\mathbb{Q}(x)[y]$ is a UFD.

Problem 5.6.37 True or false? The ring $\mathbb{Z}[x]$ is integrally closed.

Problem 5.6.38 Let R be a commutative integral domain. Prove that if R_P is integrally closed for every prime ideal P of R, then R is integrally closed.

Solution. By Problem 5.6.32, we have $R = \cap R_P$ in the ring of quotients of R. Now $\bar{R} = \overline{\cap R_P} \subseteq \cap \bar{R}_P = \cap R_P = R$.

Problem 5.6.39 Let R be a commutative noetherian local ring with maximal ideal M which satisfies $M^2 = M$. Prove that R is a field.

Solution. By Nakayama's Lemma, $M = 0$.

Problem 5.6.40 True or false? Every finite local ring is a field.

Solution. False: $\mathbb{Z}_4$.

Problem 5.6.41 True or false? A local artinian ring has finitely many prime ideals.

Solution. True. Let M be the maximal ideal of a local artinian ring R. As R is artinian, Nakayama's Lemma (or Krull's Intersection Theorem) implies that $M^n = 0$ for some n, in particular, all elements of M are nilpotent. Now, if P is a prime ideal of R, then R/P is a domain, whence $P = M$.

Problem 5.6.42 Let $R = \mathbb{R}[x, y]$,

$$I = \{f(x, y) \in R \mid f(0, 0) = \frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0\}.$$

Check that I is an ideal of R. Is it maximal, prime, primary or neither? What are the associated prime ideals of I.

Solution. That $I = (x^2, y^2, xy)$ is an ideal follows from Leibnitz's rule or by seeing directly that $I = (x^2, y^2, xy)$. Clearly, the ideal is not prime (and so not maximal). It is easy to check that I is primary and it is clear that $\sqrt{I} = (x, y)$.

Problem 5.6.43 Let M be a maximal ideal of $\mathbb{Q}[x, y, z]$. Prove that $F = \mathbb{Q}[x, y, z]/M$ is a finite algebraic extension of $\mathbb{Q}$.

Problem 5.6.44 Let R be a domain. Then R is integrally closed if and only if $R[x]$ is integrally closed.

Solution. If $R \neq \bar{R}$, then $\bar{R} \subseteq \overline{R[x]}$ implies $\bar{R}[x] \subseteq \overline{R[x]}$. Hence $R[x] \subsetneq \bar{R}[x] \subsetneq \overline{R[x]}$. Conversely, let $\bar{R} = R$, and take $\frac{f(x)}{g(x)} \in Q(R[x]) \cong Q(R)(x)$ integral over $R[x]$ and such that $\mathrm{GCD}(f, g) = 1$. So

$$\left(\frac{f(x)}{g(x)}\right)^n + h_{n-1}(x)\left(\frac{f(x)}{g(x)}\right)^{n-1} + \cdots + h_0(x) = 0,$$

whence

$$f(x)^n + h_{n-1}(x)f(x)^{n-1}g(x) + \cdots + h_0(x)g(x)^n = 0.$$

So $g(x)$ divides $f(x)^n$ (in $Q(R)(x)$), whence $g = 1$. Now need to prove

that integrality of f implies that the coefficients are in R, which is easy to do starting from the top coefficiant and going down.

Problem 5.6.45 Let $R \subseteq A$ be rings with R integrally closed in A. Suppose that $h(x)$ is a polynomial in $R[x]$ which factors in $A[x]$ as the product of two monic polynomials $h(x) = f(x)g(x)$. Show that f and g are each in $R[x]$.

Problem 5.6.46 An algebraic over $\mathbb{Q}$ element $\alpha \in \mathbb{C}$ is integral over $\mathbb{Z}$ if and only if $\operatorname{irr}(\alpha; \mathbb{Q}) \in \mathbb{Z}[x]$.

Problem 5.6.47 For each $n \in \mathbb{Z}$ find the integral closure of $\mathbb{Z}[\sqrt{n}]$ as follows:

(i) Reduce to the case where n is square-free.
(ii) Use the fact that $\sqrt{n}$ is integral to deduce that what we want is the integral closure R of $\mathbb{Z}$ in the field $\mathbb{Q}(\sqrt{n})$.
(iii) If $\alpha = a+b\sqrt{n}$ with $a, b \in \mathbb{Q}$, deduce that the minimal polynomial of α is $x^2 - \operatorname{Trace}(\alpha)x + \operatorname{Norm}(\alpha)$, where $\operatorname{Trace}(\alpha) = 2a$ and $\operatorname{Norm}(\alpha) = a^2 - b^2 n$. Thus, using Problem 5.6.46, $\alpha \in R$ if and only if $2a$ and $a^2 - b^2 n$ are integers.
(iv) Show that if $\alpha \in R$ then $a \in \frac{1}{2}\mathbb{Z}$. If $a = 0$ show that $\alpha \in R$ if and only if $b \in \mathbb{Z}$. If $a = \frac{1}{2}$ and $\alpha \in R$ show that $b \in \frac{1}{2}\mathbb{Z}$; thus, subtracting a multiple of $\sqrt{n}$, we may assume $b = 0$ or $b = \frac{1}{2}$; $b = 0$ is impossible.
(v) Conclude that the integral closure is $\mathbb{Z}[\sqrt{n}]$ if $n \not\equiv 1 \pmod 4$, and $\mathbb{Z}[\frac{1}{2} + \frac{1}{2}\sqrt{n}]$ otherwise.

Problem 5.6.48 Let $R = \mathbb{Z}[\sqrt{10}]$. Then R is integrally closed but R is not a UFD.

Hint: Integrally closed by Problem 5.6.47. On the other hand, $2 \cdot 3 = (4 + \sqrt{10})(4 - \sqrt{10})$ and the elements $2, 3, 4 \pm \sqrt{10}$ are irreducible.

Problem 5.6.49

(i) Find all prime ideals of $\mathbb{Z}[\sqrt{5}]$ which lie over the prime ideal (5) of $\mathbb{Z}$.
(ii) Find all prime ideals of $\mathbb{Z}[\sqrt{5}]$ which lie over the prime ideal (3) of $\mathbb{Z}$.
(iii) Find all prime ideals of $\mathbb{Z}[\sqrt{5}]$ which lie over the prime ideal (2) of $\mathbb{Z}$.

Problem 5.6.50 Let $P_1, \ldots, P_r$ be prime ideals of a commutative ring. Show that an ideal I which is contained in $P_1 \cup \cdots \cup P_r$ is contained in some P_i.

Problem 5.6.51 Let I be an ideal of a noetherian ring R with a reduced primary decomposition $I = Q_1 \cap \cdots \cap Q_r$. Show that every prime ideal of R which is minimal over I is the radical of some Q_i. Is the converse true?

Problem 5.6.52 Let P be a prime ideal of height r in a noetherian ring R. Show that there exists an ideal I of R with r generators over which P is minimal.

Hint. Construct $x_1, \ldots, x_r \in P$ by induction so that the prime ideals that are minimal over $Rx_1 + \cdots + Rx_i$ have height i, for every $i \leq r$.

Problem 5.6.53 Let P be a prime ideal of R. Show that the height of P is the dimension of R_P.

Problem 5.6.54 Let P be a non-zero prime ideal of R. Show that $\dim R \geq 1 + \dim R/P$.

Problem 5.6.55 Let F be an algebraically closed field. Show that the minimal prime ideals of $F[x_1, \ldots, x_n]$ are the principal ideals generated by irreducible polynomials.

Problem 5.6.56 True or false? Let F be a field. If S is an arbitrary subset of $F[x_1, \ldots, x_n]$, then there is a finite subset T of $F[x_1, \ldots, x_n]$ such that $Z(S) = Z(T)$.

Problem 5.6.57 True or false? Let F be a field. If I is any ideal of $F[x_1, \ldots, x_n]$, then $I = \{f \in F[x_1, \ldots, x_n] \mid f(a) = 0 \text{ for each } a \in Z(I)\}$.

Problem 5.6.58 True or false? Let F be algebraically closed and I, J be ideals in $F[x_1, \ldots, x_n]$. Then $Z(I) \cup Z(J) = Z(IJ)$.

Problem 5.6.59 True or false? Let F be a field, and I, J be ideals in $F[x_1, \ldots, x_n]$. Then $\sqrt{I \cap J} = \sqrt{IJ}$.

Problem 5.6.60 Let I and J be ideals of $A = \mathbb{C}[x, y]$ and $Z(I) \cap Z(J) = \varnothing$. Show that $A/(I \cap J) \cong A/I \times A/J$.

Problem 5.6.61 True or false? Let F be algebraically closed. Any decreasing sequence of algebraic sets in F^n stabilizes.

Problem 5.6.62 True or false? Let F be algebraically closed. Any increasing sequence of algebraic sets in F^n stabilizes.

Problem 5.6.63 True or false? Let F be algebraically closed. Any increasing sequence of irreducible algebraic sets in F^n stabilizes.

Problem 5.6.64 True or false? Let F be an algebraically closed field. A system of polynomial equations

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

over F has no solutions in F^n if and only if 1 can be expressed as a linear combination $1 = \sum_i p_i f_i$ with polynomial coefficients p_i.

Problem 5.6.65 True or false? The Zariski topology on F^{m+n} is the product topology of the Zariski topologies on F^m and F^n.

Problem 5.6.66 Let R be any ring. Denote by $\operatorname{Spec} R$ the set of the prime ideals of R. For an ideal $I \lhd R$ denote

$$Z(I) := \{P \in \operatorname{Spec} R \mid P \supseteq I\}.$$

Introduce the *Zariski topology* on $\operatorname{Spec} R$ by declaring the sets of the form $Z(I)$ to be closed. Define a *distinguished open set* of $\operatorname{Spec} R$ to a set of the form

$$U(f) := \{P \in \operatorname{Spec} R \mid f \notin P\}$$

where $f \in R$.

(i) Prove that this is indeed a topology.

(ii) Prove that distinguished open sets are indeed open, and moreover, they form a basis of the Zariski topology. Show that $\operatorname{Spec} R = \cup_i U(f_i)$ for some collection f_i of elements of R if and only if the ideal generated by all the f_i is R.

(iii) Prove that $\operatorname{Spec} R$ is compact in the Zariski topology.

Problem 5.6.67 Let F be an algebraically closed field of characteristic $p > 0$, and $\mathrm{Fr} : F \to F, a \mapsto a^p$ be the Frobenius homomorphism. True or false:

(i) Fr is a homeomorphism in the Zariski topology.
(ii) Fr is an isomorphism of algebraic sets.

Problem 5.6.68 Describe all automorphisms of the algebraic set F.

Problem 5.6.69 Which of the following algebraic sets over $\mathbb{C}$ are isomorphic to each other?

(i) $\mathbb{C}$;
(ii) $Z(x) \subset \mathbb{C}^2$;
(iii) $Z(y - x^2) \subset \mathbb{C}^2$;
(iv) $Z(y^2 - x^3) \subset \mathbb{C}^2$;
(v) $Z(y^2 - x^3 - x^2) \subset \mathbb{C}^2$.

Problem 5.6.70 Let W be an irreducible closed subset of an irreducible algebraic set V. Then $\dim W \leq \dim V$ and equality is attained if and only if $V = W$.

Problem 5.6.71 Let $A \supseteq R$ be an integral ring extension.

(i) If $a \in R$ is a unit in A, then a is also a unit in R.
(ii) $J(R) = R \cap J(A)$.

Solution. (i) As A/R is integral, we have

$$(a^{-1})^n + r_{n-1}(a^{-1})^{n-1} + \cdots + r_0 = 0,$$

for some $r_i \in R$. Npw multiply by a^{n-1} to get

$$a^{-1} = -r_{n-1} - \cdots - r_0 a^{n-1} \in R.$$

(ii) Follows from Theorem 5.3.20.

Problem 5.6.72 Let $A \supseteq R$ be an integral ring extension. If every non-zero prime ideal of R is a maximal ideal, then every non-zero prime ideal of A is also maximal.

Solution. By Theorem 5.3.19, if 0 is a prime ideal of R, then the only prime ideal of A lying over 0 is 0. Now it follows from Theorem 5.3.20 that if P is a non-zero prime ideal of A, then P is maximal, since $P \cap R$ is non-zero prime.

Bibliography

[AF] F. Anderson and K. Fuller, *Rings and categories of modules*, Springer-Verlag, 1974.

[Ca] P. Cameron, *Permutation groups*, Cambridge University Press, 1999.

[Hu] J.E. Humphreys, *Reflection groups and Coxeter Groups*, Cambridge University Press, 1990.

[MM] G. Malle and B.H. Matzat, *Inverse Galois theory*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1999.

[Ol] A. Yu. Olshanskii, *Geometry of defining relations in groups*, Kluwer Academic Publishers Group, Dordrecht, 1991.

[Vo] H.Völklein, *Groups as Galois groups*, Cambridge Studies in Advanced Mathematics, **53**, Cambridge University Press, Cambridge, 1996.

[Wi] H. Wielandt, *Finite Permutation groups*, Academic Press, 1964.