

# Graduate Algebra

## Definition 2.3.1

## Contents

1. Week 1 [2025-09-03]	3
1.1. Notations and recollections	3
1.2. Groups	3
1.3. Rings	4
1.4. Fields	4
1.5. Linear Algebra	4
1.6. Cayley's Theorem	5
1.7. A linear analogue of Cayley's Theorem.	6
2. Week 2 [2025-09-08]	8
2.1. The Quotient of a set by an equivalence relation	8
2.2. Sub-groups	10
2.3. Group actions	12
2.4. Quotients of groups	12
2.5. Quotients of groups and orbits.	14
2.6. The product of subgroups	15
2.7. Lagrange's Theorem	17
3. Week 3 [2025-09-15]	19
3.1. Normal subgroups	19
3.2. Quotient groups	21
3.3. First isomorphism theorem	23
3.4. Groups acting on Groups	23
3.5. $p$ -groups	25
4. Week 4 [2025-09-22]	27
4.1. Sylow subgroups	27
4.2. Applications of Sylow's Theorem	29
4.3. Rings	33
4.4. Ring homomorphisms and kernels	34
5. Week 5 [2025-09-29]	36
5.1. Quotient rings	36
5.2. Categories	38
5.3. Modules	42
5.4. The direct sum of $R$ -modules.	43
5.5. Free modules	44
6. Week 6 [2025-10-06]	46
6.1. Algebras	50
6.2. Integral Domains and prime ideals	50
6.3. Monoid algebras	52
6.4. The polynomial ring over $R$	55
6.5. Zorn's Lemma	56
6.6. Existence of maximal ideals.	57

7. Week 7 [2025-10-14]	59
7.1. The rank of a free $R$ -module	59
7.2. Principal Ideal Domains	59
7.3. Euclidean domains	61
7.4. The polynomial ring over a field.	62
7.5. The Chinese Remainder Theorem	65
8. Week 9 [2025-10-27]	69
8.1. Generation of algebras	69
8.2. Unique factorization	69
8.3. PIDs have unique factorization	72
8.4. Rings of fractions	74
8.5. Unique factorization in polynomial rings	78
9. Week 10 [2025-11-03]	83
9.1. Eisenstein's irreducibility criterion	83
9.2. Extension of scalars: an important special case of the tensor product	84
9.3. Functors between categories	87
9.4. Bilinear mappings	88
9.5. The tensor product of modules	89
9.6. Tensor products and direct sums.	91
9.7. Tensor products for PIDs.	92
10. Week 11 [2025-11-10]	93
10.1. Hom and tensor functors	93
10.2. The Jordan-Hölder theorem and composition series for a module	95
10.3. Localization at a prime ideal, local rings and Nakayama's lemma	97
11. week 12 [2025-11-17]	102
11.1. Modules over a principal ideal domain	102
11.2. Exactness of tensoring with a localization of $A$	106
11.3. The torsion submodule	109
11.4. Finitely generated torsion modules over a PID.	110
11.5. Describing $p$ -primary torsion modules over a PID.	113
12. week 13 [2025-12-01]	115
12.1. Recapitulating: finitely generated modules over a PID	115
12.2. Finitely generated modules over a PID via Invariant factors	115
12.3. Jordan decomposition of a matrix over a field	117
12.4. Algebraic extensions of fields	122

## 1. Week 1 [2025-09-03]

We'll begin by recalling some basic sorts of algebra that you more-or-less encountered before.

### 1.1. Notations and recollections

We reserve the following letters:

- $\mathbb{N}$  for the set of *natural numbers*  $0, 1, 2, \dots$
- $\mathbb{Z}$  for the set of *integers*, i.e. for all  $\pm n$  for  $n \in \mathbb{N}$
- $\mathbb{Q}$  for the set of *rational numbers*  $m/n$  for  $m, n \in \mathbb{Z}$  with  $n \neq 0$
- $\mathbb{R}$  for the set of *real numbers*, and
- $\mathbb{C}$  for the set of *complex numbers*  $a + bi$  for  $a, b \in \mathbb{R}$ .

In this first lecture, I want to recall some of the main objects of study in algebra, including: groups, rings and fields. Ultimately, the goal today is to prove an analogue of Cayley's Theorem - see [Theorem 1.6.1](#) and [Theorem 1.7.1](#) about embedding arbitrary groups in some standard groups.

### 1.2. Groups

Recall that a group is a set  $G$  together with a binary operation  $\cdot : G \times G \rightarrow G$  satisfying the following:

- associativity:  $\forall x, y, z \in G, (xy)z = x(yz)$
- identity:  $\exists e \in G, xe = ex = x$ .
- inverses:  $\forall x \in G, \exists y \in G, xy = yx = 1$ .

*Remark 1.2.1:*

- We usually write  $1$  or sometimes  $1_G$  rather than  $e$  for the identity element of  $G$ . + we usually write  $x^{-1}$  for the inverse of  $x \in G$
- there are *uniqueness* results that I'm eliding here; the identity  $1$  of  $G$  is unique, and the inverse  $x^{-1}$  of an element is unique. These statements are *consequences* of the above axioms (they don't require additional assumption.)
- A group is abelian if  $\forall a, b \in G, ab = ba$
- Sometimes we write groups additively; in that case,  $0$  is the identity element and the inverse of  $a \in G$  is  $-a \in G$ . We always insist that additive groups are abelian.

**Definition 1.2.2:** For groups  $G$  and  $H$ , a function  $\varphi : G \rightarrow H$  is a **group homomorphism** provided that  $\forall x, y \in G, \varphi(xy) = \varphi(x)\varphi(y)$ .

**Definition 1.2.3:** Let  $\varphi : G \rightarrow H$  be a group homomorphism. The **kernel** of  $\varphi$  is

$$\ker \varphi = \{g \in G \mid \varphi(g) = 1.\}$$

*Remark 1.2.4:* If  $\varphi : G \rightarrow H$  is a group homomorphism,  $\ker \varphi$  is a subgroup of  $G$  - i.e.  $\ker \varphi$  is non-empty, and is closed under multiplication and under taking inverses.

**Proposition 1.2.5:** Let  $\varphi : G \rightarrow H$  be a group homomorphism. Then  $\varphi$  is an injective (or one-to-one) function if and only if  $\ker \varphi = \{1_G\}$ .

### 1.3. Rings

**Definition 1.3.1:** A **ring** is an additive abelian group  $R$  together with a binary operation of multiplication

$$\cdot : R \times R \rightarrow R$$

which satisfies the following:

- multiplication is associative:  $\forall a, b, c \in R, (ab)c = a(bc)$ .
- there is a multiplicative identity:  $\exists 1 \in R, \forall a \in R, 1a = a1 = a$ .
- distribution laws:  $\forall a, b, c \in R, a(b + c) = ab + ac$  and  $(b + c)a = ba + ca$ .

The ring  $R$  is **commutative** provided that  $\forall a, b \in R, ab = ba$ .

*Example 1.3.2:*

- $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  are commutative rings.
- For a natural number  $n > 1$ , the ring  $\text{Mat}_n(\mathbb{Z})$  of  $n \times n$  matrices with coefficients in  $\mathbb{Z}$  is a non-commutative ring.

**Definition 1.3.3:** For a commutative ring  $R$ , an element  $a \in R$  is a **unit** provided that  $\exists v \in R, uv = vu = 1$ .

The set  $R^\times$  of units in  $R$  is a group under the multiplication of  $R$ .

### 1.4. Fields

**Definition 1.4.1:** A **field** is a commutative ring  $F$  such that  $\forall a \in F, a \neq 0 \Rightarrow a$  is a unit.

*Example 1.4.2:*  $\mathbb{Q}, \mathbb{R}, \mathbb{C}$  are fields, but  $\mathbb{Z}$  is not a field.

### 1.5. Linear Algebra

**Definition 1.5.1:** If  $F$  is a field, a vector space over  $F$  – or an  $F$ -vector space – is an additive abelian group  $V$  together with an operation of scalar multiplication

$$F \times V \rightarrow V$$

written  $(t, v) \mapsto tv$ , subject to the following:

- identity:  $\forall v \in V, 1v = v$ .
- associativity:  $\forall a, b \in F$  and  $v \in V, a(bv) = (ab)v$ .
- distributive laws:  $\forall a, b \in F$  and  $v, w \in V, (a + b)v = av + bv$  and  $a(v + w) = av + aw$ .

*Remark 1.5.2:* Probably in a linear algebra class you saw results stated for vector spaces over  $\mathbb{R}$  or  $\mathbb{C}$ ; however, “most” results in linear algebra remain valid for vector space over  $F$ .

*Example 1.5.3:* Let  $I$  be any set, and let  $V$  be the set of all functions  $f : I \rightarrow F$  which have finite support. Recall that the support of  $f$  is  $\{x \in I \mid f(x) \neq 0\}$ .

Then  $V$  is a vector space. (The addition and scalar multiplication operations are define “pointwise” – see homework.)

*Remark 1.5.4:* Recall that a **basis** of a vector space is subset  $B$  of  $V$  which is linearly indepent and spans  $V$ .

The vector space of finitely supported functions  $I \rightarrow F$  has a basis  $B = \{\delta_i \mid i \in I\}$ , where

$$\delta_i : I \rightarrow F$$

is the function defined by  $\delta_i(j) = 0$  if  $i \neq j$  and  $\delta_i(i) = 1$ .

**Definition 1.5.5:** If  $V$  and  $W$  are  $F$ -vector spaces, an  $F$ -linear map  $\varphi : V \rightarrow W$  is a homomorphism of additive groups which satisfies the condition

$$\forall t \in F, \forall v \in V, \varphi(tv) = t\varphi(v).$$

**Definition 1.5.6:** If  $V$  is an  $F$ -vector space, the general linear group  $\text{GL}(V)$  is the set

$$\{\varphi : V \rightarrow V \mid \varphi \text{ is } F\text{-linear and invertible.}\}$$

$\text{GL}(V)$  is a group whose operation is given by composition of linear transformations.

*Remark 1.5.7:* If  $V$  is finite dimensional, so that  $V$  is isomorphic to  $F^n$  as  $F$ -vector spaces, linear algebra shows that  $\text{GL}(V)$  is isomorphic to the group  $\text{GL}_n$  of  $n \times n$  matrices with non-zero determinant, where  $n = \dim_F V$  and where the operation in  $\text{GL}_n$  is given by matrix multiplication.

## 1.6. Cayley’s Theorem

Let  $\Omega$  be any set. The set  $S(\Omega)$  of all bijective functions  $\psi : \Omega \rightarrow \Omega$  is a group whose operation is composition of functions.

**Theorem 1.6.1** (Cayley’s Theorem): Let  $G$  be any group. Then  $G$  is isomorphic to a subgroup of  $S(\Omega)$  for some  $\Omega$ .

*Proof:* Let  $\Omega = G$ . For  $g \in G$ , define a mapping  $\lambda_g : G \rightarrow G$  by the rule

$$\lambda_g(h) = gh.$$

We are going to argue that the mapping  $g \mapsto \lambda_g$  defines an injective group homomorphism  $G \rightarrow S(\Omega) = S(G)$ .

First of all, we note that  $\lambda_1 = \text{id}$ . Indeed, to check this identity of functions, let  $h \in \Omega = G$ . Then

$$\lambda_1(h) = 1h = h = \text{id}(h);$$

this confirms  $\lambda_1 = \text{id}$ .

Next, we note that for  $g_1, g_2 \in G$ , we have  $(*) \quad \lambda_{g_1} \circ \lambda_{g_2} = \lambda_{g_1 g_2}$ . Again, to confirm this identify of functions, we let  $h \in \Omega = G$ . Then

$$(\lambda_{g_1} \circ \lambda_{g_2})h = \lambda_{g_1}(\lambda_{g_2}(h)) = \lambda_{g_1}(g_2h) = g_1(g_2h) = (g_1g_2)h = \lambda_{g_1g_2}(h)$$

as required.

Now, using (\*) we see for  $g \in G$  that  $\lambda_g \circ \lambda_{g^{-1}} = \lambda_1 = \text{id} = \lambda_{g^{-1}} \circ \lambda_g$ , which proves that  $\lambda_g$  is bijective; thus indeed  $\lambda_g \in S(\Omega) = S(G)$ .

Moreover, (\*) shows that the mapping  $\lambda : G \rightarrow S(G)$  given by  $g \mapsto \lambda_g$  is a group homomorphism.

It remains to see that  $\lambda$  is injective. If  $g \in \ker \lambda$ , then  $\lambda_g = \text{id}$ . Thus  $1 = \text{id}(1) = \lambda_g(1) = g1 = g$ . Thus  $g = 1$  so that  $\ker \lambda = \{1\}$  which confirms that  $\lambda$  is injective by [Proposition 1.2.5](#). This completes the proof. ■

## 1.7. A linear analogue of Cayley's Theorem.

Let  $F$  be a field.

**Theorem 1.7.1:** Let  $G$  be any group. Then  $G$  is isomorphic to a subgroup of  $\text{GL}(V)$  for some  $F$ -vector space  $V$ .

*Proof:* The proof is quite similar to the proof of Cayley's Theorem.

Let  $V$  be the vector space of all finitely supported functions  $f : G \rightarrow F$ . Recall that  $V$  has a basis  $B = \{\delta_g \mid g \in G\}$ .

We are going to define an injective group homomorphism  $G \rightarrow \text{GL}(V)$ .

For  $g \in G$ , we may define an  $F$ -linear mapping  $\lambda_g : V \rightarrow V$  by defining the value of  $\lambda_g$  at each vector in  $B$ . We set  $\lambda_g(\delta_h) = \delta_{gh}$ .

Recall that a typical element  $v$  of  $V$  has the form

$$v = \sum_{i=1}^n t_i \delta_{h_i}$$

for scalars  $t_i \in F$  and elements  $g_i \in G$ ; since  $\lambda_g$  is  $F$ -linear, we have

$$\lambda_g(v) = \sum_{i=1}^n t_i \delta_{gh_i}.$$

We now show that  $\lambda_1 = \text{id}$ . To prove this, since the functions  $V \rightarrow V$  are linear, it is enough to argue that the functions agree at each element of the basis  $B$  of  $V$ . Well, for  $h \in G$ ,

$$\lambda_1(\delta_h) = \delta_{1h} = \delta_h = \text{id}(\delta_h)$$

as required.

We next show for  $g_1, g_2 \in G$  that (\*)  $\lambda_{g_1} \circ \lambda_{g_2} = \lambda_{g_1g_2}$ . Again, it suffices to argue that these functions agree at each element  $\delta_h$  of  $B$ . For  $h \in G$  we have:

$$(\lambda_{g_1} \circ \lambda_{g_2})(\delta_h) = \lambda_{g_1}(\lambda_{g_2}\delta_h) = \lambda_{g_1}(\delta_{g_2h}) = \delta_{g_1(g_2h)} = \delta_{(g_1g_2)h} = \lambda_{g_1g_2}\delta_h$$

as required.

Now, for  $g \in G$  we see that by  $(*)$  that

$$\text{id} = \lambda_1 = \lambda_g \circ \lambda_{g^{-1}}$$

which proves that  $\lambda_g$  is invertible and hence in  $\text{GL}(V)$ .

Moreover,  $(*)$  shows that the assignment  $\lambda : G \rightarrow \text{GL}(V)$  given by the rule  $g \mapsto \lambda_g$  is a group homomorphism.

It remains to argue that  $\lambda$  is injective. Suppose that  $x \in \ker \lambda$ , so that  $\text{id} = \lambda_x$ .

Then  $\delta_1 = \text{id}(\delta_1) = \lambda_x(\delta_1) = \delta_{x1} = \delta_x$ . This implies that  $1 = x$  so that indeed the kernel of  $\lambda$  is trivial and thus  $\lambda$  is injective by [Proposition 1.2.5](#).

■

## 2. Week 2 [2025-09-08]

This week, we'll discuss **quotients**, and we'll begin our discussion of **group actions**.

### 2.1. The Quotient of a set by an equivalence relation

Let  $S$  be a set and let  $R$  be a relation on  $S$ . Formally,  $R$  is an assignment  $R : S \times S \rightarrow \text{Prop}$  – in other words, for  $a, b \in S$ ,  $R(a, b)$  is the **proposition** that  $a$  and  $b$  are related; of course  $R(a, b)$  may or may not hold.

We often use a symbol  $\sim$  or  $\sim_R$  to indicate this proposition; thus  $R(a, b) \Leftrightarrow a \sim_R b$ .

**Definition 2.1.1:** The relation  $\sim$  is an **equivalence relation** if the following properties hold:

- **reflexive:**  $\forall s \in S, s \sim s$ .
- **symmetric:**  $\forall s_1, s_2 \in S, s_1 \sim s_2 \Rightarrow s_2 \sim s_1$
- **transitive:**  $\forall s_1, s_2, s_3 \in S, s_1 \sim s_2 \text{ and } s_2 \sim s_3 \Rightarrow s_1 \sim s_3$

**Definition 2.1.2:** If  $\sim$  is an equivalence relation on the set  $S$ , a **quotient** of  $S$  by  $\sim$  is a set  $\bar{S}$  together with a surjective function  $\pi : S \rightarrow \bar{S}$  with the following properties:

(Quot 1)  $\forall a, b \in S, a \sim b \Rightarrow \pi(a) = \pi(b)$

(Quot 2) Let  $T$  be any set and let  $f$  be any function  $f : S \rightarrow T$  such that  $\forall a, b \in S, a \sim b \Rightarrow f(a) = f(b)$ . Then there is a function  $\bar{f} : \bar{S} \rightarrow T$  for which  $f = \bar{f} \circ \pi$ .

**Proposition 2.1.3:** Suppose that  $(\bar{S}_1, \pi_1)$  and  $(\bar{S}_2, \pi_2)$  are two quotients of the set  $S$  by the equivalence relation  $\sim$ . Let

$$\bar{\pi}_2 : \bar{S}_1 \rightarrow \bar{S}_2$$

be the mapping determined by the quotient property for  $(\bar{S}_1, \pi_1)$  using

$$T = \bar{S}_2 \text{ and } f = \pi_2 : S \rightarrow \bar{S}_2,$$

and let

$$\bar{\pi}_1 : \bar{S}_2 \rightarrow \bar{S}_1$$

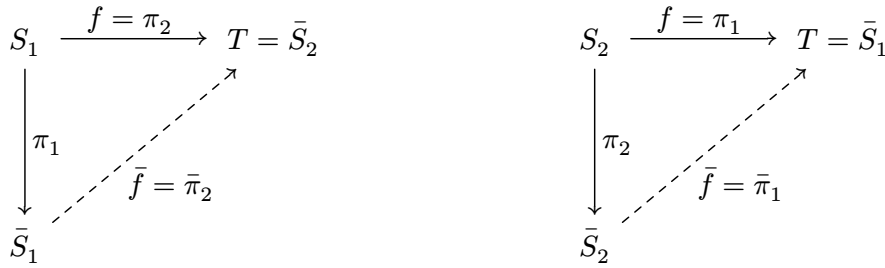
be the mapping determined by the quotient property for  $(\bar{S}_2, \pi_2)$  using

$$T = \bar{S}_1 \text{ and } f = \pi_1 : S \rightarrow \bar{S}_1.$$

Then the maps  $\pi'_2 : \bar{S}_1 \rightarrow \bar{S}_2$  and  $\pi'_1 : \bar{S}_2 \rightarrow \bar{S}_1$  are inverse to one another, and in particular  $\pi'_1$  and  $\pi'_2$  are bijections.

*Proof:* By the definition of quotients, we have commutative diagrams





In particular, we have  $\pi_2 = \bar{\pi}_2 \circ \pi_1$  and  $\pi_1 = \bar{\pi}_1 \circ \pi_2$

Substitution now yields

$$\pi_1 = \bar{\pi}_1 \circ \bar{\pi}_2 \circ \pi_1$$

and

$$\pi_2 = \bar{\pi}_2 \circ \bar{\pi}_1 \circ \pi_2$$

Since  $\pi_1$  and  $\pi_2$  are surjective, we conclude that  $\text{id} = \bar{\pi}_1 \circ \bar{\pi}_2$  and  $\text{id} = \bar{\pi}_2 \circ \bar{\pi}_1$  so indeed the indicated functions are inverse to one another. ■

*Remark 2.1.4:* The point of the Proposition is that a quotient is completely determined by the property indicated in the definition – this property is an example of what is known as a **universal property** or sometimes as a **universal mapping property**. The conclusion of the Proposition shows that any two ways of constructing a quotient are equivalent in a strong sense.

One way of constructing the quotient is by considering equivalence classes, as follows:

**Definition 2.1.5:** For an equivalence relation  $\sim$  on a set  $S$ , the **equivalence class**  $[s]$  of an element  $s \in S$  is the subset of  $S$  defined by

$$[s] = \{x \in S \mid x \sim s\}.$$

**Proposition 2.1.6:** Equivalence classes for the equivalence relation  $\sim$  have the following properties for arbitrary  $s, s' \in S$ :

- a.  $s \sim s' \Leftrightarrow [s] = [s']$
- b.  $[s] \neq [s'] \Leftrightarrow [s] \cap [s'] = \emptyset$

*Proof:* Review! ■

**Theorem 2.1.7 (Existence of quotients):** For any equivalence relation  $\sim$  on a set  $S$ , there is a quotient  $(\bar{S}, \pi)$ .

*Proof:* We consider the set  $\bar{S} = \{[s] \mid s \in S\}$  of equivalence classes and the mapping  $\pi : S \rightarrow \bar{S}$  given by the rule  $\pi(s) = [s]$ .

[Proposition 2.1.6](#) confirms condition (a) of [Definition 2.1.2](#).

For condition (b) of [Definition 2.1.2](#) suppose that  $T$  is a set and that  $f : S \rightarrow T$  is a function with the property that  $\forall a, b \in S, a \sim b \Rightarrow f(a) = f(b)$ . We must exhibit a function  $\bar{f} : \bar{S} \rightarrow T$  with the property  $f = \bar{f} \circ \pi$ . If  $\bar{f}$  exists, it must satisfy  $\bar{f}([a]) = f(a)$  for  $a \in S$ . On the other hand, in view of [Proposition 2.1.6](#) (a), the rule  $[a] \mapsto f(a)$  indeed determines a well-defined function  $\bar{f} : \bar{S} \rightarrow T$ . Moreover, the identity  $f = \bar{f} \circ \pi$  evidently holds.

■

*Remark 2.1.8:* We gave an explicit construction of the quotient using equivalence classes. On the other hand, if one has a quotient  $(\bar{S}, \pi)$ , the equivalence class  $[x]$  of an element  $x \in S$  is equal to  $\pi^{-1}(\pi(x))$ .

**Proposition 2.1.9:** If  $\sim$  is an equivalence relation on the set  $S$ , then  $S$  is the disjoint union of the equivalence classes.

*Proof:* Each element  $x \in S$  is contained in the equivalence classes  $[x]$ , so it only remains to prove that if two equivalence classes have a common element, they are equal. For this, let  $x, y \in S$  and suppose that  $z \in [x] \cap [y]$ . Then  $x \sim z$  and  $y \sim z$  so that  $x \sim y$  by transitivity; thus  $[x] = [y]$ .

■

## 2.2. Sub-groups

Let  $G$  be a group (when giving definitions, we'll write  $G$  multiplicatively).

**Definition 2.2.1:** A **subgroup** of  $G$  is a non-empty subset  $H \subseteq G$  such that  $H$  is closed under the operations of multiplication in  $G$  and inversion in  $G$ . In other words,

$$\forall a, b \in G, ab \in H \text{ and } a^{-1} \in H$$

*Example 2.2.2:* Consider the group  $G = \mathbb{Z} \times \mathbb{Z}$  where the operation is componentwise addition. Check the following!

- $H_1 = \{(a, b) \in G \mid 2a + 3b = 0\}$  is a subgroup.
- $H_2 = \{n(2, 2) + m(1, 2) \mid n, m \in \mathbb{Z}\}$  is a subgroup.

The collection of subgroups of  $G$  has a natural partial order given by *containment*.

**Proposition 2.2.3:** (Constructing subgroups)

- If  $H_i$  for  $i \in I$  is a family of subgroups of  $G$ , indexed by some set  $I$ , then the intersection  $\bigcap_{i \in I} H_i$  is again a subgroup of  $G$ .
- Let  $S \subseteq G$  be a subset. There is a unique smallest subgroup  $H(S) = \langle S \rangle$  containing  $S$ . In other words, for any subgroup  $H'$  of  $G$  with  $S \subseteq H'$ , we have  $\langle S \rangle \subseteq H'$ .

*Remark 2.2.4:*

- If  $S, T \subseteq G$  are subsets, we often write  $\langle S, T \rangle$  for  $\langle S \cup T \rangle$ . If  $S = \{s_1, s_2, \dots, s_n\}$  we often write  $\langle S \rangle = \langle s_1, s_2, \dots, s_n \rangle$ .

- b. The subgroup in [Example 2.2.2\(b\)](#) is precisely  $\langle (2, 2), (1, 2) \rangle$ .
- c. For any group  $G$  and  $a \in G$ ,  $\langle a \rangle := \langle \{a\} \rangle$  is the **cyclic subgroup** generated by  $a$ . If  $G$  is multiplicative, then  $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$  while if  $G$  is additive then  $\langle a \rangle = \{na \mid n \in \mathbb{Z}\}$ .

**Proposition 2.2.5:** If  $\varphi : G \rightarrow H$  is a group homomorphism, then  $\ker \varphi$  is a subgroup of  $G$ .

**Proposition 2.2.6:** If  $X \subseteq G$  is a non-empty subset of  $G$ , then  $X$  is a subgroup if and only if  
(\*)  $\forall a, b \in X, ab^{-1} \in X$ .

*Proof:* ( $\Rightarrow$ ): Immediate from the definition of a subgroup.

( $\Leftarrow$ ): Assume that (\*) holds. We must show that  $X$  is a subgroup.

We first argue that  $X$  contains the identity element. Since  $X$  is non-empty, there is an element  $x \in X$ . Condition (\*) then shows that  $xx^{-1} = 1 \in X$  as required.

We now show that  $X$  is closed under inversion. Let  $x \in X$ . Since  $1 \in X$ , we apply (\*) with  $a = 1$  and  $b = x$  to learn that  $1x^{-1} = x^{-1} \in X$ , as required.

Finally, we show that  $X$  is closed under multiplication. Let  $x, y \in X$ . We have already seen that  $y^{-1} \in X$ . Now apply (\*) with  $a = x$  and  $b = y^{-1}$  to learn that

$$ab^{-1} = x(y^{-1})^{-1} = xy \in X$$

as required. ■

**Proposition 2.2.7:** Let  $f : G \rightarrow H$  be a group homomorphism and let  $S \subseteq G$  be any subset. Then

$$f(\langle S \rangle) \subseteq \langle f(S) \rangle.$$

*Proof:* Since  $f$  is a homomorphism, for any subgroup  $K \subseteq G$ , the image

$$f(K) = \{f(x) \mid x \in K\}$$

is a subgroup of  $H$ . Thus  $f(\langle S \rangle)$  is a subgroup containing  $\langle f(S) \rangle$  of  $H$  where  $f(S)$  is the image of the set  $S$  via the function  $f$ . It now follows from [Proposition 2.2.3](#) that  $f(\langle S \rangle)$  is contained in the subgroup  $\langle f(S) \rangle$  generated by  $f(S)$ , as required. ■

## 2.3. Group actions

**Definition 2.3.1:** Let  $G$  be a group and let  $\Omega$  be a set. An **action** of  $G$  on  $\Omega$  is a mapping

$$G \times \Omega \rightarrow \Omega \text{ written } (g, x) \mapsto gx$$

such that for each  $x \in \Omega$  we have

- $1x = x$
- $\forall g, h \in G, (gh)x = g(hx)$ .

For brevity, sometimes we say that  $\Omega$  is a  $G$ -space.

**Proposition 2.3.2:** An action of a group  $G$  on a set  $\Omega$  determines a homomorphism  $f : G \rightarrow S(\Omega)$  such that  $f(g)(x) = gx$  for  $g \in G$  and  $x \in \Omega$ .

Conversely, given a homomorphism  $f : G \rightarrow S(\Omega)$ , there is an action of  $G$  on  $\Omega$  given by  $gx = f(g)(x)$  for each  $g \in G$  and  $x \in \Omega$ .

**Definition 2.3.3:** Suppose that  $\Omega$  is a  $G$ -space. The  $G$ -conjugacy relation on  $\Omega$  is defined as follows: for  $x, y \in \Omega$ ,  $x \sim_G y$  provided that  $\exists g \in G, gx = y$ .

**Proposition 2.3.4:** The  $G$ -conjugacy relation on  $\Omega$  is an equivalence relation.

**Definition 2.3.5:** Let  $\Omega$  be a  $G$ -space, and let  $\varphi : \Omega \rightarrow \Omega / \sim$  be the quotient mapping for the  $G$ -conjugacy relation; see [Definition 2.1.2](#). For  $x \in \Omega$ , the **orbit**  $\mathcal{O}_x = Gx$  of  $G$  through  $x$  is the subset of  $\Omega$  defined by

$$\mathcal{O}_x = \varphi^{-1}(\varphi(x)).$$

Thus the  $G$ -orbits are the equivalence classes for the relation  $\sim_G$ ; see [Remark 2.1.8](#).

Equivalently, we have  $\mathcal{O}_x = \{gx \mid g \in G\}$

**Proposition 2.3.6:**  $\Omega$  is the disjoint union of the  $G$ -orbits in  $\Omega$ .

*Proof:* This follows from [Proposition 2.1.9](#). ■

*Remark 2.3.7:* Each orbit  $\mathcal{O}_x$  is itself a  $G$ -set.

## 2.4. Quotients of groups

Let  $G$  be a group and let  $H$  be a subgroup of  $G$ . There is an action of  $H$  on the set  $G$  by right multiplication: for  $h \in H$  and  $g \in G$  we can define  $h \cdot g = gh^{-1}$ .

We are going to consider the quotient of  $G$  by the equivalence relation of  $H$ -conjugacy; this equivalence relation is defined by

$$g \sim g' \Leftrightarrow \exists h \in H, g = g'h.$$

**Definition 2.4.1:** The **left quotient of  $G$  by  $H$**  is the quotient  $(\pi, G/H)$  of  $G$  by the equivalence relation of  $H$ -conjugacy defined using the action of  $H$  on  $G$  by right multiplication as described above.

*Remark 2.4.2:*

- Of course, one can use an explicit model for the quotient by taking  $G/H$  to be the set of equivalence classes in  $G$  for the  $H$ -conjugacy relation.
- The equivalence classes for the relation of  $H$ -conjugacy defined by the action of right multiplication are precisely the **left cosets** of  $H$  in  $G$ . The class of  $x \in G$  has the form

$$xH = \{xh \mid h \in H\}$$

.

For  $x \in G$ ,

$$\pi^{-1}(\pi(x)) = xH.$$

- We can also consider the action of  $H$  on  $G$  by left multiplication. This action determines an equivalence relation of  $H$ -conjugacy, and the quotient of  $G$  by this equivalence relation is called the **right quotient of  $G$  by  $H$**  and is written  $(\pi, H \backslash G)$ . In this case, the equivalence classes are the **right cosets** where the class of  $x \in G$  has the form  $Hx = \{hx \mid h \in H\}$ .

For  $x \in G$ , we have  $\pi^{-1}(\pi(x)) = Hx$ .

**Proposition 2.4.3:** There is an action

$$\alpha : G \times G/H \rightarrow G/H$$

of the group  $G$  on the set  $G/H$  such that

$$\forall g, x \in G, \text{ we have } \alpha(g, \pi(x)) = \pi(gx)$$

where  $\pi : G \rightarrow G/H$  is the quotient map.

*Proof:* To define the action map  $\alpha$ , first fix  $g \in G$ . We are going to define the mapping

$$\alpha(g, -) : G/H \rightarrow G/H.$$

Consider the mapping  $\pi_g : G \rightarrow G/H$  given by the rule  $\pi_g(x) = \pi(gx)$ . This mapping has the property that  $x \sim_H x' \Rightarrow \pi_g(x) = \pi_g(x')$ . Indeed,

$$x \sim_H x' \Rightarrow \exists h, x = x'h \Rightarrow \pi_g(x) = \pi(gx) = \pi(gx'h) = \pi(gx') = \pi_g(x')$$

by the defining property of  $\pi$ ; see [Definition 2.1.2](#). Again using [Definition 2.1.2](#) we find the desired mapping  $\alpha(g, -) : G/H \rightarrow G/H$  with the property that

$$(\clubsuit) \quad \alpha(g, -) \circ \pi = \pi_g.$$

We now assemble the mappings  $\alpha(g, -)$  to get a mapping  $\alpha : G \times G/H \rightarrow G/H$  which satisfies  $\alpha(g, \pi(x)) = \pi(gx)$  for each  $g, x \in G$ , and it remains to check that  $\alpha$  determines an action as in [Definition 2.3.1](#).

Of course, using  $(\clubsuit)$ , we have  $\alpha(1, -) \circ \pi = \pi_1 = \pi = \text{id} \circ \pi$ ; since  $\pi$  is surjective, it follows that  $\alpha(1, -) = \text{id}$ . Thus  $\alpha(1, z) = z$  for each  $z \in G/H$ , which shows that  $\alpha$  satisfies the first requirement of [Definition 2.3.1](#).

Now suppose that  $g_1, g_2 \in G$ . To complete the proof, we must verify the remaining requirement of [Definition 2.3.1](#); thus we must show that

$$(\heartsuit) \quad \alpha(g_1, \alpha(g_2, -)) = \alpha(g_1 g_2, -)$$

On the one hand, using  $(\clubsuit)$  we find that

$$\alpha(g_1 g_2, -) \circ \pi = \pi_{g_1 g_2};$$

on the other hand, for  $z \in G$  we have

$$\begin{aligned} (\alpha(g_1, \alpha(g_2, -)) \circ \pi)(z) &= \alpha(g_1, \alpha(g_2, \pi(z))) \\ &= \alpha(g_1, \pi_{g_2}(z)) && \text{by } (\clubsuit) \\ &= \alpha(g_1, \pi(g_2 z)) \\ &= \pi(g_1(g_2 z)) && \text{by } (\clubsuit) \end{aligned}$$

Since  $\pi$  is surjective,  $(\heartsuit)$  follows at once. This completes the proof. ■

## 2.5. Quotients of groups and orbits.

**Definition 2.5.1:** Suppose that  $G$  acts on  $\Omega_1$  and on  $\Omega_2$ . A morphism of  $G$ -sets  $\varphi : \Omega_1 \rightarrow \Omega_2$  is a function  $\varphi$  with the property that  $\forall g \in G$  and  $\forall x \in \Omega_1$ , we have  $\varphi(gx) = g\varphi(x)$ .

The morphism of  $G$ -sets  $\varphi$  is an isomorphism (of  $G$ -sets) if there is a morphism of  $G$ -sets  $\psi : \Omega_2 \rightarrow \Omega_1$  such that  $\varphi \circ \psi = \text{id}$  and  $\psi \circ \varphi = \text{id}$ .

Suppose that  $G$  acts on  $\Omega$  and let  $x \in \Omega$ .

**Definition 2.5.2:** The **stabilizer of  $x$  in  $G$**  is the subgroup  $\text{Stab}_G(x) = \{g \in G \mid gx = x\}$ .

**Proposition 2.5.3:** Write  $H = \text{Stab}_G(x)$  and recall that  $\pi : G \rightarrow G/H$  is the quotient mapping. There is a unique isomorphism of  $G$ -sets  $\gamma : G/H \rightarrow \mathcal{O}_x$  with the property that

$$\gamma(\pi(1)) = x.$$

*Proof:* The rule  $g \mapsto gx$  determines a surjective mapping  $\alpha_x : G \rightarrow \mathcal{O}_x$ . Recall that the action of  $H$  on  $G$  by right multiplication determines an equivalence relation  $\sim$  on  $G$  used to construct the quotient  $G/H$ .

For  $g_1, g_2 \in G$  we find that

$$g_1 \sim g_2 \Rightarrow \exists h \in H, g_1 = g_2 h \Rightarrow \alpha(g_1) = \alpha(g_2 h) = g_2 h x = g_2 x = \alpha(g_2)$$

since  $h \in H = \text{Stab}_G(x) \Rightarrow hx = x$ .

Thus [Definition 2.1.2](#) shows that there is a mapping  $\gamma : G/H \rightarrow \mathcal{O}_x$  such that  $\gamma \circ \pi = \alpha_x$ . To see that  $\gamma$  is a morphism of  $G$ -sets, it suffices to show that  $(\clubsuit) \quad \forall g, g' \text{ we have}$

$$\gamma(g \cdot \pi(g')) = g \cdot \gamma(\pi(g')).$$

Now by the definition of the  $G$ -action on  $G/H$  we have  $g \cdot \pi(g') = \pi(gg')$ ; see [Proposition 2.4.3](#). Thus  $\gamma(g \cdot \pi(g')) = \gamma(\pi(gg')) = \alpha_x(gg') = gg' \cdot x$ . On the other hand,  $g \cdot \gamma(\pi(g')) = g \cdot \alpha_{x(g')} = g \cdot g' \cdot x$  which confirms  $(\clubsuit)$ . This shows that  $\gamma$  is indeed a morphism of  $G$ -sets.

Since  $\alpha_x$  is surjective and  $\gamma \circ \pi = \alpha_x$ , also  $\gamma$  is surjective. It only remains to see that  $\gamma$  is injective. Suppose that  $z, z' \in G/H$  such that  $\gamma(z) = \gamma(z')$ . Since  $\pi : G \rightarrow G/H$  is surjective, we may choose  $g, g' \in G$  with  $z = \pi(g)$  and  $z' = \pi(g')$ . Now

$$\gamma(z) = \gamma(z') \Rightarrow \gamma(\pi(g)) = \gamma(\pi(g')) \Rightarrow \alpha_{x(g)} = \alpha_{x(g')} \Rightarrow gx = g'x.$$

We now conclude that  $g^{-1}gx = x$  so that  $g^{-1}g \in \text{Stab}_G(x) = H$ . Since the quotient mapping  $\pi$  is constant on  $H$ -orbits,  $z = \pi(g) = \pi(gg^{-1}g') = \pi(g') = z'$ . This shows that  $\gamma$  is injective and completes the proof.  $\blacksquare$

**Definition 2.5.4:** The action of  $G$  on  $\Omega$  is **transitive** if there is a single  $G$ -orbit on  $\Omega$ . Equivalently, the action is transitive if the quotient  $\Omega / \sim$  is a singleton set.

*Example 2.5.5:* Let  $I$  be a set and let  $G = S(I)$  be the group of permutations of  $I$ . Fix  $x \in I$  and let  $H = \text{Stab}_G(x)$ . Notice that  $G$  acts on  $I$ . Moreover, the  $G$ -orbit of  $x$  is precisely  $I$  - in other words, the action of  $G$  on  $I$  is transitive.

Notice that  $H = S(I - \{x\})$ .

Now [Proposition 2.5.3](#) gives an isomorphism of  $G$ -sets  $G/H \rightarrow I$ ; i.e.  $S(I)/S(I - \{x\}) \rightarrow I$ .

## 2.6. The product of subgroups

**Definition 2.6.1:** If  $H, K \subseteq G$  are two subgroups, then  $H$  **normalizes**  $K$  if for each  $g \in H$  we have  $\text{Inn}_g K \subseteq K$  (in other words,  $\forall x \in K, gxg^{-1} \in K$ ).

**Definition 2.6.2:** Let  $H, K$  be subsets of  $G$ . The product of  $H$  and  $K$  is the subset

$$HK := \{xy \mid x \in H, y \in K\}$$

**Proposition 2.6.3:** Suppose that  $H, K$  are subgroups of  $G$  and that  $H$  normalizes  $K$ . Then  $\langle H, K \rangle = HK$ . In particular,  $HK$  is a subgroup of  $G$ .

*Proof:* Let  $X = HK$ . Since any subgroup of  $G$  which contains both  $H$  and  $K$  clearly contains  $X$ , it only remains to argue that  $X$  is a subgroup. For this, we use [Proposition 2.2.6](#). First note that  $1 = 1 \cdot 1 \in X$ , so  $X$  is non-empty. Now, let  $a_1, b_2 \in X$ . We must argue that  $a_1 a_2^{-1} \in X$ . By definition, there are elements  $x_1, x_2 \in H$  and  $y_1, y_2 \in K$  with  $a_i = x_i y_i$  for  $i = 1, 2$ . We now compute

$$a_1 a_2^{-1} = x_1 y_1 (x_2 y_2)^{-1} = x_1 y_1 y_2^{-1} x_2^{-1} = (x_1 x_2^{-1}) \cdot (x_2 y_1 y_2^{-1} x_2^{-1}).$$

We notice that  $x_1 x_2^{-1} \in H$ . Moreover,  $y_1 y_2^{-1} \in K$ ; since  $H$  normalizes  $K$  it follows that  $x_2 y_1 y_2^{-1} x_2^{-1} \in K$ .

We have now argued that  $a_1 a_2^{-1}$  has the form  $xy$  for  $x \in H$  and  $y \in K$  so that  $a_1 a_2^{-1} \in X$ . Now [Proposition 2.2.6](#) indeed shows that  $X = HK$  is a subgroup. ■

**Proposition 2.6.4:** Let  $H, K$  be subgroups of  $G$  and let  $\varphi : H \times K \rightarrow HK$  be the natural mapping given by  $\varphi(h, k) = hk$ .

- For each  $\alpha \in HK$ , the set  $\varphi^{-1}(\alpha)$  is in bijection with  $H \cap K$ .
- In particular, if  $H \cap K = \{1\}$ , then  $\varphi$  is bijective.

*Proof:* Let  $\alpha = hk \in HK$ . Note for any  $x \in H \cap K$  that  $\varphi(hx, x^{-1}k) = \alpha$  so that  $(hx, x^{-1}k) \in \varphi^{-1}(\alpha)$ . We argue that the mapping

$$\gamma : H \cap K \rightarrow \varphi^{-1}(\alpha) \text{ given by } \gamma(x) = (hx, x^{-1}k)$$

is bijective. Well, if  $(h_1, k_1) \in \varphi^{-1}(\alpha)$  then  $\varphi(h_1, k_1) = \varphi(h, k)$  so that  $h_1 k_1 = hk$  and thus  $h^{-1} h_1 = k k_1^{-1}$ . Now set  $x = h^{-1} h_1 = k k_1^{-1} \in H \cap K$  and observe that  $(h_1, k_1) = \gamma(x)$ . This shows that  $\gamma$  is surjective. To see that  $\gamma$  is injective, suppose that  $\gamma(x) = \gamma(x')$  for  $x \in H \cap K$ . Then

$$(hx, x^{-1}k) = (hx', x'^{-1}k) \Rightarrow hx = hx' \Rightarrow x = x'.$$

So  $\gamma$  is injective and the proof of a. is complete.

Now, the mapping  $\varphi$  is surjective by the definition of  $HK$ . To prove b. we suppose that

$$H \cap K = \{1\}.$$

According to a. the fiber  $\varphi^{-1}(\alpha)$  is a singleton for each  $\alpha \in HK$ ; this shows that  $\varphi$  is injective and confirms b. ■

**Corollary 2.6.5:** If  $G$  is a finite group and  $H, K$  subgroups of  $G$ , then

$$|HK| = |H| \cdot |K| / |H \cap K|.$$



*Proof:* This is a consequence of [Proposition 2.6.4](#). ■

Let's introduce some examples of groups in order to investigate this a bit more.

*Example 2.6.6:* For  $n \in \mathbb{N}$  with  $n \geq 1$ , consider the symmetric group  $S = S_n$  viewed as  $S(\mathbb{Z}/n\mathbb{Z})$  where  $\mathbb{Z}/n\mathbb{Z}$  denotes the collection of integers modulo  $n$ .

Consider the elements  $\sigma, \tau \in S$  defined by the rules  $\sigma(i) = i + 1$  and  $\tau(i) = -i$  where the addition and negation occurs in  $\mathbb{Z}/n\mathbb{Z}$ .

Viewed as permutations,  $\sigma$  identifies with an  $n$ -cycle and  $\tau$  identifies with a product of disjoint transpositions:

$$\sigma = (1, 2, \dots, n) \text{ and } \tau = (1, n-1)(2, n-2)\dots = \prod_{i=1}^{\lfloor \frac{n}{2} \rfloor} (i, n-i).$$

In particular,  $\sigma$  has order  $n$  and  $\tau$  has order 2. Moreover,

$$(\heartsuit) \quad \tau\sigma\tau = \sigma^{-1}$$

Condition  $(\heartsuit)$  shows that the subgroup  $\langle \tau \rangle$  normalizes the subgroup  $\langle \sigma \rangle$ . Thus [Proposition 2.6.3](#) shows that

$$\langle \sigma, \tau \rangle = \langle \sigma \rangle \langle \tau \rangle.$$

We call  $D = \langle \sigma, \tau \rangle$  the **dihedral group** of order  $n$ . Note that  $(\heartsuit)$  shows that  $\langle \tau \rangle$  normalizes  $\langle \sigma \rangle$  so that  $D = \langle \tau \rangle \cdot \langle \sigma \rangle$ .

We claim:

- $|D| = 2n$ . In fact,  $D$  is usually written  $D_{2n}$ .

To prove the claim, we apply [Corollary 2.6.5](#); we just need to argue that

$$(\clubsuit) \quad \langle \sigma \rangle \cap \langle \tau \rangle = \{1\}.$$

Since  $\sigma$  has order  $n$  and  $\tau$  has order 2,  $(\clubsuit)$  is immediate if  $n$  is odd.

Now suppose that  $n = 2k$  is even. The unique subgroup of order 2 in  $\langle \sigma \rangle$  is generated by  $\sigma^k$ . To prove  $(\clubsuit)$  we must argue that  $\sigma^k \neq \tau$ .

Suppose the contrary. If  $\sigma^k = \tau$  then  $\sigma(n) = \tau(n) \in \mathbb{Z}/n\mathbb{Z}$ . Since  $\sigma^k(n) \equiv n + k \pmod{n}$  while  $\tau(n) = -n \equiv n \pmod{n}$ , we conclude that  $n + k \equiv n \pmod{n}$ ; thus  $k \equiv 0 \pmod{n}$  i.e.  $2k \mid k$ , which yields a contradiction as  $k \geq 1$ . This completes the proof  $(\clubsuit)$ .

## 2.7. Lagrange's Theorem

Let  $H$  be a subgroup of the group  $G$  and write  $G/H$  for the (left) quotient, as above. Recall that the  $H$ -cosets  $xH$  are the  $H$ -orbits for this action.

**Theorem 2.7.1:** There is a bijection  $\varphi : (G/H) \times H \rightarrow G$  for which  $\{\varphi(z, h) \mid h \in H\}$  is an  $H$ -orbit (i.e. a left  $H$ -coset) for each  $z \in G/H$ .

*Proof:* Indeed, using the axiom of choice we select for each  $z \in G/H$  an element  $g_z \in \pi^{-1}(z)$  where  $\pi : G \rightarrow G/H$  is the quotient map.

Now define  $\varphi : (G/H) \times H \rightarrow G$  by the rule  $\varphi(z, h) = g_z h$ .

To see that  $\varphi$  is onto, let  $g \in G$ . One then knows that  $g \sim g_z$  for some  $z \in G/H$ . Since  $\pi^{-1}(z) = g_z H$  it follows that  $g = g_z h$  for some  $h \in H$ , so  $g = \varphi(z, h)$ .

To see that  $\varphi$  is injective, suppose that  $\varphi(z, h) = \varphi(z', h')$ . Then  $g_z h = g_{z'} h'$  so that

$$(g_{z'})^{-1} g_z \in H \Rightarrow g_z \sim g_{z'} \Rightarrow z = z'.$$

Now  $g_z h = g_{z'} h' \Rightarrow h = h'$  which completes the proof that  $\varphi$  is injective. The remaining assertion follows from the definition of  $\varphi$ . ■

**Corollary 2.7.2:** Suppose that  $G$  is a finite group and that  $H$  is a subgroup of  $G$ . Then

$$|G| = |G/H| \cdot |H|.$$

*Proof:* Indeed, for finite sets  $X$  and  $Y$ , we have  $|X \times Y| = |X| |Y|$ . ■

### 3. Week 3 [2025-09-15]

#### 3.1. Normal subgroups

Subgroups of the form  $\ker \varphi$  have a property that ordinary subgroups might lack; in this section we describe this property.

**Proposition 3.1.1:** Let  $G$  be a group.

- a. For  $g \in G$ , the assignment  $x \mapsto gxg^{-1}$  determines a group isomorphism

$$\text{Inn}_x : G \rightarrow G$$

- b. The assignment  $x \mapsto \text{Inn}_x$  determines a group homomorphism  $G \rightarrow \text{Aut}(G)$  where  $\text{Aut}(G)$  is the group of *automorphisms* of  $G$ .

*Proof sketch:*

- First check that  $\text{Inn}_x$  is a group homomorphism.
- Then check that  $(\blacklozenge) \text{Inn}_x \circ \text{Inn}_y = \text{Inn}_{xy}$  for all  $x, y \in G$ .
- Next, check that  $\text{Inn}_1 = \text{id}$ . Using  $(\blacklozenge)$ , this shows that  $(\text{Inn}_x)^{-1} = \text{Inn}_{x^{-1}}$  so indeed  $\text{Inn}_x$  is an *automorphism* of  $G$ .
- Finally,  $(\blacklozenge)$  shows that  $\text{Inn}$  is a group homomorphism.

■

**Definition 3.1.2:** A subset  $N \subseteq G$  is a **normal subgroup** of  $G$  if  $N$  is a subgroup of  $G$  and if for any  $g \in G$  and for any  $x \in N$ , we have  $gxg^{-1} \in N$ .

Using earlier notation, a subgroup  $N$  is normal if  $\forall g \in G, \text{Inn}_g N \subseteq N$ .

*Remark 3.1.3:* If  $N$  is a normal subgroup then for every  $g \in G$  we have  $\text{Inn}_g N = N$ .

Indeed, our assumption means for every  $g$  that  $\text{Inn}_g N \subseteq N$ . Thus  $\text{Inn}_g^{-1} \circ \text{Inn}_g N \subseteq \text{Inn}_g^{-1} N$  so that  $N \subseteq \text{Inn}_g^{-1} N$ . Since this holds for every  $g$ , we find that  $\text{Inn}_g N \subseteq N \subseteq \text{Inn}_g N$  for every  $g$ ; this confirms the assertion.

**Proposition 3.1.4:** Let  $H$  be a subgroup of  $G$ .

- a. Suppose  $G = \langle S \rangle$  for some subset  $S \subseteq G$ . Then  $H$  is normal in  $G$  if and only if  $\text{Inn}_x H = H$  for each  $x \in S$ .
- b. If  $H = \langle T \rangle$  for some subset  $T \subseteq H$ , then  $H$  is normal in  $G$  if and only if  $\forall t \in T, \forall x \in G, \text{Inn}_x t \in H$ .

*Proof:*

- a.  $(\Rightarrow)$ : This follows from the definition of normal subgroup.

( $\Leftarrow$ ): Write  $N = \{g \in G \mid \text{Inn}_g H = H\}$  and **check** that  $N$  is a subgroup of  $G$ . It is clear that  $H \subseteq N$  and by construction  $H$  is a normal subgroup of  $N$ . Now our assumption shows that  $S \subseteq N$  so that  $G = \langle S \rangle \subseteq N \Rightarrow N = G$  and thus  $H$  is normal in  $G$ .

b. ( $\Rightarrow$ ): Again, this implication follows from the definition of normal subgroup.

( $\Leftarrow$ ): Fix  $x \in G$ ; we must argue that  $\text{Inn}_x H \subseteq H$ . We know that  $\text{Inn}_x$  is a group homomorphism; see [Proposition 3.1.1](#). It follows from [Proposition 2.2.7](#)

$$\text{Inn}_x(\langle T \rangle) \subseteq \langle \text{Inn}_x(T) \rangle$$

which indeed shows that  $\text{Inn}_x H \subseteq H$ . ■

### Proposition 3.1.5:

Let  $N = \ker \varphi$  where  $\varphi : G \rightarrow H$  is a group homomorphism. Then  $N$  is a normal subgroup of  $G$ .

*Proof:* We have already observed that  $N$  is a subgroup. Now let  $g \in G$  and  $x \in N$  so that  $\varphi(x) = 1$ . Now

$$\varphi(\text{Inn}_g(x)) = \varphi(gxg^{-1}) = \varphi(g)\varphi(x)\varphi(g^{-1}) = \varphi(g)\varphi(g)^{-1} = 1$$

so that  $\text{Inn}_g N \subseteq N$  as required. ■

### Example 3.1.6:

Consider the group  $\text{GL}_2(\mathbb{Q})$ . For  $x \in \mathbb{Q}$  write

$$\alpha(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Note for  $x, y \in \mathbb{Q}$  that  $\alpha(x+y) = \alpha(x)\alpha(y)$ ; thus  $\alpha : \mathbb{Q} \rightarrow \text{GL}_2(\mathbb{Q})$  is an injective group homomorphism whose image

$$U_{\mathbb{Q}} = \{\alpha(x) \mid x \in \mathbb{Q}\} = \left\{ \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \mid x \in \mathbb{Q} \right\}$$

is a subgroup of  $\text{GL}_2(\mathbb{Q})$ .

Observe that  $U_{\mathbb{Z}} = \{\alpha(x) \mid x \in \mathbb{Z}\}$  is a subgroup of  $U_{\mathbb{Q}}$ .

For  $t \in \mathbb{Q}^{\times} = \mathbb{Q} \setminus \{0\}$ , write

$$h(t) = \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}.$$

For  $t, s \in \mathbb{Q}^{\times}$ , we have  $h(ts) = h(t)h(s)$  so that  $h : \mathbb{Q}^{\times} \rightarrow \text{GL}_2(\mathbb{Q})$  is an injective group homomorphism whose image

$$H = \{h(t) \mid t \in \mathbb{Q}^{\times}\} = \left\{ \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \mid t \in \mathbb{Q}^{\times} \right\}$$

is a subgroup of  $\text{GL}_2(\mathbb{Q})$ .

We observe for  $t \in \mathbb{Q}^\times$  and  $x \in \mathbb{Q}$  that

$$h(t)\alpha(x)h(t)^{-1} = \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t^{-1} & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & tx \\ 0 & 1 \end{pmatrix} = \alpha(tx).$$

This shows that  $\forall h \in H, \text{Inn}_h U_{\mathbb{Q}} \subseteq U_{\mathbb{Q}}$  so that  $H$  normalizes  $U_{\mathbb{Q}}$ .

Let  $n \in \mathbb{Z}$  with  $n > 1$  and consider the subgroup  $C = C_n = \langle h(n) \rangle$  of  $H$ .

The generator  $h(n)$  satisfies  $\text{Inn}_{h(n)} U_{\mathbb{Z}} \subset U_{\mathbb{Z}}$  since for  $x \in \mathbb{Z}$

$$h(n)\alpha(x)h(n)^{-1} = \alpha(nx) \in U_{\mathbb{Z}}.$$

Note however that  $\text{Inn}_{h(n)} U_{\mathbb{Z}}$  is a proper subset of  $U_{\mathbb{Z}}$ ; indeed, identifying  $U_{\mathbb{Z}}$  with  $\mathbb{Z}$ , the image subgroup  $\text{Inn}_{h(n)} U_{\mathbb{Z}}$  identifies with  $n\mathbb{Z}$  and of course  $n\mathbb{Z}$  has index  $n$  in  $\mathbb{Z}$ .

The group  $U_{\mathbb{Z}}$  is not normalized by  $C = \langle h(n) \rangle$  e.g. since  $\text{Inn}_{h(n)}^{-1} U_{\mathbb{Z}} = \text{Inn}_{h(n^{-1})} U_{\mathbb{Z}} \not\subseteq U_{\mathbb{Z}}$ ; indeed

$$\text{Inn}_{h(n^{-1})} \alpha(1) = \alpha\left(\frac{1}{n}\right) = \begin{pmatrix} 1 & \frac{1}{n} \\ 0 & 1 \end{pmatrix} \notin U_{\mathbb{Z}}.$$

This example shows the following: there exists a group  $G$ , subgroups  $H, K$  of  $G$ , and a subset  $S \subseteq G$  for which

$H = \langle S \rangle$  such that  $\text{Inn}_x K \subseteq K$  for all  $x \in S$  even though  $H$  does not normalize  $K$ .

Of course, if we insist that  $\text{Inn}_x K = K$  for all  $x \in S$  then  $H$  will normalize  $K$ ; see [Proposition 3.1.4](#).

In the proof of [Proposition 3.1.4](#) we gave a definition of the normalizer of  $K$  in  $H$ , namely

$$N_H(K) = \{h \in H \mid \text{Inn}_h K = K\}.$$

This example shows why in that definition one needs to insist that  $\text{Inn}_h K = K$  for each  $h \in H$  rather than simple  $\text{Inn}_h K \subseteq K$ .

### 3.2. Quotient groups

**Theorem 3.2.1:** Let  $N$  be a subgroup of  $G$ , and write  $(\pi_{G/N}, G/N)$  for the quotient. If  $N$  is a normal subgroup, then  $G/N$  is a group for which

a. the multiplication  $\mu : G/N \times G/N \rightarrow G/N$  satisfies

$$\forall g, g' \in G, \pi(g)\pi(g') = \pi(gg')$$

b. the identity is given by  $1_{G/N} = \pi(1_G)$ ,

c. inversion satisfies  $\forall g \in G, \pi(g)^{-1} = \pi(g^{-1})$ .

Moreover, the quotient map  $\pi_{G/N} : G \rightarrow G/N$  is a group homomorphism.

*Proof:* We first confirm that there is a mapping  $\mu : G/N \times G/N \rightarrow G/N$  satisfying the condition in a.

We observe that  $G/N \times G/N$  may be viewed as the quotient of the product group  $G \times G$  by the subgroup  $N \times N$ ; i.e. as  $(G \times G)/(N \times N)$ .

Consider the function

$$\varphi : G \times G \rightarrow G/N$$

given by

$$\varphi(g, g') = \pi_{G/N}(gg').$$

We claim that  $\varphi$  is constant on the  $N \times N$  orbits in  $G \times G$ . Indeed, suppose that  $(g, g') = (g_1, g'_1)(h, h')$  for  $g, g', g_1, g'_1 \in G$  and  $h, h' \in N$ . Thus  $g = g_1h$  and  $g' = g'_1h'$ . Then

$$\varphi(g, g') = \pi_{G/N}(gg') = \pi_{G/N}(g_1h \cdot g'_1h') = \pi_{G/N}(g_1g'_1g_1^{-1}hg'_1h') = \pi_{G/N}(g_1g'_1) = \varphi(g_1, g'_1)$$

since  $N$  a normal subgroup

$$\Rightarrow g_1^{-1}hg'_1 \in N \Rightarrow g_1^{-1}hg'_1h' \in N.$$

Thus there is a mapping  $\mu : G/N \times G/N \rightarrow G/N$  which satisfies  $\mu \circ \pi_{G \times G/N \times N} = \varphi$  and  $\mu$  clearly satisfies a.

Next we confirm that there is an inversion mapping  $G/N \rightarrow G/N$  that satisfies b. For this, one just checks that the mapping  $G \rightarrow G/N$  given by  $g \mapsto \pi_{G/N}(g^{-1})$  is constant on  $N$ -orbits. Let  $g, g' \in G$  and  $h \in N$  and suppose that  $g = g'h$ . We must argue that

$$\pi_{G/N}(g^{-1}) = \pi_{G/N}(g'^{-1}).$$

We have

$$(g'h)^{-1} = h^{-1}g'^{-1} = g'^{-1}g'h^{-1}g'^{-1}$$

so indeed

$$\pi_{G/N}(g^{-1}) = \pi_{G/N}((g'h)^{-1}) = \pi_{G/N}(g'^{-1}g'h^{-1}g'^{-1}) = \pi_{G/N}(g'^{-1})$$

since  $g'h^{-1}g'^{-1} \in N$  by the normality of  $N$  in  $G$ .

It remains to confirm that the group axioms hold.

To confirm associativity in  $G/N$ , let  $z, z', z'' \in G/N$ . We must argue that  $(zz')z'' = z(z'z'')$ . Since  $\pi$  is surjective we can write  $z = \pi(g)$ ,  $z' = \pi(g')$  and  $z'' = \pi(g'')$  for  $g, g', g'' \in G$ . Now we see using a. twice that

$$(zz')z'' = (\pi(g)\pi(g'))\pi(g'') = \pi(gg')\pi(g'') = \pi((gg')g'').$$

A similar calculation shows that

$$z(z'z'') = \pi(g(g'g''))$$

and now the result follows by associativity in  $G$ .

Similar calculations confirm that the  $\pi_{G/N}(1)$  acts as an identity and that  $\pi_{G/N}(g^{-1})$  is the inverse of  $\pi_{G/N}(g)$ .

Finally, it follows from the definitions that  $\pi_{G/N}$  is a group homomorphism. ■

*Example 3.2.2:*

If  $G$  is an abelian group, then  $\text{Inn}_x$  is the trivial homomorphism for each  $x \in G$ , and in particular every subgroup of  $G$  is normal.

Let's consider an additive abelian group  $A$  and  $B$  any subgroup. Write  $\pi : A \rightarrow A/B$  for the quotient mapping.

For  $a \in A$ , we often view  $\pi(a)$  as the coset  $a + B = \{a + x \mid x \in B\}$ .

We see for  $a, a' \in A$  that  $\pi(a) = \pi(a') \Leftrightarrow a - a' \in B$ .

### 3.3. First isomorphism theorem

#### **Theorem 3.3.1:**

Let  $\varphi : G \rightarrow H$  be a group homomorphism, and let  $K = \ker \varphi$ . Assume that  $\varphi$  is surjective. Then there is a unique isomorphism of groups  $\bar{\varphi} : G/K \rightarrow H$  such that  $\varphi = \bar{\varphi} \circ \pi$  where  $\pi : G \rightarrow G/K$  is the quotient homomorphism.

*Proof:* We first observe that – provided it exists –  $\bar{\varphi}$  is unique. Indeed, for any  $z \in G/K$  we may write  $z = \pi(g)$  for  $g \in G$  and then our assumption guarantees that

$$(*) \quad \bar{\varphi}(z) = \bar{\varphi}(\pi(g)) = \varphi(g).$$

So it just remains to argue that  $(*)$  determines a group isomorphism.

We first check that  $(*)$  determines a group homomorphism. Indeed, for  $z, z' \in G/K$  with  $z = \pi(g)$  and  $z' = \pi(g')$  for  $g, g' \in G$ , we have

$$\bar{\varphi}(zz') = \bar{\varphi}(\pi(g)\pi(g')) = \bar{\varphi}(\pi(gg')) = \varphi(gg') = \varphi(g)\varphi(g') = \bar{\varphi}(\pi(g))\bar{\varphi}(\pi(g')) = \bar{\varphi}(z)\bar{\varphi}(z').$$

Now we observe that since  $\varphi$  is surjective, and since  $\pi : G \rightarrow G/K$  is surjective, then  $\bar{\varphi}$  is surjective.

Finally, we check that  $\varphi$  is injective. For this, it suffices to show that  $\ker \varphi = \{1\}$ ; see [Proposition 1.2.5](#).

So, let  $z \in \ker \varphi \subseteq G/K$  and write  $z = \pi(g)$  for  $g \in G$ . We know that

$$1_H = \bar{\varphi}(z) = \bar{\varphi}(\pi(g)) = \varphi(g)$$

and we conclude that  $\varphi(g) = 1 \Rightarrow g \in \ker \varphi$ . Since  $g \in \ker \varphi$ , we know that  $\pi(g) = \pi(1)$ , in other words,  $z = \pi(g)$  is the identity element of the quotient group  $G/K$ . This proves that  $\ker \bar{\varphi}$  is trivial so that  $\bar{\varphi}$  is injective. ■

### 3.4. Groups acting on Groups

Let  $G$  and  $H$  be groups and suppose that  $G$  acts on the set  $H$ .

**Definition 3.4.1:**

We say that  $G$  acts by automorphisms on  $H$  if for each  $g \in G$ , the mapping

$$h \mapsto g \cdot h : H \rightarrow H$$

is an automorphism of the group  $H$ .

*Remark 3.4.2:*

To give an action of  $G$  on  $H$  by automorphisms is the same as to give a group homomorphism  $G \rightarrow \text{Aut}(H)$ .

**Proposition 3.4.3:**

If  $G$  acts on  $H$  by automorphisms, the set of fixed points

$$H^G = \{x \in H \mid \forall g \in G, g \cdot x = x\}$$

is a subgroup of  $H$ .

*Proof:* Notice that  $1 \in H^G$  since each group automorphism  $\psi : H \rightarrow H$  satisfies  $\psi(1) = 1$ .

Let  $x, y \in H^G$ . We must argue that  $x^{-1}y \in H^G$ .

We first argue that  $x^{-1} \in H^G$ . For this, let  $g \in G$ . Since the action of  $g$  is an automorphism of  $H$  and since  $g \cdot x = x$ , we see that

$$1 = g \cdot 1 = g \cdot xx^{-1} = (g \cdot x)(g \cdot x^{-1}) = x(g \cdot x^{-1}).$$

This shows that  $g \cdot x^{-1} = x^{-1}$  so that  $x^{-1} \in H^G$ .

Now again let  $g \in G$ . We must argue that  $g \cdot x^{-1}y = x^{-1}y$ . Since  $g$  acts as an automorphism of  $H$  we see that

$$g \cdot x^{-1}y = (g \cdot x^{-1})(g \cdot y) = x^{-1}y$$

since  $x^{-1}, y \in H^G$ .

■

*Example 3.4.4:*

$G$  acts in itself by inner automorphisms. This action is determined by the group homomorphism  $\text{Inn} : G \rightarrow \text{Aut}(G)$ .

In this case, the subgroup  $G^G = G^{\text{Inn}(G)}$  of fixed points is precisely the center  $Z = Z(G)$ :

$$Z = \{x \in G \mid \text{Inn}_g x = x\} = \{x \in G \mid gx = xg \quad \forall g \in G\}$$

For  $x \in G$ , the stabilizer  $\text{Stab}_G(x)$  is known as the centralizer:

$$\text{Stab}_G(x) = C_G(x) = \{g \in G \mid \text{Inn}_g x = x\} = \{g \in G \mid gx = xg\}.$$

And the orbit of  $x$  is known as the conjugacy class of  $x$ :



$$\mathcal{O}_x = \text{Cl}(x) = \{\text{Inn}_g x \mid g \in G\} = \{gxg^{-1} \mid g \in G\}.$$

[Proposition 2.5.3](#) gives a bijection

$$\text{Cl}(x) \simeq G/C_G(x).$$

**Proposition 3.4.5:** The center of a group  $G$  is a normal subgroup of  $G$ .

### 3.5. $p$ -groups

**Definition 3.5.1:**

For a prime number  $p$ , a finite  $p$ -group is a finite group  $G$  whose order is a power of  $p$ .

Let  $G$  be a finite  $p$ -group and suppose that  $G$  acts on the finite set  $E$ , and write  $E^G$  for the set of elements of  $E$  fixed by the action of  $G$ ; thus  $E^G = \{x \in E \mid \forall g \in G, g \cdot x = x\}$ .

**Proposition 3.5.2:**

With notation as above, we have  $|E| \equiv |E^G| \pmod{p}$ .

*Proof:* Indeed, the complement  $E \setminus E^G$  is the disjoint union of non-trivial orbits of  $G$ , each of which has order divisible by  $p$ . ■

**Proposition 3.5.3:**

Suppose that  $G$  acts by automorphisms on a second  $p$ -group  $H$ . The fixed points  $H^G$  form a non-trivial subgroup.

*Proof:* First of all, the fixed points form a subgroup because the action of an element  $g \in G$  is a group automorphism of  $H$ . In more detail, since  $H^G$  is a non-empty subset of  $G$ , it is enough to argue that for every  $x, y \in H^G$ , we have  $x^{-1}y \in H^G$ .

We first argue that  $x^{-1} \in H^G$ . For  $g \in G$ , we have

$$1 = g \cdot 1 = g \cdot xx^{-1} = (g \cdot x)(g \cdot x^{-1}) = x(g \cdot x^{-1}).$$

Thus  $g \cdot x^{-1}$  is an inverse of  $x$  so indeed  $x^{-1} = g \cdot x^{-1}$ . We now show that  $x^{-1}y \in H^G$ . For this again let  $g \in G$  be arbitrary. We have

$$g \cdot x^{-1}y = (g \cdot x^{-1})(g \cdot y) = x^{-1}y$$

which shows that  $x^{-1}y \in H^G$ .

Now [Proposition 3.5.2](#) shows that  $p$  divides the order of the subgroup  $H^G$ , so  $H^G$  is indeed non-trivial. ■

**Theorem 3.5.4:** The center of a non-trivial  $p$ -group is non-trivial.

*Proof:* If  $G$  is a non-trivial  $p$ -group, consider the action of  $G$  on itself by conjugation. The subgroup of fixed points is precisely the center of  $G$ , and [Proposition 3.5.3](#) implies that this subgroup is non-trivial. ■

**Corollary 3.5.5:**

Let  $G$  be a finite  $p$ -group with  $|G| = p^n$ . There is a series of subgroups

$$\{1\} = G_n \subset G_{n-1} \subset \dots \subset G_0 = G$$

such that  $G_i$  is normal in  $G$  for each  $0 \leq i < n$  and such that  $G_i/G_{i+1}$  is cyclic of order  $p$  for  $0 \leq i < n - 1$ .

*Proof sketch:* We proceed by induction on  $|G|$ . If  $|G| = 1$  so that  $G$  is the trivial group, the assertion is immediate.

Now suppose given a non-trivial  $p$ -group  $G$  and suppose that the result holds for all  $p$ -groups of order  $< |G|$ .

Let  $Z$  be the center of  $G$ . Then  $Z$  is non-trivial by [Theorem 3.5.4](#). Thus  $G/Z$  is  $p$  group with order  $< |G|$ .

By induction there is a sequence of subgroups

$$\{1\} = H_m \subset H_{m-1} \subset \dots \subset H_0 = G/Z.$$

such that  $H_i$  is normal in  $G/Z$  and  $H_i/H_{i+1}$  is cyclic of order  $p$  for each  $i < m$ .

Let  $G_i = \pi^{-1}(H_i) \subset G$ , where  $\pi : G \rightarrow G/Z$  is the quotient homomorphism.

One must check the following:

- $G_i$  is a normal subgroup of  $G$ ,
- $G_i/G_{i+1} \simeq H_i/H_{i+1}$  is cyclic of order  $p$  for each  $i$ .

Since  $G_m = \ker \pi = Z$  we have the sequence in  $G$ :

$$\{1\} \subset Z = G_m \subset G_{m-1} \subset \dots \subset G_1 \subset G_0 = G.$$

Thus to complete the proof of the Theorem, we must demonstrate that  $Z$  has a suitable sequence of subgroup.

Thus it remains to prove the Theorem in case  $G$  is an *abelian*  $p$ -group. This proof is addressed in the homework. ■

## 4. Week 4 [2025-09-22]

### 4.1. Sylow subgroups

Let  $G$  be a finite group of order  $n = p^m q$  with  $p$  a prime and with  $\gcd(p, q) = 1$ .

**Theorem 4.1.1** (Sylow's Theorem):

There exists a subgroup of  $G$  having order  $p^m$ ; such a subgroup is known as a *Sylow subgroup*, or a *Sylow  $p$ -subgroup*. Moreover:

- Any two Sylow  $p$ -subgroups are conjugate by an element of  $G$ .
- Any  $p$ -subgroup of  $G$  is contained in a Sylow  $p$ -subgroup.
- If  $r$  denotes the number of  $p$ -Sylow subgroups, then  $r \equiv 1 \pmod{p}$  and  $r \mid q$ .
- 

For the proof, we consider the set  $E$  of all subsets of  $G$  having order  $p^m$ . The action of  $G$  on itself by translation induces an action of  $G$  on  $E$ : for  $X \in E$ , evidently  $g \cdot X \in E$  where  $g \cdot X = \{g \cdot x \mid x \in X\}$ .

One knows that  $|E| = \binom{|G|}{p^m} = \binom{p^m q}{p^m}$ .

**Proposition 4.1.2:**  $\binom{p^m q}{p^m} \equiv q \pmod{p}$ .

*Proof:* Let  $X$  and  $Y$  be indeterminants; we work in the polynomial ring  $(\mathbb{Z}/p\mathbb{Z})[X, Y]$ . Write  $n = p^m q$  and consider

$$(X + Y)^n = ((X + Y)^{p^m})^q = (X^{p^m} + Y^{p^m})^q = \sum_{i=0}^q \binom{q}{i} (X^{p^m})^i (Y^{p^m})^{q-i}.$$

On the other hand, we have

$$(X + Y)^n = \sum_{i=0}^n \binom{n}{i} X^i Y^{n-i}.$$

and the required result follows by comparing the coefficient of  $X^{p^m} Y^{(q-1)p^m}$  in the two expressions. ■

**Definition 4.1.3:** Let  $G$  be a group and let  $H \subseteq G$  be a subgroup. The **normalizer of  $H$  in  $G$**  is the subgroup

$$N_G(H) = \{g \in G \mid \text{Inn}_g H = H\};$$

it is the stabilizer of  $H$  in  $G$  for the conjugation action of  $G$  on the set of subgroups of  $G$ .

Notice that  $G/N_G(H)$  is in bijection with the set of all conjugates  $\{gHg^{-1} \mid g \in G\}$ .

For the proof of the Theorem, we are going to use the following:

**Proposition 4.1.4:** Let  $P$  be a Sylow  $p$ -subgroup of  $G$  and let  $Q$  be any  $p$ -subgroup of  $G$ . Then

$$N_Q(P) = Q \cap P.$$

*Proof:* By definition,  $N_Q(P) = Q \cap N_G(P)$ , so we must show that  $Q \cap N_G(P) = Q \cap P$ .

Let  $H = Q \cap N_G(P)$ . Since  $P \subseteq N_G(P)$ , it is clear that  $Q \cap P \subseteq H = Q \cap N_G(P)$ . It remains to establish the reverse inclusion. Since  $H \subseteq Q$  by definition, it only remains to prove that  $H \subseteq P$ .

For this, we first claim that  $PH$  is a  $p$ -subgroup of  $G$ . Assume for the moment that this claim has been established. Since  $PH$  contains  $P$  and since  $P$  is a  $p$ -subgroup of maximal possible order, we conclude that  $P = PH$  and hence that  $H \subseteq P$  as required.

Since  $H \subseteq N_G(P)$ , the product  $PH = \{xh \mid x \in P \text{ and } h \in H\}$  is a subgroup of  $G$ . Moreover, we know that

$$|PH| = \frac{|P||H|}{|P \cap H|};$$

see [Corollary 2.6.5](#). Since  $|P|$  and  $|H|$  are powers of  $p$ ,  $PH$  is a  $p$ -subgroup. ■

Finally, we now give:

*Proof of Sylow's Theorem:* [Proposition 4.1.2](#) shows that  $|E| \not\equiv 0 \pmod{p}$ . Thus there must be some  $X \in E$  for which the orbit  $G \cdot X$  satisfies  $|G \cdot X| \not\equiv 0 \pmod{p}$ . If  $H$  is the stabilizer in  $G$  of  $X$ , there is of course a bijection between  $G \cdot X$  and  $G/H$ . In particular,

$$|G/H| \not\equiv 0 \pmod{p}.$$

Since  $|G| = |H| \cdot |G/H|$ , conclude that  $p^m$  divides the order of  $H$ .

On the other hand, fix  $x \in X$ . We claim that  $H \subseteq X \cdot x^{-1}$ . Indeed, for  $h \in H$ , since  $h$  stabilizes  $X$  we have

$$hx = x' \text{ for some } x' \in X.$$

Then  $h = x'x^{-1} \in X \cdot x^{-1}$  as required.

Concluding, we find that  $|H| \leq |X \cdot x^{-1}| = |X| = p^m$  and thus  $|H| = p^m$ . In particular,  $H$  is a Sylow subgroup.

Now let  $H'$  be any  $p$ -subgroup of  $G$  and consider the action of  $H'$  on the quotient  $G/H$  determined by left-multiplication. Since  $|G/H| = q$  is not divisible by  $p$ , [Proposition 3.5.2](#) shows that  $(G/H)^{H'} \neq \emptyset$ . Suppose that the coset  $gH \in G/H$  is fixed by  $H'$ . We claim that

$$H' \subset gHg^{-1}.$$

Indeed, since  $gH$  is fixed by  $H'$ , we have

$$x \in H' \Rightarrow xgH = gH \Rightarrow g^{-1}xgH = H \Rightarrow g^{-1}xg \in H \Rightarrow x \in gHg^{-1}.$$

This confirms that  $H' \subseteq gHg^{-1}$ . Thus any  $p$ -subgroup of  $G$  is contained in a Sylow subgroup. This proves (b).

Applying the argument of the preceding paragraph to the case where  $H'$  is a Sylow subgroup we see that  $H' = gHg^{-1}$ ; this shows that any two Sylow subgroups are conjugate, proving (a).

To prove (c), let  $P$  be a Sylow  $p$ -subgroup. Note that  $P$  acts by conjugation on the set of all Sylow  $p$ -subgroups of  $G$ . We choose Sylow  $p$ -subgroups  $Q_1, Q_2, \dots, Q_s$  which form a system of representatives of the  $P$ -orbits for this action. We may and will take  $Q_1 = P$ .

For  $1 \leq i < s$ , we write  $\mathcal{O}_i = P \cdot Q_i = \{xQ_ix^{-1} \mid x \in P\}$  for the  $P$ -orbit of  $Q_i$ . Recall that  $\mathcal{O}_i$  is in bijection with the quotient  $P/N_P(Q_i)$  where  $N_P(Q_i)$  is the *normalizer* of  $Q_i$  in  $P$ .

For  $1 \leq i \leq s$  [Proposition 4.1.4](#) shows that  $N_P(Q_i) = Q_i \cap P$ .

In particular, it follows that  $N_P(Q_1) = P \cap P = P$  so that  $|\mathcal{O}_1| = 1$ . For all  $2 \leq i \leq s$  we have  $P \neq Q_i$  so that  $N_P(Q_i) = Q_i \cap P \subsetneq P$ . Thus  $|\mathcal{O}_i| = [P : Q_i \cap P] > 1$  so that

$$|\mathcal{O}_i| \equiv 0 \pmod{p}.$$

Finally, the number  $r$  of Sylow  $p$ -subgroups satisfies

$$r = \sum_{i=1}^s |\mathcal{O}_i| = 1 + \sum_{i=2}^s |\mathcal{O}_i| \equiv 1 \pmod{p}$$

which proves the first assertion of (c). The second assertion of (c) follows since

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

and since  $P \subseteq N_G(P)$ . ■

For a finite group  $G$  and a prime number  $p$ , write  $n_p = n_p(G)$  for the number of  $p$ -Sylow subgroups of  $G$  (this is the number  $r$  from [Theorem 4.1.1](#)).

#### Corollary 4.1.5:

Let  $P \in \text{Syl}_p(G)$ . Then  $P$  is normal if and only if  $n_p = 1$ .

*Proof:* Indeed,  $P$  is normal if and only if  $\text{Inn}_G P = P$ . Since all  $p$ -Sylow subgroups are conjugate, the result is immediate. ■

## 4.2. Applications of Sylow's Theorem

#### Definition 4.2.1:

Let  $G$  be a group. A subgroup  $H$  of  $G$  is **characteristic** if for every automorphism  $\varphi : G \rightarrow G$ , we have  $\varphi(H) = H$ .

A characteristic subgroup  $H$  is always normal, since  $H$  is invariant under all the inner automorphisms  $\text{Inn}_g$  for  $g \in G$ .

*Example 4.2.2:*

For a prime number  $p$ , let  $G = \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ ; we may identify the group  $G$  with the quotient of  $\mathbb{Z} \times \mathbb{Z}$  by the subgroup  $N = \langle (p, 0), (0, p) \rangle$ . Write  $e = (1, 0) + N$  and  $f = (0, 1) + N$  so that  $e, f \in G$  are elements of order  $p$  and

$$G = \langle e, f \rangle = \langle e \rangle + \langle f \rangle \text{ and } \langle e \rangle \cap \langle f \rangle = \{0\}.$$

Since  $G$  is commutative,  $H = \langle e \rangle$  is a normal subgroup of  $G$ . But  $H$  is not characteristic in  $G$ . Indeed, the automorphism  $(a, b) \mapsto (b, a)$  of  $\mathbb{Z} \times \mathbb{Z}$  induces an automorphism  $\varphi$  of  $G = (\mathbb{Z} \times \mathbb{Z})/N$  for which

$$\varphi(e) = f \text{ and } \varphi(f) = e.$$

Thus  $\varphi(H) \neq H$ .

**Definition 4.2.3:**

A group  $G$  is **simple** if for any normal subgroup  $N \subseteq G$ , either  $N = \{1\}$  or  $N = G$ .

Any group of prime order is simple. Below we are going to find the first example of a non-abelian simple group.

Throughout the remainder of this section,  $G$  denotes a finite group.

For a prime number  $p$  we write  $n_p = n_p(G)$  for the number of  $p$ -Sylow subgroups in  $G$ , and  $\text{Syl}_p = \text{Syl}_p(G)$  for the set of Sylow  $p$ -subgroups. Recall that

$$n_p \equiv 1 \pmod{p} \text{ and } p \mid |G/P| \text{ for } P \in \text{Syl}_p(G)$$

.

**Proposition 4.2.4:**

- If  $n_p = 1$  for some prime  $p$ , then  $P \in \text{Syl}_p(G)$  is characteristic – and in particular normal – in  $G$ .
- Suppose that  $H \subseteq G$  is a normal subgroup, and suppose that  $P \in \text{Syl}_p(H)$  is a normal  $p$ -Sylow subgroup of  $H$  for some prime  $p$ . Then  $P$  is a normal subgroup of  $G$ .

*Proof:*

- For any automorphism  $\varphi$  of  $G$ ,  $\varphi(P)$  is again a  $p$ -Sylow subgroup of  $G$ . Since  $n_p = 1$ , we then  $\varphi(P) = P$  so that  $P$  is indeed characteristic.
- For  $g \in G$ ,  $\text{Inn}_g$  determines an automorphism of  $H$ . Thus by (a),  $\text{Inn}_g P = P$  which shows that  $P$  is normal in  $G$ .

■

**Proposition 4.2.5:**

Suppose that  $|G| = pq$  for distinct prime numbers  $p < q$ .

- $n_q = 1$  so that  $G$  has a normal Sylow  $q$ -subgroup.
- If  $p$  does not divide  $q - 1$  then  $n_p \equiv 1 \pmod{p}$  so that  $G$  has a normal Sylow  $p$ -subgroup. In this case,  $G$  is a cyclic group.

*Proof:*

- By [Theorem 4.1.1](#) we have  $n_q \equiv 1 \pmod{q}$  and  $n_q \mid p$ . Since  $q > p$  it follows that  $n_q = 1$ .
- Again we have  $n_p \equiv 1 \pmod{p}$  and  $n_p \mid q$ . Since  $q$  is prime, the only possibilities are  $n_p = 1$  or  $n_p = q$ .

If  $n_p = q$  then  $q \equiv 1 \pmod{p}$  so that  $p$  divides  $q - 1$ .

If  $n_p = 1$  then  $G = PQ$  where  $P$  is a Sylow  $p$ -subgroup and  $Q$  is a Sylow  $q$ -subgroup. You will prove for homework that since  $P$  and  $Q$  are both normal in  $G$ ,  $G$  is isomorphic to the direct product  $P \times Q$ . Thus

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \simeq \mathbb{Z}/pq\mathbb{Z}$$

is indeed cyclic. ■

**Proposition 4.2.6:**

Suppose that the size of a  $p$ -Sylow subgroup of  $G$  is  $p$ . Then  $G$  contains exactly  $(p - 1) \cdot n_p$  elements of order  $p$ .

*Proof:* Indeed, suppose that  $P, Q \in \text{Syl}_p(G)$  with  $P \neq Q$ . Since  $|P| = |Q| = p$  is prime we know that  $P \cap Q = \{1\}$ . Thus

$$\left| \bigcup_{P \in \text{Syl}_p} P \setminus \{1\} \right| = \sum_{P \in \text{Syl}_p} (p - 1) = n_p(p - 1).$$

■

**Proposition 4.2.7:**

Suppose that  $|G| = 12$ . Then either  $n_3 = 1$  or  $G$  is isomorphic to the alternating group  $A_4$ , and in that case,  $n_2 = 1$ .

*Proof:*

We know that  $n_3 \equiv 1 \pmod{3}$  and  $n_3 \mid 4$ . Thus  $n_3$  is either 1 or 4.

To complete the proof we must show that if  $n_3 = 4$  then  $G$  is isomorphic to  $A_4$ .

The group  $G$  acts by conjugation on the set  $\Omega = \text{Syl}_3$  of 3-Sylow subgroups, and  $|\Omega| = 4$ . This action determines a group homomorphism

$$\varphi : G \rightarrow S(\Omega) = S_4.$$

The kernel  $K$  of  $\varphi$  consists of all  $g \in G$  which normalize each Sylow 3-subgroup. In particular for  $P \in \Omega$ ,  $K \subseteq N_G(P) = P$ . Since  $K$  is normal, since  $P$  is not normal, and since  $|P| = 3$ , we conclude that  $K = \{1\}$ .

Thus  $\varphi$  is an isomorphism from  $G$  to its image in  $S_4$ .

Since  $n_3 = 4$ , [Proposition 4.2.6](#) shows that  $G$  contains exactly 8 elements of order 3. On the other hand, there are exactly 8 three-cycles in  $S_4$  and all are contained in  $A_4$ .

Thus the intersection of the image of  $\varphi$  with  $A_4$  is a group containing at least 8 elements. Since both groups have order 12, they must coincide.

Finally, notice that  $K = \langle (12)(34), (13)(24) \rangle$  is a normal 2-sylow subgroup of  $A_4$ ; thus  $n_2(A_4) = 1$ .

■

#### Proposition 4.2.8:

Suppose that  $|G| = 30$ . Then  $n_5 = 1$  so that  $G$  has a normal Sylow 5-subgroup.

*Proof:* Let  $P \in \text{Syl}_5$  and  $Q \in \text{Syl}_3$ .

First suppose that neither  $P$  nor  $Q$  is normal in  $G$ . In that case,  $n_5 = 6$  and  $n_3 = 10$ . By counting elements of order 3 and of order 5, it now follows from [Proposition 4.2.6](#) that

$$|G| \geq 6 \cdot 4 + 10 \cdot 2 = 44 > 30.$$

This contradiction proves that at least one of  $P$  or  $Q$  is normal in  $G$ .

Now if either  $P$  or  $Q$  is normal in  $G$ , then  $PQ$  is a subgroup of order 15. It follows from [Proposition 4.2.5](#) that any group of order 15 is cyclic. Using [Proposition 4.2.4](#) we then conclude that both  $P$  and  $Q$  are normal in  $G$ , and the result follows.

■

#### Proposition 4.2.9:

Suppose that  $|G| = 60$ . If  $n_5 > 1$  then  $G$  is a simple group.

*Proof:* For any group of order 60,  $n_5$  is either 1 or 6; thus we suppose  $n_5 = 6$ .

Let  $P \in \text{Syl}_5(G)$ . If  $N_G(P)$  is the normalizer of  $P$ , then  $|G/N_G(P)| = 6$  so that  $|N_G(P)| = 10$ .

We proceed by contradiction; thus, we suppose that  $\{1\} \neq H \subsetneq G$  is a normal subgroup of  $G$ .

First suppose that  $5 \mid |H|$ . Then  $H$  contains a Sylow 5-subgroup of  $G$ ; since  $H$  is normal,  $H$  contains all six Sylow 5-subgroups of  $G$ . Counting elements of order 5 in  $H$ , it follows from [Proposition 4.2.6](#) that



$$|H| \geq 6 \cdot 4 = 24.$$

Since the only divisor  $d$  of 60 with  $d \geq 24$  is  $d = 30$ , we conclude that  $|H| = 30$ . Now [Proposition 4.2.8](#) shows that  $H$  has a normal 5-Sylow subgroup  $Q \in \text{Syl}_5(H)$ , and [Proposition 4.2.4](#) shows that  $Q$  is normal in  $G$ . But this contradicts the assumption  $n_5 > 1$ .

This shows that

$$(\clubsuit) \quad G \text{ has no normal subgroup } H \text{ for which } 5 \mid |H|.$$

Thus we may suppose that  $|H|$  is a divisor of  $60/5 = 12$ .

If  $|H| = 12$ , it follows from [Proposition 4.2.7](#) that  $G$  has a normal Sylow  $p$ -subgroup for either  $p = 2$  or  $p = 3$ . In view of [Proposition 4.2.4](#), it follows that  $G$  has a normal subgroup of order 4 or 3.

If  $|G| = 6$ , then  $G$  has a normal Sylow 3-subgroup by [Proposition 4.2.5](#).

Thus we may suppose that  $|H|$  is one of 2, 3, 4.

Write  $\overline{G} = G/H$  for the quotient group, so that  $\overline{G} = 30, 20$  or 15.

We claim in each case that  $\overline{G}$  has a normal subgroup  $Q$  of order 5, i.e. that  $n_5(\overline{G}) = 1$ .

If  $|\overline{G}| = 30$  this claim follows from [Proposition 4.2.8](#). If  $|\overline{G}| = 20$  note that  $n_5 \mid 4$  and  $n_5 \equiv 1 \pmod{5}$  shows that  $n_5 = 1$ . Finally, if  $|\overline{G}| = 15$ , then  $n_5$  divides 3 and  $n_5 \equiv 1 \pmod{5}$  again shows that  $n_5 = 1$ .

If  $\pi : G \rightarrow \overline{G} = G/H$  is the quotient mapping, let  $H_1 = \pi^{-1}(Q)$  be the inverse image of the normal subgroup  $Q$  of order 5. You will prove for homework that  $H_1$  is a normal subgroup of  $G$  containing  $H$ ; since  $H_1/H \simeq Q$  it follows that  $5 \mid |H_1|$ . This contradicts  $(\clubsuit)$  and completes the proof of the Proposition. ■

#### Corollary 4.2.10:

The alternating group  $A_5$  is a simple group of order 60.

*Proof:* Indeed, the subgroups  $\langle (1, 2, 3, 4, 5) \rangle$  and  $\langle (1, 3, 2, 4, 5) \rangle$  are two distinct 5-Sylow subgroups of  $A_5$  so that  $n_5(A_5) > 1$ . ■

### 4.3. Rings

Let  $R$  be a ring and recall that we only consider rings with identity.

#### Definition 4.3.1:

A left ideal  $I$  of  $R$  is an additive subgroup of  $R$  that is closed under multiplication on the left by elements of  $R$ .

More precisely, for each  $x \in I$  and each  $a \in R$ , we have  $ax \in I$ .

*Remark 4.3.2:*

- a. There is an obvious related notion of right ideal.
- b. If  $R$  is commutative, then  $I$  is a left ideal if and only if  $I$  is a right ideal, and in that case we simply call  $I$  an ideal.
- c. For non-commutative  $R$ , we reserve the term ideal for an additive subgroup  $I$  which is both a left ideal and a right ideal. Sometimes we say that such an  $I$  is a two-sided ideal, for emphasis.

Let us now suppose that  $R$  is commutative.

**Proposition 4.3.3:**

- a. Let  $I$  and  $J$  be ideals of  $R$ . The intersection  $I \cap J$  is an ideal.
- a. More generally, if  $I_x$  is an ideal of  $R$  for each  $x$  in the index set  $X$ , then  $\bigcap_{x \in X} I_x$  is an ideal of  $R$ .

*Proof:* Since  $b \Rightarrow a$ , we prove  $b$ . Note that  $0 \in I_x$  for each  $x$  so that the intersection is non-empty.

Let  $a, b \in \bigcap_{x \in X} I_x$ . We must show that  $a - b$  is in the intersection. But for  $x \in X$ ,  $a, b \in I_x \Rightarrow a - b \in I_x$  so that indeed  $a - b \in \bigcap_{x \in X} I_x$ . ■

**Proposition/Definition 4.3.4:**

Let  $S \subset R$  be a subset. The ideal generated by  $S$ , written  $\langle S \rangle$  or  $RS$  is defined to be

$$\bigcap_{S \subseteq I} I,$$

the intersection taken over ideals of  $R$  containing  $S$ .

*Proof:* This intersection is an ideal by [Proposition 4.3.3](#). ■

**Definition 4.3.5:**

For  $a \in R$ , the principal ideal generated by  $a$  is the ideal  $\langle \{a\} \rangle$  and is denoted  $Ra$  or  $\langle a \rangle$ .

## 4.4. Ring homomorphisms and kernels

**Definition 4.4.1:**

If  $R$  and  $S$  are commutative rings, a function  $f : R \rightarrow S$  is a ring homomorphism provided that  $f$  is a homomorphism of additive groups and that

- a.  $f(ab) = f(a)f(b)$  for every  $a, b \in R$ , and
- b.  $f(1_R) = 1_S$ .

**Proposition 4.4.2:**

If  $f : R \rightarrow S$  is a ring homomorphism, the kernel  $\ker f$  is a ideal of  $R$ .

**Proposition 4.4.3:**

If  $I$  is an ideal of  $R$ , write  $\pi : R \rightarrow R/I$  for the quotient homomorphisms (where  $R/I$  is the quotient additive group).

Then there is a unique ring structure on the quotient group  $R/I$  with the property that quotient mapping  $\pi : R \rightarrow R/I$  is a ring homomorphism.

## 5. Week 5 [2025-09-29]

Parts of the material to be discussed this week are covered in the text [Dummit-Foote, “Abstract Algebra”]:

- quotient rings [Dummit-Foote] §7.3, 7.4
- modules; products and direct sums [Dummit-Foote] §10.1, 10.2, 10.3
- [Dummit-Foote] doesn’t use the language of categories.

### 5.1. Quotient rings

In this section,  $R$  will denote a ring (with identity as always, but not necessarily commutative).

By a (two-sided) ideal of  $R$ , we mean an additive subgroup  $I$  of  $R$  that is closed under multiplication with  $R$  on the left and on the right.

More precisely:  $I$  is an ideal if  $\forall x \in I, \forall r \in R, rx \in I$  and  $xr \in I$ .

If  $I$  is an ideal, then  $R/I$  is an additive abelian group, and the quotient mapping  $\pi : R \rightarrow R/I$  can be viewed as the mapping  $\pi(r) = r + I$ .

#### Theorem 5.1.1:

Let  $I$  be a two-sided ideal of  $R$ . Then the quotient group  $R/I$  has the structure of a ring with identity where the multiplication satisfies

$$(a + I)(b + I) = ab + I \text{ for } a, b \in R.$$

In particular, the quotient mapping  $\pi : R \rightarrow R/I$  is a surjective ring homomorphism. If  $I \neq R$ , then  $1_{R/I} \neq 0_{R/I}$ .

#### Theorem 5.1.2 (First isomorphism theorem for rings):

Let  $\varphi : R \rightarrow S$  be a surjective ring homomorphism. Recall that  $\varphi$  induces an isomorphism of additive groups  $\bar{\varphi} : R/I \rightarrow S$  for which  $\bar{\varphi}(a + I) = \varphi(a)$  for  $a \in R$ . Then  $\bar{\varphi}$  is an isomorphism of rings.

*Example 5.1.3:*

Let  $R$  be a commutative ring. One checks that

$$S = \left\{ \begin{pmatrix} a & d \\ 0 & b \end{pmatrix} : a, b, c \in R \right\}$$

is a subring of  $\text{Mat}_3(R)$  which is not commutative if  $1_R \neq 0_R$ . The mapping

$$f : S \rightarrow R \times R \text{ given by } f\left(\begin{pmatrix} a & d \\ 0 & b \end{pmatrix}\right) = (a, b)$$

is a surjective ring homomorphism with kernel  $K = \left\{ \begin{pmatrix} 0 & \alpha \\ 0 & 0 \end{pmatrix} \mid \alpha \in R \right\}$ . According to the theorem,  $f$  induces an isomorphism

$$S/K \simeq R \times R.$$

(Note that  $K$  is a two-sided ideal since  $K = \ker f$ . On the other hand, it is easy to check directly that  $K$  is a two-sided ideal of  $S$ .)

## 5.2. Categories

**Definition 5.2.1:** A category  $\mathcal{C}$  consists of the following:

- a class  $\text{Ob}(\mathcal{C})$  of objects,
- a class  $\text{Mor}(\mathcal{C})$  of morphisms together with class functions

$$\text{dom} : \text{Mor}(\mathcal{C}) \rightarrow \text{Ob}(\mathcal{C}) \text{ and } \text{codom} : \text{Mor}(\mathcal{C}) \rightarrow \text{Ob}(\mathcal{C})$$

for domain and codomain.

Denote by  $\text{Mor}(X, Y) = \text{Mor}_{\mathcal{C}}(X, Y)$  the subclass of  $\text{Mor}(\mathcal{C})$  consisting of morphisms  $f$  in  $\text{Mor}(\mathcal{C})$  with  $\text{dom}(f) = X$  and  $\text{codom}(f) = Y$ .

- for every three objects  $X, Y, Z$  there is a binary operation

$$(f, g) \mapsto g \circ f : \text{Mor}(X, Y) \times \text{Mor}(Y, Z) \rightarrow \text{Mor}(X, Z)$$

This data is required to satisfy:

- associativity:* For  $f$  in  $\text{Mor}(X, Y)$ ,  $g$  in  $\text{Mor}(Y, Z)$  and  $h$  in  $\text{Mor}(Z, W)$  we have

$$(h \circ g) \circ f = h \circ (g \circ f).$$

- identity:* For every object  $Z$ , there is  $\text{id}_Z$  in  $\text{Mor}(Z, Z)$  such that every  $f$  in  $\text{Mor}(Z, X)$  satisfies  $f \circ \text{id}_Z = f$  and every  $g$  in  $\text{Mor}(X, Z)$  satisfies  $\text{id}_Z \circ g = g$ .

*Remark 5.2.2:*

We often use function notation to represent morphisms – thus  $f : A \rightarrow B$  denotes the morphism  $f$  in  $\text{Mor}(A, B)$ . Be careful, though – in general, morphisms need not be functions.

*Example 5.2.3:* Here are some examples of categories.

- The category **Set** of all sets, with morphisms given by functions.
- The category **Grp** of all groups, with morphisms given by group homomorphisms.
- The category **Ab** of all abelian groups, with morphisms given by group homomorphisms.
- The category **Top** of topological spaces, with morphisms given by continuous functions.
- The category **Ring** of rings with morphisms given by ring homomorphisms.

**Definition 5.2.4:**

Let  $\mathcal{C}$  be a category. An object  $I$  of  $\mathcal{C}$  is said to be **initial** if for each object  $X$  of  $\mathcal{C}$  there is a unique morphism in  $\text{Mor}(I, X)$ .

An object  $T$  of  $\mathcal{C}$  is said to be **terminal** if for each object  $X$  of  $\mathcal{C}$  there is a unique morphism in  $\text{Mor}(X, T)$ .

*Example 5.2.5:*

- The empty set is an initial object in **Set**. Every singleton set is a terminal object in **Set**.

- b. The trivial group  $\{1\}$  is both an initial and a terminal object in Grp
- c. The trivial group  $\{0\}$  is both an initial and a terminal object in Ab

**Definition 5.2.6:**

Let  $\mathcal{C}$  be a category and let  $X$  and  $Y$  in  $\text{Ob}(\mathcal{C})$ . Then  $X$  and  $Y$  are **isomorphic** provided that there are morphisms  $f \in \text{Mor}(X, Y)$  and  $g \in \text{Mor}(Y, X)$  such that  $f \circ g = \text{id}_Y$  and  $g \circ f = \text{id}_X$ .

One says that  $f$  and  $g$  are isomorphisms between  $X$  and  $Y$ .

**Proposition 5.2.7:** Let  $\mathcal{C}$  be a category.

- a. If  $I, I'$  are initial objects, there is a unique isomorphism  $I \rightarrow I'$ .
- b. If  $T, T'$  are terminal objects, there is a unique isomorphism  $T \rightarrow T'$ .

*Proof:* We prove a; the proof of b is essentially the same..

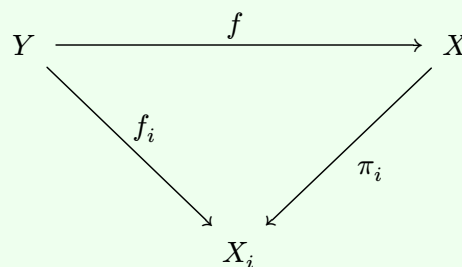
For a, since  $I$  is initial, there is a unique morphism  $f$  in  $\text{Mor}(I, I')$  and since  $I'$  is initial there is a unique morphism  $g$  in  $\text{Mor}(I', I)$ .

Now  $f \circ g$  is a morphism in  $\text{Mor}(I', I')$ . Since  $I'$  is initial,  $\text{id}_{I'}$  is the unique morphism in  $\text{Mor}(I', I')$  and we conclude that  $\text{id}_{I'} = f \circ g$ .

Similarly,  $g \circ f$  is a morphism in  $\text{Mor}(I, I)$ . Since  $I$  is initial,  $\text{id}_I$  is the unique morphism in  $\text{Mor}(I, I)$  and we conclude that  $\text{id}_I = g \circ f$ . Thus we have proved that  $f : I \rightarrow I'$  is the required unique isomorphism. ■

**Definition 5.2.8:**

Let  $\mathcal{C}$  be a category, let  $I$  be an index set, and let  $X_i$  be an object of  $\mathcal{C}$  for each  $i \in I$ . A **product** of the  $X_i$  is an object  $X$  of  $\mathcal{C}$  together with morphisms  $\pi_i : X \rightarrow X_i$  for  $i \in I$  such that given any object  $Y$  of  $\mathcal{C}$  together with morphisms  $f_i : Y \rightarrow X_i$  there is a unique morphism  $f : Y \rightarrow X$  such that  $f_i = \pi_i \circ f$  for each  $i \in I$ ; i.e. the diagram



commutes for each  $i \in I$ .

**Proposition 5.2.9:**

Let  $\mathcal{C}$  be a category, let  $I$  an index set, and let  $X_i$  be objects of  $\mathcal{C}$  for  $i \in I$ . If a product  $(X, \pi_i)$  of the  $X_i$  exists in  $\mathcal{C}$  where  $\pi_i : X \rightarrow X_i$  for  $i \in I$ , it is unique up to a unique isomorphism.

In other words, if  $(X', \pi_{i'})$  is a second product in  $\mathcal{C}$ , there is a unique isomorphism  $f : X \rightarrow X'$  in  $\mathcal{C}$  with the property that  $\pi_i = \pi_{i'} \circ f$ .

*Proof:* We introduce a new category  $\mathcal{D}$  depending on  $\mathcal{C}$ ,  $I$ , the family  $X_i$ . An object of  $\mathcal{D}$  is an object  $X$  of  $\mathcal{C}$  together with morphisms  $h_i : X \rightarrow X_i$  for each  $i \in I$ .

A morphism between objects  $(X, h_i)$  and  $(X', h'_i)$  of  $\mathcal{D}$  is a morphism  $\varphi$  in  $\text{Mor}_{\mathcal{C}(X, X')}$  such that  $h_i = h'_i \circ \varphi$ ; i.e. the diagram

$$\begin{array}{ccc} X & \xrightarrow{\varphi} & X' \\ & \searrow h_i & \swarrow h'_i \\ & X_i & \end{array}$$

commutes for each  $i \in I$ .

One checks that  $\mathcal{D}$  is a category. It is then clear that to give a terminal object in  $\mathcal{D}$  is the same as to give a product of the  $X_i$  in  $\mathcal{C}$ . Thus the uniqueness follows from [Proposition 5.2.7](#). ■

*Remark 5.2.10:*

If the objects  $X_i$  for  $i \in I$  have a product in the category  $\mathcal{C}$ , we write

$$\prod_{i \in I} X_i$$

for the product, keeping in mind that the morphisms

$$\pi_j : \prod_{i \in I} X_i \rightarrow X_j$$

are part of the data determining a product.



**Definition 5.2.11:**

Let  $\mathcal{C}$  be a category, let  $I$  be an index set, and let  $X_i$  be an object of  $\mathcal{C}$  for each  $i \in I$ . A **co-product** of the  $X_i$  is an object  $X$  of  $\mathcal{C}$  together with morphisms  $\gamma_i : X_i \rightarrow X$  for  $i \in I$  such that given any object  $Y$  of  $\mathcal{C}$  together with morphisms  $g_i : X_i \rightarrow Y$  there is a unique morphism  $g : X \rightarrow Y$  such that  $g_i = g \circ \gamma_i$  for each  $i \in I$ ; i.e. the diagram

$$\begin{array}{ccc}
 Y & \xleftarrow{g} & X \\
 & \nwarrow \gamma_i & \nearrow \gamma_i \\
 & X_i &
 \end{array}$$

commutes for each  $i \in I$ .

**Proposition 5.2.12:**

Let  $\mathcal{C}$  be a category, let  $I$  an index set, and let  $X_i$  be objects of  $\mathcal{C}$  for  $i \in I$ . If a co-product  $(X, \gamma_i)$  of the  $X_i$  exists in  $\mathcal{C}$  where  $\gamma_i : X_i \rightarrow X$ , it is unique up to a unique isomorphism.

In other words, if  $(X', \gamma'_i)$  is a second co-product in  $\mathcal{C}$ , there is a unique isomorphism  $f : X \rightarrow X'$  in  $\mathcal{C}$  with the property that  $\gamma_i = f \circ \gamma'_i$ .

*Remark 5.2.13:* If the objects  $X_i$  for  $i \in I$  have a co-product in the category  $\mathcal{C}$ , we write

$$\coprod_{i \in I} X_i$$

for the co-product, keeping in mind that the morphisms

$$\gamma_j : X_j \rightarrow \coprod_{i \in I} X_i$$

are part of the data determining a co-product.

### 5.3. Modules

Let  $R$  be a ring.

#### Definition 5.3.1:

A **left  $R$ -module**  $M$  is an additive abelian group  $M$  together with an operation of scalar multiplication  $R \times M \rightarrow M$  satisfying

- a. identity:  $1 \cdot m = m$  for every  $m \in M$ .
- b. associativity:  $(ab)m = a(bm)$  for every  $a, b \in R$  and  $m \in M$ .
- c. bilinearity:
  - $(a + b)m = am + bm$  for every  $a, b \in R$  and  $m \in M$
  - $a(m + n) = am + an$  for every  $a \in R$  and  $m, n \in M$ .

*Remark 5.3.2:*

There is a notion of right  $R$ -module  $M$ : for  $r \in R$  and  $m \in M$  the scalar multiplication is written  $m \cdot r$  and this scalar multiplication must satisfy analogous of the conditions required for a left module. When  $R$  is commutative, any left module can be viewed as a right module – for  $a \in R$  and  $m \in M$  just define the right module action via  $m \odot a = a \cdot m$  – and vice versa, so we may just speak of “ $R$ -modules” in this case.

*Example 5.3.3:*

- a. If  $R = F$  is a field, then the  $F$ -modules are precisely the  $F$ -vector spaces.
- b. Any abelian group is a  $\mathbb{Z}$ -module, and vice-versa.
- c. If  $R$  is a subring of some ring  $S$ , then  $S$  has the structure of an  $R$ -module.
- d. Any ideal  $I$  of  $R$  is an  $R$ -module (in particular,  $I$  is an  $R$ -submodule of the  $R$ -module  $R$ ).

**Proposition 5.3.4:** The data of an  $R$ -module  $M$  is equivalent to the data of an additive abelian group  $M$  together with a ring homomorphism  $R \rightarrow \text{End}_{\mathbb{Z}}(M)$ , where  $\text{End}_{\mathbb{Z}}(M)$  is the ring of additive endomorphisms of  $M$ .

#### Definition 5.3.5:

If  $M$  and  $N$  are left  $R$ -modules, a function  $\varphi : M \rightarrow N$  is a homomorphism of  $R$ -modules provided that

- $\varphi$  is a homomorphism of additive groups, and
- $\varphi(rm) = r\varphi(m)$  for every  $r \in R$  and every  $m \in M$ .

*Remark 5.3.6:*

- a. If  $R = F$  is a field, then a homomorphism of  $R$ -modules  $\varphi : M \rightarrow N$  is the same as a linear map of vector spaces.

- b. If  $A$  and  $B$  are abelian groups, a function  $\varphi : A \rightarrow B$  is a group homomorphism if and only if it is a homomorphism of  $\mathbb{Z}$ -modules.

**Definition 5.3.7:** Let  $M$  be an  $R$ -module. By an  $R$ -submodule of  $M$ , we mean an additive subgroup  $N \subseteq M$  such that  $\forall x \in N, \forall r \in R, rx \in N$ ; i.e. such that  $R \cdot N \subseteq N$ .

**Definition 5.3.8:** If  $R$  is a commutative ring, there is a category  $R\text{-mod}$  whose objects are the  $R$ -modules and whose morphisms are the  $R$ -module homomorphisms.

## 5.4. The direct sum of $R$ -modules.

Let  $I$  be an index set and suppose that  $M_i$  is an  $R$ -module for each  $i \in I$ .

**Proposition/Definition 5.4.1:** The **direct sum**  $\bigoplus_{i \in I} M_i$  of the  $R$ -modules  $M_i$  is the set of all finitely supported functions  $f : I \rightarrow \bigcup_{i \in I} M_i$  with the property that  $f(j) \in M_j$  for  $j \in I$ .

- $\bigoplus_{i \in I} M_i$  is an  $R$ -module with pointwise addition and scalar multiplication.
- Define  $\iota_j : M_j \rightarrow \bigoplus_{i \in I} M_i$  by setting  $\iota_j(m)$  to be the finitely supported function on  $I$  whose support is  $\{j\}$  and whose value at  $j$  is  $m$ .

For each  $j \in I$ , the map  $\iota_j$  is an  $R$ -module homomorphism.

*Proof:* The straightforward checks are left to the reader. ■

### Proposition 5.4.2:

Let  $I$  be an index set and let  $M_i$  be an  $R$ -module for each  $i \in I$ . Write  $M = \bigoplus_{i \in I} M_i$  and  $\iota_j : M_j \rightarrow M$   $j \in I$  as in [Proposition/Definition 5.4.1](#).

Then  $M = \bigoplus_{i \in I} M_i$  together with the  $\iota_j$  is a coproduct of the  $M_j$  in the category  $R\text{-mod}$ .

Recall that this means: Given any  $R$ -module  $N$  and  $R$ -module homomorphisms  $f_j : M_j \rightarrow N$ , there is a unique  $R$ -module homomorphism

$F : M \rightarrow N$  such that  $f_j = F \circ \iota_j$ .

$$\begin{array}{ccc}
 M = \bigoplus_{i \in I} M_i & \xrightarrow{F} & N \\
 & \nwarrow \iota_j \quad \nearrow f_j & \\
 & M_j &
 \end{array}$$

*Proof:* Let  $N$  and  $f_j : M_j \rightarrow N$  for each  $j \in I$  be given.

We first prove uniqueness of the mapping  $F$ . Consider an element

$$m \in M = \bigoplus_{i \in I} M_i.$$

Since  $m$  has finite support, we see that there is a finite subset  $J \subseteq I$  and for each  $j \in J$  an element  $m_j \in M_j$  for which

$$m = \sum_{j \in J} \iota_j(m_j).$$

Now we see that

$$(\heartsuit) \quad F(m) = \sum_{j \in I} (F \circ \iota_j)(m_j) = \sum_{j \in I} f_j(m_j).$$

This proves the uniqueness once we shows that  $(\heartsuit)$  defines an  $R$ -module homomorphism. But this follows from the definition of the  $R$ -module structure on  $\bigoplus_{i \in I} M_i$  and the fact that the  $f_j$  are  $R$ -module homomorphisms. ■

## 5.5. Free modules

Let  $R$  be a ring.

### Definition 5.5.1:

Let  $F$  be a left  $R$ -module, let  $B$  be a set and let  $\beta : B \rightarrow F$  be a function. Then  $F$  is a **free** left  $R$ -module on  $\beta$  provided that for any  $R$ -module  $X$  and any function  $j : B \rightarrow X$ , there is a unique  $R$ -module homomorphism  $\varphi : F \rightarrow X$  such that  $j = \varphi \circ \beta$ .

### Proposition 5.5.2:

Suppose that  $F$  is a free  $R$ -module on  $\beta : B \rightarrow F$ . Then the function  $\beta$  is injective (i.e. one-to-one).

*Proof:* Let  $b_1, b_2 \in B$  and suppose that  $b_1 \neq b_2$ . We must show that  $\beta(b_1) \neq \beta(b_2)$ . To this end, let  $f : B \rightarrow R$  be the function defined by

$$f(b) = \begin{cases} 1 & \text{if } b = b_1 \\ 0 & \text{otherwise} \end{cases}.$$

Since  $F$  is a free  $R$ -module on  $\beta$ , there is an  $R$ -module homomorphism  $\varphi : F \rightarrow R$  such that  $\varphi \circ \beta = f$ .

Then  $\varphi(\beta(b_1)) = f(b_1) = 1$  while  $\varphi(\beta(b_2)) = f(b_2) = 0$ . Since  $\varphi(\beta(b_1)) \neq \varphi(\beta(b_2))$  we must have  $\beta(b_1) \neq \beta(b_2)$  as required. ■

We are going to argue that the free  $R$ -modules are precisely those  $R$ -modules which have a basis.

Recall that for a set  $I$  and an additive abelian group  $A$ , the support of a function

$$f : I \rightarrow A \text{ is given by } \text{Supp}(f) = \{i \in I \mid f(i) \neq 0_A\}.$$

Then  $f$  has **finite support** provided that  $\text{Supp}(f)$  is a finite set.

Here is the definition:

**Definition 5.5.3:**

If  $M$  is an  $R$ -module, a function  $\beta : B \rightarrow M$  for some set  $B$  is an  $R$ -**basis**  $M$  provided that the following hold:

- $\beta$  is **linearly independent** i.e. if  $a : B \rightarrow R$  is a finitely supported function and if

$$\sum_{b \in B} a(b)\beta(b) = 0 \text{ then } a = 0.$$

- $\beta$  **spans**  $M$ ; i.e. for  $x \in M$  there is a finitely supported function  $a : B \rightarrow R$  such that

$$x = \sum_{b \in B} a(b)\beta(b).$$

(Observe that the sum is defined because  $a$  is finitely supported).

Notice that if  $\beta : B \rightarrow M$  is an  $R$ -basis then every element  $x \in M$  can be written in the form

$$x = \sum_{b \in B} a(b)\beta(b)$$

for a unique finitely supported function  $a : B \rightarrow R$ .

*Remark 5.5.4:* It might seem clumsy that we insist on defining bases etc. as functions from a set to an  $R$ -module. However, the reader will easily be able to pass back and forth between this language and other formulations.

For example, one says that an  $R$ -module  $F$  is free on a subset  $B \subset F$  if it is free on the inclusion mapping  $\iota : B \rightarrow F$ . Similarly, one says that a subset  $B \subset F$  is a basis if the inclusion mapping  $\iota : B \rightarrow F$

*Example 5.5.5:*

In general, an  $R$ -module  $M$  need not have a basis. For example, for  $n \in \mathbb{N}, n > 1$ , the  $\mathbb{Z}$ -module  $M = \mathbb{Z}/n\mathbb{Z}$  has no  $\mathbb{Z}$ -basis, since for any  $x \in M, nx = 0$  but  $0 \neq n \in \mathbb{Z}$ . This shows that there can be no  $\mathbb{Z}$ -linearly independent function from a non-empty set to  $M$ .

## 6. Week 6 [2025-10-06]

We continue the discussion of free modules.

### Proposition 6.1:

Let  $B$  be a set and let  $F = F(B, R)$  be the  $R$ -module consisting of all finitely supported maps  $a : B \rightarrow R$ . Consider the function  $\beta_0 : B \rightarrow F(B, R)$  where  $\beta_0(b)$  is the function

$$\beta_0(b)(b') = \begin{cases} 1 & \text{if } b = b' \\ 0 & \text{otherwise.} \end{cases}$$

- a. Then  $F$  is a free  $R$ -module on  $\beta_0$ .
- b.  $\beta_0$  is an  $R$ -basis for  $F$ .

*Proof:* For any finitely supported function  $a : B \rightarrow R$  we note that

$$(\heartsuit) \quad a = \sum_{b \in B} a(b) \beta_0(b).$$

- a. To see that  $F$  is a free module on the indicated data, let  $N$  be an arbitrary  $R$ -module and let  $\varphi : B \rightarrow N$  be any function. We must show that there is a unique  $R$ -module mapping  $\Phi : F \rightarrow N$  such that

$$(*) \quad \Phi \circ \beta_0 = \varphi.$$

We first treat uniqueness. Thus we suppose that there is a linear mapping  $\Phi : F \rightarrow N$  for which  $\Phi \circ \beta_0 = \varphi$ .

Using  $(\heartsuit)$ , the  $R$ -linearity of  $\Phi$ , and the requirement  $(*)$ , we see that

$$(\clubsuit) \quad \Phi(a) = \sum_{b \in B} a(b) \Phi(\beta_0(b)) = \sum_{b \in B} a(b) \varphi(b).$$

To complete the proof, it only remains to observe that the rule specified by  $(\clubsuit)$  indeed determines an  $R$ -module homomorphism; this follows from the definition of the  $R$ -module structure on  $F = F(B, R)$ .

- b. We first prove that  $\beta_0$  is  $R$ -linearly independent. We suppose that  $a : B \rightarrow R$  is a finitely supported function such that

$$\sum_{b \in B} a(b) \beta_0(b) = 0.$$

Then  $(\heartsuit)$  shows that  $a = 0$ ; this proves the linear independence.

Finally, we show that  $\beta_0$  spans  $F(B, R)$ . Let  $a \in F(B, R)$ . Then  $(\heartsuit)$  again shows that  $a$  has the required form.

■

**Proposition 6.2:** Let  $B$  be any set, consider for  $b \in B$  the  $R$ -module  $M_b = R$ , and let  $C$  together with  $\iota_b : R \rightarrow C$  be the coproduct (direct sum) of the modules  $M_b = R$  for  $b \in B$ .

With notation  $\beta_0$  as in [Proposition 6.1](#), there is an  $R$ -module isomorphism

$$\Psi : C \rightarrow F(B, R)$$

such that for each  $b \in B$ ,

$$(\Psi \circ \iota_b)(1) = \beta_0(b).$$

*Proof:* For each  $b \in B$ , there is an  $R$ -module homomorphism

$$f_b : R \rightarrow F(B, R) \text{ defined by } f_b(t) = t\beta_0(b) \text{ for } t \in R.$$

Using the defining property of the direct sum (coproduct), there is a unique  $R$ -module homomorphism  $\Psi : C \rightarrow F(B, R)$  such that  $(\Psi \circ \iota_b)(t) = f_b(t)$  for each  $b \in B$ . By the definition of the maps  $f_b$  we see that indeed  $\Psi \circ \iota_b(1) = \beta_0(b)$ .

To see that  $\Psi$  is an isomorphism, we construct the inverse homomorphism. For this, consider the function  $j : B \rightarrow C$  defined by  $j(b) = \iota_b(1)$ . According to [Proposition 6.1](#),  $F(B, R)$  is a free  $R$ -module on  $\beta_0$ . Using the defining property of a free  $R$ -module, we find a unique  $R$ -module homomorphism

$$\Phi : F(B, R) \rightarrow C \text{ such that } \Phi \circ \beta_0 = j.$$

We must now check that  $\Phi \circ \Psi = \text{id}_C$  and  $\Psi \circ \Phi = \text{id}_{F(B, R)}$ .

Now,

$$(\Phi \circ \Psi \circ \iota_b)(1) = \Phi(\beta_0(b)) = j(b) = \iota_b(1)$$

so that for  $\Phi \circ \Psi \circ \iota_b = \iota_b$ . But according to the defining property of the direct sum (coproduct), the unique  $R$ -module homomorphism  $f : C \rightarrow C$  for which  $f \circ \iota_b = \iota_b$  is  $f = \text{id}_C$ . Thus we conclude that  $\Phi \circ \Psi = \text{id}_C$ .

Similarly,

$$(\Psi \circ \Phi \circ \beta_0)(b) = \Psi(j(b)) = \Psi(\iota_{b(1)}) = \beta_0(b)$$

so that  $\Psi \circ \Phi \circ \beta_0 = \beta_0$ . But according to the defining property of free  $R$ -modules, if  $f : F(B, R) \rightarrow F(B, R)$  is an  $R$ -module homomorphism for which  $f \circ \beta_0 = \beta_0$  then  $f = \text{id}_{F(B, R)}$ . This shows that  $\Psi \circ \Phi = \text{id}_{F(B, R)}$  and completes the proof that  $\Phi$  and  $\Psi$  are inverse isomorphisms. ■

**Corollary 6.3:** If  $R$  is a commutative ring and if  $B$  is any set, there is an  $R$ -module  $F$  which is free on the set  $B$ .

*Proof:* Indeed, take  $F = F(B, R)$ . According to [Proposition 6.1](#),  $F$  together with the mapping  $\beta_0 : B \rightarrow F(B, R)$  determines a free  $R$ -module. ■

**Theorem 6.4:**

Let  $M$  be an  $R$ -module, let  $B$  be a set and let  $\beta : B \rightarrow F$  be a function. For  $b \in B$  consider the  $R$ -module homomorphism  $\iota_b : R \rightarrow M$  given by  $\iota_b(r) = r\beta(b)$ .

The following are equivalent:

- a.  $\beta$  is an  $R$ -basis for  $M$
- b.  $M$  is a free  $R$ -module on  $\beta : B \rightarrow M$ .
- c.  $M$  together with the  $\iota_b$  form a co-product of the  $R$ -modules  $M_b = R$ .

*Proof:* ( $a \Rightarrow b$ ): Let  $\beta$  be a basis; we show that  $M$  is free on  $\beta$ . Since  $\beta$  is a basis, we know that any  $x \in M$  may be written uniquely in the form  $x = \sum_{b \in B} a(b)\beta(b)$  for some  $a \in F(B, R)$  where  $F(B, R)$  is the  $R$ -module of all finitely supported functions  $a : B \rightarrow R$ .

Thus the assignment  $x \mapsto a$  defines an isomorphism of  $R$ -modules  $\Psi : M \rightarrow F(B, R)$ ; moreover, in the notation of [Proposition 6.1](#), we see that  $\Psi \circ \beta = \beta_0$ . Now the fact that  $M$  is free on  $\beta$  follows at once from [Proposition 6.1](#).

( $b \Rightarrow c$ ): Suppose that  $M$  is free on  $\beta$ . We are going to argue that  $M$  together with the  $\iota_b$  form a co-product of the modules  $M_b = R$  in the category  $R\text{-mod}$ . Thus we suppose that  $N$  is any  $R$ -module and that  $f_b : R \rightarrow N$  is an  $R$ -module map for each  $b \in B$ .

We form the function  $\varphi : B \rightarrow N$  defined by  $\varphi(b) = f_b(1)$ .

We claim: ( $\spadesuit$ ) a linear map  $\Phi : M \rightarrow N$  satisfies  $\Phi \circ \beta = \varphi$  if and only if it satisfies

$$\Phi \circ \iota_b = f_b \text{ for all } b \in B.$$

Indeed, from definitions we have

$$\Phi \circ \beta = \varphi \Leftrightarrow \forall b \in B, (\Phi \circ \beta)(b) = \varphi(b) \Leftrightarrow \forall b \in B, (\Phi \circ \iota_b)(1) = f_b(1).$$

Now for each  $b \in B$ , the  $R$ -module homomorphisms  $\Phi \circ \iota_b : R \rightarrow N$  and  $f_b : R \rightarrow N$  are equal if and only if they agree at  $1 \in R$ . This proves the claim.

Since  $M$  is free on  $\beta$ , there is a unique linear mapping  $\Phi : M \rightarrow N$  such that  $\Phi \circ \beta = \varphi$ . In view of ( $\spadesuit$ ) it follows that  $\Phi$  is the unique linear map satisfying  $\forall b \in B, \Phi \circ \iota_b = f_b$  as well. This proves that  $M$  is a coproduct of the  $M_b = R$  as required.

( $c \Rightarrow a$ ): Assume that  $(M, \iota_b)$  is a co-product of the modules  $M_b = R$  for  $b \in B$ . We must show that  $\beta$  is a basis.

According to [Proposition 6.2](#), there is an  $R$ -module homomorphism  $\Psi : M \rightarrow F(B, R)$  such that

$$(\Psi \circ \iota_b)(1) = \beta_0(b)$$

where  $F(B, R)$  is the  $R$ -module of finitely supported functions  $B \rightarrow R$  and where  $\beta$  is the mapping defined in [Proposition 6.1](#).

For  $b \in B$ , observe that  $\iota_{b(1)} = \beta(b) \Rightarrow \Psi(\beta(b)) = \beta_0(b)$ ; thus  $\Psi \circ \beta = \beta_0$ .



On the other hand, according to [Proposition 6.1](#),  $F(B, R)$  is a free  $R$ -module on  $\beta_0$ . Apply the defining property of a free module – see [Definition 5.5.1](#) – to the function  $\beta : B \rightarrow M$  to obtain an  $R$ -module homomorphism  $\Phi : F(B, R) \rightarrow M$  with the property that  $\Phi \circ \beta_0 = \beta$ .

We claim that the  $R$ -module homomorphisms  $\Phi$  and  $\Psi$  are inverse to one another. Once this claim is established, we see that  $\beta_0$  is a basis of  $F(B, R)$  implies that  $\beta = \Phi \circ \beta_0$  is a basis of  $M$ .

To prove the claim, first note that

$$\Phi \circ \Psi : M \rightarrow M$$

satisfies

$$\Phi \circ \Psi \circ \iota_b = \iota_b;$$

on the other hand, since  $M$  is a coproduct,  $\text{id}_M$  is the unique  $R$ -module map such that  $\text{id}_M \circ \iota_b = \iota_b$ . Thus  $\Phi \circ \Psi = \text{id}_M$ .

Finally note that

$$\Psi \circ \Phi : F(B, R) \rightarrow F(B, R)$$

satisfies

$$\Psi \circ \Phi \circ \beta_0 = \beta_0.$$

Since  $F(B, R)$  is a free  $R$ -module on  $\beta_0$ ,  $\text{id}_{F(B, R)}$  is the unique  $R$ -module map such that  $\text{id}_{F(B, R)} \circ \beta_0 = \beta_0$ .

Thus  $\Psi \circ \Phi = \text{id}_{F(B, R)}$  as required. This completes the proof.

■

## 6.1. Algebras

We now return to rings.

**Definition 6.1.1:** If  $S$  is any ring, the **center**  $Z(S)$  of  $S$  is the subring

$$Z(S) = \{r \in S \mid \forall x \in S, rx = xr\}$$

or  $R$ .

Observe that  $Z(S)$  is a commutative ring.

*Example 6.1.2:* If  $F$  is a field,  $n \in \mathbb{N}$ ,  $n \geq 1$  and  $S = \text{Mat}_n(F)$ , then  $Z(S) = F \cdot \text{id} \simeq F$ ; in words, the only matrices which commute with every matrix are scalar multiples of the identity.

**Definition 6.1.3:** If  $R$  is a commutative ring, an  **$R$ -algebra** is a ring  $A$  together with a ring homomorphism  $R \rightarrow Z(A)$ .

*Remark 6.1.4:*

- If  $A$  is an  $R$ -algebra, then  $A$  is an  $R$ -module. In particular, if  $F$  is a field, then any  $F$ -algebra is an  $F$ -vector space.
- If  $F$  is a field, for any  $F$ -algebra  $A$  the mapping  $F \rightarrow A$  is injective.

**Definition 6.1.5:** If  $A$  and  $B$  are  $R$ -algebras, a **homomorphism of  $R$ -algebras** is a ring homomorphism  $\varphi : A \rightarrow B$  such that  $\varphi$  is also a homomorphism of  $R$ -modules. In symbols:

$$\forall a \in A, \forall r \in R, \varphi(ra) = r\varphi(a).$$

*Remark 6.1.6:* Observe that for any  $R$ -algebra  $A$  there is a central subring that is the homomorphic image  $\overline{R}$  of  $R$ .

With this notation, for  $R$ -algebras  $A, B$  a ring homomorphism  $\varphi : A \rightarrow B$  is a homomorphism of  $R$ -algebras  $\Leftrightarrow$  the restriction of  $\varphi$  to the image  $\overline{R}$  in  $Z(A)$  satisfies  $\varphi(\bar{r}) = \bar{r}$  for  $r \in R$ .

*Example 6.1.7:*

- Any commutative ring has the structure of a  $\mathbb{Z}$ -algebra.

## 6.2. Integral Domains and prime ideals

Let  $R$  be a commutative ring.

**Definition 6.2.1:** An element  $a \in R$  is said to be a **zero-divisor** if  $a \neq 0$  and if  $\exists b \in R, b \neq 0$  such that  $ab = 0$ .

**Definition 6.2.2:**  $R$  is an **integral domain** provided that  $R$  has no zero-divisors.

*Example 6.2.3:*

- a. The ring  $\mathbb{Z}$  of integers is an integral domain. For  $n \in \mathbb{Z}$ , the quotient ring  $\mathbb{Z}/n\mathbb{Z}$  is an integral domain if and only if  $n$  is a prime number.
- b. Any field is an integral domain.

**Definition 6.2.4:** An ideal  $I \subset R$  is a **prime ideal** if  $I \neq R$  and if

$$a, b \in R, a \cdot b \in I \Rightarrow a \in I \text{ or } b \in I.$$

**Proposition 6.2.5:** Let  $I \subset R$  be an ideal. Then  $I$  is prime if and only if  $R/I$  is an integral domain.

*Proof:*  $(\Rightarrow)$  : Suppose that  $I$  is prime. We must show that  $R/I$  has no zero-divisors. To this end, let  $\alpha \in R/I$  be a non-zero element. Thus  $\alpha = a + I$  for  $a \in R, a \notin I$ .

Now suppose that  $\beta \in R/I$  and  $\alpha \cdot \beta = 0$ . To prove that  $\alpha$  is not a zero divisor, we must argue that  $\beta = 0$ . Write  $\beta = b + I$  for  $b \in R$ . To see that  $\beta = 0$ , we must argue that  $b \in I$ . But  $\alpha\beta = 0 \Rightarrow$  that  $ab + I$  is zero in  $R/I$  which shows that  $ab \in I$ . Since  $I$  is prime and since  $a \notin I$ , conclude that  $b \in I$  as required.

$(\Leftarrow)$ : Suppose that  $R/I$  is an integral domain. To prove that the ideal  $I$  is prime, let  $a, b \in R$  and suppose that  $ab \in I$ . We must show that  $a \in I$  or  $b \in I$ , so suppose that  $a \notin I$ . We will argue that  $b \in I$ .

Write  $\alpha = a + I, \beta = b + I \in R/I$ . Then  $ab \in I$  shows that  $\alpha\beta = ab + I$  is zero in  $R/I$ . Moreover,  $a \notin I$  shows that  $\alpha \neq 0$  in  $R/I$ . Since  $R/I$  is an integral domain, we conclude that  $\beta = 0$ . Thus  $b \in I$  as required. ■

**Definition 6.2.6:** An ideal  $I$  of  $R$  is **maximal** if  $I \neq R$  and if any ideal  $J$  with  $I \subseteq J \subseteq R$  satisfies  $J = I$  or  $J = R$ .

**Proposition 6.2.7:** If  $I \subset R$  is an ideal of  $R$  then  $I$  is a maximal ideal if and only if  $R/I$  is a field.

*Proof:*  $(\Rightarrow)$  : Suppose that  $I$  is a maximal ideal. To show that  $R/I$  is a field, let  $\alpha \in R/I, \alpha \neq 0$ . We must argue that  $\alpha$  is a unit (see [Definition 1.3.3](#)). For this, write  $\alpha = a + I$  for  $a \in R$ . Since  $\alpha \neq 0, a \notin I$ . Now consider the ideal  $J = \langle I, a \rangle \subseteq R$ . Since  $a \notin I, J \neq I$ . Since  $I$  is maximal, [Definition 6.2.6](#) shows that  $J = R$ . Thus  $1 \in \langle I, a \rangle$ . This means that  $1 = x + ab$  for some  $x \in I$  and some  $b \in R$ . Setting  $\beta = b + I$  we see that  $\alpha\beta = ab + I = 1 + I$  so  $\alpha$  is indeed a unit.

$(\Leftarrow)$  : Suppose that  $R/I$  is a field. We must argue that  $I$  is a maximal ideal. If  $\pi : R \rightarrow R/I$  is the quotient mapping, the assignment

$$J \mapsto \pi^{-1}(J)$$

determines a bijection between the ideals of  $R/I$  and the ideals of  $R$  containing  $I$ . Since the only ideals of the field  $R/I$  are 0 and  $R/I$ , the only ideals of  $R$  containing  $I$  are  $I$  and  $R$ . This proves that  $I$  is a maximal ideal of  $R$  as required. ■

**Corollary 6.2.8:** Any maximal ideal is prime.

*Proof:* Since any field is an integral domain, the Corollary follows from [Proposition 6.2.7](#) and [Proposition 6.2.5](#). ■

### 6.3. Monoid algebras

**Definition 6.3.1:** A **monoid**  $M$  is a set  $M$  equipped with a binary operation  $M \times M \rightarrow M$  such that

- the operation is associative:  $\forall x, y, z \in M, (xy)z = x(yz)$ .
- there is an identity element  $1 \in M$  such that  $\forall x \in M, 1x = x1 = x$ .

**Definition 6.3.2:** If  $M$  and  $N$  are monoids, a function  $f : M \rightarrow N$  is a **monoid homomorphism** if  $f(xy) = f(x)f(y)$  for every  $x, y \in M$ .

*Example 6.3.3:*

- Monoids are “not quite groups” – there is no requirement that elements have inverses.
- An example of a monoid that is not a group is the set of natural numbers  $\mathbb{N}$  under addition. More generally, for  $n \in \mathbb{N}$ , the set  $\mathbb{N}^n$  is a monoid under addition.
- If  $R$  is a ring, then  $(R, \times)$  – i.e. the set  $R$  with the operation of multiplication – forms a monoid.

**Proposition 6.3.4:** Let  $S$  be any ring. Let  $M$  be a monoid, let

$$a, b : M \rightarrow Z(S)$$

be finitely supported functions, and let

$$e : M \rightarrow (S, \times)$$

be a monoid homomorphism. Then

$$(\clubsuit) \quad \left( \sum_{m \in M} a_m e(m) \right) \left( \sum_{m \in M} b_m e(m) \right) = \sum_{m \in M} c_m e(m)$$

where  $c_m = \sum_{st=m} a_s b_t \in Z(S)$ , the sum being taken over all elements  $s, t \in M$  for which  $st = m$ .

*Proof:* Indeed, since the  $a$  and  $b$  take values in the center  $Z(S)$  and since  $e$  is a monoid homomorphism, we see that

$$\left( \sum_{m \in M} a_m e(m) \right) \left( \sum_{n \in M} b_n e(n) \right) = \sum_{m \in M} \sum_{n \in M} a_m b_n \cdot e(m) \cdot e(n) = \sum_{m \in M} \sum_{n \in M} a_m b_n \cdot e(m \cdot n).$$

Now the result follows from the observation that

$$\sum_{m \in M} \sum_{n \in M} a_m b_n \cdot e(m \cdot n) = \sum_{p \in M} \sum_{st=p} a_s b_t \cdot e(p).$$

■

Given a monoid  $M$ , we now *define* an  $R$ -algebra which as  $R$ -module has a basis indexed by  $M$  with multiplication defined by ( $\clubsuit$ ) from [Proposition 6.3.4](#).

**Proposition/Definition 6.3.5:** Let  $R$  be a commutative ring and let  $M$  be a monoid. The **monoid ring**  $R[M]$  over  $R$  is defined to be the free  $R$ -module with a basis  $\{e(m) : m \in M\}$  with the multiplication defined as follows: each  $\alpha, \beta \in R[M]$  may be written

$$\alpha = \sum_{m \in M} a_m e(m) \text{ and } \beta = \sum_{m \in M} b_m e(m)$$

where  $a_m, b_m \in R$  and all but finitely many of the coefficients  $a_m$  and  $b_m$  are 0.

Define

$$\alpha \cdot \beta = \sum_{m \in M} c_m e(m)$$

where

$$c_m = \sum_{s \cdot t = m} a_s \cdot b_t, \text{ the sum over } s, t \in M \text{ satisfying } s \cdot t = m.$$

With this multiplication, the  $R$ -module  $R[M]$  is a  $R$ -algebra and the assignment  $e : M \rightarrow R[M]$  is a monoid homomorphism. The algebra  $R[M]$  is commutative if and only if  $M$  is a commutative monoid. The homomorphism  $R \rightarrow R[M]$  is given by  $r \mapsto r \cdot e(1)$ .

*Proof sketch:* First notice that under the indicated rule for multiplication, for  $m, n \in M$ , we have

$$(\spadesuit)e(m)e(n) = e(mn).$$

In particular,  $e$  will be a monoid homomorphism  $M \rightarrow (R[M], \times)$  once we show that  $R[M]$  is a ring.

One first needs to argue that  $R[M]$  is a ring.

First of all, the multiplication  $(*)$  is well-defined. For this, notice that all but finitely many of the coefficients  $c_j \in R$  are 0, since that is true of  $a$  and  $b$ .

Now,  $1 = e(1)$  is the multiplicative identity of  $S$ . Indeed, every element of  $R[M]$  is an  $R$ -linear combination of elements of the form  $e(m)$ , so it suffices to prove that  $e(1)$  acts as the identity on  $e(m)$  for any  $m \in M$ . But

$$e(1)e(m) = e(1m) = e(m) \text{ and } e(m)e(1) = e(m1) = e(m).$$

Next, note that multiplication is associative. Indeed, let  $\alpha, \beta, \gamma \in R[M]$  with

$$\alpha = \sum_{m \in M} a_m e(m), \beta = \sum_{m \in M} b_m e(m), \gamma = \sum_{m \in M} c_m e(m).$$

For  $m \in M$ , we define the coefficient  $d_m = \sum_{stu=m} a_s b_t c_u \in R$ . We claim that

$$(\alpha\beta)\gamma = \sum_{m \in M} d_m e(m) = \alpha(\beta\gamma).$$

To prove the first equality, just note that using  $(\blacklozenge)$  we find

$$(\alpha\beta)\gamma = \left( \sum_m \sum_{st=m} a_s b_t e(m) \right) \left( \sum_n c_n e(n) \right) = \sum_{m,n} \sum_{st=m} a_s b_t c_n e(m) e(n) = \sum_m \sum_{stu=p} d_p e_p.$$

A similar calculation proves the second equality. This shows that the multiplication is associative.

Next we need to confirm that multiplication distributes over addition, i.e. that

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma \text{ and } (\beta + \gamma)\alpha = \beta\alpha + \gamma\alpha \text{ for } \alpha, \beta, \gamma \in R[M].$$

Since every element of  $R[M]$  is an  $R$ -linear combination of the elements  $e(m)$  for  $m \in M$ , it is enough to confirm these statements when  $\beta = r \cdot e(m)$  and  $\gamma = s \cdot e(n)$  for  $r, s \in R$  and  $m, n \in M$ .

Now, if  $\alpha = \sum_p a_p e(p)$  for a finitely supported function  $a : M \rightarrow R$ , then by definition we have

$$\begin{aligned} \alpha \cdot (r \cdot e(m) + s \cdot e(n)) &= \sum_p \left( \sum_{im=p} a_i r + \sum_{jn=p} a_j s \right) e(p) \\ &= \sum_p \left( \sum_{im=p} a_i r \right) + \sum_p \left( \sum_{jn=p} a_j s \right) e(p) \\ &= \alpha \cdot r \cdot e(m) + \alpha \cdot s \cdot e(n). \end{aligned}$$

A similar calculation shows that  $(r \cdot e(m) + s \cdot e(n))\alpha = r \cdot e(m)\alpha + s \cdot e(n)\alpha$ . This confirms the distributive law.

Finally, it is straightforward to confirm that  $r \mapsto re(1)$  defines a ring homomorphism  $R \rightarrow Z(R[M])$  so that  $R[M]$  is indeed an  $R$ -algebra.  $\blacksquare$

**Proposition 6.3.6:** Let  $S$  be a ring and let  $f : M \rightarrow (S, \times)$  be a monoid homomorphism.

a. Let  $\iota : R \rightarrow Z(S)$  be a ring homomorphism. Then there is a unique ring homomorphism

$$\varphi : R[M] \rightarrow S \text{ such that } \varphi|_R = \iota \text{ and } \varphi(e(m)) = f(m).$$

b. If  $S$  is an  $R$ -algebra, there is a unique homomorphism of  $R$ -algebras

$$\psi : R[M] \rightarrow R \text{ such that } \psi(e(m)) = f(m).$$

*Proof:* b. is an immediate consequence of a.

To prove a., note that using the ring homomorphism  $\iota$ , we may view  $S$  as an  $R$ -module. Now the function  $f : M \rightarrow S$  determines a unique  $R$ -module homomorphism  $\varphi : R[M] \rightarrow S$  such that  $\varphi|_R = \iota$  and  $\varphi \circ e = f$ .

It only remains to argue that  $\varphi$  is indeed a ring homomorphism; i.e. that  $\varphi$  preserves multiplication. In view of the definition of multiplication in  $R[M]$ , this follows from [Proposition 6.3.4](#). ■

## 6.4. The polynomial ring over $R$

For  $m \in \mathbb{N}$ , the polynomial ring over  $R$  in  $m$  variables is the monoid algebra constructed from the monoid  $\mathbb{N}^m$ . Let us write

$$\delta_j = (0, 0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{N}^m$$

where 1 in the  $j$ th position.

**Proposition 6.4.1:** Let  $S$  be a ring, let  $m \in \mathbb{N}$ , and let  $a_1, \dots, a_m \in Z(S)$ . The function

$$\varphi : \mathbb{N}^m \rightarrow (S, \times) \text{ given by } \varphi(\vec{i}) = a_1^{i_1} \dots a_m^{i_m}$$

is the unique monoid homomorphism for which  $\varphi(\delta_j) = T_j$  for  $1 \leq j \leq m$ , where

*Proof:* The uniqueness statement follows from the observation that

$$\vec{i} = \sum_{j=1}^m i_j \delta_j \text{ so that } \varphi(\vec{i}) = \sum_{j=1}^m \varphi(\delta_j)^{i_j}.$$

Since the  $a_1, \dots, a_m$  are central in  $S$ , for  $\vec{i}, \vec{j} \in \mathbb{N}^m$  we have

$$\varphi(\vec{i} + \vec{j}) = a_1^{i_1+j_1} \dots a_m^{i_m+j_m} = a_1^{i_1} \dots a_m^{i_m} a_1^{j_1} \dots a_m^{j_m} = \varphi(\vec{i}) \cdot \varphi(\vec{j}).$$

Thus  $\varphi$  is indeed a monoid homomorphism. ■

Let  $R$  be a commutative ring. Consider the additive monoid  $\mathbb{N}^m$  for  $m \in \mathbb{N}$

**Definition 6.4.2:** The polynomial ring over  $R$  in  $m$  variables is the monoid algebra of the additive monoid  $\mathbb{N}^m$ .

We write  $R[T_1, \dots, T_m]$  for the polynomial ring over  $R$ . For an element  $\vec{i} \in \mathbb{N}^m$  we write  $T_1^{i_1} \dots T_m^{i_m}$  for the element  $e(\vec{i})$  of the monoid algebra.

*Remark 6.4.3:*

- Compare with [Dummit-Foote, Ch. 7.2]
- Notice that the polynomial ring  $R[T_1, \dots, T_m]$  is a commutative  $R$ -algebra (since  $\mathbb{N}^m$  is a commutative monoid).
- Note for  $\vec{i}, \vec{j} \in \mathbb{N}^m$  that  $T^{\vec{i}} T^{\vec{j}} = T^{\vec{i}+\vec{j}}$ .
- The *variables*  $T_i$  of the polynomial ring correspond to the elements  $e(\delta_i)$  of the monoid algebra, where  $\delta_i = (0, 0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{N}^m$  has 1 in the  $i$ th position.
- Elements  $f, g \in S = R[T_1, \dots, T_n]$ , the elements may be written uniquely in the form

$$f = \sum_{\vec{i} \in \mathbb{N}^m} a_{\vec{i}} T^{\vec{i}} \text{ and } g = \sum_{\vec{i} \in \mathbb{N}^m} b_{\vec{i}} T^{\vec{i}}$$

for coefficients  $a_{\vec{i}}, b_{\vec{i}} \in R$  where all but finitely many of the  $a_{\vec{i}}$  are 0, and all but finitely many of the  $b_{\vec{i}}$  are 0.

The product of  $f$  and  $g$  is given by the formula

$$(*) \quad fg = \sum_{\vec{i}} \left( \sum_{\vec{s} + \vec{t} = \vec{i}} s_{\vec{s}} \cdot b_{\vec{t}} \right) T^{\vec{i}}.$$

**Proposition 6.4.4:** Let  $S$  be a ring and suppose that  $a_1, \dots, a_n \in Z(S)$ .

a. If  $\iota : R \rightarrow Z(S)$  is a ring homomorphism. Then there is a unique ring homomorphism

$$\varphi : R[T_1, \dots, T_n] \rightarrow S$$

with the property that  $\varphi|_R = \iota$  and  $\varphi(T_i) = a_i$  for  $1 \leq i \leq n$ .

b. If  $S$  is an  $R$ -algebra, there is a unique  $R$ -algebra homomorphism  $\psi : R[T_1, \dots, T_n] \rightarrow S$  such that  $\psi(T_i) = a_i$  for  $1 \leq i \leq n$ .

*Proof:* First observe that according to [Proposition 6.4.1](#) there is a unique monoid homomorphism  $\gamma : \mathbb{N}^m \rightarrow (S, \times)$  with the property that  $\gamma(\delta_j) = a_j$ .

Now the existence of the required ring and algebra homomorphisms follows from [Proposition 6.3.6](#). ■

## 6.5. Zorn's Lemma

Let  $(A, \leq)$  be a partially ordered set; thus the relation  $\leq$  is reflexive, antisymmetric and transitive.

**Definition 6.5.1:** Let  $B \subseteq A$ .

$B$  is a **chain in**  $A$  if  $\forall x, y \in A, x \leq y$  or  $y \leq x$ .

An element  $u \in A$  is an **upper bound** for  $B$  if  $\forall x \in B, x \leq u$ .

An element  $m \in A$  is a **maximal in**  $A$  if  $\forall x \in A, m \leq x \Rightarrow x = m$ .

*Example 6.5.2:*

a. Consider the partially ordered set  $A = \text{cal}[P](X)$ , the power set of some set  $X$ , where  $\leq \subseteq$  is set-containment.

An example of a chain is a collection of subsets  $X_i$  for  $i \in \mathbb{Z}$  with

$$\dots \subseteq X_{-1} \subseteq X_0 \subseteq X_1 \subseteq X_2 \subseteq \dots$$

Any subset  $B$  of  $A$  has an upper bound, namely

$$u = \bigcup_{b \in B} b$$

is an upper bound for  $B$ .



In this case,  $A$  has a (unique) maximal element, namely  $X$ .

b. Let  $A$  be the set of proper subsets of  $\mathbb{N}$ . Then

$$\{0\} \subseteq \{0, 1\} \subseteq \{0, 1, 2\} \subseteq \dots$$

describes a chain in  $A$  which has no upper bound in  $A$ .

On the other hand,  $A$  has (lots of) maximal elements, namely  $\mathbb{N} \setminus \{n\}$  for each  $n \in \mathbb{N}$ .

**Zorn's Lemma** is the following statement:

If  $A$  is a non-empty partially ordered set in which every chain has an upper bound, then  $A$  has a maximal element.

The **Axiom of choice** is the statement: if  $I$  is a non-empty index set and if  $A_i$  is a set for each  $i \in I$ , then the cartesian product  $\prod_{i \in I} A_i$  is non-empty. Thus, there is a choice function  $f : I \rightarrow \cup_{i \in I} A_i$  with  $f(j) \in A_j$  for each  $j \in I$ .

**Theorem 6.5.3:** Assuming usual axioms of set theory, Zorn's Lemma and the Axiom of Choice are logically equivalent.

*Remark 6.5.4:* In fact, the Axiom of Choice is independent of the axioms of set theory; namely the axioms of set theory together with the Axiom of choice are consistent, and the axioms of set theory together the negation of the axiom of choice are consistent.

In this course, we use the axioms of set theory together with the axiom of choice.

## 6.6. Existence of maximal ideals.

**Theorem 6.6.1:** Let  $A$  be a ring (with identity). Then every proper ideal of  $A$  is contained in a maximal ideal.

*Proof:* Let  $I \subset A$  be a proper ideal. Since  $1 \notin I$ , we know that  $1 \neq 0$  and so  $A \neq \{0\}$ .

Let  $\mathcal{S}$  be the set of proper ideals of  $R$  which contain  $I$ . Since  $I \in \mathcal{S}$ ,  $\mathcal{S} \neq \emptyset$ . Moreover,  $\mathcal{S}$  is partially ordered by inclusion.

If  $C$  is a chain in  $\mathcal{S}$ , define the ideal  $J = J_C$  to be

$$J_C = \bigcup_{I \in C} I.$$

We first show that  $J = J_C$  is an ideal of  $R$ . If  $a, b \in J$ , there are ideals  $A, B$  in  $C$  so that  $a \in A$  and  $b \in B$ . Since  $C$  is a chain, either  $A \subseteq B$  or  $B \subseteq A$ . In either case,  $a - b \in J$ . Since each  $A \in C$  is closed under left and right multiplication by  $R$ , so is  $J$ . This proves that  $J$  is an ideal of  $R$ .

Notice that  $I \subseteq J$  since  $I \subseteq A$  for any  $A \in C$ . Finally we claim that  $J$  is a proper ideal. For this, suppose by way of contradiction that  $J = R$ . Then  $1 \in J$  so that  $1 \in I$  for some  $I \in C$ . In that case,  $I = R$ , contrary to the fact that  $\mathcal{S}$  consists of proper ideals.

Note that  $J$  is an upper bound for the chain  $C$ . Thus we have confirmed that each chain in  $\mathcal{S}$  has an upper bound in  $\mathcal{S}$ .

By Zorn's Lemma, the partially ordered set  $\mathcal{S}$  has a maximal element, which is a maximal ideal of  $R$  containing  $I$ . ■

*Example 6.6.2:* Let  $k$  be a field, let  $k[T_1, \dots, T_n]$  be the polynomial ring over  $k$ , and let  $a_1, \dots, a_n \in k$ .

## 7. Week 7 [2025-10-14]

### 7.1. The rank of a free $R$ -module

Let  $R$  be a commutative ring.

**Definition 7.1.1:** If  $F$  is a free  $R$ -module on  $\beta : B \rightarrow F$ , the **rank** of the  $R$ -module  $F$  is the cardinality of  $B$ , denoted  $\text{rank}_R(F)$ .

**Theorem 7.1.2:** Let  $F$  be an  $R$ -module and let  $\beta_i : B_i \rightarrow F$  for  $i = 1, 2$  be functions for sets  $B_1$  and  $B_2$ . Suppose that  $F$  is a free  $R$ -module on  $\beta_1$ , and that  $F$  is a free  $R$ -module on  $\beta_2$ . Then the cardinality of  $B_1$  is equal to the cardinality of  $B_2$ .

In particular,  $\text{rank}_R(F)$  is a well-defined cardinal.

*Proof:* We may suppose that  $R \neq \{0\}$ . According to [Theorem 6.6.1](#), we may find a maximal ideal  $M$  of  $R$ . You showed in a homework problem that  $F$  is a free  $R$ -module on  $\beta_i$  if and only if  $F/MF$  is a free  $R/M$ -module on  $\beta_i$ .

Thus we have reduced the proof to the case in which  $R = K$  is a *field* in this case, the result is known from *linear algebra*. ■

*Remark 7.1.3:* Let  $K$  be a field, let  $V$  be a vector space over  $K$ , and let  $\beta_i : B_i \rightarrow V$  be bases for  $i = 1, 2$ .

To prove that the cardinality of  $B_1$  is equal to that of  $B_2$ , one can proceed as follows.

In case  $|B_1| = n$  and  $|B_2| = m$ , one uses the *Replacement Theorem* (see e.g. [Dummit-Foote, § 11.1 Theorem 3]) which states that for a suitable ordering  $B_1 = \{a_1, \dots, a_n\}$  and  $B_2 = \{b_1, \dots, b_m\}$  that  $\{b_1, \dots, b_k, a_{k+1}, \dots, a_n\}$  is a basis for each  $k \in \{1, 2, \dots, m\}$ . In particular,  $n \geq m$ . Using this result twice establishes  $n = m$ .

Finally, suppose that *a priori* only one of the basis is known to be finite. If  $|B_1| = n$ , one shows that any set of  $n + 1$  vectors is linearly independent; this forces  $B_2$  to be finite as well, so that  $|B_1| = |B_2|$  by the Replacement Theorem.

It remains true that bases have the same cardinality even when they are infinite, though I won't give a proof here.

For vector spaces, the *rank* is known as *dimension* and one usually writes

$$\dim_F V = \text{rank}_F V.$$

### 7.2. Principal Ideal Domains

**Definition 7.2.1:** A commutative ring  $A$  is a **principal ideal domain** (or **PID**) if  $A$  is an integral domain and if every ideal of  $A$  is a principal ideal.

**Definition 7.2.2:** Let  $A$  be a commutative ring and let  $a, b \in A$ .

- $a$  **divides**  $b$  – in symbols,  $a \mid b$  – provided that  $\exists x \in A, b = ax$ .
- a **greatest common divisor** of  $a$  and  $b$  is an element  $c \in A$  such that
  - $c \mid a$  and  $c \mid b$ , and
  - if  $d \mid a$  and  $d \mid b$  for  $d \in A$ , then  $d \mid c$ .

We write  $\gcd(a, b)$  for a greatest common divisor of  $a$  and  $b$ , provided it exists.

**Proposition 7.2.3:** Let  $A$  be a PID, let  $a, b \in A$ .

- An element  $d \in A$  is a greatest common divisor of  $a$  and  $b$  – i.e.  $d = \gcd(a, b)$  – if and only if  $d$  is a generator for the ideal  $\langle a, b \rangle$  – i.e.  $\langle d \rangle = \langle a, b \rangle$ .
- Any  $d = \gcd(a, b)$  has the form  $d = xa + yb$  for suitable  $x, y \in R$ .
- If  $d, d'$  are gcds of  $a$  and  $b$ , then  $d' = ud$  for a unit  $u \in R^\times$ .

*Proof:*

- If  $d$  is a generator for  $\langle a, b \rangle$  then  $a, b \in \langle d \rangle$  so that  $d \mid a$  and  $d \mid b$ . Now suppose that  $e \in A$  and that  $e \mid a$  and  $e \mid b$ . It follows that  $\langle a, b \rangle \subseteq \langle e \rangle$  so that  $\langle d \rangle \subseteq \langle e \rangle$ . We deduce that  $e \mid d$  and conclude that  $d = \gcd(a, b)$ .

On the other hand, suppose  $d = \gcd(a, b)$ . We prove that  $d$  is a generator for  $\langle a, b \rangle$ . By assumption we know  $d \mid a$  and  $d \mid b$ . This shows that  $\langle a, b \rangle \subseteq \langle d \rangle$ ; to prove that equality holds, it remains to show that  $\langle d \rangle \subseteq \langle a, b \rangle$ .

Let  $e$  be a generator for  $\langle a, b \rangle = \langle e \rangle$ . Since  $a, b \in \langle e \rangle$  deduce that  $e \mid a$  and  $e \mid b$ . Since  $d$  is a gcd, conclude that  $e \mid d$ . This proves that  $\langle d \rangle \subseteq \langle e \rangle = \langle a, b \rangle$  as required.

- The assertion follows since a gcd is a generator for the ideal  $\langle a, b \rangle$  by a.
- In view of a., the assumption means that

$$\langle d \rangle = \langle a, b \rangle = \langle d' \rangle.$$

Now  $d \in \langle d' \rangle \Rightarrow d = d'\alpha$  for  $\alpha \in R$  and  $d' \in \langle d \rangle \Rightarrow d' = d\beta$  for  $\beta \in R$ .

Now we see that  $d = d'\alpha = d\alpha\beta$ . Since  $A$  is an integral domain, we may cancel the factor of  $d$  and conclude that  $1 = \alpha\beta$ . Thus  $\alpha, \beta \in R^\times$  and the proof of c. is complete. ■

**Proposition 7.2.4:** Let  $A$  be a PID. If  $P$  is a non-zero prime ideal of  $A$ , then  $P$  is a maximal ideal. Thus  $A/P$  is a field.

*Proof:* Let  $P = \langle p \rangle$  be a non-zero prime ideal  $A$ , for  $0 \neq p \in A$ .

Let  $I = \langle m \rangle$  be any ideal containing  $P$ . We must show that  $I = P$  or  $I = A$ .

Since  $P = \langle p \rangle \subseteq I = \langle m \rangle$ , we know that  $m \mid p$ ; i.e.  $p = m\alpha$  for some  $\alpha \in A$ . Since  $m\alpha \in P$  and since  $P$  is a prime ideal, either  $m \in P$  or  $\alpha \in P$ .

If  $m \in P = \langle p \rangle$  then  $m = \beta p$  for some  $\beta \in R$ . Thus

$$p = m\alpha = p\alpha\beta.$$

Since  $A$  is an integral domain, conclude that  $1 = \alpha\beta$  so that  $\beta \in R^\times$ . This implies that  $p$  and  $m = \beta p$  generate the same ideal; i.e.  $P = I$ .

Finally, if  $\alpha \in P = \langle p \rangle$  then  $\alpha = p\beta$  for some  $\beta \in R$ . Now,  $p = m\alpha = mp\beta$  and since  $A$  is an integral domain, we find that  $1 = m\beta$ . This shows that  $m \in R^\times$  so that  $I = \langle m \rangle = R$ , as required. ■

We are going to show in the next section that examples of PIDs are obtained from so-called Euclidean domains; these will include the integers  $\mathbb{Z}$  and the ring  $K[T]$  of polynomials where  $K$  is any field.

### 7.3. Euclidean domains

Let  $A$  be a commutative ring.

**Definition 7.3.1:** A function  $N : A \rightarrow \mathbb{N} \cup \{-\infty\}$  is called a **norm** provided that  $N(x) = -\infty \Leftrightarrow x = 0$ .

**Definition 7.3.2:**  $A$  is called a **Euclidean domain** provided that  $A$  is an integral domain and there is a norm  $N$  on  $A$  such that for any pair of elements  $a, b \in A$  with  $b \neq 0$ ,  $\exists q, r \in R$  such that

$$a = qb + r \text{ with } N(r) < N(b).$$

**Proposition 7.3.3:** Suppose that  $A$  is an integral domain and suppose that  $a, b \in R$  with  $a = qb + r$ . Then

$$\langle a, b \rangle = \langle b, r \rangle$$

*Proof:* We must argue that  $a \in \langle b, r \rangle$  and that  $r \in \langle a, b \rangle$ . But the equation  $a = qb + r$  makes these assertions evident. ■

**Algorithm 7.3.4** (*The Euclidean algorithm*): Let  $A$  be a Euclidean domain with norm  $N$ , and let  $a, b \in A$  with  $b \neq 0$ .

Write  $a = qb + r_0$  with  $N(r_0) < N(b)$  so that  $\langle a, b \rangle = \langle b, r_0 \rangle$  by [Proposition 7.3.3](#).

Now suppose for  $m \in \mathbb{N}$  that  $r_0, r_1, \dots, r_m \in R$  have been chosen so that

$$(\clubsuit) \quad N(r_m) < N(r_{m-1}) < \dots < N(r_1) < N(r_0) < N(b)$$

and so that

$$(\spadesuit) \quad \langle a, b \rangle = \langle b, r_0 \rangle = \langle r_0, r_1 \rangle = \dots = \langle r_{m-1}, r_m \rangle.$$

Since  $A$  is a Euclidean domain, if  $r_m \neq 0$  we may write

$$r_{m-1} = qr_m + r_{m+1} \text{ for } q, r_{m+1} \in A \text{ with } N(r_{m+1}) < N(r_m);$$

then  $\langle a, b \rangle = \langle r_m, r_{m+1} \rangle$ .

Continuing in this fashion, so long as  $r_m \neq 0$  we may inductively construct a sequence of elements of  $R$  satisfying  $(\clubsuit)$  and  $(\spadesuit)$ . Since  $N(x) \in \mathbb{N}$  for  $x \neq 0$ , and since there can be no infinite decreasing sequence of natural numbers,  $(\clubsuit)$  shows that eventually we must have  $r_n = 0$  for some  $n \in \mathbb{N}$ .

Thus  $\langle a, b \rangle = \langle r_{n-1}, r_n \rangle = \langle r_{n-1} \rangle$  which shows that  $\langle a, b \rangle$  is a principal ideal.

**Proposition 7.3.5:** Let  $A$  be a Euclidean domain. Then  $A$  is a principal ideal domain.

*Proof:* Let  $0 \neq I \subseteq A$  be an ideal of  $A$ , and let  $z \in I \setminus \{0\}$  be an element for which  $N(z)$  is as small as possible.

We claim that  $I = \langle z \rangle$ . We must prove that  $I \subseteq \langle z \rangle$ . Let  $x \in I$  and write

$$x = qz + r \text{ for } q, r \in A \text{ with } N(r) < N(z).$$

Then  $r \in \langle x, z \rangle \subseteq I$ . If  $r \neq 0$ , the condition  $N(r) < N(z)$  contradicts the choice of  $z$  as an element for which  $N(z)$  is minimal among  $\{N(x) \mid x \in I \setminus \{0\}\}$ . Conclude that  $r = 0$  so that  $x = qz \in \langle z \rangle$  as required. ■

*Example 7.3.6:* The ring  $A = \mathbb{Z}$  of ordinary integers is a Euclidean domain. The norm is given by

$$N(x) = \begin{cases} |x| & \text{if } x \neq 0 \\ -\infty & \text{if } x = 0. \end{cases}$$

In particular,  $\mathbb{Z}$  is a principal ideal domain.

## 7.4. The polynomial ring over a field.

Let  $K$  be a field and let  $A = K[T]$  be the polynomial ring in a single variable over  $K$ . We are going to prove that  $A$  is a Euclidean domain.

Recall that an element  $f$  of  $A = K[T]$  has the form

$$f = \sum_{i=0}^N a_i T^i \text{ for } a_i \in K$$

or

$$f = \sum_{i \in \mathbb{N}} a_i T^i \text{ where } a : \mathbb{N} \rightarrow K \text{ has finite support.}$$

**Definition 7.4.1:** The **degree** of a non-zero element  $f = \sum_{i \in \mathbb{N}} a_i T^i$  – written  $\deg(f)$  – is the maximum  $i$  for which  $a_i \neq 0$ . Put another way,  $\deg(f)$  is the maximal element of the support of  $a$ .

We set  $\deg(0) = -\infty$ .

**Proposition 7.4.2:** If  $f, g \in K[T]$  then  $\deg(fg) = \deg(f) + \deg(g)$ .

*Proof:* Write  $f = \sum_{i \in \mathbb{N}} a_i T^i$  and  $g = \sum_{i \in \mathbb{N}} b_i T^i$  for finitely supported functions  $a, b : \mathbb{N} \rightarrow K$ . Then

$$fg = \sum_{i \in \mathbb{N}} \left( \sum_{s+t=i} a_s b_t \right) T^i$$

and the formula for the degree of  $fg$  follows at once. ■

**Proposition 7.4.3:** If  $f, g \in k[T]$  are non-zero polynomials, then  $\deg(f + g) \leq \max(\deg(f), \deg(g))$ . If  $\deg(f) \neq \deg(g)$  then equality holds.

*Proof:* Write  $f = \sum_{i \in \mathbb{N}} a_i T^i$  and  $g = \sum_{i \in \mathbb{N}} b_i T^i$  for finitely supported functions  $a, b : \mathbb{N} \rightarrow K$ . Then

$$f + g = \sum_{i \in \mathbb{N}} (a_i + b_i) T^i.$$

Thus  $\deg(f) \leq n$  and  $\deg(g) \leq m$  implies that  $\deg(f + g) \leq n + m$  which proves the first assertion.

If  $\deg(f) = n > m = \deg(g)$ , then the coefficient of  $T^n$  in  $f + g$  is just  $a_n$  so indeed

$$\deg(f + g) = n.$$

■

**Theorem 7.4.4:**  $A = K[T]$  is a Euclidean domain for the norm  $N = \deg : K[T] \rightarrow \mathbb{N} \cup \{-\infty\}$ . In particular,  $K[T]$  is a PID.

More precisely, if  $f, g \in K[T]$  with  $g \neq 0$ , there exist *unique* elements  $q, r \in K[T]$  with

$$f = qg + r \text{ and } N(r) < N(g).$$

*Proof:* To prove that  $A$  is a Euclidean domain and hence a PID, it suffices to prove the last assertion of the Theorem statement.

Let  $f, g \in K[T]$  with  $g \neq 0$ . If  $f = 0$ , take  $q = 0$  and  $r = 0$  and observe that the required properties hold.

We now suppose that  $f \neq 0$ . If  $\deg f < \deg g$ , we may take  $q = 0$  and  $r = f$ ; then  $N(r) < N(g)$  and  $f = 0 \cdot g + r$ .

Thus we may and will suppose that  $\deg g \leq \deg f$ . We now proceed by induction on  $n = \deg f$ .

If  $\deg f = 0$ , then  $f$  and  $g$  are both non-zero constant polynomials, say  $f = a_0$  and  $g = b_0$  with  $a_0, b_0 \in K^\times$ .

Then  $f = a_0 = (a_0/b_0)b_0 + 0$  so we may take  $q = a_0/b_0 \in K$  and  $r = 0$  to satisfy the requirements.

Thus the result holds in the base case where  $\deg f = 0$ .

Now suppose that  $m \in \mathbb{N}$  and that the result is known whenever  $f$  has degree  $\leq m$ .

Suppose that  $f$  has degree  $m + 1$  and that  $g$  has degree  $\leq m + 1$ . Write

$$f = \sum_{i=0}^{m+1} a_i T^i \text{ and } g = \sum_{i=0}^k b_i T^i$$

where  $a_{m+1} \neq 0$  and  $b_k \neq 0$  and  $k \leq m + 1$ .

Let

$$f' = f - (a_{m+1}/b_k)T^{m+1-k}g;$$

then  $\deg f' \leq m$ .

Applying the induction hypothesis, we may write

$$f' = qg + r \text{ where } \deg(r) < \deg(g) = k.$$

Now we find

$$f = f' + (a_{m+1}/b_k)T^{m+1-k}g = (q + (a_{m+1}/b_k)T^{m+1-k})g + r$$

so that  $f = q_1g + r$  with  $q_1 = (q + (a_{m+1}/b_k)T^{m+1-k})$  and  $\deg(r) < \deg(g)$ .

This proves that  $K[T]$  is a Euclidean domain. It remains to prove the uniqueness assertion. So suppose that  $f = qg + r = q_1g + r_1$  with  $\deg(r) < \deg(g)$  and  $\deg(r_1) < \deg(g)$ .



Then

$$(q - q_1)g = r_1 - r$$

so that  $g$  divides  $r_1 - r$ . Now, [Proposition 7.4.3](#) shows that  $\deg(r_1 - r) < \deg(g)$ .

On the other hand, [Proposition 7.4.2](#) shows that  $\deg(g) + \deg(q - q_1) = \deg(r_1 - r)$ . Since  $\deg(r_1 - r) < \deg(g)$ , this is only possible if  $q - q_1 = r - r_1 = 0$ .

■

**Proposition 7.4.5:** Let  $K$  be a field and let  $g = \sum_{i=0}^n a_i T^i \in K[T]$  with  $a_n \neq 0$ . Then  $A = K[T]/\langle g \rangle$  is a  $K$ -algebra with  $\dim_K A = n$ .

*Proof:* Write  $I = \langle g \rangle$ . We claim that  $B = \{T^i + I \mid 0 \leq i < n\}$  forms a  $K$ -basis for  $A = K[T]/I$ .

For  $\alpha \in A$ , write  $\alpha = f + I$  for some  $f \in K[T]$ . Now use [Theorem 7.4.4](#) to write  $f = qg + r$  for some  $r \in K[T]$  with  $\deg r < \deg g$ .

Since  $I = \langle g \rangle$ , we see that  $f - r \in I$  so that  $\alpha = f + I = r + I$ .

Since  $\deg(r) < n$ , clearly  $r$  is in the span of  $B$ . Thus we have proved that  $B$  spans  $A = K[T]/I$  as a  $K$ -module.

To see that  $B$  is linearly independent over  $F$ , suppose that  $(*) \sum_{i=0}^{n-1} a_i T^i = 0$  for  $a_i \in K$ .

Put  $h = \sum_{i=0}^{n-1} a_i T^i \in K[T]$ . Then  $(*)$  implies that  $h + I$  is the zero element of  $K[T]/I$ . Thus  $h \in I$  so that  $g \mid h$ .

Since  $\deg(g) = n$  and since  $\deg h < n$ , this is only possible if  $h = 0$  in  $K[T]$ . Since  $\{T^i \mid i \in \mathbb{N}\}$  is a  $K$ -basis for  $K[T]$  we now conclude that  $a_i = 0$  for each  $i$ . This confirms that  $B$  is linearly independent over  $K$  as required. ■

*Example 7.4.6:* Let  $K$  be a field and let  $\alpha \in K$ . Let  $\varphi : K[T] \rightarrow K$  be the unique homomorphism of  $K$ -algebras for which  $\varphi(T) = \alpha$ ; see [Proposition 6.4.4](#).

We claim that  $\ker \varphi = \langle T - \alpha \rangle$ . First of all,  $\varphi(T - \alpha) = 0$  so that  $\langle T - \alpha \rangle \subseteq \ker \varphi$ .

On the other hand, [Proposition 7.4.5](#) shows that  $K[T]/\langle T - \alpha \rangle$  is a  $K$ -algebra of dimension 1 as  $K$ -vector space. Thus the natural map  $K \rightarrow K[T]/\langle T - \alpha \rangle$  is an isomorphism.

This shows that  $\langle T - \alpha \rangle$  is a maximal ideal. Since  $\langle T - \alpha \rangle \subseteq \ker \varphi$  we conclude that equality holds:

$$\langle T - \alpha \rangle = \ker \varphi.$$

Finally, if  $\alpha, \beta \in K$  and  $\alpha \neq \beta$  then  $\langle T - \alpha \rangle \neq \langle T - \beta \rangle$ .

Indeed, let  $\varphi_\alpha : K[T] \rightarrow K$  be the evaluation homomorphism with  $\varphi_\alpha(T) = \alpha$ .

Notice that  $\varphi_{\alpha(T-\beta)} = \alpha - \beta \neq 0$  since  $\alpha \neq \beta$ . Thus  $T - \beta \notin \ker \varphi_\alpha = \langle T - \alpha \rangle$

## 7.5. The Chinese Remainder Theorem

Let  $R$  be a commutative ring.

**Definition 7.5.1:** Two ideals  $A$  and  $B$  of  $R$  are said to be **comaximal** if  $A + B = R$ .

*Remark 7.5.2:* If  $R_i$  for  $i \in I$  are rings, the product  $\prod_{i \in I} R_i$  in the category of rings is the product of the additive abelian groups together with a suitable multiplication. Identifying the product of the abelian groups as functions  $f : I \rightarrow \cup_{i \in I} R_i$  for which  $f(j) \in R_j$ , the product of functions  $f, g$  in the ring  $\prod_{i \in I} R_i$  is the *pointwise product* of  $f$  and  $g$ ; thus for  $j \in I$ ,  $(fg)(j) = f(j)g(j)$  where the product is taken in  $R_j$ .

*Remark 7.5.3:* For ideals  $A, B$  of  $R$  recall that the product  $AB$  is the ideal generated by  $\{ab \mid a \in A, b \in B\}$ . We always have  $AB \subseteq A \cap B$ .

**Proposition 7.5.4:** Let  $A, B$  be ideals of  $R$  which are comaximal and consider the mapping

$$\Phi : R \rightarrow R/A \times R/B \text{ given by } \Phi(r) = (r + A, r + B).$$

Then  $\Phi$  is a surjective ring homomorphism and  $\ker \Phi = A \cap B = AB$ .

*Proof:*  $\Phi$  is the map to  $R/A \times R/B$  determined by the mapping property of the product using the quotient maps  $\pi_A : R \rightarrow R/A$  and  $\pi_B : R \rightarrow R/B$ ; in particular,  $\Phi$  is a ring homomorphism.

Since  $A$  and  $B$  are co-maximal, we may write  $1 = a + b$  for some  $a \in A$  and  $b \in B$ .

Let us show that  $\Phi$  is surjective. Let  $(\alpha, \beta) \in R/A \times R/B$ . Thus  $\alpha = s + A$  and  $\beta = t + B$  for  $s, t \in R$ .

Let  $r = sa + tb \in R$ . Notice the following:

$$s = s1 = sa + sb \equiv sa \pmod{B}$$

and

$$t = t1 = ta + tb \equiv tb \pmod{A}.$$

This shows that  $\Phi(r) = (a + A, b + B) = (\alpha, \beta)$  so that  $\Phi$  is surjective.

Finally, we argue that  $\ker \Phi = AB$ . For  $a \in A$  and  $b \in B$ , notice that  $\Phi(ab) = 0$ ; since  $AB$  is generated by such products, we see that  $AB \subseteq \ker \Phi$ .

Finally, notice that by the definition of the product,  $\ker \Phi = \ker \pi_A \cap \ker \pi_B = A \cap B$ .

It remains to argue that  $A \cap B = AB$ .

As was pointed out in [Remark 7.5.3](#), we know that  $AB \subseteq A \cap B$ . To prove equality, suppose that  $x \in A \cap B$ .

Now note that  $x = x \cdot 1 = xa + xb$ . Since  $x \in B$ ,  $xa \in AB$ . Since  $x \in A$ ,  $xb \in AB$ . Thus  $x = xa + xb \in AB$  and the proof is complete. ■

**Theorem 7.5.5:** Let  $n \geq 1$  and let  $A_i$  be an ideal of  $R$  for each  $i \in I_n = \{0, 1, \dots, n-1\}$ .

a. The mapping

$$\Phi : R \rightarrow \prod_{i \in I_n} R/A_i \text{ given by } \Phi(r) = (j \mapsto r + A_j)$$

is a ring homomorphism whose kernel is  $\bigcap_{i \in I} A_i = A_0 \cap A_1 \cap \dots \cap A_{n-1}$ .

b. Assume that  $\forall i, j \in I_n, i \neq j \Rightarrow A_i, A_j$  are comaximal. Then  $\Phi$  is surjective and moreover

$$\bigcap_{i \in I} A_i \text{ is equal to the product } A_0 A_1 \dots A_{n-1}.$$

*Proof:* First of all, the mapping  $\Phi : R \rightarrow \prod_{i \in I} R/A_i$  is just precisely the mapping determined by the universal property of the product using the quotient maps  $R \rightarrow R/A_j$  for each  $j \in I$ . Thus  $\Phi$  is a ring homomorphism.

We give the proof by induction on  $n$ . When  $n = 1$  the map  $\Phi$  is just the quotient map  $R \rightarrow R/A_0$  and there is nothing to prove in either a. or b.

Now suppose that  $m > 0$  that  $A_0, A_1, \dots, A_{m+1}$  are ideals.

For the proof of (a), the induction hypothesis is that the homomorphism

$$\Phi_m : R \rightarrow \prod_{i \in I_m} R/A_i$$

has kernel  $\bigcap_{i \in I_m} A_i$ , and we must show that

$$\Phi_{m+1} : R \rightarrow \prod_{i \in I_{m+1}} R/A_i$$

has kernel  $\bigcap_{i \in I_{m+1}} A_i$ .

We may identify

$$\prod_{i \in I_{m+1}} R/A_i \text{ with } \left( \prod_{i \in I_m} R/A_i \right) \times R/A_m,$$

and then  $\Phi_{m+1}$  is the mapping

$$(\Phi_m, \pi) : R \rightarrow \left( \prod_{i \in I_m} R/A_i \right) \times R/A_m$$

determined by  $\Phi_m$  and the quotient map  $\pi : R \rightarrow R/A_m$ . Thus  $\ker \Phi_{m+1} = \ker \Phi_m \cap \ker \pi$ . So the induction hypothesis implies that

$$\ker \Phi_{m+1} = \ker \Phi_m \cap A_m = \left( \bigcap_{i \in I_m} A_i \right) \cap A_m = \bigcap_{i \in I_{m+1}} A_i.$$

Now the assertion of (a) follows by induction.

For the proof of (b), we suppose that  $\forall i, j \in I_{m+1}, i \neq j \Rightarrow A_i + A_j = R$ . The induction hypothesis is that  $\Phi_m$  is surjective and that  $\ker \Phi_m$  is the product  $A_0 A_1 \dots A_{m-1}$ . We must prove that  $\Phi_{m+1}$  is surjective and that  $\ker \Phi_{m+1}$  is the product  $A_0 A_1 \dots A_m$ .

Again identify

$$\prod_{i \in I_{m+1}} R/A_i \text{ with } \left( \prod_{i \in I_m} R/A_i \right) \times R/A_m$$

and  $\Phi_{m+1}$  with

$$(\Phi_m, \pi_{A_m}) : R \rightarrow \left( \prod_{i \in I_m} R/A_i \right) \times R/A_m.$$

Put  $B = \ker \Phi_m$ . According to the induction hypothesis, we have

$$B = A_0 A_1 \dots A_{m-1}.$$

For  $0 \leq i \leq m-1$ ,  $A_i$  and  $A_m$  are comaximal, so we may write  $1 = a_i + c_i$  for  $a_i \in A_i$  and  $c_i \in A_m$ .

Then

$$1 = 1^m = \prod_{i=0}^{m-1} (a_i + c_i) = a_0 a_1 \dots a_{m-1} + c \text{ for some } c \in A_m.$$

This shows that  $B$  and  $A_m$  are comaximal.

Now we may apply [Proposition 7.5.4](#) to see that  $\Phi_{m+1}$  is surjective and that

$$\ker \Phi_{m+1} = B \cap A_m = BA_m = (A_0 A_1 \dots A_{m-1}) A_m$$

as required.

Thus the Theorem follows by induction. ■

*Remark 7.5.6:* If the ideal  $A_i = Ra_i$  are all **principal**, then

$$A_0 A_1 \dots A_{m-1} = Ra_0 a_1 \dots a_{m-1}.$$

*Example 7.5.7:* Let  $K$  be a field. For  $m \in \mathbb{N}$  let  $\alpha_1, \dots, \alpha_m \in K$  be *distinct* elements. We claim that

$$(\heartsuit) \quad K[T]/\langle (T - \alpha_1)(T - \alpha_2) \dots (T - \alpha_m) \rangle \simeq \prod_{i=1}^m K.$$

Indeed for  $i \neq j$ ,  $\langle T - \alpha_i \rangle$  and  $\langle T - \alpha_j \rangle$  are distinct maximal ideals of  $K[T]$  by [Example 7.4.6](#). In particular these ideals are co-maximal. Now  $(\heartsuit)$  follows from [Theorem 7.5.5](#).

## 8. Week 9 [ 2025-10-27]

### 8.1. Generation of algebras

Let  $R$  be a commutative ring and let  $A$  and an  $R$ -algebra.

If  $x_1, x_2, \dots, x_n \in A$  recall that there is an  $R$ -algebra homomorphism  $\varphi : R[T_1, T_2, \dots, T_n] \rightarrow A$  with the property that  $\varphi(T_i) = x_i$ ; see [Proposition 6.4.4](#).

**Definition 8.1.1:** The  $R$ -subalgebra of  $A$  generated by  $x_1, x_2, \dots, x_n$  is defined to be the image of the algebra homomorphism  $\varphi$ , and it is denoted  $R[x_1, x_2, \dots, x_n]$ .

*Remark 8.1.2:* Be careful: the notation  $A = R[x_1, x_2, \dots, x_n]$  looks a bit like the notation for the polynomial ring  $R[T_1, T_2, \dots, T_n]$ , but  $A$  need not be isomorphic to a polynomial ring.

*Example 8.1.3:*

- a. Let  $i \in \mathbb{C}$  be a square root of  $-1$ . Then  $\mathbb{Z}[i]$  is the  $\mathbb{Z}$ -subalgebra of  $\mathbb{C}$  generated by  $i$ .

By definition  $\mathbb{Z}[i]$  consists of the complex numbers  $f(i)$  where  $f \in \mathbb{Z}[T]$  is a polynomial.

Since  $T^2 + 1$  is a monic polynomial, an analogue of [Theorem 7.4.4](#) shows for any  $f \in \mathbb{Z}[T]$  that we may write

$$f = q \cdot (T^2 + 1) + r \text{ with } q, r \in \mathbb{Z}[T], \deg r < 2.$$

Thus  $f(i) = r(i)$  which shows that  $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$ .

It is easy to argue that  $\mathbb{Z}[i]$  is a free  $\mathbb{Z}$ -module with basis  $\{1, i\}$ .

- b. You can describe  $\mathbb{Z}[\sqrt{2}]$  or more generally  $\mathbb{Z}[\sqrt{d}]$  as in a.
- c. An  $\alpha \in \mathbb{C}$  is *transcendental* over  $\mathbb{Q}$  if it is not the root of any polynomial  $f \in \mathbb{Q}[T]$ . For example, the real numbers  $e$  and  $\pi$  are known to be transcendental.

For  $\alpha$  transcendental, the subalgebra  $\mathbb{Z}[\alpha]$  is isomorphic to the polynomial ring  $\mathbb{Z}[T]$ .

Indeed, the evaluation mapping

$$\mathbb{Z}[T] \rightarrow \mathbb{Z}[\alpha]$$

is surjective by definition, and for transcendental  $\alpha$  the kernel of this mapping is trivial.

### 8.2. Unique factorization

Let  $A$  be a commutative ring.

**Definition 8.2.1:** An element  $a \in A$  is **irreducible** if  $a \neq 0$ ,  $a \notin A^\times$  and whenever  $a = xy$  with  $x, y \in A$  then either  $x \in A^\times$  or  $y \in A^\times$ .

An element  $a \in A$  is **prime** if the principle ideal  $\langle a \rangle = A \cdot a$  is a prime ideal of  $A$ .

Elements  $a, b \in A$  are **associate** if  $\exists u \in A^\times$  for which  $a = ub$ .

**Proposition 8.2.2:** If  $A$  is an integral domain, a prime element of  $A$  is irreducible.

*Proof:* Let  $a \in A$  be prime, and suppose that  $a = xy$  for  $x, y \in A$ . Then  $xy \in \langle a \rangle$ ; so the definition of prime ideal shows that one of  $x$  or  $y$  is in  $\langle a \rangle$ ; without loss of generality, suppose  $x \in \langle a \rangle$ . To show that  $a$  is irreducible, we will now argue that  $y \in A^\times$ .

Since  $x \in \langle a \rangle$ , we have  $a \mid x$  so that  $x = az$  for some  $z \in A$ . Now we see that  $a = xy = azy$ . Since  $A$  is an integral domain, we find that  $1 = zy$  so  $y \in A^\times$ , as required. ■

**Proposition 8.2.3:** If  $A$  is a PID, then an element of  $A$  is prime if and only if it is irreducible.

*Proof:* Let  $a \in A$ . In view of [Proposition 8.2.2](#), the Proposition will follow if we show  $a$  irreducible  $\Rightarrow a$  prime.

So assume that  $a$  is irreducible. To show that  $\langle a \rangle$  is prime, it suffices by [Corollary 6.2.8](#) to show that  $\langle a \rangle$  is a maximal ideal.

Suppose that  $M$  is an ideal of  $A$  and that  $\langle a \rangle \subseteq M$ . Since  $A$  is a PID, we know that  $M = \langle m \rangle$  for some  $m \in M$ . Since  $\langle a \rangle \subseteq \langle m \rangle$ , we know that  $m \mid a$ . Thus  $a = mx$  for some  $x \in A$ . Since  $\langle a \rangle$  is a prime ideal containing  $mx$ , we conclude that either  $m \in \langle a \rangle$  or  $x \in \langle a \rangle$ .

If  $m \in \langle a \rangle$  then  $\langle m \rangle \subseteq \langle a \rangle$  and we conclude that  $M = \langle m \rangle = \langle a \rangle$ .

If  $x \in \langle a \rangle$  then  $x = ay$  for some  $y \in A$ . Then  $a = mx = may$ . Thus  $1 = my$  which shows that  $m \in A^\times$  so that  $M = A$ .

This shows that the ideal  $\langle a \rangle$  is maximal, and thus prime. So  $a$  is indeed a prime element. ■

*Example 8.2.4:* Consider the ring  $A = \mathbb{Z}[\sqrt{-5}]$ .

Notice that the mapping  $\alpha \mapsto \bar{\alpha}$  given for  $\alpha = a + b \cdot \sqrt{-5}$  by  $\bar{\alpha} = a - b \cdot \sqrt{-5}$  is a ring automorphism of  $A$ .

Now, the function  $N : A \rightarrow \mathbb{Z}$  given by  $N(\alpha) = \alpha \cdot \bar{\alpha}$  – i.e.

$$N(a + b \cdot \sqrt{-5}) = (a + b \cdot \sqrt{-5})(a - b \cdot \sqrt{-5}) = a^2 + 5b^2$$

– is a monoid homomorphism. This means  $N(\alpha\beta) = N(\alpha)N(\beta)$  for all  $\alpha, \beta \in A$ .

Further note that if  $\alpha \in A$  and  $N(\alpha) = \pm 1$  then  $\alpha \in A^\times$ . Indeed, we see for such an element  $\alpha$  that  $\alpha^{-1} = \pm \bar{\alpha} \in A$ .

We claim that  $A$  is not a PID. Indeed, first note that 3 is irreducible in  $A$ . Indeed, if  $3 = \alpha\beta$  then  $9 = N(\alpha)N(\beta)$ . If  $N(\alpha) = 9$  then  $\beta \in A^\times$  and similarly if  $N(\beta) = 9$  then  $\alpha \in A^\times$ .

On the other hand, there are no elements  $\alpha \in A$  with  $N(\alpha) = 3$ . Indeed  $N(\alpha) = a^2 + 5b^2$  and the equation  $a^2 + 5b^2 = 3$  has no integer solutions. This proves the irreducibility of 3 in  $A$ .

On the other hand, 3 is not prime in  $A$ . Indeed,

$$9 = (2 + \sqrt{-5}) \cdot (2 - \sqrt{-5}) \in A \cdot 3 \text{ but } 2 \pm \sqrt{-5} \notin A \cdot 3$$

since

$$(\clubsuit) \quad A \cdot 3 = \{3x + 3y\sqrt{-5} \mid x, y \in \mathbb{Z}\}.$$

In particular, the ideal  $I = \langle 3, 2 + \sqrt{-5} \rangle$  is not principal. Indeed, suppose that  $I = A \cdot \alpha$ . Then  $\alpha \mid 3$  and  $\alpha \mid 2 + \sqrt{-5}$ .

Since 3 is irreducible,  $\alpha = 3u$  where  $N(u) = \pm 1$ . But then  $A \cdot \alpha = A \cdot 3$  so that  $2 + \sqrt{-5} \in I = A \cdot 3$ . But this is impossible by  $(\clubsuit)$ .

**Definition 8.2.5:** An integral domain  $A$  is a **unique factorization domain** (or **UFD**) if  $\forall a \in A$  for which  $a \neq 0$  and  $a \notin A^\times$ , we have the following:

- There are finitely many irreducible elements  $p_1, \dots, p_n \in A$  such that  $a = p_1 \dots p_n$ .
- The expression in a. is unique up to associates: namely, if also  $a = q_1 \dots q_m$  for irreducible elements  $q_j \in A$ , then  $n = m$  and – after possibly renumbering the factors – the elements  $p_i$  and  $q_i$  are associate for each  $i = 1, 2, \dots, n$ .

*Example 8.2.6:*

- Any field is a UFD (vacuously)
- We will show shortly that any PID is a UFD.
- And we will show that if  $A$  is a UFD, then  $A[T]$  is a UFD. Thus by induction we see for example that  $F[T_1, \dots, T_d]$  is a UFD for any field and any  $d \in \mathbb{N}$ .
- The ring  $\mathbb{Z}[\sqrt{-5}]$  is not a UFD. Indeed,  $6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$ . We have seen before that 3 is irreducible. Similar arguments show that 2 is irreducible and that the elements  $1 \pm \sqrt{-5}$  are irreducible. But 3 is not associate to either of  $1 \pm \sqrt{-5}$ .

**Proposition 8.2.7:** Let  $A$  be a UFD.

- Let  $p_i \in A$  be irreducibles for  $i = 1, \dots, n$ . If  $p$  is irreducible and  $p \mid p_1 \dots p_n$  then for some  $1 \leq i \leq n$ ,  $p$  and  $p_i$  are associates.
- If  $p \in A$  a non-zero non-unit, then  $p$  is prime if and only if it is irreducible.

*Proof:*

- The assumption means that  $px = p_1 \dots p_n$  for some  $n$ . If  $x$  is a unit, the definition of UFD shows that  $n = 1$  and  $p$  is associate with  $p_1$ . Otherwise, we may write  $x = q_1 \dots q_s$  for irreducibles  $q_j \in A$ . Now  $p q_1 \dots q_s = p_1 \dots p_n$ , and the definition of UFD shows that  $s + 1 = n$  and for some  $i$ ,  $p$  and  $p_i$  are associate. This completes the proof of a.
- Let  $p \in A$  be a non-zero non-unit. In view of [Proposition 8.2.2](#), we must argue that if  $p$  is irreducible, then  $p$  is prime.

Thus we suppose that  $p \mid ab$  for  $a, b \in A$ . We must show that  $p \mid a$  or  $p \mid b$ .

If  $a$  is a unit, we see that  $p \mid b$ , and similarly if  $b$  is a unit, then  $p \mid a$ . Thus we may suppose that neither  $a$  nor  $b$  is a unit. We now find irreducible elements  $p_1, \dots, p_n$  and  $q_1, \dots, q_m$  in  $A$  such that  $a = p_1 \dots p_n$  and  $b = q_1 \dots q_m$ . Thus

$$p \mid p_1 \cdots p_n \cdot q_1 \cdots q_m.$$

Now the result in part a. shows that  $p$  is associate to some  $p_i$  or some  $q_j$ . But it is then immediate that  $p$  divides  $a$  or  $p$  divides  $b$ . This completes the proof. ■

### 8.3. PIDs have unique factorization

**Proposition 8.3.1:** Let  $A$  be a PID. If

$$I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$$

is a sequence of ideals of  $A$ , then  $\exists N$  such that for  $n \geq N$  we have  $I_n = I_N$ .

*Proof:* Let  $I = \bigcup_{j \in \mathbb{N}} I_j$ , and check that  $I$  is an ideal of  $A$ .

Since  $A$  is a PID,  $I = A \cdot a$  for some  $a \in A$ . Now, by the definition of  $I$ , we see that  $a \in I_j$  for some  $j \in \mathbb{N}$ . But then for any  $\ell \geq j$  we have  $I_\ell \subseteq I = A \cdot a \subseteq I_\ell$  so that  $I = I_\ell = I_j$  as required. ■

**Proposition 8.3.2:** Let  $A$  be a PID and let  $p \in A$  be prime.

- If  $q$  is prime and  $p \mid q$  then  $p$  and  $q$  are associate.
- let  $a_1, \dots, a_m \in A$  and suppose that  $p \mid a_1 \cdots a_m$ . Then  $p \mid a_j$  for some  $j$ .

*Proof:*

- Write  $px = q$  for some  $x \in A$ . Since  $q$  is prime, either  $q \mid p$  or  $q \mid x$ . If  $qy = x$  then  $pqy = q$  so that  $py = 1$ ; thus  $p$  is a unit, contrary to the assumption that  $p$  is prime.

If  $qy = p$  then  $qyx = q$  which shows that  $yx = 1$ . Thus  $y \in A^\times$  so that  $p$  and  $q$  are associate.

- Proceed by induction on  $m$ . When  $m = 1$  there is nothing to do. Now suppose that  $n \in \mathbb{N}_{>0}$  and the result holds for  $n$  elements of  $A$ . Let  $a_1, \dots, a_{n+1} \in A$  and suppose that

$$p \mid a_1 \cdots a_n \cdot a_{n+1} = (a_1 \cdots a_n) \cdot a_{n+1}.$$

We see that  $(a_1 \cdots a_n) \cdot a_{n+1} \in A \cdot p$ , a prime ideal. Thus, either  $a_1 \cdots a_n \in A \cdot p$  or  $a_{n+1} \in A \cdot p$ .

In the first case,  $p \mid a_1 \cdots a_n$  and we may apply induction to learn that  $p \mid a_j$  for some  $1 \leq j \leq n$ . And in the second case,  $p \mid a_{n+1}$ . Thus the conclusion of the Proposition holds in all cases. ■

**Theorem 8.3.3:** Let  $A$  be a PID. Then  $A$  is a UFD.

*Proof:* Let  $a \in A$  be a non-zero non-unit.

We first argue that  $a$  may be expressed as a product of irreducible elements of  $A$ .



To give the proof, we proceed by contradiction. Thus, we suppose that  $A$  has no such factorization.

First note that  $a$  is not irreducible (otherwise,  $a$  can be written as a product of irreducibles)

Now set  $a = a_0$  and write  $a_0 = x \cdot y$  where neither  $x$  nor  $y$  is a unit. Now, at least one of  $x$  or  $y$  can not be written as a product of irreducibles; without loss of generality we suppose it to be  $x$  and we set  $a_1 = x$ . Since  $y$  is not a unit, there is a proper inclusion of ideals  $A \cdot a_0 \subset A \cdot a_1$ .

Since  $a_1$  can't be written as a product of irreducibles, we can repeat this argument with  $a_1$  playing the role of  $a_0$ . Continuing in this manner, we produce a sequence of elements  $a_0, a_1, a_2, \dots, a_n, \dots$  such that

$$A \cdot a_0 \subset A \cdot a_1 \subset \dots \subset A \cdot a_n \subset \dots$$

where each inclusion is proper.

Now, the existence of such a sequence contradicts [Proposition 8.3.1](#). This contradiction proves that every element of  $A$  has a factorization as a product of irreducibles.

It remains to prove the uniqueness of factorization.

So, suppose that  $a = p_1 \dots p_n = q_1 \dots q_m$  where  $p_i$  and  $q_j$  are irreducible in  $A$  and where  $n \leq m$ .

We proceed by induction on  $n$ . When  $n = 1$ , we have  $p_1 = q_1 \dots q_m$  and [Proposition 8.3.2](#) shows for some  $j$  that  $p_1$  and  $q_j$  are associate; i.e.  $q_j = up_1$  for  $u \in A^\times$ . Reordering the terms we may suppose that  $j = 1$ .

Now

$$p_1 = up_1 q_2 \dots q_m \Rightarrow 1 = uq_2 \dots q_m.$$

This shows that  $q_j$  is a unit for  $j = 2, \dots, m$ . Since the  $q_j$  are primes, this means that  $j = 1$ . This completes the proof when  $n = 1$ .

Now assume the result is known for  $n$ , and suppose that

$$p_1 p_2 \dots p_{n+1} = q_1 \dots q_m \text{ for some } m \geq n + 1.$$

Then  $p_{n+1}$  divides  $q_1 \dots q_m$  and [Proposition 8.3.2](#) shows that  $p_{n+1}$  is associate with some  $q_j$ . Re-ordering the factors, we may suppose that  $j = m$ ; write  $q_m = up_{n+1}$  for some  $u \in A^\times$ .

Thus

$$p_1 \dots p_n p_{n+1} = q_1 \dots q_{m-1} up_{n+1} \Rightarrow p_1 \dots p_n = uq_1 \dots q_{m-1}.$$

Now it follows by the induction hypothesis that  $n = m - 1$  and that after possibly re-ordering the  $q_i$  one knows that  $p_i$  and  $q_i$  are associate for all  $i = 1, \dots, n$ . Since also  $p_{n+1}$  is associate with  $q_m$ , the result follows by induction. ■

**Proposition 8.3.4:** Let  $A$  be a UFD and let  $a, b \in A \setminus \{0\}$ . Suppose that

$$a = u \cdot p_1^{e_1} p_2^{e_2} \dots p_n^{e_n} \text{ and } b = v \cdot p_1^{f_1} p_2^{f_2} \dots p_n^{f_n}$$

where  $u, v \in A^\times$ , the  $p_i$  are irreducible elements of  $A$  which are pairwise not associate, and  $e_i, f_j \in \mathbb{N}$ .

Set

$$d = p_1^{\min(e_1, f_1)} p_2^{\min(e_2, f_2)} \dots p_n^{\min(e_n, f_n)}.$$

Then  $d$  is a gcd of  $a, b$ .

In particular, gcds exists in UFDs.

*Proof:* The condition on the exponents of the  $p_i$  shows that  $d \mid a$  and  $d \mid b$ .

Now suppose that  $c$  is any common divisor of  $a$  and  $b$ . and write

$$c = w q_1^{g_1} q_2^{g_2} \dots q_m^{g_m}$$

for irreducibles  $q_j$  in  $A$  and a unit  $w \in A^\times$ . Fix  $1 \leq i \leq m$ . Since  $q_i$  divides  $c$  and hence divides  $a$  and  $b$ , and since irreducibles are prime in a UFD by [Proposition 8.2.7](#), we see that  $p_i$  and  $q_j$  are associate for some  $j$ .

After replacing  $q_i$  by an associate (and possibly modifying  $w \in A^\times$ ) we may suppose that  $\{q_1, \dots, q_m\} \subseteq \{p_1, \dots, p_n\}$ . Thus after re-indexing the irreducibles  $p_i$ , we may write

$$c = w p_1^{g_1} \dots p_n^{g_n}.$$

Now  $c \mid a \Rightarrow g_i \leq e_i$  for each  $i$ , and  $c \mid b \Rightarrow g_i \leq f_i$  for each  $i$ . Thus  $c \mid d$  which completes the proof that  $d = \gcd(a, b)$ . ■

*Example 8.3.5:* In general, for a UFD  $A$  it is not true that  $\langle a, b \rangle = \langle \gcd(a, b) \rangle$ .

We are going to show below that for a field  $F$ , the two-variable polynomial ring  $F[T, U]$  is a UFD. Since  $T$  is irreducible in  $F(U)[T]$ ,  $T$  is irreducible in  $F[T, U]$ . Similarly,  $U$  is irreducible in  $F[T, U]$ .

Thus  $\gcd(T, U) = 1$ . But  $\langle T, U \rangle \neq F[T, U]$ . To see this, note that for any  $f \in \langle T, U \rangle$ , evaluation at  $T \mapsto 0$  and  $U \mapsto 0$  gives  $f(0, 0) = 0$ . This shows that  $1 \notin \langle T, U \rangle$ .

## 8.4. Rings of fractions

Let  $A$  be a commutative ring.

When  $A$  is an integral domain, we would like to know that there is a field  $K$  containing  $A$  in the same way that  $\mathbb{Q}$  contains  $\mathbb{Z}$ .

For this, we'll use a more general construction known as localization. Compare with [Dummit-Foote], § 7.5 and § 15.4.

**Definition 8.4.1:** A subset  $S \subset A$  is said to be **multiplicatively closed** if  $0 \notin S$ ,  $1 \in S$ , and  $a, b \in S \Rightarrow a \cdot b \in S$ .

Notice that  $S$  is a (multiplicative) moniod; in fact, it is a submonoid of  $(A, \times)$ .

**Theorem 8.4.2:** Let  $S$  be a multiplicatively closed subset of  $A$ . Then there is a commutative ring  $A[S^{-1}]$  together with a ring homomorphism  $\iota : A \rightarrow A[S^{-1}]$  such that  $\iota(s) \in A[S^{-1}]^\times$  for each  $s \in S$ .

Every element of  $A[S^{-1}]$  can be written in the form  $\iota(a) \cdot \iota(s)^{-1}$  for some  $a \in A$  and  $s \in S$ .

The ring  $A[S^{-1}]$  satisfies the universal mapping property: if  $B$  is any commutative ring and if  $\varphi : A \rightarrow B$  is a ring homomorphism such that  $\varphi(s) \in B^\times$  for each  $s \in S$ , then there is a unique ring homomorphism  $\psi : A[S^{-1}] \rightarrow B$  such that  $\psi \circ \iota = \varphi$ .

*Proof:* Consider the set  $D = \{(a, s) \mid a \in A, s \in S\}$  and define a relation  $\sim$  on  $D$  by the condition:

$$(a, s) \sim (b, t) \Leftrightarrow \exists x \in S, x(at - bs) = 0.$$

Note that if  $A$  is an integral domain, then

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

We now confirm that  $\sim$  is an equivalence relation: reflexive and symmetric are immediate. To see that  $\sim$  is transitive, let  $(a, s), (b, t), (c, u) \in D$  and suppose that  $(a, s) \sim (b, t)$  and  $(b, t) \sim (c, u)$ . We show that  $(a, s) \sim (c, u)$ . We know that  $x(at - bs) = 0$  and  $y(bu - ct) = 0$  for some  $x, y \in S$ . Multiplying the first equation by  $yu$  and the second by  $xs$  we find that

$$xatyu = xbsyu = yctsx.$$

We conclude that  $au(xty) = cs(xty)$ ; since  $xty \in S$ , it follows that  $(a, s) \sim (c, u)$ .

Write  $a/s = \frac{a}{s}$  for the equivalence class of the element  $(a, s) \in D$ . Note that for  $t \in S$ , we have  $a/s = at/st$ .

For  $\alpha = a/s$  and  $\beta = b/t$  we define

$$\alpha + \beta = \frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st} \text{ and } \alpha\beta = \frac{a}{s} \cdot \frac{b}{t} = \frac{ab}{st}.$$

Write  $Q$  for the set of equivalence classes in  $D$  for  $\sim$ ; we claim that  $Q$  becomes a commutative ring with the indicated operations.

There are a number of things to check(!):

- the operations  $+$  and  $\times$  on  $Q$  are well-defined – i.e. they are independent of the chosen representatives of the equivalence classes.
- $Q$  is an abelian group under addition with identity element  $\frac{0}{1} = \frac{0}{s}$  for any  $s \in S$ . The operation of taking additive inverses is given by

$$-\left(\frac{a}{b}\right) = \frac{-a}{b}.$$

- c. multiplication in  $Q$  is associative, commutative and distributive. Moreover,  $\frac{1}{1} = \frac{s}{s}$  for any  $s \in S$  is a multiplicative identity.

The proofs of these assertions are straightforward and omitted.

Next, we define  $\iota : A \rightarrow Q$  by

$$\iota(a) = \frac{a}{1} = \frac{as}{s} \text{ for any } s \in S.$$

One checks that this mapping is a ring homomorphism; again the proof is omitted.

Finally, let  $\varphi : A \rightarrow B$  is a ring homomorphism for which  $\varphi(s) \in B^\times$  for  $s \in S$ . We must find a unique  $\psi : A[S^{-1}] \rightarrow B$  with the required property.

We first treat uniqueness: For  $a/s \in A[S^{-1}]$ , note that  $a/s = a/1 \cdot 1/s$  so that

$$\psi(a/s) = \psi(a/1)\psi(1/s) = \psi(\iota(a)) \cdot \psi(s)^{-1} = \varphi(a)\varphi(s)^{-1}.$$

Thus it only remains to check that this formula yields a well-defined ring homomorphism. The proof is routine. ■

**Proposition 8.4.3:** Let  $A$  be a commutative ring and let  $S$  a multiplicatively closed subset of  $A$ .

- a. The kernel of the map  $\iota : A \rightarrow A[S^{-1}]$  is given by

$$\ker \iota = \{a \in A \mid \exists s \in S, sa = 0\}.$$

- b. If  $A$  is an integral domain, then  $\iota$  is injective.

**Definition 8.4.4:** The ring constructed in [Theorem 8.4.2](#) is called the ring of fractions of  $A$  for the multiplicatively closed subset  $S$ .

**Proposition 8.4.5:** If  $A$  is an integral domain and  $S = A \setminus \{0\}$ , then the ring of fractions  $A[S^{-1}]$  is a field, called the **field of fractions of  $A$** . We usually write  $\text{Frac}(A)$  for this field.

*Proof:* We only must confirm that any  $\alpha \in A$  with  $\alpha \neq 0$  is a unit. But  $\alpha$  has the form  $\frac{a}{s}$  for  $a \in A$  and  $s \in S$ . And  $\frac{a}{s} \neq 0$  implies that  $a \neq 0$ . Thus  $a \in S = A \setminus \{0\}$  so that  $\frac{s}{a} \in A[S^{-1}]$ . Now it is straightforward to see that

$$\frac{a}{s} \cdot \frac{s}{a} = \frac{1}{1} = 1.$$

Thus  $\alpha$  is a unit; this confirms that  $A[S^{-1}]$  is a field. ■

*Example 8.4.6:*

- a. We have  $\text{Frac}(\mathbb{Z}) = \mathbb{Q}$ .  
b. Let  $K$  be a field. We write  $K(T) = \text{Frac}(K[T])$  for the field of fractions of the polynomial ring  $K[T]$ .

More generally, we write  $K(T_1, T_2, \dots, T_n) = \text{Frac}(K[T_1, T_2, \dots, T_n])$  for the field of fractions of the multi-variable polynomial ring  $K[T_1, T_2, \dots, T_n]$ .

The ring of fractions satisfies the following universal mapping property:

**Theorem 8.4.7:** Let  $A$  an integral domain and  $S$  a multiplicatively closed subset of  $A$ . If  $B$  is any ring and  $\varphi : A \rightarrow B$  any ring homomorphism such that  $\varphi(s) \in B^\times$  for each  $s \in S$ , then there is a unique ring homomorphism

$$\psi : A[S^{-1}] \rightarrow B$$

such that  $\psi|_A = \varphi$ .

*Proof:* We first prove uniqueness of  $\psi$ . Since any element  $\alpha$  of  $A[S^{-1}]$  may be written in the form  $\alpha = a \cdot s^{-1}$  for  $a \in A$  and  $s \in S$  and since  $\varphi(s)$  is a unit in  $B$ , we must have

$$(\clubsuit) \quad \psi(\alpha) = \psi(a)\psi(s^{-1}) = \psi(a)\psi(s)^{-1} = \varphi(a)\varphi(s)^{-1}.$$

Thus once we check that  $(\clubsuit)$  indeed defines a ring homomorphism, we see that there is only one possible definition of  $\psi$ .

It remains to check that  $(\clubsuit)$  indeed defines a ring homomorphism. First, we must confirm that this definition of  $\psi$  is well defined.

We suppose that  $\frac{a}{s} = \frac{a'}{s'}$  in  $A[S^{-1}]$ ; thus  $as' = a's$ . Applying  $\varphi$  we see that

$$\varphi(a)\varphi(s') = \varphi(a')\varphi(s)$$

so that

$$(\diamond) \quad \varphi(a)\varphi(s)^{-1} = \varphi(a')\varphi(s')^{-1}.$$

Now,  $\psi(a/s) = \varphi(a)\varphi(s)^{-1}$  and  $\psi(a'/s') = \varphi(a')\varphi(s')^{-1}$  so  $(\diamond)$  shows that  $\psi(a/s) = \psi(a'/s')$ .

It is now straightforward to check that  $\psi$  is additive and multiplicative.

For example

$$\begin{aligned} \psi\left(\frac{a}{s} + \frac{a'}{s'}\right) &= \psi\left(\frac{as' + a's}{ss'}\right) \\ &= \varphi(as' + a's) \cdot \varphi(ss')^{-1} \\ &= \varphi(as')\varphi(ss')^{-1} + \varphi(a's)\varphi(ss')^{-1} \\ &= \varphi(a)\varphi(s)^{-1} + \varphi(a')\varphi(s')^{-1} \\ &= \psi\left(\frac{a}{s}\right) + \psi\left(\frac{a'}{s'}\right). \end{aligned}$$

■

**Proposition 8.4.8:** Let  $A$  be an integral domain and let  $S$  be a mult. closed subset of  $A$ .

- An ideal  $I$  of  $A$ , generates an ideal  $I^e = A[S^{-1}] \cdot I$  of  $A[S^{-1}]$
  - An ideal  $J$  of  $A[S^{-1}]$  determines an ideal  $J^c$  of  $A$  by the rule  $J^c = \{x \in J \mid x \in A\} = A \cap J$ .
- a. For any ideal  $J$  of  $A[S^{-1}]$  we have  $J = (J^c)^e$ .  
b. For an ideal  $I$  of  $A$  we have

$$(I^e)^c = \{a \in A \mid sa \in I \text{ for some } s \in S\}.$$

- c. The assignment  $I \mapsto I^e$  determines a bijection

$$\{\text{prime ideals } I \text{ of } A \text{ with } I \cap S = \emptyset\} \rightarrow \{\text{prime ideals of } A[S^{-1}]\}$$

with inverse given by  $J \mapsto J^c$ .

*Proof:* Write  $\iota : A \rightarrow A[S^{-1}]$  for the natural mapping  $\iota(a) = a/1$ .

- a. It is true by definition that  $(J^c)^e \subseteq J$ . For the reverse inclusion, suppose that  $a/s \in J$ . Then  $a/1 = s \cdot (a/s) \in J$ . This shows that  $a \in J^c$ . Now  $a/s = (a/1)(1/s) \in (I^e)^c$  as required.

- b. Let  $I' = \{a \in A \mid sa \in I \text{ for some } s \in S\}$ . We first argue that  $I' \subseteq (I^e)^c$ .

If  $a \in I'$  choose  $s \in S$  with  $sa = b \in I$ . Then  $a = \frac{b}{s} \in I^e$  which shows that  $a \in (I^e)^c$ .

On the other hand, let  $a \in (I^e)^c$ . Thus  $a/1 = b/s$  for  $b \in I$  and  $s \in S$ . Thus  $sa = b \cdot 1 \in I$  which shows that  $a \in I'$ . This shows that  $(I^e)^c \subseteq I'$  so that equality holds, as required.

- c. Let's write

$$\mathcal{P} = \{\text{prime ideals } I \text{ of } A \text{ with } I \cap S = \emptyset\} \text{ and } \mathcal{Q} = \{\text{prime ideals of } A[S^{-1}]\}$$

and let  $e : \mathcal{P} \rightarrow \mathcal{Q}$  be the mapping  $I \mapsto I^e$  and  $c : \mathcal{Q} \rightarrow \mathcal{P}$  be the mapping  $J \mapsto J^c$ .

Part a. shows that  $e \circ c = \text{id}_{\mathcal{Q}}$ . Now suppose that  $I \in \mathcal{P}$ .

We must argue that  $(c \circ e)(I) = I$ . According to b., we know that  $(c \circ e)(I) = I'$  where  $I' = \{a \in A \mid sa \in I \text{ for some } s \in S\}$ .

We claim that since  $I$  is prime and  $I \cap S = \emptyset$  that  $I = I'$ . Evidently  $I \subseteq I'$  so we must prove the  $I' \subseteq I$ . Let  $a \in I'$ . By the definition of  $I'$  we know that  $sa \in I$  for some  $s \in S$ . Since  $I$  is prime, either  $s \in I$  or  $a \in I$ . But since  $I \cap S = \emptyset$  we know that  $s \notin I$  so that  $a \in I$ . This shows that  $I = I'$ , so that  $(c \circ e)(I) = I$ .

■

## 8.5. Unique factorization in polynomial rings

Let  $A$  be a UFD. In this section we are going to prove that the polynomial ring  $A[T]$  is a UFD as well.

Write  $K$  for the field of fraction and notice that  $A[T] \subset K[T]$ .

Now,  $K[T]$  is a PID, hence in particular a UFD. We are going to use unique factorization in  $K[T]$  together with unique factorization in  $A$  to deduce unique factorization in  $A[T]$ .

An important tool here is *Gauss' Lemma* comparing factorization in  $A[T]$  with that in  $K[T]$ .

Recall that [Proposition 8.3.4](#) we can compute gcds for a pair of elements of a UFD. If  $a_0, a_1, \dots, a_d \in A$  then we set

$$\gcd(a_0, \dots, a_d) = \gcd(\gcd(a_0, \dots, a_{d-1}), a_d).$$

**Definition 8.5.1:** Let  $f \in A[T]$  be a non-zero polynomial and let  $c = \gcd(\text{coefficients of } f)$ . Then  $c$  is known as the **content** of  $f$ , denoted  $\text{content}(f)$ .

The polynomial  $f$  is **primitive** if  $\text{content}(f)$  is a unit in  $A$ .

**Proposition 8.5.2:** Let  $f \in A[T]$  be a non-zero polynomial, and write  $c = \text{content}(f) \in A$ . Then there is a primitive polynomial  $g \in A[T]$  for which  $f = c \cdot g$ .

*Proof:* Write

$$f = \sum_{i=0}^n a_i T^i \text{ with } a_i \in A.$$

Then by definition  $c = \gcd(a_0, a_1, \dots, a_n)$ . In particular, for each  $i$ ,  $c \mid a_i$  so we may write  $a_i = cb_i$  for  $b_i \in A$ .

Set  $g = \sum_{i=0}^n b_i T^i$  so that indeed  $f = c \cdot g$ .

Now observe that

$$\text{content}(g) = \gcd(b_0, b_1, \dots, b_n) = \gcd\left(\frac{a_0}{c}, \frac{a_1}{c}, \dots, \frac{a_n}{c}\right) = \frac{c}{c} = 1$$

so that  $g$  is primitive, as required. ■

**Proposition 8.5.3:** Let  $p \in A$  be irreducible and consider the assignment

$$h \mapsto \bar{h} : A[T] \rightarrow (A/A \cdot p)[T]$$

defined for  $h = \sum_{i=0}^m a_i T^i \in A[T]$  by

$$\bar{h} = \sum_{i=0}^m (a_i + A \cdot p) T^i = \sum_{i=0}^m \bar{a}_i T^i.$$

- The assignment  $h \mapsto \bar{h}$  is a ring homomorphism, and
- For  $h \in A[T]$ ,  $\bar{h} = 0$  if and only if  $p \mid \text{content}(h)$ .

*Proof:*

- In fact, the mapping  $h \mapsto \bar{h}$  is the ring homomorphism given by [Proposition 8.2.2](#) for the quotient mapping  $\pi : A \rightarrow A/A \cdot p$  and the assignment  $T \mapsto T$ .

b. Just observe that  $\bar{h} = 0$  if and only if  $p \mid a_i$  for each  $i$ .

■

**Proposition 8.5.4** (Gauss' Lemma): Let  $f, g \in A[T]$  be non-zero polynomials.

- a.  $c(f \cdot g) = c(f) \cdot c(g)$ .
- b. In particular, if  $f$  and  $g$  are primitive, then  $f \cdot g$  is primitive.

*Proof:* First note that a. is an immediate consequence of b. Indeed, write  $f = cf_0$  and  $g = dg_0$  for primitive  $f_0, g_0 \in A[T]$ . Then  $fg = c \cdot d \cdot f_0 \cdot g_0$ . By b.  $f_0 \cdot g_0$  is primitive, and we see that  $\text{content}(fg) = cd$  as required.

We now prove b. We proceed by contradiction. Thus, we suppose that  $f, g$  are primitive but that  $f \cdot g$  is not primitive.

Write  $d = \text{content}(f \cdot g)$ . Since  $f \cdot g$  is not primitive, there is an irreducible element  $p \in A$  for which  $p \mid d$ .

Since  $p$  is irreducible, it is prime [Proposition 8.2.7](#) so that  $A \cdot p$  is a prime ideal. Thus  $A/A \cdot p$  is an integral domain, and it follows that  $(A/A \cdot p)[T]$  is an integral domain.

Now consider the homomorphism  $h \mapsto \bar{h} : A[T] \rightarrow (A/A \cdot p)[T]$  of [Proposition 8.5.3](#).

Since  $p \mid \text{content}(fg)$  the proposition shows that  $\overline{f \cdot g} = 0$ . On the other hand, since  $f$  and  $g$  are primitive, the proposition shows that  $\bar{f} \neq 0$  and  $\bar{g} \neq 0$ .

Since  $\overline{f \cdot g} = \bar{f} \cdot \bar{g}$  and since  $(A/A \cdot p)[T]$  is an integral domain, we have arrived at a contradiction. This proves that  $\text{content}(f \cdot g) = 1$ , and the proof is complete. ■

**Corollary 8.5.5:** Let  $f \in A[T]$  be primitive. If  $f = g \cdot h$  for non-constant polynomials  $g, h \in K[T]$ , there are non-zero elements  $r, s \in K$  such that  $r \cdot g \in A[T]$ ,  $s \cdot h \in A[T]$  and  $f = gh$ .

In particular, if  $f$  is reducible in  $K[T]$ , then  $f$  is reducible in  $A[T]$ .

*Proof:* Use [Proposition 8.5.2](#) to write

$$g = \left(\frac{x}{y}\right)g_0 \text{ and } h = \left(\frac{z}{w}\right)h_0 \text{ for primitive polys } g_0, h_0 \in A[T],$$

where  $x, y, z, w \in R$  with  $y \cdot w \neq 0$ .

Then we find that

$$f = gh = \left(\frac{xz}{yw}\right) \cdot g_0 h_0$$

so that

$$yw f = xz g_0 h_0.$$



Since  $f$  is primitive,  $\text{content}(ywf) = yw$ . Moreover, [Proposition 8.5.4](#) shows that  $g_0h_0$  is primitive, so that  $\text{content}(xzg_0h_0) = xz$ . It follows that  $yw$  and  $xz$  are associate; i.e.  $ywu = xz$  for  $u \in A^\times$ .

Using cancellation in the integral domain  $A$ , we finally see that

$$ywf = ywug_0h_0 \Rightarrow f = (ug_0) \cdot h_0$$

which proves the result. ■

**Corollary 8.5.6:** Let  $A$  be a UFD with field of fractions  $K$  and let  $f \in A[T]$ . Suppose that  $\text{content}(f) = 1$ . Then  $f$  is irreducible in  $K[T]$  if and only if  $f$  is irreducible in  $A[T]$ .

In particular, if  $f \in A[T]$  is monic and irreducible in  $F[T]$  then  $f$  is irreducible in  $K[T]$ .

*Proof:*  $(\Rightarrow)$ : By [Corollary 8.5.5](#), we know that  $f$  reducible in  $K[T]$  implies that  $f$  is reducible in  $A[T]$ .

$(\Leftarrow)$ : Since the content of  $f$  is a unit, if  $f = gh$  for  $g, h \in A[T]$ , then  $g, h$  are non-constant - i.e.  $g, h \notin A$ . But this factorization shows that  $f$  is reducible in  $K[T]$ . ■

**Theorem 8.5.7:** An integral domain  $A$  is a UFD if and only if  $A[T]$  is a UFD.

*Proof:*  $(\Leftarrow)$ : Suppose  $A[T]$  is a UFD. Viewing  $A$  as degree 0 polynomials in  $A[T]$  and using the fact that  $\deg(f \cdot g) = \deg(f) + \deg(g)$ , it is straightforward to see that  $A$  itself is a UFD.

$(\Rightarrow)$ : Let  $f \in A[T]$ , and let  $c = \text{content}(f) \in A$ , so that  $f = c \cdot f_0$  for a primitive polynomial  $f_0 \in A[T]$ . Such a factorization is unique up to multiplication by a unit in  $A$ . Since  $c$  can be factored uniquely into a product of irreducibles in  $A$ , it suffices to prove that  $f_0$  can be factored uniquely into a product of irreducibles. Thus we may and will suppose that  $f$  is primitive.

Since  $K[T]$  is a PID,  $f$  may be written as a scalar times a product of irreducibles in  $K[T]$ . By [Corollary 8.5.5](#), there are polynomials  $p_1, \dots, p_n \in A[T]$  such that

$$f = p_1 \cdots p_n$$

such that  $p_i$  is irreducible in  $K[T]$ . Now [Corollary 8.5.6](#) shows that each  $p_i$  is irreducible in  $A[T]$ .

This completes the proof that any element of  $A[T]$  has a factorization as a product of irreducibles.

We now prove that such a factorization is unique. Suppose that

$$f = p_1 \cdots p_m = q_1 \cdots q_n$$

for irreducibles  $p_i, q_j \in A[T]$ . Since  $f$  is primitive,  $p_i$  and  $q_j$  are primitive  $\forall i, j$ . According to [Corollary 8.5.6](#), the  $p_i, q_j$  remain irreducible in  $K[T]$ . Since  $K[T]$  is a PID hence a UFD, we conclude that  $n = m$  and that after possibly re-ordering,  $p_i$  is associate with  $q_i$  in  $K[T]$ .

Since  $p_i$  and  $q_i$  are primitive, it follows that  $p_i$  is associate with  $q_i$  in  $A[T]$  and the proof is complete. ■

**Corollary 8.5.8:** Let  $A$  be a UFD. For any  $n \in \mathbb{N}$ ,  $A[T_1, T_2, \dots, T_n]$  is a UFD.

*Proof:* For  $n \geq 1$ , note that  $A[T_1, T_2, \dots, T_n] = A[T_1, T_2, \dots, T_{n-1}][T_n]$ . Thus the result follows by induction on  $n$ . ■

## 9. Week 10 [2025-11-03]

### 9.1. Eisenstein's irreducibility criterion

Let  $A$  be a PID, let  $K$  be the field of fractions of  $A$ , and let  $0 \neq p \in A$  be a prime (equivalently: irreducible) element.

**Theorem 9.1.1** (Eisenstein's criteria): Let  $n \in \mathbb{N}_{>0}$  and suppose

$$f = T^n + \sum_{i=0}^{n-1} a_i T^i \in A[T] \text{ where } p \mid a_i \text{ for each } 0 \leq i \leq n-1 \text{ and } p^2 \nmid a_0.$$

Then  $f$  is irreducible in  $A[T]$  and in  $K[T]$ .

*Proof:* Note that the content  $\text{content}(f) = 1$  since  $f$  is monic.

Now suppose that  $f = \alpha\beta$  for  $\alpha, \beta \in A[T]$  and that  $\deg \alpha > 0$ . According to [Corollary 8.5.6](#), to prove that  $f$  is irreducible in  $A[T]$ , it is enough to argue that  $f$  is irreducible in  $K[T]$ . For this, we only must argue that  $\deg \beta = 0$ .

Now, let  $(g \mapsto \bar{g}) : A[T] \rightarrow (A/A \cdot p)[T]$  be the “reduction mod  $p$ ” homomorphism.

Then we have

$$T^n = \bar{f} = \bar{\alpha}\bar{\beta} = \bar{\alpha}\bar{\beta}.$$

Since  $A$  is a PID, we know that  $A/A \cdot p$  is a field. We now use unique factorization in the UFD  $(A/A \cdot p)[T]$ . Note that  $T$  is irreducible. Thus it follows that  $\bar{\alpha}$  is associate to  $T^m$  and  $\bar{\beta}$  is associate to  $T^\ell$  for some  $m, \ell \in \mathbb{N}$  with  $m + \ell = n$ . Thus  $\bar{\alpha} = \bar{c}T^m$  and  $\bar{\beta} = \bar{d}T^\ell$  for some  $c, d \in A$ .

Moreover, we have  $\deg \alpha = \deg \bar{\alpha} = m$  and  $\deg \beta = \deg \bar{\beta} = \ell$ .

If  $m, \ell > 0$  then the constant terms of  $\alpha$  and of  $\beta$  are contained in  $A \cdot p$ . Thus the constant term  $a_0$  of  $f$  is contained in  $A \cdot p^2$ . But this contradicts the hypothesis that  $p^2 \nmid a_0$ . This contradiction proves that  $m = 0$  or  $\ell = 0$  as required. ■

## 9.2. Extension of scalars: an important special case of the tensor product

Let  $A$  be a commutative ring, and let  $B$  be an  $A$ -algebra.

Given an  $A$ -module  $M$ , we would like to construct a  $B$ -module  $B \otimes_A M$  corresponding to  $M$  together with a homomorphism of  $A$ -modules  $M \rightarrow B \otimes_A M$ .

**Proposition 9.2.1:** There is a  $B$ -module  $B \otimes_A M$  together with an  $A$ -module homomorphism  $\iota : M \rightarrow B \otimes_A M$  such that for any  $B$ -module  $N$  and any  $A$ -module homomorphism  $\varphi : M \rightarrow N$  there is a unique  $B$ -module homomorphism  $\psi : B \otimes_A M \rightarrow N$  such that

$$\psi \circ \iota = \varphi.$$

Moreover, the pair  $B \otimes_A M, \iota$  is uniquely determined up to isomorphism by these conditions.

*Proof:* Uniqueness follows because the pair  $(B \otimes_A M, \iota)$  is an initial object in the category whose objects are pairs  $(C, \tau)$  where  $C$  is a  $B$ -module, and  $\tau : M \rightarrow C$  an  $A$ -module homomorphism, and where a morphism  $(C, \tau) \rightarrow (C', \tau')$  is a  $B$ -module homomorphism  $f : C \rightarrow C'$  such that  $f \circ \tau = \tau'$ .

To construct  $B \otimes_A M$ , consider the free  $\mathbb{Z}$ -module  $F$  on the set  $B \times M$ . We just write  $(b, m)$  for a basis element of  $F$  determined by  $(b, m) \in B \times M$ . Thus a typical element of  $F$  is a finitely supported function  $a : B \times M \rightarrow \mathbb{Z}$  which we write in the form

$$\sum_{(b,m) \in B \times M} a_{b,m}(b, m) \text{ for } a_{b,m} \in \mathbb{Z}.$$

Consider the submodule  $R$  of  $F$  generated by all elements of the following forms:

- $(b_1 + b_2, m) - (b_1, m) - (b_2, m)$
- $(b, m_1 + m_2) - (b, m_1) - (b, m_2)$
- $(ab, m) - (b, am)$

for all  $b, b_1, b_2 \in B, m, m_1, m_2 \in M$  and  $a \in A$ .

We set  $B \otimes_A M = F/R$  to be the quotient, and for  $b \in B$  and  $m \in M$ , we write  $b \otimes m$  for the image of  $(b, m)$  in  $F/R$ . We say that  $b \otimes m$  is a **simple tensor**.

The form of elements in  $R$  implies the following relations in  $B \otimes_A M$  (♥):

- $(b_1 + b_2) \otimes m = b_1 \otimes m + b_2 \otimes m.$
- $b \otimes (m_1 + m_2) = b \otimes m_1 + b \otimes m_2$
- $ab \otimes m = b \otimes am$

for all  $b, b_1, b_2 \in B, m, m_1, m_2 \in M$  and  $a \in A$ .

Notice that an element of  $B \otimes_A M$  has the form

$$\sum_{\text{finite}} b_i \otimes m_i.$$

We now show that  $B \otimes_A M$  is naturally a  $B$ -module. An element  $b_0 \in B$  determines a group homomorphism  $\lambda_{b_0} : F \rightarrow F$  given by

$$(b, m) \mapsto (b_0 \cdot b, m).$$

We must argue that this mapping satisfies  $\lambda_b R \subseteq R$ .

Notice that for all  $b, b_1, b_2 \in B, m, m_1, m_2 \in M$  and  $a \in A$  we have:

- $\lambda_{b_0}((b_1 + b_2, m) - (b_1, m) - (b_2, m)) = (b_0(b_1 + b_2), m) - (b_0 b_1, m) - (b_0 b_2, m)$   
 $= (b_0 b_1 = b_0 b_2, m) - (b_0 b_1, m) - (b_0 b_2, m) \in R.$
- $\lambda_{b_0}((b, m_1 + m_2) - (b, m_1) - (b, m_2)) = (b_0 b, m_1 + m_2) - (b_0 b, m_1) - (b_0 b, m_2) \in R.$
- $\lambda_{b_0}((ab, m) - (b, am)) = (ab_0 b, m) - (b_0 b, am) \in R.$

This shows that  $\lambda_{b_0}(R) \subseteq R$  so that  $\lambda_{b_0}$  induces a well defined mapping

$$\lambda_{b_0} : B \otimes_A M \rightarrow B \otimes_A M \text{ satisfying } \lambda_{b_0}(b \otimes m) = b_0 b \otimes m.$$

We must now confirm that the mapping  $\lambda : B \rightarrow \text{End}_{\mathbb{Z}}(B \otimes_A M)$  is a ring homomorphism. But this is a consequence of the relations ( $\heartsuit$ ). For example to see that  $\lambda$  is additive, notice that

$$\begin{aligned} (\lambda_{b_1} + \lambda_{b_2})(b \otimes m) &= \lambda_{b_1}(b \otimes m) + \lambda_{b_2}(b \otimes m) = b_1 b \otimes m + b_2 b \otimes m \\ &= (b_1 b + b_2 b) \otimes m = (b_1 + b_2)b \otimes m = \lambda_{b_1+b_2}(b \otimes m). \end{aligned}$$

This proves that  $\lambda_{b_1} + \lambda_{b_2} = \lambda_{b_1+b_2}$ .

One shows similarly that  $\lambda_{b_1} \circ \lambda_{b_2} = \lambda_{b_1 \cdot b_2}$ .

The ring homomorphism  $\lambda$  gives  $B \otimes_A M$  the structure of a  $B$ -module.

Now define  $\iota : M \rightarrow B \otimes_A M$  by the rule  $\iota(m) = 1 \otimes m$ . The relations ( $\heartsuit$ ) show that  $\iota$  is an  $A$ -module homomorphism.

Now let  $N$  be a  $B$ -module and  $\varphi : M \rightarrow N$  an  $A$ -module homomorphism. We must show find a unique  $B$ -module homomorphism  $\psi$  as in the statement.

We first argue that if  $\psi : B \otimes_A M \rightarrow N$  exists, it is unique. Indeed, for  $b \in B$  and  $m \in M$  we must have

$$\psi(b \otimes m) = \psi(b \cdot (1 \otimes m)) = b \cdot \psi(1 \otimes m) = b \cdot \psi(\iota(m)) = b \cdot \varphi(m)$$

for each  $b \in B$  and each  $m \in M$ .

It remains to argue that this rule determines a group homomorphism. Well, we can define

$$\tilde{\psi} : F \rightarrow N \text{ by } \tilde{\psi}((b, m)) = b \cdot \varphi(m).$$

We now argue that  $\tilde{\psi}(R) = 0$ .

- $\tilde{\psi}((b_1 + b_2, m) - (b_1, m) - (b_2, m)) = (b_1 + b_2) \cdot \varphi(m) - b_1 \cdot \varphi(m) - b_2 \cdot \varphi(m) = 0.$
- $\tilde{\psi}((b, m_1 + m_2) - (b, m_1) - (b, m_2)) = b \cdot \varphi(m_1 + m_2) - b \cdot \varphi(m_1) - b \cdot \varphi(m_2) = 0.$
- $\tilde{\psi}((ab, m) - (b, am)) = ab \cdot \varphi(m) - b \cdot \varphi(am) = 0.$

Thus  $\tilde{\psi}$  induces a group homomorphism  $\psi : B \otimes_A M \rightarrow N$  with the property that  $\psi(b \otimes m) = b \cdot \varphi(m)$  for  $b \in B$  and  $m \in M$ .

Since by construction  $\psi(bx) = b\psi(x)$  for  $b \in B$  and  $x \in B \otimes_A M$  we see that  $\psi$  is the required  $B$ -module homomorphism and the proof is complete. ■

**Proposition 9.2.2:** Let  $\varphi : M \rightarrow M'$  be a homomorphism of  $A$ -modules. There is a unique homomorphism of  $B$ -modules  $1 \otimes \varphi : B \otimes_A M \rightarrow B \otimes_A M'$  such that the following diagram of  $A$ -module homomorphisms is commutative:

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M' \\ \downarrow \iota_M & & \downarrow \iota_{M'} \\ B \otimes_A M & \xrightarrow{1 \otimes \varphi} & B \otimes_A M' \end{array}$$

*Proof:* Indeed, one defines  $1 \otimes \varphi$  to be the  $B$ -module homomorphism determined by the  $A$ -module homomorphism  $\iota_{M'} \circ \varphi : M \rightarrow B \otimes_A M'$ .

■

*Remark 9.2.3:* As in the proof of [Proposition 9.2.1](#), for  $x \in M$ , we write  $1 \otimes x = \iota(x) \in B \otimes_A M$ . We call these elements **pure tensors**.

The pure tensors satisfy the properties indicated in the proof of [Proposition 9.2.1](#). In fact, these properties can be deduced directly using the  $A$ -homomorphism  $\iota$  and the  $B$ -module structure on  $B \otimes_A M$  as follows:

Indeed, let  $b, b_1, b_2 \in B, m, m_1, m_2 \in M$ , and  $a \in A$ . Then

- $(b_1 + b_2) \otimes m = (b_1 + b_2) \cdot 1 \otimes m = (b_1 + b_2) \cdot \iota(m) = b_1 \cdot \iota(m) + b_2 \cdot \iota(m) = b_1 \otimes m + b_2 \otimes m.$
- $b \otimes (m_1 + m_2) = b \cdot 1 \otimes (m_1 + m_2) = b \cdot \iota(m_1 + m_2) = b \cdot (\iota(m_1) + \iota(m_2)) = b \cdot \iota(m_1) + b \cdot \iota(m_2) = b \otimes m_1 + b \otimes m_2.$
- $(ab) \otimes m = ab \cdot 1 \otimes m = ab \cdot \iota(m) = b \cdot \iota(am) = b \otimes am.$

*Remark 9.2.4:* The  $B$ -module  $B \otimes_A M$  is generated by all pure tensors  $1 \otimes m$  for  $m \in M$  – i.e. by the image of the map  $\iota$ .

This fact follows from the construction given in [Proposition 9.2.1](#), but can be seen just from the mapping property, as follows:

Let  $X$  be the  $B$ -submodule of  $B \otimes_A M$  generated by the image of  $\iota$ . We argue that  $X = B \otimes_A M$  or equivalently that  $B \otimes_A M / X = 0$ .

For this, it suffices to show that for any  $B$ -module  $Y$  and any  $B$ -module homomorphism

$$\gamma : B \otimes_A M \rightarrow Y \text{ such that } X \subseteq \ker \gamma$$

then  $\gamma = 0$ .

But notice that  $\gamma \circ \iota : M \rightarrow Y$  is an  $A$ -module homomorphism which according to [Proposition 9.2.1](#) induces a *unique*  $B$ -module mapping  $\psi : B \otimes_A M \rightarrow Y$  such that  $\psi \circ \iota = \gamma \circ \iota$ . But  $\psi = 0$  and  $\psi = \gamma$  both have this property; thus uniqueness implies that  $\gamma = 0$  as required.

**Proposition 9.2.5:** Suppose that  $F$  is a free  $A$ -module on  $\beta : C \rightarrow F$  for some set  $C$ . Then  $B \otimes_A F$  is a free  $B$ -module on  $\beta' = \iota \circ \beta : C \rightarrow B \otimes_A F$ .

*Proof:* Let  $X$  be a  $B$ -module and suppose that  $f : C \rightarrow X$  is a function. We must argue that

- ( $\clubsuit$ ): there is a unique  $B$ -module homomorphism  $\Phi : B \otimes_A F \rightarrow X$  such that  $\Phi \circ \beta' = f$ .

Since  $F$  is a free  $A$ -module, viewing  $X$  as an  $A$ -module, there is a unique  $A$ -module homomorphism  $\Psi : F \rightarrow X$  such that  $\Psi \circ \beta = f$ .

Now, using the mapping property of scalar extension [Proposition 9.2.1](#), there is a unique homomorphism of  $B$ -modules  $1 \otimes \Psi : B \otimes_A F \rightarrow X$  such that  $(1 \otimes \Psi) \circ \iota = \Psi$ .

Taking  $\Phi = 1 \otimes \Psi$ , we note that

$$\Phi \circ \beta' = (1 \otimes \Psi) \circ \iota \circ \beta = \Psi \circ \beta = f.$$

Thus  $\Phi$  is the unique  $B$ -module mapping satisfying ( $\clubsuit$ ). ■

**Corollary 9.2.6:**

- If  $F$  is a free  $A$ -module of rank  $n \in \mathbb{N}$ , then  $B \otimes_A F$  is a free  $B$ -module of rank  $n$ .
- If  $K \subseteq L$  is a field extension and if  $V$  is a  $K$ -vector space with  $\dim_K V = n \in \mathbb{N}$ , then  $\dim_L(L \otimes_K V) = n$ .

An  $A$ -module  $M$  is said to be a **torsion module** if  $\forall x \in M, \exists a \in A, a \neq 0$  such that  $a \cdot x = 0$ .

**Proposition 9.2.7:** Let  $A$  be an integral domain with field of fractions  $K$ . If  $M$  is a torsion  $A$ -module, then  $K \otimes_A M = 0$ .

*Proof:* We know that the  $K$ -vector space  $K \otimes_A M$  is spanned by elements of the form  $1 \otimes x$  for  $x \in M$  – see [Remark 9.2.4](#). So it suffices to show that  $1 \otimes x = 0$  in  $K \otimes_A M$  for each  $x \in M$ .

Fix  $x \in M$ . Since  $M$  is a torsion module, we may choose  $0 \neq a \in A$  such that  $a \cdot x = 0$ .

Now, note that  $1 = a/a$  in  $K$ . Thus using [Remark 9.2.3](#) we see that

$$1 \otimes x = (a/a) \otimes x = (1/a) \otimes a \cdot x = (1/a) \otimes 0 = 0.$$

■

### 9.3. Functors between categories

Let  $\mathcal{C}$  and  $\mathcal{D}$  be categories.

**Definition 9.3.1:** A **covariant functor**  $\mathcal{F} : \mathcal{C} \rightarrow \mathcal{D}$  is an assignment

$$X \mapsto \mathcal{F}(X) : \text{Ob}(\mathcal{C}) \rightarrow \text{Ob}(\mathcal{D})$$

together with an assignment

$$f \mapsto \mathcal{F}(f) : \text{Mor}(\mathcal{C}) \rightarrow \text{Mor}(\mathcal{D})$$

such that

- a. if  $f : X \rightarrow Y$  is a morphism in  $\mathcal{C}$  then  $\mathcal{F}(f) : \mathcal{F}(X) \rightarrow \mathcal{F}(Y)$  is a morphism in  $\mathcal{D}$ .
- b. if  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are morphisms in  $\mathcal{C}$  then  $\mathcal{F}(f \circ g) = \mathcal{F}(f) \circ \mathcal{F}(g)$ .

A **contravariant functor** is defined similarly, but a. and b. are replaced by the conditions

- a. if  $f : X \rightarrow Y$  is a morphism in  $\mathcal{C}$ , then  $\mathcal{F}(f) : \mathcal{F}(Y) \rightarrow \mathcal{F}(X)$  is a morphism in  $\mathcal{D}$ .
- b. if  $f : X \rightarrow Y$  and  $g : Y \rightarrow Z$  are morphisms in  $\mathcal{C}$  then  $\mathcal{F}(f \circ g) = \mathcal{F}(g) \circ \mathcal{F}(f)$ .

*Example 9.3.2:* Let  $A$  be a commutative ring and fix an  $A$ -module  $X$ .

- a. For any  $A$ -module  $M$ , notice that the space of  $A$ -linear maps  $\text{Hom}_A(X, M)$  is again an  $A$ -module.

Moreover, if  $\varphi : M_1 \rightarrow M_2$  is an  $A$ -linear map, then the assignment

$$(f \mapsto \varphi \circ f) : \text{Hom}_A(X, M_1) \rightarrow \text{Hom}_A(X, M_2)$$

is an  $A$ -module homomorphism.

Thus  $\text{Hom}_A(X, -)$  defines a covariant functor  $A\text{-mod} \rightarrow A\text{-mod}$ .

- b. For any  $A$ -module  $M$ , the space of  $A$ -linear maps  $\text{Hom}_A(M, X)$  is again an  $A$ -module.

Moreover, if  $\varphi : M_1 \rightarrow M_2$  is an  $A$ -linear map, then the assignment

$$(f \mapsto f \circ \varphi) : \text{Hom}_A(M_2, X) \rightarrow \text{Hom}_A(M_1, X)$$

is an  $A$ -module homomorphism.

Thus  $\text{Hom}_A(-, X)$  defines a contravariant functor  $A\text{-mod} \rightarrow A\text{-mod}$ .

**Proposition 9.3.3:** Let  $A$  be a commutative ring and let  $B$  be an  $A$ -algebra. The assignment  $M \mapsto B \otimes_A M$  determines a functor  $A\text{-mod} \rightarrow B\text{-mod}$  where to an  $A$ -module homomorphism  $\varphi : M \rightarrow M'$  the functor assigns the  $B$ -module homomorphism  $1 \otimes \varphi : B \otimes_A M \rightarrow B \otimes_A M'$ .

*Proof:* This follows from [Proposition 9.2.1](#) and [Proposition 9.2.2](#) ■

## 9.4. Bilinear mappings

Let  $A$  be a commutative ring and let  $M, N, P$  be  $A$ -modules.

**Definition 9.4.1:** A function  $\beta : M \times N \rightarrow P$  is said to be  **$A$ -bilinear** provided that:

- a. for fixed  $x_0 \in M$  the mapping  $y \mapsto \beta(x_0, y) : N \rightarrow P$  is  $A$ -linear, and that for fixed  $y_0 \in N$ , the mapping  $x \mapsto \beta(x, y_0) : M \rightarrow P$  is  $A$ -linear, and
- b. for  $\alpha \in A$  and  $(x, y) \in M \times N$ , we have  $\beta(\alpha x, y) = \beta(x, \alpha y)$ .



If  $I$  is an index set and if  $M_i$  is an  $A$ -module for each  $i \in I$ , one can of course give an analogous definition for the notion of an  $A$ -**multilinear** function  $\mu : \prod_{i \in I} M_i \rightarrow P$ .

*Example 9.4.2:*

- a. Let  $M \in \text{Mat}_{n \times m}(A)$  be an  $n \times m$  matrix with coefficients in  $A$ . Then  $M$  determines a bilinear mapping

$$\beta_M : A^n \times A^m \rightarrow A \text{ by the rule } \beta(x, y) = x^T \cdot M \cdot y$$

where we view  $x \in A^n$  and  $y \in A^m$  as *column vectors* (so that the transpose  $x^T$  is a row vector).

- b. Let  $B$  be an  $A$ -algebra, let  $M$  be an  $A$ -module,  $N$  a  $B$ -module and  $\varphi : M \rightarrow N$  an  $A$ -module homomorphism. Then the mapping

$$B \times M \rightarrow N \text{ given by } (b, m) \mapsto b \cdot \varphi(m)$$

is  $A$ -bilinear.

- c. Let  $F$  be a field and suppose that  $E$  is an extension field of  $F$ . Let  $V$  be an  $F$ -vector space, let  $W$  be an  $E$ -vector space and suppose that  $\varphi : V \rightarrow W$  is an  $F$ -linear mapping (we can of course view  $W$  as an  $F$ -vector space).

Consider the mapping

$$\sigma : E \times V \rightarrow W \text{ given by } (\alpha, v) \mapsto \alpha \cdot \varphi(v).$$

Then  $\sigma$  is  $F$ -bilinear.

## 9.5. The tensor product of modules

Let  $A$  be a commutative ring and let  $M$  and  $N$  be  $A$ -modules.

**Proposition 9.5.1:** There is an  $A$ -module  $M \otimes_A N$  together with an  $A$ -bilinear mapping

$$M \times N \rightarrow M \otimes_A N \text{ written } (m, n) \mapsto m \otimes n$$

with the following property:

given any  $A$ -module  $X$  and an  $A$ -bilinear mapping  $\beta : M \times N \rightarrow X$ , there is a unique homomorphism of  $A$ -modules  $\varphi : M \otimes_A N \rightarrow X$  such that for every  $m \in M$  and every  $n \in N$  we have  $\varphi(m \otimes n) = \beta(m, n)$ .

Moreover, this property uniquely determines  $M \otimes_A N$  up to an isomorphism of  $A$ -modules.

*Proof:* The proof is similar to that of [Proposition 9.2.1](#), with a few small differences.

In this case, uniqueness follows since  $M \otimes_A N$  is an initial object in the category whose objects are pairs  $(C, \beta)$  where  $C$  is an  $A$ -module and  $\beta : M \times N \rightarrow C$  is an  $A$ -bilinear map, where a morphism  $(C, \beta) \rightarrow (C', \beta')$  is an  $A$ -module homomorphism  $f : C \rightarrow C'$  such that  $f \circ \beta = \beta'$ .

For the construction, let  $F$  be the free  $A$ -module on the set  $M \times N$ ; thus  $F$  has an  $A$ -basis of elements  $(m, n)$  for  $m \in M$  and  $n \in N$ .

Let  $R \subseteq F$  be the  $A$ -submodule generated by elements of the form

- $(x_1 + x_2, y) - (x_1, y) - (x_2, y)$
- $(x, y_1 + y_2) - (x, y_1) - (x, y_2)$

- $(ax, y) - (x, ay)$

for all  $x, x_1, x_2 \in M, y, y_1, y_2 \in N$  and  $a \in A$ .

Now set  $M \otimes_A N = F/R$  and write  $x \otimes y$  for the image of the basis element  $(x, y)$  in  $F/R$ .

Thus  $M \otimes_A N$  is an  $A$ -module, and the relations  $R$  confirm that

- $(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$
- $x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$
- $ax \otimes y = x \otimes ay$

for all  $x, x_1, x_2 \in M, y, y_1, y_2 \in N$  and  $a \in A$ .

This proves that  $(x, y) \mapsto x \otimes y$  is an  $A$ -bilinear mapping.

Now suppose that  $X$  is any  $A$ -module and that  $\beta : M \times N \rightarrow C$  is an  $A$ -bilinear mapping.

Since  $F$  is a free module, there is an  $A$ -module homomorphism

$$\tilde{\varphi} : F \rightarrow X \text{ such that } \tilde{\varphi}((x, y)) = \beta(x, y) \text{ for } (x, y) \in M \times N.$$

One now confirms that  $\tilde{\varphi}(R) \subseteq R$ ; this is straightforward and we omit the details – compare with the proof of [Proposition 9.2.1](#).

Thus  $\tilde{\varphi}$  induces a mapping  $\varphi : M \otimes_A N \rightarrow X$  such that

$$\varphi(x \otimes y) = \beta(x, y) \text{ for } (x, y) \in M \times N$$

as required. ■

*Remark 9.5.2:* The tensor product  $M \otimes_A N$  comes equipped with a bilinear mapping  $M \times N \rightarrow M \otimes_A N$  which is written  $(m, n) \mapsto m \otimes n$ . Elements of the form  $m \otimes n$  for  $(m, n) \in M \times N$  are known as **pure tensors**.

As in [Remark 9.2.4](#), one knows that the pure tensors generate  $M \otimes_A N$  as an  $A$ -module.

**Proposition 9.5.3:** Let  $B$  be an  $A$ -algebra and  $M$  an  $A$ -module. When viewed as an  $A$ -module, the scalar extension  $B \otimes_A M$  is the tensor product of the  $A$ -modules  $B$  and  $M$ .

*Proof:* Just for this proof, let us write  $B \otimes'_A M$  for  $B$ -module obtained by extension of scalars as in [Proposition 9.2.1](#).

Then as in [Example 9.4.2](#), one knows that the  $A$ -module mapping  $\iota : M \rightarrow B \otimes'_A M$  determines an  $A$ -bilinear mapping

$$B \times M \rightarrow B \otimes'_A M \text{ given by } (b, m) \mapsto b \cdot \iota(m).$$

thus the mapping property of the tensor product – see [Proposition 9.5.1](#) – gives a unique  $A$ -module homomorphism  $\Phi : B \otimes_A M \rightarrow B \otimes'_A M$  such that  $\Phi(b \otimes m) = b \cdot \iota(m)$ .

On the other hand, the  $A$ -bilinear mapping  $B \times M \rightarrow B \otimes_A M$  determines a  $B$ -module structure on  $B \otimes_A M$ , and the mapping  $m \mapsto 1 \otimes m$  is an  $A$ -module homomorphism.

Now it follows from the mapping property of extension of scalars – see [Proposition 9.2.1](#) – that there is a unique  $B$ -module homomorphism  $\Psi : B \otimes'_A M \rightarrow B \otimes_A M$  such that  $\Psi \circ \iota$  is the mapping  $m \mapsto 1 \otimes m : M \rightarrow B \otimes_A M$ .

It is easy to see that  $\Phi$  and  $\Psi$  are inverse to one another. ■

**Proposition 9.5.4:** Let  $E$  and  $F$  be free  $A$ -modules with basis  $\beta : B \rightarrow E$  and  $\beta' : B' \rightarrow F$ . Then  $E \otimes_A F$  is a free  $A$ -module with basis

$$B \times B' \rightarrow E \otimes_A F \text{ given by } (b, b') \mapsto \beta(b) \otimes \beta'(b').$$

*Proof:* Let  $X$  be an  $A$ -module and let  $\tau : B \times B' \rightarrow X$  be any function.

Recall that we can view an element  $e \in E$  as a finitely supported function  $e : B \rightarrow A$  and an element of  $f \in F$  as a finitely supported function  $f : B' \rightarrow A$ .

The map  $\tau$  determines an  $A$ -bilinear form  $\gamma : E \times F \rightarrow X$  by the following rule:

$$\gamma(e, f) = \sum_{b \in B} \sum_{b' \in B'} e(b) f(b') \tau(b, b').$$

Now the universal property of the tensor product gives a unique  $A$ -module homomorphism

$$\Phi : E \otimes_A F \rightarrow X \text{ such that } \Phi(e \otimes f) = \gamma(e, f) \text{ for } e, f \in E \times F.$$

Now, an element  $b \in B$  determines a function  $\delta_b : B \rightarrow A$  with support  $\{b\}$  and value 1 at  $b$ . Moreover, under our identifications,  $\beta(b) = \delta_b$  and  $\beta'(b') = \delta_{b'}$ .

For  $b \in B$  and  $b' \in B'$  a small calculation shows that

$$\Phi(\beta(b) \otimes \beta'(b')) = \Phi(\delta_b \otimes \delta_{b'}) = \gamma(\delta_b, \delta_{b'}) = \tau(b, b').$$

Thus we have confirmed that  $E \otimes_A F$  satisfies the universal property of the free module as required. ■

**Corollary 9.5.5:** If  $E$  and  $F$  are free  $A$ -modules of rank  $e \in \mathbb{N}$  and  $f \in \mathbb{N}$  respectively, then  $E \otimes_A F$  is a free  $A$ -module of rank  $e \cdot f$ .

## 9.6. Tensor products and direct sums.

Let  $A$  be a commutative ring, and let  $M, N, P$  be  $A$ -modules.

**Proposition 9.6.1:** There is a natural isomorphism of  $A$ -modules

$$M \otimes_A (N \oplus P) \xrightarrow{\sim} (M \otimes_A N) \oplus (M \otimes_A P).$$

*Proof:* This will be in the next homework assignment! ■

## 9.7. Tensor products for PIDs.

**Proposition 9.7.1:** Let  $A$  be a PID and let  $p, q \in A$  be elements with  $\gcd(p, q) = 1$ . Then  $(A/pA) \otimes_A (A/qA) = 0$ .

*Proof:* Write  $1 = up + vq$  for  $u, v \in A$ . For any  $x \in A/pA$  and  $y \in A/qA$  we see that

$$\begin{aligned} x \otimes y &= 1 \cdot x \otimes y = (up + vq) \cdot x \otimes y \\ &= up \cdot x \otimes y + vq \cdot x \otimes y \\ &= u \cdot (p \cdot x) \otimes y + x \otimes (vqy) \\ &= 0 \otimes y + x \otimes 0 = 0 + 0 = 0. \end{aligned}$$

■

## 10. Week 11 [2025-11-10]

### 10.1. Hom and tensor functors

Let  $A$  be a commutative ring.

(The following is [Atiyah-MacDonald, Prop. 2.9])

**Proposition 10.1.1:**

a. A sequence of  $A$ -modules

$$(\clubsuit) \quad 0 \rightarrow X \xrightarrow{u} Y \xrightarrow{v} Z$$

is exact if and only if

$$(\spadesuit) \quad 0 \rightarrow \operatorname{Hom}_A(M, X) \xrightarrow{f \mapsto u \circ f} \operatorname{Hom}_A(M, Y) \xrightarrow{g \mapsto v \circ g} \operatorname{Hom}_A(M, Z)$$

is exact for every  $A$ -module  $M$ .

b. A sequence of  $A$ -modules

$$(\diamondsuit) \quad X \xrightarrow{u} Y \xrightarrow{v} Z \rightarrow 0$$

is exact if and only if

$$(\heartsuit) \quad 0 \rightarrow \operatorname{Hom}_A(Z, M) \xrightarrow{f \mapsto f \circ v} \operatorname{Hom}_A(Y, M) \xrightarrow{g \mapsto g \circ u} \operatorname{Hom}_A(X, M)$$

is exact for every  $A$ -module  $M$ .

*Proof:* For a. you proved  $\Rightarrow$  in the homework.

Let's prove b; the proof of a is completely similar.

We focus on  $\Leftarrow$  (the other implication is routine). Assume  $(\heartsuit)$  is exact for every  $A$ -module  $M$ . To prove  $(\diamondsuit)$  is exact, we first prove exactness at  $Z$  - i.e. we show that  $v$  is surjective.

Let  $C = Z / \operatorname{im} v$  be the cokernel of  $v$  and write  $\pi : Z \rightarrow C$  for the quotient mapping. Now notice that  $\pi \circ v = 0$ . Since  $(\heartsuit)$  is exact with  $M = C$ , we deduce that  $\pi = 0$ . But this shows that  $C = 0$  so indeed  $v$  is surjective.

We now consider exactness at  $Y$ . For any  $M$  we know by exactness that  $f \circ v \circ u = 0$  for every  $f : Z \rightarrow M$ . Taking  $M = Z$  and  $f = \operatorname{id}$  we deduce that  $v \circ u = 0$ .

This shows that  $\operatorname{im} u \subseteq \ker v$ . Let  $C = \operatorname{coker} u = Y / \operatorname{im} u$  and let  $\pi : Y \rightarrow C$  be the quotient map. Then  $\pi \circ u = 0$ . Now the exactness of  $(\heartsuit)$  with  $M = C$  shows that there is  $\gamma : Z \rightarrow C$  such that  $\pi = \gamma \circ v$ . Thus  $\operatorname{im} u = \ker \pi \supseteq \ker v$ . ■

**Proposition 10.1.2:** Let  $M, N, P$  be  $A$ -modules. There is a natural isomorphism

$$\Phi : \text{Hom}_A(M \otimes_A N, P) \xrightarrow{\sim} \text{Hom}_A(M, \text{Hom}_A(N, P))$$

such that

$$\Phi(f)(m)(n) = f(m \otimes n) \text{ for } f \in \text{Hom}_A(M \otimes_A N, P), m \in M, n \in N.$$

*Proof:* For  $f, m$  as in the statement, notice first that  $n \mapsto f(m \otimes n) : N \rightarrow P$  is  $A$ -linear, then that  $m \mapsto f(m \otimes -) : M \rightarrow \text{Hom}_A(N, P)$  is  $A$ -linear, and finally that  $\Phi$  is  $A$ -linear.

Now define  $\Psi : \text{Hom}_A(M, \text{Hom}_A(N, P)) \rightarrow \text{Hom}_A(M \otimes_A N, P)$  as follows.

An element  $g : \text{Hom}_A(M, \text{Hom}_A(N, P))$  determines an  $A$ -bilinear mapping

$$(m, n) \mapsto g(m)(n) : M \times N \rightarrow P,$$

and thus an  $A$ -module homomorphism

$$\Psi(g) : M \otimes_A N \rightarrow P \text{ for which } \Psi(g)(m \otimes n) = g(m)(n).$$

One readily checks that  $\Psi(g)$  is  $A$ -linear.

We claim that  $\Psi$  is the inverse of  $\Phi$ .

To see that  $\Psi \circ \Phi = \text{id}$ , let  $f \in \text{Hom}_A(M \otimes_A N, P)$ . Then for  $m, n \in M \times N$ , we see that

$$(\Psi \circ \Phi)(f)(m \otimes n) = \Phi(\Psi(f))(m \otimes n) = \Phi(f)(m)(n) = f(m \otimes n).$$

Thus  $(\Psi \circ \Phi)(f) = f$ .

To see that  $\Phi \circ \Psi = \text{id}$ , let  $g \in \text{Hom}_A(M, \text{Hom}_A(N, P))$ .

Then for  $m, n \in M \times N$  we see that

$$(\Phi \circ \Psi)(g)(m)(n) = \Phi(\Psi(g))(m)(n) = \Psi(g)(m \otimes n) = g(m)(n).$$

Thus  $(\Phi \circ \Psi)(g) = g$ . ■

**Proposition 10.1.3:**

- There is a functor  $A\text{-mod} \rightarrow A\text{-mod}$  given on objects by  $N \mapsto M \otimes_A N$  and for a morphism  $N_1 \xrightarrow{\varphi} N_2$  we have a morphism  $1 \otimes \varphi : M \otimes_A N_1 \rightarrow M \otimes_A N_2$ .
- This functor is *right exact*. i.e. if  $0 \rightarrow X \xrightarrow{i} Y \xrightarrow{\pi} Z \rightarrow 0$  is a short exact sequence of  $A$ -modules, then the sequence of  $A$ -modules and  $A$ -module homomorphisms

$$(\clubsuit) \quad M \otimes_A X \xrightarrow{1 \otimes i} M \otimes_A Y \xrightarrow{1 \otimes \pi} M \otimes_A Z \rightarrow 0$$

is exact.

*Proof:*

- An  $A$ -module homomorphism  $N_1 \xrightarrow{\varphi} N_2$ , determines an  $A$ -bilinear mapping

$$M \times N_1 \rightarrow M \otimes_A N_2 \text{ via } (m, n) \mapsto m \otimes \varphi(n)$$

and hence a unique  $A$ -module homomorphism

$$1 \otimes \varphi : M \otimes_A N_1 \rightarrow M \otimes_A N_2 \text{ with } (1 \otimes \varphi)(m \otimes n) = m \otimes \varphi(n).$$

It is straightforward to check that if  $N_2 \xrightarrow{\psi} N_3$  is another  $A$ -module homomorphism, then  $(1 \otimes \psi) \circ (1 \otimes \varphi) = 1 \otimes (\psi \circ \varphi)$ .

b. It follows from [Proposition 10.1.1\(b\)](#) that  $(\clubsuit)$  is exact if and only if

$$0 \rightarrow \text{Hom}_A(M \otimes_A X, N) \rightarrow \text{Hom}_A(M \otimes_A Y, N) \rightarrow (\text{Hom}_A(M \otimes_A Z, N))$$

is exact for every  $A$ -module  $N$ .

It now follows from [Proposition 10.1.2](#) that  $(\clubsuit)$  is exact if and only if  $(*)$

$$0 \rightarrow \text{Hom}_A(X, \text{Hom}_A(M, N)) \rightarrow \text{Hom}_A(Y, \text{Hom}_A(M, N)) \rightarrow \text{Hom}_A(Z, \text{Hom}_A(M, N))$$

is exact for every  $N$ .

Finally, [Proposition 10.1.1\(a\)](#) shows that indeed  $(*)$  is exact for every  $N$ , as required. ■

## 10.2. The Jordan-Hölder theorem and composition series for a module

Let  $A$  be a ring.

**Definition 10.2.1:** An  $A$ -module  $S$  is **simple** – or **irreducible** – if  $S$  is non-zero and the only  $A$ -submodules of  $S$  are  $0$  and  $S$ .

**Proposition 10.2.2:** Let  $A$  be a commutative ring and  $S$  be a simple  $A$ -module. Then  $S \simeq A/\mathfrak{m}$  for some maximal ideal  $\mathfrak{m}$  of  $A$ .

*Proof:* Let  $0 \neq x \in S$  and consider the  $A$ -module homomorphism  $\varphi : A \rightarrow S$  given by  $a \mapsto a \cdot x$ .

Since  $S$  is a simple  $A$ -module,  $S$  is equal to the image  $A \cdot x$  of this homomorphism. Let  $\mathfrak{m} = \ker \varphi$ . The first isomorphism Theorem shows that  $S = A \cdot x \simeq A/\mathfrak{m}$ .

There is a bijection between submodules of the quotient  $A/\mathfrak{m}$  and submodules  $I$  of  $A$  with  $\mathfrak{m} \subseteq I \subseteq A$ . Since  $A$  is commutative, submodules of  $A$  are precisely the ideals of  $A$ . Since the  $A$ -module  $S \simeq A/\mathfrak{m}$  is simple, it follows that  $\mathfrak{m}$  is a maximal ideal. ■

**Definition 10.2.3:** Let  $M$  be an  $A$ -module. A **composition series** of length  $d \in \mathbb{N}$  for  $M$  is a sequence of  $A$ -submodules  $0 = M_0 \subset M_1 \subset \dots \subset M_d = M$  such that  $M_i/M_{i-1}$  is a simple  $A$ -module for  $i = 1, \dots, d$ .

The **composition factors** of the composition series are the quotients  $M_i/M_{i-1}$  for  $1 \leq i \leq d$ .

The  $A$ -module  $M$  has **length**  $d \in \mathbb{N}$  if there is a composition series of length  $d$ .

We say that  $M$  has finite length if it has length  $d$  for some  $d \in \mathbb{N}$ .

*Example 10.2.4:* If  $K$  is a field, the length of the  $K$ -module  $V$  is equal to  $\dim_K V$ .

**Proposition 10.2.5:** Let  $M$  be an  $A$ -module of finite length. If if

$$0 = M_0 \subset M_1 \subset \dots \subset M_d = M \text{ and } 0 = N_0 \subset N_1 \subset \dots \subset N_e = M$$

are composition series, then  $d = e$  and for some permutation  $\pi$  of  $\{1, 2, \dots, d\}$  there are  $A$ -module isomorphisms  $M_i/M_{i-1} \simeq N_{\pi(i)}/N_{\pi(i)-1}$  for each  $i$ .

*Proof:* We proceed by induction on  $d$ . If  $d = 1$ , then  $M = M_1$  is a simple  $A$ -module. It is then easy to see that  $N_1 = M$  so that  $e = 1$  and the assertion about isomorphism classes is trivial.

Now suppose that  $d \geq 2$  and that the conclusion is valid for module of length  $< d$ . Note that  $M_1 \subseteq M_1 \cap N_e$ . Since  $M_1$  is simple, for  $0 < j \leq e$ , either  $M_1 \cap N_j = 0$  or  $M_1 \subseteq N_j$ . Thus, we may choose the smallest  $i$  with  $0 < i \leq e$  and  $M_1 \cap N_i \neq 0$ ; in particular  $M_1 \cap N_j = 0$  for  $0 \leq j < i$ .

We claim that  $N_i = N_{i-1} + M_1$ . Indeed,  $M_1 + N_{i-1}/N_{i-1}$  is a non-zero submodule of the simple module  $N_i/N_{i-1}$ . Note in particular that

$$(\clubsuit) \quad M_1 \simeq N_i/N_{i-1}.$$

We now have composition series for  $M/M_1$ ; namely

$$(\heartsuit) \quad 0 \subset M_2/M_1 \subset \dots \subset M_d/M_1 = M/M_1$$

of length  $d - 1$ , and

$$(\diamondsuit) \quad 0 \subset (N_2 + M_1)/M_1 \subset \dots \subset (N_{i-1} + M_1)/M_1 \subset N_{i+1}/M_1 \subset \dots \subset N_e/M_1 = M/M_1.$$

of length  $e - 1$ .

Note that the composition factors for  $(\heartsuit)$  are

$$(M_j/M_1)/(M_{j-1}/M_1) \simeq M_j/M_{j-1} \text{ for } j = 2, \dots, d$$

and the composition factors for  $(\diamondsuit)$  are

$$\begin{aligned} ((N_j + M_1)/M_1)/((N_{j-1} + M_1)/M_1) &\simeq (N_j + M_1)/(N_{j-1} + M_1) \\ &\simeq N_j/N_{j-1} \text{ for } j = 1, \dots, i-1 \end{aligned}$$

together with

$$(N_j/M_1)/(N_{j-1}/M_1) \simeq N_j/N_{j-1} \text{ for } j = i+1, \dots, e.$$

It now follows by induction that  $d - 1 = e - 1$  and that the multi-sets of isomorphism classes of composition factors for  $(\heartsuit)$  and  $(\diamondsuit)$  coincide: i.e.

$$\{[M_j/M_{j-1}] \mid j = 2, \dots, d\} = \{[N_j/N_{j-1}] \mid j = 1, \dots, i-1, i+1, \dots, d\}.$$

In view of  $(\clubsuit)$  this completes the proof. ■



### 10.3. Localization at a prime ideal, local rings and Nakayama's lemma

Let  $A$  be a commutative ring and  $S$  a multiplicatively closed subset. Suppose that  $I$  is an ideal of  $A$ . Then  $I^e = I \cdot A[S^{-1}]$  is an ideal of  $A[S^{-1}]$  and every element of  $I^e$  has the form  $a/b$  for  $a \in I$  and  $b \in S$ .

Write  $j : A \rightarrow A[S^{-1}]$  for the natural ring homomorphism  $j(a) = a/1$ . For an ideal  $I$  of  $A$ , write  $I^e = I \cdot A[S^{-1}]$  for the ideal  $I$  extended to  $A[S^{-1}]$ .

For an ideal  $J$  of  $A[S^{-1}]$  write  $J^c = j^{-1}(J)$  for the *contraction* of the ideal  $J$  to  $A$ .

**Proposition 10.3.1:**

- Let  $J$  be an ideal of  $A[S^{-1}]$ . Then  $J = (J^c)^e$ . In particular, every ideal of  $A[S^{-1}]$  is the extension of some ideal of  $A$  and distinct ideals of  $A[S^{-1}]$  have distinct contractions to  $A$ .
- Let  $I$  be an ideal of  $A$ .
  - We have  $(I^e)^c = \{a \in A \mid sa \in I \text{ for some } s \in S\}$ .
  - $I^e = A[S^{-1}]$  if and only if  $I \cap S \neq \emptyset$ .

*Proof:*

- We always have  $(J^c)^e \subseteq J$ . Suppose that  $a/s \in J$ . Then  $a/1 = s(a/s) \in J$  so that  $a \in j^{-1}(J) = J^c$ . This shows that  $a/1 \in (J^c)^e$  and hence also  $a/s \in (J^c)^e$ . This completes the proof of a.
- First, write  $I' = \{a \in A \mid sa \in I \text{ for some } s \in S\}$ . We first show that  $I' \subseteq (I^e)^c$ . Let  $a \in I'$  and choose  $s \in S$  with  $sa = b \in I$ . Then  $a/1 = b/s \in I^e$ , so that  $j(a) \in I^e \Rightarrow a \in (I^e)^c$ .

To show that  $(I^e)^c \subseteq I'$ , let  $a \in (I^e)^c$  so that  $a/1 = b/s$  for some  $b \in I$  and  $s \in S$ . Then  $x(as - b)$  for some  $x \in S$  so that  $xsa = xb \in I$ . Since  $xs \in S$  it follows that  $a \in I'$ , as required.

Finally, note that  $I^e = A[S^{-1}]$  if and only if  $1/1 \in I^e$  if and only if  $1 \in (I^e)^c = I'$ . Now the final assertions follows from the definition of  $I'$ .

■

**Proposition 10.3.2:** The assignment  $Q \mapsto Q^e$  determines a bijection

$$\{\text{prime ideals } Q \text{ of } A \text{ with } Q \cap S = \emptyset\} \xrightarrow{Q \mapsto Q^e} \{\text{prime ideals of } A[S^{-1}]\}$$

with inverse given by  $R \mapsto R^c = j^{-1}(R)$

*Proof:* We first note that if  $R$  is a prime ideal of  $A[S^{-1}]$  then  $R^c = j^{-1}(R)$  is a prime ideal of  $A$ . To see that  $S \cap R^c = \emptyset$ , suppose that  $s \in S \cap R^c$ . Then  $j(s) = s/1 \in R$ . But  $s/1$  is a unit in  $A[S^{-1}]$  whence  $R = A[S^{-1}]$ , contrary to the assumption that  $R$  is prime.

On the other hand, if  $Q$  is a prime ideal of  $A$  with  $Q \cap S = \emptyset$ , we show that  $Q^e = Q \cdot A_P$  is a prime ideal of  $A[S^{-1}]$ . For this, suppose that  $a/s, b/t \in A[S^{-1}]$  and that  $(a/s)(b/t) = (ab/st) \in Q^e$ . Thus, there is  $q \in Q$  and  $u \in S$  such that  $abuv = stqv$  for some  $v \in S$ . It is enough to argue

that  $a \in Q$  or  $b \in Q$ . Suppose that  $b \notin Q$ . Thus  $abuv \in Q$ . Since  $uv \in S$  we know that  $uv \notin Q$  and thus  $uvb \notin Q$ . Since  $Q$  is prime, conclude that  $a \in Q$ .

It remains to see that  $Q \mapsto Q^e$  and  $R \mapsto R^c$  are inverse to one another.

It follows from [Proposition 10.3.1\(b\)](#) that  $Q = (Q^e)^c$  for every prime ideal  $Q$  of  $A$  disjoint from  $S$ .

Moreover, for a prime ideal  $R$  of  $A[S^{-1}]$ , [Proposition 10.3.1\(b\)](#) implies that  $R = (R^c)^e$  as required. ■

Let  $P$  be a prime ideal of  $A$  and write  $S = A \setminus P$ . Since  $P$  is prime, the subset  $S$  is multiplicatively closed in  $A$ .

We write  $A_P$  for the ring  $A[S^{-1}]$  and call this ring the **localization** of  $A$  at  $P$ .

**Proposition 10.3.3:**

- a.  $A_P$  is a local ring with unique maximal ideal  $P^e = P \cdot A_P$ .
- b. There is a bijection between the prime ideals of  $A_P$  and the prime ideals of  $A$  contained in  $P$ .

*Proof:* This follows from [Proposition 10.3.2](#). ■

**Proposition 10.3.4:** Let  $A$  be a local ring with unique maximal ideal  $\mathfrak{m}$ .

- a. If  $a \in A$ , then  $a \notin \mathfrak{m} \Leftrightarrow a \in A^\times$ .
- b. For  $a \in \mathfrak{m}$ ,  $1 + a \in A^\times$ .

*Proof:* To prove a., let  $a \in A \setminus \mathfrak{m}$ . Then by uniqueness,  $a$  is contained in no maximal ideal of  $A$ . It follows that  $A \cdot a = A$  which shows that  $a \in A^\times$ . Conversely, if  $a \in A^\times$ , then  $A \cdot a = A \Rightarrow a \notin \mathfrak{m}$ .

Now b. follows since  $1 + a \notin \mathfrak{m}$  for any  $a \in \mathfrak{m}$ . ■

*Example 10.3.5:* Let  $A$  be a PID and let  $p \in A$  be a prime element. Write  $K$  for the field of fractions of  $A$ .

Then  $S = A \setminus pA$  consists of all  $a \in A$  for which  $a \notin pA$  – i.e. so that  $p \nmid a$ .

And  $A_{pA} = \{a/b \in K \mid a, b \in A, p \nmid b\}$ .

Moreover,  $pA_{pA} = \{a/b \in A_{pA} \mid a, b \in A, p \nmid b, p \mid a\}$ . In the notation of [Proposition 10.3.1](#),  $pA_{pA} = (pA)^e$ .

Important cases are  $\mathbb{Z}_{p\mathbb{Z}}$  for a prime number  $p$ , and  $K[T]_{pK[T]}$  for a field  $K$  and an irreducible polynomial  $p \in K[T]$ .

**Proposition 10.3.6 (Nakayama's Lemma):** Let  $A$  be a local ring with unique maximal ideal  $\mathfrak{m}$ . If  $M$  is a finitely generated  $A$ -module for which  $\mathfrak{m} \cdot M = M$  then  $M = 0$ .

*Proof:* Let  $n \in \mathbb{N}$  be the smallest number such that  $M$  is generated as  $A$ -module by  $n$  elements. To show that  $M = 0$ , we must show that  $n = 0$ .

We proceed by contradiction; thus we suppose  $n > 0$  and we let  $x_1, \dots, x_n \in M$  be a system of  $A$ -module generators.

Since  $\mathfrak{m} \cdot M = M$ , there are elements  $a_1, \dots, a_n \in \mathfrak{m}$  such that

$$x_n = \sum_{i=1}^n a_i x_i.$$

But then

$$(1 - a_n)x_n = \sum_{i=1}^{n-1} a_i x_i$$

and since  $1 - a_n \in A^\times$  by [Proposition 10.3.4](#), we see that

$$x_n = \sum_{i=1}^{n-1} a_i (1 - a_n)^{-1} x_i.$$

This shows that  $x_1, \dots, x_{n-1}$  generate  $M$  as  $A$ -module, contradicting the minimality of  $n$ . This contradiction proves that  $n = 0$ , and thus  $M = 0$  as required. ■

*Remark 10.3.7:* From Matsumura's text "Commutative Ring Theory" (p. 9):

"This Theorem is usually referred to as Nakayama's Lemma, but the late Professor Nakayama maintained that it should be referred to as a theorem of Krull and Azumaya ... we refer to it as NAK in this book."

**Proposition 10.3.8:** Let  $A$  be a local ring with unique maximal ideal  $\mathfrak{m}$  and write  $k = A/\mathfrak{m}$ . If  $M$  is an  $A$ -module, there is an isomorphism

$$k \otimes_A M \xrightarrow{\sim} M/\mathfrak{m}M.$$

In particular, if  $M$  is finitely generated then  $k \otimes_A M = 0 \Rightarrow M = 0$ .

*Proof:* We show that  $M/\mathfrak{m}M$  satisfies the universal property of the scalar extension [Proposition 9.2.1](#) for the  $A$ -module homomorphism  $\pi : M \rightarrow M/\mathfrak{m}M$  given by the quotient mapping.

Thus, let  $X$  be any  $k$ -module and let  $\varphi : M \rightarrow X$  be an  $A$ -module homomorphism. Then  $\varphi(\mathfrak{m}M) = 0$  since  $X$  is a  $k$ -module, so  $\varphi$  induces a unique  $A$ -module homomorphism

$$\psi : M/\mathfrak{m}M \rightarrow X \text{ such that } \psi \circ \pi = \varphi.$$

Now observe that  $\psi$  is a  $k$ -module homomorphism.

Thus  $M/\mathfrak{m}M$  satisfies the required mapping property, so  $M/\mathfrak{m}M \simeq k \otimes_A M$ .

The final assertion is now an immediate consequence of the Nakayama Lemma [Proposition 10.3.6](#). ■

**Proposition 10.3.9:** Let  $A$  be a local ring with unique maximal ideal  $\mathfrak{m}$  and set  $k = A/\mathfrak{m}$ . Suppose that  $\varphi : M_1 \rightarrow M_2$  is an  $A$ -module homomorphism and that  $M_2$  is finitely generated. Then  $\varphi$  is a surjective if and only if  $1 \otimes \varphi : k \otimes_A M_1 \rightarrow k \otimes_A M_2$  is surjective.

*Proof:* If  $\varphi$  is an surjective, it follows from [Proposition 10.1.3](#) that  $1 \otimes \varphi$  is surjective.

On the other hand, suppose that  $1 \otimes \varphi$  is an surjective. Write  $C = \text{coker } \varphi = M_2 / \text{im } \varphi$ . To see that  $\varphi$  is surjective, we must shows that  $C = 0$ .

Since  $M_2$  is a finitely generated  $A$ -module, evidently  $C$  is a finitely generated  $A$ -module. We claim that  $k \otimes_A C$  is the cokernel of  $1 \otimes \varphi$ .

Indeed, applying the functor  $k \otimes_A -$  to the short exact sequence

$$0 \rightarrow M_1 / \ker \varphi \rightarrow M_2 \rightarrow C \rightarrow 0$$

gives an exact sequence

$$k \otimes_A (M_1 / \ker \varphi) \xrightarrow{1 \otimes \bar{\varphi}} k \otimes_A M_2 \rightarrow k \otimes_A C \rightarrow 0$$

where  $\bar{\varphi} : M_1 / \ker \varphi \rightarrow M_2$  is the induced mapping. Note that  $\text{coker } \varphi = \text{coker } \bar{\varphi}$ .

Since  $1 \otimes \varphi$  is an isomorphism,  $k \otimes_A C = \text{coker } 1 \otimes \varphi = 0$  and since  $C$  is finitely generated, the Nakayama Lemma [Proposition 10.3.6](#) – see [Proposition 10.3.8](#) – implies that  $C = 0$  as required. ■

In the problem sets, we introduced the notion of a projective  $A$ -module; one characterization is that  $P$  is a projective  $A$ -module if and only if it is isomorphic to a direct summand of a free  $A$ -module.

**Proposition 10.3.10:** Let  $A$  be a local ring with unique maximal ideal  $\mathfrak{m}$  and let  $M$  be a finitely generated projective  $A$ -module. Then  $M$  is free as  $A$ -module.

*Proof:* Write  $k = A/\mathfrak{m}$ , and choose a basis  $x_1, \dots, x_n$  for the finite dimensional  $k$ -vector space  $k \otimes_A M = M/\mathfrak{m}M$ .

Choose elements  $y_1, \dots, y_n \in M$  such that  $x_i = y_i + \mathfrak{m}M = 1 \otimes y_i$  for each  $i$ .

Let  $F \simeq A^n$  be the free  $A$ -module on a basis  $b_1, \dots, b_n$  and let

$$\varphi : F \rightarrow M \text{ be given by } \varphi(b_i) = y_i.$$

Since  $\{x_i\}$  is a  $k$ -basis for  $k \otimes_A M$ ,  $1 \otimes \varphi$  is an isomorphism of  $k$ -vector spaces.

In particular,  $1 \otimes \varphi$  is surjective so [Proposition 10.3.9](#) shows that  $\varphi$  is surjective.

Write  $K = \ker \varphi$ . The Proposition will follow if we prove that  $K = 0$ .

Since  $M$  is projective, there is an  $A$ -module homomorphism  $\psi : M \rightarrow F$  such that  $\varphi \circ \psi = \text{id}_M$ .

Now

$$F \simeq K \oplus \text{im } \psi$$

so that  $K$  is the cokernel of  $\psi$ . Since  $1 \otimes \psi$  is the inverse to  $1 \otimes \varphi$ , it follows that  $1 \otimes \psi$  is surjective so that  $\text{coker } 1 \otimes \psi = 0$ . Thus [Proposition 10.3.9](#) implies that  $K = \ker \varphi = \text{coker } \psi = 0$ . ■

## 11. week 12 [2025-11-17]

### 11.1. Modules over a principal ideal domain

**Definition 11.1.1:** Let  $A$  be a ring and  $M$  an  $A$ -module. For  $m \in M$ , the **annihilator** of  $m$  is the left ideal  $\text{Ann}_A(m) = \{a \in A \mid am = 0\}$ .

**Definition 11.1.2:** Let  $A$  be an integral domain with field of fractions  $K$ . An  $A$ -module  $M$  is **torsion free** if  $\forall 0 \neq x \in M, \text{Ann}_A(x) = 0$ .

**Proposition 11.1.3:** Let  $A$  be an integral domain and let  $M$  be a finitely generated torsion free  $A$ -module. Then  $M$  is isomorphic to a submodule of a finitely generated free  $A$ -module.

*Proof:* Let  $v_1, v_2, \dots, v_n \in M$  be a set of  $A$ -module generators for  $M$ . After possibly re-ordering the  $v_i$ , we may arrange that for some  $s \leq n$ ,  $v_1, \dots, v_s$  are  $A$ -linearly independent and that for all  $s + 1 \leq i \leq n$ , the set  $\{v_1, \dots, v_s, v_i\}$  is linearly dependent. Note that

$$F = \langle v_1, \dots, v_s \rangle = \sum_{i=1}^s Av_i \simeq A^s$$

is a free  $A$ -module. So if  $s = n$ , the proof is complete.

Now suppose  $s < n$ . For each  $i$ ,  $s < i \leq n$  the linear dependence of  $\{v_1, \dots, v_s, v_i\}$  implies that there is  $a_i \in A$  for which  $a_i v_i \in F$ .

Now put  $a = \prod_{i=s+1}^n a_i \in A$ . Then  $aM \subseteq F$ .

Since  $M$  is torsion free, the mapping  $\lambda_a : M \rightarrow M$  given by multiplication with  $a$  is injective.

Thus  $\lambda_a$  defines an isomorphism from  $M$  to a submodule of the finitely generated free module  $F$ . ■

**Definition 11.1.4:** Let  $A$  be an integral domain, and let  $M$  be a finitely generated  $A$ -module. Then **rank** of  $M$  – denoted  $\text{rank}_A M$  – is the maximal number of  $A$ -linearly independent elements of  $M$ .

Note that the preceding definition is compatible with [Definition 7.1.1](#) (where it is defined only for free  $A$ -modules).

**Proposition 11.1.5:** Let  $A$  be an integral domain, let  $N$  be a finitely generated  $A$ -module, and let  $M \subseteq N$  an  $A$ -submodule. Assume there is an element  $x \in N$  with  $\text{Ann}_A x = 0$  such that  $N$  is the internal direct sum

$$N = A \cdot x \oplus M.$$

Then  $\text{rank}_A M < \text{rank}_A N$ .

*Proof:* Indeed, let  $y_1, \dots, y_n$  be  $A$ -linearly independent elements of  $M$ . Then the elements  $x, y_1, \dots, y_n$  are  $A$ -linearly independent in  $N$ . ■

Throughout the remainder of this section,  $A$  denotes a PID.

**Proposition 11.1.6:** A non-empty set of ideals of the PID  $A$  has a maximal element.

Compare [Dummit-and-Foote, Theorem 12.1.4]

*Proof:* Assume the contrary. Thus there is a non-empty set of ideals  $S$  in  $A$  with no maximal element. Then we can construct a strictly increasing chain of ideals in  $S$ :

$$I_1 \subset I_2 \subset \dots \subset I_n \subset \dots$$

contrary to [Proposition 8.3.1](#). ■

**Proposition 11.1.7:** Let  $F$  be a free  $A$ -module of finite rank  $n \in \mathbb{N}$  and let  $M \subseteq F$  be an  $A$ -submodule. Then:

- a.  $M$  is a free  $A$ -module, and
- b. there is a basis  $\{e_1, \dots, e_n\}$  of  $F$  and elements  $a_1, a_2, \dots, a_s \in A$  such that

$$\{a_1 e_1, a_2 e_2, \dots, a_s e_s\}$$

is a basis for  $M$  for some  $s \leq n$ .

*Proof:*

We proceed a. and b. by induction on the rank of  $M$ . Since  $M$  is torsion free, if  $M$  has rank 0 then  $M = 0$  and then a. and b. are trivial.

We now suppose that  $M$  has rank  $m > 0$  and that a. and b. are known to hold for a submodule of rank  $< m$  of a finitely generated free module.

For  $\varphi \in \text{Hom}_A(F, A)$ , the image  $\varphi(M)$  is an ideal of  $A$  which therefore has the form

$$\varphi(M) = \langle a_\varphi \rangle = a_\varphi \cdot A \text{ for some } a_\varphi \in A.$$

Write  $\Sigma = \{\varphi(M) \mid \varphi \in \text{Hom}_A(F, A)\}$ ; since  $0 \in \Sigma$ , this is a non-empty set of ideals. Thus by [Proposition 11.1.6](#) there is a maximal element of  $\Sigma$ .

Let  $\nu \in \text{Hom}_A(F, A)$  be chosen so that  $\nu(M) = \langle a_\nu \rangle$  is maximal in  $\Sigma$ . Write  $a = a_\nu \in A$  and choose  $x \in M$  so that  $\nu(x) = a$ .

We claim that  $a \neq 0$ . For this, choose a basis  $\mathcal{B} = \{b_1, b_2, \dots, b_n\}$  of the free  $A$ -module  $F$ , and for  $1 \leq i \leq n$  write

$$\pi_i : F \simeq A^n \rightarrow A$$

for the projection mapping onto the  $i$ th coordinate with respect to this basis.

Since  $M \neq 0$ , we know that for some  $1 \leq i \leq n$ ,  $\pi_i(M) \neq 0$ . Since  $\pi_i(M) \in \Sigma$  it follows that  $a \neq 0$ .

We now observe that

$$(\clubsuit) \quad \forall \varphi \in \text{Hom}_A(M, A) \text{ we have } a \mid \varphi(x).$$

Indeed, let  $d$  be a generator for the ideal of  $A$  generated by  $a$  and  $\varphi(x)$ . Thus

$$d = ra + s\varphi(x) \text{ for } r, s \in A.$$

Set  $\psi = r\nu + s\varphi \in \text{Hom}_A(F, A)$  and notice that

$$\psi(x) = r\nu(x) + s\varphi(x) = ra + s\varphi(x) = d$$

which shows that  $d \in \psi(M)$ , so that  $\langle d \rangle \subseteq \psi(M)$ .

On the other hand,  $d$  is a divisor of  $a$  so that  $\langle a \rangle \subseteq \langle d \rangle \subseteq \psi(M)$ .

In particular,  $\langle a \rangle \subseteq \psi(M)$  and by maximality of  $\langle a \rangle$  in  $\Sigma$  we conclude that

$$\langle a \rangle = \langle d \rangle = \psi(M).$$

Since  $\langle a \rangle = \langle d \rangle$  we see that

$$a \mid d \Rightarrow a \mid \varphi(x)$$

proving  $(\clubsuit)$ .

Recalling the basis  $\mathcal{B}$  we have chosen, we may apply  $(\clubsuit)$  to  $\pi_i \in \text{Hom}_A(F, A)$  for each  $i$ .

Thus for each  $i$  we have

$$\pi_i(x) = a \cdot \gamma_i \text{ for } \gamma_i \in A.$$

Now define

$$x_1 = \sum_{i=1}^n \gamma_i b_i \in F.$$

Since  $y = \sum_{i=1}^n \pi_i(y) b_i$  for any  $y \in F$  we have  $x = ax_1$ . Moreover, we have  $\nu(x_1) = 1$ .

We now claim that

$$(\heartsuit_1) \quad F = Ax_1 \oplus \ker \nu \text{ and } (\heartsuit_2) \quad M = Ax \oplus (M \cap \ker \nu).$$

We first prove that  $Ax_1 \cap \ker \nu = 0$ . Indeed, if  $ax_1 \in \ker \nu$  then  $0 = \nu(ax_1) = a\nu(x_1) = a \cdot 1 = a$  so  $a = 0$  and hence  $ax_1 = 0$ .

Since  $Ax \subset Ax_1$  it now follows that  $Ax \cap (M \cap \ker \nu) = 0$  as well.



We now show that  $F = Ax_1 + \ker \nu$ . let  $y \in F$  be arbitrary and notice that

$$y = \nu(y)x_1 + (y - \nu(y)x_1).$$

We must argue that  $y - \nu(y)x_1 \in \ker \nu$ . But

$$\nu(y - \nu(y)x_1) = \nu(y) - \nu(y)\nu(x_1) = \nu(y) - \nu(y) \cdot 1 = 0.$$

This prove ( $\heartsuit_1$ ).

Now suppose that  $y' \in M$ . According to ( $\clubsuit$ ),  $\nu(y')$  is divisible by  $a$ ; say  $\nu(y') = \gamma a$  for  $\gamma \in A$ .

Then the decomposition used above has the form

$$y' = \nu(y')x_1 + (y' - \nu(y')x_1) = \gamma ax_1 + (y' - \gamma ax_1)$$

where the first summand is in  $Ax_1$  and the second is in  $\ker \nu \cap M$ . This prove ( $\heartsuit_2$ ).

It now follows from [Proposition 11.1.5](#) that  $\text{rank}_A(\ker \nu \cap M) < \text{rank}_A M$ .

We now prove a. By the induction hypothesis,  $(\ker \nu) \cap M$  is a free  $A$ -module. It then follows from ( $\heartsuit_2$ ) that  $M = A \cdot x \oplus ((\ker \nu) \cap M)$  is a free  $A$ -module. Thus a. follows by induction.

Finally, we prove b. First note that  $\ker \nu$  is a submodule of the free module  $F$ , so by a. we know that  $\ker \nu$  is a free  $A$ -module. Since  $F = A \cdot x_1 \oplus \ker \nu$  we even know that  $\ker \nu$  has rank  $n - 1$ .

Now we may apply the induction hypothesis to the pair  $\ker \nu \cap M \subset \ker \nu$ . We find a basis  $\{b_1, b_2, \dots, b_{n-1}\}$  for  $\ker \nu$  and elements  $\alpha_1, \dots, \alpha_t \in A$  such that  $\{\alpha_1 b_1, \dots, \alpha_t b_t\}$  is a basis for  $\ker \nu \cap M$ , for some  $t \leq n - 1$ .

Now the proof is completed by observing that  $\{b_1, b_2, \dots, b_{n-1}, x_1\}$  is a basis of  $F$  and that  $\{\alpha_1 b_1, \dots, \alpha_t b_t, ax_1 = x\}$  is a basis for  $M$ .

■

**Corollary 11.1.8:** Let  $A$  be a PID. Any finitely generated torsion free  $A$ -module is free.

In particular, any finitely generated projective  $A$ -module is free.

*Proof:* Let  $M$  be a finitely generated torsion free  $A$ -module. According to [Proposition 11.1.3](#), there is an injective homomorphism  $0 \rightarrow M \rightarrow F$  where  $F$  is a finitely generated free  $A$ -module. Now the result follows from [Proposition 11.1.7](#)

■

**Definition 11.1.9:** For any ring  $A$ , an  $A$ -module  $N$  is **cyclic** if  $N$  is isomorphic to  $A/I$  for some left ideal  $I$  of  $A$ .

*Remark 11.1.10:* Equivalently,  $M$  is a cyclic  $A$ -module if  $M = A \cdot m$  for some  $m \in M$ . Note that there is an isomorphism of  $A$ -modules

$$A/\text{Ann}_A(m) \simeq A \cdot m.$$

**Corollary 11.1.11:** Let  $A$  be a PID and let  $M$  be a finitely generated  $A$ -module. Then  $M$  is a direct sum of cyclic  $A$ -modules.

*Proof:* Let  $x_1, \dots, x_n \in M$  be  $A$ -module generators, and let  $\varphi : A^n \rightarrow M$  be the surjective  $A$ -module homomorphism determined by  $e_i \mapsto x_i$  where the  $e_i$  are the standard basis for  $A^n$ .

Thus  $M$  is isomorphic as  $A$ -module to  $A^n / \ker \varphi$ . According to [Proposition 11.1.7](#), there is a basis (possibly not the standard basis!)  $b_1, \dots, b_n$  of  $A^n$  and elements  $a_1, \dots, a_n \in A$  such that  $\{a_i b_i \mid 1 \leq i \leq s\}$  is a basis for  $\ker \varphi$ , for some  $s \leq n$ . Then

$$M \simeq A^n / \ker \varphi \simeq A^{n-s} \oplus \left( \bigoplus_{i=1}^s A / A \cdot a_i \right).$$

This confirms that  $A$  is a direct sum of cyclic  $A$ -modules. ■

## 11.2. Exactness of tensoring with a localization of $A$

**Proposition 11.2.1:** Let  $A$  be a commutative ring, let  $S \subset A$  a multiplicatively closed subset, and let  $M$  be an  $A$ -module.

a. Every element  $\alpha \in A[S^{-1}] \otimes_A M$  may be written in the form

$$\alpha = (1/s) \otimes m \text{ for } s \in S, m \in M.$$

b. If  $\alpha = (1/s) \otimes m \in A[S^{-1}] \otimes_A M$ , then  $\alpha = 0$  if and only if  $tm = 0$  in  $M$  for some  $t \in S$ .

c. Let  $0 \rightarrow M \xrightarrow{\varphi} N$  be an injective  $A$ -module homomorphism. Then

$$0 \rightarrow A[S^{-1}] \otimes_A M \xrightarrow{\text{id} \otimes \varphi} A[S^{-1}] \otimes_A N$$

is injective.

*Proof:* Write  $B = A[S^{-1}]$  for the localization.

a. Since any tensor product is generated by pure tensors, we may write

$$\alpha = \sum_{i=1}^n (a_i / s_i) \otimes m_i \text{ for elements } a_i \in A, s_i \in S, m_i \in M.$$

Using bilinearity, we find

$$\alpha = \frac{1}{s_1 \dots s_n} \otimes \left( \sum_{i=1}^n S_i m_i \right) \text{ where } S_i = \prod_{j \neq i} s_j \text{ for } i = 1, \dots, n;$$

this proves the a.

b. First of all, since  $s$  is a unit in  $B$ , we may and will suppose that  $s = 1$ ; i.e. we suppose that  $\alpha = 1 \otimes m$ .

Choosing a set of  $A$ -module generators for  $M$ , we may find a free  $A$ -module and a surjective homomorphism  $\pi : F \rightarrow M$ . Thus

$0 \rightarrow K \xrightarrow{\iota} F \xrightarrow{\pi} M \rightarrow 0$  is exact, where  $K = \ker \pi$ .

It follows from [Proposition 10.1.3](#) that

$$B \otimes_A K \xrightarrow{\text{id} \otimes \iota} B \otimes_A F \xrightarrow{\text{id} \otimes \pi} B \otimes_A M \rightarrow 0$$

is exact. Now,  $m = \pi(x)$  for some  $x \in F$ . And if  $1 \otimes m = 0$ , then

$$1 \otimes x \in \ker(\text{id} \otimes \pi).$$

By exactness, there is  $\beta \in K \otimes_A F$  such that  $(\text{id} \otimes \iota)(\beta) = 1 \otimes x$ .

By a.,  $\beta$  has the form  $\beta = (1/t) \otimes y$  for  $y \in K$ . Then

$$1 \otimes x = (\text{id} \otimes \iota)(\beta) = (1/t) \otimes \iota(y).$$

This is an equality in the  $B$ -module  $B \otimes_A F$ . This  $B$ -module is free – see [Proposition 9.2.5](#) – and it follows that  $tx = \iota(y)$  in  $F$ . But then  $tm = \pi(tx) = \pi(\iota(y)) = 0$  which completes the proof.

- c. Now if  $\alpha \in \ker(\text{id} \otimes \varphi)$ , use a. to write  $\alpha = (1/s) \otimes m$  for  $s \in S$  and  $m \in M$ . Since  $s$  is a unit in  $B$ , we know that  $\alpha \in \ker(\text{id} \otimes \varphi)$  if and only if  $1 \otimes m \in \ker(\text{id} \otimes \varphi)$ ; thus we may suppose  $s = 1$ .

Now,

$$1 \otimes \varphi(m) = 0 \text{ in } B \otimes_A N$$

and so by b. there is  $t \in S$  for which  $0 = t \cdot \varphi(m) = \varphi(tm)$ . But then  $tm \in \ker \varphi \Rightarrow tm = 0$  in  $M$ . And this show that  $1 \otimes m = 0$  in  $B \otimes_A M$ .

■

*Remark 11.2.2:* With notation as in the proposition, it follows that the functor  $A[S^{-1}] \otimes_A -$  from  $A\text{-mod}$  to  $A[S^{-1}]\text{-mod}$  is exact.

**Proposition 11.2.3:** Let  $A$  be an integral domain with field of fractions  $K$ . If  $M$  is a finitely generated torsion free module with  $M \neq 0$ , then  $K \otimes_A M \neq 0$ .

*Proof:* According to [Proposition 11.1.3](#), there is an injective homomorphism of  $A$ -modules  $0 \rightarrow M \rightarrow F$  for a finitely generated free  $A$ -module  $F$ .

Writing  $F \simeq A^n$ , we consider the projections  $\pi_i : A^n \rightarrow A$  for  $i = 1, \dots, n$ . Since  $M$  identifies with a non-zero submodule of  $A^n$ , there is some  $j$  for which  $\pi_j(M) \neq 0$ .

Now note that  $\text{id} \otimes \pi_j : K \otimes_A F \rightarrow K \otimes_A A = K$  is a  $K$ -module mapping with

$$(\text{id} \otimes \pi_j)(K \otimes_A M) \neq 0.$$

In particular, this shows that  $K \otimes_A M \neq 0$ .

■

**Proposition 11.2.4:** Let  $A$  be an integral domain with field of fractions  $K$ , and let  $0 \neq I \subseteq A$ . Then  $K \otimes_A I \simeq K$ .

*Proof:*

According to [Proposition 11.2.1](#), applying the functor  $K \otimes_A -$  to the exact sequence

$$0 \rightarrow I \rightarrow A$$

gives an exact sequence

$$0 \rightarrow K \otimes_A I \rightarrow K \otimes_A A = K.$$

Now, [Proposition 11.2.3](#) shows that  $K \otimes_A I \neq 0$ . Thus  $K \otimes_A I$  is a non-zero subspace of the 1 dimensional vector space  $K$ ; thus  $K \otimes_A I$  is a 1 dimensional vector space over  $K$ . ■

*Example 11.2.5 (Modules which are not the direct sum of cyclic submodules):*

If  $A$  is not a PID, then a finitely generated  $A$ -module need not be a direct sum of cyclic modules.

Indeed, consider the two-variable polynomial ring  $A = F[T, S]$  for some field  $F$ . and let  $I$  be the non-principal ideal  $I = \langle T, S \rangle$

- a.  $I$  is a finitely generated torsion-free  $A$ -module which is not a direct sum of cyclic  $A$ -modules. Indeed, let  $K = F(T, S)$  be the field of fractions of  $F[T, S]$ .

According to [Proposition 11.2.4](#),  $K \otimes_A I$  is a one-dimensional  $K$ -vector space.

Since  $I$  is not a principal ideal,  $I$  is not a cyclic  $A$ -module.

Thus, to see that  $I$  is not a direct sum of cyclic modules, it suffices to show that  $I$  can't be written as a direct sum of 2 non-zero submodules.

Suppose that  $I = I_1 \oplus I_2$  is a direct sum of  $A$ -modules with  $I_j \neq 0$  for  $j = 1, 2$ . Then

$$K \otimes_A I = K \otimes_A I_1 \oplus K \otimes_A I_2$$

shows that  $\dim_K K \otimes_A I \geq 2$  which is a contradiction.

- b.  $M = I/I^3$  is a finitely generated, torsion  $A$ -module which is not a direct sum of cyclic  $A$ -modules.

Write  $t, s$  for the image in  $M$  of  $T$  resp.  $S$ . Notice that  $I/I^3$  has a basis as an  $F$ -vector space formed by the elements

$$t, s, t^2, t \cdot s, s^2.$$

In particular,  $\dim_F M = 5$ .

If  $\alpha \in M$  and  $\alpha + I \cdot M$  is non-zero in  $M/I \cdot M$ , then  $\dim_F A \cdot \alpha = 3$ .

Now, suppose that  $M = \bigoplus_{i=1}^n A \cdot \alpha_i$  for elements  $\alpha_i \in M$ . Then in particular,

$$M/I \cdot M = \sum_{i=1}^n A \cdot \alpha_i / I \cdot \alpha_i.$$

Since  $\dim_F M/I \cdot M = 2$  at least two of the  $\alpha_i$  – must have non-zero image in  $M/I \cdot M$ . After re-ordering, we may suppose that  $\alpha_1, \alpha_2 \in M \setminus I \cdot M$ .

But then  $A \cdot \alpha_1$  and  $A \cdot \alpha_2$  both have dimension 3, so that  $A \cdot \alpha_1 \cap A \cdot \alpha_2 \neq 0$ . This contradicts the assumption that  $M = \bigoplus_{i=1}^n A \cdot \alpha_i$  is direct. This contradiction shows that  $M$  is not the direct sum of cyclic submodules.

### 11.3. The torsion submodule

Let  $A$  be a commutative ring and let  $M$  be an  $A$ -module.

**Definition 11.3.1:**  $M$  is a **torsion**  $A$ -module if  $\forall x \in M, \text{Ann}_A(x) \neq 0$ . In other words, for every  $x \in M$ , there is an element  $0 \neq a \in A$  such that  $a \cdot x = 0$ .

**Proposition 11.3.2:** Suppose that  $A$  is an integral domain, let  $M$  be an  $A$ -module, and let

$$M_t = \{x \in M \mid \text{Ann}_A(x) \neq 0\}.$$

- a.  $M_t$  is a submodule of  $M$ , called the **torsion submodule**.
- b.  $M/M_t$  is torsion-free.

*Proof:*

- a. Note that  $0 \in M_t$ . Let  $x, y \in M_t$  and  $a \in A$ ; we must argue that  $ax + y \in M_t$ . By assumption, there are elements  $b_1, b_2 \in A \setminus \{0\}$  for which  $b_1x = 0$  and  $b_2y = 0$ . Since  $A$  is an integral domain,  $b_1b_2 \neq 0$ . Now note that

$$b_1b_2(ax + y) = a(b_1b_2x + b_1b_2y) = 0.$$

This proves that  $ax + y \in M_t$ .

- b. Suppose that  $\alpha \in M/M_t$  is non-zero. We must show that  $\text{Ann}_A \alpha = 0$ . Let  $a \in A$  with  $a \cdot \alpha = 0$ ; we must show that  $a = 0$ . Write  $\alpha = m + M_t$  for  $m \in M$ . Since  $a \cdot \alpha = 0$ , we know that  $a \cdot m \in M_t$ . Thus there is  $0 \neq b \in A$  such that  $ba \cdot m = 0$ ; in particular,

$$ba \in \text{Ann}_A(m).$$

Since  $\alpha \neq 0$ , we know that  $m \notin M_t$ . Thus  $\text{Ann}_A(m) = 0$ . This proves that  $ba = 0$ . Since  $b \neq 0$  and since  $A$  is an integral domain, conclude that  $a = 0$  as required. ■

**Proposition 11.3.3:** Let  $A$  be a PID, and let  $M$  be a finitely generated  $A$ -module. Then

$$M \simeq M_t \bigoplus F$$

for a finitely generated free  $A$ -module  $F$ .

*Proof:* Since  $M$  is finitely generated, it follows from [Proposition 11.3.2](#) that  $M/M_t$  is a finitely generated, torsion free  $A$ -module.

Now [Corollary 11.1.8](#) shows that  $M/M_t$  is a finitely generated free  $A$ -module.

Consider the short exact sequence of  $A$ -modules

$$0 \rightarrow M_t \rightarrow M \rightarrow M/M_t \rightarrow 0.$$

Since  $M/M_t$  is free, this sequence has a *section*  $\sigma : M/M_t \rightarrow M$  and it follows that

$$M \simeq M_t \oplus (M/M_t),$$

as required. ■

Notice that in order to describe all finitely generated modules over a PID  $A$ , by the previous Proposition, it only remains to describe the finitely generated *torsion*  $A$ -modules.

#### 11.4. Finitely generated torsion modules over a PID.

Let  $A$  be a commutative ring let  $M$  be an  $A$ -module.

**Definition 11.4.1:** The **annihilator** of  $M$  is the ideal

$$\text{Ann}_A(M) = \{a \in A \mid a \cdot M = 0\} = \{a \in A \mid \forall x \in M, a \cdot x = 0\} = \bigcap_{m \in M} \text{Ann}_A m.$$

Now let  $A$  be a PID and let  $M$  be a torsion  $A$ -module.

**Definition 11.4.2:** For a prime element  $p \in A$ , the  **$p$ -primary submodule** of  $M$  is the submodule

$$M_p = \{x \in M \mid p^a x = 0 \text{ for some } a \in \mathbb{N}\}.$$

Equivalently,

$$M_p = \{x \in M \mid \text{Ann}_A x = p^a \cdot A \text{ for some } a \in \mathbb{N}\}.$$

*Example 11.4.3:* Let  $p \in A$  be a prime in the PID  $A$  and consider the  $A$ -module

$$M = \bigoplus_{i=1}^{\infty} A/p^i A.$$

Then  $M$  is a torsion  $A$ -module which is not finitely generated,  $M = M_p$ , and  $\text{Ann}_A M = 0$ .

**Proposition 11.4.4:** Suppose that  $M$  is a finitely generated  $A$ -module such that  $M = M_p$  for some prime  $p \in A$ . Then  $\text{Ann}_A M = p^e \cdot A$  for some  $e \in \mathbb{N}$ , and

*Proof:* Write  $\text{Ann}_A M = d \cdot A$  for some  $d$ , and write  $d = p^e \cdot c$  for some  $c \in A$  with  $\gcd(p, c) = 1$ . We must argue that  $c$  is a unit in  $A$ , so suppose the contrary. In that case, we may choose  $q \neq p$  a prime divisor of  $c$

Now  $d/q \in A$  and  $d/q \notin \text{Ann}_A M$  so there is  $x \in M$  with

$$(*) \quad y = (d/q)x \neq 0$$

Note that  $qy = 0$ . Since  $y \in M$ , there is  $a \in \mathbb{N}$  for which  $p^a y = 0$ .

Now  $\gcd(q, p^a) = 1$  so there are elements  $u, v \in A$  with  $uq + vp^a = 1$ . But then

$$y = 1 \cdot y = (uq + vp^a)y = uqy + vp^a y = 0 + 0 = 0$$

contradicting  $(*)$ . This contradiction proves that  $c$  is a unit, and thus that  $\text{Ann}_A M = p^e \cdot A$ . ■

**Proposition 11.4.5:** Let  $A$  be a PID and let  $0 \neq d \in A$ . Write  $d = p_1^{e_1} p_2^{e_2} \dots p_n^{e_n}$  for pairwise non-associate primes  $p_1, \dots, p_n \in A$  and exponents  $e_1, e_2, \dots, e_n \in \mathbb{N}_{>0}$ .

a. There are elements  $x_1, \dots, x_n \in A$  such that for each  $i$ ,

$$j \neq i \Rightarrow x_i \equiv 0 \pmod{p_j^{e_j}} \text{ and } x_i \equiv 1 \pmod{p_i^{e_i}},$$

and such that  $\sum_{i=1}^n x_i \equiv 1 \pmod{d}$ .

b.  $p_i^{e_i} x_i \equiv 0 \pmod{d}$  for each  $i$ .

c. For each  $i, j \neq i \Rightarrow x_i x_j \equiv 0 \pmod{d}$  while  $x_i x_i \equiv x_i \pmod{d}$ .

*Proof:* By the Chinese Remainder Theorem [Theorem 7.5.5](#), there is an isomorphism

$$\Phi : A/dA \xrightarrow{\sim} \prod_{i=1}^n A/p_i^{e_i} A.$$

Thus for  $1 \leq i \leq n$  we may choose  $x_i \in A$  such that  $\Phi(x_i) \in \prod_{i=1}^n A/p_i^{e_i} A$  is the element whose  $j$ -th component is 0 if  $j \neq i$  and is 1 if  $j = i$ .

Now the elements  $x_1, \dots, x_n \in A$  have the required properties. ■

**Proposition 11.4.6:** Let  $A$  be a PID and let  $M$  be a finitely generated torsion  $A$ -module.

a. Then  $\text{Ann}_A(M) = A \cdot d$  for some  $0 \neq d \in A$ .

b. Write  $d = p_1^{e_1} p_2^{e_2} \dots p_n^{e_n}$  for pairwise non-associate primes  $p_1, \dots, p_n \in A$  and exponents  $e_1, e_2, \dots, e_n \in \mathbb{N}_{>0}$ .

Then  $M$  is the internal direct sum

$$M = \bigoplus_{i=1}^n M_{p_i}.$$

c. There are elements  $x_1, \dots, x_n \in A$  such that  $M_{p_i} = x_i \cdot M$  with  $x_i m = m$  for  $m \in M_{p_i}$  and  $x_i x_j M = 0$  for  $i \neq j$ .

d.  $\text{Ann}_A(M_{p_i}) = p_i^{e_i} \cdot A$ .

*Proof:* a. follows since  $A$  is a PID.

Now let  $x_1, \dots, x_n \in A$  be elements satisfying the conclusion of [Proposition 11.4.5](#) for  $d$ .

We first claim that  $x_i \cdot M \subseteq M_{p_i}$ . Well, we know that  $p_i^{e_i} x_i \equiv 0 \pmod{d}$ . Since  $dM = 0$ , we see for  $m \in M$  that

$$p_i^{e_i} x_i m = 0 \text{ so that } x_i m \in M_{p_i}.$$

Since  $\sum_i x_i \equiv 1 \pmod{d}$  and since  $dM = 0$ , we see that

$$m = 1 \cdot m = \left( \sum_{i=1}^n x_i \right) m = \sum_{i=1}^n x_i m \in \sum_{i=1}^n x_i \cdot M.$$

This shows that  $M = \sum_i x_i \cdot M$  and hence that  $M = \sum_i M_{p_i}$ .

Finally, we show that the sum  $M = \sum_{i=1}^n M_{p_i}$  is direct. Suppose that

$$0 = \sum_{i=1}^n m_i \text{ where } m_i \in M_{p_i} \text{ for } i = 1, 2, \dots, n.$$

We must argue that  $m_j = 0$  for each  $j$ .

For  $i \neq j$ , we have  $x_j \equiv 0 \pmod{p_i^{e_i}}$ . It follows that  $x_j M_{p_i} = 0$ .

Moreover, since  $x_j \equiv 1 \pmod{p_j^{e_j}}$  it follows that  $x_j m = m$  for all  $m \in M_{p_j}$ .

Thus we see that

$$0 = \sum_{i=1}^n m_i \Rightarrow 0 = x_j \sum_{i=1}^n m_i = \sum_{i=1}^n x_j m_i = m_j.$$

This completes the proof that  $M = \bigoplus M_{p_i}$  is a direct sum. Since

$$x_i M \subseteq M_{p_i} \text{ for each } i, \text{ and } M = \sum_i x_i M$$

it shows also that  $x_i M = M_{p_i}$  for each  $i$ . Thus we have proved b. and c.

For d., we first note that by [Proposition 11.4.4](#),  $\text{Ann}_A M_{p_i}$  is generated by  $p^f$  for some  $f \in \mathbb{N}$ . since  $M_{p_i} = x_i M$  and since the element  $x_i$  satisfies  $p_i^{e_i} x_i \equiv 0 \pmod{d}$ , we deduce that  $p_i^{e_i} \in \text{Ann}_A M_{p_i}$ . Thus  $f \leq e_i$ . To prove equality, we proceed as follows.

Since  $d/p_i \in A$  and  $\text{Ann}_A M = A \cdot d$ , there is an element  $z \in M$  with  $(d/p_i)z \neq 0$ .

Writing  $z = \sum_j x_j z$  we note that for  $j \neq i$  that since  $p_j^{e_j} x_j \equiv 0 \pmod{d}$ , we have  $0 = (d/p_i) x_j z$  so that

$$0 \neq (d/p_i)z = \sum_{i=1}^n (d/p_i) x_j z = (d/p_i) x_i z.$$

For  $j \neq i$ ,  $p_j$  acts invertibly on  $M_{p_i}$ ; it follows that  $p_i^{e_i-1} x_i z \neq 0$ . This proves that  $f = e_i$  so that  $\text{Ann}_A M_{p_i} = p_i^{e_i} \cdot A$  as required.

■



## 11.5. Describing $p$ -primary torsion modules over a PID.

Let  $A$  be a PID, let  $p \in A$  be a prime element. We write  $k_p = A/p \cdot A$  for the residue field for the prime ideal  $p \cdot A$ .

Let  $M$  be a finitely generated, torsion  $A$ -module with  $M = M_p$ . Our remaining task is to describe the possibilities for  $M$  up to isomorphism.

Recall [Definition 10.2.3](#) that we have defined the **composition length** – or just **length** – of a finitely generated torsion  $A$ -module.

**Definition 11.5.1:** For  $d \in \mathbb{N}$ , a partition of  $d$  is a sequence  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_e)$  with  $\lambda_i \in \mathbb{N}_{>0}$ ,  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_e$  and  $\sum_i \lambda_i = d$ .

We write  $|\lambda| = \sum_{i=1}^e \lambda_i = d$ .

For a partition  $\lambda$  of  $d \in \mathbb{N}$  and a prime  $p \in A$ , define the  $A$ -module  $M_{p,\lambda}$  to be

$$M_{p,\lambda} = \bigoplus_{i=1}^e A/p^{\lambda_i} \cdot A.$$

**Definition 11.5.2:** Let  $M$  be a finitely generated,  $p$ -primary torsion  $A$ -module. For  $i \in \mathbb{N}$ , the  **$p$ -torsion submodule** is the submodule

$$M[p^i] = \{x \in M \mid p^i x = 0\}.$$

Note that  $M[p] = M[p^1]$  can be viewed as a vector space over the field  $k_p$ .

**Proposition 11.5.3:** Let  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_e)$  be a partition of  $d$ .

- The  $A$ -module  $M_{p,\lambda}$  has length  $d$ .
- $M_{p,\lambda}[p]$  has dimension  $e$  as a  $k$ -vector space.
- $\text{Ann}_A M_{p,\lambda} = p^{\lambda_e} \cdot A$ .

Now we can state a solution to the following problem: for  $d \in \mathbb{N}$ , describe all  $p$ -primary torsion  $A$ -modules  $M$  with composition length  $d$ .

**Theorem 11.5.4:** Let  $M$  be a  $p$ -primary torsion  $A$ -module with composition length  $d$ . Then  $M$  is isomorphic as an  $A$ -module to  $M_{p,\lambda}$  for some partition  $\lambda$  of  $d$ .

Moreover, if  $\lambda$  and  $\mu$  are partitions of  $d$ , then  $M_{p,\lambda} \simeq M_{p,\mu}$  if and only if  $\lambda = \mu$ .

*Example 11.5.5:* Let  $p$  be a prime in  $A$ . Then [Theorem 11.5.4](#) shows that the following is a full list of  $p$ -primary  $A$ -modules having composition length 4.

- $A/p^4 A$  determined by the partition  $\lambda = (4)$ .
- $A/p^3 A \oplus A/pA$  determined by the partition  $\lambda = (1, 3)$ .

- $A/p^2 A \oplus A/p^2 A$  determined by the partition  $\lambda = (2, 2)$ .
- $A/p^2 A \oplus A/pA \oplus A/pA$  determined by the partition  $\lambda = (1, 1, 2)$ .
- $A/pA \oplus A/pA \oplus A/pA \oplus A/pA$  determined by the partition  $\lambda = (1, 1, 1, 1)$ .

*Proof of Theorem 11.5.4:* We have already seen – see [Corollary 11.1.11](#) – that  $M$  is a direct sum of cyclic submodules, say

$$M = A \cdot x_1 \oplus A \cdot x_2 \oplus \dots \oplus A \cdot x_m \text{ for } x_i \in M.$$

Since  $M$  is  $p$ -primary,  $\text{Ann}_A x_i = p^{e_i} \cdot A$  for some  $e_i \in \mathbb{N}$ . We may and will suppose that

$$e_1 \leq e_2 \leq \dots \leq e_n.$$

Thus

$$M = \bigoplus_{i=1}^n A/p^{e_i} \cdot A,$$

so that  $M \simeq M_{p,\lambda}$  where  $\lambda = (e_1, e_2, \dots, e_n)$ .

Now suppose that  $M_{p,\lambda} \simeq M_{p,\mu}$ . Since isomorphic modules have the same length, conclude from [Proposition 11.5.3](#) that there is some  $n \in \mathbb{N}$  for which  $\lambda$  and  $\mu$  are both partitions of  $n$ ; i.e.  $|\lambda| = n = |\mu|$ .

Since  $M_{p,\lambda}[p]$  is isomorphic to  $M_{p,\mu}[p]$ , it follows that

$$\dim_{k_p} M_{p,\lambda}[p] = \dim_{k_p} M_{p,\mu}[p].$$

Now we may conclude from [Proposition 11.5.3](#) that the partitions  $\lambda$  and  $\mu$  each have the same number of entries; thus  $\lambda = (\lambda_1, \dots, \lambda_e)$  and  $\mu = (\mu_1, \dots, \mu_e)$ .

We are prove that  $\lambda = \mu$  by induction on  $n = |\lambda| = |\mu|$ .

If  $n = 1$  then  $\lambda = (1) = \mu$  and the result is trivial. Now suppose that  $n > 1$  and that whenever  $\lambda', \mu'$  are partitions with  $|\lambda'| = |\mu'| < n$  for which  $M_{p,\lambda'} \simeq M_{p,\mu'}$  then  $\lambda' = \mu'$ .

Let

$$\begin{aligned} s &= \dim_{k_p} M_{p,\lambda}[p] - \dim_{k_p} (M_{p,\lambda}[p^2]/M_{p,\lambda}[p]) \\ &= \dim_{k_p} M_{p,\mu}[p] - \dim_{k_p} (M_{p,\mu}[p^2]/M_{p,\mu}[p]) \end{aligned}$$

and notice that

$$s = \#\{i \mid \lambda_i = 1\} = \#\{i \mid \mu_i = 1\}.$$

Thus  $M_{p,\lambda}/M_{p,\lambda}[p]$  is isomorphic to  $M_{p,\lambda'}$  where

$$\lambda' = (\lambda_{s+1} - 1, \lambda_{s+2} - 1, \dots, \lambda_e - 1).$$

Similarly,  $M_{p,\mu}/M_{p,\mu}[p]$  is isomorphic to  $M_{p,\mu'}$  where

$$\mu' = (\mu_{s+1} - 1, \mu_{s+2} - 1, \dots, \mu_e - 1).$$

Now,

$$M_{p,\lambda} \simeq M_{p,\mu} \Rightarrow M_{p,\lambda}/M_{p,\lambda}[p] \simeq M_{p,\mu}/M_{p,\mu}[p] \Rightarrow M_{p,\lambda'} \simeq M_{p,\mu'}$$

So by the induction hypothesis we have  $\lambda' = \mu'$ .

This shows that  $\lambda_j = \mu_j$  for  $s+1 \leq j \leq e$ . Now  $\lambda_j = \mu_j = 1$  for  $j \leq s$  so indeed  $\lambda = \mu$  as required. ■

## 12. week 13 [2025-12-01]

### 12.1. Recapitulating: finitely generated modules over a PID

Let  $A$  be a PID.

**Theorem 12.1.1:** Let  $M$  be a finitely generated  $A$ -module. Then

$$M \simeq F \oplus \bigoplus_{p \text{ prime}} M_p$$

where  $F$  is a free module of rank  $r \in \mathbb{N}$  and for a prime  $p$  of  $A$ ,  $M_p$  is the  $p$ -primary component of  $M$ .

For each prime  $p$  there is a partition  $\lambda(p) = (\lambda(p)_1 \leq \lambda(p)_2 \leq \dots \leq \lambda(p)_e)$  such that

$$M_p \simeq M_{p,\lambda(p)} = \bigoplus_{i=1}^e A/p^{\lambda(p)_i} \cdot A.$$

The natural number  $r = \text{rank } M$  and for each prime  $pA$  of  $A$  the partition  $\lambda(p)$  of  $M_p$  are uniquely determined by  $M$ .

The length of the torsion submodule  $M_t$  is given by  $\sum_p |\lambda(p)|$ .

### 12.2. Finitely generated modules over a PID via Invariant factors

Let  $A$  be a PID.

**Theorem 12.2.1:** Let  $M$  be a finitely generated  $A$ -module. Then there are elements  $a_1, a_2, \dots, a_d \in A$  such that  $a_i \mid a_{i+1}$  for each  $i$  and such that

$$M \simeq F \oplus \bigoplus_{i=1}^d A/a_i \cdot A.$$

Moreover, the list of ideals

$$a_1 \cdot A \supseteq a_2 \cdot A \supseteq \dots \supseteq a_d \cdot A$$

is uniquely determined by  $M$ .

*Proof:* You will deduce this result from [Theorem 12.1.1](#) for homework. ■

*Remark 12.2.2:* The elements  $a_1, \dots, a_d \in A$  as in the Theorem are known as the **elementary divisors** of  $A$ .

*Example 12.2.3:* Let  $p, q$  be non-associate primes of  $A$ , let  $\lambda = (1, 4)$  and  $\mu = (1, 2, 3)$ .

Consider the  $A$ -module  $M = M_{p,\lambda} \oplus M_{q,\mu}$ , so that  $M$  is an  $A$ -module of length

$$|\lambda| + |\mu| = 5 + 6 = 11.$$

The invariant factors are

$$a_1 = q, a_2 = p \cdot q^2, a_3 = p^4 \cdot q^3.$$

Indeed, evidently  $a_1 \mid a_2 \mid a_3$ . And using the Chinese Remainder Theorem – [Theorem 7.5.5](#) – we see that

$$\begin{aligned} & (A/q \cdot A) \oplus (A/p \cdot q^2 \cdot A) \oplus (A/p^4 \cdot q^3 \cdot A) \\ & \simeq (A/pA) \oplus (A/p^4A) \oplus (A/qA) \oplus (A/q^2A) \oplus (A/q^3A) \\ & \simeq M_{p,\lambda} \oplus M_{q,\mu} = M. \end{aligned}$$

### 12.3. Jordan decomposition of a matrix over a field

Let  $K$  be a field and let  $V$  be a finite dimension  $K$ -vector space. The group  $G = \text{GL}(V)$  of  $K$ -linear automorphisms of  $V$  acts by conjugation on the set  $\text{End}_K(V)$  of  $K$ -linear endomorphisms of  $V$ : for  $g \in \text{GL}(V)$  and  $X \in \text{End}_K(V)$  we have

$$g \cdot X \cdot g^{-1} \in \text{End}_K(V).$$

A choice of a basis  $\mathcal{B} = \{e_1, e_2, \dots, e_n\}$  of  $V$  determines an isomorphism  $\text{GL}(V) \simeq \text{GL}_n(K)$  where  $\text{GL}_n(K)$  denotes the group of  $n \times n$  matrices with non-zero determinant. And the choice of  $\mathcal{B}$  determines a (compatible!) isomorphism  $\text{End}_K(V) \simeq \text{Mat}_n(K)$ . Of course,  $\text{GL}_n(K)$  acts by conjugation on  $\text{Mat}_n(K)$ .

We are interested in finding representatives for the  $\text{GL}(V)$ -orbits in  $\text{End}_K(V)$ , or what amounts to the same thing, for the  $\text{GL}_n(K)$ -orbits in  $\text{Mat}_n(K)$ .

We record the following:

**Proposition 12.3.1:** Let  $\varphi \in \text{End}_K(V)$ . Then  $V$  becomes a module for the polynomial ring  $K[T]$  where the module structure is determined by the  $K$ -algebra mapping

$$F_\varphi : K[T] \rightarrow \text{End}_K(V) \text{ where } T \mapsto \varphi.$$

Let's write  $V_\varphi$  for  $V$  viewed as a  $K[T]$ -module using the endomorphism  $\varphi$  as in the Proposition.

**Proposition 12.3.2:** Let  $\varphi, \psi \in \text{End}_K(V)$ . Then  $\varphi = g\psi g^{-1}$  for some  $g \in \text{GL}(V)$  if and only if  $V_\psi \simeq V_\varphi$  as  $K[T]$ -modules.

*Proof:* ( $\Rightarrow$ ): Suppose that  $g\psi g^{-1} = \varphi$  for  $g \in \text{GL}(V)$ , and view  $g : V \rightarrow V$  as a vector space isomorphism. We argue that  $g : V_\psi \rightarrow V_\varphi$  is isomorphism of  $K[T]$ -modules. Since  $g$  is invertible, it is enough to see for each  $v \in V$  that

$$(*) \quad g(T \cdot v) = T \cdot gv.$$

But on  $V_\varphi$  the action of  $T$  is by the map  $\varphi$ , while on  $V_\psi$  that action is given by  $\psi$ . Thus  $(*)$  amounts to

$$(\clubsuit) \quad g(\psi(v)) = \varphi(g(v)).$$

But the assumption  $g\psi g^{-1} = \varphi$  implies that  $g \circ \psi = \varphi \circ g$  which confirms  $(\clubsuit)$ .

( $\Leftarrow$ ): Suppose that  $\lambda : V_\psi \rightarrow V_\varphi$  is an isomorphism of  $K[T]$ -modules.

Ignoring the  $K[T]$ -module structure for the moment,  $\lambda$  is in particular an invertible linear map  $V \rightarrow V$  – i.e.  $\lambda \in \text{GL}(V)$ . Since  $\lambda$  is a  $K[T]$ -module map,  $\lambda \circ \psi = \varphi \circ \lambda$ . This shows that

$$\lambda \circ \psi \circ \lambda^{-1} = \varphi$$

as required. ■

One invariant of this  $K[T]$ -module is of course the annihilator  $\text{Ann}_{K[T]}(V_\varphi)$ , or what amounts to the same thing, the kernel of the algebra homomorphism  $F_\varphi$ .

**Definition 12.3.3:** The **minimal polynomial** of  $\varphi \in \text{End}_K(V)$  over  $K$  is the monic generator  $m_\varphi \in K[T]$  of the ideal  $\text{Ann}_{K[T]}(V_\varphi) = \ker F_\varphi$  in  $K[T]$ .

*Remark 12.3.4:* More generally, for any finite dimensional  $K$ -algebra  $A$  and element  $a \in A$ , the minimal polynomial of  $a$  over  $K$  is the monic generator for the ideal  $\ker(T \mapsto a) \subseteq K[T]$ . Note that the assumption  $\dim_K A < \infty$  implies that this ideal must be non-zero!

**Definition 12.3.5:** The **characteristic polynomial** of  $\varphi \in \text{End}_K(V)$  over  $K$  is the polynomial

$$\chi_\varphi = \det(T \cdot \text{id} - \varphi) \in K[T].$$

We are going to use the classification of modules over a PID to get information about  $\varphi$ .

We first investigate the case where  $V_\varphi$  is a **cyclic** module for  $K[T]$ . This means that there is  $v \in V_\varphi$  such that  $V_\varphi = K[T] \cdot v$ , or equivalently that  $V_\varphi \simeq K[T]/f \cdot K[T]$  for some  $f \in K[T]$ .

If  $V_\varphi = K[T] \cdot v$ , write  $\text{Ann}_{K[T]}(v) = f \cdot A$ ; then

$$(g \mapsto g.v) : K[T]/f \cdot K[T] \xrightarrow{\sim} V_\varphi$$

is an isomorphism.

**Proposition 12.3.6:** Let  $V \simeq K[T]/f \cdot K[T]$  be a cyclic  $K[T]$ -module where

$$f = a_0 + a_1 T + a_2 T^2 + \dots + a_{d-1} T^{d-1} + T^d \in K[T]$$

with  $a_i \in K$ . Write  $\varphi : V \rightarrow V$  for the  $K$ -linear map given by multiplication with  $T$ , and let  $\mathcal{B} = \{b_0, b_1, \dots, b_{d-1}\}$  be the basis of  $V$  given by

$$b_i = T^i + f \cdot K[T], \quad i = 0, 1, \dots, d.$$

Then the matrix of  $\varphi$  in the basis  $\mathcal{B}$  is given by the so-called *companion matrix*

$$M_f = \begin{pmatrix} 0 & 0 & 0 & \dots & -a_0 \\ 1 & 0 & 0 & \dots & -a_1 \\ 0 & 1 & 0 & \dots & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -a_{d-1} \end{pmatrix}$$

*Proof sketch:* Just note for  $0 \leq i < d-1$ ,  $\varphi(b_i) = b_{i+1}$  while  $\varphi(b_{d-1}) = \sum_i -a_i b_i$ . ■

We can compute readily the characteristic polynomial of a companion matrix:

**Proposition 12.3.7:**  $\chi_{M_f} = \det(T \cdot \text{id} - M_f) = f$ .

*Proof:* We proceed by induction on  $\deg f = d$ . When  $d = 1$ ,  $f = T + a$  and  $M_f = (-a)$  so that

$$\chi_{M_f} = \det(T \text{id} - (-a)) = \det(T + a) = T + a.$$

Now suppose that  $d > 1$  and that the result is known for polynomials of degree  $< d$ .

Let  $f$  have degree  $d$ . Using column expansion on the first column, we see that

$$(\clubsuit) \quad \det(T \cdot \text{id} - M_f) = T \cdot \det M_1 - (-1) \cdot \det M_2$$

where

$$M_1 = \begin{pmatrix} T & 0 & 0 & \dots & a_1 \\ -1 & T & 0 & \dots & a_2 \\ 0 & -1 & T & \dots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & T + a_{d-1} \end{pmatrix} = T \cdot \text{id} - M_{(f-a_0)/T}$$

and

$$M_2 = \begin{pmatrix} 0 & 0 & 0 & \dots & a_0 \\ -1 & T & 0 & \dots & a_2 \\ 0 & -1 & T & \dots & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & T + a_{d-1} \end{pmatrix}.$$

Now,  $\det M_1 = (f - a_0)/T$  by the induction hypothesis.

And using row expansion on the first row we see that

$$\det M_2 = (-1)^n a_0 \cdot \det \begin{pmatrix} -1 & T & 0 & 0 & \dots & 0 \\ 0 & -1 & T & 0 & \dots & 0 \\ 0 & 0 & -1 & T & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & T \\ 0 & 0 & 0 & 0 & \dots & -1 \end{pmatrix} = (-1)^n a_0 \cdot (-1)^n = a_0.$$

Thus  $(\clubsuit)$  shows that

$$\det(T \cdot \text{id} - M_f) = T \cdot (f - a_0)/T + a_0 = f.$$

■

**Proposition 12.3.8:** Let  $V \simeq K[T]/f \cdot K[T]$  be a cyclic  $K[T]$ -module where  $f \in K[T]$  is monic polynomial. Write  $\varphi : V \rightarrow V$  for the  $K$ -linear map given by multiplication with  $T$ .

Then  $f = m_\varphi = \chi_\varphi$ .

*Proof:* We first argue that  $f = m_\varphi$ ; i.e. that  $\text{Ann}_{K[T]}(V) = f \cdot K[T]$ . Since  $V \simeq K[T]/f \cdot K[T]$ , we evidently have  $f \in \text{Ann}_{K[T]}(V)$ . On the other hand, if  $g \in \text{Ann}_{K[T]}(V)$ , then

$$g + f \cdot K[T] = g \cdot (1 + f \cdot K[T]) = 0$$

so that  $g \in f \cdot K[T]$  and hence  $f \mid g$ . This proves that  $f = m_\varphi$ .

To see that  $f = \chi_\varphi$ , let  $d = \deg f$ .

As in [Proposition 12.3.6](#), we know that the matrix of  $\varphi$  in the basis  $\mathcal{B} = \{b_0, b_1, \dots, b_{d-1}\}$  of  $V$  given by

$$b_i = T^i + f \cdot K[T], \quad i = 0, 1, \dots, d$$

is the companion matrix  $M_f$ . In [Proposition 12.3.7](#), we showed that  $\chi_\varphi = \chi_{M_f} = f$ .  $\blacksquare$

*Remark 12.3.9:* If  $V_\varphi \simeq M \oplus N$  for  $K[T]$ -submodules  $M$  and  $N$ , then  $\varphi : M \oplus N \rightarrow M \oplus N$  is given by  $\varphi(m, n) = (\varphi_1(m), \varphi_2(n))$  for linear maps  $\varphi_1 : M \rightarrow M$  and  $\varphi_2 : N \rightarrow N$ . Moreover, the characteristic polynomial of  $\varphi$  is given by

$$\chi_\varphi = \chi_{\varphi_1} \cdot \chi_{\varphi_2}$$

while the minimal polynomial is given by

$$m_\varphi = \text{lcm}(m_{\varphi_1}, m_{\varphi_2}).$$

**Theorem 12.3.10** (Cayley-Hamilton): For  $\varphi \in \text{End}_K(V)$ , the minimal polynomial of  $\varphi$  divides the characteristic polynomial of  $\varphi$ ; i.e.  $m_\varphi \mid \chi_\varphi$  in  $K[T]$ .

Moreover, let  $p$  be an irreducible factor of the characteristic polynomial  $\chi_\varphi$ . Then  $p$  divides the minimal polynomial  $m_\varphi$ .

*Proof:* There is a  $K[T]$ -module isomorphism

$$V_\varphi = \bigoplus_{p, \lambda} V_{p, \lambda}$$

where  $p$  ranges over irreducible polynomials and  $\lambda$  ranges over partitions, where

$$V_{p, \lambda} = \bigoplus_{i=1}^e K[T]/p^{\lambda_i} K[T].$$

In view of [Remark 12.3.9](#), in order to prove the first assertion, it is enough to suppose that  $V_\varphi$  is  $p$ -primary for some irreducible polynomial  $p$ .

When  $V_\varphi = V_{p, \lambda}$  for a partition  $\lambda = (\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_e)$  [Remark 12.3.9](#) together with [Proposition 12.3.8](#) shows that the characteristic polynomial of  $\varphi$  is given by  $\chi_\varphi = \prod_{i=1}^e p^{\lambda_i} = p^{|\lambda|}$  while the minimal polynomial is given by  $m_\varphi = p^{\lambda_e}$ ; since  $\lambda_e \leq |\lambda|$ , it is then clear that  $m_\varphi$  divides  $\chi_\varphi$ .

Now suppose that  $p$  is an irreducible factor of the characteristic polynomial  $\chi_\varphi$ . Then the preceding arguments show that the  $p$ -primary component  $V_{p, \lambda}$  of  $V_\varphi$  is non-zero, and it then follows from [Remark 12.3.9](#) that  $p^{\lambda_e}$  divides the minimal polynomial  $m_\varphi$ .  $\blacksquare$



**Definition 12.3.11:** A field  $K$  is said to be **algebraically closed** if every polynomial  $f \in K[T]$  has a root in  $K$ .

Suppose that  $K$  is an algebraically closed, and let  $f \in K[T]$ . It follows from the *remainder theorem* that if  $\alpha \in K$  is a root of  $f$ , then  $T - \alpha$  divides  $f$  in  $K[T]$ . Finding a root of the polynomial  $f/(T - \alpha)$  in  $K$  and iterating the argument, we see that if  $f$  has degree  $d$ , there are elements  $c, \alpha_1, \dots, \alpha_d \in K$  for which

$$f = c \prod_{i=1}^d (T - \alpha_i).$$

In other words, if  $F$  is algebraically closed, every polynomial can be written as a product of linear factors.

In particular,

**Proposition 12.3.12:** If  $K$  is algebraically closed, the irreducible polynomials are - up to associates - precisely the polynomials

$$T - \alpha \text{ for } \alpha \in K.$$

To describe the  $\text{GL}(V)$  conjugacy classes of endomorphisms of  $V$ , it is enough to describe the action of  $T$  on  $(T - \alpha)$ -primary  $K[T]$ -modules for each  $\alpha \in K$ . In view of [Remark 12.3.9](#), it just remains to describe the action of  $T$  on a cyclic  $(T - \alpha)$ -primary  $K[T]$ -module.

For a natural number  $d$  and  $\alpha \in K$  consider the lower triangular  $d \times d$  **Jordan block** matrix

$$J_{\alpha,d} = \begin{pmatrix} \alpha & 0 & 0 & \dots & 0 & 0 \\ 1 & \alpha & 0 & \dots & 0 & 0 \\ 0 & 1 & \alpha & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & \alpha & 0 \\ 0 & 0 & 0 & \dots & 1 & \alpha \end{pmatrix}$$

having eigenvalue  $\alpha$  with multiplicity  $d$ , and 1's on the “sub-diagonal”.

**Proposition 12.3.13:** Suppose that  $\alpha \in K$  and that  $f = (T - \alpha)^e$  for some  $e \in \mathbb{N}$ . Then  $\mathcal{B} = \{(T - \alpha)^i \mid 0 \leq i < e\}$  is a basis for  $V_f = K[T]/f \cdot K[T]$  and the matrix of  $\varphi$  in  $\mathcal{B}$  is given by  $J_{d,\alpha}$

**Theorem 12.3.14** (Jordan decomposition): Let  $\varphi \in \text{End}_K(V)$ . Then there is a basis  $\mathcal{B}$  of  $V$ , elements  $\alpha_1, \alpha_2, \alpha_e \in K$  and partitions  $\lambda^1, \lambda^2, \dots, \lambda^e$  such that the matrix of  $\varphi$  in  $\mathcal{B}$  is block-diagonal, with blocks given by the Jordan block matrices  $J_{\alpha_i, \lambda_j^i}$ .

## 12.4. Algebraic extensions of fields