Graduate Algebra

Definition 2.3.1

Contents

1. Week 1 [2025-09-03] ... 2
   1.1. Notations and recollections ... 2
   1.2. Groups ... 2
   1.3. Rings ... 3
   1.4. Fields ... 3
   1.5. Linear Algebra ... 3
   1.6. Cayley's Theorem ... 4
   1.7. A linear analogue of Cayley's Theorem. ... 5
2. Week 2 [2025-09-08] ... 7
   2.1. The Quotient of a set by an equivalence relation ... 7
   2.2. Sub-groups ... 9
   2.3. Group actions ... 11
   2.4. Quotients of groups ... 11
   2.5. Quotients of groups and orbits. ... 13
   2.6. The product of subgroups ... 14
   2.7. Lagrange's Theorem ... 16
3. Week 3 [2025-09-15] ... 18
   3.1. Normal subgroups ... 18
   3.2. Quotient groups ... 20
   3.3. First isomorphism theorem ... 22
   3.4. Groups acting on Groups ... 23
   3.5. p-groups ... 24
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. Integeral 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

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* $0, 1, 2, \ldots$
- $\mathbb{Z}$ 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
- $\mathcal{C}$ 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 $\cdot : G \times G \to G$ satifsying 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:

a. We usually write 1 or sometimes 1_G rather than e for the inverse element of G. + we usually write x^{-1} for the inverse of $x \in G$

b. 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.)

c. A group is abelian if $\forall a, b \in G, ab = ba$

d. Sometimes we write groups additively; in that case, 0 is the identity element and 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 $\varphi : G \to H$ is a group homomorphism provided that $\forall x, y \in G, \varphi(xy) = \varphi(x)\varphi(y)$.

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

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

Remark 1.2.4: If $\varphi : G \to H$ 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 $\varphi : G \to H$ be a group homomorphism. Then φ 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

$$\cdot : R \times R \to R$$

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:

a. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathcal{C}$ are commutative rings.
b. For a natural number $n > 1$, the ring $\mathrm{Mat}_n(\mathbb{Z})$ 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: $\mathbb{Q}, \mathbb{R}, \mathcal{C}$ 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

$$F \times V \to V$$

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 $\mathcal{C}$; 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 $f : I \to F$ 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 $I \to F$ has a basis $B = \{\delta_i \mid i \in I\}$, where

$$\delta_i : I \to F$$

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 $\varphi : V \to W$ 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 $\mathrm{GL}(V)$ is the set

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

$\mathrm{GL}(V)$ 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 $\mathrm{GL}(V)$ is isomorphic to the group GL_n of $n \times n$ matrices with non-zero determinant, where $n = \dim_F V$ and where the operation in GL_n is given by matrix multiplication.

1.6. Cayley's Theorem

Let Ω be any set. The set $S(\Omega)$ of all bijective functions $\psi : \Omega \to \Omega$ 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 Ω.

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

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

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

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

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

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

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

$$\left(\lambda_{g_1} \circ \lambda_{g_2}\right)h = \lambda_{g_1}\left(\lambda_{g_2}(h)\right) = \lambda_{g_1}(g_2 h) = g_1(g_2 h) = (g_1 g_2)h = \lambda_{g_1 g_2}(h)$$

as required.

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

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

It remains to see that λ is injective. If $g \in \ker \lambda$, then $\lambda_g = \mathrm{id}$. Thus $1 = \mathrm{id}(1) = \lambda_g(1) = g1 = g$. Thus $g = 1$ so that $\ker \lambda = \{1\}$ which confirms that λ 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 $\mathrm{GL}(V)$ 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 $f : G \to F$. Recall that V has a basis $B = \{\delta_g \mid g \in G.\}$

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

For $g \in G$, we may define an F-linear mapping $\lambda_g : V \to V$ by defining the value of λ_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 λ_g is F-linear, we have

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

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

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

as required.

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

$$\left(\lambda_{g_1} \circ \lambda_{g_2}\right)(\delta_h) = \lambda_{g_1}\left(\lambda_{g_2}\delta_h\right) = \lambda_{g_1}\left(\delta_{g_2 h}\right) = \delta_{g_1(g_2 h)} = \delta_{(g_1 g_2)h} = \lambda_{g_1 g_2}\delta_h$$

as required.

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

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

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

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

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

Then $\delta_1 = \mathrm{id}(\delta_1) = \lambda_x(\delta_1) = \delta_{x1} = \delta_x$. This implies that $1 = x$ so that indeed the kernel of λ is trivial and thus λ 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 $R : S \times S \to \text{Prop}$ – 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 $\underset{R}{\sim}$ to indicate this proposition; thus $R(a, b) \Leftrightarrow a \underset{R}{\sim} b$.

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: $\forall s_1, s_2, s_3 \in S, s_1 \sim s_2$ 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 $\pi : S \to \bar{S}$ 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 $f : S \to T$ such that $\forall a, b \in S, a \sim b \Rightarrow f(a) = f(b)$. Then there is a function $\bar{f} : S \to T$ 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

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

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

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

and let

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

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

$$T = \bar{S}_1 \quad \text{and} \quad f = \pi_1 : S \to \bar{S}_2.$$

Then the maps $\pi_2' : \bar{S}_1 \to \bar{S}_2$ and $\pi_{1'} : \bar{S}_2 \to \bar{S}_1$ are inverse to one another, and in particular π_1' and π_2' are bijections.

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

$$\begin{tikzcd} S_1 \arrow[r, "f = \pi_2"] \arrow[d, "\pi_1"] & T = \bar{S}_2 \\ \bar{S}_1 \arrow[ur, dashed, "\bar{f} = \bar{\pi}_2"'] & \end{tikzcd} \qquad \begin{tikzcd} S_2 \arrow[r, "f = \pi_1"] \arrow[d, "\pi_2"] & T = \bar{S}_1 \\ \bar{S}_2 \arrow[ur, dashed, "\bar{f} = \bar{\pi}_1"'] & \end{tikzcd}$$

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 π_1 and π_2 are surjective, we conclude that $\mathrm{id} = \bar{\pi}_1 \circ \bar{\pi}_2$ and $\mathrm{id} = \bar{\pi}_2 \circ \bar{\pi}_1$ 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 $\left(\bar{S}, \pi\right)$.

Proof: We consider the set $\bar{S} = \{[s] \mid s \in S\}$ of equivalence classes and the mapping $\pi : S \to \bar{S}$ 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 $f : S \to T$ 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 $\bar{f} : \bar{S} \to T$ with the property $f = \bar{f} \circ \pi$. If $\bar{f}$ exists, it must satisfy $\bar{f}([a]) = a$ 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 $\bar{f} : \bar{S} \to T$. 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 $\left(\bar{S}, \pi\right)$, 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!

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

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

Proposition 2.2.3: (Constructing subgroups)

a. 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.
b. 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:

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

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 $\langle a \rangle = \{a^n \mid n \in Z\}$ while if G is addtive then $\langle a \rangle = \{na \mid n \in Z\}$.

Proposition 2.2.5: If $\varphi : G \to H$ 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 $(*) \quad \forall a, b \in X, ab^{-1} \in X$.

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 $f : G \to H$ 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 Ω be a set. An action of G on Ω is a mapping

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

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 Ω is a G-space.

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

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

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

Proposition 2.3.4: The G-conjugacy relation on Ω is an equivalence relation.

Definition 2.3.5: Let Ω be a G-space, and let $\varphi : \Omega \to \Omega/\sim$ 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 Ω defined by

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

Thus the G-orbits are the equivalence classes for the relation $\underset{G}{\sim}$; see Remark 2.1.8.

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

Proposition 2.3.6: Ω is the disjoint union of the G-orbits in Ω.

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:

a. 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.
b. 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 $x \in G$,

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

c. 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 $(\pi, H \setminus G)$. 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

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

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

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

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

Consider the mapping $\pi_g : G \to G/H$ given by the rule $\pi_g(x) = \pi(gx)$. This mapping has the property that $x \underset{H}{\sim} x' \Rightarrow \pi_g(x) = \pi_{g(x')}$. Indeed,

$$x \underset{H}{\sim} x' \Rightarrow \exists h, x = x'h \Rightarrow \pi_g(x) = \pi(gx) = \pi(gx'h) = \pi(gx') = \pi_g(x')$$

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

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

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

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

Now suppose that $g_1, g_2 \in G$. To complete the proof, we must veryify 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 (♣) 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

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

Since π is surjective, (♥) follows at once. This completes the proof.

■

2.5. Quotients of groups and orbits.

Definition 2.5.1: Suppose that G acts on Ω_1 and on Ω_2. A morphism of G-sets $\varphi : \Omega_1 \to \Omega_2$ is a function φ 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 φ is an isomorphism (of G-sets) if there is a morphism of G-sets $\psi : \Omega_2 \to \Omega_1$ such that $\varphi \circ \psi = \mathrm{id}$ and $\psi \circ \varphi = \mathrm{id}$.

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

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

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

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

Proof: The rule $g \mapsto gx$ determines a surjective mapping $\alpha_x : G \to \mathcal{O}_x$. 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 $h \in H = \mathrm{Stab}_G(x) \Rightarrow hx = x$.

Thus Definition 2.1.2 shows that there is a mapping $\gamma : G/H \to \mathcal{O}_x$ such that $\gamma \circ \pi = \alpha_x$. To see that γ is a morhpism of G-sets, it suffices to show that (♣) $\quad \forall g, g'$ 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 (♣). This shows that γ is indeed a morphism of G-sets.

Since α_x is surjective and $\gamma \circ \pi = \alpha_x$, also γ is surjective. It only remains to see that γ is injective. Suppose that $z, z' \in G/H$ such that $\gamma(z) = \gamma(z')$. Since $\pi : G \to G/H$ 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 $g^{-1}g \in \mathrm{Stab}_G(x) = H$. Since the quotient mapping π is constant on H-orbits, $z = \pi(g) = \pi(gg^{-1}g') = \pi(g') = z'$. This shows that γ is injective and completes the proof. ■

Definition 2.5.4: The action of G on Ω is transitive if there is a single G-orbit on Ω. 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 $H = \mathrm{Stab}_G(x)$. 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 $G/H \to I$; i.e. $S(I)/S(I - \{x\}) \to I$.

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 $\mathrm{Inn}_g K \subseteq K$ (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 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 $\varphi : H \times K \to HK$ be the natural mapping given by $\varphi(h, k) = hk$.

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

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

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 γ is surjective. To see that γ 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 γ is injective and the proof of a. is complete.

Now, the mapping φ 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 φ 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, σ identifies with an n-cycle and τ identifies with a product of disjoint transpositions:

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

In particular, σ has order n and τ 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 σ has order n and τ 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 σ^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 $\sigma^k(n) \equiv n + k (\texttt{mod}\ n)$ while $\tau(n) = -n \equiv n (\texttt{mod}\ n)$, we conclude that $n + k \equiv n (\texttt{mod}\ n)$; thus $k \equiv 0 (\texttt{mod}\ n)$ 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 $\varphi : (G/H) \times H \to G$ 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 $\pi : G \to G/H$ is the quotient map.

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

To see that φ 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 φ 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 φ is injective. The remaining assertion follows from the definition of φ. ■

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

Proof sketch:

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

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

Indeed, our assumption means for every g that $\operatorname{Inn}_g N \subseteq N$. Thus $\operatorname{Inn}_g^{-1} \circ \operatorname{Inn}_g N \subseteq \operatorname{Inn}_g^{-1} N$ so that $N \subseteq \operatorname{Inn}_g^{-1} N$. Since this holds for every g, we find that $\operatorname{Inn}_g N \subseteq N \subseteq \operatorname{Inn}_g N$ 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 $\operatorname{Inn}_x H = H$ 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 $\forall t \in T, \forall x \in G, \operatorname{Inn}_x t \in H$.

Proof:

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

($\Leftarrow$): Write $N = \{g \in G \mid \operatorname{Inn}_g H = H\}$ 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 $\operatorname{Inn}_x H \subseteq H$. We know that Inn_x is a group homomorphism; see Proposition 3.1.1. It follows from Proposition 2.2.7

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

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

■

Proposition 3.1.5:

Let $N = \ker \varphi$ where $\varphi : G \to H$ 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

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

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

■

Example 3.1.6:

Consider the group $\mathrm{GL}_2(\mathbb{Q})$. 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 $\alpha : \mathbb{Q} \to \mathrm{GL}_2(\mathbb{Q})$ 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 $\mathrm{GL}_2(\mathbb{Q})$.

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 $h : \mathbb{Q}^\times \to \mathrm{GL}_2(\mathbb{Q})$ 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 $\mathrm{GL}_2(\mathbb{Q})$.

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 $\forall h \in H, \operatorname{Inn}_h U_{\mathbb{Q}} \subseteq U_{\mathbb{Q}}$ 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 $\operatorname{Inn}_{h(n)} U_{\mathbb{Z}} \subset U_{\mathbb{Z}}$ since for $x \in \mathbb{Z}$

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

Note however that $\operatorname{Inn}_{h(n)} U_{\mathbb{Z}}$ is a proper subset of $U_{\mathbb{Z}}$; indeed, identifying $U_{\mathbb{Z}}$ with $\mathbb{Z}$, the image subgroup $\operatorname{Inn}_{h(n)} U_{\mathbb{Z}}$ 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 $\operatorname{Inn}_{h(n)}^{-1} U_{\mathbb{Z}} = \operatorname{Inn}_{h(n^{-1})} U_{\mathbb{Z}} \not\subseteq U_{\mathbb{Z}}$; indeed

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

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 $\operatorname{Inn}_x K \subseteq K$ for all $x \in S$ even though H does not normalize K.

Of course, if we insist that $\operatorname{Inn}_x K = K$ 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

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

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

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 $\mu : G/N \times G/N \to G/N$ 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 $\pi_{G/N} : G \to G/N$ is a group homomorphism.

Proof: We first confirm that there is a mapping $\mu : G/N \times G/N \to G/N$ 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

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

given by

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

We claim that φ 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 $h, h' \in H$. Thus $g = g_1 h$ and $g' = g_1' h'$. Then

$$\varphi(g, g') = \pi_{G/N}(gg') = \pi_{G/N}(g_1 h \cdot g_1' h') = \pi_{G/N}(g_1 g_1' g_1'^{-1} h g_1' h') = \pi_{G/N}(g_1 g_1') = \varphi(g_1, g_1')$$

since N a normal subgroup

$\Rightarrow g_1'^{-1} h g_1' \in N \Rightarrow g_1'^{-1} h g_1' h' \in N$.

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

Next we confirm that there is an inversion mapping $G/N \to G/N$ that satifies b. For this, one just checks that the mapping $G \to G/N$ 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

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

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

It remains to cofirm 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 π 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 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 $\pi : A \to A/B$ 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 $\varphi : G \to H$ be a group homomorphism, and let $K = \ker \varphi$. Assume that φ is surjective. Then there is a unique isomorphism of groups $\overline{\varphi} : G/K \to H$ such that $\varphi = \overline{\varphi} \circ \pi$ where $\pi : G \to G/K$ is the quotient homomorphism.

Proof: We first observe that – provided it exists – $\overline{\varphi}$ 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

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

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

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

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

Finally, we check that φ 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

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

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 $\ker \overline{\varphi}$ is trivial so that $\overline{\varphi}$ 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

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

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 $G \to \mathrm{Aut}(H)$.

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 each group automorphism $\psi : H \to H$ 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 $\mathrm{Inn} : G \to \mathrm{Aut}(G)$.

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

$$Z = \left\{x \in G \mid \mathrm{Inn}_g\, x = x\right\} = \{x \in G \mid gx = xg \quad \forall g \in G.\}$$

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

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

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

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

Proposition 2.5.3 gives a bijection

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

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

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

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

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

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 $\pi : G \to G/Z$ 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:

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

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:

a. Any two Sylow p-subgroups are conjugate by an element of G.
b. Any p-subgroup of G is contained in a Sylow p-subgroup.
c. If r denotes the number of p-Sylow subgroups, then $r \equiv 1 (\bmod\ p)$ and $r \mid q$.
d.

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: $\binom{p^m q}{p^m} \equiv q (\bmod\ p)$.

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

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

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 $|E| \not\equiv 0 (\texttt{mod } p)$. Thus there must be some $X \in E$ for which the orbit $G \cdot X$ satisfies $|G \cdot X| \not\equiv 0 (\texttt{mod } p)$. If H is the *stabilizer* in G of X, there is of course a bijection between $G \cdot X$ and G/H. In particular,

$$|G/H| \not\equiv 0 (\texttt{mod } p).$$

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 $Q_1, Q_2, ..., Q_s$ 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

$$|\mathcal{O}_i| \equiv 0 (\operatorname{mod} p).$$

Finally, the number r of Sylow p-subgroups satisfies

$$r = \sum_{i=1}^{s} |\mathcal{O}_i| = 1 + \sum_{i=2}^{s} |\mathcal{O}_i| \equiv 1 (\operatorname{mod} p)$$

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

Proof: Indeed, P is normal if and only if $\operatorname{Inn}_g P = P$. 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 $\varphi : G \to G$, we have $\varphi(H) = H$.

A characteristic subgroup H is always normal, since H is invariant under all the inner automorphisms 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 φ 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 $\mathrm{Syl}_p = \mathrm{Syl}_p\, G)$ for the set of Sylow p-subgroups. Recall that

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

.

Proposition 4.2.4:

a. If $n_p = 1$ for some prime p, then $P \in \mathrm{Syl}_p(G)$ is characteristic – and in particular normal – in G.

b. Suppose that $H \subseteq G$ is a normal subgroup, and suppose that $P \in \mathrm{Syl}_p(H)$ is a normal p-Sylow subgroup of H for some prime p. Then P is a normal subgroup of G.

Proof:

a. For any automorphism φ 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.

b. For $g \in G$, Inn_g determines an automorphism of H. Thus by (a), $\mathrm{Inn}_g\, P = P$ which shows that P is normal in G.

■

Proposition 4.2.5:

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

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

Proof:

a. By Theorem 4.1.1 we have $n_q \equiv 1 (\texttt{mod}\ q)$ and $n_q \mid p$. Since $q > p$ it follows that $n_q = 1$.
b. Again we have $n_p \equiv 1 (\texttt{mod}\ p)$ 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 $q \equiv 1 (\texttt{mod}\ p)$ 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 $P, Q \in \mathrm{Syl}_p(G)$ with $P \neq Q$. Since $|P| = |Q| = p$ is prime we know that $P \cap Q = \{1\}$. Thus

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

■

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 $n_3 \equiv 1 (\texttt{mod}\ 3)$ 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 $\Omega = \mathrm{Syl}_3$ of 3-Sylow subgroups, and $|\Omega| = 4$. This action determines a group homomorphism

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

The kernel K of φ 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 φ 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 φ 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 $P \in \mathrm{Syl}_5$ and $Q \in \mathrm{Syl}_3$.

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 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 $P \in \mathrm{Syl}_5(G)$. 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 $Q \in \mathrm{Syl}_5(H)$, 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 $n_5\left(\overline{G}\right) = 1$.

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

If $\pi : G \to \overline{G} = G/H$ 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 leftby 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

a ideal.

a. More generally, if I_x is an ideal of R for each x in the index set X, then $\cap_{x \in X} I_x$ is an ideal of R.

Proof: Since b ⇒ 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 $f : R \to S$ is a ring homomorphism provided that is a homomorphisms 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 $f : R \to S$ 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 $\pi : R \to R/I$ 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 $\pi : R \to R/I$ 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 $\pi : R \to R/I$ 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 $\pi : R \to R/I$ 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 $\varphi : R \to S$ be a surjective ring homomorphism. Recall that φ induces an isomorphism of additive groups $\overline{\varphi} : R/I \to S$ for which $\overline{\varphi}(a + I) = \varphi(a)$ for $a \in R$. Then $\overline{\varphi}$ 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 $\mathrm{Mat}_3(R)$ which is not commutative if $1_R \neq 0_R$. The mapping

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

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

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

  for domain and codomain.

  Denote by $\mathrm{Mor}(X, Y) = \mathrm{Mor}_{\mathcal{C}}(X, Y)$ the subclass of $\mathrm{Mor}(\mathcal{C})$ consisting of morphisms f in $\mathrm{Mor}(\mathcal{C})$ with $\mathrm{dom}(f) = X$ and $\mathrm{codom}(f) = Y$.
- for every three objects X, Y, Z there is a binary operation

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

This data is required to satisfy:

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

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

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

Remark 5.2.2:

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

Example 5.2.3: Here are some examples of categories.

a. The category Set of all sets, with morphisms given by functions.
b. The category Grp of all groups, with morphisms given by group homomorphisms.
c. The category Ab of all abelian groups, with morphisms given by group homomorphisms.
d. The category Top of topological spaces, with morphisms given by continuous functions.
e. 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 $\mathrm{Mor}(I, X)$.

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 $\mathrm{Mor}(X, I)$.

Example 5.2.5:

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

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 $I \to I'$.

b. If T, T' are terminal objects, there is a unique isomorphism $T \to T'$.

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

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

Similarly, $g \circ f$ is a morphism in $\mathrm{Mor}(I, I)$. Since I is initial, id_I is the unique morphism in $\mathrm{Mor}(I, i)$ and we conclude that $\mathrm{id}_I = g \circ f$. Thus we have proved that $f : I \to I'$ 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 $\pi_i : X \to X_i$ for $i \in I$ such that given any object Y of $\mathcal{C}$ toegether with morphisms $f_i : Y \to X_i$ there is a unique morphism $f : Y \to X$ such that $f_i = \pi_i \circ f$ for each $i \in I$; i.e. the diagram

$$\begin{tikzcd} Y \arrow[rr, "f"] \arrow[dr, "f_i"'] & & X \arrow[dl, "\pi_i"] \\ & X_i & \end{tikzcd}$$

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, π_i) of the X_i exists in $\mathcal{C}$ where $\pi_i : X \to X_i$ 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 $f : X \to X'$ 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 D is an object X of $\mathcal{C}$ together with morphisms $h_i : X \to X_i$ for each $i \in I$.

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

$$\begin{tikzcd}
X \arrow[rr, "\varphi"] \arrow[dr, "h_i"'] & & X' \arrow[dl, "h_i'"] \\
& X_i &
\end{tikzcd}$$

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

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

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 $\gamma_i : X_i \to X$ for $i \in I$ such that given any object Y of $\mathcal{C}$ toegether with morphisms $g_i : X_i \to Y$ there is a unique morphism $g : X \to Y$ such that $g_i = g \circ \gamma_i$ for each $i \in I$; i.e. the diagram

$$\begin{tikzcd} Y & & X \arrow[ll, "g"'] \\ & X_i \arrow[ul, "g_i"] \arrow[ur, "\gamma_i"'] & \end{tikzcd}$$

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, γ_i) of the X_i exists in $\mathcal{C}$ where $\gamma_i : X_i \to X$, it is unique up to a unique isomorphism.

In other words, if (X', γ_i') is a second co-product in $\mathcal{C}$, there is a unique isomorphism $f : X \to X'$ 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

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

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 $R \times M \to M$ 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 $R \to \operatorname{End}_{\mathbb{Z}}(M)$, where $\operatorname{End}_{\mathbb{Z}}(M)$ is the ring of additive endomorphisms of M.

Definition 5.3.5:

If M and N are left R-modules, a function $\varphi : M \to R$ is a homomorphism of R-modules provided that
- φ 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 $\varphi : M \to N$ is the same as a linear map of vector spaces.

b. If A and B are abelian groups, a function $\varphi : A \to B$ 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-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 $f : I \to \bigcup_{i \in I} M_i$ with the property that $f(j) \in M_j$ for $j \in I$.

a. $\bigoplus_{i \in I} M_i$ is an R-module with pointwise addition and scalar multiplication.

b. Define $\iota_j : M_j \to \bigoplus_{i \in I} M_i$ 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 ι_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 $\iota_j : M_j \to M$ $j \in I$ as in Proposition/Definition 5.4.1.

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

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

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

$$\begin{tikzcd} M = \bigoplus_{i \in I} M_i \arrow[rr, "F"] & & N \\ & M_j \arrow[ul, "\iota_j"'] \arrow[ur, "f_j"] & \end{tikzcd}$$

Proof: Let N and $f_j : M_j \to N$ 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 $\beta : B \to F$ be a fuction. Then F is a free left R-module on β provided that for any R-module X and any function $j : B \to X$, there is a unique R-module homomorphism $\varphi : F \to X$ such that $j = \varphi \circ \beta$.

Proposition 5.5.2:

Suppose that F is a free R-module on $\beta : B \to F$. Then the function β 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 $f : B \to R$ 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 β, there is an R-module homomorphism $\varphi : F \to R$ 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-moodules which have a basis.

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

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

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

Here is the definition:

Definition 5.5.3:

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

- β is linearly independent i.e. if $a : B \to R$ is a finitely supported function and if
$$\sum_{b \in B} a(b)\beta(b) = 0 \text{ then } a = 0.$$
- β spans M; i.e. for $x \in M$ there is a finitely supported function $a : B \to R$ 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 $\beta : B \to M$ 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 $a : B \to R$.

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 $\iota : B \to F$. Similarly, one says that a subset $B \subset F$ is a basis if the inclusion mapping $\iota : B \to F$

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 $a : B \to R$. Consider the function $\beta_0 : B \to F(B, R)$ 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 β_0.
b. β_0 is an R-basis for F.

Proof: For any finitely supported function $a : B \to R$ 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 $\varphi : B \to N$ be any function. We must show that there is a unique R-module mapping $\Phi : F \to N$ such that

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

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

   Using $(\heartsuit)$, the R-linearity of Φ, 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 β_0 is R-linearly independent. We suppose that $a : B \to R$ 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 β_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 $\iota_b : R \to C$ be the coproduct (direct sum) of the modules $M_b = R$ for $b \in B$.

With notation β_0 as in Proposition 6.1, there is an R-module isomorphism

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

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

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

Using the defining property of the direct sum (coproduct), there is a unique R-module homomorphism $\Psi : C \to F(B, R)$ 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 Ψ is an isomorphism, we construct the inverse homomorphism. For this, consider the function $j : B \to C$ defined by $j(b) = \iota_b(1)$. According to Proposition 6.1, $F(B, R)$ is a free R-module on β_0. Using the defining propert of a free R-module, we find a unique R-module homomorphism

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

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

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 homomomorphism $f : C \to C$ for which $f \circ \iota_b = \iota_b$ is $f - \mathrm{id}_C$. Thus we conclude that $\Phi \circ \Psi = \mathrm{id}_C$.

Similarly,

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

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

Theorem 6.4:

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

The following are equivalent:

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

Proof: $(a \Rightarrow b)$: Let β be a basis; we show that M is free on β. Since β 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 $a : B \to R$.

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

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

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

We claim: (♠) a linear map $\Phi : M \to N$ 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 $\Phi \circ \iota_b : R \to N$ and $f_b : R \to N$ are equal if and only if they agree at $1 \in R$. This proves the claim.

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

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

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

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

where $F(B, R)$ the is the R-module of finitely supported functions $B \to R$ and where β 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 β_0. Apply the defining property of a free module – see Definition 5.5.1 – to the function $\beta : B \to M$ to obtain an R-module homomorphism $\Phi : F(B,R) \to M$ with the property that $\Phi \circ \beta_0 = \beta$.

We claim that the R-module homomorphisms Φ and Ψ are inverse to one another. Once this claim is established, we see that β_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

$$\Phi \circ \Psi : M \to M$$

satisfies

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

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

Finally note that

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

satisfies

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

Since $F(B,R)$ is a free R-module on β_0, $\mathrm{id}_{F(B,R)}$ is the unique R-module map such that $\mathrm{id}_{F(B,R)} \circ \beta_0 = \beta_0$.

Thus $\Psi \circ \Phi = \mathrm{id}_{F(B,R)}$ 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 $S = \mathrm{Mat}_n(F)$, then $Z(S) = F.\,\mathrm{id} \simeq F$; 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 $R \to Z(A)$.

Remark 6.1.4:

a. 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.

b. If F is a field, for any F-algebra A the mapping $F \to A$ is injective.

Definition 6.1.5: If A and B are R-aglebras, a homomorphism of R-algebras is a ring homomorphism $\varphi : A \to B$ such that φ 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 $\varphi : A \to B$ is a homomorphism of R-algebras $\Leftrightarrow$ the restriction of φ to the image `overline`$\{R\}$ in $Z(A)$ satisfies $\varphi(\overline{r}) = \overline{r}$ for $r \in R$.

Example 6.1.7:

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

6.2. Integeral 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 α 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 arguet 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 inteegral 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 α 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 α is indeed a unit.

$(\Leftarrow)$: Suppose that R/I is a field. We must argue that I is a maximal ideal. If $\pi : R \to R/I$ 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 $M \times M \to M$ such that

- the operation is associative: $\forall x, y, z \in M, (xy)z = x(yz)$.
- there is a 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 $f : M \to N$ is a monoid homomorphism if $f(xy) = f(x)f(y)$ for every $x, y \in M$.

Example 6.3.3:

a. Monoids are "not quite groups" – there is no requirement that elements have inverses.
b. 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 $\mathbb{N}^m$ is a monoid under addition.
c. 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

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

be finitely supported functions, and let

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

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 (♣) 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 $e : M \to R[M]$ is a monoid homomorhpism. The algebra $R[M]$ is commutative if and only if M is a commutative monoid. The homomorphism $R \to R[M]$ 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

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

In particular, e will be a monoid homomorphism $M \to (R[M], \times)$ 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 $c_{\vec{\jmath}} \in R$ 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 (♦) 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 $a : M \to R$, 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 $R \to Z(R[M])$ so that $R[M]$ is indeed an R-algebra. ■

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

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

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

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

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

Proof: b. is an immediate consequence of a.

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

It only remainds to argue that φ is indeed a ring homomorphism; i.e. that φ 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

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

where 1 in the jth position.

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

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

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

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

Since the $a_1, ..., a_m$ are central in S, for $\vec{\imath}, \vec{\jmath} \in \mathbb{N}^m$ we have

$$\varphi(\vec{\imath} + \vec{\jmath}) = a_1^{i_1 + j_1} ... a_m^{i_m + j_m} = a_1^{i_1} ... a_m^{i_m} a_1^{j_1} ... a_m^{j_m} = \varphi(\vec{\imath}) \cdot \varphi(\vec{\jmath})).$$

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

Remark 6.4.3:

a. Compare with [Dummit-Foote, Ch. 7.2]
b. Notice that the polynomial ring $R[T_1, ..., T_m]$ is a commutative R-algebra (since $\mathbb{N}^m$ is a commutative monoid).
c. Note for $\vec{\imath}, \vec{\jmath} \in \mathbb{N}^m$ that $T^{\vec{\imath}} T^{\vec{\jmath}} = T^{\vec{\imath} + \vec{\jmath}}$.
d. The *variables* T_i of the polynomial ring correspond to the elements $e(\delta_i)$ of the monoid algebra, where $\delta_i = (0, 0, ..., 0, 1, 0, ..., 0) \in \mathbb{N}^m$ has 1 in the ith position.
e. Elements $f, g \in S = R[T_1, ..., T_n]$, 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 $a_1, \ldots, a_n \in Z(S)$.

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

$$\varphi : R[T_1, \ldots, T_n] \to S$$

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 homomomorphism $\psi : R[T_1, \ldots, T_n] \to S$ 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 $\gamma : \mathbb{N}^m \to (S, \times)$ 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 $A = \texttt{cal}[P](X)$, 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

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

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

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

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 $f : I \to \cup_{i \in I} A_i$ 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 negtation 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 $k[T_1, ..., T_n]$ be the polynomial ring over k, and let $a_1, ..., a_n \in k$.