Algebraic Structures Notes

Instructor: Anton Bernshteyn
Notes by: Lichen Zhang
Carnegie Mellon University
21-373 Algebraic Structures

Fall 2019

Contents

1 Sets, Functions and Equivalence Relations **3**
1.1 Basic definitions . . . 3
1.2 Higher-order functions . . . 5
1.3 Equivalence relations & quotients . . . 6

2 Functors **8**
2.1 Covariant functors . . . 8
2.2 Universal elements . . . 10
2.3 Contravariant functors . . . 12

3 Groups **13**
3.1 Concrete groups . . . 13
3.2 Abstract groups . . . 17
3.3 Abelian groups, subgroups, homomorphisms . . . 21
3.4 Actions, cosets, orbits . . . 23
3.5 Orbit-stabilizer theorem . . . 28
3.6 Normal subgroups and kernels . . . 32
3.7 Groups and counting . . . 34

4 Categories **36**
4.1 Basic definitions & examples . . . 36
4.2 Monomorphism, epimorphism . . . 38

5 Advanced Topics on Groups **41**
5.1 Product & coproduct of groups . . . 41
5.2 Free groups . . . 43
5.3 Structure of finite groups . . . 52

6 Rings and Modules **56**
6.1 Basic definitions 56
6.2 Free modules, independent, spanning, and basis 58
6.3 Commutative rings and IBN 61
6.4 Polynomials 67

1 Sets, Functions and Equivalence Relations

1.1 Basic definitions

We start these notes by reminding readers of some basic definitions of sets, functions and equivalence relations, which, though fundamental, are critical in the study of various algebraic structures. Given two sets A, B, except for the standard notation $A \cup B, A \cap B, A \setminus B, A \subseteq B, A \subset B$, we introduce the *symmetric difference* of two sets.

Definition 1.1. The *symmetric difference* of two sets A, B, denoted by $A \triangle B$, is defined as $(A \setminus B) \cup (B \setminus A)$.

Definition 1.2. The set of all functions $f : A \to B$ is denoted as B^A.

Exercise 1.3. If both A, B are finite, show that $\left|B^A\right| = |B|^{|A|}$.

Proof. To count all functions from A to B, it is useful to think of it in a *bucket-item* manner: each element of A can be viewed as a *bucket*, while each element of B is an *item*. Counting the number of functions is just counting the number of ways to put exactly one item into each bucket. For each bucket, there are $|B|$ items we can put into it. Moreover, these items can be put into as many buckets as we want, so there are $\underbrace{|B||B|\ldots|B|}_{|A| \text{ times}} = |B|^{|A|}$ functions in total. □

Definition 1.4. Let $f : A \to B$ be a function.

- f is *injective*, if for $a_1, a_2 \in A$, $a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$.
- f is *surjective*, if for any $b \in B$, $b \in \operatorname{im}(f)$, or equivalently, $B = \operatorname{im}(f)$.
- f is *bijective* if it is both injective and surjective.

Definition 1.5. Two functions f, g are *equal*, if $\operatorname{dom}(f) = \operatorname{dom}(g)$, and for any $x \in \operatorname{dom}(f)$, $f(x) = g(x)$.

Example 1.6.

1. Let A be a set, $\{0\}$ be a 1-element set, what is $A^{\{0\}}$? It's the set of all functions from $\{0\}$ to A. It is clear that there is a bijection between $A^{\{0\}}$ and A by $f \mapsto f(0)$. In such a scenario, we say A is *isomorphic* to $A^{\{0\}}$.

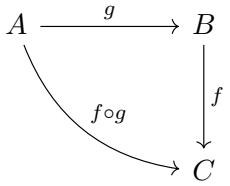
   Definition 1.7. Given two sets A, B, we say A is *isomorphic* to B, denoted by $A \cong B$, if there exists a bijection from A to B.

2. $A^{\{0,1\}} \cong A \times A$, the bijection is given by $f \mapsto (f(0), f(1))$.
3. What are $A^{\emptyset}, \emptyset^A$? They are different as long as $A \neq \emptyset$.
4. $\{0,1\}^A \cong \mathcal{P}(A)$, the bijection is $f \mapsto \{a \in A : f(a) = 1\}$.

Definition 1.8. Given three sets A, B, C, two functions $f : B \to C, g : A \to B$, define the *composition* of functions f, g, denoted by $f \circ g$, as follows:

- $\operatorname{dom}(f \circ g) = \{x \in \operatorname{dom}(g) : g(x) \in \operatorname{dom}(f)\}$.

- For $x \in \operatorname{dom}(f \circ g)$, define $(f \circ g)(x) = f(g(x))$.



Exercise 1.9 (Important)**.** Verify function composition is associative, i.e., given f, g, h three functions with well-behaved domains & codomains, verify $f \circ (g \circ h) = (f \circ g) \circ h$.

Example 1.10. Applying a function to an element can be "encoded" as a composition operation: let $f : A \to B$ be our function, consider $f(a)$, we can replace a by the function $g : \{0\} \to A$ by $0 \mapsto a$, then $f \circ g : \{0\} \to B : 0 \mapsto f(a)$.

Definition 1.11. Given a set A, the *identity function* on A, denoted by $\mathrm{id}_A : A \to A$, is the function $a \mapsto a$.

Exercise 1.12. Let $f : A \to B$, then $f \circ \mathrm{id}_A = \mathrm{id}_B \circ f = f$.

Definition 1.13. Let $f : A \to B, g : B \to A$, if $f \circ g = \mathrm{id}_B$, then f is a *left inverse* of g and g is a *right inverse* of f. If f is both a left and a right inverse of g, then f is a *two-sided inverse* of g.

Exercise 1.14. Give an example such that f is a left inverse, but not a right inverse of g.

Exercise 1.15. Let $A \neq \emptyset, B$ be two sets and $f : A \to B$. Prove the following:

1. f has a left inverse $g : B \to A$ if and only if f is injective.
2. f has a right inverse $g : B \to A$ if and only if f is surjective.

The proof is not hard, for the first part, one simply chooses g to send each element of $\operatorname{im}(f)$ to its unique preimage under f, and everything else to some fixed element of A, this is where we need $A \neq \emptyset$. For the second part, g can simply send each element to one of its preimages under f.

Definition 1.16. Given $f : A \to B$, the following are equivalent:

1. f is bijective.
2. f has a right and left inverse.
3. f has a two-sided inverse.

In this case, we denote the (two-sided) inverse of f as f^{-1}. Also, f^{-1} is a bijection from B to A and $\left(f^{-1}\right)^{-1} = f$.

Exercise 1.17. Suppose $f : A \to B, g : B \to A$ are both bijective, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

1.2 Higher-order functions

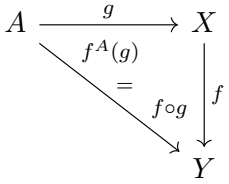
In this section, we study *higher-order functions*, which take functions as inputs and spit out functions. This idea has been heavily exploited in functional programming and the study of programming languages. We motivate with an example.

Example 1.18. Let X, Y, A be sets, then we claim $A^{X\times Y} \cong \left(A^X\right)^Y$.

Given $f \in A^{X\times Y}$, i.e., $f : X \times Y \to A$, we associate f to the function $F : Y \to A^X$ as follows: for any $y \in Y, x \in X$, $(F(y))(x) := f(x, y)$. This is called *function currying*.

The inverse of this operation is as follows: to each $F \in \left(A^X\right)^Y$, assign $f : X \times Y \to A$ by $f(x, y) := (F(y))(x)$.

Definition 1.19. Let X, Y, A be sets, to each function $f : X \to Y$ corresponds a function $f^A : X^A \to Y^A$, given by $f^A(g) := f \circ g$.



Proposition 1.20. *Let X, Y, Z be sets and $f : X \to Y, g : Y \to Z$, then $(g \circ f)^A = g^A \circ f^A$.*

Proof. First, let's compare the domains.

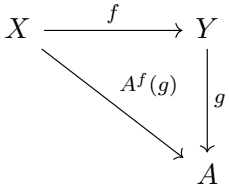
- $g \circ f : X \to Z$, so $(g \circ f)^A : X^A \to Z^A$
- $g^A : Y^A \to Z^A, f^A : X^A \to Y^A$, so $g^A \circ f^A : X^A \to Z^A$

Take any $h \in X^A$, let's do some computations.

- $(g \circ f)^A(h) = (g \circ f) \circ h$
- $g^A \circ f^A(h) = g^A(f \circ h) = g \circ (f \circ h)$

By Exercise 1.9, we know that function composition is associative, therefore, they are equal. □

Definition 1.21. Let X, Y, A be sets, to each function $f : X \to Y$ corresponds a function $A^f : A^Y \to A^X$ by $\left(A^f\right)(g) = g \circ f$.



Exercise 1.22. Let X, Y, Z be sets and $f : X \to Y, g : Y \to Z$, then $A^{g\circ f} = A^f \circ A^g$.

1.3 Equivalence relations & quotients

Definition 1.23. A *(binary) relation* on a set X is a subset $R \subseteq X \times X$, usually, we write xRy instead of $(x, y) \in R$.

Definition 1.24. An *equivalence relation* E on a set X is a relation such that

1. E is reflexive.
2. E is symmetric.
3. E is transitive.

Example 1.25. Consider the following relations on $\mathbb{R}$:

- $xEy \Leftrightarrow x - y \in \mathbb{Z}$, this *is* an equivalence relation.
- $xEy \Leftrightarrow x + y \in \mathbb{Z}$, this is *not* an equivalence relation, transitivity does not hold.

Definition 1.26. Given a function $f : X \to Y$, the *equivalence kernel* of f is the relation E_f on X, given by $xE_f y \Leftrightarrow f(x) = f(y)$, this E_f is an equivalence relation.

Conversely, given an equivalence relation E on X, for $x \in X$, let $[x]_E := \{y \in X : xEy\}$, defined as the *equivalence class* of x.

Let $X/E := \{[x]_E : x \in X\}$ be the *quotient* of X by E, the map $p_E : X \to X/E : x \mapsto [x]_E$ is called the *quotient map*.

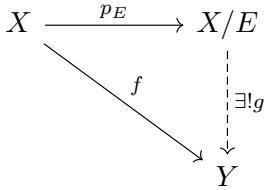
Proposition 1.27. *Let E be an equivalence relation on X, then $p_E : X \to X/E$ is a surjection, with equivalence kernel E.*

Proof. The first part is clear: any equivalence class $[x]_E$ has at least one canonical representative, namely x in it, since E is reflexive. To see the second part, consider the following chain of equivalences: $xEy \Leftrightarrow [x]_E = [y]_E \Leftrightarrow p_E(x) = p_E(y)$. Why is the first equivalence true? If xEy, then by symmetry and transitivity, $xEz \Leftrightarrow yEz$ for any $z \in X$, so $[x]_E = [y]_E$. On the other hand, if $[x]_E = [y]_E$, then $y \in [y]_E = [x]_E$ by reflexivity, i.e., xEy. □

Definition 1.28. Let E be an equivalence relation on a set X, a function $f : X \to Y$ is *E-invariant* if $xEy \Rightarrow f(x) = f(y)$, i.e., f is constant on equivalence classes.

Theorem 1.29. *Let E be an equivalence relation on X, if $f : X \to Y$ is E-invariant, then there is a unique function $g : X/E \to Y$ such that $f = g \circ p_E$.*

Proof. The theorem is equivalent to the following diagram:



Define a function g as follows: for each $C \in X/E$, let $g(C) := f(x)$, for any $x \in C$. Since f is E-invariant, this does not depend on the choice of $x \in C$, so g is well-defined.

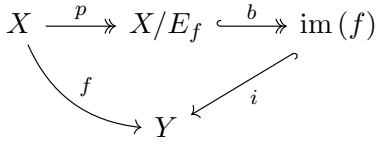
Moreover, $(g \circ p_E)(x) = g(p_E(x)) = g([x]_E) = f(x)$, this proves the existence of such a g.

To see uniqueness, if $f = g \circ p_E$, then for any x, $(g \circ p_E)(x) = g([x]_E) = f(x)$, and every $C \in X/E$ is of the form $[x]_E$, therefore, it's unique. □

Exercise 1.30. Based on the construction of g above, prove the following:

1. g is injective if and only if the equivalence kernel of f is E.
2. g is surjective if and only if f is surjective.
3. If $E_f = E$ and $\operatorname{im}(f) = Y$, then g is a bijection.

Corollary 1.31 (Canonical decomposition theorem)**.** *Let $f : X \to Y$, then f can be decomposed as follows:*



where:

- *p is the quotient map p_{E_f}, which is surjective.*
- *i is the inclusion map, which is injective.*
- *b is a bijective map, which is g from Theorem 1.29 applied with $E = E_f$ to f regarded as a map $X \to \operatorname{im}(f)$, so b is bijective by the exercise above.*

2 Functors

In this section, we study the first algebraic structure in this course, which operates on sets and functions.

2.1 Covariant functors

Definition 2.1. A *functor* (on sets to sets) is a rule $\mathcal{F}$ that assigns each set S a set $\mathcal{F}(S)$ and to each function $f : S \to T$ a function $\mathcal{F}(f) : \mathcal{F}(S) \to \mathcal{F}(T)$ such that

1. For each set S, $\mathcal{F}(\mathrm{id}_S) = \mathrm{id}_{\mathcal{F}(S)}$.
2. For any $f : S \to T$, $g : R \to S$, we have $\mathcal{F}(f \circ g) = \mathcal{F}(f) \circ \mathcal{F}(g)$.

One can think of a functor $\mathcal{F}$ as consisting of

- an "object map": $\mathcal{F}_O$: sets $\to$ sets.
- a "mapping map": $\mathcal{F}_M$: functions $\to$ functions.

Let's look at some examples of functors to form a better sense of the definition.

Example 2.2.

1. The identity functor, given by $\mathrm{Id}(S) = S, \mathrm{Id}(f) = f$.
2. Fix a set A, let $\mathcal{F}(S) := A$ for any set S, and let $\mathcal{F}(f) := \mathrm{id}_A$, for any f.

   Exercise 2.3. Are there any other functors with $\mathcal{F}(S) = A$, for any set S?

3. The powerset functor, given by $\mathcal{F}(S) = \mathcal{P}(S)$, and for function $f : S \to T$, we need to have $\mathcal{F}(f) : \mathcal{P}(S) \to \mathcal{P}(T)$. For any subset $A \subseteq S$, let $(\mathcal{F}(f))(A) := \{f(a) : a \in A\}$, i.e., the image of A under f, or $f(A)$. Let's verify it's a functor.

   Proof. Notice $\mathcal{F}(\mathrm{id}_S)(A) = \mathrm{id}_S(A) = A$, and $\mathrm{id}_{\mathcal{F}(S)}(A) = A$, so they are equal. For composition, we use f_* as a shorthand notation for $\mathcal{F}(f)$, then

   - $(f \circ g)_*(A) = \{(f \circ g)(a) : a \in A\} = (f \circ g)(A) = f(g(A))$
   - $f_* \circ g_*(A) = f_*(g_*(A)) = f_*(g(A)) = f(g(A))$

   Again, the composition is preserved. □

4. Fix a set A, define the functor to raise sets and functions to the power of A: $S \mapsto S^A$, $f \mapsto f^A$. This defines a functor, since in Proposition 1.20, we have shown that $(f \circ g)^A = f^A \circ g^A$. Identity is easy to verify.
5. For each set S, let $[S]^{<\infty}$ be the set of all *finite* subsets of S, define a functor as follows: for each set S, $\mathcal{F}(S) := [S]^{<\infty}$, and for each function $f : S \to T$, $\mathcal{F}(f) : \mathcal{F}(S) \to \mathcal{F}(T)$, given by $(\mathcal{F}(f))(A) := \{f(a) : a \in A\}$. Since the input subset A is finite, the output is also finite, so this functor is well-defined.

Definition 2.4. A functor $\mathcal{G}$ is called a *subfunctor* of $\mathcal{F}$ if for any set S, $\mathcal{G}(S) \subseteq \mathcal{F}(S)$, and for any $f : S \to T$, $\mathcal{G}(f) = (\mathcal{F}(f))|_{\mathcal{G}(S)}$.

A quick observation: suppose $\mathcal{F}$ is a functor and let $\mathcal{G}_0$ be a mapping that assigns each set S a subset $\mathcal{G}_0(S) \subseteq \mathcal{F}(S)$, then there is a unique subfunctor $\mathcal{G}$ of $\mathcal{F}$ whose object map is $\mathcal{G}_0$ if and only if $\operatorname{im}\left(\mathcal{F}(f)|_{\mathcal{G}_0(S)}\right) \subseteq \mathcal{G}_0(T)$, for any $f : S \to T$.

Example 2.5. Fix a set A, let $\mathcal{F}$ be the functor given by $\mathcal{F}(S) := S^A$, and for each $f : S \to T$, define $\mathcal{F}(f) := f^A$. Let E be an equivalence relation on A, let $\mathcal{G}_0(S)$ be the set of all E-invariant functions $f : A \to S$. This gives rise to a subfunctor of $\mathcal{F}$: first, it's clear that for any set S, we have $\mathcal{G}_0(S) \subseteq S^A$, our goal is to show that $f^A|_{\mathcal{G}_0(S)} : \mathcal{G}_0(S) \to \mathcal{G}_0(T)$ is well-defined and well-behaved, i.e., $f^A|_{\mathcal{G}_0(S)}(g) = f \circ g$, is this function E-invariant? Notice that by definition, g is E-invariant, so $a_1 E a_2 \Rightarrow g(a_1) = g(a_2) \Rightarrow f(g(a_1)) = f(g(a_2))$.

2.2 Universal elements

Definition 2.6. A *universal element* for a functor $\mathcal{F}$ is a pair (R, u), where R is a set, $u \in \mathcal{F}(R)$ such that for any set S and each $s \in \mathcal{F}(S)$, there exists a unique function $h : R \to S$, satisfying $(\mathcal{F}(h))(u) = s$.

Example 2.7. Fix a set A and an equivalence relation E, we have a functor $\mathcal{F}$ that maps S to $\mathcal{F}(S) = \{E\text{-invariant functions } g : A \to S\}$, and each $f : S \to T$ to $\mathcal{F}(f)$ by $g \mapsto f \circ g$.

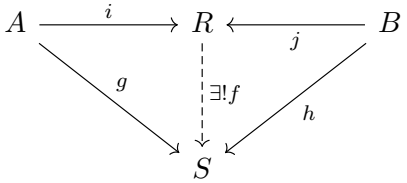
What is the universal element of this functor? We need a set R, and $u \in \mathcal{F}(R)$, where u is an E-invariant function from A to R, our goal is to satisfy the *universal property*, i.e., for all sets S and $s : A \to S$ which is E-invariant, there is a unique function $h : R \to S$ such that $\mathcal{F}(h)(u) = h \circ u = s$.

Let's set $R = A/E$, and $u = p_E$, set $h([x]_E) = s(x)$. Given such u, h is well-defined and must be unique by Theorem 1.29, so $(A/E, p_E)$ is a universal element of the functor $\mathcal{F}$.

Example 2.8. Fix sets A, B, define $\mathcal{F}(S) = S^A \times S^B$, given $f : S \to T$, define $\mathcal{F}(f)$ by $(g, h) \mapsto (f \circ g, f \circ h)$.

Exercise 2.9. Prove it's a functor.

What is the universal element? It's called the *disjoint union* of A and B. Let R be our target set, $u = (i, j) \in \mathcal{F}(R) = R^A \times R^B$, for a set S and $s = (g, h) \in \mathcal{F}(S)$, there is a unique function $f : R \to S$ such that $\mathcal{F}(f)(i, j) = (g, h) = (f \circ i, f \circ j)$.



If $A \cap B = \emptyset$, then we can simply set $R = A \cup B$ and let i, j be the inclusion maps. Set

$$f(x) = \begin{cases} g(x), & \text{if } x \in A \\ h(x), & \text{if } x \in B \end{cases}$$

On the other hand, if $A \cap B \neq \emptyset$, then let A', B' be $A' := A \times \{0\}, B' = B \times \{1\}$, let $R := A' \cup B'$, and i, j map each element with an extra bit. By this bit, f has the ability to decode the input. Set

$$f(x) = \begin{cases} g\left(i^{-1}(x)\right), & \text{if } x \in A' \\ h\left(j^{-1}(x)\right), & \text{if } x \in B' \end{cases}$$

In both cases, this f is the only possible choice: $f \circ i = g$ and $f \circ j = h$ fix the value of f at every element of $R = i(A) \cup j(B)$. This construction is referred to as the *disjoint union* of A and B.

Proposition 2.10. *Let (R, u) be a universal element of $\mathcal{F}$, suppose $b : R \to R'$ is a bijection, then (R', u') is also a universal element, where $u' = \mathcal{F}(b)(u)$.*

Proof. Take any set S and $s \in \mathcal{F}(S)$, we need to show that there exists a unique $h' : R' \to S$ such that $\mathcal{F}(h')(u') = s$. We separately show the existence and uniqueness.

Existence: since (R, u) is universal, there exists a unique $h : R \to S$ with $\mathcal{F}(h)(u) = s$. Define $h' = h \circ b^{-1}$, then

$$\begin{aligned}
\mathcal{F}\left(h'\right)\left(u'\right) &= \mathcal{F}\left(h'\right)\left(\mathcal{F}(b)(u)\right) \\
&= \mathcal{F}\left(h'\right) \circ \mathcal{F}(b)(u) \\
&= \mathcal{F}\left(h' \circ b\right)(u) \\
&= \mathcal{F}\left(h \circ b^{-1} \circ b\right)(u) \\
&= \mathcal{F}(h)(u) \\
&= s
\end{aligned}$$

Uniqueness: Notice that for any h' with $\mathcal{F}(h')(u') = s$, we must have $\mathcal{F}(h' \circ b)(u) = \mathcal{F}(h)(u)$: the first three lines of the computation above give $\mathcal{F}(h' \circ b)(u) = \mathcal{F}(h')(u') = s = \mathcal{F}(h)(u)$. By uniqueness of h, we must have $h' \circ b = h$, so the only choice for h' is $h \circ b^{-1}$. □

The next theorem is the converse of the above proposition.

Theorem 2.11 (Uniqueness of universal element)**.** *Let $\mathcal{F}$ be a functor, if $(R, u), (R', u')$ are universal for $\mathcal{F}$, then there exists a unique bijection $b : R \to R'$ such that $u' = \mathcal{F}(b)(u)$.*

Proof. Since (R, u) is universal, there exists a unique $b : R \to R'$ such that $\mathcal{F}(b)(u) = u'$. Similarly, (R', u') is universal, so there exists a unique b' such that $\mathcal{F}(b')(u') = u$. It suffices to show b is a bijection, and b' is its inverse, i.e., $b \circ b' = \mathrm{id}_{R'}, b' \circ b = \mathrm{id}_R$.

$$\begin{aligned}
\mathcal{F}\left(b' \circ b\right)(u) &= \mathcal{F}\left(b'\right) \circ \mathcal{F}(b)(u) \\
&= \mathcal{F}\left(b'\right)\left(u'\right) \\
&= u
\end{aligned}$$

Apply the universal property to u itself, there exists a unique $h : R \to R$ such that $\mathcal{F}(h)(u) = u$. First observe that $\mathrm{id}_{\mathcal{F}(R)}(u) = u$, and $\mathcal{F}(\mathrm{id}_R) = \mathrm{id}_{\mathcal{F}(R)}$, so h must be id_R. On the other hand, $\mathcal{F}(b' \circ b)(u) = u$, which implies $b' \circ b = h = \mathrm{id}_R$. The other identity is similar. □

2.3 Contravariant functors

Definition 2.12. A *contravariant functor* $\mathcal{C}$ assigns to each set S a set $\mathcal{C}(S)$ and to each $f : S \to T$ a function $\mathcal{C}(f) : \mathcal{C}(T) \to \mathcal{C}(S)$ such that

- $\mathcal{C}(\mathrm{id}_S) = \mathrm{id}_{\mathcal{C}(S)}$
- If $f : S \to T, g : R \to S$, then $\mathcal{C}(f \circ g) = \mathcal{C}(g) \circ \mathcal{C}(f)$

Example 2.13 (Contravariant powerset functor)**.** Define $\mathcal{C}$ as follows: $S \mapsto \mathcal{P}(S)$, $f \mapsto f_*$ by $f_* : \mathcal{P}(T) \to \mathcal{P}(S)$, for each set $B \subseteq T$, $f_*(B) = \{a \in S : f(a) \in B\}$, i.e., the *preimage* of B under f.

Exercise 2.14. Check it's a contravariant functor.

Example 2.15. Fix a set A, the following is contravariant: $S \mapsto A^S, f \mapsto A^f$, i.e., $f : S \to T, A^f : A^T \to A^S$, and $A^f(g) = g \circ f$.

Definition 2.16. Let $\mathcal{C}$ be a contravariant functor, a *universal element* for $\mathcal{C}$ is a pair (R, u), where R is a set and $u \in \mathcal{C}(R)$, such that for any S and $s \in \mathcal{C}(S)$, there exists a unique $h : S \to R$ such that $\mathcal{C}(h)(u) = s$.

Universal elements for contravariant functors are very similar to universal elements for covariant functors, except for the function h, now its domain is S and codomain is R.

Example 2.17. Fix sets A, B, define a contravariant functor $\mathcal{C}$ as follows: $\mathcal{C}(S) := A^S \times B^S$, for $f : S \to T$, define $\mathcal{C}(f) : \mathcal{C}(T) \to \mathcal{C}(S)$ by $\mathcal{C}(f)(g, h) := (g \circ f, h \circ f)$.

What is the universal element? We want (R, u), $u \in \mathcal{C}(R)$, so $u = (p, q) \in A^R \times B^R$, for any S and $s = (g, h) \in \mathcal{C}(S)$, there exists a unique $f : S \to R$ such that $s = \mathcal{C}(f)(p, q) = (p \circ f, q \circ f) = (g, h)$. Set $R = A \times B$, $p(a, b) = a, q(a, b) = b$, and $f(x) = (g(x), h(x))$. It's not hard to check this is a universal element.

Similar to covariant functors, contravariant functors also have a uniqueness theorem for universal elements.

Theorem 2.18. *If $(R, u), (R', u')$ are universal elements for a contravariant functor $\mathcal{C}$, then there exists a unique bijection $b' : R' \to R$ such that $u' = \mathcal{C}(b')(u)$.*

Proof. Similar to the covariant case, since (R, u) and (R', u') are universal, there exist a unique $b' : R' \to R$ with $\mathcal{C}(b')(u) = u'$ and a unique $b : R \to R'$ with $\mathcal{C}(b)(u') = u$. Then $\mathcal{C}(b' \circ b)(u) = \mathcal{C}(b)(\mathcal{C}(b')(u)) = u = \mathcal{C}(\mathrm{id}_R)(u)$, so $b' \circ b = \mathrm{id}_R$ by the uniqueness in the universal property of (R, u). The other identity $b \circ b' = \mathrm{id}_{R'}$ is similar, so b' is a bijection. □

3 Groups

In this section, we study groups, one of the most famous algebraic structures, with a simple definition but surprisingly complicated dynamics.

3.1 Concrete groups

Definition 3.1. A *concrete group* on a set X is a set G of bijections $X \to X$, such that

1. $\mathrm{id}_X \in G$
2. if $f \in G$, then $f^{-1} \in G$
3. if $f, g \in G$, then $f \circ g \in G$

Example 3.2.

1. The set $\left\{f \in X^X : f \text{ is a bijection}\right\}$ is a concrete group on X, we call it the *symmetric group* on X, denoted by $\mathrm{Sym}(X)$. The elements of $\mathrm{Sym}(X)$ are called *permutations*.
2. $\{\mathrm{id}_X\}$ is a concrete group.
3. $I = \{f \in \mathrm{Sym}(\mathbb{R}) : \forall x, y \in \mathbb{R}, x < y \Rightarrow f(x) < f(y)\}$ is a concrete group on $\mathbb{R}$.
   (a) $\mathrm{id}_{\mathbb{R}} \in I$
   (b) $f \in I \Rightarrow f^{-1} \in I$. Take $x, y \in \mathbb{R}$ with $x < y$, our goal is to show that $f^{-1}(x) < f^{-1}(y)$. Suppose not, then since f^{-1} is injective, $f^{-1}(x) > f^{-1}(y)$, but f is increasing, so $f\left(f^{-1}(x)\right) > f\left(f^{-1}(y)\right) \Rightarrow x > y$, contradicts $x < y$.
   (c) $f, g \in I \Rightarrow f \circ g \in I$, take $x, y \in \mathbb{R}$ with $x < y$, then $g(x) < g(y)$, so $f(g(x)) < f(g(y))$.
4. Let $X \subseteq \mathbb{R}^n$, a bijection $f : X \to X$ is an *isometry* if for all $x, y \in X$, $\|x - y\|_2 = \|f(x) - f(y)\|_2$. $\mathrm{Iso}(X) := \{f \in \mathrm{Sym}(X) : f \text{ is an isometry}\}$ is a concrete group.
5. Let $X \subseteq \mathbb{R}^n$, a bijection $f : X \to X$ is a *homeomorphism* if f and f^{-1} are continuous. $\mathrm{Homeo}(X) := \{f \in \mathrm{Sym}(X) : f \text{ is a homeomorphism}\}$ is a concrete group on X.

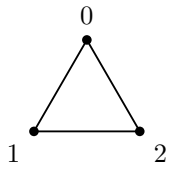
   Exercise 3.3. Show that $\{f \in \mathrm{Sym}(X) : f \text{ is continuous}\}$ might not be a concrete group.

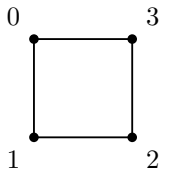
6. Graphs: Let $G = (V, E)$, an *automorphism* of G is a bijection $f : V \to V$, such that for any $x, y \in V$, $\{x, y\} \in E \Leftrightarrow \{f(x), f(y)\} \in E$. The set

$$\mathrm{Aut}(G) := \{f \in \mathrm{Sym}(V) : f \text{ is an automorphism}\}$$

   is a concrete group on V. Below are several graphs illustrating this idea:

① $G =$ $\Rightarrow$ a long path

② $G =$  $\Rightarrow$ triangle / 3-cycle

③ $G =$  $\Rightarrow$ square / 4-cycle

The first graph is just a long path, there are only two automorphisms, namely, the identity and flipping the entire path. Notice no matter the length of this path, there are only these two kinds of automorphisms.

The second graph is a triangle, and $\mathrm{Aut}(G) = \mathrm{Sym}(\{0,1,2\})$. Why? No matter how we permute the three vertices, it remains an automorphism, thus, the automorphisms are exactly all permutations.

The last one is a 4-cycle, and $|\mathrm{Aut}(G)| = 8$. To do the counting, we can first put vertex 0: 4 positions to put. Then there are 2 ways to put 1. After that, everything is fixed, so in total $4 \times 2 = 8$ automorphisms.

Two quick definitions: $C_n :=$ cycle of length n, and $D_{2n} := \mathrm{Aut}(C_n)$. Perhaps another name for D_{2n} is much more famous: it is the *dihedral group* of order $2n$, and for $n \geq 3$, $|\mathrm{Aut}(C_n)| = 2n$, we typically use $|D_{2n}|$ to denote this number.

7. Let G be a concrete group on a set X, a function $f \in \mathrm{Sym}(G)$ is an *automorphism* of G if for any $g, h \in G$, $f(g \circ h) = f(g) \circ f(h)$. It is not hard to show that if f is an automorphism of G, then $f(\mathrm{id}_X) = \mathrm{id}_X$, and id_G, f^{-1} (apply f^{-1} to $f(f^{-1}(g) \circ f^{-1}(h)) = g \circ h$) and the composite of two automorphisms are automorphisms, therefore, $\mathrm{Aut}(G) := \{f \in \mathrm{Sym}(G) : f \text{ is an automorphism of } G\}$ is a concrete group on G.

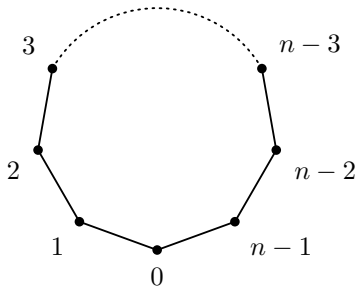
8. Let X be a set and $f \in \mathrm{Sym}(X)$, for $n \in \mathbb{N}$, we write $f^n := \underbrace{f \circ f \circ \ldots \circ f}_{n}$, $f^0 := \mathrm{id}_X$, $f^{n+1} := f \circ f^n$. Also, let $f^{-n} := \left(f^{-1}\right)^n$.

   Exercise 3.4. For any $n, m \in \mathbb{Z}$, show that:

   (a) $f^{-n} = (f^n)^{-1}$
   (b) $(f^n)^m = f^{nm}$
   (c) $f^n \circ f^m = f^{n+m}$

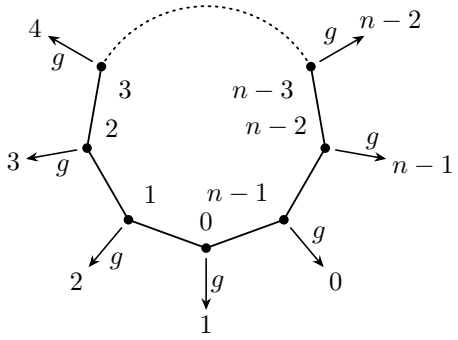
   From this, it follows that $\langle f \rangle := \{f^n : n \in \mathbb{Z}\}$ is a concrete group on X, this group is called the *cyclic group* generated by f.

Now let's dive deeper into the study of the dihedral group. It is useful to keep the following picture of the dihedral group D_{2n} in mind:



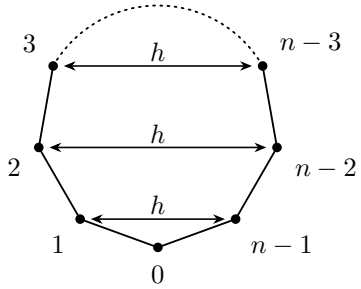
D_{2n} is just all automorphisms on the cycle of length n, and for $n \geq 3$, we have $|D_{2n}| = 2n$. Consider the following 2 elements of D_{2n}:

- $g :=$



  i.e., g rotates the cycle to the left, by 1.

- $h :=$



  i.e., h flips the cycle along the fixed 0 axis.

Proposition 3.5. *Elements of* D_{2n} *are* $\mathrm{id} = g^0, g^1, g^2, \ldots, g^{n-1}$, *and* $h = g^0 \circ h, g^1 \circ h, g^2 \circ h, \ldots, g^{n-1} \circ h$.

Proof. Take any $f \in D_{2n}$, let $k := f(0)$, we wish to express f either as g^k or $g^k \circ h$, let $f' := g^{-k} \circ f$, clearly, $f'(0) = g^{-k} \circ f(0) = g^{-k}(k) = 0$, so f' is either id or h: f' must send 1, a neighbor of 0, to a neighbor of 0, i.e., 1 or $n-1$, and then the images of the other vertices are forced one by one around the cycle. Moreover, we can write f as $g^k \circ f'$, this completes the proof. □

Exercise 3.6. For each k, ℓ, compute $\left(g^k \circ h^\ell\right)(0)$ and $\left(g^k \circ h^\ell\right)(1)$, and conclude that the $2n$ functions listed in the claim are distinct.

Proof Sketch. For different k, $\left(g^k \circ h^\ell\right)(0)$ must be different. For fixed k, $g^k(1) \neq g^k \circ h(1)$. □

Observe that D_{2n} is the smallest group containing g, h. Indeed, if G is a group containing g, h, then for any k, ℓ, $g^k \circ h^\ell \in G$, so $D_{2n} \subseteq G$. We say that D_{2n} is *generated* by g, h.

Lemma 3.7. *Let X be a set, $\mathcal{G}$ be a set of concrete groups on X, define*

$$\bigcap \mathcal{G} := \{f \in \mathrm{Sym}(X) : \forall G \in \mathcal{G}, f \in G\}.$$

Then $\bigcap \mathcal{G}$ is also a concrete group.

Proof. 1. $\mathrm{id}_X \in \bigcap \mathcal{G}$, since for any G, $\mathrm{id}_X \in G$.

2. if $f \in \bigcap \mathcal{G}$, then $f^{-1} \in \bigcap \mathcal{G}$, to see this, notice that for any G, $f \in G$, and G is a concrete group, so for any G, $f^{-1} \in G \Rightarrow f^{-1} \in \bigcap \mathcal{G}$.

3. if $f, g \in \bigcap \mathcal{G}$, then apply the same argument, $f \circ g \in \bigcap \mathcal{G}$.

□

Definition 3.8. If $S \subseteq \mathrm{Sym}(X)$, then

$$\langle S \rangle := \bigcap \{G : G \text{ is a concrete group on } X \text{ such that } S \subseteq G\}$$

is a concrete group by Lemma 3.7, furthermore, $S \subseteq \langle S \rangle$, so $\langle S \rangle$ is the smallest group containing S as a subset. $\langle S \rangle$ is called the *group generated by* S.

For example, if $f \in \mathrm{Sym}(X)$, then $\langle f \rangle = \langle \{f\} \rangle$, and more generally, we write $\langle \{f_1, f_2, \ldots, f_n\} \rangle$ as $\langle f_1, f_2, \ldots, f_n \rangle$, e.g., with $g, h \in D_{2n}$, we have $D_{2n} = \langle g, h \rangle$.

The next proposition will show that we can write elements of $\langle S \rangle$ in a more structured way.

Proposition 3.9. *Let X be a set, $S \subseteq \mathrm{Sym}(X)$, then*

$$\langle S \rangle = \left\{ g_1^{\epsilon_1} \circ g_2^{\epsilon_2} \circ \ldots \circ g_k^{\epsilon_k} : k \in \mathbb{N}, g_1, \ldots, g_k \in S, \epsilon_1, \ldots, \epsilon_k \in \{-1, 1\} \right\}$$

Proof. Let H denote the set on RHS, clearly, $H \subseteq \langle S \rangle$, and also, $S \subseteq H$. So to argue for the other inclusion, it suffices to show H is a group, then we will have $\langle S \rangle \subseteq H$ automatically.

- Take $k = 0$, we get the empty composition, which is id_X, so $\mathrm{id}_X \in H$.
- Suppose $f = g_1^{\epsilon_1} \circ \cdots \circ g_k^{\epsilon_k} \in H$, then $f^{-1} = g_k^{-\epsilon_k} \circ g_{k-1}^{-\epsilon_{k-1}} \circ \cdots \circ g_2^{-\epsilon_2} \circ g_1^{-\epsilon_1} \in H$.
- If $f, g \in H$, then it's clear $f \circ g \in H$: we just express it by adding a $\circ$ between the representation of f and g.

□

Let's look at some examples of groups generated by a set of elements.

Example 3.10. 1. Rubik's cube: let X be a set of cardinality 48, and $S :=$ set of 6 individual rotations, viewed as permutations on X. The Rubik's group is $\langle S \rangle$.

2. The *lamplighter group* is defined as follows: $X = \{0, 1\}^{\mathbb{Z}} \times \mathbb{Z}$, and there are two functions:

   - g : move to the next light, i.e., $g(x, n) := (x, n + 1)$.
   - h : switch the light the person is currently standing by, i.e., $h(x, n) := (x', n)$, where $x'_i = \begin{cases} 1 - x_i, & \text{if } i = n \\ x_i, & \text{otherwise} \end{cases}$

   We use $\langle g, h \rangle$ to denote the lamplighter group.

3.2 Abstract groups

Definition 3.11. A *binary operation* on a set S is a function $\star : S \times S \to S$.

Example 3.12.
1. If G is a concrete group on X, then $\circ : G \times G \to G$ is a binary operation, where $\circ(f, g) = f \circ g$. Note: $\star$ is always written as an infix operator.
2. There are tons of binary operations on $\mathbb{Z}$: $+$ on $\mathbb{Z}$, $*$ on $\mathbb{Z}$, and $\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z} : (n, m) \mapsto n$.

Definition 3.13. A binary operation $\star : S \times S \to S$ is *associative* if $\forall a, b, c \in S : (a \star b) \star c = a \star (b \star c)$.

Definition 3.14. Let $\star$ be a binary operation on S, then an element $e \in S$ is an *identity* for $\star$ if for all $a \in S$, $a \star e = e \star a = a$.

Lemma 3.15. *Let $\star$ be a binary operation on S, then $\star$ has at most one identity.*

Proof. Suppose both e and e' are identities for $\star$, then $e \star e' = e = e'$. □

Definition 3.16. Let $\star$ be a binary operation with identity element e on S, an *inverse* of an element a is an element $b \in S$ such that $a \star b = b \star a = e$.

Exercise 3.17. Give an example of a binary operation $\star$ with identity such that there is an element with more than one inverse.

You can intentionally assign two elements as inverses to some element, then force those two elements to be different from each other.

Lemma 3.18. *If $\star$ is associative, with identity e, then for any $a \in S$, if it has an inverse, then it has at most one inverse.*

Proof. Suppose b, b' are inverses of a, then compute $b \star a \star b' = (b \star a) \star b' = e \star b' = b'$, on the other hand, $b \star a \star b' = b \star (a \star b') = b \star e = b$, by associativity, we have $b = b'$. □

After setting up enough definitions, we are ready to define abstract groups.

Definition 3.19. An *abstract group*, or *group* for short, is a set G with a binary operation $\star$ such that

- $\star$ is associative.
- $\star$ has an identity.
- each $a \in G$ has an inverse with respect to $\star$.

The main example: if G is a concrete group on X, then $(G, \circ)$ is a group.

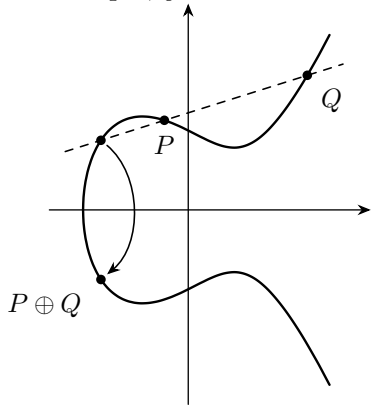
Some quick examples: $(\mathbb{N}, +)$ is not a group, but $(\mathbb{Z}, +)$ is a group. $(\mathbb{R}, *)$ is not a group, but $(\mathbb{R} \setminus \{0\}, *)$ is a group.

Exercise 3.20. Let X be a set, recall $\triangle$ is the symmetric difference operation on two sets, then $(\mathcal{P}(X), \triangle)$ is a group.

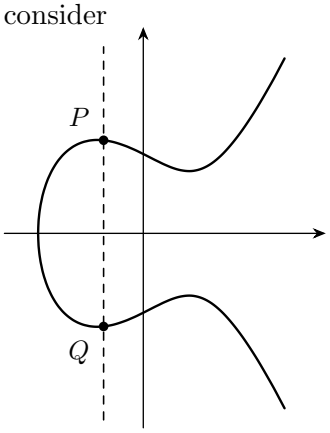
Some more examples of groups:

Example 3.21. 1. $\mathbb{Z}/n\mathbb{Z}$: let $n \geq 2$ be an integer, and $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, $+_n$: addition modulo n, i.e., for $x, y \in \mathbb{Z}_n$, define $x +_n y$ to be the remainder of $x + y$ after division by n. Then, $(\mathbb{Z}_n, +_n)$ is a group. Please verify this. The nasty part is to verify that $+_n$ is associative.

2. Elliptic curves: An elliptic curve is the set of points $(x, y) \in \mathbb{R}^2$, satisfying an equation of the form $y^2 = x^3 + ax + b$ with $4a^3 + 27b^2 \neq 0$ (so that the curve has no cusps or self-intersections), for example, $y^2 = x^3 - x + 1$. We denote the set by E. Let's illustrate this with a diagram:



Let $*$ be an extra element not in E, define an operation $\oplus$ on $E \cup \{*\}$ as follows: $P \oplus Q \Rightarrow$ take the line PQ (the tangent line at P if $P = Q$), find its third intersection with E (a point of tangency counts twice, or three times if it is an inflection point), then flip that point with respect to the x axis. What if PQ is vertical, so it intersects E only at P and Q? For example, consider



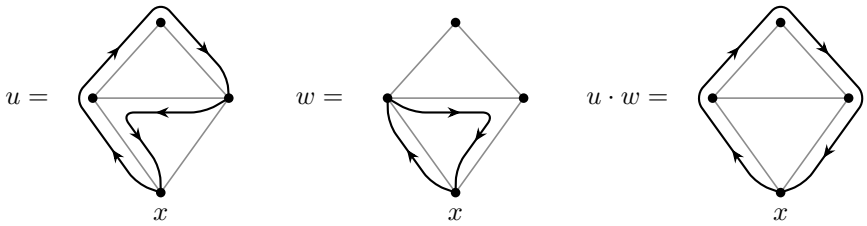
Then define $P \oplus Q = *$. For all $P \in E$, define $* \oplus P = P \oplus * = P$, and $* \oplus * = *$, i.e., $*$ is the identity element. This turns $E \cup \{*\}$ into a group.

3. Let G be a connected graph, fix a vertex $x \in V(G)$, define $\Pi_1(G, x)$, the *homotopy group* of G, as follows:

$$\Pi_1(G, x) := \{\text{all non-backtracking closed walks in } G, \text{ starting and ending at } x\}$$

Here, non-backtracking just means if you have used edge (u, v), you cannot immediately take the edge (v, u).

Given two walks u and w, $u \cdot w$ is the walk obtained from u and w by concatenating them and removing all backtrack edges. For example,



This defines a group.

Theorem 3.22 (Cayley)**.** *Every abstract group is isomorphic to a concrete group, i.e., for all groups $(G, \star)$, there exists a set X, a concrete group H on X, and a bijection $\varphi : G \to H$, such that for all $f, g \in G$, we have $\varphi(f \star g) = \varphi(f) \circ \varphi(g)$.*

Proof. Take $X = G$. For each $g \in G$, let $\varphi(g) : G \to G$ be the function given by $\varphi(g)(h) := g \star h$. We prove a sequence of propositions, to show that this φ gives us the desired result.

Proposition 3.23 (Closure of φ)**.** *For all $f, g \in G$, $\varphi(f \star g) = \varphi(f) \circ \varphi(g)$.*

Proof of Closure. Take $h \in G$. $\varphi(f \star g)(h) = (f \star g) \star h$, while $\varphi(f) \circ \varphi(g)(h) = \varphi(f)(g \star h) = f \star (g \star h)$. By associativity of $\star$, the identity holds. □

The next part will focus on proving φ is an isomorphism, i.e., it is a bijection. By Proposition 3.23, we need to show the following:

Proposition 3.24. *Let e be the identity of G, then $\varphi(e) = \mathrm{id}_G$.*

Proof of identity. Take $h \in G$, $\varphi(e)(h) = e \star h = h$, for all h, so $\varphi(e) = \mathrm{id}_G$. □

The next proposition will be useful in finding our concrete group H.

Proposition 3.25. *For all $g \in G$, $\varphi(g)$ is a bijection.*

Proof of $\varphi(g)$ bijectivity. We claim that the inverse of $\varphi(g)$ is $\varphi\left(g^{-1}\right)$:

$$\varphi(g) \circ \varphi\left(g^{-1}\right) = \varphi\left(g \star g^{-1}\right) = \varphi(e) = \mathrm{id}_G,$$
$$\varphi\left(g^{-1}\right) \circ \varphi(g) = \varphi\left(g^{-1} \star g\right) = \varphi(e) = \mathrm{id}_G.$$

□

So we can view φ as a function from G to $\mathrm{Sym}(G)$, let $H := \mathrm{im}(\varphi) = \{\varphi(g) : g \in G\}$.

Proposition 3.26. *H is a concrete group on G.*

Proof of H is a concrete group.

- $\mathrm{id}_G \in H$, since $e \in G$, $\varphi(e) \in H$.
- If $f \in H$, then $f^{-1} \in H$: suppose $f = \varphi(g)$ for some $g \in G$, then $f^{-1} = \varphi\left(g^{-1}\right) \in H$.
- If $f_1, f_2 \in H$, then $f_1 \circ f_2 \in H$: since $\varphi(g_1) \circ \varphi(g_2) = \varphi(g_1 \star g_2) \in H$.

□

Finally, we are ready to show φ is a bijection.

Proposition 3.27. *$\varphi : G \to H$ is a bijection.*

Proof. Surjectivity comes from the definition of H: it is $\operatorname{im}(\varphi)$. What about injectivity? If $\varphi(g_1) = \varphi(g_2)$, then $\varphi(g_1)(e) = g_1 \star e = g_1 = \varphi(g_2)(e) = g_2 \star e = g_2$. □

This completes the proof of Cayley's theorem. □

Let's examine Theorem 3.22 on the dihedral group. Recall $D_{2n} = \langle g, h \rangle$, where g is the rotation by 1, and h is the flip with respect to the 0 axis. What is $\varphi(g)$? If we apply it to other elements, the result will be further shifted by 1. What about $\varphi(h)$? It will flip all elements.

3.3 Abelian groups, subgroups, homomorphisms

In this section, we expand more on structures of abstract groups. We introduce three important concepts that will be used across all later sections.

We first introduce some conventions, to simplify the notation. Often, we just say "G is a group" without mentioning the operation. In such a case, the operation is called "multiplication" and denoted by $a \cdot b$ or simply, ab. There's a notable exception, i.e., groups with addition, such as $(\mathbb{Z}, +)$.

Definition 3.28. A binary operation $\star$ on set S is *commutative* if order does not matter, i.e., $a \star b = b \star a$, for any $a, b \in S$.

Definition 3.29. If $(G, \star)$ is a group and $\star$ is commutative, we say that $(G, \star)$ is an *abelian group*. In this case, often the operation is denoted by "+". We write $-x$ for x^{-1}, and 0 as the identity. Also, nx is used to denote x^n.

An exception to the above convention: $(\mathbb{Q} \setminus \{0\}, \cdot)$ is abelian.

Definition 3.30. Let $(G, \star)$, $(H, \cdot)$ be groups, a function $\varphi : G \to H$ is a *homomorphism* if $\varphi(g \star h) = \varphi(g) \cdot \varphi(h)$ for all $g, h \in G$.

Exercise 3.31. If $\varphi : G \to H$ is a homomorphism, where G, H are groups with identities e_G, e_H, then $\varphi(e_G) = e_H$, and for all $g \in G$, $\varphi\left(g^{-1}\right) = (\varphi(g))^{-1}$.

It's not hard to see that, for any $g \in G$, $\varphi(g) = \varphi(g \cdot e_G) = \varphi(g) \cdot \varphi(e_G)$, so $\varphi(e_G) = e_H$. A similar argument can be used to show the inverse part.

Given the definition of homomorphism, it is natural to ask: if we have two groups G, H, does there exist a homomorphism from G to H? The answer is yes. Simply map every element of G to e_H, then this is a valid homomorphism. Remarkably, one can show that this is the only valid homomorphism between certain groups.

Exercise 3.32. Consider $(\mathbb{Z}_n, +)$ as a group, show that there is no non-trivial homomorphism from $\mathbb{Z}_4$ to $\mathbb{Z}_7$.

We have seen the definition of isomorphism between two sets before, here we extend it to groups.

Definition 3.33. A bijective homomorphism from G to H is called an *isomorphism*. If an isomorphism from G to H exists, we say that G and H are *isomorphic*, denoted by $G \cong H$.

Exercise 3.34. Show that the inverse of an isomorphism from G to H is an isomorphism from H to G.

Recall that in Theorem 3.22, our argument has shown that the map from G to $\mathrm{Sym}(G)$ by $g \mapsto (h \mapsto gh)$ is an injective homomorphism. We abstract this intuition further.

Definition 3.35. Let $(G, \star)$ be a group. Given a subset $S \subseteq G$, define $\star_S : S \times S \to G$ by $a \star_S b = a \star b$. If $\star_S$ is a binary operation on S, i.e., $\mathrm{im}(\star_S) \subseteq S$, and $(S, \star_S)$ is a group, then $(S, \star_S)$ or S is called a *subgroup* of G, denoted by $S \leq G$.

Lemma 3.36. *Let G be a group, $S \subseteq G$, then $S \leq G$ if and only if the following conditions hold:*

1. $e \in S$
2. *if* $g \in S$, *then* $g^{-1} \in S$
3. *if* $g, h \in S$, *then* $gh \in S$

In particular, a concrete group on set X *is the same as a subgroup of* $\mathrm{Sym}(X)$.

Proof. ($\Leftarrow$) : By the three conditions, together with the associativity of the operation of G, S is a group, and S is closed under the group operation, so $S \leq G$.

($\Rightarrow$) : Suppose $S \leq G$, condition 3 holds by definition, it suffices to show the other two conditions.

1. We need to show $e_S = e$. The first observation is, for any element $g \in G$, if $g^2 = g$, then $g = e$. Why? Given $g^2 = g$, we can multiply both sides by g^{-1}, get $g = e$. Indeed, $e_S^2 = e_S$, so $e_S = e$.
2. We need to show that, if $g \in S$ and h is the inverse of g in S, then $h = g^{-1}$. Indeed, $gh = hg = e_S = e$, so $h = g^{-1}$ by the uniqueness of inverses, as desired. □

Proposition 3.37. *Suppose* $\varphi : G \to H$ *is a homomorphism, then* $\mathrm{im}(\varphi) \leq H$.

Proof. It suffices to check the three conditions in Lemma 3.36.

- Since $\varphi(e_G) = e_H$, we have $e_H \in \mathrm{im}(\varphi)$.
- Suppose $h \in \mathrm{im}(\varphi)$, then there exists some $g \in G$ such that $\varphi(g) = h$. Since $(\varphi(g))^{-1} = \varphi\left(g^{-1}\right)$, $h^{-1} \in \mathrm{im}(\varphi)$.
- Suppose $h_1, h_2 \in \mathrm{im}(\varphi)$, then there exist $g_1, g_2 \in G$ being mapped to them by φ. Since φ is a homomorphism, we have $h_1 h_2 = \varphi(g_1 g_2)$, so $h_1 h_2 \in \mathrm{im}(\varphi)$. □

Observe that if φ is injective, then trivially, we have $G \cong \mathrm{im}(\varphi)$.

3.4 Actions, cosets, orbits

Groups are interesting structures with nice properties we can leverage. Can we extend them to arbitrary sets? We do so by defining *actions*.

Definition 3.38. Let G be a group, an *action* on a set X is a homomorphism $\alpha : G \to \mathrm{Sym}(X)$.

For notation, we usually denote the domain and codomain of α as $\alpha : G \curvearrowright X$, and often, instead of $\alpha(g)(x)$, we write $g \cdot_\alpha x$ or $g \cdot x$ if α is clear from context.

Example 3.39.

1. Consider $\alpha : G \curvearrowright G$ given by $g \cdot h = gh$, this is called the *left multiplication action*.
2. $\alpha : G \curvearrowright G$ given by $g \cdot h = hg^{-1}$, this is called the *right multiplication action*. Let's verify it is an action. Suppose $g_1, g_2, h \in G$, then $(g_1 g_2) \cdot h = h(g_1 g_2)^{-1} = h g_2^{-1} g_1^{-1} = g_1 \cdot \left(h g_2^{-1}\right) = g_1 \cdot (g_2 \cdot h)$.

   Moreover, the map $G \to G : h \mapsto hg^{-1}$ is a bijection, we leave it to the reader to check this.
3. The trivial action: $G \curvearrowright X : g \cdot x = x$, for all $g \in G, x \in X$.

Definition 3.40. A *right action* of G on X is a function $\beta : G \to \mathrm{Sym}(X)$, such that for all $g_1, g_2 \in G$, $\beta(g_1 g_2) = \beta(g_2) \circ \beta(g_1)$. We usually write it as $\beta : X \curvearrowleft G$, and $x \cdot_\beta g$ or $x \cdot g$ for $\beta(g)(x)$. With this notation, the requirement of β can be written as $(x \cdot g_1) \cdot g_2 = x \cdot (g_1 g_2)$.

So we can define the *right multiplication action* via $h \cdot g = hg$, also, $h \cdot g = g^{-1} h$.

Definition 3.41. The *shift action* $\sigma : G \curvearrowright \mathcal{P}(G)$, defined as $g \cdot S := gS := \{gh : h \in S\}$.

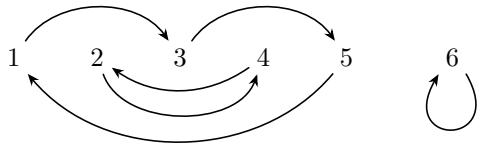
Notice that $\mathcal{P}(G) \cong \{0,1\}^G$, so we can view this action as $G \curvearrowright \{0,1\}^G$. More generally, there is a shift action $G \curvearrowright X^G$ for any set X: take $g \in G$, $f \in X^G$, then $g \cdot f \in X^G$ is a function $G \to X$, given by $(g \cdot f)(h) := f\left(g^{-1} h\right)$, for all $h \in G$. It is instructive for the reader to verify this is in fact an action.

Even more generally, given any action $G \curvearrowright Y$, we can "lift" it to an action $G \curvearrowright X^Y$: for $f \in X^Y$, define $g \cdot f : y \mapsto f\left(g^{-1} \cdot y\right)$.

An important type of action to study is actions of $(\mathbb{Z}, +)$. What does this kind of action look like? Let $\alpha : \mathbb{Z} \curvearrowright X$, then $\alpha(0) = \mathrm{id}_X$, suppose $\alpha(1) := f \in \mathrm{Sym}(X)$, then $\alpha(n) = f^n$, for all $n \in \mathbb{Z}$. This is true, since we can view $\mathbb{Z}$ as generated by 1, and since α is a homomorphism, it must preserve such a structure. On the other hand, for any $f \in \mathrm{Sym}(X)$, $n \mapsto f^n$ is a homomorphism $\mathbb{Z} \to \mathrm{Sym}(X)$, since $f^n \circ f^m = f^{n+m}$ by the exercise on powers above. Thus, there is a natural bijection between $\{\text{actions of } \mathbb{Z} \curvearrowright X\}$ and $\mathrm{Sym}(X)$, given by $\alpha \mapsto \alpha(1)$.

Example 3.42. Consider $\alpha : \mathbb{Z} \curvearrowright [6]$, given by $\alpha(1) = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 2 \\ 1 \\ 6 \end{bmatrix}$. We write it in a vector form, since essentially, elements of $\mathrm{Sym}([6])$ are permutations, so we just need to specify where each number

goes. The following diagram illustrates this action:



Definition 3.43. Let $\alpha : G \curvearrowright X, x \in X$. The *orbit* of x under α is the set $G \cdot_\alpha x = G \cdot x := \{g \cdot x : g \in G\}$.

Definition 3.44. The *orbit equivalence relation*, E_α, is the binary relation on set X, given by $xE_\alpha y \Leftrightarrow y \in G \cdot_\alpha x$.

Exercise 3.45. Verify this is an equivalence relation.

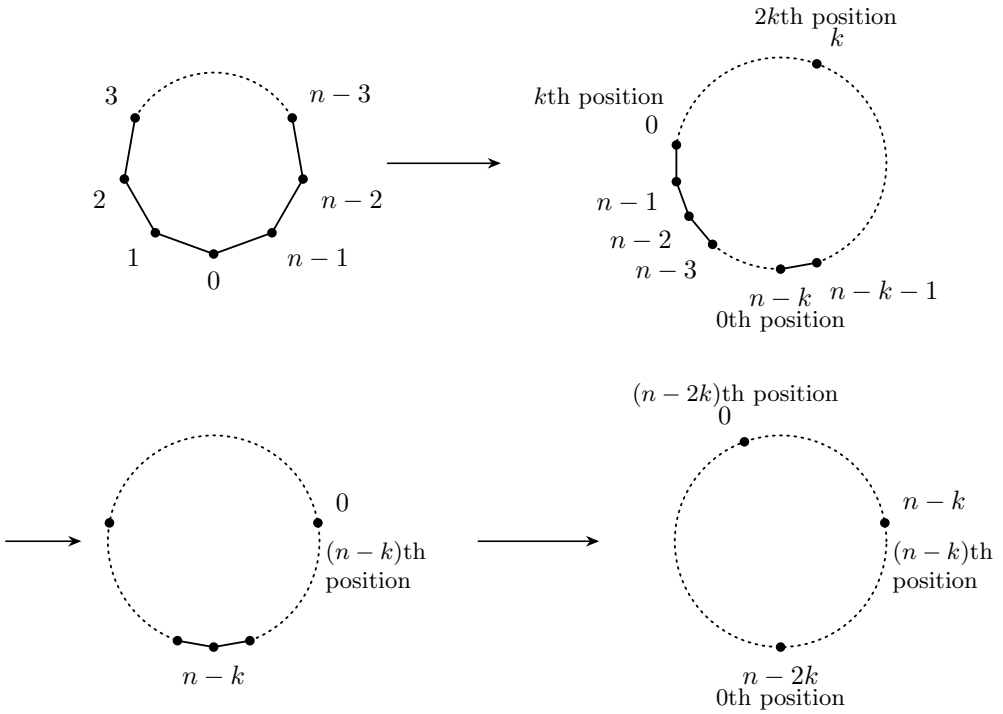
Exercise 3.46. For actions of $(\mathbb{Z}, +)$, orbits are of the following types: single element orbit, two-element orbit, three-element orbit, etc., and another type: a long chain that does not end.

The next example introduces an important type of action, which plays a major role in a bunch of concepts in abstract groups.

Example 3.47. The *conjugation action*, $\alpha_{conj} : G \curvearrowright G : g \cdot h := ghg^{-1}$. Let's verify it is a homomorphism: $\alpha_{conj}(g_1g_2)(h) = g_1g_2h(g_1g_2)^{-1} = g_1g_2hg_2^{-1}g_1^{-1} = g_1\alpha_{conj}(g_2)(h)g_1^{-1} = \alpha_{conj}(g_1) \circ \alpha_{conj}(g_2)(h)$. Recall that $\mathrm{Aut}(G) = \{\text{automorphisms of } G\} \leq \mathrm{Sym}(G)$, and the definition of automorphism can be found in item 7 of Example 3.2, i.e., functions that preserve compositions of elements. We can effectively view $\alpha_{conj} : G \to \mathrm{Aut}(G)$. What we need to argue is: for all $g \in G$, $\alpha_{conj}(g) : G \to G$ is a homomorphism, i.e., for any $h_1, h_2 \in G$, $\alpha_{conj}(g)(h_1h_2) = \alpha_{conj}(g)(h_1)\alpha_{conj}(g)(h_2)$.

$$\begin{aligned}\alpha_{conj}(g)(h_1h_2) &= gh_1h_2g^{-1}\\ &= gh_1g^{-1}gh_2g^{-1}\\ &= \left(gh_1g^{-1}\right)\left(gh_2g^{-1}\right)\\ &= \alpha_{conj}(g)(h_1)\alpha_{conj}(g)(h_2)\end{aligned}$$

Let's examine the conjugation action on the dihedral group D_{2n}. What is the orbit of h under conjugation? Clearly, $hhh = h$, and for any rotation g^k, $g^{-k}hg^k$ is as follows: first rotate 0 to k, and $n-k$ to 0, then flip it with respect to the fixed 0 axis, which results in k going to $n-k$, which can be viewed as the original 0 going to $n-k$. Finally, g^{-k} rotates $n-k$ to $n-2k$ and k to 0. It is a bit abstract to talk in this way, the following diagram illustrates the behavior:



One can view this as the flip with respect to the $n-k$ axis.

Similar to E_α, we use E_{conj} to denote the orbit equivalence relation under conjugation action, i.e., g and h are conjugate if and only if $gE_{conj}h$.

A quick observation on abelian groups: G is abelian if and only if the conjugation action is trivial.

For the forward implication, notice that $\alpha(g)(h) = ghg^{-1} = gg^{-1}h = h$. For the converse, $ghg^{-1} = h \to gh = hg$.

Definition 3.48. Let $\mathrm{Sub}(G) := \{H \subseteq G : H \leq G\}$, instead of conjugating elements, we can extend the definition by conjugating the entire subgroup, i.e., $\alpha_{conj} : G \curvearrowright \mathrm{Sub}(G)$ by $g \cdot H = gHg^{-1} := \left\{ghg^{-1} : h \in H\right\}$. Moreover, $g \cdot H$ is actually a subgroup of G, so we can further extend $E_{conj} :=$ the orbit equivalence relation on $\mathrm{Sub}(G)$. We say two subgroups H_1, H_2 are *conjugate* if $H_1 E_{conj} H_2$, i.e., $gH_1g^{-1} = H_2$ for some $g \in G$. This even gives an isomorphism between these two subgroups, the mapping is $h_1 \mapsto gh_1g^{-1}$, so $H_1 E_{conj} H_2 \Rightarrow H_1 \cong H_2$.

Definition 3.49. A subgroup $H \leq G$ is *normal* if $G \cdot_{conj} H = \{H\}$, notation: $H \trianglelefteq G$.

Example 3.50.

1. $\{e\}$, the trivial subgroup.
2. If G is abelian, then every subgroup is normal.
3. $G = D_{2n}$, the dihedral group. $\langle h \rangle = \{h, e\}$ is not a normal subgroup of D_{2n}, define $h_i :=$ flip with respect to axis i, then h is conjugate to h_i by g^i. So $\langle h \rangle$ is conjugate to $\langle h_i \rangle$, but they are not equal for $i = 1$: $h_1(0) = 2 \neq 0$ since $n \geq 3$, so $h_1 \notin \langle h \rangle$.

   On the other hand, $\langle g \rangle \trianglelefteq D_{2n}$ is a normal subgroup: any g^k acting on it results in the same group, and $h \cdot g^i = g^{n-i}$.

Definition 3.51. Let $\alpha : G \curvearrowright X, \beta : G \curvearrowright Y$, a *homomorphism* from α to β is a function $\varphi : X \to Y$ such that for any $g \in G, x \in X$, we have $\varphi(g \cdot_\alpha x) = g \cdot_\beta \varphi(x)$. Another term to describe φ is, it is G-equivariant. A diagram:

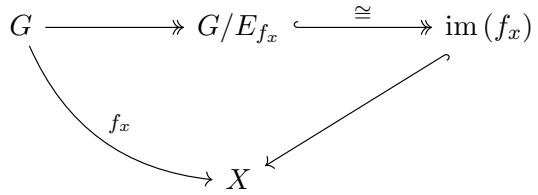
$$\begin{array}{ccc} X & \xrightarrow{\alpha(g)} & X \\ \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi} \\ Y & \xrightarrow{\beta(g)} & Y \end{array}$$

If φ is bijective, then φ is an *isomorphism*.

Example 3.52. Consider the following two $\mathbb{Z}$ actions on $[5]$, i.e., permutations on $[5]$: $\begin{bmatrix} 3 \\ 4 \\ 5 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ 4 \\ 5 \\ 3 \end{bmatrix}$, these two $\mathbb{Z}$ actions are isomorphic. How to figure it out? Notice in the first action, there are two orbits: $1-3-5$ and $2-4$, while in the second action, there are also two orbits: $1-2$ and $3-4-5$, so we can build a bijection between these two orbits. More generally, two $\mathbb{Z}$-actions α, β are isomorphic $\Leftrightarrow$ there exists a bijection $b : X/E_\alpha \to Y/E_\beta$, such that each orbit $O \in X/E_\alpha$ has the same type, or size, as $b(O)$, i.e., $|O| = |b(O)|$. To generalize this idea further, instead of saying the size of the orbit is the same, we formalize it using actions: for any group G, two actions α, β are isomorphic if and only if there exists a bijection $b : X/E_\alpha \to Y/E_\beta$ such that for each orbit $O \in X/E_\alpha$, the induced actions $G \curvearrowright O$ and $G \curvearrowright b(O)$ are isomorphic.

This helps to reduce the problem to studying actions with just one orbit, and such actions are called *transitive*.

In the next part, it is useful to recall Corollary 1.31, the canonical decomposition. Let $\alpha : G \curvearrowright X$ be an action, pick $x \in X$, define $f_x : G \to X : g \mapsto g \cdot x$. Apply Corollary 1.31 to f_x, we get the following diagram:



Where E_{f_x} is the equivalence kernel of f_x: Definition 1.26, i.e., $g_1 E_{f_x} g_2 \Leftrightarrow f_x(g_1) = f_x(g_2) \Leftrightarrow g_1 \cdot x = g_2 \cdot x \Leftrightarrow \left(g_2^{-1} g_1\right) \cdot x = x$. Let's leverage this special structure by defining the stabilizer of x.

Definition 3.53. The *stabilizer* of x is $\mathrm{Stab}(x) := \{g \in G : g \cdot x = x\}$.

Proposition 3.54. $\mathrm{Stab}(x) \leq G$.

Proof. Let's check the three properties in Lemma 3.36.

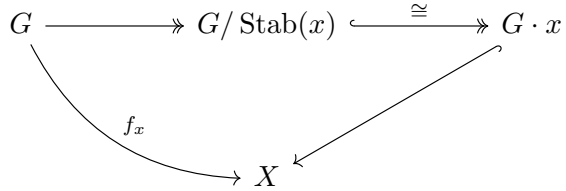
- First, $e \cdot x = x$. To see this, notice any action is a homomorphism, so $\alpha(e) = \mathrm{id}_X$, therefore, $e \cdot x = x$.
- If $g \cdot x = x$, then by the property of homomorphism, $g^{-1} \cdot x = g^{-1} \cdot (g \cdot x) = (g^{-1}g) \cdot x = e \cdot x = x$, so $g^{-1} \in \mathrm{Stab}(x)$.
- If g, h both fix x, then $(gh) \cdot x = g \cdot h \cdot x = g \cdot x = x$, so $gh \in \mathrm{Stab}(x)$. □

Definition 3.55. Let $H \leq G$, the *left coset equivalence relation* of H, denoted by LC_H on G, is given by $g_1 LC_H g_2 \Leftrightarrow g_2^{-1} g_1 \in H$.

LC_H is in fact an equivalence relation, and LC_H is the orbit equivalence relation under the right action $G \curvearrowleft H : g \cdot h = gh$. The orbits of this action are called *left-cosets*. We say that $g_1 \in$ orbit of $g_2 \Leftrightarrow g_2^{-1} g_1 \in H \Leftrightarrow g_1 \in g_2 H$. Consider $G/LC_H = \{gH : g \in G\}$, we use G/H as a shorthand notation for G/LC_H.

Similarly, *right-cosets* of H are the orbits of the left action $H \curvearrowright G$ by $h \cdot g = hg$, notation: $H \backslash G = G/RC_H = \{Hg : g \in G\}$.

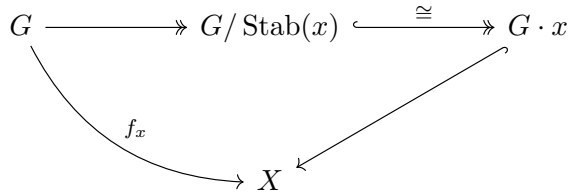
With this definition, we can write G/E_{f_x} as $G/\mathrm{Stab}(x)$. The next part in the diagram is $\mathrm{im}(f_x)$. What is $\mathrm{im}(f_x)$? $\mathrm{im}(f_x) = \{f_x(g) : g \in G\} = \{g \cdot x : g \in G\} = G \cdot x$, i.e., orbit of x. So we can complete the diagram as follows:



This is the *orbit-stabilizer theorem.*

3.5 Orbit-stabilizer theorem

Theorem 3.56 (Orbit-Stabilizer Theorem)**.** *Let G be a group, $G \curvearrowright X$ be an action, fix $x \in X$, let $f_x : G \to X : g \mapsto g \cdot x$. The following diagram holds:*



Proof. Direct application of the canonical decomposition theorem, with $G/E_{f_x} = G/\operatorname{Stab}(x)$ and $\operatorname{im}(f_x) = G \cdot x$. □

Let's further explore this diagram. Suppose our action is transitive, i.e., there is only one orbit, then $G \cdot x = X$. By orbit-stabilizer, it is isomorphic to $G/\operatorname{Stab}(x)$, i.e., G quotient a subgroup of G. This observation leads to the following theorem:

Theorem 3.57. *Let $\alpha : G \curvearrowright X$ be a transitive action, then α is isomorphic to an action $\beta : G \curvearrowright G/H$, for a subgroup $H \leq G$.*

Proof. The action we consider is given by $g \mapsto (aH \mapsto gaH)$, it is clearly an action: $(g_1 g_2) \cdot_\beta aH = g_1 g_2 aH = g_1 \cdot_\beta (g_2 aH) = g_1 \cdot_\beta (g_2 \cdot_\beta aH)$.

Let's fix an $x \in X$. Since α is transitive, $G \cdot_\alpha x = X$, and by Theorem 3.56, $G \cdot_\alpha x \cong G/\operatorname{Stab}(x)$. Set $H = \operatorname{Stab}(x)$, define $\varphi : G/H \to X : gH \mapsto g \cdot_\alpha x$. We will show φ is an isomorphism by showing it's a homomorphism and a bijection. But first of all, let's show φ is a well-defined function.

- Suppose $g_1 H = g_2 H$, we need to show that $g_1 \cdot_\alpha x = g_2 \cdot_\alpha x$. Since $g_1 H = g_2 H$, we have $g_2^{-1} g_1 H = H$, this means $g_2^{-1} g_1 \in H$. Why? Let's prove a lemma for this.

  Lemma 3.58. *Let H be a subgroup of G, and $g \in G$, then $gH = H \Rightarrow g \in H$.*

  Proof of Lemma. Let's do the contrapositive: if $g \notin H$, then $gH \neq H$. This is clear: since H is a subgroup, $e \in H$, so $ge = g \in gH$, but $g \notin H$, so $gH \neq H$. □

  So $g_2^{-1} g_1 \in H$, recall that $H = \operatorname{Stab}(x)$, which means $\left(g_2^{-1} g_1\right) \cdot_\alpha x = x \Rightarrow g_1 \cdot_\alpha x = g_2 \cdot_\alpha x$.

- Homomorphism: Let $y \in X$, our goal: $\varphi(g_1 \cdot_\beta g_2 H) = g_1 \cdot_\alpha \varphi(g_2 H)$. LHS is just $\varphi(g_1 g_2 H) = (g_1 g_2) \cdot_\alpha x$, while RHS is $g_1 \cdot_\alpha g_2 \cdot_\alpha x$, since α is an action, the identity holds.

- Bijection: Define $\phi : X \to G/H$ by $y \mapsto gH$, where $g \cdot_\alpha x = y$ (such a g exists since α is transitive). This is well-defined, since if $g_1 \cdot_\alpha x = g_2 \cdot_\alpha x = y$, then $g_2^{-1} g_1 \cdot_\alpha x = x$, which means $g_2^{-1} g_1 \in \operatorname{Stab}(x)$, so $g_1 \in g_2 H$. Similarly, $g_2 \in g_1 H$, this shows the mapping is well-defined. Moreover, it is clear that ϕ is the inverse of φ, so φ is a bijection. □

Theorem 3.59. *Given 2 subgroups $H_1, H_2 \leq G$, the actions $G \curvearrowright G/H_1, G \curvearrowright G/H_2$ given by $g \mapsto (aH_i \mapsto gaH_i)$ are isomorphic, if and only if H_1, H_2 are conjugate to each other.*

Proof. Exercise. □

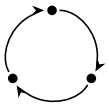
Combine Theorem 3.57 and Theorem 3.59, we obtain a 1-to-1 correspondence between

$$\{\text{isomorphism classes of transitive actions of } G\}$$

and

$$\{\text{conjugacy classes of subgroups of } G\}$$

Example 3.60. Consider $G = (\mathbb{Z}, +)$. In Exercise 3.46, we have studied orbits of $(\mathbb{Z}, +)$, now let's establish connections between isomorphism classes of transitive actions or "the shape of orbits" and conjugacy classes of subgroups of $\mathbb{Z}$.

Isomorphism classes of transitive actions	Conjugacy classes of subgroups of $\mathbb{Z}$
	$\mathbb{Z}$
	$2\mathbb{Z}$
	$3\mathbb{Z}$
$\vdots$	$\vdots$
an infinite chain	$\{0\}$

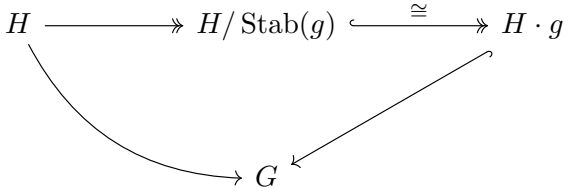
One way to think of this: an action of $\mathbb{Z}$ is determined by $\alpha(1) = f$, and the corresponding subgroup we choose is $\mathrm{Stab}(x)$, i.e., the integer powers k such that $f^k(x) = x$. If the orbit only contains one element, then any f^k has the property that $f^k(x) = x$, so $\mathrm{Stab}(x) = \mathbb{Z}$. Similarly, if orbits contain 2 elements, then $f^{2k}(x) = x$, so all even powers work. If orbit is an infinite chain, then any f^k for $k \neq 0$ will not become id_X, so only $\{0\}$ works.

Exercise 3.61. Recall: a group G is cyclic if $G = \langle g \rangle$ for some $g \in G$, i.e., G is generated by a single element. Show that every subgroup of a cyclic group is cyclic.

For this exercise, pick the least positive integer power of g in the subgroup, prove it generates the entire subgroup.

Next, we talk about some important applications and corollaries of orbit-stabilizer.

1. Let $H \leq G$, $H \curvearrowright G$ by left multiplication:$h \cdot g := hg$. Take some $g \in G$, consider the diagram:



The orbit $H \cdot g$ is the right-coset of H corresponding to g: Hg. $\mathrm{Stab}(g) = \{h \in H : hg = g\} =$

$\{e\}$, so we further have $Hg \cong H/\operatorname{Stab}(g) \cong H$, i.e., every right-coset of H is in bijection with H.

Theorem 3.62 (Lagrange)**.** *If G is a finite group, $H \leq G$, then $|G| = |H||G/H|$, where we usually use $|G : H|$ to denote $|G/H|$, called the index of H in G.*

Proof. By our above observation, each right coset of H is in bijection with H, and similarly, using the right action $G \curvearrowleft H : g \cdot h = gh$, whose orbits are the left cosets gH and whose stabilizers are trivial, each left coset is in bijection with H, which means they all have cardinality $|H|$. Moreover, either left-cosets or right-cosets partition G, and there are $|G/H|$ or $|H\backslash G|$ such cosets in total, so $|G| = |H||G/H|$. □

Corollary 3.63. $|H| \mid |G|$.

Proof. By Theorem 3.62, $|G| = |H||G/H|$, and $|G/H|$ is a positive integer. □

For instance, if $|G|$ is prime, then subgroups of G are either $\{e\}$ or G.

Exercise 3.64. Show that the only group of size p a prime, up to isomorphism, is $\mathbb{Z}_p$.

One can show such a group must be cyclic: suppose not, then you can construct another subgroup of size smaller than p, contradicts Theorem 3.62.

Corollary 3.65. *Let G be a finite group, $g \in G$.*

(a) $\{n \in \mathbb{N}^+ : g^n = e\} \neq \emptyset$

(b) Let $|g| := \min\{n \in \mathbb{N}^+ : g^n = e\}$*, called the order of g, then $g^n = e$ if and only if $|g| \mid n$.*

(c) $|g| \mid |G|$*, i.e., $g^{|G|} = e$.*

Proof. (a) Since G is finite, we cannot keep generating new elements by repeatedly powering g, so there must exist some positive $i, j \in \mathbb{N}^+$ with $j > i$, such that $g^j = g^i \Rightarrow g^{j-i} = e$.

(b) ($\Leftarrow$) : If $|g| \mid n$, then $n = k|g|$ for some integer k, so $g^n = g^{k|g|} = \left(g^{|g|}\right)^k = e^k = e$.
($\Rightarrow$) : Use the division algorithm. Suppose $|g| \nmid n$, so $n = k|g| + r$, for some $0 < r < |g|$, so $g^n = g^{k|g|+r} = g^{k|g|}g^r = eg^r = g^r = e$, however, $r < |g|$, so contradicts the definition of order.

(c) Consider the subgroup $\langle g \rangle$, it has size $|g|$: every element in $\langle g \rangle$ can be expressed as g^k, for $0 \leq k < |g|$ (write $g^n = g^{q|g|+r} = g^r$ with $0 \leq r < |g|$), and these $|g|$ powers are distinct: $g^i = g^j$ with $0 \leq i < j < |g|$ would give $g^{j-i} = e$ with $0 < j - i < |g|$. By Corollary 3.63, $|g| \mid |G|$, as desired.

□

Exercise 3.66. Let p be a prime number, consider $G = (\mathbb{Z}_p \setminus \{0\}, \cdot)$.

(a) Show this is an abelian group.

(b) Conclude that for every $a \in \mathbb{Z}$, $a^p \equiv a \pmod{p}$, this is *Fermat's little theorem.*

The first part is a little bit painstaking to show. The second part is just applying Corollary 3.65 to G (which has $p-1$ elements), with the case $p \mid a$ handled separately.

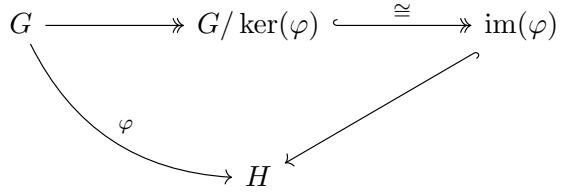
2. Let $\varphi : G \to H$ be a group homomorphism, it gives rise to an action $\alpha_\varphi : G \curvearrowright H$, given by $(\alpha_\varphi(g))(h) := \varphi(g)\,h$. Let e_H be the identity of H, the orbit of e_H is $\operatorname{im}(\varphi)$, since $G \cdot_\alpha e_H = \{\varphi(g) : g \in G\}$. $\operatorname{Stab}(e_H) = \{g \in G : \varphi(g)\, e_H = e_H\} = \{g \in G : \varphi(g) = e_H\}$, i.e., the preimage of $\{e_H\}$ under φ, we call it the *kernel* of φ, denoted by $\ker(\varphi)$.

   Theorem 3.67 (First Isomorphism Theorem)**.** *Let $\varphi : G \to H$ be a group homomorphism, then* $\operatorname{im}(\varphi) \cong G/\ker(\varphi)$.

   Proof. Apply orbit-stabilizer, with element e_H, so $\operatorname{Stab}(e_H) = \ker(\varphi)$, and $G \cdot e_H = \operatorname{im}(\varphi)$. So we get a bijection $\psi : G/K \to \operatorname{im}(\varphi)$, $gK \mapsto \varphi(g)$, where $K := \ker(\varphi)$. If $g_1K = g_1'K$ and $g_2K = g_2'K$, then $\psi((g_1g_2)K) = \varphi(g_1)\varphi(g_2) = \varphi(g_1')\varphi(g_2') = \psi((g_1'g_2')K)$, so $(g_1g_2)K = (g_1'g_2')K$ since ψ is injective, i.e., $(g_1K)(g_2K) := (g_1g_2)K$ is well-defined. Moreover, $\psi((g_1K)(g_2K)) = \psi(g_1K)\psi(g_2K)$, so ψ "copies" the group structure of $\operatorname{im}(\varphi)$ to G/K, i.e., G/K is a group and ψ is an isomorphism. We will look at this group structure on G/K in the next section. □

3.6 Normal subgroups and kernels

In this part we focus on normal subgroups and kernels of homomorphisms. Let $\varphi : G \to H$ be a homomorphism, recall Theorem 3.67, combine with Theorem 3.56, we have the following diagram:



In this diagram, $\operatorname{im}(\varphi)$ is a subgroup of H, by the isomorphism between $G/\ker(\varphi)$ and $\operatorname{im}(\varphi)$, we can "copy" the group structure and operation to $G/\ker(\varphi)$. This makes all the arrows in the diagram group homomorphisms. Let $K := \ker(\varphi)$, what is the group structure of G/K? For each choice $C_1, C_2 \in G/K$, let's do the most natural thing. Let $C_1 = g_1K, C_2 = g_2K$, then define $C_1C_2 := (g_1K)(g_2K) = (g_1g_2)K$. So the group operation on G/K is denoted by $(g_1K)(g_2K) = (g_1g_2)K$.

Theorem 3.68. *Let $K \leq G$, then there is a group operation on G/K defined by $(g_1K)(g_2K) = (g_1g_2)K$ if and only if K is a **normal** subgroup.*

Proof. ($\Leftarrow$) : Assume K is normal, we need to check that the group operation is well-defined, i.e., if $g_1K = g_1'K, g_2K = g_2'K$, then $(g_1g_2)K = (g_1'g_2')K$. Since $g_1K = g_1'K$, we have $g_1^{-1}g_1'K = K$, so $g_1^{-1}g_1' \in K$, similarly, $g_2^{-1}g_2' \in K$. Since K is normal, $K = g_2^{-1}Kg_2$, so $g_2^{-1}(g_1^{-1}g_1')g_2 \in K$, and therefore $g_2^{-1}g_1^{-1}g_1'g_2' = (g_2^{-1}g_1^{-1}g_1'g_2)(g_2^{-1}g_2') \in K$, which means $(g_1g_2)^{-1}(g_1'g_2') \in K$, therefore, $(g_1g_2)^{-1}(g_1'g_2')K = K \Rightarrow (g_1g_2)K = (g_1'g_2')K$, as desired.

($\Rightarrow$) : Suppose G/K is a group, we want to show that $K \trianglelefteq G$, this means for any $g \in G$, $K = gKg^{-1}$. Take $g \in G$.

($\subseteq$) : We can reuse some arguments in the converse part, i.e., suppose $g_1K = g_1'K, g_2K = g_2'K$, then $(g_1g_2)K = (g_1'g_2')K$, this means that $K = (g_1g_2)^{-1}(g_1'g_2')K = g_2^{-1}g_1^{-1}g_1'g_2'K$, so $g_2^{-1}g_1^{-1}g_1'g_2' \in K$, $g_1^{-1}g_1'g_2'g_2^{-1} \in g_2Kg_2^{-1}$. Moreover, this holds whenever $g_1^{-1}g_1' \in K$ and $g_2^{-1}g_2' \in K$, and we use it to show $K \subseteq gKg^{-1}$. Now, for any $k \in K$, let's take $g_1 = e$, $g_1' = k$ and $g_2 = g_2' = g$, then $g_1^{-1}g_1'g_2'g_2^{-1} = e^{-1}kgg^{-1} = k$, and clearly, $e^{-1}k \in K$ and $g^{-1}g = e \in K$, so we have $k \in gKg^{-1}$.

($\supseteq$) : Since $K \leq G, K \subseteq gKg^{-1}$, we have $Kg \subseteq gK \Rightarrow g^{-1}Kg \subseteq K$. Set $h = g^{-1}$, so we have $hKh^{-1} \subseteq K$, and since g is arbitrary, replacing g by g^{-1} gives $gKg^{-1} \subseteq K$, as desired. □

We abstract the argument in the $\supseteq$ part into a Lemma:

Lemma 3.69. *If $K \leq G$ satisfies $K \subseteq gKg^{-1}$ for all $g \in G$ (equivalently, $gKg^{-1} \subseteq K$ for all $g \in G$), then K is normal.*

In fact, if $\alpha : G \curvearrowright X$ and $Y \subseteq X$ such that $Y \subseteq g \cdot_\alpha Y$ for all $g \in G$, then $g \cdot_\alpha Y = Y$ for all $g \in G$.

Corollary 3.70. *A subgroup $K \leq G$ is a kernel of some homomorphism if and only if K is normal in G.*

Proof. ($\Rightarrow$) : Let $K = \ker(\varphi)$ with $\varphi : G \to H$, by Theorem 3.67, $G/K \cong \operatorname{im}(\varphi)$ as groups. By Theorem 3.68, $K \trianglelefteq G$.

($\Leftarrow$) : Suppose K is normal, by Theorem 3.68, G/K is a group and $G \twoheadrightarrow G/K$, the quotient map, is a homomorphism, we claim the kernel is K. What is the identity of G/K? It is K, since

$(gK)(eK) = (ge)K = gK$. Moreover, what elements $g \in G$ satisfy $gK = K$? Exactly the elements in K. □

Let $\{*\}$ be a one-element trivial group, if $K \trianglelefteq G$, consider the following sequence:

$$\{*\} \longrightarrow K \hookrightarrow G \twoheadrightarrow G/K \longrightarrow \{*\}$$

In this sequence, the image of every homomorphism is equal to the kernel of the next.

- $\operatorname{im}(\{*\} \to K) = \{e_K\}$, $\ker(K \hookrightarrow G) = \{e_K\}$
- $\operatorname{im}(K \hookrightarrow G) = K$, $\ker(G \twoheadrightarrow G/K) = K$
- $\operatorname{im}(G \twoheadrightarrow G/K) = G/K$, $\ker(G/K \to \{*\}) = G/K$

Definition 3.71. The sequence with the above property is called an *exact* sequence.

Exercise 3.72. Let K, G, H be groups, suppose we have a sequence of homomorphisms which is exact:

$$\{*\} \longrightarrow K \longrightarrow G \longrightarrow H \longrightarrow \{*\}$$

Show that $K \cong K' \trianglelefteq G, H \cong G/K'$.

Exercise 3.73 (important)**.** Suppose $\varphi : G \to H$ is a homomorphism, then

1. φ is surjective $\Leftrightarrow \operatorname{im}(\varphi) = H$
2. φ is injective $\Leftrightarrow \ker(\varphi) = \{e_G\}$

Both parts are not hard to show, but the result is very useful.

Definition 3.74. A group G is *simple* if the only normal subgroups of G are G and $\{e\}$.

Exercise 3.75. $(\mathbb{Z}_n, +_n)$ is simple if and only if n is prime.

Notice $(\mathbb{Z}_n, +_n)$ is abelian, so every subgroup is normal. If n is not prime, then it has subgroups other than G and $\{e\}$, so it's not simple.

3.7 Groups and counting

In this section, we use the structure of groups to facilitate counting, especially, we study the famous necklace problem.

Lemma 3.76 (Burnside)**.** *Let $\alpha : G \curvearrowright X$ be an action of a finite group G, on a finite set X. Then*

$$\underbrace{|X/E_\alpha|}_{\text{number of orbits}} = \frac{1}{|G|} \sum_{g \in G} |\underbrace{\{x \in X : g \cdot x = x\}}_{\text{points fixed by } g}|$$

Before proving the lemma, let's first look at an example.

Example 3.77. Suppose $\alpha : G \curvearrowright X$ is an action such that for any $x \in X$, $\operatorname{Stab}(x) = \{e\}$, such actions are called *free*. By orbit-stabilizer, every orbit is isomorphic to $G/\operatorname{Stab}(x) \cong G$, so every orbit has size $|G|$, and there are $\frac{|X|}{|G|}$ orbits in total. For each $g \in G$, $|\{x \in X : g \cdot x = x\}| = \begin{cases} 0, & \text{if } g \neq e \\ |X|, & \text{if } g = e \end{cases}$, so RHS is also $\frac{|X|}{|G|}$.

Proof. Let's do some calculations.

$$\begin{aligned} |X/E_\alpha| &= \sum_{O \in X/E_\alpha} 1 \\ &= \sum_{O \in X/E_\alpha} \sum_{x \in O} \frac{1}{|O|} \\ &= \sum_{x \in X} \frac{1}{|G \cdot x|} && \text{orbits partition } X \end{aligned}$$

By orbit-stabilizer, $G \cdot x \cong G/\operatorname{Stab}(x)$, so $|G \cdot x| = |G/\operatorname{Stab}(x)| = |G|/|\operatorname{Stab}(x)|$, where the last equality is due to Theorem 3.62, so we can write $|X/E_\alpha|$ as:

$$\begin{aligned} |X/E_\alpha| &= \sum_{x \in X} \frac{1}{|G \cdot x|} \\ &= \sum_{x \in X} \frac{|\operatorname{Stab}(x)|}{|G|} \\ &= \frac{1}{|G|} \sum_{x \in X} |\operatorname{Stab}(x)| \\ &= \frac{1}{|G|} \sum_{x \in X} \sum_{g \in G} \delta(g \cdot x = x) && \delta(\cdot) \text{ is the indicator function} \\ &= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X} \delta(g \cdot x = x) \\ &= \frac{1}{|G|} \sum_{g \in G} |\{x \in X : g \cdot x = x\}| && \square \end{aligned}$$

Example 3.78. How many different types of necklaces with n gems of k kinds are there, if the necklaces are the same up to rotation? Let's first look at $n = 4, k = 2$:

$n = 4,\ k = 2$: ①  ②  ③  ④  ⑤  ⑥ 

Let G : all rotations, $X : k^n$ configurations of necklaces, so the problem reduces to counting the number of orbits of the action $G \curvearrowright X$ by rotating one configuration.

In this setting, $|G| = n$: we can rotate one configuration clockwise up to $n-1$ steps. $|X| = k^n$, so in general, number of orbits $= \frac{1}{n} \sum_{s \in \mathbb{Z}_n} |\{x \in X : s \cdot x = x\}|$. Let's examine our $n = 4, k = 2$ case.

- For $s = 0$, all $2^4 = 16$ configurations are fixed.
- For $s = 1$, only 2 configurations, namely, all squares or all triangles are fixed points.
- For $s = 2$, diagonals must be the same, so 4 configurations.
- For $s = 3$, only 2 configurations.

By Lemma 3.76, the count is $\frac{1}{4}(16 + 2 + 4 + 2) = 6$.

Exercise 3.79. Show that $|\{x \in X : s \cdot x = x\}| = k^{\gcd(n,s)}$.

Exercise 3.80. What is the count, if you can flip the necklace?

For the second exercise, choose $G = D_{2n}$.

Corollary 3.81 (Pólya)**.** *Let G be a finite group, X, Y finite sets, and $\alpha : G \curvearrowright X$ be an action. For each $g \in G$, let $c(g)$ be the number of cycles in the permutation $\alpha(g)$ on X, i.e., the number of orbits in the action $\mathbb{Z} \curvearrowright X : 1 \mapsto \alpha(g)$. The action α gives rise to an action $G \curvearrowright Y^X$ by $(g \cdot f)(x) := f\left(g^{-1} \cdot x\right)$, i.e., the shift action. Then the number of orbits of this action is equal to $\frac{1}{|G|} \sum_{g \in G} |Y|^{c(g)}$.*

Proof.

$$\begin{aligned}
\text{number of orbits} &= \frac{1}{|G|} \sum_{g \in G} |\{f \in Y^X : g \cdot f = f\}| \\
&= \frac{1}{|G|} \sum_{g \in G} |\{f \in Y^X : f(x) = f\left(g^{-1} \cdot x\right), \forall x \in X\}| && \text{this means } f \text{ is constant on cycles in } \alpha(g) \\
&= \frac{1}{|G|} \sum_{g \in G} |Y|^{c(g)} && \square
\end{aligned}$$

4 Categories

A category is a more abstract definition of "collections" of objects. Categories are useful because *the same definition* can work in several categories.

4.1 Basic definitions & examples

Definition 4.1. A *category* $\mathcal{C}$ consists of

- a class $\mathrm{Obj}(\mathcal{C})$ of *objects*
- for all $A, B \in \mathrm{Obj}(\mathcal{C})$, a set $\mathrm{Hom}_{\mathcal{C}}(A, B)$, or simply $\mathrm{Hom}(A, B)$ of *morphisms* from A to B
- for all $A, B, C \in \mathrm{Obj}(\mathcal{C})$, a function $\circ : \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$, called *composition*, such that:

  1. composition is associative: Let $A, B, C, D \in \mathrm{Obj}(\mathcal{C})$, $f \in \mathrm{Hom}(A, B)$, $g \in \mathrm{Hom}(B, C)$, $h \in \mathrm{Hom}(C, D)$, then
  $$h \circ (g \circ f) = (h \circ g) \circ f$$
  2. identities: For each $A \in \mathrm{Obj}(\mathcal{C})$, there is a $1_A \in \mathrm{Hom}(A, A)$, such that for all $B \in \mathrm{Obj}(\mathcal{C}), f \in \mathrm{Hom}(A, B), g \in \mathrm{Hom}(B, A)$:
  $$f \circ 1_A = f, 1_A \circ g = g$$
  Notice for fixed A, 1_A is unique: duplicate the argument that proves the identity of a binary operation is unique.

Example 4.2. 1. The most basic example is the category $\underline{\mathrm{Set}}$ of sets:

- $\mathrm{Obj}(\underline{\mathrm{Set}})$ = all sets
- $\mathrm{Hom}_{\underline{\mathrm{Set}}}(A, B)$ = functions from A to B
- $\circ$ = composition of functions

2. One can obtain different categories by taking fewer morphisms:

   (a)
   - Obj = all sets
   - $\mathrm{Hom}(A, B)$ = bijections from A to B
   - $\circ$ = composition

   (b)
   - Obj = all sets
   - $\mathrm{Hom}(A, B) = \begin{cases} \emptyset, & \text{if } A \not\subseteq B \\ \{\text{the inclusion map } A \hookrightarrow B\}, & \text{if } A \subseteq B \end{cases}$
   - $\circ$ = composition

3. $\mathrm{Obj} = \mathbb{Z}, \mathrm{Hom}(n, m) = \begin{cases} \emptyset, & \text{if } n > m \\ \text{one element set } \{f_{nm}\}, & \text{if } n \leq m \end{cases}$

   For composition, we need $\circ : \mathrm{Hom}(n, m) \times \mathrm{Hom}(k, n) \to \mathrm{Hom}(k, m)$.

- If $k \leq n \leq m$, $f_{nm} \circ f_{kn} := f_{km}$.
- Otherwise, there's nothing to define: simply $\emptyset$.

4. The category of graphs, $\underline{\text{Graph}}$:
   - $\text{Obj}\left(\underline{\text{Graph}}\right) = \{\text{all graphs}\}$
   - $\text{Hom}\left(G, H\right) = \{\text{all homomorphisms from } G \text{ to } H\}$. Let's recall the definition of a homomorphism of graphs: it is a function $f : V\left(G\right) \to V\left(H\right)$ such that if $\{x, y\} \in E\left(G\right)$, then $\{f\left(x\right), f\left(y\right)\} \in E\left(H\right)$.
   - $\circ$ = composition
5. The category of groups, $\underline{\text{Grp}}$:
   - $\text{Obj}\left(\underline{\text{Grp}}\right) = \{\text{all groups}\}$
   - $\text{Hom}\left(G, H\right) = \{\text{all group homomorphisms from } G \text{ to } H\}$
   - $\circ$ = composition

4.2 Monomorphism, epimorphism

In this section, we introduce some general concepts for different categories, which are useful in understanding more about groups.

Definition 4.3. Let $\mathcal{C}$ be a category, and $A, B \in \mathrm{Obj}(\mathcal{C})$.

- $f \in \mathrm{Hom}(A, B)$ is a *monomorphism* if for all $C \in \mathrm{Obj}(\mathcal{C})$ and $g, h \in \mathrm{Hom}(C, A)$,

$$f \circ g = f \circ h \Rightarrow g = h$$

- Similarly, f is an *epimorphism* if for all $C \in \mathrm{Obj}(\mathcal{C})$ and $g, h \in \mathrm{Hom}(B, C)$,

$$g \circ f = h \circ f \Rightarrow g = h$$

In the category $\underline{\mathrm{Set}}$, it is not hard to show that

- monomorphism $\Leftrightarrow$ injective
- epimorphism $\Leftrightarrow$ surjective

What about in $\underline{\mathrm{Grp}}$? Clearly, injectivity implies monomorphism. What about monomorphism implies injectivity?

Proof. Suppose $f : A \to B$ is a group homomorphism, take $a, a' \in A$ with $f(a) = f(a')$. Then there are homomorphisms $g, h : \mathbb{Z} \to A$ given by $g(n) := a^n, h(n) := (a')^n$. Let's verify f and g, h satisfy the monomorphism condition:

$$\begin{aligned}
(f \circ g)(n) &= f(g(n)) \\
&= f(a^n) \\
&= f(a)^n && \text{since } f \text{ is a homomorphism} \\
&= f\left(a'\right)^n \\
&= f\left(\left(a'\right)^n\right) \\
&= f(h(n)) \\
&= (f \circ h)(n)
\end{aligned}$$

So $f \circ g = f \circ h$, combined with f being a monomorphism, we have $g = h \Rightarrow g(1) = h(1) \Rightarrow a = a'$, therefore, f is injective. □

So indeed, monomorphism in $\underline{\mathrm{Grp}}$ $\Leftrightarrow$ injective.

Note: in this proof, we have used the property of integers: for every group A and $a \in A$, there is a unique homomorphism $\mathbb{Z} \to A$ sending 1 to a.

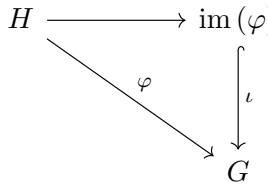
What about epimorphisms? Surjectivity implies epimorphisms. How about the converse? In the category of $\underline{\mathrm{Grp}}$, the converse is true, however, it is a non-trivial task to show.

Theorem 4.4 (Schreier)**.** *If a group homomorphism $\varphi : H \to G$ is an epimorphism in the category $\underline{\mathrm{Grp}}$, then φ is surjective.*

Notice it is not true in every category.

Example 4.5. Consider the category of subsets of $\mathbb{R}^n$, $n \in \mathbb{N}$, with continuous functions as morphisms. The embedding $\iota : \mathbb{Q} \hookrightarrow \mathbb{R}$ is, surprisingly, an epimorphism. Given $g, h : \mathbb{R} \to \mathbb{R}^n$ as continuous functions, suppose $g \circ \iota = h \circ \iota$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, a continuous function on $\mathbb{R}$ is determined by its values on rational points. Therefore, $g = h$, and ι is an epimorphism. But clearly, ι is not surjective.

Proof of Schreier. We can first decompose φ into the following diagram:



The claim is, if φ is an epimorphism, then so is ι. Let $g, h : G \to K$ be group homomorphisms, if $g \circ \iota = h \circ \iota$, then $g \circ \varphi = h \circ \varphi$: for any $x \in H$, let $t = \varphi(x)$, then $g(\varphi(x)) = (g \circ \iota)(t) = (h \circ \iota)(t) = h(\varphi(x))$. So $g = h$ since φ is an epimorphism. Therefore, it suffices to consider the case where $H \leq G$, and $\varphi : H \hookrightarrow G$ is the inclusion map. Since φ is the inclusion map, φ is surjective if and only if $H = G$, and we will show this.

Lemma 4.6. *If H is normal in G, then $H = G$.*

Proof of Lemma. Consider the two homomorphisms from G to G/H:

- the quotient homomorphism $q : G \twoheadrightarrow G/H$
- the trivial homomorphism $t : G \to G/H : g \mapsto H$

The claim is $q \circ \varphi = t \circ \varphi$, since for all $h \in H$, we have $(q \circ \varphi)(h) = q(h) = H$, and $(t \circ \varphi)(h) = t(h) = H$. Since φ is an epimorphism, we have $q = t$, i.e., the quotient map is nothing but the trivial map, this means $G/H = \{H\}$, i.e., $G = H$, as desired. □

Corollary 4.7. *In such a scenario, we can assume $|G/H| \geq 3$.*

Proof of Corollary. If $|G/H| = 1$, then $G = H$. Otherwise, if $|G/H| = 2$, then there are exactly two cosets: H and gH, for some $g \in G$. Suppose H is not normal, i.e., there is some $a \in G$ such that $aHa^{-1} \neq H$, which means, after replacing a by a^{-1} if necessary, there is some $h \in H$, with $aha^{-1} \notin H$. Clearly, $a \notin H$, so we can write $a = g\ell$, for some $\ell \in H$, and $aha^{-1} \notin H$ implies that $aha^{-1} = gh'$, for some $h' \in H$. Multiply both sides by g^{-1}, we get $g^{-1}aha^{-1} = h' \Rightarrow g^{-1}g\ell h\ell^{-1}g^{-1} = h' \Rightarrow \ell h\ell^{-1}g^{-1} = h'$, since $\ell, h, h' \in H$, it must be the case that $g^{-1} \in H \Rightarrow g \in H$, contradicts that gH and H partition G. Thus, H must be normal, and then $H = G$ by the lemma. □

Toward a contradiction, let's assume $H \neq G$, so H cannot be normal. By the above corollary, it's safe to say $|G/H| \geq 3$. We will show that in this case, it's possible to construct two homomorphisms $\alpha, \beta : G \to \mathrm{Sym}(G)$, such that $\alpha \circ \varphi = \beta \circ \varphi$, but $\alpha \neq \beta$, which contradicts that φ is an epimorphism. Define $\alpha =$ left multiplication action, and $\beta = \alpha$ conjugated by "swapping" two of the cosets of H. More specifically, since $|G/H| \geq 3$, there are two elements $u, v \in G$ such that H, Hu, Hv are all

distinct: the map $gH \mapsto Hg^{-1}$ is a well-defined bijection from left cosets to right cosets, so H also has at least 3 right cosets. Let $\pi : G \to G$ be the bijection such that

$$\pi(g) := \begin{cases} hv, & \text{if } g = hu \text{ for some } h \in H \\ hu, & \text{if } g = hv \text{ for some } h \in H \\ g, & \text{if } g \notin Hu \cup Hv \end{cases}$$

Define an action $\beta : G \curvearrowright G$ by $\beta(g) := \pi \circ \alpha(g) \circ \pi^{-1} = \pi \circ \alpha(g) \circ \pi$. We need to check α, β agree on H. Take $h, h' \in H$, then

- $\alpha(h)(h') = hh', \beta(h)(h') = \pi \circ \alpha(h) \circ \pi(h') = \pi \circ \alpha(h)(h') = \pi(hh') = hh'$
- $\alpha(h)(h'u) = hh'u, \beta(h)(h'u) = \pi \circ \alpha(h) \circ \pi(h'u) = \pi \circ \alpha(h)(h'v) = \pi(hh'v) = hh'u$
- $\alpha(h)(h'v) = hh'v, \beta(h)(h'v) = \pi \circ \alpha(h) \circ \pi(h'v) = \pi \circ \alpha(h)(h'u) = \pi(hh'u) = hh'v$
- if $g \notin H \cup Hu \cup Hv$, then $hg \notin Hu \cup Hv$ either since $Hhg = Hg$, so $\alpha(h)(g) = hg$ and $\beta(h)(g) = \pi(\alpha(h)(\pi(g))) = \pi(hg) = hg$

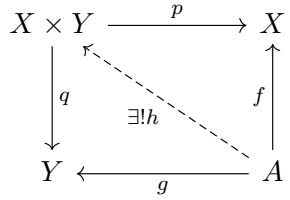
Thus, $\alpha|_H = \beta|_H \Rightarrow \alpha \circ \varphi = \beta \circ \varphi$, but $\alpha \neq \beta$, in particular, $\alpha(u) \neq \beta(u) : \alpha(u)(e) = u, \beta(u)(e) = v$. This completes the contradiction, and thus completes the proof of the theorem. $\square$

5 Advanced Topics on Groups

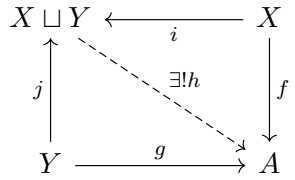
In this section, we study more advanced topics on groups, including products & coproducts, free groups, and Sylow groups.

5.1 Product & coproduct of groups

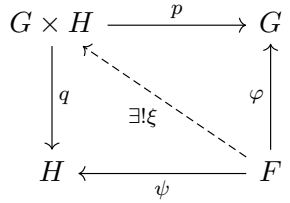
Before studying products & coproducts of groups, let's first recall them in Set. It is best to illustrate the definitions of products & coproducts with diagrams. The first is what we need for the product of sets:



The second one is what we need for the coproduct of sets, and the coproduct is sometimes called *disjoint union*:



In Grp: Given groups G, H, we want a group $G \times H$ and homomorphisms $p : G \times H \to G, q : G \times H \to H$, such that



As a set, $G \times H$ is just the Cartesian product of G and H, where operations are defined as $(g_1, h_1) \cdot (g_2, h_2) := (g_1 g_2, h_1 h_2)$, i.e., pointwise multiplication. Define homomorphisms p, q as $p(g, h) = g, q(g, h) = h$. Given homomorphisms $\varphi : F \to G, \psi : F \to H$, there exists a unique $\xi : F \to G \times H$, such that $p \circ \xi = \varphi, q \circ \xi = \psi$, by $\xi(f) = (\varphi(f), \psi(f))$.

Definition 5.1. $G \times H$ is called the *direct product of G and H*.

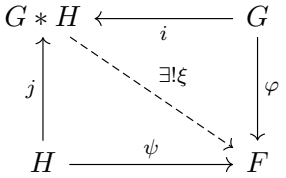
Example 5.2.

1. $G \times$ trivial group $\cong G$
2. $\mathbb{Z} \times \mathbb{Z} = \mathbb{Z}^2$
3. $\mathbb{Z}_3 \times \mathbb{Z}_7 \cong \mathbb{Z}_{21}$

Exercise 5.3. Show that if $n, m \geq 2$, $\gcd(m, n) = 1$, then $\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{nm}$.

However, the condition $\gcd(m, n) = 1$ cannot be dropped: for example, $\mathbb{Z}_2 \times \mathbb{Z}_2 \ncong \mathbb{Z}_4$, since $\mathbb{Z}_4$ is cyclic but the former is not.

Now let's switch gears to look at coproducts. In Grp, what we want is:



Recall that in Set: Example 2.8, we have shown that if $\operatorname{im}(i) \cap \operatorname{im}(j) = \emptyset$, then we can simply take the union of G and H. However, $\operatorname{im}(i) \cap \operatorname{im}(j) \neq \emptyset$, since $i(e_G) = j(e_H)$. The question is, is there any group K such that there are homomorphisms $i : G \to K, j : H \to K$, i, j are injective, and $\operatorname{im}(i) \cap \operatorname{im}(j) = \{e_K\}$? Consider $K = G \times H$, and $i : g \mapsto (g, e_H), j : h \mapsto (e_G, h)$.

Proposition 5.4. *In the category Ab, a.k.a. abelian groups, $G \times H$ is the coproduct of G and H.*

Proof. Take an abelian group F with homomorphisms $\varphi : G \to F, \psi : H \to F$, we want to show there exists a unique $\xi : G \times H \to F$ such that the diagram holds. Simply define $\xi(g, h) := \varphi(g) + \psi(h)$, this is a homomorphism, since

$$\begin{aligned}
\xi((g_1, h_1) + (g_2, h_2)) &= \xi(g_1 + g_2, h_1 + h_2) \\
&= \varphi(g_1 + g_2) + \psi(h_1 + h_2) \\
&= \varphi(g_1) + \varphi(g_2) + \psi(h_1) + \psi(h_2) \\
&= \varphi(g_1) + \psi(h_1) + \varphi(g_2) + \psi(h_2) \\
&= \xi(g_1, h_1) + \xi(g_2, h_2)
\end{aligned}$$

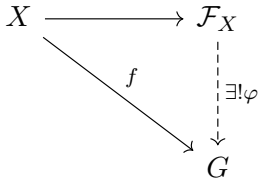
Notice that $\xi(g, 0) = \varphi(g) + \psi(0) = \varphi(g)$, so $\xi \circ i = \varphi$, and similarly $\xi \circ j = \psi$. We leave the uniqueness part as an exercise to the reader. □

So in abelian groups, we call coproducts *direct sums*.

What about Grp? We need to introduce the notion of *free groups*.

5.2 Free groups

What are free groups? We use a diagram to introduce the idea. Let X be a set and G be a group, then the free group $\mathcal{F}_X$ is the group that satisfies the following diagram:



Let's first do the easy case, say $X = \{x\}$, then we claim $\mathcal{F}_X \cong \mathbb{Z}$, and $\varphi : \mathbb{Z} \to G : n \mapsto (f(x))^n$, and for the map from $\{x\}$ to $\mathcal{F}_X$, we simply map x to 1. Given such a setting, φ is clearly unique, and is also a homomorphism.

What if $X = \{x, y\}$, in such a scenario, what should $\mathcal{F}_{\{x,y\}}$ look like? To address this problem, we need to introduce the concept of *reduced words*.

Definition 5.5. Let X be a set (generators), define

$$X' = \left\{x' : x \in X \text{ is associated with } x', \text{ serve as } \textit{inverse}\right\}$$

The *words* over $X \cup X'$ are simply $(X \cup X')^*$, where A^* is the *Kleene star*, i.e., all words over alphabet A with finite length. The *reduced words* are a subset of words over $X \cup X'$, with no x and x' adjacent to each other.

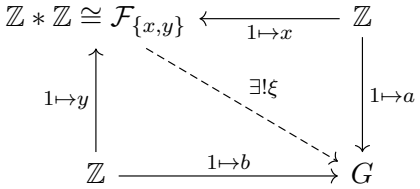
Definition 5.6. The *free group* over a set X, denoted by $\mathcal{F}_X$, is defined as the set of all reduced words over $X \cup X'$. The group operation is concatenation+reduction, i.e., for two reduced words w_1, w_2, we first concatenate them, then cancel out all adjacent x and x'.

Exercise 5.7. Show that such an operation is associative. In other words, the result of reducing a word is independent of the order of reductions.

For simplicity, we write x^{-1} for x'.

Whenever G is a group, $f : X \to G$, then there exists a unique homomorphism $\varphi : \mathcal{F}_X \to G$ that completes the diagram in Section 5.2: let $w = x_1 x_2 \dots x_n$ be a reduced word, define $\varphi(w) = f(x_1) f(x_2) \dots f(x_n)$, and define $f\left(x^{-1}\right) := (f(x))^{-1}$.

What is the coproduct $\mathbb{Z} * \mathbb{Z}$?



Since $\varphi : \mathbb{Z} \to G, \psi : \mathbb{Z} \to G$, we can abstract them as $1 \mapsto a, 1 \mapsto b$, so ξ is defined as $\xi(x) = \varphi(1) = a, \xi(y) = \psi(1) = b$.

Example 5.8. Consider the dihedral group, $D_{2n} = \langle g, h \rangle$, where $g^n = e, h^2 = e, hghg = e$.

Proposition 5.9. *Suppose G is a group, $x, y \in G$ such that $x^n = y^2 = yxyx = e$, then there exists a unique homomorphism: $\varphi : D_{2n} \to G$, such that $\varphi(g) = x, \varphi(h) = y$.*

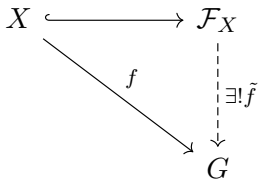
Proof. Since D_{2n} is generated by g, h, every element of D_{2n} is of the form $g^{m_1}h^{\ell_1}\dots g^{m_k}h^{\ell_k}$, where $m_i, \ell_i \in \mathbb{Z}$. Then $\varphi\left(g^{m_1}h^{\ell_1}\dots g^{m_k}h^{\ell_k}\right) = x^{m_1}y^{\ell_1}\dots x^{m_k}y^{\ell_k}$, this proves uniqueness.

We need to show φ is in fact a function, and it suffices to check if $g^{m_1}h^{\ell_1}\dots g^{m_k}h^{\ell_k} = e_{D_{2n}}$, then $x^{m_1}y^{\ell_1}\dots x^{m_k}y^{\ell_k} = e_G$. We proceed by cleaning up the exponents.

1. Get rid of negative powers of g by $g^{n-1} = g^{-1}, x^{n-1} = x^{-1}$
2. Get rid of negative powers of h by $h^{-1} = h, y^{-1} = y$
3. $hghg = e \Rightarrow ghg = h \Rightarrow hg = g^{n-1}h, yx = x^{n-1}y$, so we can move all g in front of h, and all x in front of y

This results in a form of $g^m h^\ell$, where $0 \le m < n, 0 \le \ell \le 1$, and we have $x^m y^\ell$ as well, since $g^m h^\ell = e_{D_{2n}}$, we must have $m = \ell = 0 \Rightarrow x^0 y^0 = e_G$, as desired. □

If $w = x_1 \dots x_n$ is a word, let $\tilde{f}(w) := f(x_1)\dots f(x_n)$, so our diagram becomes:



Definition 5.10. Let R be a set of reduced words over X, i.e., $R \subseteq \mathcal{F}_X$, call the elements of R *relations*. Given a group G and $f : X \to G$, say (G, f) *satisfies* R if $\tilde{f}(w) = e_G$, for all $w \in R$.

Example 5.11. Let $X = \{x, y\}, G = D_{2n}, f : X \to D_{2n} : (x \mapsto g, y \mapsto h)$, then (D_{2n}, f) satisfies $\left\{x^n, y^2, xyxy\right\}$.

Note that, if (G, f) satisfies R and $\varphi : G \to H$ is a homomorphism, then $(H, \varphi \circ f)$ also satisfies R, since $\widetilde{\varphi \circ f}(w) = \varphi\left(\tilde{f}(w)\right) = \varphi(e_G) = e_H$.

Definition 5.12. Let G be a group, X be a set, $f : X \to G$, $R \subseteq \mathcal{F}_X$, we write $(G, f) \cong \langle X \mid R\rangle$, if

1. (G, f) satisfies R.
2. If there exists another (H, g) that satisfies R, then there exists a unique homomorphism $\varphi : G \to H$, such that $g = \varphi \circ f$.

We call $\langle X \mid R\rangle$ a *presentation* of G.

Exercise 5.13. If $(G, f), (H, g) \cong \langle X \mid R\rangle$, show that there exists a unique isomorphism $\varphi : G \to H$, such that $g = \varphi \circ f$.

As a hint, use the homomorphism given by definition of $(G, f) \cong \langle X \mid R\rangle$, and show the homomorphism given by $(H, g) \cong \langle X \mid R\rangle$ is its inverse.

Example 5.14.
1. $\mathcal{F}_X \cong \langle X \mid \emptyset\rangle = \langle X \mid\rangle$
2. $\langle \{x\} \mid \{x^n\}\rangle = \langle x \mid x^n\rangle \cong \mathbb{Z}_n$

Exercise 5.15. If $(G, f) \cong \langle X \mid R \rangle$, then $G = \langle \{f(x) : x \in X\} \rangle$.

3. As a shorthand notation, below we will write $G \cong \langle X \mid R \rangle$ to denote that there exists $f : X \to G$ such that $(G, f) \cong \langle X \mid R \rangle$. With this notation, we have $D_{2n} \cong \langle x, y \mid x^n, y^2, xyxy \rangle$.
4. Free abelian group over X: $\langle X \mid \{x^{-1}y^{-1}xy : x, y \in X\} \rangle$.
5. Let G be an arbitrary group, $X = G$, $\mathrm{id}_G : G \to G$, then $(G, \mathrm{id}_G) \cong \langle G \mid R_G \rangle$, where

$$R_G = \{\text{all reduced words } w \in \mathcal{F}_G \text{ that are satisfied in } (G, \mathrm{id}_G), \\ \text{i.e., } w = g_1 \dots g_n \text{ such that } g_1 \dots g_n = e_G\}$$

Corollary 5.16. *If G is finite, then G is finitely presented, i.e., it has a presentation with finitely many generators and finitely many relations.*

Proof. Notice that R_G itself is infinite, since it contains the word $e_G e_G \dots e_G$ of every length, so let's use a smaller set. Let R'_G be its subset of the words ghk^{-1} with $g, h, k \in G$ and $gh = k$ in G, it is finite since G is finite. Clearly, (G, id_G) satisfies R'_G. If (H, f) satisfies R'_G, then $f(g) f(h) = f(gh)$ for all $g, h \in G$, so f itself is the unique homomorphism $\varphi : G \to H$ with $f = \varphi \circ \mathrm{id}_G$. Thus $(G, \mathrm{id}_G) \cong \langle G \mid R'_G \rangle$, which has finitely many generators and relations, as desired. □

Theorem 5.17. *Let X be a set, $R \subseteq \mathcal{F}_X$, then there exists a group G with a function $f : X \to G$ such that $(G, f) \cong \langle X \mid R \rangle$.*

We need a Lemma/definition in order to prove this theorem.

Lemma 5.18. *Let G be a group, $R \subseteq G$, then define $N(R) := \bigcap \{H \trianglelefteq G : R \subseteq H\}$, it is a normal subgroup of G, called the normal closure of R. $N(R)$ is the smallest normal subgroup of G containing R.*

Proof. Take $g \in G$, we want to show that $N(R)$ is normal, i.e., $gN(R)g^{-1} = N(R)$. Take $h \in N(R)$, then $ghg^{-1} \in N(R)$, since h is in every normal subgroup that contains R. This proves $gN(R)g^{-1} \subseteq N(R)$. Recall in Lemma 3.69, we have shown that one side of containment suffices to show a group is normal. Thus, $N(R)$ is normal, as desired. □

Exercise 5.19. Let G be a group, $R \subseteq G$, show that $N(R) = \langle \{grg^{-1} : g \in G, r \in R\} \rangle$.

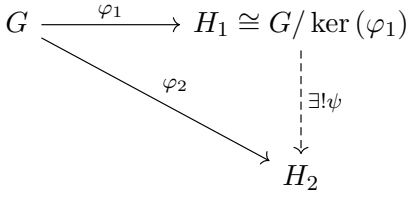
Lemma 5.20. *Let G, H_1, H_2 be groups, $\varphi_1 : G \to H_1, \varphi_2 : G \to H_2$ be homomorphisms such that $\mathrm{im}(\varphi_1) = H_1$ and $\ker(\varphi_1) \subseteq \ker(\varphi_2)$, then there exists a unique homomorphism $\psi : H_1 \to H_2$ such that $\psi \circ \varphi_1 = \varphi_2$.*

Proof. Let $K_1 := \ker(\varphi_1), K_2 := \ker(\varphi_2)$, by Theorem 3.67, $\mathrm{im}(\varphi_1) \cong G/K_1$, since $\mathrm{im}(\varphi_1) = H_1$, we can assume $H_1 = G/K_1$, so one can view φ_1 as a homomorphism from G to G/K_1, i.e., the quotient map. Define $\psi : G/K_1 \to H_2$ by $\psi(gK_1) = \psi(\varphi_1(g)) = \varphi_2(g)$, we need to check that ψ

is well-defined, i.e., if $gK_1 = g'K_1$, then $\varphi_2(g) = \varphi_2(g')$. Notice that

$$\begin{aligned} & gK_1 = g'K_1 \\ \Rightarrow\ & g^{-1}g'K_1 = K_1 \\ \Rightarrow\ & g^{-1}g' \in K_1 \\ \Rightarrow\ & g^{-1}g' \in K_2 && \text{since } K_1 \subseteq K_2 \\ \Rightarrow\ & \varphi_2\left(g^{-1}g'\right) = e_{H_2} \\ \Rightarrow\ & \varphi_2(g) = \varphi_2\left(g'\right) \end{aligned}$$

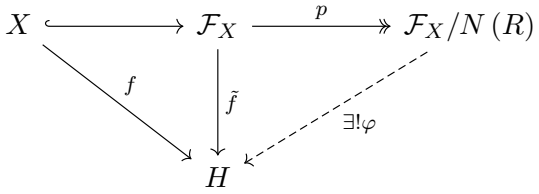
Moreover, ψ is a homomorphism, since $\psi(gK_1 \cdot g'K_1) = \psi(gg'K_1) = \varphi_2(gg') = \varphi_2(g)\varphi_2(g') = \psi(gK_1)\psi(g'K_1)$. Uniqueness is left as an exercise to the reader. This gives us the following diagram:



□

Now we are ready to prove the theorem.

Proof of Theorem. Take $G := \mathcal{F}_X/N(R)$, consider the following diagram:

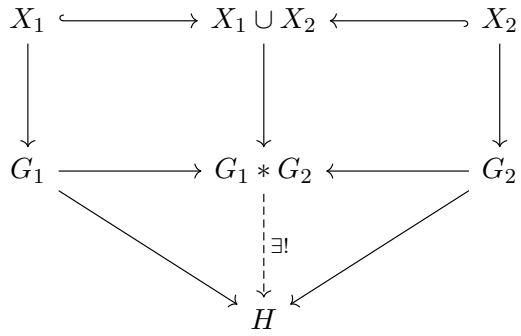


where (H, f) satisfies R. What we need to show is two things: 1). G satisfies R, and 2). there exists a unique homomorphism from G to H. For the first part, notice that $R \subseteq N(R)$, if we let ι denote the embedding of X into $\mathcal{F}_X$, then $p \circ \iota$ will satisfy R. For the second part, we apply Lemma 5.20 to $\mathcal{F}_X - \mathcal{F}_X/N(R) - H$, where we need to check that p is surjective, and $\ker(p) \subseteq \ker\left(\tilde{f}\right)$. The first part is clear, since p is the quotient map, for the second part, notice $\ker(p) = N(R)$. By our assumption, (H, f) satisfies R, so $R \subseteq \ker\left(\tilde{f}\right)$, and due to Corollary 3.70, $\ker\left(\tilde{f}\right)$ is normal, so $N(R) \subseteq \ker\left(\tilde{f}\right)$. Thus, by Lemma 5.20, there exists a unique homomorphism $\varphi : \mathcal{F}_X/N(R) \to H$ with $\varphi \circ p = \tilde{f}$, so $\varphi \circ (p \circ \iota) = \tilde{f} \circ \iota = f$. To see uniqueness, if $\varphi' \circ p \circ \iota = f$, then $\varphi' \circ p$ and $\tilde{f}$ are homomorphisms $\mathcal{F}_X \to H$ that agree on X, so $\varphi' \circ p = \tilde{f}$ by the uniqueness for free groups, then $\varphi' = \varphi$ by the uniqueness in Lemma 5.20, as desired. □

Example 5.21. Coproducts of groups. Let G_1, G_2 be groups, say $G_1 \cong \langle X_1 \mid R_1 \rangle, G_2 \cong \langle X_2 \mid R_2 \rangle$, assume $X_1 \cap X_2 = \emptyset$.

Proposition 5.22. *$\langle X_1 \cup X_2 \mid R_1 \cup R_2 \rangle = G_1 * G_2$ is a coproduct of G_1 and G_2, it is also called the free product.*

Proof. Consider the following diagram:



By the observation in Section 5.2, H satisfies both R_1 and R_2 via the composites $X_k \to G_k \to H$, so there exists a unique homomorphism from $G_1 * G_2$ to H compatible with these maps, and it makes the two lower triangles commute, since each G_k is generated by the image of X_k, by the exercise in the example above. □

This has some interesting dynamics: consider $\mathbb{Z}_2 * \mathbb{Z}_2$, we can represent the first with $\langle x \mid x^2 \rangle$ and the second with $\langle y \mid y^2 \rangle$, by Proposition 5.22, we have $\mathbb{Z}_2 * \mathbb{Z}_2 = \langle x, y \mid x^2, y^2 \rangle$, this is an infinite group: we can repeatedly form $xy, xyxy, xyxyxy, \ldots$, i.e., $(xy)^*$.

Example 5.23. Consider $\langle x, y \mid x^n, y^2, x^{-1}y^{-1}x^2y \rangle$, where the last relation can be interpreted as $x^2y = yx$, by this relation, we can swap all x in front of y, so every element can be written $x^m y^\ell$, for $0 \leq m < n, 0 \leq \ell \leq 1$. The size of this group is thus at most $2n$.

Let's do some more computations.

$$\begin{aligned} x &= y^2 x \\ &= yx^2 y \\ &= x^2 yxy \\ &= x^4 y^2 \\ &= x^4 \end{aligned}$$

So $x^3 = e$. If $\gcd(n, 3) = 1$, then $x^n = x^3 = e \Rightarrow x = e$, the group is then $\{e, y\} \cong \mathbb{Z}_2$. Notice $y \neq e$, since $x \mapsto 0, y \mapsto 1$ defines a homomorphism onto $\mathbb{Z}_2$: it sends all three relations to 0.

Exercise 5.24. What if n is divisible by 3?

Example 5.25. A famous computational problem, the *word* problem: given a finite presentation $\langle X \mid R \rangle$ and $w \in \mathcal{F}_X$, is w the identity in $\langle X \mid R \rangle$? Though the problem seems not so hard to solve, it is in fact *undecidable*.

Let's try to understand more structures about free groups.

We write $\mathcal{F}_n$ for $\mathcal{F}_X$ with $|X| = n$. $\mathcal{F}_0$ is the trivial group, $\mathcal{F}_1 \cong \mathbb{Z}$, $\mathcal{F}_2$ can be pictured as a $2d$ tree, we know that $\mathcal{F}_1 \ncong \mathcal{F}_2$, since $\mathcal{F}_1$ is abelian, but $\mathcal{F}_2$ is not. One can prove $\mathcal{F}_3 \ncong \mathcal{F}_2$ via *abelianization* or some counting argument on homomorphisms, but it is not obvious. Moreover, $\mathcal{F}_3$ is isomorphic to a subgroup of $\mathcal{F}_2$, making the structure even more mysterious. Even more, there exists an injective homomorphism from $\mathcal{F}_\infty$ to $\mathcal{F}_2$. Still, something can be said as "general property" for free groups.

Theorem 5.26 (Nielsen–Schreier)**.** *Every subgroup of a free group is free.*

The type of argument we are going to use is new to this course, it comes from a specific field called *geometric group theory*. One can view it as a general type of graphs, where multi-edges and self-loops are allowed, and for each edge, we think of it as two arcs with different orientations. We formally re-define the notion of graph we are going to use.

Definition 5.27. A graph G consists of the following:

- a set V of vertices
- a set E of edges
- a pair of functions, $start, end : E \to V$
- a bijection $flip : E \to E$, such that
  - $flip\,(flip\,(e)) = e$
  - $flip\,(e) \neq e$
  - $start\,(flip\,(e)) = end\,(e)$
  - $end\,(flip\,(e)) = start\,(e)$

Definition 5.28. For $x, y \in V$, an *xy-walk* is a sequence of edges $e_k e_{k-1} \dots e_2 e_1$ (possibly empty), where:

- $start\,(e_{i+1}) = end\,(e_i)$
- $start\,(e_1) = x, end\,(e_k) = y$

Given a graph G, we form a subset $E^+ \subseteq E$ by picking an arbitrary edge from each pair $\{e, flip\,(e)\}$. Let $E^- := E \setminus E^+$.

For an xy-walk p and a yz-walk q, one can define an xz-walk qp by concatenating the sequences, this gives rise to a category:

- Objects are vertices
- Morphisms from x to y: xy-walks
- Composition: concatenation of walks

Definition 5.29. A walk $e_k e_{k-1} \dots e_1$ is *reduced* if there is no i such that $e_{i+1} = flip\,(e_i)$.

For every xy-walk p, if p is *not* reduced, say there exists some i with $e_{i+1} = flip\,(e_i)$, then we can apply a reduction operation to p, by removing both e_i and e_{i+1} from the sequence. After finitely many reductions, we have a reduced walk, denoted by $Red\,(p)$. Define an operation on reduced walk p from x to y, reduced walk q from y to z, by $qp := Red\,(\text{concat}\,(qp))$.

Proposition 5.30. *1. Reduced walk $Red\,(p)$ does not depend on the order of reductions taken.*

2. If p is a reduced xy-walk, q is a reduced yz-walk, r is a reduced zw-walk, then $(rq)\,p = r\,(qp)$.

Proof. Consider the free group $\mathcal{F}_{E^+}$. Notice walks are directly related to words over $E^+ \cup E^-$, and reduced walks are essentially a subset of reduced words. Concatenation with reduction of reduced walks coincides with multiplication in $\mathcal{F}_{E^+}$. By using the property of $\mathcal{F}_{E^+}$, we have completed the proof of the proposition. □

This gives rise to a new category:

- Objects are vertices
- Morphisms from x to y: reduced xy-walks
- Composition: free group multiplication

In this category, every morphism has an inverse: let p be a walk from x to y, then flipping all edges along p and reversing their order gives p^{-1}, a walk from y to x.

For $x \in V$, let $\pi(x, G)$ be the group of all reduced xx-walks. Notice $\pi(x, G) \leq \mathcal{F}_{E^+}$.

Proposition 5.31. *If G is connected, then $\pi(x, G) \cong \pi(y, G)$, for any $x, y \in V$. So we can write $\pi(G)$ without mentioning any vertex, this group is called the fundamental group of G.*

Proof. Since G is connected, there exists a walk p from x to y. Define the isomorphism as follows: given an xx-walk q, maps it to pqp^{-1}, i.e., first travels from y to x, take an xx-walk, then travel back to y. To see this is a homomorphism, observe that $pq_1q_2p^{-1} = pq_1p^{-1}pq_2p^{-1}$, and the inverse of this mapping is, given a yy-walk o, map it to $p^{-1}op$. □

Theorem 5.32. *Let X be a set, $\mathcal{F}_X$ be the free group on X, and $H \leq \mathcal{F}_X$, then $H \cong \pi(G)$ for some connected graph G.*

Proof. Look at $\mathcal{F}_X/H$. Let $\alpha : \mathcal{F}_X \curvearrowright \mathcal{F}_X/H$ be the left multiplication action. Define a graph G as follows: $V = \mathcal{F}_X/H, E = \left\{(C, x) : C \in \mathcal{F}_X/H, x \in X \cup X^{-1}\right\}$, i.e., vertices are the left cosets of H, and edges comprise two bits of information: the start of the edge, C, and the end point xC. So $flip(C, x) = \left(xC, x^{-1}\right)$, $flip\left(C, x^{-1}\right) = \left(x^{-1}C, x\right)$, $start(C, x) = C, end(C, x) = xC$. Now our goal is to show $\pi(G) \cong H$. The graph G is connected: if $g = x_k \dots x_1$ with $x_i \in X \cup X^{-1}$, then the edges $(x_{i-1} \dots x_1 H, x_i)$, $i = 1, \dots, k$, form a walk from H to gH. First, $\pi(G) \cong \pi(H, G)$, what are elements of $\pi(H, G)$? That's all reduced HH-walks, i.e., sequences $(C_k, x_k) \dots (C_2, x_2)(C_1, x_1)$, where $x_1, \dots, x_k \in X \cup X^{-1}$, $C_1 = H$, $x_kC_k = H$, and $C_{i+1} = x_iC_i$, so we can write the sequence as $x_kx_{k-1} \dots x_2x_1H = H \Rightarrow x_k \dots x_2x_1 \in H$, a.k.a., each reduced word in H corresponds to exactly one reduced walk in $\pi(H, G)$, since a walk is reduced exactly when its word is: $(C_{i+1}, x_{i+1}) = flip(C_i, x_i)$ if and only if $x_{i+1} = x_i^{-1}$. This correspondence turns concatenation followed by reduction into multiplication in $\mathcal{F}_X$, so it is an isomorphism $\pi(H, G) \cong H$. This completes the proof. □

Definition 5.33. Let G be a graph, $e = \{u, v\}$ be an edge, the *edge contraction of e* is the operation that removes e and $flip(e)$ from E, and makes one super vertex w, connecting all neighbors of u and v to w. Notice edge contraction might yield a graph with self-loops, if the original graph admits multi-edges. We use G/e as a notation for G contracting edge e.

Proposition 5.34. *If $e \in E(G)$ and e is not a self-loop, then $\pi(G) \cong \pi(G/e)$.*

Proof. Let $e = \{u, v\}$, and let w be the super vertex of u and v, after contraction. We will show $\pi(u, G) \cong \pi(w, G/e)$. We classify uu-walks as follows:

Type 1 Walks that do not use edge e. For this kind of walk, we simply map it to the same sequence of edges in G/e, with all u and v substituted by w.

Type 2 Walks that use edge e. This means at some time of the walk, it takes edge e to go from u to v, since this is essentially a uu-walk, it then must take some walk p to go from v to u, so it is safe to simply remove all occurrences of edge e in the original walk, the result is still a ww-walk, and it is not hard to see that this mapping is bijective.

Thus, $\pi(u, G) \cong \pi(w, G/e)$. □

Definition 5.35. A *spanning tree* in a connected graph G is a set $T \subseteq E$ such that

1. for all $e \in T$, $flip(e) \in T$
2. for all $x, y \in V$, there is an xy-walk with edges in T
3. there are no non-trivial reduced closed walks in T, i.e., T does not contain cycles

Proposition 5.36. *If G is a connected graph, then G has a spanning tree.*

Proof. A simple procedure will work: keep adding edges that do not form a cycle, until all vertices have been covered. This works when G is finite, but we need infinite graphs below, since $\mathcal{F}_X/H$ may be infinite. In general, apply Zorn's lemma (Lemma 6.30) to the sets of edges that are closed under $flip$ and contain no non-trivial reduced closed walks, ordered by inclusion. The union of a chain is again such a set, since a closed walk uses only finitely many edges, and a maximal such set connects all vertices, otherwise we could add an edge of G joining two of its components. □

Exercise 5.37. If $T \subseteq E$ is a spanning tree in G, then for all $x, y \in V$, there is a unique reduced xy-walk in T.

Proof. Existence: by condition 2, there is an xy-walk in T, and reducing it gives a reduced xy-walk in T. For uniqueness, suppose there are at least two reduced xy-walks $q_1 \neq q_2$ in T, then one can use the first walk to travel from x to y, and use the second to travel back from y to x: this is doable since by definition of spanning tree, both e and $flip(e)$ are in T. This gives us a closed walk $q_2^{-1} q_1$ in T, whose reduction is non-trivial, since reduction is multiplication in $\mathcal{F}_{E^+}$ and $q_1 \neq q_2$, contradicts the definition of spanning tree. □

Theorem 5.38. *For every connected graph G, $\pi(G)$ is a free group.*

Proof. Pick a vertex $x \in V$, let $T \subseteq E$ be a spanning tree in G. We claim that $\pi(G) \cong \mathcal{F}_{E^+ \setminus T}$, where we interpret $flip(e)$ as e^{-1}. First of all, it is clear that $\pi(x, G) \subseteq \mathcal{F}_{E^+}$, define a homomorphism $\varphi : \mathcal{F}_{E^+} \to \mathcal{F}_{E^+ \setminus T}$, by $e \mapsto \begin{cases} e, & \text{if } e \notin T \\ \text{id}, & \text{otherwise} \end{cases}$ The goal is to show $\varphi|\pi(x, G)$ is an isomorphism from $\pi(x, G)$ to $\mathcal{F}_{E^+ \setminus T}$, i.e., for every $e_k \dots e_2 e_1 \in \mathcal{F}_{E^+ \setminus T}$, there is a unique $p \in \pi(x, G)$ such that $\varphi(p) = e_k \dots e_1$. Let's say we write $p = p_k e_k p_{k-1} e_{k-1} \dots p_2 e_2 p_1 e_1 p_0$, where $p_0, p_1, \dots, p_k$ are walks using tree edges from T. Notice that p_0 is a $(x, start(e_1))$-walk, for each $1 \leq i < k$, p_i is a $(end(e_i), start(e_{i+1}))$-walk, and p_k is a $(end(e_k), x)$-walk. These are all walks using purely T edges, so by Exercise 5.37, these are all unique. For existence, taking $p_0, \dots, p_k$ to be these reduced walks in T gives a reduced walk $p \in \pi(x, G)$ with $\varphi(p) = e_k \dots e_1$: no reduction can happen next to e_i, since $flip(e_i) \notin T$ and $e_k \dots e_1$ is reduced. This proves $\varphi|\pi(x, G)$ is a bijection, and therefore, $\pi(G) \cong \mathcal{F}_{E^+ \setminus T}$, $\pi(G)$ is a free group. □

With all these in hand, we are ready to prove Theorem 5.26.

Proof of Nielsen–Schreier. Let $H \leq \mathcal{F}_X$, then by Theorem 5.32, $H \cong \pi(G)$ for some connected graph G, and by Theorem 5.38, $\pi(G)$ is a free group, therefore, H is also a free group, as desired. $\square$

5.3 Structure of finite groups

Groups have a simple definition but complicated internal dynamics, so many people try to look at finite simple groups, and classify them based on certain criteria. In this section, we will try to do so.

Theorem 5.39 (Cauchy)**.** *If G is a finite group, p is a prime number such that $p \mid |G|$, then G has an element of order p.*

Proof. Consider the following set $S := \{(a_1, a_2, \ldots, a_p) \in G^p : a_1 a_2 \ldots a_p = e\}$. What is the cardinality of S? The claim is $|S| = |G|^{p-1}$, the idea is we can freely pick $p-1$ elements of the sequence, while the last element is fixed as $(a_1 a_2 \ldots a_{p-1})^{-1}$. The sequence $a_1 a_2 \ldots a_p$ exhibits *rotational invariance:* $a_1 a_2 \ldots a_p = e \Rightarrow a_1^{-1} a_1 a_2 \ldots a_p a_1 = a_1^{-1} a_1 = e \Rightarrow a_2 \ldots a_p a_1 = e$, so let's define an action $\alpha : \mathbb{Z}_p \curvearrowright S$ by $\alpha(1)(a_1, \ldots, a_p) := (a_2, \ldots, a_p, a_1)$, and

$$\alpha(m)(a_1, \ldots, a_p) := (a_{m+1}, \ldots, a_p, a_1, a_2, \ldots, a_m),$$

i.e., shifts the sequence to the left by m. By orbit-stabilizer, $\mathbb{Z}_p / \operatorname{Stab}(s) \cong \mathbb{Z}_p \cdot s$, since $|\mathbb{Z}_p| = p$ and $\operatorname{Stab}(s) \leq \mathbb{Z}_p$, $|\mathbb{Z}_p / \operatorname{Stab}(s)|$ is either 1 or p, which also indicates every orbit of α has size either 1 or p. Since orbits partition S, we can classify S into fixed points, i.e., elements with orbit size 1, let Z be the set of these points. Since $(e, e, \ldots, e) \in Z$, $|Z| \geq 1$. Moreover, $|S| = |Z| + kp$, where k is the number of orbits of size p, taking modulo p, we have $|S| = |Z| \bmod p$, since $|S| = |G|^{p-1}$ and $p \mid |G|$, $|Z| = 0 \bmod p$, so $|Z| \geq p \geq 2$. So let $(a_1, \ldots, a_p) \in Z$ with $(a_1, \ldots, a_p) \neq (e, e, \ldots, e)$, and by definition of fixed points, $(a_1, a_2, \ldots, a_p) = (a_2, a_3, \ldots, a_p, a_1) \Rightarrow a_1 = a_2 = a_3 = \ldots = a_p = a \neq e \Rightarrow a^p = e$, so a has order p, since its order divides the prime p and is not 1, as desired. □

Exercise 5.40. Let G be a finite group, p is a prime number such that $p \mid |G|$, then number of cyclic subgroups of G, of order p, equals 1 mod p.

Proof. We make use of the fixed points set Z in Theorem 5.39. Since $|Z| = 0 \bmod p$, we can write $|Z| = kp$ for some positive integer k, get rid of $(e, e, \ldots, e)$, there are $kp - 1$ sequences remaining. Recall from the proof of Theorem 5.39 that Z consists of the sequences $(a, a, \ldots, a)$ with $a^p = e$, so these $kp-1$ sequences are one for each element of order p. Each subgroup of order p contains exactly $p - 1$ elements of order p, every element of order p generates such a subgroup, and two distinct ones meet only in e, so the number of subgroups of order p is $N = \frac{kp-1}{p-1}$. Since $N(p-1) = kp - 1$, we get $-N = -1 \bmod p$, i.e., $N = 1 \bmod p$. □

Proposition 5.41. *If G is a finite group, $|G| = mp$ where p is a prime number and $1 < m < p$, then G is not simple.*

Proof. The claim is G has a normal subgroup of order p, suppose this is not the case, then G has ≥ 2 subgroups of order p: by Theorem 5.39, G has a subgroup of order p, and if it were the only one, it would be normal, since its conjugates are also subgroups of order p. By Exercise 5.40, the number of subgroups of order p is 1 mod p, so the number must be $\geq p + 1$. Note that if both $H_1, H_2 \leq G$ with $|H_1| = |H_2| = p$, $H_1 \neq H_2 \Rightarrow H_1 \cap H_2 = \{e\}$, so $|G| \geq 1 + (p+1)(p-1) = p^2$, contradicts $|G| = mp < p^2$. □

Example 5.42. Suppose we have a group G with $|G| = 2019$, notice $2019 = 673 \times 3$, and 673 is a prime, so there are no simple groups of order 2019.

Theorem 5.43 (Sylow's First Theorem)**.** *Let G be a finite group, p be a prime number, and $k \in \mathbb{N}$. If $p^k \mid |G|$, then G has a subgroup of order p^k.*

Proof. By induction on $|G|$. If $k = 0$, then it's trivial. Assume $k \geq 1$, in particular, $p \mid |G|$. If G has a subgroup H such that $H \neq G$ and $p^k \mid |H|$, then by induction hypothesis, we are done. Suppose G has no such subgroups, we claim that G has a normal subgroup of order p. We will make use of the *class formula* of G:

$$|G| = |Z(G)| + \sum_{a \in A} |G : Z(a)|$$

where $Z(G)$ is the *center* of G, all elements in $Z(G)$ can commute with any elements in G, and $Z(a)$ is the *centralizer* of element a, which contains elements that can commute with a. Finally, A is a subset of G formed by picking an element from each non-trivial conjugacy class. To give a short proof of class formula, consider the conjugation action $\alpha_{conj} : G \curvearrowright G$, notice the stabilizer of a is exactly $Z(a)$, and by orbit-stabilizer, each orbit has size $|G : Z(a)|$. Since orbits partition G, we get the class formula.

Notice that each $Z(a)$ is a subgroup of G other than G, since a is not in the center, so by our assumption, $p^k \nmid |Z(a)|$, which means $p \mid |G|/|Z(a)| \Rightarrow p \mid |G : Z(a)|$, so we can rewrite class formula as

$$mp^k = |Z(G)| + p \sum_{a \in A} C_a$$

where we write $|G : Z(a)|$ as $C_a p$. Taking modulo p on both sides, we get $0 = |Z(G)| + 0 \bmod p$, so $|Z(G)| = 0 \bmod p$. Moreover, $e \in Z(G)$, so $|Z(G)| \geq 1 \Rightarrow p \mid |Z(G)|$. By Theorem 5.39, there exists a subgroup of $Z(G)$ with order p, which is also normal: this comes from the fact that $Z(G)$ is the center, so any subgroup of $Z(G)$ is normal.

Let $N \trianglelefteq G$ be the normal subgroup of G with order p, consider G/N. Notice that $p^{k-1} \mid |G/N|$, so there is a subgroup $P \leq G/N$ of order p^{k-1}, by induction hypothesis. Let Q be the preimage of P, under the quotient map $G \twoheadrightarrow G/N$, one can verify that Q is a subgroup of G which contains N, so $N \trianglelefteq Q$, we can write $P \cong Q/N \Rightarrow |Q| = |P||N| = p^{k-1}p = p^k$, and we get our desired subgroup. □

Definition 5.44. A *p-group* is a group of size equal to a power of p. A *Sylow p-subgroup* of a finite group G is a subgroup $P \leq G$ such that $|P|$ = the largest power of p dividing $|G|$.

Remark 5.45. One can interpret Theorem 5.43 as saying that every finite group G has a Sylow p-subgroup, if $p \mid |G|$.

Theorem 5.46 (Sylow's Second Theorem)**.** *Let G be a finite group, p be a prime number, $H \leq G$ be a p-group and $P \leq G$ be a Sylow p-subgroup, then H is contained in a conjugate of P, i.e., there exists $g \in G$ such that $H \subseteq gPg^{-1}$.*

Before proving the theorem, let's first prove a handy Lemma.

Lemma 5.47. *Let G be a p-group, S be a finite set, and Z denote the set of fixed points under any action $G \curvearrowright S$, then $|Z| = |S| \mod p$.*

Proof. Given an element $s \in S$, let O_s denote the orbit contains s, by orbit-stabilizer, $|O_s| = |G/\operatorname{Stab}(s)| = |G|/|\operatorname{Stab}(s)| \Rightarrow |O_s| \mid |G|$, since G is a p-group, let $p^k := |G|$, so $|O_s| \mid p^k$. Since orbits partition S, we can write $|S| = |Z| + \sum_{s \in A} |O_s|$, where A is a subset of S formed by picking

an element from each non-trivial orbit. For each non-trivial orbit, $|O_s| \geq 2 \Rightarrow p \mid |O_s|$, so if we take modulo p on both sides, we get $|Z| = |S| \bmod p$, as desired. □

Proof of Sylow's Second Theorem. Let $H \curvearrowright G/P$ by left multiplication, by orbit-stabilizer, every orbit has size equal to $|H/\operatorname{Stab}(gP)|$, so every orbit has size a divisor of $|H|$, hence a power of p. But $|G/P| = |G|/|P| \neq 0 \bmod p$, since P is Sylow p-subgroup, so at least one orbit has size $p^0 = 1$, which means there is some $g \in G$, such that for any $h \in H$, $hgP = gP \Rightarrow g^{-1}hgP = P \Rightarrow g^{-1}hg \in P \Rightarrow H \subseteq gPg^{-1}$. □

Theorem 5.48 (Sylow's Third Theorem)**.** *Let G be a finite group with $|G| = mp^k$, $m \neq 0$ mod p, where p is a prime number. Then number of Sylow p-subgroups of G divides m, and is 1 mod p.*

Definition 5.49. Let $H \leq G$, the *normalizer* of H in G is $N_G(H) := \left\{g \in G : gHg^{-1} = H\right\}$, i.e., the stabilizer of H under conjugation action $G \curvearrowright \operatorname{Sub}(G)$, several properties of $N_G(H)$:

- $N_G(H) \leq G$
- $H \trianglelefteq N_G(H)$
- Orbit-stabilizer: number of subgroups of G that are conjugate to H, i.e., size of orbit containing H, is equal to $|G : N_G(H)|$

Proof of Sylow's Third Theorem. We first prove the number of Sylow p-subgroups of G divides m. Let P be any Sylow p-subgroup of G, we use $n_P :=$ number of subgroups conjugate to $P = |G : N_G(P)| = |G|/|N_G(P)|$. By Theorem 5.46, every Sylow p-subgroup Q satisfies $Q \subseteq gPg^{-1}$ for some $g \in G$, so $Q = gPg^{-1}$, since both have p^k elements. Thus n_P is the number of all Sylow p-subgroups of G. Notice $|G| = mp^k$, and $|N_G(P)|$ is divisible by $|P| = p^k$, suppose $|N_G(P)| = \ell p^k$, then $n_P = \frac{mp^k}{\ell p^k} = \frac{m}{\ell} \Rightarrow n_P\ell = m, n_P \mid m$.

Next, we prove $n_P = 1 \bmod p$. Consider the following:

$$\begin{aligned} m &= |G : P| \\ &= |G : N_G(P)||N_G(P) : P| \\ &= n_P|N_G(P) : P| \end{aligned}$$

Lemma 5.50. *Let G be a finite group, p be a prime number, $H \leq G$ be a p-group, then $|N_G(H) : H| = |G : H|$ mod p.*

Proof of Lemma. Consider the action $H \curvearrowright G/H$ by left multiplication. Orbits have sizes which are powers of p. By Lemma 5.47, $|G : H| = |G/H| =$ (number of fixed points of this action) mod p, so it suffices to show $|N_G(H) : H|$ equals the number of fixed points, modulo p. Consider the following chain of logical equivalence:

$$\begin{aligned} gH \text{ is a fixed point} &\Leftrightarrow \forall h \in H, hgH = gH \\ &\Leftrightarrow g^{-1}hgH = H \\ &\Leftrightarrow g^{-1}hg \in H \\ &\Leftrightarrow H = gHg^{-1} && \text{since } g^{-1}Hg \subseteq H \text{ and both have } |H| \text{ elements} \\ &\Leftrightarrow g \in N_G(H) \end{aligned}$$

So gH is a fixed point if and only if $g \in N_G(H)$, i.e., it corresponds to a coset of $N_G(H)/H$. Thus, we have the identity. □

By Lemma, our original equation becomes $m = |G : P| = |N_G(P) : P| \bmod p$, i.e., $n_P | N_G(P) : P| = |N_G(P) : P| \bmod p$, since P is Sylow p-subgroup, $|N_G(P) : P|$ is not divisible by p, so we must have $n_P = 1 \bmod p$, as desired. □

Proposition 5.51. *There are no simple groups of order 12.*

Proof. $12 = 2^2 \times 3$. Number of Sylow 2-subgroups: 1 or 3, number of Sylow 3-subgroups: 1 or 4. Assume for a contradiction that there is a simple group of order 12, then there must be 3 Sylow 2-subgroups and 4 Sylow 3-subgroups. Why? A unique Sylow p-subgroup would be normal, since its conjugates are Sylow p-subgroups too. So there are 3 subgroups of order 4, 4 subgroups of order 3. How many elements are there of order 3? Exactly $2 \times 4 = 8$ elements, since each subgroup of order 3 contains two elements of order 3, each element of order 3 generates a subgroup of order 3, and two distinct subgroups of order 3 meet only in e. So there remain $12 - 8 = 4$ elements, and we cannot fit 3 subgroups of order 4 into 4 elements. □

Proposition 5.52. *There are no simple groups of order 24.*

Proof. $24 = 2^3 \times 3$. Number of Sylow 3-subgroups: 1 or 4, number of Sylow 2-subgroups: 1 or 3, again assume for a contradiction that there is a simple group of order 24, so number of Sylow 3-subgroups is 4, and number of Sylow 2-subgroups is 3. Let $G \curvearrowright \{\text{Sylow 2-subgroups}\}$ by conjugation, we can view it as $\alpha : G \to \operatorname{Sym}(\{1,2,3\})$. Now let's examine $\ker(\alpha)$, since $|G| = 24, |\operatorname{Sym}(\{1,2,3\})| = 6$, α is not injective, so $\ker(\alpha) \neq \{e\}$. On the other hand, since α is conjugation, clearly $\ker(\alpha) \neq G$: there are 3 Sylow 2-subgroups and they are all conjugate to each other by Theorem 5.46, so the action is not trivial. Thus, we have found a non-trivial normal subgroup of G, which is $\ker(\alpha)$. □

6 Rings and Modules

6.1 Basic definitions

Rings and modules are built upon groups, they are less "abstract", in the sense that they allow more operations, so there are real-life examples and applications of these concepts.

Definition 6.1. Let A be an abelian group. An *endomorphism* of A is a homomorphism $f : A \to A$.

Example 6.2. Some examples of endomorphisms:

- identity map, $\mathbf{1}$
- zero map, i.e., map all elements to 0, $\mathbf{0}$
- inverse map: $a \mapsto -a$, denoted by $-\mathbf{1}$, this uses that A is abelian: $-\mathbf{1}(a+b) = -b-a = -a-b = -\mathbf{1}(a) + -\mathbf{1}(b)$

Definition 6.3. Let $\mathrm{End}(A) := \{\text{all endomorphisms of } A\}$, besides the standard function composition, we also equip $\mathrm{End}(A)$ with pointwise addition: for $f, g \in \mathrm{End}(A)$, define $f+g : A \to A$ by $(f+g)(a) = f(a) + g(a)$.

Proposition 6.4. *$f+g$ is an endomorphism.*

Proof.

$$\begin{aligned}(f+g)(a+b) &= f(a+b) + g(a+b)\\ &= f(a) + f(b) + g(a) + g(b)\\ &= f(a) + g(a) + f(b) + g(b)\\ &= (f+g)(a) + (f+g)(b)\end{aligned}$$

□

Exercise 6.5. Check that $(\mathrm{End}(A), +)$ is abelian.

Definition 6.6. A *concrete ring* is a subgroup $R \leq (\mathrm{End}(A), +)$, such that for all $f, g \in R$, $f \circ g \in R$. If $\mathrm{id}_A = \mathbf{1} \in R$, then R is called a *unital* ring.

Example 6.7.

1. $\mathrm{End}(A)$, analogous to $\mathrm{Sym}(A)$, called the *endomorphism ring* of A.
2. $\{\mathbf{0}\} \leq \mathrm{End}(A)$ is a ring.
3. What is $\mathrm{End}(\mathbb{Z})$? For $f \in \mathrm{End}(\mathbb{Z})$, all its values are determined by $f(1)$, for each $n \in \mathbb{Z}$, $\exists! f \in \mathrm{End}(\mathbb{Z})$ with $f(1) = n, f(m) = nm$. Let f_n denote the function that $f(1) = n$, so $\mathrm{End}(\mathbb{Z}) = \{f_n : n \in \mathbb{Z}\}$, and $f_n + f_k = f_{n+k}, f_n \circ f_k = f_{nk}$, so $(\mathrm{End}(\mathbb{Z}), +, \circ) \cong (\mathbb{Z}, +, \cdot)$.
4. $\{f_r : r \in \mathbb{R}\} \subseteq \mathrm{End}(\mathbb{R})$, for each $r \in \mathbb{R}$, $f_r : \mathbb{R} \to \mathbb{R}$ is given by $f_r(s) = rs$, then $f_r \in \mathrm{End}(\mathbb{R})$, $f_r + f_s = f_{r+s}, f_r \circ f_s = f_{rs}$, so $\{f_r : r \in \mathbb{R}\}$ is a concrete ring on $\mathbb{R}$, and it is isomorphic to $(\mathbb{R}, +, \cdot)$. However, it is not the entire $\mathrm{End}(\mathbb{R})$: there exists $f \in \mathrm{End}(\mathbb{R})$ such that $f(1) = \pi, f\left(\sqrt{2}\right) = e$.

Definition 6.8. An *(abstract) ring* is a set R with two binary operations $+, \cdot$, such that $(R, +)$ is an abelian group with identity $0 \in R$. $\cdot$ is associative and for all $a, b, c \in R$, $a \cdot (b+c) = a \cdot b + a \cdot c$ and $(a+b) \cdot c = a \cdot c + b \cdot c$. If $\cdot$ has an identity element 1, then R is called *unital*. If $\cdot$ is commutative, then R is called a *commutative ring*.

Analogous to concrete group and group, if $R \subseteq \operatorname{End}(A)$ is a concrete ring, $f, g, h \in R$, then

$$\begin{aligned}(f \circ (g+h))(a) &= f \circ (g+h)(a) \\ &= f(g(a) + h(a)) \\ &= f(g(a)) + f(h(a)) \\ &= f \circ g(a) + f \circ h(a)\end{aligned}$$

The other distributive law, $((g+h) \circ f)(a) = g(f(a)) + h(f(a))$, holds by the definition of $+$.

Theorem 6.9. *Let $(R, +, \cdot)$ be a unital ring, for each $r \in R$, let $f_r : R \to R : a \mapsto ra$, then $f_r \in \operatorname{End}(R, +)$, $\tilde{R} := \{f_r : r \in R\}$ is a concrete ring on $(R, +)$ and the map $R \to \tilde{R} : r \mapsto f_r$ is an isomorphism from $(R, +, \cdot)$ to $\left(\tilde{R}, +, \circ\right)$.*

Proof. One can easily duplicate the argument from Theorem 3.22. We won't repeat it here. □

Example 6.10. $\operatorname{End}\left(\mathbb{Z}^2\right)$. Let $f : \mathbb{Z}^2 \to \mathbb{Z}^2$ be an endomorphism, suppose $f(1,0) = (a,b)$, $f(0,1) = (c,d)$, then $f(x,y) = xf(1,0) + yf(0,1) = (xa, xb) + (yc, yd) = (xa + yc, xb + yd)$. It is best to write it in matrix notation:

$$f(x,y) = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

So $\operatorname{End}\left(\mathbb{Z}^2\right)$ is all 2×2 integer matrices.

Exercise 6.11. Verify that in $\operatorname{End}\left(\mathbb{Z}^2\right)$, ring addition is matrix addition, and function composition is matrix multiplication. It's worth noticing that this ring is not commutative.

Definition 6.12. Let R, S be rings, a *homomorphism* from R to S is a function $f : R \to S$ such that

- f is a homomorphism of abelian groups $(R, +_R), (S, +_S)$
- $\forall a, b \in R$, $f(ab) = f(a) f(b)$

If R, S are unital, then a homomorphism $f : R \to S$ is called *unital* if $f(1_R) = 1_S$.

Definition 6.13. An *R-module* or a *module over R* is an abelian group M, with a homomorphism $\alpha : R \to \operatorname{End}(M)$. If R is unital, then R-module $\alpha : R \to \operatorname{End}(M)$ is unital, if $\alpha(1_R) = \mathrm{id}_M$. Often, we say "$M$ is an R-module" and write ra for $\alpha(r)(a)$.

Example 6.14.

1. Let R be a ring, then $R \to \operatorname{End}(R) : r \mapsto (a \mapsto ra)$ is a homomorphism, so R is a module over itself.
2. Let R be a ring, M be any abelian group, the trivial homomorphism $R \to \operatorname{End}(M) : r \mapsto \mathbf{0}$.
3. $\mathbb{Z}$-module: Let M be an abelian group, consider the following homomorphism: $\mathbb{Z} \to \operatorname{End}(M) : 0 \mapsto \mathbf{0}, 1 \mapsto \mathrm{id}_M, n \mapsto n \cdot \mathrm{id}_M = \underbrace{\mathrm{id}_M + \mathrm{id}_M + \ldots + \mathrm{id}_M}_{n}$, i.e., $(n \mathrm{id}_M)(a) = \underbrace{a + a + \ldots + a}_{n} = na$, $\forall n \in \mathbb{Z}, a \in M$.

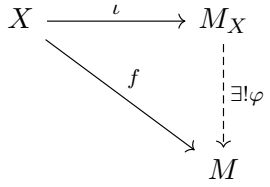
   Exercise 6.15. Show that this is a ring homomorphism, i.e., it preserves multiplication.

   Remark 6.16. This establishes a relation: unital $\mathbb{Z}$-modules correspond to abelian groups.
4. Let R be a ring, R^n(as abelian group) can be equipped with an R-module structure as follows: $r(a_1, a_2, \ldots, a_n) = (ra_1, \ldots, ra_n)$.

6.2 Free modules, independent, spanning, and basis

Definition 6.17. Let R be a unital ring and X be a set, a *free R-module over X* is a (unital) R-module M_X, together with the function $\iota : X \to M_X$, such that for all (unital) R-modules M and functions $f : X \to M$, there exists a unique R-module homomorphism $\varphi : M_X \to M$, such that $\varphi \circ \iota = f$.

Remark 6.18. This is almost the same diagram as free group:



Definition 6.19. Let M, M' be R-modules, a function $\varphi : M \to M'$ is an *R-module homomorphism* if it is a group homomorphism and for all $r \in R$, $a \in M$, $\varphi(ra) = r\varphi(a)$.

Let's now discover the structure of M_X. If X is finite, then $M_X \cong R^n$, where $n = |X|$. The idea is to think of the embedding function ι as mapping each element of X to an indicator vector, or more formally, define $M_X = \{\omega : X \to R\}$ and $\iota : X \to M_X : x \mapsto \omega_x$, where $\omega_x(y) = \begin{cases} 1_R, & \text{if } y = x \\ 0_R, & \text{otherwise} \end{cases}$ On the other hand, if X is infinite, we just change the definition of M_X a little bit: $M_X = \{\omega : X \to R : \text{there are finitely many } x \in X \text{ such that } \omega(x) \neq 0_R\}$.

For each $r \in R, \omega \in M_X$, define $r \cdot \omega \in M_X$ by $(r \cdot \omega)(y) = r(\omega(y))$. Then the definition of φ becomes obvious, let's first write elements of M_X as $\sum_{x \in X} a_x \omega_x$, where $a_x \in R$, so $\varphi(\omega) := \varphi\left(\sum_{x \in X} a_x \omega_x\right) = \sum_{x \in X} a_x f(x)$.

Exercise 6.20. Check this is a homomorphism, and it is unique.

We will extend on φ in free modules, to introduce some concepts the reader might recall from linear algebra.

Definition 6.21. Let $X \subseteq M$, and let $\varphi : M_X \to M$ be the R-module homomorphism given by the inclusion $X \hookrightarrow M$.

- X is *independent* if φ is injective: φ is injective $\to \ker(\varphi) = \{0\} \to$ for any coefficients a_x, $\sum_{x \in X} a_x \cdot x = 0$ implies $a_x = 0, \forall x \in X$.
- X is *spanning* if φ is surjective: φ is surjective $\to \forall m \in M, \exists$ coefficients $a_x, x \in X$, only finitely many of which are non-zero, such that $\sum_{x \in X} a_x \cdot x = m$. We write $\text{span}(X)$ for the span of X.
- X is a *basis*, if φ is bijective. In this case, M is a *free R-module*.

Before looking at free modules, let's first check out some examples on non-free modules.

Definition 6.22. Let G be a group, G is *torsion-free*, if every $g \in G$ with $g \neq e$ has infinite order, i.e., for $k \in \mathbb{Z} \setminus \{0\}$, $g^k \neq e$.

Example 6.23. 1. $\mathbb{Z}_n$ $(n \geq 2)$ is a $\mathbb{Z}$-module that has no basis, there are no non-empty independent sets. $\forall x \in \mathbb{Z}_n$, $n \cdot x = 0 \Rightarrow \{x\}$ is not independent. Note that $(\mathbb{Z}_n, +, \cdot)$ is a ring, and as a $\mathbb{Z}_n$-module, $\mathbb{Z}_n$ is free, since $\{1\}$ is a basis. Observe that any free $\mathbb{Z}$-module is *torsion-free*, i.e., any element in module M has infinite order: $\forall r \in \mathbb{Z} \setminus \{0\}$ and $x \in M \setminus \{0\}$, $r \cdot x \neq 0$. Let B be a basis of M. For basis elements, this is trivially true, so suppose x is not a basis element, we can write it as $\sum_{b \in B} a_b \cdot b$, so $r \cdot x = \sum_{b \in B} r a_b \cdot b$. Since $x \neq 0$, some a_b is non-zero, then $r a_b \neq 0$, since $r \neq 0$ and $\mathbb{Z}$ has no zero divisors, so $r \cdot x \neq 0$ by the independence of B.

2. $(\mathbb{Q}, +)$ is a torsion-free abelian group. As a $\mathbb{Z}$-module, it has no basis: suppose $X \subseteq \mathbb{Q}$ is a basis for $\mathbb{Q}$ as a $\mathbb{Z}$-module, pick $x \in X, x \neq 0$, then $\frac{1}{2}x = \sum_{y \in X} c_y \cdot y$, with only finitely many non-zero $\Rightarrow 1 \cdot x = \sum_{y \in X} (2c_y) \cdot y$, the coefficient of x is odd, but coefficients on RHS are even, so X cannot be a basis.

   Fact: If X is a basis and $\sum_{x \in X} a_x \cdot x = \sum_{x \in X} b_x \cdot x$, then $a_x = b_x, \forall x \in X$.

3. If X is infinite, then $\mathbb{Z}^X$ is not a free $\mathbb{Z}$-module.

Definition 6.24. A *division ring* is a unital ring R in which every non-zero element has multiplicative inverse. A non-zero commutative division ring is called a *field*.

Example 6.25. $\mathbb{R}, \mathbb{Q}, \mathbb{C}, \mathbb{Z}_p$ where p is a prime number, $\left\{a + b\sqrt{2} : a, b \in \mathbb{Q}\right\} \subset \mathbb{R}$. Let's verify the last one for the multiplicative inverse part: $\frac{1}{a+b\sqrt{2}} = \frac{a-b\sqrt{2}}{a^2-2b^2} = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}$, notice $a^2 - 2b^2 \neq 0$ for $(a, b) \neq (0, 0)$, since $\sqrt{2}$ is irrational.

In the next part, we will try to prove that every module over a division ring is free. From now on, modules over a unital ring are unital, i.e., $1 \cdot x = x$. We need this: if $M \neq \{0\}$ and $r \cdot x = 0$ for all r and x, then $\operatorname{span}(X) = \{0\}$ for all $X \subseteq M$, so M has no basis.

Exercise 6.26. If every R-module is free, then R is a division ring.

Lemma 6.27. *Let R be a division ring, M be an R-module, if $I \subseteq M$ is an inclusion maximal independent set, i.e., if $I \subseteq J$ and J is independent, then $I = J$. Then I is a basis.*

Proof. We need to show $\operatorname{span}(I) = M$. Take $x \in M \setminus I$, we want $x \in \operatorname{span}(I)$. Notice by definition of I, $I \cup \{x\}$ is not independent, so for some $(c_y : y \in I \cup \{x\}) \in R^{I \cup \{x\}}$, only finitely many (but at least one) are non-zero, we have $\sum_{y \in I \cup \{x\}} c_y \cdot y = 0 \Rightarrow c_x \cdot x + \sum_{y \in I} c_y \cdot y = 0$, since I is independent, $c_x \neq 0$, so $c_x \cdot x = \sum_{y \in I} -c_y \cdot y \Rightarrow x = \sum_{y \in I} -\left(c_x^{-1} c_y\right) \cdot y$. □

Definition 6.28. Let P be a set, $\preceq$ be a binary relation on P, we say that $\preceq$ is a *partially-ordered relation* if

- $\preceq$ is reflexive
- $\preceq$ is antisymmetric, i.e., for $x, y \in P$, if $x \preceq y$ and $y \preceq x \Rightarrow x = y$
- $\preceq$ is transitive

We say $(P, \preceq)$ is a *partially-ordered set*, or *poset* for short.

Definition 6.29. Let $(P, \preceq)$ be a poset, a subset $C \subseteq P$ is a *chain* if for every $x, y \in C$, $x \preceq y$ or $y \preceq x$. An *upper bound* for C is an element $x \in P$ such that $\forall y \in C$, $y \preceq x$.

Lemma 6.30 (Zorn)**.** *If $(P, \preceq)$ is a poset such that every chain has an upper bound, then P has a maximal element.*

Proof. We refer readers to any online resource for the proof, since it is out of the scope of this course. □

Proposition 6.31. *Let R be a ring, and R be a module over itself by left multiplication, i.e., $\alpha : R \to \mathrm{End}(R)$, $\alpha(r)(a) = ra$. If R is unital, then α is injective.*

Proof. Let $r, s \in R$ with $r \neq s$, then $\alpha(r)(1) = r \cdot 1 = r, \alpha(s)(1) = s \cdot 1 = s$, so $\alpha(r) \neq \alpha(s)$. □

Theorem 6.32. *All modules over a division ring are free.*

Proof. Let R be a division ring, M be an R-module, let $P = \{I \subseteq M : I \text{ is independent}\}$, consider the poset $(P, \subseteq)$, if $(P, \subseteq)$ has a maximal element, then by Lemma 6.27, M has a basis. Let $C \subseteq P$ be a chain, we want to show that C has an upper bound, the claim is $\bigcup C = \bigcup_{I \in C} I$ is the upper bound we want, we need to show that $\bigcup C$ is independent. Suppose not, then $\sum_{x \in \bigcup C} c_x \cdot x = 0$ is a nontrivial linear combination, we write it as $c_1 \cdot x_1 + c_2 \cdot x_2 + \ldots + c_k \cdot x_k = 0$, let $x_i \in I_i \in C$, reorder them as $I_1 \subseteq I_2 \subseteq \ldots \subseteq I_k$, then $x_1, \ldots, x_k \in I_k$, but I_k is independent, a contradiction.

By Lemma 6.30, $(P, \subseteq)$ has a maximal element. □

Next, we want to define $\mathrm{rank}(R^n)$ to be n, but what if $R^n \cong R^m$ for $n \neq m$?

Definition 6.33. A unital ring R has the *Invariant Basis Number*, or *IBN* for short, if for all $n, m \in \mathbb{N}^+, R^n \cong R^m \Rightarrow n = m$.

Theorem 6.34. *Every non-zero division ring has IBN.*

Lemma 6.35 (Exchange property)**.** *Let R be a division ring, let M be an R-module, let $I \subseteq M$ be independent, $B \subseteq M$ a basis. Then for all $x \in I \setminus B$, there exists $y \in B \setminus I$ such that $(I \setminus \{x\}) \cup \{y\}$ is independent.*

Proof of Exchange property. Suppose no such y, this means for any $y \in B \setminus I$, $(I \setminus \{x\}) \cup \{y\}$ is not independent, i.e., $\sum_{a \in (I \setminus \{x\}) \cup \{y\}} c_a \cdot a = 0$, for some coefficients c_a, not all zero. By an argument analogous to Lemma 6.27, we have $y \in \mathrm{span}(I \setminus \{x\})$, so $B \setminus I \subseteq \mathrm{span}(I \setminus \{x\})$, also $B \cap I \subseteq I \setminus \{x\} \subseteq \mathrm{span}(I \setminus \{x\})$, so $B \subseteq \mathrm{span}(I \setminus \{x\})$, since $\mathrm{span}(B) = M$, so $I \setminus \{x\}$ spans M. However, $x \in \mathrm{span}(I \setminus \{x\})$, contradicts I is independent. □

Proof of Theorem. Suppose for a contradiction that $R^n \cong R^m$ but $n \neq m$, without loss of generality, assume $n < m$. Let B_n be a basis for R^n, and B_m be a basis for R^m, so $|B_n| = n, |B_m| = m$. An isomorphism $\varphi : R^m \to R^n$ carries B_m to a basis $\varphi(B_m)$ of R^n with m elements, so both bases now lie in the same module R^n. We then apply the exchange property with $I = \varphi(B_m)$ and $B = B_n$, to "swap" elements of $I \setminus B$ out and replace them with elements of $B \setminus I$. Each swap keeps I independent with m elements and puts one more element of B into I, so after at most m swaps, $I \subseteq B$. Then $m = |I| \leq |B| = n$, contradicting $n < m$. □

6.3 Commutative rings and IBN

In this section, we will prove an important theorem.

Theorem 6.36. *Every non-zero commutative unital ring R has IBN.*

Example 6.37. $\mathbb{Z}$ has the IBN. However, as a $\mathbb{Z}$-module, it has maximal independent set, say $\{2\}$, but it is not a basis. So the idea is to choose a correct module. We provide two proofs.

Proof #1. Suppose $\mathbb{Z}^n \cong \mathbb{Z}^m$ with $n < m$. Notice $\mathbb{Z} \subseteq \mathbb{Q}$, and $\mathbb{Q}$ is a field. Let $B_n := \{x_1, \ldots, x_n\}$ be a basis of $\mathbb{Z}^m$ of size n, so $\text{span}_{\mathbb{Z}}(B_n) = \mathbb{Z}^m$, and $\text{span}_{\mathbb{Q}}(B_n) = \mathbb{Q}^m$. Let's show this. Let $y = \left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \in \mathbb{Q}^m$, let $\tilde{y} = (q_1 \ldots q_m)\, y \in \mathbb{Z}^m$, so $\tilde{y} = c_1 x_1 + \ldots + c_n x_n \Rightarrow y = \frac{c_1}{q_1 \ldots q_m} x_1 + \ldots + \frac{c_n}{q_1 \ldots q_m} x_n$. However, since $\mathbb{Q}$ is a division ring, by Theorem 6.34, $\text{rank}(\mathbb{Q}^m) = m$, contradicts $\mathbb{Q}^m$ can be spanned by B_n: a maximal independent subset J of B_n is a basis of $\mathbb{Q}^m$, since every element of B_n lies in $\text{span}(J)$ as in the proof of Lemma 6.27, so $\mathbb{Q}^m \cong \mathbb{Q}^{|J|}$ with $|J| \leq n < m$. □

Proof #2. Let p be a prime number, $\pi_p : \mathbb{Z} \to \mathbb{Z}_p : k \mapsto k \bmod p$, again let $B_n = \{x_1, \ldots, x_n\}$ be a basis of $\mathbb{Z}^m$ with n elements, define

$$\pi_p^m : \mathbb{Z}^m \to \mathbb{Z}_p^m : (a_1, a_2, \ldots, a_m) \mapsto (\pi_p(a_1), \pi_p(a_2), \ldots, \pi_p(a_m)),$$

so $\text{span}\left(\pi_p^m(B_n)\right) = \mathbb{Z}_p^m$ (here we overload π_p^m to apply on each element of B_n): take $y \in \mathbb{Z}_p^m$, let $\tilde{y} \in \mathbb{Z}^m$ be the element such that $\pi_p^m(\tilde{y}) = y$, write $\tilde{y} = c_1 x_1 + \ldots + c_n x_n$, so $y = \pi_p(c_1)\, \pi_p^m(x_1) + \ldots + \pi_p(c_n)\, \pi_p^m(x_n)$, however, $\mathbb{Z}_p$ is a field, so this contradicts $\mathbb{Z}_p$ has IBN, as at the end of Proof #1. □

We can extend from $\mathbb{Z}^n$ to show something about $\mathcal{F}_n$, the free group.

Corollary 6.38. *Let $n, m \in \mathbb{N}$, then $\mathcal{F}_n \cong \mathcal{F}_m \Rightarrow n = m$.*

Proof. The *abelianization* of $\mathcal{F}_n$ is $\mathbb{Z}^n$, then $\mathcal{F}_n \cong \mathcal{F}_m \Rightarrow \mathbb{Z}^n \cong \mathbb{Z}^m \Rightarrow n = m$. □

For the definition and properties of abelianization, we refer the reader to any online resource.

Let R be a non-zero, commutative, unital ring. For Proof #1 in Example 6.37 to work, we need to embed R into a field, but this might not work: for example, $\mathbb{Z}_6$ cannot be embedded into a field, since $2 \times 3 = 0 \bmod 6$, suppose we can embed $\mathbb{Z}_6$ into some field K, then in K, 3 has inverse 3^{-1}, so $2 \times 3 \times 3^{-1} = 2 \times \left(3 \times 3^{-1}\right) = 2 = (2 \times 3) \times 3^{-1} = 0$, which is not possible. For this to work, one requires additional structure on R.

Definition 6.39. A commutative, non-zero, unital ring R is called *integral domain* if $\forall a, b \in R, a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$.

Theorem 6.40. *A commutative, non-zero, unital ring R can be embedded into a field if and only if R is integral domain.*

Proof. ($\Rightarrow$): We have already shown the contrapositive: the argument for $\mathbb{Z}_6$ works whenever $ab = 0$ with $a, b \neq 0$, since in a field, $a = abb^{-1} = 0$.

($\Leftarrow$): Given an integral domain R, we shall construct a field $\text{Frac}(R)$, called the *field of fractions* of R. Every element of $\text{Frac}(R)$ is a fraction of the form $\frac{a}{b}, a, b \in R, b \neq 0$.

Idea: we want $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$, $\frac{a}{c} \cdot \frac{b}{d} = \frac{ab}{cd}$. Let's try to do so. Define $\tilde{+}, \tilde{\cdot}$ be binary operations on $R \times (R \setminus \{0\})$ via $(a, b) \tilde{+} (c, d) = (ad + bc, bd)$, $(a, b) \tilde{\cdot} (c, d) = (ac, bd)$.

- $\tilde{+}, \tilde{\cdot}$ are commutative, by definition of R
- $\tilde{+}, \tilde{\cdot}$ are associative:

$$\begin{aligned}\big[(a,b)\,\tilde{+}\,(c,d)\big]\,\tilde{+}\,(e,f) &= (ad+bc,bd)\,\tilde{+}\,(e,f)\\ &= (adf+bcf+bde,bdf)\\ (a,b)\,\tilde{+}\,\big[(c,d)\,\tilde{+}\,(e,f)\big] &= (a,b)\,\tilde{+}\,(cf+de,df)\\ &= (adf+bcf+bde,bdf)\end{aligned}$$

  One can follow the same strategy to show $\tilde{\cdot}$ is associative.

- identity for $\tilde{+}$: $(0,1)$, for $\tilde{\cdot}$: $(1,1)$
- distributivity, however, does not hold:

$$\begin{aligned}(a,b)\,\tilde{\cdot}\,\big[(c,d)\,\tilde{+}\,(e,f)\big] &= (a,b)\,\tilde{\cdot}\,(cf+de,df)\\ &= (acf+ade,bdf)\\ [(a,b)\,\tilde{\cdot}\,(c,d)]\,\tilde{+}\,[(a,b)\,\tilde{\cdot}\,(e,f)] &= (ac,bd)\,\tilde{+}\,(ae,bf)\\ &= \big(acbf+aebd,b^2df\big)\end{aligned}$$

The problem is different pairs may represent the same fraction, i.e., $\frac{a}{b}=\frac{c}{d}\Leftrightarrow ad=bc$. Define a binary relation $\sim$ on $R\times(R\setminus\{0\})$ via $(a,b)\sim(c,d)\Leftrightarrow ad=bc$.

Proposition 6.41. *$\sim$ is an equivalence relation.*

Proof. Reflexive and symmetric are obvious, let's focus on transitivity. Suppose $(a,b)\sim(c,d)$ and $(c,d)\sim(e,f)$, then $ad=bc, cf=de$, we want to show: $af=be$.

$$\begin{aligned}ad &= bc\\ adef &= bcef\\ afde &= becf\\ afde &= bede\\ afde-bede &= 0\\ (af-be)\,de &= 0\end{aligned}$$

Since R is an integral domain, either $af-be$ is 0 or de is 0. In the former case, we are done, so let's check for the latter case. If de is 0, then e must be 0, since $cf=de$, c must also be 0, and again, since $ad=bc$, a must be 0, so $af=0=be$, as desired. □

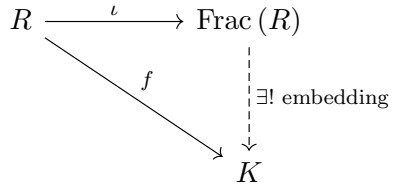
Let's now formally define $\mathrm{Frac}\,(R):=R\times(R\setminus\{0\})\,/\sim$, notation: we write $\frac{a}{b}$ for the $\sim$ class of the pair (a,b).

Define $\frac{a}{b}+\frac{c}{d}=\frac{ad+bc}{bd}$, $\frac{a}{b}\cdot\frac{c}{d}=\frac{ac}{bd}$, one can verify this is well-defined. Now $+,\cdot$ on $\mathrm{Frac}\,(R)$ are associative, commutative, distributive, and have identities:

- Additive inverse: $-\frac{a}{b}=\frac{-a}{b}$: $\frac{a}{b}+\frac{-a}{b}=\frac{0}{1}$, $\frac{ab+b(-a)}{b^2}=\frac{0}{b^2}$, but $0\cdot 1=0\cdot b^2$, so $\frac{0}{b^2}=\frac{0}{1}$.
- Multiplicative inverse: $\left(\frac{a}{b}\right)^{-1}=\frac{b}{a}$ for $\frac{a}{b}\neq\frac{0}{1}$, i.e., $a\neq 0$, the multiplicative identity is $\frac{1}{1}$, $\frac{1}{1}\neq\frac{0}{1}$ since $0\neq 1$ in R.

So $\mathrm{Frac}(R)$ is a field. Define $\iota : R \to \mathrm{Frac}(R) : a \mapsto \frac{a}{1}$, this is an injective unital ring homomorphism. □

Remark 6.42. $\mathrm{Frac}(R)$ is the smallest field containing R, we always have the following diagram:



where f is a unital embedding, K is a field. The embedding from $\mathrm{Frac}(R)$ to K is given by $\frac{a}{b} \mapsto f(a) f(b)^{-1}$.

Corollary 6.43. *Integral domains have IBN.*

Proof. Repeat Proof #1 for $\mathbb{Z}$, with the embedding $\mathbb{Z} \hookrightarrow \mathbb{Q}$ replaced by the embedding of the integral domain into its field of fractions. □

To achieve our goal, i.e., prove commutative rings have IBN, we need to adapt the idea in $\mathbb{Z} \twoheadrightarrow \mathbb{Z}_p$, i.e., find a surjective homomorphism $R \twoheadrightarrow F$, where F is a field.

Let R be a ring, M, N be R-modules, $\varphi : M \to N$ be an R-module homomorphism, we have the following diagram:

$$M \twoheadrightarrow M/\ker(\varphi) \xrightarrow{\cong} \mathrm{im}(\varphi) \hookrightarrow N$$

This is called the *isomorphism theorem for R-modules*. In fact, $\mathrm{im}(\varphi), \ker(\varphi)$ and $M/\ker(\varphi)$ are all R-modules.

- $\mathrm{im}(\varphi)$ is a submodule of N: if $x \in \mathrm{im}(\varphi), r \in R$, then there exists some $y \in M$ with $x = \varphi(y)$, so $r \cdot x = \varphi(r \cdot y) \in \mathrm{im}(\varphi)$.
- $\ker(\varphi)$ is a submodule of M: if $x \in \ker(\varphi), r \in R$, then $\varphi(x) = 0, \varphi(r \cdot x) = r \cdot \varphi(x) = r \cdot 0 = 0$, so $r \cdot x \in \ker(\varphi)$.

Generally speaking, if $M' \subseteq M$ is a submodule, then M/M' becomes an R-module under $r \cdot (x + M') := (r \cdot x) + M'$. This is well-defined: if $x + M' = y + M' \Rightarrow x - y \in M' \Rightarrow r \cdot (x - y) \in M' \Rightarrow r \cdot (x - y) + M' = M' \Rightarrow (r \cdot x) + M' = (r \cdot y) + M'$.

Exercise 6.44. Check this makes M/M' an R-module.

Remark 6.45. If R is a unital ring, M, N be *unital* R-modules, i.e., $1 \cdot x = x$, $\varphi : M \to N$ be a module homomorphism, then $M \twoheadrightarrow M/\ker(\varphi) \cong \mathrm{im}(\varphi) \hookrightarrow N$, all intermediate modules are also unital.

Definition 6.46. Let R be a ring, R is an R-module under left multiplication, submodules of R, viewed as an R-module, are called *(left) ideals*. In other words, $I \subseteq R$ is a (left) ideal if

1. I is a subgroup of $(R, +)$
2. for all $x \in I, r \in R$, we have $r \cdot x \in I$

Note that an ideal is a subring.

Example 6.47.
- For every ring R: $R, \{0\}$.
- $\mathbb{Z}$: $\{0\}, \mathbb{Z}, n\mathbb{Z}$, i.e., every subgroup of $\mathbb{Z}$ is an ideal, but not necessarily unital, for example, $2\mathbb{Z}$ is not unital.

Proposition 6.48. *Let K be a field, the only ideals in K are $\{0\}, K$.*

Proof. Suppose $I \subseteq K$ is an ideal, $\{0\} \subset I \subset K$, take any $x \in I \setminus \{0\}$, then for all $y \in K$, $y = \left(yx^{-1}\right) x \in I$, so $I = K$. □

Proposition 6.49. *Let R be a non-zero commutative, unital ring. R is a field if and only if the only ideals are R and $\{0\}$.*

Proof. ($\Rightarrow$): Proposition 6.48.

($\Leftarrow$): Suppose R is not a field, take $x \neq 0$ with no multiplicative inverse, consider $I := \{rx : r \in R\}$, $I \neq R$, since $1 \notin I$, and it is also non-zero, since $x \in I$. Let's check I is an ideal:

1. $0 \in I$, $rx + sx = (r+s) x \in I, -(rx) = (-r) x \in I$.
2. $s \cdot rx = (sr) x \in I$.

□

Proposition 6.50. *Let I be an ideal in a unital ring R, then $I \neq R$ if and only if $1 \notin I$, i.e., I is a proper ideal.*

Proof. ($\Rightarrow$) : Let's do the contrapositive: if $1 \in I$, then for any $r \in R, r \cdot 1 = r \in I$, so $I = R$.

($\Leftarrow$) : Clear, since $1 \in R$. □

Let R, Q be rings, $\varphi : R \to Q$ be a ring homomorphism, consider the diagram:

$R \twoheadrightarrow R/\ker(\varphi) \xrightarrow{\cong} \operatorname{im}(\varphi) \hookrightarrow Q$

where $\ker(\varphi)$ is an ideal of R: it's a subgroup, and for $x \in \ker(\varphi), \varphi(x) = 0 \Rightarrow \varphi(rx) = \varphi(r)\varphi(x) = \varphi(r) \cdot 0 = 0$, so $rx \in \ker(\varphi)$. $\operatorname{im}(\varphi)$ is a subring of Q: for $x, y \in \operatorname{im}(\varphi), \varphi(a) = x, \varphi(b) = y \Rightarrow \varphi(ab) = xy$, so $xy \in \operatorname{im}(\varphi)$.

Remember the group situation: $\varphi : G \to H$, then

$G \twoheadrightarrow G/\ker(\varphi) \xrightarrow{\cong} \operatorname{im}(\varphi) \hookrightarrow H$ where $\ker(\varphi)$ is normal, so $G/\ker(\varphi)$ is a group, and $\operatorname{im}(\varphi)$ is a subgroup of H. In our ring analogy, $\ker(\varphi)$ is not only a left ideal, but also a right ideal: for $x \in \ker(\varphi), r \in R$, $xr \in \ker(\varphi)$. In fact, $\ker(\varphi)$ is a *two-sided* ideal, both left and right ideal.

Let R be a ring, $I \subseteq R$ be a two-sided ideal, then R/I becomes a ring under multiplication: $(x+I)(y+I) := (xy) + I$, let's check it's well-defined: suppose $x + I = x' + I, y + I = y' + I$, want to show that $(xy) + I = (x'y') + I$. Notice $x - x', y - y' \in I$, so $(x - x') y = xy - x'y \in I, x'(y - y') = x'y - x'y' \in I \Rightarrow xy - x'y + x'y - x'y' = xy - x'y' \in I \Rightarrow (xy) + I = (x'y') + I$.

Exercise 6.51. Check R/I is a ring.

Theorem 6.52. *Let R be a non-zero, commutative, unital ring, then there exists a surjective homomorphism $R \twoheadrightarrow K$, where K is a field.*

For what ideals $I \subseteq R$ is R/I a field?

Example 6.53. $R = \mathbb{Z}$, then $\mathbb{Z}/I$ is a field for $I = p\mathbb{Z}$, where p is a prime number. $R/I \neq \{0\} \Leftrightarrow I \neq R$, so I is a proper ideal.

Lemma 6.54. *Let R be a commutative, unital ring, $I \subseteq R$ be an ideal, then there is a one-to-one correspondence between $\{J \subseteq R/I, \text{an ideal}\}$ and $\{I' \subseteq R, \text{an ideal}, I \subseteq I'\}$, the bijective mapping is given by $J \mapsto \{x \in R : x + I \in J\} = \bigcup J$, and the inverse is $I' \mapsto \{x + I : x \in I'\}$.*

Proof. We'll verify the direction $I' \mapsto \{x + I : x \in I'\}$, let $I \subseteq I' \subseteq R$, and $J := \{x + I : x \in I'\}$, we need to check that J is an ideal of R/I. This is the case, since J is the image of I' under quotient map $R \to R/I$, so J is a subgroup of R/I. To see it's an ideal, take $x + I \in J, x \in I'$, and $y + I \in R/I$, then $(y + I)(x + I) = (yx) + I \in J$, since $yx \in I'$.

Now let's consider the inverse, let $I'' = \{y \in R : y + I \in J\}$, the goal is to show $I' = I''$. Since $x \in I' \Rightarrow x + I \in J$, so $x \in I'' \Rightarrow I' \subseteq I''$.

For the other inclusion, suppose $y + I \in J$, then there exists $x \in I'$ such that $x + I = y + I \Rightarrow x - y \in I \subseteq I'$, so $y \in I' \Rightarrow I'' \subseteq I'$. Thus, this is a valid bijection.

The other direction is left as an exercise to the reader. □

Let $I \subseteq R$ be a proper ideal, consider the following chain of logical equivalence:

$$\begin{aligned} R/I \text{ is a field} &\Leftrightarrow \text{the only ideals in } R/I \text{ are } \{0\}, R/I \\ &\Leftrightarrow \text{the only ideals in } R \text{ containing } I \text{ are } I \text{ and } R \\ &\Leftrightarrow \text{the only proper ideal containing } I \text{ is } I \\ &\Leftrightarrow I \text{ is inclusion maximal among all proper ideals} \end{aligned}$$

Theorem 6.55 (Krull)**.** *Let R be a commutative, unital, non-zero ring, then R has a maximal proper ideal.*

Proof. We will use Lemma 6.30. Let $P := \{I \subset R : I \text{ is a proper ideal}\}$, consider the poset $(P, \subseteq)$, we want to show every chain has an upper bound in this poset. Take $C \subseteq P$ be a chain, if $C = \emptyset$, then $\{0\}$ is an upper bound, otherwise, $C \neq \emptyset$, then $\bigcup C$ is an upper bound for C. Let's check $\bigcup C$ is an ideal, take $x \in \bigcup C, r \in R$, suppose $x \in I_i$, then $r \cdot x \in I_i \subseteq \bigcup C$. Take $x, y \in \bigcup C$, $x \in I_1 \subseteq \bigcup C, y \in I_2$, so $I_2 \subseteq I_1$ or $I_1 \subseteq I_2$, this means $x + y$ belongs to the larger ideal, which is a subset of $\bigcup C$. So $\bigcup C$ is an ideal.

Is $\bigcup C$ proper? Notice $1 \notin \bigcup C$, this is true because for any $I \in C, 1 \notin I$.

So every chain in $(P, \subseteq)$ has an upper bound, by Zorn, it has a maximal element. □

Corollary 6.56. *If I is a proper ideal, then there is a maximal ideal I' containing I, i.e., $I \subseteq I'$.*

Proof. Apply Theorem 6.55 to R/I, which is non-zero since I is proper, obtaining a maximal proper ideal in R/I, call it J. Using the bijective mapping from the ideals of R/I to the ideals of R containing I in Lemma 6.54, map J to $\bigcup J$, clearly, $I \subseteq \bigcup J$, so it suffices to show $\bigcup J$ is a maximal proper ideal in R. Suppose it is not maximal, i.e., $\bigcup J \subsetneq I'' \subsetneq R$ for some ideal I''. The mapping in Lemma 6.54 and its inverse preserve inclusions, so the ideal $\{x + I : x \in I''\}$ of R/I lies strictly between J and R/I, which contradicts J is maximal. $\bigcup J$ is proper since J is proper. □

Let's now prove Theorem 6.52.

Proof. Use Theorem 6.55 to get a maximal proper ideal $I \subset R$, consider R/I, it is a field by our above reasoning. The surjective homomorphism: $R \twoheadrightarrow R/I$ is then given by the quotient map. □

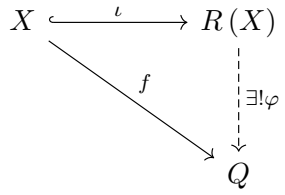
Now we have enough tools to prove Theorem 6.36.

Proof. Suppose for a contradiction that $R^m \cong R^n$ but $n < m$, take a basis $x_1, \ldots, x_n \in R^m$, fix a surjective homomorphism $\pi : R \twoheadrightarrow K$ with K a field, set $\pi^m : R^m \to K^m : (a_1, a_2, \ldots, a_m) \mapsto (\pi(a_1), \pi(a_2), \ldots, \pi(a_m))$, then the set $\{\pi^m(x_1), \ldots, \pi^m(x_n)\}$ spans K^m, but $n < m$, K is a field, so it's impossible, as at the end of Proof #1 in Example 6.37.

For the spanning part, take any $y \in K^m$, π is surjective, so $y = \pi^m(z)$ for some $z \in R^m$, where $z = a_1 x_1 + \ldots + a_n x_n$, for some $a_1, \ldots, a_n \in R$, therefore $y = \pi(a_1)\pi^m(x_1) + \ldots + \pi(a_n)\pi^m(x_n)$, so the set is spanning. □

6.4 Polynomials

We have studied free groups and free modules, so it's time to consider what are free (unital, commutative) rings? The diagram always looks similar:



where $R(X)$ is a commutative unital ring, φ is a unital ring homomorphism, and Q is also a commutative unital ring.

Let's first consider $X = \emptyset$, then $R(\emptyset) \cong \mathbb{Z}$, since the only restriction has become φ to be a unique unital ring homomorphism, and we achieve so by $0 \mapsto 0_Q, 1 \mapsto 1_Q$. What if $X = \{x\}$? We will show $R(\{x\}) \cong \mathbb{Z}[x]$, the *polynomial ring in one variable over* $\mathbb{Z}$.

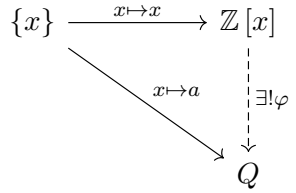
Definition 6.57. A *polynomial* over $\mathbb{Z}$ in one variable x is a sequence $p = (a_0, a_1, \ldots) \in \mathbb{Z}^{\mathbb{N}}$, such that only finitely many a_i's are nonzero, usually, we write $a_0 + a_1 x + a_2 x^2 + \ldots = \sum_{i=0}^{\infty} a_i x^i = (a_0, a_1, \ldots)$.

The set of all such polynomials is denoted by $\mathbb{Z}[x]$. $\mathbb{Z}[x]$ is an abelian group under $+$, defined by $(a_0, a_1, \ldots) + (b_0, b_1, \ldots) = (a_0 + b_0, a_1 + b_1, \ldots)$. As an abelian group, $\mathbb{Z}[x] \cong \bigoplus_{\mathbb{N}} \mathbb{Z}$, we turn $\mathbb{Z}[x]$ into a commutative, unital ring by defining a product. For notation, let $[x^n](a_0, a_1, \ldots) := a_n$, for example, $\left[x^2\right]\left(1 + 3x - x^2 + 7x^8\right) = -1$. The product is defined as follows: for $p, q \in \mathbb{Z}[x]$, $[x^n](pq) := \sum_{i+j=n, i,j \in \mathbb{N}} \left[x^i\right](p) \cdot \left[x^j\right](q)$, or we can write it as $\left(\sum_i a_i x^i\right)\left(\sum_j b_j x^j\right) = \sum_n c_n x^n$, where $c_n = \sum_{i+j=n} a_i b_j$.

Multiplicative identity: $(1, 0, 0, 0, \ldots)$, and $x \mapsto (0, 1, 0, 0, \ldots)$. Observe that $(0, 1, 0, 0, \ldots)^n = \left(0, 0, 0, \ldots, 0, \underbrace{1}_{n}, 0, \ldots\right)$.

Exercise 6.58. Check $\mathbb{Z}[x]$ is indeed a commutative ring.

We can now complete the diagram:



For $p = (c_0, c_1, c_2, \ldots) \in \mathbb{Z}[x]$, define $\varphi(p) = c_0 + c_1 a + c_2 a^2 + \ldots$, with only finitely many c_i's are nonzero. By our definition, $\varphi(p)$ is essentially $p(a)$, polynomial p evaluates at point a.

$\mathbb{Z}[x]$ is a free abelian group, generated by $(1, 0, 0, \ldots), (0, 1, 0, \ldots), (0, 0, 1, 0, \ldots), \ldots$.

What if $X = \{x, y\}$? We can extend to $\mathbb{Z}[x, y]$, *integer polynomials over 2 variables.*

Definition 6.59. A *polynomial over* $\mathbb{Z}$ *in two variables* x, y is denoted by $\mathbb{Z}[x, y]$, defined as the set $\left\{\text{functions } c : \mathbb{N} \times \mathbb{N} \to \mathbb{Z}, \text{ such that } \left\{(i, j) \in \mathbb{N}^2 : c(i, j) \neq 0\right\} \text{ is finite}\right\}$.

Addition is defined pointwise, so $\mathbb{Z}[x, y] \cong \bigoplus_{\mathbb{N} \times \mathbb{N}} \mathbb{Z}$, and multiplication is defined as follows: for $p, q \in \mathbb{Z}[x, y]$, $[x^n y^m](pq) := \sum_{a,b,c,d \in \mathbb{N}, (a,b)+(c,d)=(n,m)} \left[x^a y^b\right](p) \cdot \left[x^c y^d\right](q)$.

Define the *evaluation of p on point* (a,b), for $p \in \mathbb{Z}[x,y]$ and $a,b \in Q$, by

$$p(a,b) := \sum_{(n,m)\in\mathbb{N}^2} ([x^n y^m](p))\, a^n b^m.$$

Exercise 6.60. Check the map $\mathbb{Z}[x,y] \to Q : p \mapsto p(a,b)$ is a unital ring homomorphism.

If X is infinite, say $X = \{x_0, x_1, x_2, \ldots\}$, then we can characterize the monomials as $x_0^{t_0} x_1^{t_1} x_2^{t_2} \ldots$, where $t_0, t_1, t_2, \ldots \in \mathbb{N}$, finitely many of them are nonzero.

Definition 6.61. Let R be a commutative, unital ring, and X be a set, a *polynomial over R in variable X*, denoted by $R[X]$, is a polynomial with coefficients in R, with variable set X, and finitely many coefficients are nonzero.

Definition 6.62. Let M be an R-module, where R is a commutative, unital ring. An *R-endomorphism* of M is a homomorphism $f : M \to M$ of R-modules, i.e.

- f is a homomorphism of groups $(M,+) \to (M,+)$.
- for all $r \in R, x \in M$, $f(r \cdot x) = r \cdot f(x)$.

One can interpret an R-endomorphism as a standard endomorphism, with the additional constraint that f is linear in $r \cdot x$.

Let $\mathrm{End}_R(M) := \{\text{all } R\text{-endomorphisms of } M\} \subseteq \mathrm{End}(M,+)$.

- It is a ring: closed under addition, additive inverses, contains 0, and closed under composition.
- It is an R-module: define for $r \in R$ and $f \in \mathrm{End}_R(M)$, $r \cdot f : M \to M$ via $(r \cdot f)(x) = r \cdot (f(x))$.

  –

$$\begin{aligned}(r \cdot f)(x+y) &= r \cdot (f(x+y)) \\ &= r \cdot (f(x) + f(y)) \\ &= r \cdot f(x) + r \cdot f(y) \\ &= (r \cdot f)(x) + (r \cdot f)(y)\end{aligned}$$

  –

$$\begin{aligned}(r \cdot f)(s \cdot x) &= r \cdot (f(s \cdot x)) \\ &= r \cdot (s \cdot f(x)) \\ &= r \cdot s \cdot f(x) \\ &= s \cdot r \cdot f(x) \\ &= s \cdot (r \cdot f)(x)\end{aligned}$$

Exercise 6.63. Finish the check that $\mathrm{End}_R(M)$ is an R-module.

Definition 6.64. A *concrete R-algebra* is a subset $A \subseteq \mathrm{End}_R(M)$ that is both a subring and submodule of $\mathrm{End}_R(M)$.

An *(abstract) R-algebra* is a set A together with a ring structure $(A,+,\cdot)$ and a (unital) R-module structure such that for all $r,s \in R, x,y \in A, (r \cdot x) \cdot (s \cdot y) = (rs) \cdot (xy)$.

So $\mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^2\right) \cong 2 \times 2$ real matrices, since any $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^2\right)$ must preserve linear scaling, we won't have something like $f(1) = \pi, f\left(\sqrt{2}\right) = e$.

Proposition 6.65. $\mathrm{End}_{\mathbb{Z}}\left(\mathbb{R}^2\right) = \mathrm{End}_{\mathbb{Q}}\left(\mathbb{R}^2\right)$.

Proof. $(\supseteq)$: Clear, if f is linear on all rationals, then it must be linear on all integers.

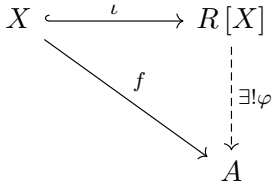
$(\subseteq)$: Take $f \in \mathrm{End}_{\mathbb{Z}}\left(\mathbb{R}^2\right)$, we want to show: $f \in \mathrm{End}_{\mathbb{Q}}\left(\mathbb{R}^2\right)$, i.e., for all $q \in \mathbb{Q}, x \in \mathbb{R}^2, f(qx) = qf(x)$. Write $q = \frac{s}{t}$ for $s, t \in \mathbb{Z}, t \neq 0$, so the question is whether $\frac{s}{t} f(x) = f\left(\frac{s}{t}x\right)$. Notice that $f(sx) = f\left(\frac{s}{t}tx\right) = sf(x)$, and $f\left(\frac{s}{t}tx\right) = tf\left(\frac{s}{t}x\right)$, so $sf(x) = tf\left(\frac{s}{t}x\right) \Rightarrow \frac{s}{t}f(x) = f\left(\frac{s}{t}x\right)$. □

Exercise 6.66. Show that $\mathbb{R}$ as a $\mathbb{Q}$-vector space is infinitely dimensional, in fact, each of the following infinite subsets of $\mathbb{R}$ is $\mathbb{Q}$-linearly independent:

- (Tricky): $1, \sqrt{2}, \sqrt{3}, \sqrt{5}, \sqrt{7}, \ldots$
- (Easy): $\ln 2, \ln 3, \ln 5, \ln 7, \ldots$
- (Medium): Give an example of an uncountable such set.

Notice that $\mathrm{End}_{\mathbb{R}}(\mathbb{R}) \cong \mathbb{R}$, but $\mathrm{End}_{\mathbb{Q}}(\mathbb{R})$ is much bigger.

Analogous to $\mathbb{Z}[X]$ as the free ring over set X, $R[X]$ is a *free* commutative, unital R-algebra:



Here A is a commutative, unital R-algebra, and φ is a unital R-algebra homomorphism.

From now on, we focus on the ring $R[x]$.

Definition 6.67. Let R be a ring, $r \in R$ is called *unit* if it has multiplicative inverse.

If R is an integral domain, then in $R[x]$, units are exactly the nonzero constants whose coefficients are units, since $\deg(fg) = \deg(f) + \deg(g)$ for nonzero f, g. This fails in general: in $\mathbb{Z}_4[x]$, $(1+2x)(1+2x) = 1$.

Recall the fundamental theorem of arithmetic: every natural number $n \geq 1$ admits a unique prime factorization. In the proof, we first use a strong induction to show every natural number *has* a prime factorization, then we use a contradicting argument to show it's unique. We remind reader the proof:

Theorem 6.68 (Fundamental Theorem of Arithmetic)**.** *Every natural number $n \geq 1$ has a unique prime factorization, up to the order of the factors.*

Proof. **Existence:** Induction on n. For $n = 1$, it's clear. Otherwise, either n is a prime number, then we are done, or n is a composite, let p be one of its prime divisors, then by induction hypothesis, n/p has a prime factorization, plus p, we get a prime factorization for n.

Uniqueness: Suppose otherwise, and let n be the smallest natural number with two different prime factorizations: $n = p_1 \ldots p_s = q_1 \ldots q_t$. Then $n > 1$, so $s \geq 1$, and $p_1 \mid n \Rightarrow p_1 \mid q_1 \ldots q_t$. Recall *Euclid's Lemma*: if p is a prime, $p \mid ab \Rightarrow p \mid a$ or $p \mid b$, so $p_1 \mid q_1$ or $p_1 \mid q_2 \ldots q_t$, we iteratively apply this argument to the product part, finally we will have $p_1 \mid q_j$ for some $1 \leq j \leq t$, so $p_1 = q_j$, since q_j is prime. Cancelling $p_1 = q_j$ from both factorizations gives two different prime factorizations of $n/p_1 < n$, contradicting the minimality of n. □

To this end, we will adapt the same argument to show that any nonzero polynomial in $F[x]$, where F is a field, admits a unique "prime" factorization.

Definition 6.69. Let R be a ring, a nonunit $a \in R$ is *irreducible* if it cannot be written as $a = bc$, for nonunits b, c.

Definition 6.70. A nonzero nonunit $a \in R$ is *prime* if $\forall b, c \in R, a \mid bc \Rightarrow a \mid b$ or $a \mid c$.

Definition 6.71. Let $a, b \in R$, the *greatest common divisors of a,b*, denoted by $\gcd(a, b)$, is the set $\{g : g \mid a \wedge g \mid b, \forall h, h \mid a \wedge h \mid b \Rightarrow h \mid g\}$.

Proposition 6.72. *Let R be an integral domain. Any two gcd's are related by units.*

Proof. Suppose $g_1, g_2 \in \gcd(a, b)$, then $g_1 \mid g_2, g_2 \mid g_1$, so there exists k, ℓ, $g_1 = kg_2, g_2 = \ell g_1 \Rightarrow g_2 = k\ell g_2$. If $g_2 = 0$, then $g_1 = kg_2 = 0 = g_2$. Otherwise, we can cancel $g_2 \neq 0$, since R is an integral domain, which gives $k\ell = 1 \Rightarrow k, \ell$ are units. □

In our familiar $\mathbb{Z}$, we have $\gcd(a, b) \neq \emptyset, \gcd(a, b) = \gcd(a - kb, b), \gcd(a, 0) = \{\pm a\}$. Moreover, for any $n \in \mathbb{Z}$, n is prime $\Leftrightarrow$ n is irreducible. We prove so by using the *Bézout identity*.

Lemma 6.73 (Bézout identity)**.** *For any $a, b \in \mathbb{Z}$ and any $g \in \gcd(a, b)$, there exist $x, y \in \mathbb{Z}$ such that $ax + by = g$.*

Lemma 6.74 (Euclid's Lemma)**.** *In $\mathbb{Z}$, prime and irreducible are equivalent.*

Proof. Suppose p is irreducible in $\mathbb{Z}$ and $p \mid ab$, without loss of generality, assume $p \nmid a$, then $1 \in \gcd(a, p)$, since the only divisors of p are ± 1 and $\pm p$, so by Lemma 6.73, there exist x, y such that $px + ay = 1 \Rightarrow pbx + aby = b$, since $p \mid ab, ab = kp$ for some k, so $bxp + kyp = b \Rightarrow p \mid b$. For the other direction, if p is prime and $p = bc$, then $p \mid b$ or $p \mid c$, say $b = pd$, then $p = pdc$, and cancelling $p \neq 0$ gives $dc = 1$, so c is a unit, i.e., p is irreducible. □

How to prove the Bézout identity? We will make use of *Euclidean algorithm*, which relies on *division algorithm*, or more concretely, *division with remainders*.

Consider $R = F[x]$, where F is some field, we need to define division with remainders for polynomials.

Lemma 6.75 (Division algorithm)**.** *For any $a(x), b(x) \in F[x]$, $b(x) \neq 0$, there exist $q(x), r(x)$ such that $a(x) = q(x)b(x) + r(x)$, such that $\deg(r) < \deg(b)$ or $r = 0$.*

Proof. If $a = 0$, take $q = r = 0$. Otherwise, let $n := \deg(a), m := \deg(b)$, we prove by induction on n. Consider two cases.

Case 1 $m > n$, then $q(x) = 0, r(x) = a(x)$.

Case 2 $m \leq n$, write $a(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, with $a_n \neq 0$, and $b(x) = b_0 + b_1x + b_2x^2 + \ldots + b_mx^m$, consider $a'(x) := a(x) - \frac{a_n}{b_m}x^{n-m}b(x)$, so $\deg(a') < n$ or $a' = 0$, by induction hypothesis or the case $a' = 0$ above, $a'(x) = q'(x)b(x) + r(x)$, take $q(x) = q'(x) + \frac{a_n}{b_m}x^{n-m}$, then $a(x) = q(x)b(x) + r(x)$.

□

The proof also gives an algorithm, which is referred as the *division algorithm*.

```
Algorithm Division:
     input: a(x), b(x) != 0;
     output: q(x), r(x);
     n = deg(a);
     m = deg(b);
     // Base case
     if n < m:
          q(x) = 0;
          r(x) = a(x);
     else:
          a'(x) = a(x) - (a_n / b_m)x^(n-m) b(x);
          // Degree of a' < a
          (q'(x), r(x)) = Division(a'(x), b(x));
          q(x) = q'(x) + (a_n / b_m)x^(n-m);
     return (q(x), r(x));
```

Let's now prove Euclidean algorithm works in $F[x]$.

Lemma 6.76 (Euclidean algorithm)**.** *For any $a(x), b(x), q(x)$,*

$$\gcd(a(x), b(x)) = \gcd(a(x) - q(x)\, b(x), b(x)).$$

Proof. $(\subseteq)$: Given $g \in \gcd(a, b)$, $g \mid a, b \Rightarrow g \mid qb, \forall q$, so $g \mid a - qb$.

$(\supseteq)$: $g \mid a - qb, g \mid b \Rightarrow g \mid a$. So the pairs a, b and $a - qb, b$ have the same common divisors, thus the same greatest common divisors. □

Now we are ready to prove Bézout identity for polynomials.

Lemma 6.77 (General Bézout identity)**.** *For any $a(x), b(x) \in F[x]$, there exist $w(x), z(x)$, such that $a(x)\, w(x) + b(x)\, z(x) = g(x)$, where $g(x) \in \gcd(a(x), b(x))$.*

Proof. We will perform induction on the degree of b, with $b = 0$ as the base case: then $a \in \gcd(a, 0)$ and $a = 1 \cdot a + 0 \cdot b$. If $b \neq 0$, using Lemma 6.75, we write $a(x) = q(x)\, b(x) + r(x)$ with $r = 0$ or $\deg(r) < \deg(b)$, by Lemma 6.76, $\gcd(a(x), b(x)) = \gcd(r(x), b(x))$, so by induction hypothesis applied to the pair b, r,

$$\begin{aligned} g(x) &= w(x)\, r(x) + z(x)\, b(x) \\ &= w(x)\,(a(x) - q(x)\, b(x)) + z(x)\, b(x) \\ &= w(x)\, a(x) - w(x)\, q(x)\, b(x) + z(x)\, b(x) \\ &= w(x)\, a(x) + (z(x) - w(x)\, q(x))\, b(x) \end{aligned}$$

□

With Bézout in hand, we get a result similar to the one for integers: irreducible polynomials are primes, i.e., $p(x) \mid a(x)\, b(x) \Rightarrow p(x) \mid a(x)$ or $p(x) \mid b(x)$.

Theorem 6.78. *Let F be a field, then for any nonzero $f(x) \in F[x]$, $f(x)$ admits a unique prime factorization, up to the order of the factors and multiplication by units.*

In the last part of these notes, we will take a look at *roots* or *zeros* of polynomials.

Theorem 6.79. *Let K be a field, $p \in K[x]$ be a monic polynomial of degree ≥ 1, where monic means the coefficient of highest degree is 1. Then there is a field L such that $K \subseteq L$, p has a root in L.*

Proof. Let's introduce a new variable t, and consider $K[t]$. Let I be the ideal in $K[t]$, generated by $p(t)$, i.e., $I := \{q(t)\, p(t) : q(t) \in K[t]\}$. Look at $K[t]/I$, define $\varphi(t) = t + I \in K[t]/I$, so

$$\begin{aligned}
p(t+I) &= a_0 + a_1\varphi(t) + a_2\varphi(t)^2 + \ldots + \varphi(t)^d \\
&= a_0 + a_1\varphi(t) + a_2\varphi\left(t^2\right) + \ldots + \varphi\left(t^d\right) && \text{since the quotient map is a homomorphism} \\
&= a_0 + \varphi(a_1 t) + \varphi\left(a_2 t^2\right) + \ldots + \varphi\left(t^d\right) \\
&= \varphi\left(a_0 + a_1 t + a_2 t^2 + \ldots + t^d\right) \\
&= \varphi(p(t)) \\
&= p(t) + I \\
&= I && \text{since } p(t) \in I \\
&= 0_{K[t]/I}
\end{aligned}$$

Example 6.80. Consider $p(x) = x^2 + 1 \in \mathbb{R}[x]$, then the field we consider is $\mathbb{R}[t]/\langle t^2+1\rangle$, elements are of the form $q(t)+I$, and $p(t+I) = (t+I)^2+(1+I) = t^2+I+1+I = \left(t^2+1\right)+I = I$.

It remains to show $K[t]/I$ is a field.

- $K[t]/I$ is nonzero, especially, $1 \notin I$. Notice $\deg(p(t)) \geq 1$, so for any nonzero $q(t) \in K[t]$, $\deg(q(t)\, p(t)) \geq 1$, but 1 has degree 0, so $1 \neq q(t)\, p(t)$, also $0 \cdot p(t) = 0 \neq 1$. Thus, $1 \notin I$, and $1 + I \in K[t]/I$.
- $K[t]/I$, in fact, might fail to be a field, unless $p(x)$ is irreducible. More formally, if $p(x)$ is an irreducible polynomial with degree ≥ 1, then the ideal in $K[t]$ generated by $p(t)$ is maximal. Suppose not, let J be a proper ideal of $K[t]$ such that $\langle p(t)\rangle \subset J$, take any $q(t) \in J \setminus \langle p(t)\rangle$, let $g(t)$ be a gcd of $p(t), q(t)$ as in Lemma 6.77, then $g(t)$ is not a unit multiple of $p(t)$, since $q(t) \notin \langle p(t)\rangle$ and $g \mid q$. Since $g \mid p$ and p is irreducible, $g(t)$ is a unit, but by Lemma 6.77, $g(t) = s(t)\, p(t) + f(t)\, q(t) \in \langle p(t), q(t)\rangle \subseteq J$, which means $1 \in J$, J is no longer proper, a contradiction.

By Theorem 6.78, we factor $p(x) = p_1(x)\, p_2(x) \ldots p_k(x)$, where all of them are irreducible, then pick one of $p_i(x)$ to generate $\langle p_i(t)\rangle$, and take a quotient, we get our desired result. □

Exercise 6.81. Every ideal in $K[t]$ is generated by one element.

We can even embed K into $K[t]/I$. Let p be an irreducible polynomial, then $K[t]/\langle p(t)\rangle$ is a field, moreover, the embedding $K \hookrightarrow K[t]/\langle p(t)\rangle : a \mapsto (a, 0, 0, \ldots) + I$, is injective. The injectivity comes from the fact that, except for 0, all other polynomials of degree 0 are not in $\langle p(t)\rangle$.

Theorem 6.82. *Let K be a field, there is a field L with $K \subseteq L$, such that L is algebraically closed, i.e., every polynomial $p \in L[x]$ of degree ≥ 1 has a root in L.*

Proof. Iteratively apply Theorem 6.79, until every polynomial has a root. This is only a sketch: each new field has new polynomials, so this process need not end, and a complete proof needs a construction that handles all polynomials at once, e.g., with Zorn's lemma (Lemma 6.30). □

Finally, we are ready to state the *Fundamental Theorem of Algebra*, though it is not that fundamental, and it is not a theorem of pure algebra: every proof uses some analysis, such as the completeness of $\mathbb{R}$. We state it without proof.

Theorem 6.83 (Fundamental Theorem of Algebra)**.** $\mathbb{C}$ *is algebraically closed.*