MATH 4580
SPRING 2026 LECTURE NOTES

RICHARD HABURCAK

Contents

Forward

These are lecture notes from a Spring semester 2026 course of Math 4580, Undergraduate Abstract Algebra 1, at The Ohio State University. This course is the first in a sequence, followed by Math 4581, Undergraduate Abstract Algebra 2. The main topics of this course are an introduction to groups, their structure, and examples, as well as some ring and field theory, both through the isomorphism theorems, with some extra topics included, for example the Jordan–Hölder theorem, unique factorization domains, and some facts about polynomial rings over fields.

Acknowledgments. These notes are collected from a few sources, one of my favorite being Serge Lang's *Undergraduate Algebra*, third edition, as well as Serge Lang's *Algebra*, and notes from a previous course graciously given to me by Léo Jimenez. I am also indebted to the students for carefully reading the notes and pointing out typos and other improvements.

Remarks. When possible, I tried to make these notes self-contained, assuming a background in writing proofs, though many facts or details will be recalled or explained when needed. I will sometimes write "recall", which should be understood as "if this is familiar, move on; and if not, then I suggest you look it up before proceeding". But don't let it stop you from reading through the notes! Black-boxing is an important skill to develop, though a dangerous one. There may be typos or mistakes scattered throughout these notes, though I hope none are completely derailing. If you notice any mistakes, please let me know. And if you found these notes useful, I would also appreciate any feedback.

Introduction

The purpose of this course is to introduce a few central topics in modern mathematics, namely the concepts of a *group* and a *ring*, which both arose in the study of solving polynomial equations, and have now cemented themselves as foundational concepts in practically all fields of modern mathematics.

1. Sets, Functions, and Relations

Most of the objects we'll be studying are sets with some *extra structure*, making them a bit more interesting and useful. Let's recall some facts about sets and maps that we'll use throughout the course, most of this section should be review from a previous proof-based math class. If you find that any of the topics here are not comfortable, I highly recommend the book *The Art of Proof* by Matthias Beck and Ross Geoghegan, available as a pdf on the first author's website here https://matthbeck.github.io/aop.html.

1.1. Sets.

This will likely be our least precise definition, as taking the time to give a proper definition would be a course in and of itself! For the *very interested and motivated* reader, I encourage you to delve into Zermelo–Fraenkel set theory (ZFC) (with the axiom of choice), though it is much beyond the scope of this course. And in practice, we won't concern ourselves too much with the set-theoretic foundations of the subject.

Definition 1.1.1. A *set* is a collection of objects called its *elements*. If X is a set, and x is an element of the set X, we write $x \in X$ to mean "x is an element of X". If an object z is not an element of X, we write $z \notin X$.

For example, the integers,

$$\mathbb{Z} \coloneqq \{\dots, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$$

are a set. Another useful set is $\emptyset \coloneqq \{\}$, which is called the *empty set*, and has no elements.

Definition 1.1.2. If X and X' are sets, and if every element of X' is an element of X, then we say X' is a *subset* of X, and write $X' \subseteq X$.

Remark 1.1.3. Note that our definition of subset does not exclude the possibility that $X' = X$, that is the sets X and X' have precisely the same elements.

Definition 1.1.4. If $X' \subseteq X$, and $X' \neq X$, then we say X' is a *proper subset of* X, and usually write $X' \subset X$ or $X' \subsetneq X$ to emphasize that X' is a subset of X but not equal to X. We will also commonly say that "the set X' is contained in X".

Definition 1.1.5. If X and Y are sets, then we write $X = Y$ when $X \subseteq Y$ and $Y \subseteq X$.

Example 1.1.6. For example, the set of non-negative whole numbers $\{0, 1, 2, 3, \dots\}$ is a proper subset of $\mathbb{Z}$. To avoid any confusion, we'll usually denote the set of non-negative integers by

$$\mathbb{Z}_{\geq 0} \coloneqq \{0, 1, 2, 3, \dots\} \subset \mathbb{Z}.$$

The set of positive integers

$$\mathbb{Z}_{>0} \coloneqq \{1, 2, 3, \dots\}$$

is a subset of both $\mathbb{Z}$ and $\mathbb{Z}_{\geq 0}$.

Definition 1.1.7. If X_1 and X_2 are sets, the *intersection of* X_1 *and* X_2, denoted by

$$X_1 \cap X_2,$$

is the set of elements which lie in both X_1 and in X_2. The *union of* X_1 *and* X_2, denoted by

$$X_1 \cup X_2,$$

is the set of elements which lie in X_1 or in X_2.

Example 1.1.8. For instance, let $X_1 = \{\dots, -3, -2, -1, 0\}$ and let $X_2 = \mathbb{Z}_{\geq 0}$. Then $X_1 \cap X_2 = \{0\}$ and $X_1 \cup X_2 = \mathbb{Z}$.

Example 1.1.9. If $X_3 = \{\dots, -3, -2, -1\}$, then $X_2 \cap X_3 = \emptyset$.

A common way of writing sets is as a collection of objects satisfying some condition using *set-builder notation.* If P is some property that objects can have, we can form a collection of objects satisfying the property P by writing

$$\{x \mid x \text{ satisfies property } P\}.$$

A word of caution: this might not always be a set! So to make sure the collection of objects we work with form sets, we'll usually use set-builder notation starting with objects x is some set X, and write

$$\{x \in X \mid x \text{ satisfies property } P\}$$

to mean the subset of X of elements that have the property P.

Example 1.1.10. We can write the set of positive even numbers in many ways, for example as

$$\{n \in \mathbb{Z}_{>0} \mid 2 \text{ divides } n\}.$$

Definition 1.1.11. If $Y \subseteq X$ are sets, we denote by $X \setminus Y$ the subset of elements of X that are not in Y, called the *complement of* Y *in* X. In set-builder notation, we have

$$X \setminus Y \coloneqq \{x \in X \mid x \notin Y\}.$$

Another way of making new sets out of old sets is by forming the *product.*

Definition 1.1.12. If X and Y are sets, we denote by $X \times Y$ the set of all pairs (x, y) with $x \in X$ and $y \in Y$, called the *Cartesian product of* X *and* Y. In set-builder notation, we would write

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

Definition 1.1.13. if X is a set, we will denote by $|X|$ or $\#X$ the *cardinality of* X, or the number of elements of X. If X has finitely many elements, we also call $|X|$ the *order* of X.

Let's recall some basic facts and operations on sets.

Proposition 1.1.14 (Properties of inclusion)**.** *For any sets A, B, C, we have*

- $A \subseteq A$,
- *if* $A \subseteq B$ *and* $B \subseteq A$, *then* $A = B$,
- *if* $A \subseteq B$ *and* $B \subseteq C$, *then* $A \subseteq C$.

Definition 1.1.15. If X is a set, the set of all subsets of X is also a set, called the *power set of* X and denoted by

$$\mathcal{P}(X) \coloneqq \{X' \mid X' \subseteq X\}.$$

Note that $X \in \mathcal{P}(X)$, and not $X \subseteq \mathcal{P}(X)$.

Proposition 1.1.16 (Properties of set operations)**.** *Let A, B, and C be sets. Then*

- $(A \cap B) \cap C = A \cap (B \cap C)$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- *if* $A, B \subseteq C$, *then* $C \setminus (A \cup B) = (C \setminus A) \cap (C \setminus B)$
- *if* $A, B \subseteq C$, *then* $C \setminus (A \cap B) = (C \setminus A) \cup (C \setminus B)$.

Proposition 1.1.17 (Properties of Cartesian products)**.** *Let A, B, C, and D be sets. Then*

- $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$
- $(A \cup B) \times C = (A \times C) \cup (B \times C)$
- $A \times \emptyset = \emptyset$
- *if* $A \times C = B \times C$ *and* $C \neq \emptyset$, *then* $A = B$.

1.2. **Functions (maps).**

Definition 1.2.1. Let X and Y be sets. A *map* or *function* f from X to Y is an association to every element of X a unique element $f(x) \in Y$. We write $f : X \to Y$ to mean that "f is a map from X to Y, and call X the *domain* of the map f, and call Y the *target* or *codomain* of f. We call $f(x)$ the value of f at x, or the *image* of x under f, and we usually write

$$f : X \to Y, \ x \mapsto f(x)$$

to denote that f maps x to $f(x)$. The *image of* f is the subset

$$\mathrm{im} f \coloneqq \{f(x) \in Y \mid x \in X\} \subseteq Y,$$

the subset of elements of Y that are values of the map f.

If $A \subseteq X$ is a subset, we write

$$f(A) \coloneqq \{f(a) \in Y \mid a \in A\} \subseteq \mathrm{im} f \subseteq Y,$$

to denote the *image of the subset A under f.*

If $y \in Y$, we write

$$f^{-1}(y) \coloneqq \{x \in X \mid f(x) = y\} \subseteq X,$$

called the *pre-image of y under f*. If $B \subseteq Y$ is a subset, we write

$$f^{-1}(B) \coloneqq \{x \in X \mid f(x) \in B\} \subseteq X,$$

called the *pre-image of B under f*.

Remark 1.2.2. Note that a map $f : X \to Y$ has three pieces, the domain X, the target Y, and a rule for associating $f(x) \in Y$ to $x \in X$. Changing any of these is a different map!

Definition 1.2.3. An alternative definition of a map $f : X \to Y$ is as a subset

$$\Gamma_f \subseteq X \times Y$$

satisfying the property that for any $x \in X$, there is a unique y such that $(x, y) \in \Gamma_f$. The way to go back and forth between the two definitions of a map is via the *graph of a map* $f : X \to Y$, defined by

$$\Gamma_f \coloneqq \{(x, y) \in X \times Y \mid y = f(x)\}.$$

Example 1.2.4. Let $\mathbb{R}$ denote the set of real numbers. Consider the map $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$. We have $\mathrm{im} f = \{x \in \mathbb{R} \mid x \geq 0\}$, and for $y > 0$, we have $f^{-1}(y) = \{\sqrt{y}, -\sqrt{y}\}$.

Consider the map $g : \mathbb{R} \to \mathbb{R}_{\geq 0} = \{x \in \mathbb{R} \mid x \geq 0\}, x \mapsto x^2$. In principle, this is practically the same map as f, but the target is different!

Consider the map $h : \mathbb{R}_{\geq 0} \to R, x \mapsto x^2$. We see that $\mathrm{im} g = \mathrm{im} h$, though the functions are different (they have a different domain)!

Definition 1.2.5. Let $f : X \to Y$ be a map. We say that

- f is *injective* if for all $a, b \in X$, $f(a) = f(b)$ implies $a = b$. We write $f : X \hookrightarrow Y$ if f is injective.
- f is *surjective* if $\mathrm{im} f = Y$. That is, if for every $y \in Y$, there is an $x \in X$ such that $f(x) = y$. We write $f : X \twoheadrightarrow Y$ is f is surjective.
- f is *bijective* if f is both injective and surjective.

Definition 1.2.6. Two maps $f : A \to B$ and $g : A \to B$ are *equal* if for all $a \in A$ we have $f(a) = g(a)$.

Definition 1.2.7 (Composition)**.** Let $f : A \to B$ and $g : B \to C$ be maps. We define a new map

$$g \circ f : A \to C, a \mapsto g(f(a)),$$

called g *composed with* f, and say that $g \circ f$ is the function "f followed by g".

. . . . End of Class 1 (1/12) . . . .

Exercise 1.2.8. Let $f : A \to B$, $g : B \to C$, $h : C \to D$ be maps. Show that

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

Definition 1.2.9. If $X' \subseteq X$, we define a map $\iota_{X'} : X' \hookrightarrow X, x \mapsto x$ called the *inclusion map of* X' *into* X. By definition, $\iota_{X'}$ is injective.

Exercise 1.2.10. Let $f : X \to Y$ be a map. Show that f can be written as a composition of an surjective map followed by an injective map. Hint: consider $\mathrm{im} f \subseteq Y$.

Definition 1.2.11. If $f : X \to Y$ is a map and $X' \subseteq X$ is a subset, we define

$$f|_{X'} : X' \to Y, x \mapsto f(\iota_{X'}(x)),$$

called the *restriction of* f *to* X'.

Example 1.2.12. The map $h : \mathbb{R}_{\geq 0} \to \mathbb{R}, x \mapsto x^2$ is the restriction of $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ to the subset $\mathbb{R}_{\geq 0} \subset \mathbb{R}$.

Exercise 1.2.13. Let $f : A \to B$ and $g : B \to C$ be maps. Then

(1) if both f and g are injective, so is $g \circ f$.
(2) if both f and g are surjective, so is $g \circ f$.
(3) if $g \circ f$ is surjective, then g is surjective.
(4) if $g \circ f$ is injective, then f is injective.

Definition 1.2.14. Let X be a set. Define the map $\mathrm{id}_X : X \to X, x \mapsto x$, called the *identity map of X*. Note that id_X is bijective.

Exercise 1.2.15. Let $f : X \to Y$ be a map. Show that

$$f \circ \mathrm{id}_X = f \text{ and } \mathrm{id}_Y \circ f = f.$$

We can sometimes invert a map. Suppose that $f : X \to Y$ is a bijection. Then since f is surjective, for every $y \in Y$, there is some $x \in X$ such that $y = f(x)$. Since f is injective, this x is unique.

Definition 1.2.16. If $f : X \to Y$ is bijective. We define a map

$$f^{-1} : Y \to X, y \mapsto x \text{ where } f(x) = y$$

called the *inverse of f*.

Exercise 1.2.17. Prove the following proposition.

Proposition 1.2.18. *Let $f : X \to Y$ and $g : Y \to X$ be maps such that*

$$f \circ g = \mathrm{id}_Y \text{ and } g \circ f = \mathrm{id}_X.$$

Use the first equality to show that f is surjective. Use the second equality to show that f is injective. Similarly, show that g is bijective. Thus there are inverse maps f^{-1} and g^{-1}. Show that $f = g^{-1}$ and $g = f^{-1}$.

We can summarize all of these into a theorem, which I encourage you to prove in detail using the exercises above as inspiration.

Theorem 1.2.19. *Let $f : X \to Y$ be a map. Then f is bijective if and only if there is a function $f^{-1} : Y \to X$ such that*

$$f \circ f^{-1} = \mathrm{id}_Y \text{ and } f^{-1} \circ f = \mathrm{id}_X.$$

Moreover, f^{-1} is unique if it exists.

Example 1.2.20. The function $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, x \mapsto x^2$ is bijective, and the inverse is given by $f^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, y \mapsto \sqrt{y}$.

We'll conclude with a proposition, which you are also encouraged to prove in detail.

Proposition 1.2.21. *Let $f : A \to B$ and $g : B \to C$ be bijections. Then $g \circ f : A \to C$ is a bijection with inverse $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

1.3. **Relations.**

Definition 1.3.1. A *relation* between two sets A and B is a subset $R \subseteq X \times Y$. If A is a set, then a relation on A is a subset $R \subseteq A \times A$.

For $(a, b) \in R \subseteq A \times B$, we write aRb or sometimes $a \sim b$ if the relation R is understood.

Just as with functions, we don't think of relations as subsets, and instead as a rule that tells us when a is related to b.

Example 1.3.2. The relation $\leq$ is a relation on $\mathbb{Z}$. Try writing down the subset that defines the relation.

Definition 1.3.3. Let A be a set and R a relation on A. We say that the relation R is

- *reflexive* if for all $a \in A$, $a \sim a$;
- *antireflexive* if for all $a \in A$, $a \nsim a$ ($(a, a) \notin R$);
- *symmetric* if for all $a, b \in A$, $a \sim b$ implies $b \sim a$;
- *antisymmetric* if for all $a, b \in A$, $a \sim b$ and $b \sim a$ implies $a = b$;
- *transitive* if for all $a, b, c \in A$, $a \sim b$ and $b \sim c$ implies $a \sim c$.

Example 1.3.4.

- The relation $\leq$ on $\mathbb{Z}$ is reflexive, antisymmetric, and transitive.
- The relation $=$ is a relation on any set. It is reflexive, symmetric, antisymmetric, and transitive.
- The relation $\neq$ is also a relation on any set. It is symmetric and antireflexive.

Most commonly, the type of relation we will use is one that shares many of the properties of $=$.

Definition 1.3.5. Let A be a set, a reflexive, symmetric, and transitive relation $\sim$ on A is called an *equivalence relation.* For $a \in A$, the *equivalent class of* a is the set

$$[a]_\sim := \{x \in A \mid x \sim a\},$$

or simply $[a]$ when the relation $\sim$ is understood.

For $a, b \in A$ such that $a \sim b$, we clearly have $[a] = [b]$.

Definition 1.3.6. The set of equivalence classes under $\sim$ is denoted by

$$A/\sim := \{[a] \mid a \in A\}.$$

Let's see some examples.

Example 1.3.7 (Congruence modulo n)**.** Let $n \in \mathbb{Z}$. We define an equivalence relation on $\mathbb{Z}$, called congruence modulo n by

$$a \equiv b \mod n \text{ if } n \text{ divides } b - a.$$

To check that this is an equivalence relation, note that

- $a - a = 0$, and n divides 0, thus $a \equiv a \mod n$;
- if $a \equiv b \mod n$, then n divides $b - a$, thus divides $a - b$ as well, thus $b \equiv a \mod n$;
- if $a \equiv b \mod n$ and $b \equiv c \mod n$, then n divides $b - a$ and $c - b$, whence n divides $c - b + b - a = c - a$, thus $a \equiv c \mod n$.

The set of equivalence classes is usually denoted by $\mathbb{Z}/n\mathbb{Z}$, and can be identified with the set of remainders after division by n, i.e. $\{0, 1, \ldots, n-1\}$.

Example 1.3.8. Let $f : A \to B$ be a map. We can define an equivalence relation on A by $x \sim y$ if and only if $f(x) = f(y)$. The equivalence class of $x \in A$ is $f^{-1}(\{f(x)\})$.

In fact, all equivalence relations arise in this way.

Proposition 1.3.9. *Let A be a set and $\sim$ an equivalence relation on A. Then there is a surjective map $\pi_A \to A/\sim, a \mapsto [a]$. the equivalence relation $\sim$ is exactly the one obtained from the map π.*

The nice thing about an equivalence relation is that it breaks up A into pieces.

Definition 1.3.10. A *partition* of a set A is a collection of subsets $\{S_i\}_{i\in I}$ of A such that every $a \in A$ is contained in exactly one set S_i. That is, a partition is a set of subsets $\{S_i\}_{i\in I}$ such that A is the disjoint union of the subsets S_i,

$$A = \bigsqcup_{i\in I} S_i.$$

In fact, partitions and equivalence relations are equivalent.

Theorem 1.3.11. *Let A be a set. If $\sim$ is an equivalence relation on A, then the set of equivalence classes form a partition of A. Conversely, if $\{S_i\}_{i\in I}$ is a partition of A, then the relation $x \sim y$ if and only if $x, y \in S_i$ for some i is an equivalence relation.*

2. Integers, induction, and well-ordering

We'll review the concept of mathematical induction, with which we assume the reader is familiar, having seen it in a proof-based math class.

Recall 2.0.1 (Mathematical induction)**.** Let $P(n)$ be a statement about n. Suppose that

- $P(1)$ is true, and
- for all $n \in \mathbb{Z}_{>0}$, if $P(n)$ is true, then $P(n+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{Z}_{>0}$.

.... End of Class 2 (1/14) ....

Remark 2.0.2. Note that by shifting n, we can start at any $m \in \mathbb{Z}$. Thus, if

- $P(m)$ is true, and
- for all $n \geq m$, if $P(n)$ is true, then $P(n+1)$ is true,

then $P(n)$ is true for all $n \geq m$.

Proofs by induction generally follow the same format (we'll start at m here)

- precisely state the property $P(n)$ that you wish to prove.
- Base case: prove that $P(m)$ holds.
- Inductive step: Suppose that $n \geq m$ and that $P(n)$ holds. Show that $P(n+1)$ holds. Remember to state clearly that you are assuming that $P(n)$ holds, and clearly explain where it is used in the proof.
- Conclusion: by induction, $P(n)$ holds for all $n \geq m$.

The main proofs where induction is useful are if you are asked to prove a formula or some property, or if there is a way of breaking up the proof into smaller pieces each of which can further be broken down.

Example 2.0.3. For all $n \in \mathbb{Z}_{>0}$, we have

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Proof. Let $P(n)$ be the property that $1 + \cdots + n = \frac{n(n+1)}{2}$. we will prove this by induction.

For the base case, we compute $\frac{1(1+1)}{2} = \frac{2}{2} = 1$. Since $1 = \frac{1(1+1)}{2}$, we see that $P(1)$ holds.

For the inductive step, we suppose that $P(n)$ holds. We compute

$$\begin{aligned} 1+2+\cdots+n+(n+1) &= \frac{n(n+1)}{2}+(n+1) \text{ by the inductive hypothesis} \\ &= \frac{n(n+1)+2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}, \end{aligned}$$

which is $P(n+1)$, thus $P(n+1)$ is true.

By induction, $P(n)$ holds for all $n \in \mathbb{Z}_{>0}$. □

We also have the principle of *strong* mathematical induction.

Recall 2.0.4 (Strong mathematical induction)**.** Let $P(n)$ be a statement about n. Suppose that

- $P(1)$ is true, and
- for all $n \in \mathbb{Z}_{>0}$, if $P(k)$ is true for all $k \leq n$, then $P(n+1)$ is true.

Then $P(n)$ is true for all $n \in \mathbb{Z}_{>0}$.

As you may have seen before, mathematical induction and strong mathematical induction are equivalent, that is, we can deduce one from the other in both directions. Interestingly, they are both equivalent to the *well-ordering principle.*

Recall 2.0.5 (Well-Ordering principle)**.** Let $S \subseteq \mathbb{Z}$ be a non-empty subset bounded from below, i.e. there is an $a \in \mathbb{Z}$ such that $a \leq s$ for all $s \in S$. Then S has a *least element*, some $s_0 \in S$ such that $s_0 \leq s$ for all $s \in S$.

Remark 2.0.6. The well-ordering principle is a property of the relation $\leq$ on $\mathbb{Z}$. It is not true for all relations. For example, consider the subset $(0,1) \subset \mathbb{R}$, and the relation $\leq$ on $(0,1)$. Even though $(0,1)$ is bounded from below, there is no smallest element of $(0,1)$.

Theorem 2.0.7. *The well-ordering principle and induction are equivalent.*

Proof. We first show that we can derive well-ordering from induction. Let $S \subset \mathbb{Z}$ be a non-empty subset of $\mathbb{Z}$ that is bounded from below by a. Assume, for contradiction that S has no smallest element. Thus, for all $n \in \mathbb{Z}$, either $n \notin S$ or there is some $s \in S$ such that $s < n$. We will use induction to show that for all $n \in \mathbb{Z}$, $n \notin S$, which is a contradiction as $S \subset \mathbb{Z}$ is non-empty. If $n < a$, then $n \notin S$, as a is a lower bound for S.

Base case: if $n = a$, then $n \notin S$, as a is a lower bound for S.

Inductive step: Let $n \in \mathbb{Z}$, suppose that $n \geq a$ and suppose that k is not in S for any $k \leq n$. By assumption, either $n+1 \notin S$, or there is some $s \in S$ such that $s < n+1$. In the later case $s \leq n$ is in S, which we assumes does not occur. Thus $n \notin S$.

Thus, by induction, $S = \emptyset$, which is a contradiction. So S must have a least element.

We now show that the well-ordering principle implies induction. So we assume that the well–ordering principle is true. Let $P(n)$ be a property of integers, and assume that

- $P(1)$ is true, and
- that for all $n \in \mathbb{Z}_{>0}$, if $P(n)$ is true then $P(n+1)$ is true.

We want to show that $P(n)$ is true for all $n \in \mathbb{Z}_{>0}$. Towards a contradiction, assume that $P(n)$ is false for some n. Consider the set

$$S = \{n \in \mathbb{Z}_{>0} \mid P(n) \text{ is false}\}.$$

By assumption, S is non-empty. Thus the well-ordering principle shows that S has a smallest element, say s_0. Thus $P(s_0)$ is false. By assumption, $1 \notin S$, as $P(1)$ we have assumed that is true, thus $s_0 \neq 1$. Since s_0 is the least element of S, we have $s_0 - 1 \notin S$. Therefore $P(s_0 - 1)$ is true, and thus $P(s_0)$ is true. This is a contradiction, and thus $P(n)$ must be true for all $n \in \mathbb{Z}_{>0}$, as desired. □

Exercise 2.0.8 (Fundamental Theorem of Arithmetic)**.** We call a number $p \in \mathbb{Z}_{>0}$ *prime* if $p \geq 2$ and if $p = mn$ for some $m, n \in \mathbb{Z}_{>0}$, then $n = 1$ or $m = 1$.

- Show that p is prime if and only if whenever p divides mn, then p divides n or p divides m.
- Using induction, show that every positive integer $n \geq 2$ can be expressed as a product of prime numbers (not necessarily distinct) $n = p_1 \cdots p_r$.
- Using the first statement, show that the primes are unique up to reordering.

Remark 2.0.9. The first few prime numbers are $2, 3, 5, 7, 11, \ldots$. The fact that 2 is prime is no accident. Indeed, suppose that $2 = mn$ for some $m, n \in \mathbb{Z}_{>0}$. Then if both $n, m \neq 1$, both must be at least 2, and thus $mn \geq 4$, which is a contradiction. Hence one of m or n must be 1, and so 2 is prime.

Theorem 2.0.10 (Euclid's Theorem)**.** *There are infinitely many primes.*

Proof. Suppose that there are finitely many primes, $p_1, \ldots, p_r$. Then $n = 1 + p_1 \cdots p_r$ is not divisible by any of the p_i. However, n can be written as the product of primes, thus there must be a prime $p \mid n$ not on our list, which is a contradiction. □

3. Integers, number theory

This section is also assumed to be mostly review from a proof-based math class, though perhaps presented in slightly different language that we'll develop later. If you find that any of the topics here are not comfortable, I highly recommend the book *The Art of Proof* by Matthias Beck and Ross Geoghegan, available as a pdf on the first author's website here `https://matthbeck.github.io/aop.html`.

3.1. Euclidean Algorithm, Greatest Common Divisors.

Recall 3.1.1.

Theorem 3.1.2 (Division algorithm in $\mathbb{Z}$)**.** *Given $a, b \in \mathbb{Z}$, there are unique $q, r \in \mathbb{Z}$ such that $b = qa + r$ with $|r| < |a|$ or $r = 0$.*

The proof of this is fairly straightforward, and you are encouraged to find a direct proof, or find a proof using the well-ordering principle.

Remark 3.1.3. You'll see later that we can do much the same with polynomials, which will be very useful in 4581.

Notation 3.1.4. For $a, b \in \mathbb{Z}$, we will write $a \mid b$ if a divides b. That is, if there is some $n \in \mathbb{Z}$ such that $na = b$.

Definition 3.1.5. Let $a, b \in \mathbb{Z}$. The *greatest common divisor of a and b* is some $d \in \mathbb{Z} \setminus \{0\}$ such that

- $d \geq 0$,
- $d \mid a$ and $d \mid b$, and
- if $e \in \mathbb{Z} \setminus \{0\}$ also divides both a and b, then $e \mid d$.

We write $\gcd(a, b)$ for the greatest common divisor of a and b.

Exercise 3.1.6. Show that if $\gcd(a, b)$ exists, then it is unique.

Let us show that greatest common divisors exist.

Definition 3.1.7. Let $I \subseteq \mathbb{Z}$. We say that I is an *ideal* if

- $0 \in I$,
- if $m, n \in I$, then $m + n \in I$, and
- if $n \in I$, and $a \in \mathbb{Z}$, then $an \in I$.

Example 3.1.8. Let $m_1 \dots, m_r$ be integers. Let

$$I = \{x_1m_1 + \cdots x_rm_r \ \mid \ x_i \in \mathbb{Z}\}.$$

Then I is an ideal. Indeed, $0 = 0m_1 + \cdots 0m_r \in I$. If $x_1m_1 + \cdots x_rm_r, y_1m_1 + \cdots y_rm_r \in I$, then

$$(x_1m_1 + \cdots x_rm_r) + (y_1m_1 + \cdots y_rm_r) = (x_1 + y_1)m_1 + \cdots + (x_r + y_r)m_r \in I.$$

And if $n \in \mathbb{Z}$, then

$$n(x_1m_1 + \cdots x_rm_r) = (nx_1)m_1 + \cdots + (nx_r)m_r \in I.$$

Hence I is indeed an ideal.

Definition 3.1.9. Let $I \subset \mathbb{Z}$ be an ideal. If there exist $m_1 \dots, m_r \in I$ such that every element $a \in I$ can be written in the form $x_1m_1 + \cdots x_rm_r$ for some integers $x_1, \dots, x_r \in \mathbb{Z}$, we say that I is *generated* by $m_1, \dots, m_r$, and that $m_1, \dots, m_r$ are the *generators* of I. We write $I = (m_1, \dots, m_r)$ if I is generated by $m_1, \dots, m_r$.

Example 3.1.10. Note that $\{0\} \subset \mathbb{Z}$ it an ideal, generated by 0. Also, $\mathbb{Z}$ is an ideal, generated by 1.

Theorem 3.1.11. *Let I be an ideal of $\mathbb{Z}$. Then there exists an integer d which generates I. If $I \neq \{0\}$, then we may take d to be the smallest positive integer in I.*

Proof. If $I = \{0\}$, then I is generated by 0. Suppose that $I \neq \{0\}$. If $n \in I$, then $-n = (-1)n \in I$. So I must contain some positive integer. By the well-ordering principle, there is a smallest positive integer in I, call it d. We claim that d generates I. To see this, let $n \in I$ and using the division algorithm (Theorem 3.1.2), write

$$n = dq + r,$$

with $0 \leq r < d$. Then $r = n - dq \in I$, and since $r < d$, we must have $r = 0$ as d is the smallest positive integer in I. Thus $n = dq$, and thus d generates I, as claimed. □

Example 3.1.12. For $a, b \in \mathbb{Z}$, the ideal generated by a and b is

$$(a, b) := \{n \in \mathbb{Z} \mid n = ax + by \text{ for some } x, y \in \mathbb{Z}\}.$$

Theorem 3.1.13. *Let $a, b \in \mathbb{Z}$. Let d be a non-negative generator of the ideal $(a, b) \subseteq \mathbb{Z}$. Then $d = \gcd(a, b)$.*

Proof. By assumption, $d \geq 0$. Since $a, b \in (a, b) = (d)$, we have $a = q_1d$ and $b = q_2d$ for some $q_1, q_2 \in \mathbb{Z}$. Thus $d \mid a$ and $d \mid b$. Let e be a non-zero integer dividing both a and b, say

$$a = h_1e \text{ and } b = h_2e,$$

for some integers $h_1, h_2 \in \mathbb{Z}$. Since d generates (a, b), there are integers $x, y \in \mathbb{Z}$ such that $d = ax + by$, thus

$$d = ax + by = (h_1e)x + (h_2e)y = (xh_1 + yh_2)e,$$

and so e divides d. Thus $d = \gcd(a, b)$, as claimed. □

Remark 3.1.14. Exactly the same proof works if we want to find the greatest common divisor of more than two integers. For example, if $m_1, \dots, m_r$ are non-zero integers, and $(d) = (m_1, \dots, m_r)$, with $d > 0$, then $d = \gcd(m_1, \dots, m_r)$.

Definition 3.1.15. We say that integers $m_1, \dots, m_r$ are *relatively prime* if $\gcd(m_1, \dots, m_r) = 1$. In that case, since $1 \in (m_1, \dots, m_r)$, there are integers $x_1 \dots, x_r \in \mathbb{Z}$ such that

$$1 = x_1 m_1 + \cdots x_r m_r.$$

Exercise 3.1.16. Show that $\gcd(a, 0) = a$, and that $\gcd(a, b) = \gcd(b, a)$

Example 3.1.17 (Computing $\gcd(a, b)$)**.** We know that $\gcd(a, b)$ is the smallest non-negative element of $(a, b) = \{ax + by \mid a, x \in \mathbb{Z}\}$. If either a or b are zero, say $b = 0$, then $\gcd(a, 0) = a$. Thus suppose that $a, b > 0$ and $a > b$. Then we can find divide and write

$$a = bq + r,\ 0 \leq r < b.$$

But since $a, b \in (a, b)$, we have $r = a - bq \in (a, b)$. Moreover, we see that

$$(a, b) = (b, r),$$

thus

$$\gcd(a, b) = \gcd(b, r),$$

and we can continue dividing and replacing the larger entry with the remainder until the remainder is zero! And then we'll have found the generator of the ideal! Let's do an explicit example.

To compute $\gcd(67, 941)$, we compute $941 = 14 \times 67 + 3$, so

$$(67, 941) = (3, 67).$$

Continuing, as $67 = 22 \times 3 + 1$, so

$$(67, 941) = (3, 67) = (1, 3),$$

and since $3 = 3 \times 1$, we have

$$(67, 941) = (3, 67) = (1, 3) = (1, 0) = (1).$$

Thus $\gcd(67, 941) = 1$.

Theorem 3.1.18 (Euclidean Algorithm for gcd)**.** *Let $a, b \in \mathbb{Z}$, we can find $\gcd(a, b)$ using the following algorithm.*

```
Step 0: Change $a, b$ into $-a$ or $-b$ so that both $a$ and $b$ are non-negative.
Step 1: Set $r_0 = a$, $r_1 = b$.
Step 2: Divide $r_{n-1}$ by $r_n$ and write $r_{n-1} = r_n q_n + r_{n+1}$.
            – if $r_{n+1} \neq 0$, repeat;
            – else, output $r_n$.
```

.... End of Class 3 (1/16) ....

Proof. This is exactly the process we've described above. Each step preserves the ideal $(a, b) = (r_0, r_1) = (r_1, r_2) = \cdots = (r_n, r_{n+1})$, and if it terminates, then $r_{n+1} = 0$, and $(a, b) = (r_n)$. Thus it remains to show that this process indeed terminates.

But since $r_1 > r_2 > \cdots > r_n \cdots \geq 0$, for all n, by the division algorithm (Theorem 3.1.2), this process cannot continue forever, as the set of remainders forms a non-empty subset of $\mathbb{Z}$ that is bounded from below, and thus has a least element. □

Theorem 3.1.19 (Chinese Remainder Theorem)**.** *Let n, n' be relatively prime positive integers. Let $a, b \in \mathbb{Z}$. Then there is some $x \in \mathbb{Z}$ such that*

$$x \equiv a \mod n,$$
$$x \equiv b \mod n'.$$

More generally, if $n_1, \ldots, n_r$ are pairwise relatively prime, that is, $\gcd(n_i, n_j) = 1$ for $i \neq j$, and $a_1, \ldots, a_r \in \mathbb{Z}$, then there is a some $x \in \mathbb{Z}$ such that

$$x \equiv a_i \mod n_i \text{ for all } 1 \leq i \leq r.$$

Additionally, all solutions are equivalent modulo $n_1 \cdots n_r$.

Proof. We give a proof of the general case by induction on r.

Base case: $(r = 2)$ Since $\gcd(n_1, n_2) = 1$, there are $x_1, x_2 \in \mathbb{Z}$ such that $1 = x_1 n_1 + x_2 n_2$. Then we have $x_1 n_1 \equiv 1 \mod n_2$, and $x_2 n_2 \equiv 1 \mod n_1$. Thus

$$x = a_2 x_1 n_1 + a_1 x_2 n_2$$

is our desired solution.

Inductive step: Suppose the claim holds for r, we show it holds for $r + 1$. Since $\gcd(n_i, n_j) = 1$ for $i \neq j$, we have $\gcd(n_1 n_2 \cdots, n_r, n_{r+1}) = 1$. And by the inductive hypothesis, there is some $x_0 \in \mathbb{Z}$ such that $x_0 \equiv a_i \mod n_i$ for $1 \leq i \leq r$. Now consider the two congruences

$$x \equiv x_0 \mod n_1 \cdots n_r, \; x \equiv a_{r+1} \mod n_{r+1}.$$

By the base case, we can find $x \in \mathbb{Z}$ satisfying these two equivalences. By the division algorithm (Theorem 3.1.2), since

$$x \equiv x_0 \mod n_1 \cdots n_r \text{ and } x_0 \equiv a_i \mod n_i \text{ for } 1 \leq i \leq r$$

we have $x \equiv a_i \mod n_i$ for $1 \leq i \leq r + 1$, as desired.

Thus, by induction, the claim holds for all r.

We now show that all solutions are congruent modulo $n_1 \cdots n_r$. Suppose a and b both satisfy $x \equiv a_i \mod n_i$ for relatively prime n_i, then we have $a \equiv b \mod n_i$ for each i, so $n_i \mid (a - b)$ for each i. Thus $n_1 \cdots n_r \mid (a - b)$, hence $a \equiv b \mod n_1 \cdots n_r$. Thus all of the solutions are equivalent modulo $n_1 \cdots n_r$. □

Remark 3.1.20. The notion of ideals will return when we discuss *rings.*

We'll end with a useful fact.

Lemma 3.1.21. *Let $n_1, n_2 \in \mathbb{Z}$. Then the least common multiple of n_1 and n_2 (which exists by well-ordering) is $\frac{n_1 n_2}{\gcd(n_1, n_1)}$.*

Proof. Let $d = \gcd(n_1, n_2)$. Note that $\frac{n_1}{d}, \frac{n_2}{d} \in \mathbb{Z}$, and $\gcd\left(\frac{n_1}{d}, \frac{n_2}{d}\right) = 1$, else d would be larger. Thus

$$\frac{n_1 n_2}{d} = \frac{n_2}{d} n_1 = \frac{n_1}{d} n_2$$

is a common multiple of both n_1 and n_2.

Now let n be a common multiple of both n_1 and n_2. We aim to show that $\frac{n_1 n_2}{d}$ divides n. We may write $n = k_1 n_1 = k_2 n_2$ for some integers $k_1, k_2 \in \mathbb{Z}$. Hence, as $d \mid n_1$ and $d \mid n_2$, we have $d \mid n$, and so

$$\frac{n}{d} = k_1 \frac{n_1}{d} = k_2 \frac{n_2}{d} \in \mathbb{Z}.$$

Thus $\frac{n_1}{d}$ divides $k_2 \frac{n_2}{d}$. Since $\gcd\left(\frac{n_1}{d}, \frac{n_2}{d}\right) = 1$, $\frac{n_1}{d}$ divides k_2 whereby

$$n = k_2 n_2 = q \frac{n_1}{d} n_2 = q \frac{n_1 n_2}{d}.$$

Therefore $\frac{n_1 n_2}{d} = \frac{n_1 n_2}{\gcd(n_1, n_1)}$ divides n, as desired. □

. . . . End of Class 4 (1/21) . . . .

4. Groups

We now turn to one of the main topics of the course, the notion of a *group*. Historically, the concept of a group arose in the study of symmetries of roots of polynomial equations, and slowly developed into an abstract notion of symmetry. We are privileged that the notion of a group has become standard, early definitions of a group were as functions that satisfy certain properties. Let us take some inspiration from the historical definition, and come up with the definition of group ourselves!

For example, consider a square. There are many symmetries that move the square around, but return it to the same place. For example D_4 is the symmetry group of the square. There is the "do nothing symmetry" which we'll call e, rotation by 90° which is denoted by r, and repeated rotation which we can denote by powers r^2, r^3 and $r^4 = e$, as well as reflections s_h, s_v through the horizontal and vertical axes, and reflections across the diagonals s_d and $s_{d'}$, as shown in Figure 4. Labeling the vertices you can find relations among these symmetries. For example, $(s_v)^2 = e$.

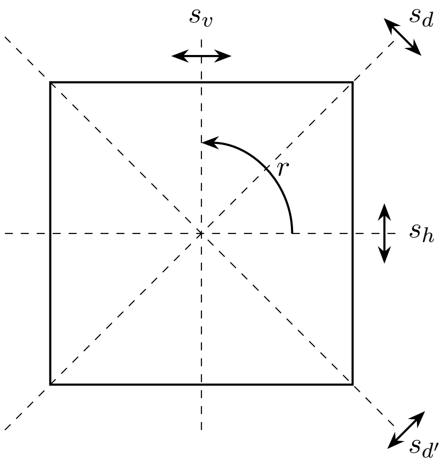


Figure 1. How elements of D_4 act on the square

We want to capture the idea of an object having some symmetry, that is, some object X and some maps $\{f_i : X \to X\}$, such that the maps f_i preserve some properties of X. Well, clearly id_X should be such a map, the "do nothing map". The maps f_i should also be invertible, after all they should preserve X, and their inverses should also preserve the properties of X. And finally, we should be able to compose the symmetries and still get symmetries that preserve X and its properties. This is all a group is! But we'll abstract away all of the unnecessary details, like X and the mysterious properties we are trying to preserve.

Since we'll no longer be dealing with composing maps, we'll need a more general notion of "combining to things to get another thing".

Definition 4.0.1. Let A be a set. A *binary operation* on A is a map

$$* : A \times A \to A, \ (a, b) \mapsto a * b.$$

Example 4.0.2.

(1) The usual addition $(+)$ and multiplication $(\times)$ operations on $\mathbb{Z}$ are binary operations. They are special, since $a + b = b + a$ and $a \times b = b \times a$.
(2) Let $M_n(\mathbb{R}) \coloneqq \{$ square $n \times n$ matrices with entries in $\mathbb{R}\}$. Matrix multiplication is a binary operation on $M_n(\mathbb{R})$, this time $AB \neq BA$ in general.
(3) Let A be a set, then $\cup$ and $\cap$ are binary operations on $\mathcal{P}(A)$.

(4) Let A be a set, denote by
$$A^A := \{\text{maps } A \to A\}.$$
Then function composition is a binary operation on A^A.

Many of these binary operations have additional properties.

Definition 4.0.3. We say that a binary operation $*$ on A is

- *associative* if for all $a, b, c \in A$, we have
$$a * (b * c) = (a * b) * c.$$
- *commutative* if for all $a, b \in A$, we have
$$a * b = b * a.$$

We say that $*$ has an *identity element* if there is some $e \in A$ such that for all $a \in A$,
$$e * a = a * e = a,$$
and the element $e \in A$ is called the *identity* for $*$.

If e is an identity element for $*$, we say that $b \in A$ is *an inverse for* $a \in A$ if
$$a * b = e = b * e.$$

In fact, identity elements are unique, if they exist.

Proposition 4.0.4. *Let $*$ be a binary operation on A. If $*$ has an identity element, then it is unique.*

Proof. Suppose that e_1 and e_2 are both identity elements for $*$. Then
$$e_1 = e_1 * e_2 = e_2,$$
the left equality coming from the fact that e_2 is an identity, and the right coming from the fact that e_1 is an identity. □

Recall 4.0.5. Thus, if $*$ has an identity, we will refer to it as *the* identity.

Example 4.0.6.

(1) Consider the binary operations $+$ and $\times$ on $\mathbb{Z}$.
- The binary operation $+$ on $\mathbb{Z}$ is associative and commutative, and has an identity element 0. It also has inverses, the inverse for $+$ of a is $-a$.
- The binary operation $\times$ on $\mathbb{Z}$ is associative and commutative, and also has an identity 1. However, the only elements that have inverses for $\times$ are ± 1.

(2) Matrix multiplication of square $n \times n$ matrices is associative but not commutative, and has an identity, namely the identity matrix
$$\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$$
and matrices with non-zero determinant have inverses.

(3) The binary operations $\cap$ and $\cup$ are both associative and commutative on $\mathcal{P}(A)$. The identity for $\cup$ is $\emptyset$, and the identity for $\cap$ is A. Try to work out which elements of $\mathcal{P}(A)$ have inverses.

(4) For a set A, function composition on A^A is associative, but not commutative in general. The identity is id_A, and the elements with inverses are exactly the bijections $A \to A$.

Associative binary operations also have some other uniqueness properties.

Proposition 4.0.7. *Let $*$ be an associative binary operation on A. Then*

(i) *if $a \in A$ has an inverse, then the inverse is unique. We will denote it by a^{-1}.*
(ii) *if $a \in A$ has inverse a^{-1}, then the inverse of a^{-1} is a, i.e. $(a^{-1})^{-1} = a$.*
(iii) *if both $a, b \in A$ have inverses, then the inverse of $a * b$ is $b^{-1} * a^{-1}$, i.e.*

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

Proof. Let $a, b \in A$, and assume they both have inverses. Let $e \in A$ be the identity element for $*$.

To show (i), suppose that a has inverses x_1 and x_2, then

$$\begin{aligned} x_1 &= x_1 * e \\ &= x_1 * (a * x_2) \\ &= (x_1 * a) * x_2 \text{ by associativity} \\ &= e * x_2 \\ &= x_2, \end{aligned}$$

and so $x_1 = x_2$.

To show (ii), we since $a * a^{-1} = a^{-1} * a = e$, the inverse of a^{-1} is a.

To show (iii), we compute

$$\begin{aligned} (a * b) * (b^{-1} * a^{-1}) &= a * (b * b^{-1}) * a^{-1} \\ &= a * e * a^{-1} \\ &= a * a^{-1} \\ &= e, \end{aligned}$$

and a similar computation shows that $(b^{-1} * a^{-1}) * (a * b) = e$. Since inverses are unique, we have $(a * b)^{-1} = b^{-1} * a^{-1}$, as desired. □

Note that a set with an associative binary operation that has an identity and inverses is exactly what we wanted a group to be! We're now ready to give the abstract definition of a group, which will keep us occupied for a majority of the course.

4.1. Definition of a group and examples.

Definition 4.1.1. A *group* G is a set together with a binary operation $* : G \times G \to G$ (called the group operation) satisfying the following properties.

(**GR 1**) $*$ is associative, i.e. for all $a, b, c \in G$

$$(a * b) * c = a * (b * c).$$

(**GR 2**) $*$ has an identity element $e \in G$ such that for all $x \in G$

$$e * x = x * e = x.$$

(**GR 3**) Every element of G has an inverse for $*$, i.e., for all $x \in G$, there is some $x^{-1} \in G$ such that

$$x * x^{-1} = x^{-1} * x = e.$$

Note that by Proposition 4.0.4 and Proposition 4.0.7, the identity element e is unique, and inverses are also unique.

Definition 4.1.2. A group G is called *abelian* or *commutative* or *additive* if $*$ is commutative.

Example 4.1.3.

(1) The set $\mathbb{Z}$ together with the binary operation $+$ is a group, in fact an abelian group. However, $\mathbb{Z}$ with the binary operation $\times$ is not a group.
(2) $M_n(\mathbb{R})$ does not form a group under matrix multiplication, as not every matrix has an inverse. However, the set $\mathrm{GL}_n(\mathbb{R})$ of invertible $n \times n$ matrices with real entries is a group with group operation matrix multiplication.

(3) The symmetries of the square, D_4 shown above in Figure 4 form a group under composition. D_4 is not abelian.

(4) The set of matrices

$$\left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right\}$$

is a group under matrix multiplication.

(5) The set of matrices

$$\left\{\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}, \begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}\right\}$$

is a group under matrix multiplication.

.... End of Class 5 (1/23) ....

Notation 4.1.4. When G is an abelian group, we will usually (but not always) denote the group operation by $+$, and so $a+b=b+a$ for all $a,b\in G$. We will also generally denote the identity of an abelian group by 0, and the inverse of $a\in G$ by $-a$, so

$$a+(-a)=(-a)+a=0.$$

And when we use $+$ to denote the group operation, we usually call G an *additive group*. We will only use the additive notation when G is an abelian group.

Remark 4.1.5. When the binary operation $*$ of a group G is understood, we will generally write $x*y=xy$, unless we want to emphasize that G is abelian and use additive notation. When we write the group operation as $x*y=xy$, we sometimes call G a *multiplicative group*.

Let's see some more examples of groups.

Example 4.1.6.

(1) $\mathbb{Z},\mathbb{Q},\mathbb{R}$, and $\mathbb{C}$ are all abelian groups under $+$.

(2) $\mathbb{Q}^\times := \mathbb{Q}\setminus\{0\}$, $\mathbb{R}^\times := \mathbb{R}\setminus\{0\}$, and $\mathbb{C}^\times := \mathbb{C}\setminus\{0\}$ are groups under $\times$.

(3) Any vector space is a group under $+$. This includes $M_n(\mathbb{R})$ with $+$.

There are also some important, though maybe less familiar examples.

Definition 4.1.7. Let A be a set, the set of bijections $A\to A$ is denoted by

$$\operatorname{Perm}(A) := \{\text{bijections } A\to A\}.$$

Proposition 4.1.8. *Let A be a set, then* $\operatorname{Perm}(A)$ *is a group under function composition, with identity* id_A.

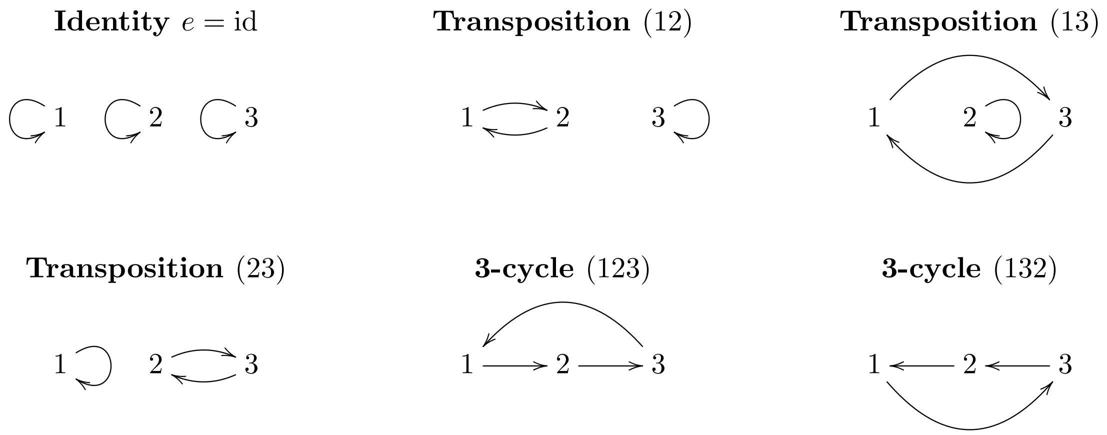
Proof. The composition of two bijections is a bijection, hence function composition is a binary operation on $\operatorname{Perm}(A)$. That function composition is associative is clear, as is the fact that id_A is an identity for composition. Finally, since every $f\in\operatorname{Perm}(A)$ is a bijection, there exists $f^{-1}: A\to A$ such that $f\circ f^{-1} = f^{-1}\circ f = \mathrm{id}_A$, and by Proposition 1.2.18, f^{-1} is also a bijection. Thus every $f\in\operatorname{Perm}(A)$ has an inverse $f^{-1}\in\operatorname{Perm}(A)$. □

Definition 4.1.9. When $A=\{1,2,\dots,n\}$, $\operatorname{Perm}(A)$ is usually called S_n, the *permutation group*.

Example 4.1.10. Let's explore some small permutation groups.

- When $n=1$, $A=\{1\}$, there is only one bijection of $\{1\}$ with itself, namely id_A. So $S_1=\{e=\mathrm{id}_A\}$.

- When $n = 2$, $A = \{1, 2\}$ and there are now two bijections $A \to A$, id_A and
$$\sigma : A \to A \ , \ \begin{matrix} \sigma(1) = 2 \\ \sigma(2) = 1. \end{matrix}$$
Thus $S_2 = \{e = \mathrm{id}, \sigma\}$ and we can check that $\sigma \circ \sigma = \mathrm{id}$.
- When $n = 3$, $A = \{1, 2, 3\}$ and there are many bijections. We'll write them using pictures.



Identity $e = \mathrm{id}$

Transposition (12)

Transposition (13)

Transposition (23)

3-cycle (123)

3-cycle (132)

We can check, for example, that $(12) \circ (13) = (132)$, since
$$\begin{aligned} (12) \circ (13)1 &= (12)3 = 3, \\ (12) \circ (13)2 &= (12)2 = 1, \\ (12) \circ (13)3 &= (12)1 = 2, \end{aligned}$$
which is exactly what (132) does.
- The group $S_4 = \{\text{bijections } \{1, 2, 3, 4\} \to \{1, 2, 3, 4\}\}$ has 24 elements, which is just a few too many to draw a diagram for each. If you're interested though, it is a good example.

Exercise 4.1.11. Prove the following proposition.

Proposition 4.1.12. *Show that S_3 is not abelian by finding two elements $\sigma, \tau \in S_3$ such that $\sigma \circ \tau \neq \tau \circ \sigma$. More generally, show that S_n is abelian if and only if $n = 1, 2$.*

Remark 4.1.13. Note that if we think of the standard unit vectors
$$\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} , \ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} , \ e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$
as a set of three elements, then the group of 3×3 matrices in Example 4.1.3(5) can be thought of as the group S_3, each matrix acting as a bijection by standard matrix acting on a vector. For example
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
switches $e_1 \leftrightarrow e_2$ and preserves e_3. Try to match up the matrices with the elements of S_3 based on how they act on the standard unit vectors.

We'll explore permutation groups in detail later.

Example 4.1.14.

(1) Let $S^1 := \{z \in \mathbb{C} \mid |z| = 1\} \subset \mathbb{C}$, called the *circle group*, which is a group under ×. Using Euler's Formula ($e^{i\theta} = \cos(\theta) + i\sin(\theta)$), we see that $S^1 = \{e^{i\theta} \in \mathbb{C} \mid \theta \in \mathbb{R}\}$.

(2) The set $\{1, -1\} \subset \mathbb{Z}$ forms a group under ×.

(3) The set $\{1, -1, i, -i\} \subset \mathbb{C}$ forms a group under ×

(4) The set $Q = \{1, -1, i, j, k, -i, -j, -k\}$ forms a group under multiplication with identity 1 and the rules

$$i^2 = j^2 = k^2 = ijk = -1, ij = k, ji = -ij, ik = -ki, kj = -jk, (-1)(i) = (i)(-1), \ldots, (-1)(k) = (k)(-1).$$

This is called the Quaternion group, in honor of Hamilton's quaternions, which we will see again later when we study rings.

.... End of Class 6 (1/26) ....

One nice thing about groups is that, since every element has an inverse, we can sometimes cancel and simplify expressions.

Proposition 4.1.15 (Cancellation laws)**.** *Let G be a group. For all $a, b, c \in G$,*

- *if $ab = ac$, then $b = c$;*
- *if $ba = ca$, then $b = c$.*

Proof. To prove the first statement, suppose that $ab = ac$. Then, multiplying both sides on the left by a^{-1} we obtain

$$\begin{aligned} a^{-1}ab &= a^{-1}ac \\ eb &= ec \\ b &= c, \end{aligned}$$

as claimed. The second statement follows by multiplying by a^{-1} on the right. □

Proposition 4.1.16. *Let G be a group, and $g, h \in G$. Then there is a unique $x \in G$ such that $gx = h$. Similarly, there is a unique $y \in G$ such that $yg = h$.*

Proof. Let $x = g^{-1}h$. Then

$$\begin{aligned} gx &= gg^{-1}h \\ &= eh \\ &= h. \end{aligned}$$

To show that x is unique, suppose that $gx_1 = h = gx_2$, then, canceling g on the left, we obtain $x_1 = x_2$. Similarly, $y = hg^{-1}$. □

Exercise 4.1.17. Let G be a group. Define a map

$$L : G \to \mathrm{Perm}(G),\ g \mapsto (L_g : G \to G,\ h \mapsto L_g(h) = gh).$$

Show that $L : G \to \mathrm{Perm}(G)$ is an injective map.

Definition 4.1.18 (Trivial group)**.** The trivial group is the group on the set $\{e\}$, with $e * e = e$.

Remark 4.1.19. Note that in the trivial group, calling the single element e is arbitrary (although it is the identity), and groups with one element will be denoted by $\{e\}$, $\{1\}$ or $\{0\}$.

Exercise 4.1.20. Let $X = \{e, x\}$ be a set with 2 elements. Show that if $*$ is a binary operation on X that makes X a group with identity e, then $x^2 = e$.

Definition 4.1.21. Direct Product Let G, G' be groups. Let

$$G \times G' := \{(g, g') \mid g \in G,\ g' \in G'\}$$

be the Cartesian product. We can define a group operation on $G \times G'$ component wise, namely

$$(g_1, g_1')(g_2, g_2') := (g_1 g_2, g_1' g_2').$$

Then $G \times G'$ is a group called the *direct product of* G *and* G' with identity $(e_G, e_{G'})$. More generally, if $G_1, \dots, G_n$ are groups, we let

$$\prod_{i=1}^{n} G_i := G_1 \times G_2 \times \cdots \times G_n = \{(g_1, \dots, g_n) \mid g_i \in G_i\},$$

and define the group operation componentwise, called the *direct product.* If $e_i \in G_i$ is the identity, then the identity of the direct product is $(e_1, \dots, e_n)$.

Example 4.1.22. We've seen that $\mathbb{R}$ is a group under $+$. And real n-dimensional space

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-times}}$$

is an abelian group with addition defined componentwise.

Definition 4.1.23. If G is a group and $|G| < \infty$ is a finite set, then G is called a *finite group.* And $|G|$ is called the *order of* G.

Example 4.1.24.

(1) The order of the trivial group $\{1\}$ is 1.
(2) The order of the group $\{1, -1\}$ under multiplication is 2.
(3) The order of the group D_4 is 8, namely

$$D_4 = \{e, r, r^2, r^3, s_v, s_h, s_d, s_{d'}\}.$$

(4) The orders of the first few symmetric groups are

$$|S_1| = 1,$$
$$|S_2| = 2,$$
$$|S_3| = 6.$$

Can you guess the order of S_n?

Since group operations are associative, we set some notation to avoid unnecessary parentheses.

Notation 4.1.25. Let G be a group, and $x_1 \dots, x_n \in G$. We recursively define

$$x_1 \cdots x_n = (x_1 \cdots x_{n-1}) x_n.$$

Since the group operation is associative, we can place parentheses anywhere in the product and still obtain the same product. We also write

$$\prod_{i=1}^{n} x_i = x_1 \cdots x_n.$$

For $x \in G$, we define

$$x^n = \prod_{i=1}^{n} x = \underbrace{x \cdots x}_{n\text{-times}},$$

if $n = 0$ then we define $x^0 = e$. If $n < 0$, then we define $x^n = (x^{-1})^{-n}$, e.g.

$$x^{-2} = x^{-1}x^{-1}.$$

Thus, as you can verify using induction, we have

$$x^{m+n} = x^m x^n \text{ for all } m, n \in \mathbb{Z}$$
$$\text{and } (x^n)^m = x^{nm} \text{ for all } m, n \in \mathbb{Z}.$$

Remark 4.1.26. Since G may not be abelian, for $x, y \in G$ in general $(xy)^n \neq x^n y^n$.

Notation 4.1.27. If the group G is written additively, using $+$, then we have

$$\sum_{i=1}^{n} x_i = (x_1 + \cdots + x_{n-1}) + x_n = x_1 + \cdots + x_n,$$

if $n > 0$ then

$$nx = \sum_{i=1}^{n} x = \underbrace{x + \cdots + x}_{n\text{-times}},$$

and we agree that $0x = 0$, and if $n < 0$, then

$$nx = (-n)(-x).$$

And again, using induction, you can readily verify that

$$(m+n)x = mx + nx \text{ and } (mn)x = m(nx).$$

4.2. **Cyclic Groups.**

Definition 4.2.1. We say a group G is *cyclic* if there is some $a \in G$ such that every element $g \in G$ can be written in the form $g = a^n$ for some integer n. We call the element a a *generator* of G.

Proposition 4.2.2. *Let G be a cyclic group with generator a. Then G is abelian.*

Proof. Suppose $x, y \in G$. Since G is cyclic, we have $x = a^n$ and $y = a^m$ for some integers $n, m \in \mathbb{Z}$. Thus

$$xy = (a^n)(a^m) = a^{n+m} = a^{m+n} = (a^m)(a^n) = yx,$$

whence G is abelian. □

Example 4.2.3. The group $\{1, -1\}$ under multiplication is cyclic, generated by -1.

Example 4.2.4. If G is a group, and $a \in G$, consider the subset

$$\langle a \rangle := \{a^n \mid n \in \mathbb{Z}\} \subseteq G.$$

Then $\langle a \rangle$ is a cyclic group, with identity $a^0 = e$ and group operation given by the group operation of G.

Proposition 4.2.5. *Let G be a cyclic group with generator a, then there is a surjective map*

$$\varphi : \mathbb{Z} \to G, \ n \mapsto a^n.$$

Moreover, G is infinite if and only if φ is injective.

Proof. Since G is cyclic with generator a, every element of G can be written as a^n for some integer $n \in Z$. By definition, $a^n = \varphi(n)$, thus φ is surjective.

Now suppose that G is infinite. We claim that φ is injective, hence a bijection. We prove the contrapositive, that if φ is not injective, then G is finite. Thus suppose that $\varphi(n) = \varphi(m)$ for some $n \neq m$. Without loss of generality, suppose that $n > m$. Since

$$\varphi(n-m) = a^{n-m} = a^n a^{-m} = a^n (a^m)^{-1} = a^n (a^n)^{-1} = a^n a^{-n} = a^{n-n} = a^0 = e,$$

by the well-ordering principle, there is some smallest positive integer d such that $a^d = e$. Then for $n \in \mathbb{Z}$, writing $n = qd + r$ with $0 \leq r < d$, we have

$$a^n = a^{qd+r} = a^{qd}a^r = (a^d)^q a^r = e^q a^r = ea^r = a^r,$$

and so

$$G = \{e = a^0, \ldots, a^{d-1}\},$$

which are all distinct as d is the smallest positive integer with $a^d = e$. Thus G is finite. Therefore, if G is not finite, then φ is injective.

Conversely, if φ is injective, then $|\mathbb{Z}| \leq |G|$, and so G is infinite. □

Remark 4.2.6. Let G be a cyclic group with generator a, and $\varphi : \mathbb{Z} \to G$ the map of Proposition 4.2.5. You can check that the subset of $\mathbb{Z}$ such that $\varphi(n) = e$ is an ideal of $\mathbb{Z}$, as in Definition 3.1.7. And d, the smallest positive integer such that $a^d = e$, is the generator of that ideal.

Remark 4.2.7. Note that the map $\varphi : \mathbb{Z} \to G, n \mapsto a^n$ of Proposition 4.2.5 satisfies

$$\varphi(n + m) = \varphi(n)\varphi(m).$$

In essence, φ plays nicely with the group structures of $\mathbb{Z}$ and G.

Remark 4.2.8. If G is an infinite cyclic group, then G is basically the same as the group $\mathbb{Z}$ under addition. The bijection $\varphi : \mathbb{Z} \to G$ really just changes addition in $\mathbb{Z}$ into addition of the exponents $(a^n a^m = a^{n+m})$ in G.

What happens if G is finite? Well, the proof already showed what happens.

Definition 4.2.9. Let G be a group, and $g \in G$. The *order* of g, denoted $\mathrm{ord}(g)$ is the smallest $n \in \mathbb{Z}_{>0}$ such that $g^n = e$, if such an n exists, else $\mathrm{ord}(g) = \infty$. We say that a has *finite order* if $\mathrm{ord}(a)$ is finite.

. . . . End of Class 7 (1/28) . . . .

Proposition 4.2.10. *Let G be a finite group. Then every $g \in G$ has finite order.*

Proof. For $g \in G$, there are two possibilities:

(1) for all $n, m \in \mathbb{Z}_{>0}$, if $n \neq m$, then $g^n \neq g^m$.
(2) there are $n, m \in \mathbb{Z}_{>0}$ such that $g^n = g^m$ but $n \neq m$.

These two options are exactly whether the map $\varphi : \mathbb{Z} \to \langle g \rangle \subseteq G$ is injective or not injective. In the first case, there would be an injection $\mathbb{Z} \to G$, which is a contradiction as G is finite.

Therefore, there are some $n, m \in \mathbb{Z}_{>0}$ such that $g^n = g^m$ and $n \neq m$. Without loss of generality, we may assume $n < m$, and so

$$g^{m-n} = g^m g^{-n} = g^n g^{-n} = g^{n-n} = g^0 = e,$$

hence as $m - n \in \mathbb{Z}_{>0}$, g has finite order. □

Corollary 4.2.11. *Let G be a group and $a \in G$. Suppose that a has finite order. If $a^n = e$, then* $\mathrm{ord}(a) \mid n$. *And*

$$\langle a \rangle = \{e = a^0, a^1, \ldots, a^{\mathrm{ord}(a)-1}\}.$$

Proof. Let $I = \{n \in \mathbb{Z} \mid a^n = e\}$. We claim that $I \subseteq \mathbb{Z}$ is an ideal. Indeed, we have $0 \in I$ as $a^0 = e$ by definition. If $n, m \in I$, then

$$a^{n+m} = a^n a^m = ee = e,$$

so $n + m \in I$. Also, if $n \in I$, then

$$a^{-n} = (a^n)^{-1} = e^{-1} = e,$$

so $-n \in I$. Hence I is an ideal. Thus, by Theorem 3.1.11, there is some $d \in I$, the smallest non-negative integer in I, such that $I = (d)$. We cannot have $d = 0$, as a has finite order. Thus d is the smallest non-negative integer such that $a^d = e$, thus $d = \operatorname{ord}(a)$. Therefore, as $I = (d)$, for $n \in I$, $d \mid n$, as claimed.

Exercise 4.2.12. Let G be a group, and $g \in G$. Suppose that g has finite order. Show that $\langle g \rangle = \{e = g^0, g, \ldots, g^{\operatorname{ord}(g)-1}\}$.

□

Exercise 4.2.13. Let G be a group, and $g \in G$. Suppose that g has finite order. Show that $g^{-1} = g^n$ for some $n \geq 1$.

Exercise 4.2.14. Let G be a group such that for all $g \in G$, $g^2 = e$. Prove that G is abelian.

4.2.1. *The groups $\mathbb{Z}/n\mathbb{Z}$.* Let's see some finite cyclic groups in a more concrete way.

We'll need to recall some facts about modular arithmetic.

Recall 4.2.15. Let $n \in \mathbb{Z}$. For integers a, b we write $a \equiv b \mod n$ if $n \mid (b - a)$. This is an equivalence relation, and we denote the equivalence class of a by $\overline{a}$. Thus

$$\overline{a} = \{b \in \mathbb{Z} \mid b \equiv a \mod n\}.$$

The equivalence classes modulo n are

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}.$$

Proposition 4.2.16. *For $a, b, c, d \in \mathbb{Z}$, if $a \equiv c \mod n$ and $b \equiv d \mod n$, then*

(1) $a + b \equiv c + d \mod n$.
(2) $ab \equiv cd \mod n$.

In terms of equivalence classes, Proposition 4.2.16(1) says that we can define a binary operation on the equivalence classes modulo n by

$$\overline{a} + \overline{b} = \overline{a + b},$$

which is well defined since

$$\text{if } \overline{a} = \overline{c} \text{ and } \overline{b} = \overline{d}, \text{ then } \overline{a} + \overline{b} = \overline{a+b} = \overline{c+d} = \overline{c} + \overline{d}.$$

We needed to be careful and make sure that our definition of adding equivalence classes did not depend on the representative we chose.

Proposition 4.2.17. *Let $n \in \mathbb{Z}_{>0}$. The set of residues modulo n, $\{\overline{0}, \ldots, \overline{n-1}\}$ with the operation $\overline{a} + \overline{b} = \overline{a + b}$ is an abelian cyclic group of order n.*

Proof. Associativity and commutativity come from the associativity and commutativity of $+$ in $\mathbb{Z}$. For example, to show associativity, we compute

$$\begin{aligned}\overline{a} + (\overline{b} + \overline{c}) &= \overline{a} + \overline{b + c} \\ &= \overline{a + (b + c)} \\ &= \overline{(a + b) + c} \\ &= \overline{a + b} + \overline{c} \\ &= (\overline{a} + \overline{b}) + \overline{c}.\end{aligned}$$

Commutativity is similar. Finally, $\overline{0}$ is the identity element, and the inverse of $\overline{a}$ is $\overline{-a}$. Lastly, to see that this group is cyclic, note that $\overline{a} = n(\overline{a})$. □

Definition 4.2.18. The group $\{\overline{0}, \dots, \overline{n-1}\}$ under addition is called the *cyclic group of order* n, denoted by $\mathbb{Z}/n\mathbb{Z}$, or sometimes by $\mathbb{Z}_n$. One generator of $\mathbb{Z}/n\mathbb{Z}$ is $\overline{1}$, though there may sometimes be others.

Example 4.2.19. The group $\mathbb{Z}/2\mathbb{Z} = \{\overline{0}, \overline{1}\}$ under addition is a finite cyclic group of order 2.

Example 4.2.20. The group $\mathbb{Z}/8\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}, \dots, \overline{7}\}$ is a finite cyclic group of order 8. Note that by Proposition 4.2.10, every element has order dividing 8.

We've seen that $\overline{1}$ is a generator for $\mathbb{Z}/8\mathbb{Z}$. But in fact, so is $\overline{3}$, or $\overline{5}$, or $\overline{7}$. For example, we have

$$\begin{aligned}\overline{3} &= \overline{3} \\ 2(\overline{3}) &= \overline{6} \\ 3(\overline{3}) &= \overline{1} \\ 4(\overline{3}) &= \overline{4} \\ 5(\overline{3}) &= \overline{7} \\ 6(\overline{3}) &= \overline{2} \\ 7(\overline{3}) &= \overline{5} \\ 8(\overline{3}) &= \overline{0}.\end{aligned}$$

In fact, this is a very general feature.

Proposition 4.2.21. *Let G be a finite cyclic group of order n, and let a be a generator.*

(i) *If $r \in \mathbb{Z}_{>0}$ with $\gcd(r, n) = 1$, then a^r is a generator of G.*
(ii) *Conversely, if a^r is a generator of G, then $\gcd(r, n) = 1$.*

Proof. Since $\gcd(r, n) = 1$, there are some $x, y \in \mathbb{Z}$ such that $xr + yn = 1$. Thus for $g = a^k \in G$, we have

$$a^k = a^{k \times 1} = a^{kxr + kyn} = a^{kxr}a^{kyn} = (a^r)^{xk}e = (a^r)^{xk}.$$

Therefore a^r is also a generator of G, proving (i).

To prove (ii), let $b \in G$. Then $\langle b \rangle = \{e = b^0, b, b^2, \dots, b^{\operatorname{ord}(b)-1}\} \subseteq G$, and by comparing sizes of $\langle b \rangle$ and G, b is a generator if and only if $\operatorname{ord}(b) = n$.

Now suppose that b is a generator of G. Since a is a generator, $b = a^r$ for some r. Suppose for contradiction that $\gcd(r, n) = d > 1$. Then $r = md$, and so $b = a^r = (a^m)^d$. However, since $d > 1$, $n = kd$ for some $k < n$. Thus, since

$$b^k = ((a^m)^d)^k = (a^m)^{kd} = a^{nm} = (a^n)^m = e^m = e,$$

$\operatorname{ord}(b) = \operatorname{ord}(a^m)^d \leq k < n$. Hence $\operatorname{ord}(b) \leq k < n$, which is a contradiction to n being a generator. Thus $\gcd(r, n) = 1$, and every generator of G is of the form a^r for some r coprime to n. □

Writing things additively, this immediately gives the following characterization of generators of $\mathbb{Z}/n\mathbb{Z}$.

Corollary 4.2.22. *Let $n \in \mathbb{Z}_{>0}$. Then $\overline{a}$ is a generator of $\mathbb{Z}/n\mathbb{Z}$ if and only if* $\gcd(n, a) = 1$.

Proof. A generator of the finite cyclic group of order n, $\mathbb{Z}/n\mathbb{Z}$, is $\overline{1}$. Thus, by Proposition 4.2.21, $\overline{a}$ is a generator of $\mathbb{Z}/n\mathbb{Z}$ if and only if $\overline{a} = r\overline{1}$ for some $r \in \mathbb{Z}_{>0}$ with $\gcd(r, n) = 1$. □

Another example of groups comes from the residues modulo n that are invertible under multiplication. By Proposition 4.2.16(2), we have

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b},$$

thus we can define a binary operation on $\mathbb{Z}/n\mathbb{Z}$ by multiplying residue classes modulo n. Note, however, that not every residue class modulo n is invertible.

Definition 4.2.23. Let $n \in \mathbb{Z}_{>0}$. We define

$$(\mathbb{Z}/n\mathbb{Z})^\times := \{\overline{a} \in \mathbb{Z}/n\mathbb{Z} \ \mid \ \text{there exists } \overline{c} \in \mathbb{Z}/n\mathbb{Z} \text{ with } \overline{a} \cdot \overline{c} = \overline{1}\},$$

which are the residue classes modulo n with a multiplicative inverse.

Proposition 4.2.24. *The set $(\mathbb{Z}/n\mathbb{Z})^\times$ is a group under multiplication, the identity is $\overline{1}$.*

Proof. We must show that multiplication of residue classes defines a binary operation on $(\mathbb{Z}/n\mathbb{Z})^\times$. Namely, that is $\overline{a}, \overline{b} \in (\mathbb{Z}/n\mathbb{Z})^\times$, then $\overline{a} \cdot \overline{b} = \overline{ab} \in (\mathbb{Z}/n\mathbb{Z})^\times$. Since $\overline{a}, \overline{b} \in (\mathbb{Z}/n\mathbb{Z})^\times$, there are $\overline{c}, \overline{d} \in (\mathbb{Z}/n\mathbb{Z})^\times$ such that

$$\overline{a} \cdot \overline{c} = \overline{1} \text{ and } \overline{b} \cdot \overline{d} = \overline{1}.$$

Thus

$$\begin{aligned} \overline{a} \cdot \overline{b} \cdot \overline{c} \cdot \overline{d} &= \overline{a \cdot b \cdot c \cdot d} \\ &= \overline{(a \cdot c) \cdot (b \cdot d)} \\ &= (\overline{a} \cdot \overline{c}) \cdot (\overline{b} \cdot \overline{d}) \\ &= \overline{1} \cdot \overline{1}, \end{aligned}$$

whereby $\overline{a} \cdot \overline{c} = \overline{ac} \in (\mathbb{Z}/n\mathbb{Z})^\times$.

That multiplication of residue classes is associative is clear, and that $\overline{1}$ is the identity is also clear. By definition, every element has an inverse under multiplication. Thus $(\mathbb{Z}/n\mathbb{Z})^\times$ is a group. □

Proposition 4.2.25. *For $n \in \mathbb{Z}_{>0}$, $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$ if and only if* $\gcd(a, n) = 1$.

Proof. Suppose that $a \in (\mathbb{Z}/n\mathbb{Z})^\times$. Then there is some $\overline{c}$ such that

$$\overline{a} \cdot \overline{c} = \overline{1}.$$

Reading it modulo n this means

$$ac \equiv 1 \mod n,$$

i.e. $ac - 1 = kn$ for some $k \in \mathbb{Z}$. Thus $1 = ac - kn$, and so $\gcd(a, n) = 1$.

Conversely, suppose that $\gcd(a, n) = 1$, so there is some $c, k \in \mathbb{Z}$ such that $ac + kn = 1$, reading this modulo n gives

$$\overline{a} \cdot \overline{c} = \overline{1},$$

and thus $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$. □

Thus we have seen that the residue classes coprime to n are both the generators of $\mathbb{Z}/n\mathbb{Z}$ as well as the elements of $(\mathbb{Z}/n\mathbb{Z})^\times$.

Definition 4.2.26. For $n \in \mathbb{Z}_{>0}$, the *Euler totient function*, $\varphi(n)$, is defined by

$$\varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times| = \#\{1 \leq a \leq n \ \mid \ \gcd(a, n) = 1\}.$$

Example 4.2.27. $\varphi(8) = 4$, as we have seen that $(\mathbb{Z}/8\mathbb{Z})^\times = \{\overline{1}, \overline{3}, \overline{5}, \overline{7}\}$.

Proposition 4.2.28. *For p prime, $|(\mathbb{Z}/p\mathbb{Z})^\times| = p - 1$.*

Proof. Clearly if $0 \leq d < p$, then $\gcd(d, p) = 1$. □

Example 4.2.29. More generally, you can show that for a prime p, $\varphi(p^a) = p^a - p^{a-1} = p^{a-1}(p-1)$. and that φ is *multiplicative*, in the sense that

$$\varphi(ab) = \varphi(a)\varphi(b) \text{ if } \gcd(a, b) = 1.$$

Together, this gives the general formula for $\varphi(n)$, namely if $n = p_1^{a_1} \cdots p_r^{a_r}$, then

$$\begin{aligned} \varphi(n) &= \varphi(p_1^{a_1}) \cdots \varphi(p_r^{a_r}) \\ &= p_1^{a_1 - 1}(p_1 - 1) \cdots p_r^{a_r - 1}(p_r - 1). \end{aligned}$$

Example 4.2.30. Let's do some hands on computation with the group $(\mathbb{Z}/8\mathbb{Z})^\times$. We can check that every element has order 2, for example $\overline{3} \cdot \overline{3} = \overline{9} = \overline{1}$. Thus, while $(\mathbb{Z}/8\mathbb{Z})^\times$ is abelian, it is not cyclic (as any generator would have to have order 4)! In fact, we'll see later that $(\mathbb{Z}/8\mathbb{Z})^\times$ looks like the group

$$(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z}).$$

....................................End of Class 8 (1/30) ....

4.3. Subgroups.

Let G be a group. In thinking about groups, we've many times made use of the set

$$\langle g \rangle = \{e = g^0, g, g^2, \dots\} \subseteq G \text{ for } g \in G.$$

This is an example of a much more general phenomenon.

Definition 4.3.1. Subgroup Let $(G, *)$ be a group with identity e. A subset $H \subseteq G$ is called a *subgroup* if

(**SG 1**) $e \in H$, (contains the identity)
(**SG 2**) for all $g, h \in H$, $g * h \in H$, (closed under multiplication)
(**SG 3**) for all $h \in H$, $h^{-1} \in H$. (closed under inverses)

When H is a subgroup of G, we will write $H \leq G$. If we want to emphasize that $H \neq G$, then we will write $H \lneq G$ or $H < G$, and call H a *proper subgroup.*

Remark 4.3.2. When $H \leq G$ is a subgroup, then H is itself a group, the group operation is the same as that of G, just restricted to only the elements of H.

Proposition 4.3.3. *Let G be a group, then $\{e\} \leq G$, called the trivial subgroup, and $G \leq G$ is a subgroup.*

Proof. That G is a subgroup is clear, as the conditions to be a subgroup are the same as those of being a group. To show that $\{e\}$ is a subgroup, we just note that $e \in \{e\}$, $e * e = e$, and $e^{-1} = e$. □

Example 4.3.4.

(1) The additive group of rational numbers $\mathbb{Q} < \mathbb{R}$ is a subgroup of the real numbers under addition.
(2) The group $\{1, -1\}$ is a subgroup of $\{1, -1, i, -i\}$.
(3) Let $S^1 = \{z \in \mathbb{C} \mid |z| = 1\} \subseteq \mathbb{C}^\times$. Then $S^1 < \mathbb{C}^\times$ is a subgroup. (Recall: $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$ is a group under multiplication)
(4) Let V be a vector space, then a subspace $W \subseteq V$ is a subgroup.
(5) Let G be a group and $g \in G$. Then $\langle g \rangle = \{g^n \mid n \in \mathbb{Z}\} \leq G$ is a subgroup.
(6) The symmetric groups S_n are subgroups of each other, simply by fixing the elements larger than n. Thus $S_1 < S_2 < S_3 < S_4 < \cdots < S_n < S_{n+1} < \cdots$.

(7) If you label the vertices of a square with $1, 2, 3, 4$, the dihedral group D_4 from Figure 4 is a subgroup of S_4, with each element of D_4 changing the labels of the vertices as it moves the square around.

There is a general way of obtaining subgroups from a group G.

Definition 4.3.5. Let G be a group, and $S \subseteq G$ a non-empty subset. Let

$$\langle S\rangle := \left\{ \prod_{i=1}^{n} x_i \ \mid\ x_i \in S \text{ or } x_i^{-1} \in S \right\},$$

with the empty product being e. Then $\langle S\rangle$ is a subgroup of G, called the *subgroup generated by* S, and we call S a set of *generators* of $\langle S\rangle$.

Exercise 4.3.6. Let G be a group, and $S \subseteq G$ be a non-empty subset. Then $\langle S\rangle$ is a subgroup of G.

Proposition 4.3.7. *Let G be a group and $S \subseteq G$ be a non-empty subset. Then $\langle S\rangle$ is the smallest subgroup of G containing S. That is, if $H \leq G$ is any subgroup containing S, then $\langle S\rangle \leq H$.*

Proof. Suppose $H \leq G$ is a subgroup such that $S \subseteq H$. Then since H is a subgroup, for every $x_i \in S$, $x_i^{-1} \in H$. And since H is a subgroup, for each $x_i, x_j \in S$, $x_i x_j \in H$ and $x_i x_j^{-1} \in H$. Thus every finite product of element of S and their inverses are in H. Hence, as every element of $\langle S\rangle$ is a finite product of elements of S or their inverses, we have $\langle S\rangle \leq H$. □

Exercise 4.3.8. Prove the following proposition.

Proposition 4.3.9. *Let $H, K \leq G$ be subgroups, then $H \cap K \leq G$ is a subgroup.*

Remark 4.3.10. There is another way of presenting groups, using *generators and relations*. Let S be a set and R a set of equations

$$R = \{R_i : x_1 \cdots x_r = e \ \mid\ x_i \in S \text{ or } x_i^{-1} \in S\}.$$

Then, if the set of strings of symbols in S, simplified by the relations in R (using the rule that $xe = ex = x$ for all $x \in S$), forms a group, we call

$$G = \langle S \mid R\rangle$$

a *presentation* for the group G.

RH: Caution!!! It is very hard (even impossible) to tell if given a set of generators S and relations R, you even obtain a group! Worse still, even if you know you've obtained a group, it is hard to tell whether $\langle S \mid R\rangle = \{1\}$ or not! It is also difficult (or even impossible) to show that two strings of elements of S give the same element when simplified!

For example, you can show that

$$\langle x, y \mid x^2 = y^2 = (xy)^2 = e\rangle$$

is a finite group of order 4, in fact $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. However

$$\langle a, b\rangle \mid a^3 = b^3 = (ab)^3 = e\rangle$$

is an infinite group!

Example 4.3.11. The *dihedral group* D_n of order $2n$ is a group given by the presentation

$$D_n = \langle r, s \mid r^n = s^2 = e, rs = sr^{-1}\rangle.$$

It is the symmetry group of the regular n-gon.

Let us turn to some more familiar examples.

Lemma 4.3.12. *For $n \in \mathbb{Z}$, denote by $n\mathbb{Z} \coloneqq \{nk \mid k \in \mathbb{Z}\}$. That is, $n\mathbb{Z}$ is the ideal (n). Then $n\mathbb{Z} \leq \mathbb{Z}$ is a subgroup.*

Proof. Let us check the conditions to be a subgroup. We have $0 \in n\mathbb{Z}$ as $0 = 0 \times n$. Let $x, y \in n\mathbb{Z}$, then $x = an$, $y = bn$ for some $a, b \in \mathbb{Z}$. So $x + y = n(a + b) \in n\mathbb{Z}$. Let $x \in n\mathbb{Z}$, then $x = na$ for some $a \in \mathbb{Z}$. Then $-x = n \times (-a) \in n\mathbb{Z}$. □

Note that these are exactly the conditions of an ideal of $\mathbb{Z}$.

Theorem 4.3.13. *The subgroups of $\mathbb{Z}$ are exactly the ideals $n\mathbb{Z}$.*

Proof. We've seen in Lemma 4.3.12 that the ideals $n\mathbb{Z}$ are subgroups of $\mathbb{Z}$. Thus it remains to show that any subgroup is of this form. We show that every subgroup $H \leq \mathbb{Z}$ is an ideal, as then the result follows from Theorem 3.1.11. Let $H \leq \mathbb{Z}$ be a subgroup. Then $0 \in H$, as every subgroup contains the identity. If $x, y \in H$, then $x + y \in H$, as H is closed under the group operation. Lastly, if $x \in H$, then $-x \in H$, as subgroups are closed under inverses. Hence if $a \in \mathbb{Z}$, then $ax \in H$. Thus H is an ideal. Thus, by Theorem 3.1.11, $H = n\mathbb{Z}$ for some $n \in \mathbb{Z}$, as desired. □

In particular, the subgroups of $\mathbb{Z}$ are

- $d\mathbb{Z}$ for $d \geq 1$, or
- $\{0\} = 0\mathbb{Z}$.

In the former case, the subgroup $d\mathbb{Z}$ is cyclic, generated by d, and infinite. Thus there is a surjective map $\mathbb{Z} \to d\mathbb{Z}, n \mapsto dn$, which is also injective as $d\mathbb{Z}$ is infinite (as in Proposition 4.2.5).

We can also understand the subgroups of $\mathbb{Z}/n\mathbb{Z}$.

Theorem 4.3.14. *The subgroups of $\mathbb{Z}/n\mathbb{Z}$ are exactly those of the form*

$$d\mathbb{Z}/n\mathbb{Z} \coloneqq \{d \cdot \overline{k} \mid \overline{k} \in \mathbb{Z}/n\mathbb{Z}\} \text{ such that } d \mid n.$$

Proof. First, we must show that if $d \mid n$, then $d\mathbb{Z}/n\mathbb{Z}$ is a subgroup. We check the conditions to be a subgroup.

- $\overline{0} = d\overline{0} \in d\mathbb{Z}/n\mathbb{Z}$.
- Let $x, y \in d\mathbb{Z}/n\mathbb{Z}$. Then $x = d \cdot \overline{a}$, $y = d \cdot \overline{b}$ for some $\overline{a}, \overline{b} \in \mathbb{Z}/n\mathbb{Z}$. Then $x + y = d \cdot \overline{a} + d \cdot \overline{b} = d \cdot (\overline{a + b}) \in d\mathbb{Z}/n\mathbb{Z}$.
- Let $x \in d\mathbb{Z}/n\mathbb{Z}$. Thus $x = d \cdot \overline{a}$ for some $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$. And $-x = d \cdot (\overline{-a}) \in d\mathbb{Z}/n\mathbb{Z}$.

Thus $d\mathbb{Z}/n\mathbb{Z}$ is a subgroup of $\mathbb{Z}/n\mathbb{Z}$. (Alternatively, $d\mathbb{Z}/n\mathbb{Z} = \langle \overline{d} \rangle \leq d\mathbb{Z}/n\mathbb{Z}$.)

Conversely, suppose that $H \leq \mathbb{Z}/n\mathbb{Z}$ is a subgroup. Let

$$\widetilde{H} = \{h \in \mathbb{Z} \mid \overline{h} \in H\}.$$

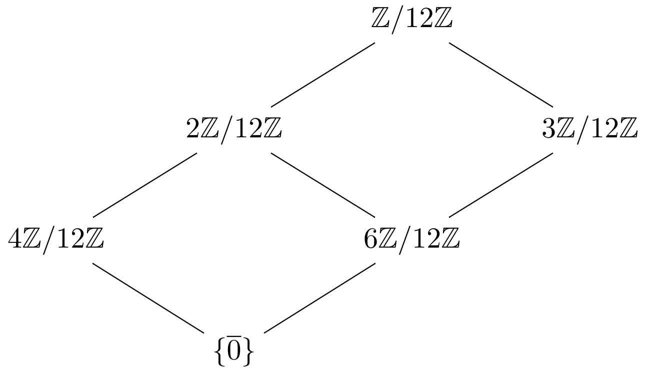
By the well-ordering principle, $\widetilde{H}$ must have a least non-negative element, call it d. We will show that $\widetilde{H}$ is the ideal $d\mathbb{Z}$. Since $d \in \widetilde{H}$, $\overline{d} \in H$, and since H is a subgroup, $q \cdot \overline{d} = \overline{qd} \in H$ for all $q \in \mathbb{Z}$, thus $d\mathbb{Z} \subseteq \widetilde{H}$. To prove the other containment, let $h \in \widetilde{H}$. By the division algorithm, we can write $h = dq + r$ with $0 \leq r < d$. Thus, as H is a subgroup, $\overline{h} - q \cdot \overline{d} = \overline{r} \in H$, hence $r \in \widetilde{H}$. Since d is the least non-negative integer in $\widetilde{H}$, $r = 0$, and so $h = qd \in d\mathbb{Z}$. Thus $\widetilde{H} = d\mathbb{Z}$. This means that for $\overline{h} \in H$, $h = dk$, and so $\overline{h} = d \cdot \overline{k}$ with $k \in \mathbb{Z}/n\mathbb{Z}$, hence $H = d\mathbb{Z}/n\mathbb{Z}$. In particular, as $\overline{0} = \overline{n} \in H$, $n \in \widetilde{H}$. Thus $n \in d\mathbb{Z}$, and $d \mid n$, as claimed. □

Example 4.3.15. Let's find all the subgroups of $\mathbb{Z}/12\mathbb{Z}$. We have one for each divisor of 12, which are $1, 2, 3, 4, 6, 12$. The subgroups are thus

- $1\mathbb{Z}/12\mathbb{Z} = \mathbb{Z}/12\mathbb{Z}$,
- $2\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{2}, \overline{4}, \overline{6}, \overline{8}, \overline{10}\}$,
- $3\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{3}, \overline{6}, \overline{9}\}$,
- $4\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{4}, \overline{8}\}$,

- $6\mathbb{Z}/12\mathbb{Z} = \{\overline{0}, \overline{6}\}$,
- $12\mathbb{Z}/12\mathbb{Z} = \{\overline{0}\}$.

We can organize them into a diagram called the *subgroup lattice* of $\mathbb{Z}/12\mathbb{Z}$:


$\mathbb{Z}/12\mathbb{Z}$
$2\mathbb{Z}/12\mathbb{Z}$
$3\mathbb{Z}/12\mathbb{Z}$
$4\mathbb{Z}/12\mathbb{Z}$
$6\mathbb{Z}/12\mathbb{Z}$
$\{\overline{0}\}$

where a line between an upper subgroup U and a lower subgroup L means $L < U$ (for example, $\{\overline{0}\} < 4\mathbb{Z}/12\mathbb{Z}$).

In general, for a finite group G, one can draw a subgroup lattice showing all subgroups of G and how they are contained. In practice, it gets quite complicated. You can even draw parts of the subgroup lattice for an infinite group, though those get complicated very quickly!

Exercise 4.3.16. Find all the subgroups of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and draw the subgroup lattice.

Exercise 4.3.17. Find all the subgroups of S_3 and draw the subgroup lattice. Hint: Writing the elements of S_3 as matrices (as in Example 4.1.3(5)) or as lists of numbers (as in Example 4.1.10) will make it much easier to do computations.

Exercise 4.3.18. Show that for subgroups $n\mathbb{Z} \le \mathbb{Z}$ and $d\mathbb{Z} \le \mathbb{Z}$, there is a containment of subgroups $n\mathbb{Z} \le d\mathbb{Z}$ if and only if $d \mid n$. Draw part of the subgroup lattice of the subgroups of $\mathbb{Z}$ for the subgroups $\mathbb{Z}, 2\mathbb{Z}, 3\mathbb{Z}, \dots, 20\mathbb{Z}$, and the trivial subgroup $\{0\} = 0\mathbb{Z}$.

.... End of Class 9 (2/2) ....

Let's see some more examples of subgroups.

Example 4.3.19. Recall that $\mathrm{GL}_n(\mathbb{R}) = \{n \times n \text{ matrices with real entries}\}$ is a group under matrix multiplication. For $n = 2$, we have

$$\mathrm{GL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \;\middle|\; a, b, c, d \in \mathbb{R},\ ad - bc \neq 0 \right\},$$

and you may recognize $ad - bc = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. There is a nice subgroup, namely

$$\mathrm{SL}_2(\mathbb{R}) := \{M \in \mathrm{GL}_2(\mathbb{R}) \mid \det M = 1\}.$$

To see that $\mathrm{SL}_2(\mathbb{R})$ is a subgroup, recall that the identity is $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, which has determinant 1, thus $I_2 \in \mathrm{SL}_2(\mathbb{R})$. Furthermore, recall that for $M, N \in \mathrm{GL}_n(\mathbb{R})$,

$$\det(MN) = \det(M) \cdot \det(N).$$

From this it follows that $\mathrm{SL}_2(\mathbb{R}) \leq \mathrm{GL}_n(\mathbb{R})$.

Exercise 4.3.20. Show that

$$\mathrm{SL}_2(\mathbb{R}) \coloneqq \{M \in \mathrm{GL}_2(\mathbb{R}) \mid \det M = 1\}$$

is a subgroup of

$$\mathrm{GL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{R},\ ad - bc \neq 0 \right\}.$$

4.4. **Homomorphisms.** In many of the proofs so far, we've made use of constructing some maps between groups, for example in Proposition 4.2.5 and secretly also in Theorem 4.3.14. Maps between groups are most interesting when they play nicely with the group structures on both sides.

Definition 4.4.1. Let $(G, *_G)$ and $(H, *_H)$ be groups. A *group homomorphism* or a *morphism of groups* is a map

$$\varphi : G \to H$$

such that for all $x, y \in G$,

$$\varphi(x *_G y) = \varphi(x) *_H \varphi(y).$$

Example 4.4.2. Let G be a group, the identity map $\mathrm{id}_G : G \to G$ is a group homomorphism.

Example 4.4.3. The map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, k \mapsto \overline{k}$ is a group homomorphism, and was used in the proof of Theorem 4.3.14.

Example 4.4.4. Let G be a group, and $g \in G$. There is a group homomorphism

$$\varphi : \mathbb{Z} \to G,\ n \mapsto g^n.$$

This is exactly the the group homomorphism that was used in Proposition 4.2.5.

Example 4.4.5. Let $e \in \mathbb{R}$ be Euler's number (the number e from calculus). Then the exponential map $\mathbb{R} \to \mathbb{R}^\times_{>0},\ x \mapsto e^x$ is a group homomorphism from the additive group of real numbers into the multiplicative group of positive real numbers. Its inverse map, the logarithm, is also a group homomorphism. This amounts to the usual exponent and logarithm rules.

Exercise 4.4.6. Let G be a group. Consider the map $f : G \to G, g \mapsto g^{-1}$.

(1) Show that f is a bijection.
(2) Show that f is a group homomorphism if and only if G is abelian.
(3) What is $f \circ f$? What is f^{-1}?

Proposition 4.4.7. *Let $\varphi : G \to G'$ and $\psi : G' \to G''$ be group homomorphisms. Then the composition $\psi \circ \varphi : G \to G''$ is a group homomorphism.*

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

$$\psi \circ \varphi(gh) = \psi(\varphi(gh)) = \psi(\varphi(g)\varphi(h)) = \psi(\varphi(g))\psi(\varphi(h)).$$ □

Proposition 4.4.8. *Let $f : G \to G'$ be a group homomorphism, and let $e \in G$ and $e' \in G'$ be the identity elements. Then $f(e) = e'$.*

Proof. As f is a group homomorphism, we have

$$f(e) = f(ee) = f(e)f(e),$$

and multiplying by $f(e)^{-1}$ on the left gives

$$e' = f(e)^{-1}f(e) = f(e)^{-1}f(e)f(e) = e'f(e) = f(e).$$ □

Proposition 4.4.9. *Let $f : G \to G'$ be a group homomorphism, and let $g \in G$. Then $f(g^{-1}) = f(g)^{-1}$.*

Proof. We compute

$$e' = f(e) = f(gg^{-1}) = f(g)f(g^{-1}),$$

thus $f(g)^{-1} = f(g^{-1})$, as desired. □

Now let us see how group homomorphisms interact with subgroups.

Theorem 4.4.10. *Let $\varphi : G \to G'$ be a group homomorphism, and let $H' \leq G'$ be a subgroup. Then the inverse image of H' under φ,*

$$H = \varphi^{-1}(H') = \{h \in G \mid \varphi(h) \in H'\}$$

is a subgroup of G.

Proof. We verify that $H = f^{-1}(H')$ satisfies the conditions of a subgroup. First, as H' is a subgroup, $e' \in H'$. By Proposition 4.4.8, $\varphi(e) = e'$, $e \in H$. Now suppose that $x, y \in H$, then by definition, $\varphi(x), \varphi(y) \in H'$. As φ is a group homomorphism, $\varphi(xy) = \varphi(x)\varphi(y)$; and as H' is a subgroup, $\varphi(x)\varphi(y) \in H'$, whereby $\varphi(xy) \in H$. Lastly, suppose $x \in H$. Thus, by definition, $\varphi(x) \in H'$. As G is a group, $x^{-1} \in G$. Since φ is a group homomorphism, Proposition 4.4.9 gives $\varphi(x^{-1}) = \varphi(x)^{-1}$. As H' is a subgroup, and $\varphi(x) \in H'$, $\varphi(x)^{-1} = \varphi(x^{-1}) \in H'$ as well. Thus $x^{-1} \in H$. Therefore, $H = \varphi^{-1}(H') \leq G$ is a subgroup. □

Remark 4.4.11. This is exactly the fact that we used in the proof of Theorem 4.3.14, the ideal $\widetilde{H}$ is just the inverse image of the subgroup $H \leq \mathbb{Z}/n\mathbb{Z}$ under the group homomorphism

$$\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, k \mapsto \overline{k}.$$

Definition 4.4.12. Let $\varphi : G \to G'$ be a group homomorphism, and let $e' \in G'$ be the identity. The *kernel of* φ is

$$\ker\varphi := \varphi^{-1}(\{e'\}) = \{g \in G \mid \varphi(g) = e'\}.$$

Proposition 4.4.13. *Let $\varphi : G \to G'$ be a group homomorphism. Then $\ker\varphi \leq G$ is a subgroup.*

Proof. This follows immediately from Theorem 4.4.10, as $\ker\varphi$ is the inverse image of the trivial subgroup of G'. □

Example 4.4.14. Let G be a group, and $g \in G$. Let $\varphi : \mathbb{Z} \to G, n \mapsto g^n$. Then $\ker\varphi$ is either $\{0\}$ if g has infinite order or $d\mathbb{Z}$ where $d = \mathrm{ord}(g)$ if g has finite order.

Example 4.4.15. The map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, k \mapsto \overline{k}$ has kernel $n\mathbb{Z}$.

Theorem 4.4.16. *Let $\varphi : G \to G'$ be a group homomorphism, and let $e \in G$ be the identity element. Then φ is injective if and only if $\ker\varphi = \{e\}$.*

Proof. Let $e' \in G'$ be the identity element. If φ is injective, then clearly $\ker\varphi = \varphi^{-1}(e') = \{e\}$.

Conversely, suppose that $\ker\varphi = \{e\}$, and suppose $g, h \in G$ with $\varphi(g) = \varphi(h)$. Then

$$e' = \varphi(g)\varphi(h)^{-1} = \varphi(g)\varphi(h^{-1}) = \varphi(gh^{-1}).$$

Hence $gh^{-1} = e$, and consequently $g = h$, thus φ is injective. □

Example 4.4.17. Let $H \leq G$ be a subgroup. The inclusion $\iota_H \hookrightarrow G$ is an injective group homomorphism.

Notation 4.4.18. When a group homomorphism $\varphi : G \to G'$ is injective, we may emphasize this by writing

$$\varphi : G \hookrightarrow G'$$

and calling φ an *inclusion* or an *embedding*.

Definition 4.4.19. Let $\varphi : G \to G'$ be a group homomorphism. The *image of* φ is

$$\mathrm{im}\varphi := \varphi(G) = \{g' \in G' \ \mid \ g' = \varphi(g) \text{ for some } g \in G\}.$$

Example 4.4.20. Let $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z},\ k \mapsto \overline{k}$. Then the image of $d\mathbb{Z}$ under φ is

$$\varphi(d\mathbb{Z}) = \begin{cases} \{0\} & \text{if } n \mid d \\ d\mathbb{Z}/n\mathbb{Z} & \text{if } d \mid n \\ \mathbb{Z}/n\mathbb{Z} & \text{if } \gcd(d,n) = 1. \end{cases}$$

Theorem 4.4.21. *Let $\varphi : G \to G'$ be a group homomorphism. Then $\mathrm{im}\varphi \leq G'$ is a subgroup.*

Proof. Since $\varphi(e) = e'$, $e' \in \mathrm{im}\varphi$. Let $x = \varphi(g), y = \varphi(h) \in \mathrm{im}\varphi$. Then

$$xy = \varphi(g)\varphi(h) = \varphi(gh) \in \mathrm{im}\varphi.$$

Lastly, let $x = \varphi(g) \in \mathrm{im}\varphi$. Then as $g^{-1} \in G$,

$$\varphi(g)^{-1} = \varphi(g^{-1}) \in \mathrm{im}\varphi,$$

as desired. □

Corollary 4.4.22. *Let $\varphi : G \to G'$ be a group homomorphism, and let $H \leq G$ be a subgroup. Then $\varphi(H) \leq G'$ is a subgroup.*

Proof. By definition $\varphi(H) = \varphi \circ \iota_H(H) = \mathrm{im}(\varphi \circ \iota_H)$. Thus, as φ and ι_H are group homomorphism, so is $\varphi \circ \iota_H$, thus $\varphi(H) = \mathrm{im}(\varphi \circ \iota_H) \leq G'$ is a subgroup, by Theorem 4.4.21. □

.... End of class 10 (2/4) ....

Definition 4.4.23. Let $\varphi : G \to G'$ be a group homomorphism. We say that φ is an *isomorphism* if there exists a *group homomorphism* $\psi : G' \to G$ such that

$$\varphi \circ \psi = \mathrm{id}_{G'} \text{ and } \psi \circ \varphi = \mathrm{id}_G.$$

When there exists an isomorphism $G \to G'$, we write $G \cong G'$ and say "G is *isomorphic* to G' (as groups)". We will also write

$$G \xrightarrow{\sim} G' \text{ or } G \xrightarrow{\cong} G'$$

to symbolize an isomorphism.

Theorem 4.4.24. *Let $\varphi : G \to G'$ be a group homomorphism. Then φ is an isomorphism if and only if φ is bijective.*

Proof. Suppose that φ is an isomorphism. By definition, there exists a group homomorphism $\psi : G' \to G$ such that

$$\varphi \circ \psi = \mathrm{id}_{G'} \text{ and } \psi \circ \varphi = \mathrm{id}_G.$$

From Proposition 1.2.18, φ is bijective.

Conversely, suppose that φ is bijective. Then there exists an inverse map $\varphi^{-1} : G' \to G$ (by Theorem 1.2.19). It remains to prove that φ^{-1} is a group homomorphism. Let $g', h' \in G'$. Since φ is surjective, there are some $g, h \in G$ such that $g' = \varphi(g)$ and $h'\varphi(h)$. Then

$$\varphi(gh) = \varphi(g)\varphi(h) = g'h'.$$

Hence, by definition,

$$\varphi^{-1}(g'h') = gh = \varphi^{-1}(g')\varphi^{-1}(h'),$$

whence φ^{-1} is a group homomorphism. □

Corollary 4.4.25. *Let $\varphi : G \to G'$ be a group homomorphism.*

(i) *If $\ker\varphi$ is trivial, then φ is an isomorphism of G with its image $\varphi : G \xrightarrow{\sim} \mathrm{im}\varphi$.*
(ii) *If $\varphi : G \to G'$ is surjective and $\ker\varphi$ is trivial, then φ is an isomorphism.*

Proof. Both follow immediately from Theorem 4.4.24. □

Remark 4.4.26. One should think of isomorphic groups are "basically the same". That is, if $G \cong G'$ are isomorphic groups, then roughly speaking, any property of G that can be defined entirely in terms of the group properties is shared by G'. For example, being abelian, being cyclic, and many others.

Theorem 4.4.27. *Let G be a cyclic group. If G is infinite, then $G \cong \mathbb{Z}$. If G is finite and $n = |G|$, then $G \cong \mathbb{Z}/n\mathbb{Z}$.*

Proof. Let G be a cyclic group, generated by $a \in G$.

Suppose first that G is infinite. We have shown in Proposition 4.2.5 that for a cyclic group G with generator a, the map $\varphi : \mathbb{Z} \to G,\ n \mapsto a^n$ is surjective, and it is injective if and only if G is infinite. Since

$$\varphi(n+m) = a^{n+m} = a^n a^m = \varphi(n)\varphi(m),$$

φ is a group homomorohism. Thus, as G is infinite, then φ is an isomorphism $\mathbb{Z} \cong G$.

Now suppose that G is finite, and $|G| = n$. Then $G = \{e = a^0, a, \dots, a^{n-1}\}$, and we can define a map

$$\varphi : \mathbb{Z}/n\mathbb{Z} \to G,\ \overline{k} \mapsto a^k.$$

We first check that φ is well-defined, i.e. does not depend on the choice of $k \mod n$. Indeed, since $a^n = e$, if $k \equiv k' \mod n$ then $k = k' + xn$ for some $x \in \mathbb{Z}$ and

$$\varphi(\overline{k}) = a^k = a^{k'+xn} = a^{k'} a^{xn} = a^{k'} e = a^{k'} = \varphi(\overline{k'}).$$

Thus φ is well-defined. Since

$$\varphi\left(\overline{k} + \overline{k'}\right) = \varphi\left(\overline{k+k'}\right) = a^{k+k'} = a^k a^{k'} = \varphi\left(\overline{k}\right)\varphi\left(\overline{k'}\right),$$

φ is a group homomorphism. Clearly φ is surjective. Since G and $\mathbb{Z}/n\mathbb{Z}$ both have n elements, φ must also be injective. Thus $\varphi : \mathbb{Z}/n\mathbb{Z} \to G$ is an isomorphism. □

Thus we are justified in calling $\mathbb{Z}/n\mathbb{Z}$ *the* finite cyclic group of order n.

Exercise 4.4.28. Prove the following proposition.

Proposition 4.4.29. *Let $\varphi : G \to H$ be a group homomorphism.*

(1) If G is abelian, then the subgroup $\mathrm{im}\varphi \leq H$ *is abelian.*
(2) If G is abelian and $\varphi : G \xrightarrow{\sim} H$ is an isomorphism, then H is abelian.

Remark 4.4.30. We'll see some ways to construct isomorphisms between groups, but it is generally quite difficult in a vacuum, you really need to know something about the structure of the groups to construct an isomorphism.

Conversely, it is usually much easier to show that two groups are *not* isomorphic. To show that two groups are not isomorphic, it suffices to find some property of groups that is preserved under isomorphism, e.g. being abelian, and show that one of the groups under consideration has this property and the other does not.

Proposition 4.4.31. *Let $\varphi : G \xrightarrow{\sim} G'$ be a group isomorphism. Then for all $g \in G$,*

$$\mathrm{ord}(g) = \mathrm{ord}(\varphi(g)).$$

Proof. First, we assume that g has infinite order. We want to show that $\varphi(g)$ also has infinite order. Assume towards a contradiction that $\mathrm{ord}(\varphi(g))$ is finite. Thus, there is some $n \in \mathbb{Z}_{>0}$ such that $\varphi(g)^n = e'$. As φ is a a group homomorphism,

$$\varphi(e) = e' = \varphi(g)^n = \varphi(g^n).$$

As φ is an isomorphism, it is in particular injective, thus $g^n = e$, hence g has finite order, which is a contradiction. Thus if g has infinite order, $\varphi(g)$ has infinite order as well.

Now suppose that $\operatorname{ord}(g) = n$ is finite. Then

$$e' = \varphi(e) = \varphi(g^n) = \varphi(g)^n,$$

and so by Corollary 4.2.11, $\operatorname{ord}(\varphi(g)) \mid n$. If $d = \operatorname{ord}(\varphi(g)) < n$, then

$$\varphi(e) = e' = \varphi(g)^d = \varphi(g^d)$$

and as φ is injective, $g^d = e$, which is a contradiction as $d < n = \operatorname{ord}(g)$. Therefore

$$\operatorname{ord}(\varphi(g)) = \operatorname{ord}(g). \qquad \square$$

This is useful for showing that some groups are not isomorphic.

Proposition 4.4.32. *The groups* $\mathbb{Z}/4\mathbb{Z}$ *and* $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ *are not isomorphic.*

Proof. We first make some observations. If $(a, b) \in \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, then

$$(a, b) + (a, b) = (a + a, b + b) = (2a, 2b) = (0, 0) = e.$$

Thus every element of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ has order 1 or 2. However, the element $\overline{1} \in \mathbb{Z}/4\mathbb{Z}$ has order 4.

Thus, suppose for contradiction that $\varphi : \mathbb{Z}/4\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is an isomorphism. Then $\varphi(\overline{1}) \in \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, and thus has order 1 or 2, contradicting Proposition 4.4.31. Therefore there is no such isomorphism. □

Example 4.4.33. The groups $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/6\mathbb{Z}$ are isomorphic. Indeed, we can construct an isomorphism by hand. Define

$$\varphi : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}, \quad \overline{a} \mapsto (a \mod 2, \quad a \mod 3).$$

We first need to check that this is well-defined! But this is clear, since if $a \equiv a' \mod 6$, then $6 \mid (a - a')$, hence $2 \mid (a - a')$ and $3 \mid (a - a')$, and so $a \equiv a' \mod 2$ and $a \equiv a' \mod 3$. Then we check that

$$\begin{aligned}
\varphi(\overline{a} + \overline{b}) &= \varphi(\overline{a + b}) \\
&= (a + b \mod 2, \quad a + b \mod 3) \\
&= (a \mod 2 \ + \ b \mod 2, \quad a \mod 3 \ + \ b \mod 3) \\
&= (a \mod 2, \quad a \mod 3) \ + \ (b \mod 2, \quad b \mod 3) \\
&= \varphi(\overline{a}) + \varphi(\overline{b}).
\end{aligned}$$

Thus φ is a group homomorphism. Now suppose that $\varphi(\overline{a}) = (0, 0) = e$. Then $a \equiv 0 \mod 2$ and $a \equiv 0 \mod 3$. Since $\gcd(2, 3) = 1$, the Chinese Remainder Theorem (Theorem 3.1.19) shows that $a \equiv 0 \mod 6$, i.e. $\overline{a} = \overline{0} \in \mathbb{Z}/6\mathbb{Z}$. Thus φ is injective. Since both $\mathbb{Z}/6\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ have order 6 (have 6 elements), φ must also be surjective. Hence

$$\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}.$$

This is an example of a much more general fact. To make the proof easier, we'll set up some notation.

Definition 4.4.34. Let X and $Y_1, \dots, Y_n$ be sets. A map

$$f : X \to Y_1 \times \cdots Y_n$$

into the product is given by n maps $f_i : X \to Y_i$ such that

$$f(x) = (f_1(x), f_2(x), \dots, f_n(x)) \text{ for all } x \in X.$$

The maps $f_i : X \to Y_i$ are called the *coordinate maps* of f.

Proposition 4.4.35. *Let G and H be groups, and suppose that*
$$H = H_1 \times \cdots \times H_n$$
is a direct product of groups H_i. Let $\varphi : G \to H$, and let $\varphi_i : G \to H_i$ be the i-th coordinate map. Then φ is a group homomorphism if and only if φ_i are all group homomorphisms.

Proof. The proof follows directly from the fact that the group structure on the product is defined componentwise. □

Theorem 4.4.36 (Chinese Remainder Theorem)**.**
$$\mathbb{Z}/n_1n_2\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \text{ if and only if } \gcd(n_1, n_2) = 1.$$

Proof. If $\gcd(n_1, n_2) = 1$, then for $(\overline{a}, \overline{b}) \in \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$, the classical Chinese Remainder Theorem (Theorem 3.1.19) shows that we can find $x \in \mathbb{Z}$ such that
$$x \equiv a \mod n_1,$$
$$x \equiv b \mod n_2.$$
And moreover, the choice of x is unique modulo n_1n_2. Let
$$\varphi_1 : \mathbb{Z}/n_1n_2\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \text{ and } \varphi_2 : \mathbb{Z}/n_1n_2\mathbb{Z} \to \mathbb{Z}/n_2\mathbb{Z}$$
be the group homomorphisms reducing modulo n_i (checking that φ_1 and φ_2 are well-defined and group homomorphisms is the same as in Example 4.4.33). Then we obtain a group homomorphism
$$\varphi : \mathbb{Z}/n_1n_2\mathbb{Z} \to \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}, \; n \mapsto (\varphi_1(n), \varphi_2(n)) = (\overline{n}, \overline{n}).$$
The classical Chinese Remainder Theorem (Theorem 3.1.19) shows that φ is bijective.

Conversely, suppose that $\mathbb{Z}/n_1n_2\mathbb{Z} \cong \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$. Now, $\mathbb{Z}/n_1n_2\mathbb{Z}$ has an element of order n_1n_2 (namely $\overline{1}$), thus by Proposition 4.4.31, there is an element $(a, b) \in \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$ of order n_1n_2. Note that for some $x \in \mathbb{Z}$, if $n_1 \mid x$ and $n_2 \mid x$, then
$$x(a, b) = (xa, xb) = (0, 0).$$
Towards a contradiction, suppose that $\gcd(n_1, n_2) = d > 1$. By Lemma 3.1.21, the least common multiple of n_1 and n_2 is $\frac{n_1n_2}{d} < n_1n_2$. Hence every element of $\mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z}$ has order at most $\frac{n_1n_2}{d} < n_1n_2$, which is a contradiction. □

More generally, let's see when a group is isomorphic to a direct product.

We've seen that if $H_1, H_2 \leq G$ are subgroups, then $H_1 \cap H_2 \leq G$ is a subgroup. Can we somehow multiply subgroups?

Definition 4.4.37. Let G be a group and let $X, Y \leq G$ be subsets. Define the subset
$$XY := \{xy \in G \mid x \in X, \; y \in Y\} \subseteq G.$$
We call $XY \subseteq G$ the *product of the subsets X and Y*. It is easy to check that if $X, Y, Z \subseteq G$ are subsets, then $(XY)Z = X(YZ)$.

Exercise 4.4.38. Let $H \leq G$ be a subgroup, and let $X \subseteq H$ be a non-empty subset. Show that $XH = H$.

Remark 4.4.39. In general, H_1H_2 is not a subgroup. We'll see when it is a subgroup shortly.

.... End of Class 11 (2/6) ....

Theorem 4.4.40. *Let G be a group. Then $G \cong G_1 \times G_2$ if and only if then there are subgroups $G_1 \cong H_1 \leq G$ and $G_2 \cong H_2 \leq G$ such that*

- $H_1 \cap H_2 = \{e\}$,

- *$G = H_1H_2$, and*
- *for all $h_1 \in H_1$ and $h_2 \in H_2$, $h_1h_2 = h_2h_1$.*

Proof. Suppose that $\varphi : G_1 \times G_2 \xrightarrow{\sim} G$ is an isomorphism. We have the coordinate inclusion group homomorphisms

$$\iota_1 : G_1 \to G_1 \times G_2,\ g_1 \mapsto (g_1, e_2)$$
$$\iota_2 : G_2 \to G_1 \times G_2,\ g_2 \mapsto (e_1, g_2).$$

Define $H_1 = \operatorname{im}(\varphi \circ \iota_1)$ and $H_2 = \operatorname{im}(\varphi \circ \iota_2)$, which are both subgroups by Theorem 4.4.21. Moreover, since ι_1, ι_2, and φ are all injective, $\varphi \circ \iota_1$ and $\varphi \circ \iota_2$ are isomorphisms of G_i with H_i by Corollary 4.4.25, thus $H_1 \cong G_1$ and $H_2 \cong G_2$.

Suppose $g \in H_1 \cap H_2$, then $g = \varphi((g_1, e_2)) = \varphi((e_1, g_2))$. As φ is injective,

$$(g_1, e_2) = (e_1, g_2),$$

so $g_1 = e_1$ and $g_2 = e_2$, so $g = \varphi((e_1, e_2)) = e$. Thus $H_1 \cap H_2 = \{e\}$.

Let $g \in G$. Then $g = \varphi(g_1, g_2)$ for some $g_i \in G_i$. Since

$$\varphi((g_1, g_2)) = \varphi((g_1, e_2)(e_1, g_2)) = \varphi((g_1, e_2))\varphi((e_1, g_2)) = \underbrace{\varphi(\iota_1(g_1))}_{\in H_1}\underbrace{\varphi(\iota_2(g_2))}_{\in H_2},$$

we have $G = H_1H_2$, as desired.

Finally, since let $h_1 = \varphi((g_1, e_2)) \in H_1$ and $h_2 = \varphi((e_1, g_2)) \in H_2$. Then

$$\begin{aligned} h_1h_2 &= \varphi((g_1, e_2))\varphi((e_1, g_2)) \\ &= \varphi((g_1, e_2)(e_1, g_2)) \\ &= \varphi((g_1, g_2)) \\ &= \varphi((e_1, g_2)(g_1, e_2)) \\ &= \varphi((e_1, g_2))\varphi((g_1, e_2)) \\ &= h_2h_1, \end{aligned}$$

as desired.

Conversely, suppose that $H_1, H_2 \leq G$ are subgroups such that $H_1H_2 = G$, $H_1 \cap H_2 = \{e\}$, and for all $h_1 \in H_1$ and $h_2 \in H_2$, $h_1h_2 = h_2h_1$. Define a map $\varphi : H_1 \times H_2 \to G, (h_1, h_2) \mapsto h_1h_2$. Since

$$\varphi((h_1, h_2)(h_1', h_2')) = \varphi((h_1h_1', h_2h_2')) = h_1\underbrace{h_1'h_2}h_2' = h_1\underbrace{h_2h_1'}h_2' = \varphi((h_1, h_2))\varphi((h_1', h_2')),$$

thus φ is a group homomorphism.

Since $H_1H_2 = G$, φ is surjective.

Suppose $(h_1, h_2) \in \ker \varphi$. Then $h_1h_2 = e$, hence $h_1 = eh_2^{-1} \in H_2$ and $h_2 = h_1^{-1}e \in H_1$. Thus $h_1, h_2 \in H_1 \cap H_2 = \{e\}$, and so $h_1 = h_2 = e$. Thus φ is injective.

Hence $\varphi : H_1 \times H_2 \to G$ is a bijective group homomorphism, and hence $H_1 \times H_2 \cong G$. □

RH: The following examples will be skipped in class, but are quite useful and important examples. I encourage you to look them over when preparing for the midterm.

Let's see some more examples of group homomorphisms.

Definition 4.4.41. Let G be a group, an *automorphism* of G is a group isomorphism $\varphi : G \xrightarrow{\sim} G$. We denote by $\operatorname{Aut}(G)$ the set of automorphisms of G.

Exercise 4.4.42. Prove the following proposition.

Proposition 4.4.43. *Let G be a group, then* $\operatorname{Aut}(G)$ *is a subgroup of* $\operatorname{Perm}(G)$*, the group of bijections of the set G, where the group law is composition of maps.*

Definition 4.4.44. Let G be a group. For $g \in G$, let

$$L_g : G \to G,\ h \mapsto gh.$$

We call L_g *left translation by g.*

Exercise 4.4.45. Prove the following proposition.

Proposition 4.4.46. *The map*

$$L : G \to \mathrm{Perm}(G),\ g \mapsto L_g$$

is an injective group homomorphism.

That is, show that $L_{gh} = L_g \circ L_h$, and that the map L is injective.

In addition, find an example of a group G and some $g \in G$ such that L_g is not a group homomorphism. Hint: consider S_3.

Definition 4.4.47. Let G be a group. For $g \in G$, let

$$\mathfrak{c}_g : G \to G,\ h \mapsto ghg^{-1}.$$

We call $\mathfrak{c}_g$ *conjugation by g.*

Exercise 4.4.48. Prove the following proposition.

Proposition 4.4.49. *The map*

$$\mathfrak{c} : G \to \mathrm{Aut}(G),\ g \mapsto \mathfrak{c}_g$$

is an injective group homomorphism.

That is, show that $\mathfrak{c}_{gh} = \mathfrak{c}_g \circ \mathfrak{c}_h$, and that the map $\mathfrak{c}$ is injective.

Exercise 4.4.50. Prove the following proposition.

Proposition 4.4.51. *Let G be a group. Then G is abelian if and only if the map*

$$\varphi : G \to G,\ g \mapsto g^{-1}$$

is a group homomorphism.

Exercise 4.4.52. Prove the following proposition.

Proposition 4.4.53. *Let G be a cyclic group, and $\varphi : G \to H$ be a group homomorphism. Show that* $\mathrm{im}\varphi$ *is cyclic.*

4.5. **Cosets.** We'll see how a subgroup can be used to "break up" a group into smaller pieces.

Definition 4.5.1. Let G be a group, and $H \leq G$ be a subgroup. A *left coset* of H in G with representative $g \in G$ is a subset of the form

$$gH := \{gh : h \in H\} \subseteq G.$$

A *right coset* is defined similarly as

$$Hg \coloneqq \{hg : h \in H\} \subseteq G.$$

An element of gH is called a *coset representative* of the coset gH.

Example 4.5.2. Let $G = \mathbb{Z}$ and $H = 3\mathbb{Z}$, the subgroup of integers divisible by 3. Then the cosets of $3\mathbb{Z}$ are given by

$$\begin{gathered} 0 + 3\mathbb{Z} \\ 1 + 3\mathbb{Z} \\ 2 + 3\mathbb{Z}. \end{gathered}$$

Remark 4.5.3. Note that there is ambiguity in how we represent a coset, as two elements $g_1, g_2 \in G$ can represent the same coset. For example,

$$1 + 3\mathbb{Z} = 4 + 3\mathbb{Z} = 7 + 3\mathbb{Z} = ... = (3n + 1) + 3\mathbb{Z}.$$

The representatives for each coset of $3\mathbb{Z}$ have the same remainders $\pmod 3$.

When dealing with cosets, you should take care to check whether the things you want to say depends on the coset representative!

Proposition 4.5.4. *Let $H \leq G$ be a subgroup, and $g_1, g_2 \in G$. the following are equivalent*

(1) $g_1H = g_2H$,
(2) $Hg_1^{-1} = Hg_2^{-1}$,
(3) $g_1H \subseteq g_2H$,
(4) $g_2 \in g_1H$,
(5) $g_1^{-1}g_2 \in H$

Proof. We will show that (1) $\iff$ (2) $\implies$ (3) $\implies$ (4) $\implies$ (5) $\implies$ (1).

(1) $\implies$ (2): Assume that $g_1H = g_2H$, and let $g \in Hg_1^{-1}$, so $g = hg_1^{-1}$ for some $h \in H$. Then $g^{-1} = g_1h^{-1} \in g_1H = g_2H$, so $g^{-1} = g_2k$ for some $k \in H$, and so

$$g = (g^{-1})^{-1} = (g_2k)^{-1} = k^{-1}g_2^{-1} \in Hg_2^{-1}.$$

Thus $Hg_1^{-1} \subseteq Hg_2^{-1}$. The other containment is similar.

(2) $\implies$ (1): Assume that $Hg_1^{-1} = Hg_2^{-1}$, and let $g \in g_1H$. Then $g = g_1h$ for some $h \in H$. Then $g^{-1} = h^{-1}g_1^{-1} \in Hg_1^{-1} = Hg_2^{-1}$, so $g^{-1} = kg_2^{-1}$ for some $k \in H$. And as before $g = g_2k^{-1} \in g_2H$. Thus $g_1H \subseteq g_2H$. The other containment is similar.

(2) $\implies$ (3): We have shown that (2) $\implies$ (1), and clearly (1) $\implies$ (3).

(3) $\implies$ (4): Assume that $g_1H \subseteq g_2H$. In particular, this means that $g_1 = g_1e \in g_1H \subseteq g_2H$. Thus there is some $h \in H$ such that $g_1 = g_2h$, and multiplying by h^{-1} on the right gives $g_2h^{-1} = g_1 \in g_1H$.

(4) $\implies$ (5): By assumption, $g_2 = g_1h$ for some $h \in H$. Multiplying by g_1^{-1} on left yields $g_1^{-1}g_2 = g_1^{-1}g_1h = eh = h \in H$.

(5) $\implies$ (1): Assume that $g_1^{-1}g_2 \in H$. Then

$$\begin{aligned} g_1H &= g_1(g_1^{-1}g_2H) \\ &= g_1g_1^{-1}g_2H \\ &= g_2H. \end{aligned}$$

□

The following is a very useful consequence.

Proposition 4.5.5. *Let aH and bH be two cosets of H in G. Then either $aH = bH$ or $aH \cap bH = \emptyset$. Thus the left cosets of H partition G. The associated relation is $g_1 \sim g_2$ if and only if $g_1^{-1}g_2 \in H$. The same is true for right cosets, the associated relation is $g_1 \sim g_2$ if and only if $g_1g_2^{-1} \in H$.*

Proof. We will do the proof for left cosets only, the proof for right cosets being similar.

We want to show that for two left cosets aH, bH, either $aH \cap bH = \emptyset$, or $aH = bH$.

We will show that if $aH \cap bH \neq \emptyset$, then they are equal. Suppose $g \in aH \cap bH$. By the definition of cosets, we have $g = ah = bh'$ for some $h, h' \in H$. Hence $ah = bh'$ and so $a = bh'h^{-1}$. But since $h'h^{-1} \in H$, we have $aH = b(h'h^{-1})H = b(h'h^{-1}H) = bH$, as desired.

The associated equivalence relation is given by $g_1 \sim g_2$ if and only if there is some $a \in G$ such that $g_1, g_2 \in aH$. By Proposition 4.5.4, this is equivalent to $g_1H = aH = g_2H$. □

Definition 4.5.6. For a group G and a subgroup $H \leq G$, the set of left cosets is denoted ${}^{G}\!/_{H}$ and the set of right cosets is denoted by ${}_{H}\backslash^{G}$. A subset $\{g_i\} \subset G$ is a set of *(left) coset representatives* if the cosets $\{g_iH\}$ are all distinct and $G = \bigsqcup_i g_iH$.

Example 4.5.7. The cosets of $3\mathbb{Z}$ partition $\mathbb{Z}$, namely

$$\mathbb{Z} = 0 + 3\mathbb{Z} \cup 1 + 3\mathbb{Z} \cup 2 + 3\mathbb{Z},$$

which is just the grouping of integers by their remainders after division by 3.

Proposition 4.5.8. *Let G be a group, and $H \leq G$ a subgroup. There is a bijection between ${}^{G}\!/_{H}$ and ${}_{H}\backslash^{G}$.*

Proof. We define a map

$$\mathrm{inv} : {}^{G}\!/_{H} \to {}_{H}\backslash^{G},\ gH \mapsto Hg^{-1}.$$

We must first check that inv is well-defined, that it does not depend on the coset representative!

Thus assume that $g_1H = g_2H$. By Proposition 4.5.4, we have

$$\mathrm{inv}(g_1H) = Hg_1^{-1} = Hg_2^{-1} = \mathrm{inv}(g_2H),$$

thus inv is well-defined.

Surjectivity is straightforward, let $Hg \in {}_{H}\backslash^{G}$, then

$$\mathrm{inv}(g^{-1}H) = Hg.$$

For injectivity, suppose that $\mathrm{inv}(g_1H) = \mathrm{inv}(g_2H)$, i.e., $Hg_1^{-1} = Hg_2^{-1}$. By Proposition 4.5.4, $g_1H = g_2H$, and so inv is injective. □

Example 4.5.9. For example, since the cosets of $3\mathbb{Z} \in \mathbb{Z}$ are

$$\{0 + 3\mathbb{Z},\ 1 + 3\mathbb{Z},\ 2 + 3\mathbb{Z}\},$$

we see that the set of left cosets of $3\mathbb{Z} \in \mathbb{Z}$ is

$$\mathbb{Z}/3\mathbb{Z} = \{0 + 3\mathbb{Z},\ 1 + 3\mathbb{Z},\ 2 + 3\mathbb{Z}\}.$$

This might get a little confusing, because we've already used the notation $\mathbb{Z}/3\mathbb{Z}$ to be the cyclic group of order 3, namely

$$\mathbb{Z}/3\mathbb{Z} = \{\overline{0}, \overline{1}, \overline{2}\}.$$

We'll see soon why these are really the same, when we talk about *quotient groups.* In particular, you could imagine adding the cosets of $3\mathbb{Z}$ much like you add the elements $\overline{0}$, $\overline{1}$, $\overline{2}$, and make the set of left cosets into a group! This works because $\mathbb{Z}$ is abelian, in general the set of left cosets might not always form a group. But we'll see soon when it does!

Definition 4.5.10. The cardinality of ${}^{G}\!/_{H}$ is the number of (left) cosets of H in G, denoted by $(G : H)$, called the *index* of H in G. The index of the trivial group $(G : 1)$ is the *order* of G.

Thus we see that $(\mathbb{Z} : 3\mathbb{Z}) = 3$, even though both $\mathbb{Z}$ and $3\mathbb{Z}$ are infinite.

Proposition 4.5.11. *Let G be a group, and H a subgroup. For all $g \in G$, there is a bijection between the cosets $eH = H$ and gH.*

Proof. We define $L_g : H \to gH,\ h \mapsto gh$, and note that L_g has inverse $L_{g^{-1}}$. □

.... End of Class 12 (2/9) ....

4.5.1. *Lagrange's Theorem.* We are now ready to prove Lagrange's Theorem.

Theorem 4.5.12 (Lagrange's theorem)**.** *Let G be a finite group and H a subgroup. Then*

$$|G| = (G : H)|H|.$$

In particular, the order of the subgroup H divides the order of G.

Proof. All cosets of H have cardinality $|H|$, as all cosets are bijective (Proposition 4.5.11). Thus, as the cosets of H partition G, and there are $(G : H)$ cosets, we obtain

$$|G| = (G : H)|H|. \qquad \square$$

We can also prove a more general form of Lagrange's theorem.

Theorem 4.5.13 (General Lagrange's Theorem)**.** *Let G be a group, let $H \leq G$ be a subgroup of G, and let $K \leq H$ be a subgroup of H. Let $\{x_i\}$ be a set of (left) coset representatives for K in H and let $\{y_j\}$ be a set of (left) coset representative for H in G, then $\{y_j x_i\}$ is a set of (left) coset representatives for K in G. In particular,*

$$(G : K) = (G : H) \cdot (H : K).$$

Proof. As cosets partition a group into disjoint pieces, we have

$$H = \bigsqcup_i x_i K,$$

and

$$G = \bigsqcup_j y_j H.$$

Hence we have

$$G = \bigsqcup_j y_j \left(\bigsqcup_i x_i K \right) = \bigcup_{j,i} y_j x_i K,$$

and we want to show that this last union is also disjoint, and so $\{y_j x_i\}$ will be a set of (left) coset representatives for K in G.

Suppose that

$$y_j x_i K = y_n x_m K,$$

we want to show that $y_j = y_n$ and $x_i = x_m$. By multiplying by H on the right (noting that since $x_i K \subset H$) we have

$$y_j H = y_j(x_i K)H = y_n(x_m K)H = y_n H,$$

and thus $y_j = y_n$, as they represent the same coset and $\{y_j\}$ was assumed to be a list of distinct coset representatives.

We now have

$$y_j x_i K = y_n x_m K = y_j x_m K,$$

and multiplying on the left by y_j^{-1}, we have

$$x_i K = x_m K,$$

and thus $x_i = x_m$, as desired.

The equation now follows from the fact that $(G : K)$ is the number of (left) cosets of K in G, which is the number of (left) coset representatives. That is,

$$(G : K) = \#\{y_j x_i\} = (\#\{y_j\}) \cdot (\#\{x_i\}) = (G : H) \cdot (H : K). \qquad \square$$

Remark 4.5.14. We note that the formula

$$(G : K) = (G : H) \cdot (H : K)$$

of the General Lagrange's Theorem works for infinite and finite quantities. Namely, if two of the indices are finite, then so is the third, and the formula holds.

Remark 4.5.15. The usual Lagrange's Theorem is obtained from the General Lagrange's Theorem by taking $K = \{e\}$.

Corollary 4.5.16. *Let G be a finite group. The order of any element of G divides the order of G.*

Proof. Let $g \in G$, and consider the subgroup $\langle g \rangle \leq G$. Then $|\langle g \rangle| = \operatorname{ord}(g)$, and Lagrange's Theorem gives

$$|G| = (G : \langle g \rangle) \operatorname{ord}(g). \qquad \square$$

Exercise 4.5.17. Prove the following proposition.

Proposition 4.5.18. *Let G be a finite group and $g \in G$. Then $g^{|G|} = e$.*

As a result, we can identify some groups.

Theorem 4.5.19. *Let p be a prime number, and G a finite group of order p. Then G is cyclic. Hence G isomorphic to $\mathbb{Z}/p\mathbb{Z}$.*

Proof. We show that G is cyclic, the remaining statement follows from Theorem 4.4.27.

Suppose G is a group of order p. Since $p \geq 2$, there is some $e \neq g \in G$. Consider the subgroup $\langle g \rangle \leq G$. Since $g \neq e$, $\{1, g\} \subseteq \langle g \rangle$, thus $|\langle g \rangle| \geq 2$. By Lagrange's Theorem,

$$p = |G| = (G : \langle g \rangle) \cdot |\langle g \rangle|.$$

Since p is prime, and $|\langle g \rangle| \geq 2$, we have $(G : \langle g \rangle) = 1$, and so $G = \langle g \rangle$. □

RH: we'll skip the example of $|S_n| = n!$ now, but we'll prove it either on (2/13) or when we talk about S_n in more detail. We're skipping it now just for time (so you see all the concepts on Homework 3 as early as possible).

We can also find the order of some groups.

Recall that the symmetric group, S_n, is

$$S_n = \{\text{bijections } \{1, \ldots, n\} \to \{1, \ldots, n\}\},$$

and element of S_n are called permutations.

Theorem 4.5.20. *Let S_n be the symmetric group. Then $|S_n| = n! = n(n-1) \cdots 2 \cdot 1$.*

Proof. We proceed by induction. Clearly $|S_1| = |\{\text{id}\}| = 1$.

Let H be the subset of S_n such that $\sigma(n) = n$. Then H is a subgroup, since all permutations in H preserve n, and so do all of their inverses. We may view H as the subgroup S_{n-1}.

We wish to describe the cosets of H. For each integer i with $1 \leq i \leq n$, let $\tau_i \in S_n$ be the permutation swapping $i \leftrightarrow n$ and leaving all other elements of $\{1, \ldots, n\}$ fixed. Note that $\tau_n = \text{id}$.

We claim that the cosets

$$\tau_1 H, \ldots, \tau_n H$$

are distinct, and are all the cosets of H in S_n.

To see this, let $\sigma \in S_n$, and suppose that $\sigma(n) = i$. Then

$$\tau_i^{-1} \sigma(n) = \tau_i^{-1}(n) = n,$$

and thus $\tau_i^{-1} \sigma \in H$, so $\sigma \in \tau_i H$. Thus every element of G lies in some coset $\tau_i H$, and so $\tau_1 H, \ldots, \tau_n H$ are all of the cosets of H. To see that these are all distinct, we note that if $i \neq j$, then for any $\sigma \in H$,

$$\tau_i \sigma(n) = \tau_i(n) = i,$$

and

$$\tau_j \sigma(n) = \tau_j(n) = j.$$

Hence every element of $\tau_i H$ sends n to i and every element of $\tau_j H$ sends n to j. Thus if $i \neq j$, then $\tau_i H$ and $\tau_j H$ cannot have any element in common.

Thus, as $H \cong S_{n-1}$, from Lagrange's Theorem,

$$|S_n| = (S_n : H)|H| = n \cdot |S_{n-1}|,$$

and by induction we have

$$|S_n| = n \cdot (n-1)! = n(n-1)\cdots 1. \qquad \square$$

Cosets also have a nice relationship with homomorphisms.

Theorem 4.5.21. *Let $\varphi : G \to G'$ be a group homomorphism. Let $H = \ker\varphi$ and let*

$$a' = \varphi(a) \in \mathrm{im}\varphi \leq G'.$$

Then

$$\varphi^{-1}(a') = \{g \in G \mid \varphi(g) = a'\} = aH.$$

Similarly, $\varphi^{-1}(a') = Ha$.

Proof. Let $g \in aH$, so $g = ah$ for some $h \in H = \ker\varphi$. Then

$$\varphi(g) = \varphi(ah) = \varphi(a)\varphi(h) = \varphi(a)e' = \varphi(a),$$

whence $g \in \varphi^{-1}(a')$.

Conversely, suppose that $g \in G$ and $\varphi(g) = a'$. Then

$$\varphi(a^{-1}g) = \varphi(a)^{-1}\varphi(g) = (a')^{-1}a' = e'.$$

Thus $a^{-1}g \in \ker\varphi = H$, and so $a^{-1}g = h$ for some $h \in H$, whence $g \in aH$.

Similarly, for $g \in Ha$, $g = ha$ and so

$$\varphi(ha) = \varphi(h)\varphi(a) = e'a' = a',$$

so $g \in \varphi^{-1}(a')$. And conversely, of $g \in \varphi^{-1}(a')$, then $\varphi(ga^{-1}) = e'$, so $ga^{-1} \in H$, and thus $g \in Ha$. □

As a corollary, kernels of group homomorphisms are very important. They have very nice cosets.

Theorem 4.5.22. *Let $\varphi : G \to G'$ be a group homomorphism, and let $H = \ker\varphi$. Then for every $g \in G$, $gH = Hg$.*

Proof. Let $g \in G$, and let $g' = \varphi(g)$. Then $gH = \varphi^{-1}(g') = Hg$. □

4.6. **Normal Subgroups.**

Definition 4.6.1. Let G be a group, and $H \leq G$ a subgroup. we say that H is *normal* if for all $g \in G$,

$$gH = Hg.$$

We write $H \trianglelefteq G$ or $H \triangleleft G$ if we want to emphasize that H is a normal subgroup of G.

Remark 4.6.2. Note that H being a normal subgroup is **NOT** the same as saying that $ghg^{-1} = h$ for all $h \in H$ when G is not abelian!

Exercise 4.6.3. Prove the following proposition.

Proposition 4.6.4. *Let G be a group and $H \leq G$ a subgroup. Then the following are equivalent*

1. *$H \trianglelefteq G$ is normal,*
2. *for all $g \in G$, $g^{-1}Hg \subseteq H$,*
3. *for all $g \in G$, $g^{-1}Hg = H$.*

Proposition 4.6.5. *Let G be an abelian group, then every subgroup is normal.*

Proof. Let $H \leq G$ be a subgroup. Then for any $g \in G$ and any $h \in H$, $gh = hg$, so $ghg^{-1} = h$, hence $gHg^{-1} = H$. □

Exercise 4.6.6. Show that S_3 has a unique normal proper non-trivial subgroup. That is, there is only one subgroup $A_3 \lhd S_3$ (and $A_3 \neq \{e\}$, S_3). Hint: Writing the elements of S_3 as matrices (as in Example 4.1.3(5)) or as lists of numbers (as in Example 4.1.10) will make it much easier to do computations.

4.6.1. *Quotient Groups.* Normal subgroups are very special, because when $H \trianglelefteq G$ is normal, we can make the set of cosets ${}^{G}/_{H}$ into a group.

Theorem 4.6.7. *Let G be a group and let $H \trianglelefteq G$ be a normal subgroup. Then for two cosets $aH, bH \in {}^{G}/_{H}$, $(aH)(bH) = abH$, the binary operation*

$$ {}^{G}/_{H} \times {}^{G}/_{H} \to {}^{G}/_{H}, \ (aH, bH) \mapsto abH $$

is well-defined and defines a group structure on ${}^{G}/_{H}$, with identity H.

Moreover, the map

$$ \pi : G \to {}^{G}/_{H}, \ g \mapsto gH $$

is a surjective group homomorphism with kernel $\ker \pi = H$.

Proof. Since $H \trianglelefteq G$ is normal, we have

$$ (aH)(bH) = a(Hb)H = abHH = abH. $$

We must make sure that this is a well-defined operation that does not depend on the choice of coset representative. So suppose that $a_1H = a_2H$ and $b_1H = b_2H$, then

$$ \begin{aligned} a_1b_1H &= a_1Hb_1 \\ &= a_2Hb_1 \\ &= a_2b_1H \\ &= a_2b_2H. \end{aligned} $$

Thus the binary operation is well-defined.

We now show that ${}^{G}/_{H}$ is a group. Clearly multiplication of cosets is associative, and for any $g \in G$, $gHeH = gHH = gH$, thus the coset H is an identity. Finally, a straightforward computation shows that $(gH)^{-1} = g^{-1}H$.

The fact that the map $\pi : G \to {}^{G}/_{H}$ is a group homomorphism is just

$$ \pi(g_1g_2) = g_1g_2H = g_1g_2HH = g_1Hg_2H = \pi(g_1)\pi(g_2), $$

and π is clearly surjective.

Lastly, we show that $H = \ker \pi$. Let $h \in H$, then $\pi(h) = hH = H$, which is the identity of ${}^{G}/_{H}$, thus $H \subseteq \ker \pi$. Conversely, if $g \in G$ such that $gH = H$, then $g = ge \in gH = H$, and thus $\ker \pi \subseteq H$. Therefore $H = \ker \pi$, as desired. □

Definition 4.6.8. For a normal subgroup $H \trianglelefteq G$, we call the group homomorphism $\pi : G \to G/H$ the *quotient homomorphism*, or the *canonical homomorphism*, and the group G/H the *quotient group* of G by H. We also call ${}^{G}/_{H}$ "G modulo H" or "G mod H".

.... End of Class 13 (2/11) ....

Theorem 4.6.9. *Let G be a group and $H \leq G$ be a subgroup. Then H is normal if and only if H is the kernel of a group homomorphism.*

Proof. We've seen in Theorem 4.5.22 that the kernel of a group homomorphism is normal. Conversely, if H is normal, then Theorem 4.6.7 shows that $H = \ker \pi$ for the quotient homomorphism $\pi : G \to G/H$. □

Example 4.6.10. Consider the subgroup $n\mathbb{Z} \trianglelefteq \mathbb{Z}$. The quotient group is $\mathbb{Z}/n\mathbb{Z}$. Indeed, the cosets of $n\mathbb{Z}$ are

$$\{k + n\mathbb{Z} \mid 0 \leq k \leq n-1\},$$

and the group operation is given by

$$(k + n\mathbb{Z}) + (k' + n\mathbb{Z}) = (k + k') + n\mathbb{Z},$$

which is exactly the group operation on $\mathbb{Z}/n\mathbb{Z}$.

Example 4.6.11. We've seen that $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is a group homomorphism. The kernel is

$$\mathrm{SL}_n(\mathbb{R}) = \ker \det = \{M \in \mathrm{GL}_n(\mathbb{R}) \mid \det(M) = 1\}.$$

Thus $\mathrm{SL}_n(\mathbb{R}) \triangleleft \mathrm{GL}_n(\mathbb{R})$ is a normal subgroup. The cosets of $\mathrm{SL}_n(\mathbb{R})$ are just $c\,\mathrm{SL}_n(\mathbb{R})$ for $c \in \mathbb{R}^\times$, thus the quotient

$$\mathrm{GL}_n(\mathbb{R}) \big/ \mathrm{SL}_n(\mathbb{R}) \cong \mathbb{R}^\times.$$

Example 4.6.12. You've seen that S_3 has a unique non-trivial proper normal subgroup

$$A_3 = \{(1), (123), (132)\} \triangleleft S_3.$$

And thus S_3/A_3 is a group of order $\left|S_3/A_3\right| = (S_3 : A_3) = \frac{|S_3|}{|A_3|} = \frac{6}{3} = 2$. Hence

$$S_3/A_3 \cong \mathbb{Z}/2\mathbb{Z}.$$

Exercise 4.6.13. Show that $A_3 \cong \mathbb{Z}/3\mathbb{Z}$. However, show that $S_3 \ncong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Conclude that even if $H \trianglelefteq G$ is a normal subgroup, then G is not always isomorphic to $H \times G/H$ as groups.

Remark 4.6.14. Observe that if $\varphi : G \to G'$ is a group homomorphism, and $H \subseteq \ker \varphi$, then for any $g \in G$ and any $h \in H$,

$$\varphi(gh) = \varphi(g)\varphi(h) = \varphi(g)e' = \varphi(g).$$

Thus φ has the same value for all $x \in gH$, which we could write as

$$\varphi(gH) = \varphi(g).$$

In effect, when H is a subgroup, we could define a map $G/H \to G'$ by the formula $gH \mapsto \varphi(g)$, and when H is normal this might even be a group homomorphism. This is no accident.

Theorem 4.6.15 (Maps from quotients, Universal Property of Quotient Groups)**.** *Let $H \trianglelefteq G$ be a normal subgroup with quotient map $\pi : G \to G/H$, and $\varphi : G \to K$ be a group homomorphism. If $H \subseteq \ker \varphi$, then there is a unique group homomorphism $\overline{\varphi} : G/H \to K$ such that $\overline{\varphi} \circ \pi = \varphi$.*

We symbolize this as saying that the diagram

$$\begin{tikzcd}
G \arrow[d, "\pi"'] \arrow[dr, "\varphi"] & \\
G/H \arrow[r, dotted, "\exists!\ \overline{\varphi}"'] & K
\end{tikzcd}$$

commutes, i.e. $\overline{\varphi} \circ \pi = \varphi$ (following the arrows in either direction is the same). We say the homomorphism $\overline{\varphi}$ is *induced by* φ.

Proof of Theorem 4.6.15. We define $\overline{\varphi} : {}^{G}/_{H} \to K$ by

$$\overline{\varphi}(gH) = \varphi(g).$$

We must check that this is well-defined. So suppose $gH = g'H$. Then $g = g'h$ for some $h \in H$, and so

$$\overline{\varphi}(gH) = \varphi(g) = \varphi(g'h) = \varphi(g')\underbrace{\varphi(h)}_{\in \ker\varphi} = \varphi(g')e_K = \varphi(g') = \overline{\varphi}(g'H),$$

so $\overline{\varphi}$ is well-defined. Moreover, $\overline{\varphi}$ is a group homomorphism as

$$\overline{\varphi}\,((aH)(bH)) = \overline{\varphi}\,(abH) = \varphi(ab) = \varphi(a)\varphi(b) = \overline{\varphi}(aH) = \overline{\varphi}(bH).$$

Clearly, by definition $\overline{\varphi} \circ \pi = \varphi$.

Now suppose that $\psi : {}^{G}/_{H} \to K$ is another group homomorphism such that $\psi \circ \pi = \varphi$. Then

$$\psi(gH) = \psi(\pi(g)) = \varphi(g) = \overline{\varphi}(gH),$$

and so $\psi = \overline{\varphi}$, thus $\overline{\varphi}$ is unique. □

Slogan. This can be remembered by the slogan "if a map collapses H, you can quotient by H". It is really more of a set-theoretic fact, though we won't go into that generality here.

4.7. **Isomorphism Theorems.**

A fundamental result in group theory is the following.

Theorem 4.7.1 (First Isomorphism Theorem)**.** *Let $\varphi : G \to G'$ be a group homomorphism, and let $\pi : G \to {}^{G}/_{\ker\varphi}$ be the quotient homomorphism. Then $\overline{\varphi} : {}^{G}/_{\ker\varphi} \to G'$ gives a group isomorphism*

$${}^{G}/_{\ker\varphi} \cong \mathrm{im}\varphi.$$

Proof. Let $H = \ker\varphi$. By Theorem 4.6.15, there is a unique group homomorphism $\overline{\varphi} : {}^{G}/_{H} \to G'$ defined by $\overline{\varphi}(gH) = \varphi(gH) = \varphi(g)$.

We first show that $\overline{\varphi}$ is injective. **RH:** We want to show that for $gH \in \ker\overline{\varphi}$, $gH = H$, thus every element of $\ker\overline{\varphi}$ is the identity element of G/H, and so $\overline{\varphi}$ is injective. Indeed, suppose that $gH \in \ker\overline{\varphi}$. Then

$$\overline{\varphi}(gH) = \varphi(g) = e',$$

and so $g \in \ker\varphi = H$, hence $gH = H$. Thus $\overline{\varphi}$ is injective.

Clearly $\mathrm{im}\overline{\varphi} = \mathrm{im}\varphi$, hence, still using the same name

$$\overline{\varphi} : {}^{G}/_{\ker\varphi} \to \mathrm{im}\varphi$$

is a surjective group homomorphism.

Thus $\overline{\varphi} : {}^{G}/_{\ker\varphi} \to \mathrm{im}\varphi$ is a bijective group homomorphism, hence an isomorphism. □

The First Isomorphism Theorem is a fundamental fact of groups. It is used repeatedly to construct isomorphisms, as it is usually easier to construct *some* group homomorphism first, and then mod out by the kernel to obtain an isomorphism.

Corollary 4.7.2. *Let $\varphi : G \to G'$ be a surjective group homomorphism. Then $G' \cong G/\ker\varphi$.*

Let's see the First Isomorphism Theorem in action with some familiar examples.

Example 4.7.3. Let G be a cyclic group with generator a. Then, as we showed in Proposition 4.2.5, there is a surjective group homomorphism $\varphi : \mathbb{Z} \to G, n \mapsto a^n$. Thus $G \cong \mathbb{Z}/\ker\varphi$. Since $\ker\varphi \leq \mathbb{Z}$ is a subgroup, $\ker\varphi = n\mathbb{Z}$ for some n which is the order of a, and so $G \cong \mathbb{Z}/n\mathbb{Z}$. If G is infinite, then $n = 0$ and $G \cong \mathbb{Z} \cong \mathbb{Z}/0\mathbb{Z}$.

Example 4.7.4 (Chinese Remainder Theorem)**.** Let $a, b \in \mathbb{Z}$ with $\gcd(a, b) = 1$. We have the quotient group homomorphisms $\pi_1 : \mathbb{Z} \to \mathbb{Z}/a\mathbb{Z}$ and $\pi_2 : \mathbb{Z} \to \mathbb{Z}/b\mathbb{Z}$. We can stitch these together into a group homomorphism

$$\varphi : \mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z},\ k \mapsto (\pi_1(k), \pi_2(k)).$$

Since $\gcd(a, b) = 1$, the classic Chinese Remainder Theorem (Theorem 3.1.19) shows that φ is surjective. Moreover, we can check that $\ker \varphi = \operatorname{lcm}(a, b) = \frac{ab}{\gcd(a,b)} = ab$, and so from the First Isomorphism Theorem (Theorem 4.7.1), we obtain

$$\mathbb{Z}/ab\mathbb{Z} \cong \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}.$$

Exercise 4.7.5. Prove the Chinese Remainder Theorem using the First Isomorphism Theorem. That is, show that

$$\mathbb{Z}/ab\mathbb{Z} \cong \mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z} \text{ if and only if } \gcd(a, b) = 1.$$

.... End of Class 14 (2/13) ....

Example 4.7.6. We've seen that $\mathbb{R}^\times \cong \mathrm{GL}_n(\mathbb{R}) \big/ \mathrm{SL}_n(\mathbb{R})$. We can now show this even more easily, since $\det : \mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}^\times$ is clearly surjective, with kernel (by definition) $\mathrm{SL}_n(\mathbb{R})$.

Example 4.7.7 ($\mathbb{R}/\mathbb{Z} \cong S^1$)**.** We can define a surjective group homomorphism

$$\exp : \mathbb{R} \to S^1,\ x \mapsto e^{2\pi i x}$$

and since $e^{2\pi i n} = 1$ if and only if $n \in \mathbb{Z}$, we see that $\ker \exp = \mathbb{Z}$. Thus, quotienting by the kernel, the First Isomorphism Theorem (Theorem 4.7.1) gives $\mathbb{R}/\mathbb{Z} \cong S^1$.

Example 4.7.8 ($\mathbb{C}^\times/\mathbb{R}^\times_{>0} \cong S^1$)**.** Using standard multiplication of complex numbers, we can define a surjective group homomorphism

$$\mathbb{C}^\times \to S^1,\ z \mapsto \frac{z}{|z|},$$

which is a group homomorphism as $|z_1 z_2| = |z_1||z_2|$. The kernel is exactly those complex numbers with $z = |z|$, which is exactly $\mathbb{R}^\times_{>0}$. Thus $\mathbb{C}^\times/\mathbb{R}^\times_{>0} \cong S^1$.

Example 4.7.9 ($\mathbb{C}^\times/S^1 \cong \mathbb{R}^\times_{>0}$)**.** Using standard multiplication of complex numbers, we can define a surjective group homomorphism

$$\mathbb{C}^\times \to \mathbb{R}^\times_{>0},\ z \mapsto |z|,$$

which is a group homomorphism as $|z_1 z_2| = |z_1||z_2|$. The kernel is exactly those complex numbers with $|z| = 1$, which is exactly S^1. Thus $\mathbb{C}^\times/S^1 \cong \mathbb{R}^\times_{>0}$.

Example 4.7.10 ($\mathbb{C}^\times \cong S^1 \times \mathbb{R}^\times_{>0}$)**.** Using the two maps defined above, we define a group homomorphism

$$\mathbb{C}^\times \to S^1 \times \mathbb{R}^\times_{>0},\ z \mapsto (\frac{z}{|z|}, |z|),$$

with inverse given simply by

$$S^1 \times \mathbb{R}^\times_{>0} \to \mathbb{C}^\times,\ (\frac{z}{|z|}, |z|) \mapsto \frac{z}{|z|}|z| = z.$$

You can check that the inverse is a group homomorphism, and thus $\mathbb{C}^\times \cong S^1 \times \mathbb{R}^\times_{>0}$. Thus is in fact the standard "polar decomposition" $z = |z|e^{i\theta_z}$.

Exercise 4.7.11. Let $f : G \to G'$ be a group homomorphism with $H = \ker f$. Suppose that G is a finite group.

(1) Show that $|G| = |\mathrm{im} f| \cdot |H|$
(2) Assume further that G' is also a finite group, and assume $\gcd(|G|, |G'|) = 1$. Show that $\mathrm{im} f = \{e'\}$, where $e' \in G'$ is the identity element.

Exercise 4.7.12. Prove the following proposition.

Proposition 4.7.13. *Let $\varphi : G \to G'$ be a surjective group homomorphism, and let $H \trianglelefteq G$ be a normal subgroup. Show that $\varphi(H)$ is a normal subgroup of G'.*

Definition 4.7.15. Let G be a group, and define the subgroup

$$[G, G] := \langle ghg^{-1}h^{-1} \;\mid\; g, h \in G\rangle \subseteq G,$$

called the *commutator subgroup*, generated by *commutators* $ghg^{-1}h^{-1}$. Prove the following proposition.

Proposition 4.7.16. *Let G be a group, show that $[G, G]$ is a normal subgroup of G, and that $G/[G, G]$ is abelian.*

Exercise 4.7.17. Prove the following proposition

Proposition 4.7.18. *Let G be a group and A an abelian group. If $\varphi : G \to A$ is a group homomorphism, then $[G, G] \leq \ker \varphi$.*

We now look at a consequence of the First Isomorphism Theorem, which is so ubiquitous that it is called the Third Isomorphism Theorem. We'll motivate it a little first.

If $N \trianglelefteq G$ is a normal subgroup, what are the subgroups of G/N?

Lemma 4.7.19. *Let $\pi : G \to G'$ be a group homomorphism. Then for any subgroup $\ker \pi \leq H \leq G$, we have $\pi^{-1}(\pi(H)) = H$.*

Proof. The inclusion $H \subseteq \pi^{-1}(\pi(H))$ is true in general for maps and subsets. Thus we only prove the inclusion $\pi^{-1}(\pi(H)) \subseteq H$. So let $g \in \pi^{-1}(\pi(H))$, thus $\pi(g) \in \pi(H)$, hence $\pi(g) = \pi(h)$ for some $h \in H$. This implies $\pi(gh^{-1}) \in \ker \pi$. Since $\ker \pi \leq H$, we have $gh^{-1} \in H$, and so $g \in Hh = H$, as desired. □

Lemma 4.7.20. *Let $\varphi : G \to G'$ be a surjective group homomorphism, and $H \leq G$ a subgroup with $\ker \varphi \leq H$. Then H is normal in G if and only if $\varphi(H)$ is normal in G'.*

Proof. If $H \trianglelefteq G$, then $\varphi(H) \trianglelefteq G'$ is Proposition 4.7.13. Conversely, suppose that $\varphi(H) \trianglelefteq G'$ is normal. Let $h \in H$ and $g \in G$. Since $\varphi(H)$ is normal in G',

$$\varphi(g)^{-1}\varphi(h)\varphi(g) = \varphi(g^{-1}hg) \in \varphi(H).$$

In other words, $g^{-1}hg \in \varphi^{-1}(\varphi(H)) = H$ by Lemma 4.7.19. Therefore, for any $g \in G$ and $h \in H$, $g^{-1}hg \in H$, hence $H \trianglelefteq G$ is normal. □

Theorem 4.7.21 (Subgroup Correspondence)**.** *Let G be a group, $N \trianglelefteq G$ a normal subgroup, and $\pi : G \to G/N$ the quotient homomorphism. Then the map*

$$\widetilde{\pi} : \{ \textit{ subgroups } H \leq G \ \mid \ N \leq H \leq G\} \to \{ \textit{ subgroups } H' \leq G/N\}, \ H \mapsto \pi(H) = H/N$$

is a bijection. Moreover, it also restricts to a bijection between the subsets of normal subgroups only.

Proof. If $H \leq G$ is a subgroup, then $\widetilde{\pi}(H) = H/N = \pi(H)$, which is the image of a subgroup under a group homomorphism, which is a subgroup. Thus $\widetilde{\pi}$ does indeed map to subgroups of G/N.

We will find an inverse, namely, define a map

$$\widetilde{\pi}^{-1} : \{ \text{ subgroups } H' \leq G/N\} \to \{ \text{ subgroups } H \leq G \ \mid \ N \leq H \leq G\}, \ H' \mapsto \pi^{-1}(H').$$

We check that $\widetilde{\pi}$ and $\widetilde{\pi}^{-1}$ are inverses, which amounts to checking that $\pi^{-1}(\pi(H)) = H$ for all $N \leq H \leq G$, and $\pi(\pi^{-1}(H')) = H'$ for all $H' \leq G/N$. The first follows from Lemma 4.7.19 and the second from the fact that π is surjective.

Finally, that $\widetilde{\pi}$ is still a bijection when considering only normal subgroups follows from Lemma 4.7.20 applied to the surjective group homomorphism $\pi : G \to G/N$. □

Thus, what if we have two normal subgroups $N \trianglelefteq K \trianglelefteq G$, how do the groups G/N, G/K and (K/N) compare? Well, we have $K/N \trianglelefteq G/N$, so it is tempting to write something like

$$\frac{G}{N} \Big/ \frac{K}{N} = \frac{G}{N} \cdot \frac{N}{K} = \frac{G}{K},$$

and we'll show that this is in fact correct!

Theorem 4.7.22 (Third Isomorphism Theorem)**.** *Let G be a group with normal subgroups*

$$N \trianglelefteq K \trianglelefteq G.$$

Then

$$G/N \Big/ K/N \cong G/K.$$

Proof. We will use the first isomorphism theorem. we have the quotient homomorphisms

$$\pi : G \to G/N,$$

and

$$\rho : G/N \to G/N \Big/ K/N ,$$

and composing them we obtain a surjective group homomorphism

$$\varphi : G \to G/N \Big/ K/N .$$

It remains to show that $K = \ker \varphi$. Note that

$$\begin{aligned} \ker \varphi &= \pi^{-1}(\ker \rho) \\ &= \pi^{-1}(K/N) \\ &= \pi^{-1}(\pi(K)) \\ &= K(\text{ by Lemma 4.7.19}), \end{aligned}$$

and the result follows from the first isomorphism theorem. □

. . . . End of Class 15 (2/16) . . . .

A second proof of the Third Isomorphism Theorem. Since K is a normal subgroup, we have the quotient map $\pi : G \to G/K$. From Theorem 4.6.15, as $N \subseteq K$, we have an induced map

$$\overline{\pi} : G/N \to G/K.$$

Since π is surjective, $\overline{\pi}$ is still surjective. We compute $\ker \overline{\pi}$. By definition

$$\ker \overline{\pi} = \{gN \mid \overline{\pi}(gN) = \pi(g) = gK = K\} = \{gN \mid g \in K\} = K/N,$$

thus by the first isomorphism theorem (Theorem 4.7.1) we have

$$G/K \cong {}^{G/N}\big/_{\ker \overline{\pi}} = {}^{G/N}\Big/_{K/N}. \qquad \square$$

You may have noticed that we skipped an isomorphism theorem. Let's motivate it now.

Lemma 4.7.23. *Let G be a group, $N \trianglelefteq G$ a normal subgroup, and $H \leq G$ a subgroup. Then*

(1) $HN = \{hn \mid h \in H,\ n \in N\}$ is a subgroup of G, and
(2) $H \cap N$ is a normal subgroup of H.

Proof. We check that HN is a subgroup. Since H and N are both subgroups, they both contain e, and so $e = ee \in HN$. Let $h_1, h_2 \in H$ and $n_1, n_2 \in N$. We have

$$h_1 \underbrace{n_1 h_2}_{NH} n_2 = h_1 \underbrace{h_2 n_3}_{HN} n_2 \in HN$$

for some $n_3 \in N$ as N is normal. Thus HN is closed under products. For inverses, we have

$$(h_1 n_1)^{-1} = n_1^{-1} h_1^{-1} = h_1^{-1} n_4 \in HN$$

for some $n_4 \in N$ as N is normal. Thus HN is a subgroup of G.

We have seen in Proposition 4.3.9 that the intersection of two subgroups is a subgroup, thus $H \cap N$ is a subgroup. It remains to show that $H \cap N$ is normal in H. Thus let $n \in H \cap N$ and $h \in H$. Then $h^{-1}nh \in N$ as N is normal in G. We also have $h^{-1}nh \in H$ as H is a subgroup and $h, n \in H$. Thus for any $h \in H$ and $n \in H \cap N$, $h^{-1}nh \in H \cap N$, therefore $H \cap N \trianglelefteq H$ is normal. □

Theorem 4.7.24 (Second Isomorphism Theorem)**.** *Let G be a group, $N \trianglelefteq G$ a normal subgroup, and $H \leq G$ a subgroup. Then*

$${}^{HN}\big/_{N} \cong {}^{H}\big/_{H \cap N}.$$

Proof. We will use the first isomorphism theorem. Define map

$$\varphi : H \to {}^{G}\big/_{N},\ h \mapsto hN,$$

which is simply the restriction of the quotient map to H. The image of φ is simply the subgroup ${}^{HN}\big/_{N} \leq {}^{G}\big/_{N}$. To find the kernel of φ, note that for any $h \in H$,

$$\begin{aligned} h \in \ker \varphi &\iff hN = N \\ &\iff h \in N \\ &\iff h \in H \cap N, \end{aligned}$$

and so $\ker \varphi = H \cap N$. Thus, by the first isomorphism theorem, we have

$${}^{HN}\big/_{N} = \mathrm{im}\varphi \cong {}^{H}\big/_{\ker \varphi} = {}^{H}\big/_{H \cap N}. \qquad \square$$

These isomorphism theorems are used over and over when considering groups. However, we'll move on to investigate an important collection of groups that is long overdo.

4.8. **Permutation Groups.** We'll investigate permutation groups, specifically the groups S_n.

Recall that if X is a set, then the set

$$\{\text{bijections } X \to X\}$$

is a group under composition. We will be mostly concerned with the case when $X = [n] := \{1, 2, \ldots, n\}$, in which case the group of bijections is denoted S_n and called the *symmetric group on n elements.*

One reason why symmetric groups are so important is because of the following theorem.

Theorem 4.8.1 (Cayley's Theorem)**.** *Let G be a finite group, then G is isomorphic to a subgroup of S_n for some n.*

Proof. Let $n = |G|$. We clearly have $\mathrm{Perm}(G) \cong S_n$, by definition. We've seen in Proposition 4.4.46 that the map

$$L : G \to \mathrm{Perm}(G), \ g \mapsto L_g$$

is an injective group homomorphism. The image of L is a subgroup (isomorphic to G by the first isomorphism theorem) of $\mathrm{Perm}(G) \cong S_n$. □

Let us understand the groups S_n more carefully.

We've seen in Theorem 4.5.20 that $|S_n| = n!$, and in Proposition 4.1.12 that S_n is not abelian for $n \geq 3$. Since permutation groups are so ubiquitous, it would be nice to have a way of writing down their elements.

Notation 4.8.2. Let $\sigma \in S_n$, say

$$\begin{aligned} \sigma : 1 &\mapsto a_1 \\ 2 &\mapsto a_2 \\ &\vdots \\ n &\mapsto a_n, \end{aligned}$$

we represent this in *two-line notation* as an array

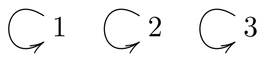
$$\begin{bmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & n \end{bmatrix},$$

which symbolizes that $i \mapsto a_i$.

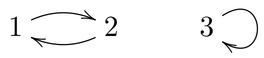
. . . . End of Class 16 (2/18) . . . .

Example 4.8.3. The elements of S_3 are:

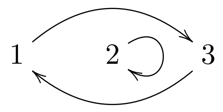
Identity $e = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$



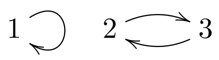
Transposition $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}$



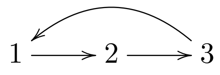
Transposition $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$



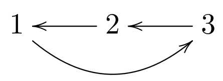
Transposition $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$



3-cycle $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$



3-cycle $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}$



One thing to notice is that every picture breaks up into *cycles*.

4.8.1. *Cycle decomposition.*

Definition 4.8.4. A cycle C for a permutation $\sigma \in S_n$ is a non-empty ordered collection of elements $a_1, \ldots, a_k \in [n]$ such that

$$\begin{aligned} \sigma(a_1) &= a_2 \\ \sigma(a_2) &= a_3 \\ &\vdots \\ \sigma(a_{k-1}) &= a_k \\ \sigma(a_k) &= a_1. \end{aligned}$$

As a picture, this is

$$\begin{tikzcd} a_1 \arrow[r] & a_2 \arrow[r] & \cdots \arrow[r] & a_k \arrow[lll, bend left] \end{tikzcd}$$

Example 4.8.5. The permutation

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{bmatrix}$$

has cycles

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[r] & 4 \arrow[ll, bend left] & \text{and} & 3 \arrow[r, bend left] & 5 \arrow[l, bend left] \end{tikzcd}$$

Example 4.8.6. If $\sigma(x) = x$, then

$$\begin{tikzcd} x \arrow[loop above] \end{tikzcd}$$

is a cycle for σ.

Theorem 4.8.7 (Cycle Decomposition of sets)**.** *Let $\sigma \in S_n$. The set $[n] = \{1, \ldots, n\}$ can be partitioned uniquely into cycles for σ.*

Proof. We first show that every $i \in \{1, \ldots, n\}$ is contained in some cycle. Consider the sequence $a_1 = i, a_2 = \sigma(i), a_3 = \sigma(a_2) \cdots$. Since $\{1, \ldots, n\}$ is finite, the sequence must eventually repeat. Let j be the firs repeated element. We claim that $j = i$.

Suppose that $j \neq i$. Then $j = a_k = \sigma(a_{k-1}) = \sigma(a_{k+m})$ for some k, m, but $a_{k-1} \neq a_{k+m}$. However, this contradicts σ being a bijection. Thus $i = j$.

It follows that $i, a_2, a_3, \ldots, a_{k-1}$ forms a cycle containing i.

This in fact proves more, that i is contained in a cycle, and that the cycle containing i is uniquely determined by i, as all other elements are obtained by applying σ to i.

Therefore, if C and C' are two cycles for σ with $C \cap C" \neq \emptyset$, then letting $i \in C \cap C'$ be any element, $C = C'$ is the cycle containing i.

Thus the cycles for σ partition $[n]$. □

Definition 4.8.8. Let $C \subseteq [n]$ be a cycle for some permutation $\sigma \in S_n$. We define

$$\sigma_C(x) = \begin{cases} \sigma(x) & \text{if } x \in C \\ x & \text{if } x \notin C \end{cases}$$

called a *cyclic permutation* which we can think of as "do σ on the cycle C, and nothing outside the cycle C". If the cycle C has k distinct element, we call σ_C a *k-cycle*.

Proposition 4.8.9. *Let $\sigma_C \in S_n$ be a k-cycle. Then σ_C has order k.*

Proof. Let $C = \{a_1, a_2, \dots, a_k\} \subseteq [n]$ be the cycle. For $i < k$, we have $\sigma_C^i(a_1) = a_{i+1} \neq a_i$, thus $\operatorname{ord}(\sigma_C) \geq k$. Moreover, we have $\sigma^k(a_i) = a_i$ for all i. And for $x \notin \{a_1, \dots, a_k\}$, $\sigma_C^i(x) = x$ for all i. Thus for all $x \in [n]$, $\sigma_C^k(x) = x$. Thus $\sigma_C^k = \mathrm{id}$, whereby $\operatorname{ord}(\sigma_C) \leq k$. Therefore, $\operatorname{ord}(\sigma_C) = k$. □

Proposition 4.8.10 (Cycle decomposition of permutations)**.** *Let $\sigma \in S_n$, and let*

$$\{1, \dots, n\} = C_1 \sqcup \cdots \sqcup C_k$$

be the cycle decomposition induced by σ. Then

$$\sigma = \sigma_{C_1} \circ \cdots \circ \sigma_{C_k}.$$

Proof. The idea is to check that both sides do the same thing to any $x \in [n]$. If $x \in C_i$, then $\sigma(x)$ is the next element in the cycle containing x. And

$$\begin{aligned} \sigma_{C_1} \circ \cdots \circ \sigma_{C_i} \circ \cdots \circ \sigma_{C_k}(x) &= \sigma_{C_1} \circ \cdots \circ \sigma_{C_i}(x) \\ &= \sigma_{C_1}(\cdots \sigma_{C_{i-1}}(\sigma_{C_i}(x))) \\ &= \sigma_{C_i}(x) \\ &= \text{ next element in the cycle containing } x. \end{aligned}$$

□

Corollary 4.8.11. *If C_i and C_j are two disjoint cycles in $\{1, \dots, n\}$, then*

$$\sigma_{C_i} \circ \sigma_{C_j} = \sigma_{C_j} \circ \sigma_{C_i}.$$

Proof. Again, they do the same thing for all $x \in [n]$. □

Notation 4.8.12 (cycle notation)**.** Let

$$\begin{tikzcd} a_1 \arrow[r] & a_2 \arrow[r] & \cdots \arrow[r] & a_k \arrow[lll, bend left] \end{tikzcd}$$

be a cycle. We denote by

$$(a_1 a_2 \cdots a_k)$$

the permutation

$$x \mapsto \begin{cases} a_2 & \text{if } x = a_1 \\ a_3 & \text{if } x = a_2 \\ & \vdots \\ a_k & \text{if } x = a_{k-1} \\ a_1 & \text{if } x = a_k \\ x & \text{if } x \notin \{a_1, \dots, a_k\} \end{cases}$$

Example 4.8.13. The permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$$

has cycles

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[r] & 4 \arrow[ll, bend left] & \text{and} & 3 \arrow[r, bend left] & 5 \arrow[l, bend left] \end{tikzcd}$$

and is written in cycle notation as

$$(124)(35) = (35)(124).$$

Definition 4.8.14. A *transposition* is an element of S_n of the form $\tau = (ij)$ for some $i \neq j \in [n]$.

Example 4.8.15. Let us see what happens when we pre-compose with a transposition.

Let $\sigma \in S_n$ and $(ij) \in S_n$ a transposition. What is $\sigma(ij)$?

If $x \notin \{i, j\}$, then $(ij)(x) = x$, so

$$\sigma(ij)(x) = \sigma(x).$$

And we can compute the remaining two values, namely

$$\begin{aligned}\sigma(ij)(i) &= \sigma(j)\\ \sigma(ij)(j) &= \sigma(i).\end{aligned}$$

Thus

$$\sigma(ij)(x) = \begin{cases}\sigma(x) & \text{if } x \notin \{i, j\}\\ \sigma(j) & \text{if } x = i\\ \sigma(i) & \text{if } x = j\end{cases}$$

Exercise 4.8.16. Prove the following proposition.

Proposition 4.8.17. *For $(a_1 \cdots a_k) \in S_n$,*

$$(a_1a_2 \cdots a_k) = (a_1a_2)(a_2a_3) \cdots (a_{k-2}a_{k-1})(a_{k-1}a_k) = (a_1a_k)(a_1a_{k-1}) \cdots (a_1a_3)(a_1a_2).$$

With this, we can show that every permutation can be written as a product of transpositions.

Proposition 4.8.18. *Every permutation in S_n can be written as a product of transpositions.*

Proof. Let $\sigma \in S_n$. From Proposition 4.8.10, we can write $\sigma = \sigma_a \circ \cdots \circ \sigma_k$ as a product of cyclic permutations. From Proposition 4.8.17, each $\sigma_i = \tau_1^{(i)} \cdots \tau_{m_i}^{(i)}$ is a product of transpositions, and so

$$\sigma = \underbrace{(\tau_1^{(1)} \cdots \tau_{m_1}^{(1)})}_{\sigma_1} \cdots \cdots \underbrace{(\tau_1^{(k)} \cdots \tau_{m_k}^{(k)})}_{\sigma_k}.$$ □

Remark 4.8.19. The expression of a permutation as a product of transpositions is highly non-unique! In fact, we can make the expression arbitrarily long.

Nevertheless, there is an important quantity that is preserved when we express a permutation as a product of transpositions.

We'll need some set-up first.

Example 4.8.20. Let $f(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$ be a real function of n variables. For $\sigma \in S_n$, we can define a new function $\sigma f : \mathbb{R}^n \to \mathbb{R}$ by

$$(\sigma f)(x_1, \ldots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}).$$

Then for $\sigma, \tau \in S_n$, we have

$$(\sigma\tau)f = \sigma(\tau f).$$

Indeed, we compute

$$\begin{aligned}\sigma(\tau f)(x_1, \ldots, x_n) &= \sigma f(x_{\tau(1)}, \ldots, x_{\tau(n)})\\ &= f(x_{\sigma(\tau(1))}, \ldots, x_{\sigma(\tau(n))})\\ &= ((\sigma\tau)f)(x_1, \ldots, x_n).\end{aligned}$$

You can easily verify that for $f, g : \mathbb{R}^n \to \mathbb{R}$, and for $\sigma \in S_n$, we have

$$\begin{aligned}\sigma(f + g) &= \sigma f + \sigma g\\ \sigma(fg) &= (\sigma f)(\sigma g),\end{aligned}$$

and thus for any $c \in \mathbb{R}$, we have

$$\sigma(cf) = c(\sigma f).$$

Theorem 4.8.21. *For each $\sigma \in S_n$, we can assign a* sign, $\operatorname{sgn}(\sigma) \in \{1, -1\}$, *such that*

(a) *If $\tau \in S_n$ is a transposition, then* $\operatorname{sgn}(\tau) = -1$

(b) *If $\sigma, \sigma' \in S_n$ are permutations, then*
$$\operatorname{sgn}(\sigma\sigma') = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma').$$

(c) *The map* $\operatorname{sgn} : S_n \to \{1, -1\}$ *is a group homomorphism.*

Proof. Let Δ be the function
$$\Delta(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$
It is clear that for $\sigma \in S_n$,
$$\sigma\Delta = \pm\Delta,$$
as the permutation either flips each term of the product or not, and writing $(x_i - x_j) = -(x_j - x_i)$, we can collect sign changes together. We define $\operatorname{sgn}(\sigma)$ by
$$\operatorname{sgn}(\sigma) = \begin{cases} +1 & \text{if } \sigma\Delta = \Delta \\ -1 & \text{if } \sigma\Delta = -\Delta. \end{cases}$$

Let $\tau \in S_n$ be a transposition, say $\tau = (rs)$ with $r < s$. Then
$$\tau\Delta(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} \tau(x_j - x_i)$$
and we compute that
$$\tau(x_s - x_r) = (x_{\tau(s)} - x_{\tau(r)}) = (x_r - x_s) = -(x_s - x_r).$$
If a factor $(x_j - x_i)$ does not contain r or s, then the factor remains unchanged. All other factors can be considered in pairs, one containing r and the other containing s as follows
$$(x_k - x_s)(x_k - x_r) \text{ if } k > s,$$
$$(x_s - x_k)(x_k - x_r) \text{ if } r < k < s,$$
$$(x_s - x_k)(x_r - x_k) \text{ if } k < r,$$
each remaining unchanged after applying τ. Hence
$$\tau\Delta = -\Delta,$$
and so $\operatorname{sgn}(\tau) = -1$ for a transposition.

For the remaining statements, we simply compute
$$\operatorname{sgn}(\sigma\sigma')\Delta = (\sigma\sigma')\Delta = \sigma(\sigma'\Delta) = \sigma(\operatorname{sgn}(\sigma')\Delta) = \operatorname{sgn}(\sigma')\sigma\Delta = \operatorname{sgn}(\sigma')\,\sigma\,\Delta,$$
and thus
$$\operatorname{sgn}(\sigma\sigma') = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma'),$$
as desired. This is exactly what it means for $\operatorname{sgn} : S_n \to \{1, -1\}$ is a group homomorphism. □

Remark 4.8.22. In particular, if $\sigma \in S_n$ is a product of m transpositions
$$\sigma = \tau_1 \cdots \tau_m,$$
then
$$\operatorname{sgn}(\sigma) = (-1)^m.$$

Definition 4.8.23. We call a permutation $\sigma \in S_n$ *even* if $\operatorname{sgn}(\sigma) = 1$, and *odd* if $\operatorname{sgn}(\sigma) = -1$. Equivalently, σ is even (resp. odd) if σ can be written as a product of an ever (resp. odd) number of permutations.

Exercise 4.8.24. Let $\sigma \in S_n$. If σ is even, then every expression of σ as a product of transpositions has an even number of transpositions. If σ is odd, then every expression of σ as a product of transpositions has an odd number of transpositions.

Corollary 4.8.25. *If $\sigma \in S_n$, then* $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Proof. We have

$$1 = \operatorname{sgn}(\mathrm{id}) = \operatorname{sgn}(\sigma\sigma^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1}),$$

hence either $\operatorname{sgn}(\sigma)$ and $\operatorname{sgn}(\sigma^{-1})$ are both equal to 1, or both equal to -1, as desired. □

Definition 4.8.26. The kernel of the group homomorphism $\operatorname{sgn} : S_n \to \{1, -1\}$ is call the *alternating group*, denoted by A_n. It consists of the even permutations.

Exercise 4.8.27. Write the following permutations in two-line notation or in cycle notation.

(1) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$

(2) (132)

(3) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix}$

(4) (1342)

(5) (145)(23)

Exercise 4.8.28. Determine the sign of the following permutations:

(1) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$

(2) (132)

(3) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix}$

(4) (1342)

(5) (145)(23)

Exercise 4.8.29. Let G be a finite group

(1) If $|G| = 2k$ for some positive integer k, show that G has an element of order 2. Hint: show that there exists $g \in G$, $g \neq e$, such that $g = g^{-1}$.

(2) Assume that k is odd. Let $a \in G$ have order 2, and let $L_a : G \to G, g \mapsto ag$ be left translation by a. Show that L_a is an odd permutation.

(3) Still assuming that k is odd, show that G has a normal subgroup of order k. Hint: use the previous parts.

Exercise 4.8.30. Show that every permutation in S_n can be written as a product of the transposition (12) and the n-cycle $(12 \cdots n)$.

Let's see if we can find some nice ways to compute in S_n.

Proposition 4.8.31. *Let $\sigma, \tau \in S_n$ and suppose that σ has cycle decomposition*

$$\sigma = (a_1 a_2 \cdots a_{k_1})(b_1 b_2 \cdots b_{k_2}) \cdots .$$

Then $\tau\sigma\tau^{-1}$ has cycle decomposition

$$(\tau(a_1)\tau(a_2)\cdots\tau(a_{k_1}))(\tau(b_1)\tau(b_2)\cdots\tau(b_{k_2}))\cdots ,$$

that is, $\tau\sigma\tau^{-1}$ is obtained from σ by replacing each every i int eh cycle decomposition by the entry $\tau(i)$.

Proof. □

Definition 4.8.32. Let G be a group, and $g \in G$. We call the set

$$\mathrm{Cl}(g) \coloneqq \{hgh^{-1} \mid h \in G\}$$

the *conjugacy class of g*. It consists of all elements of G that are conjugate to g.

Definition 4.8.33. If $\sigma \in S_n$ is a product of disjoint cycles of lengths $1 \leq n_1 \leq n_2 \leq \cdots \leq n_r$, we say that σ has *cycle type* $n_1, n_2, \ldots, n_r$.

Remark 4.8.34. If $\sigma \in S_n$, then the cycle type of σ is a *partition of* n, i.e., $1 \leq n_1 \leq n_2 \leq \cdots \leq n_r$ with $n_1 + n_2 + \cdots + n_r = n$.

Proposition 4.8.35. *Two elements of S_n are conjugate if and only if they have the same cycle type. The number of conjugacy classes of S_n equals the number of partitions of n.*

Proof. By Proposition 4.8.31, conjugate permutations have the same cycle type. Conversely, suppose that σ_1 and σ_2 have the same cycle type, and order the cycles in non-decreasing length, including 1-cycles. So perhaps σ_1 is

$$(a_1)(a_2)(a_3 a_4)\cdots(a_{n-2}a_{n-1}a_n).$$

Ignoring parentheses, each cycle decomposition is a list in which all the integers $1, \ldots, n$ appear exactly once, for example the list for σ_1 is

$$a_1, a_2, a_3, \ldots, a_n.$$

Let τ be the permutation sending the i^{th} number on the list for σ_1 to the i^{th} number on the list for σ_2. By Proposition 4.8.31, we have

$$\tau\sigma_1\tau^{-1} = \sigma_2,$$

and so σ_1 and σ_2 are conjugate.

There is a bijection between conjugacy classes of S_n and permissible cycle types. Each cycle type gives a partition of n, and each partition of n gives a cycle type by inserting parentheses into the list

$$1\ 2\ 3\ \cdots\ n,$$

and the second assertion follows. □

Remark 4.8.36. The study of the symmetric group is an interesting and exciting field of mathematics, and is the central question in an area of math called "Algebraic Combinatorics".

4.9. **Dihedral Groups.** We'll take a closer look at special subgroups of S_n, the dihedral groups D_n which are the symmetries of a regular n-gon.

Definition 4.9.2. A *regular n-gon* is an n-sided polygon with all sides of equal length and all angles between sides equal.

For example equilateral triangles, squares,... Note that for a regular n-gon to exist, we must have $n \geq 3$.

Definition 4.9.3. Let V_n be the set of vertices of a regular n-gon X. A *symmetry* of X is a bijection $f : V_n \to V_n$ such that if x and y are vertices connected by an edge, then $f(x)$ and $f(y)$ are also connected by an edge.

Proposition 4.9.4. *The set of symmetries of a regular n-gon form a group under composition.*

Proof. Let $n \in \mathbb{Z}_{>0}$, $n \geq 3$, and X be a regular n-gon.

The identity map on vertices (the "do nothing bijection") is clearly a symmetry of X.

Let f, g be symmetries of X, and let x, y be edges of X connected by an edge. Then $g(x)$ and $g(y)$ are connected by an edge, hence $f(g(x))$ and $f(g(y))$ are connected by an edge. Thus, as $f \circ g$ is a bijection, it is also a symmetry of X.

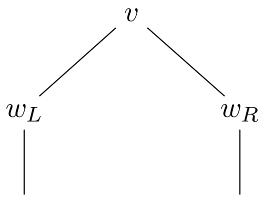
Finally, since f is a bijection, there is an inverse bijection f^{-1}. It remains to show that if x and y are vertices connected by an edge, then $f^{-1}(x)$ and $f^{-1}(y)$ are also connected by an edge. Let v, w be the two distinct vertices connected by edges to $f^{-1}(x)$. Then $f(v)$ and $f(w)$ are connected by an edge to $f(f^{-1}(x)) = x$. Thus $f(v) = y$ or $f(w) = y$. Thus $v = f^{-1}(y)$ or $w = f^{-1}y$, i.e., that $f^{-1}(y)$ is connected by an edge to $f^{-1}x$. Thus f^{-1} is also a symmetry of X. □

Definition 4.9.5. Let D_n be the group of symmetries of a regular n-gon.

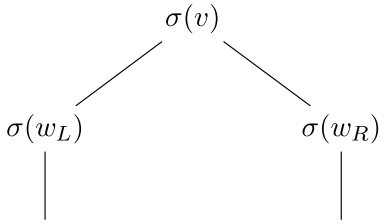
Can we describe the structure of D_n?

Lemma 4.9.6. *The order of D_n is $2n$.*

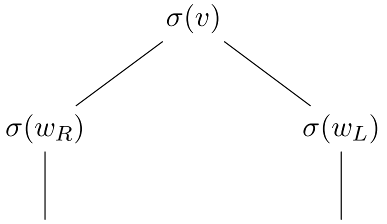
Proof. Let v be any vertex of a regular n-gon X, and denote its adjacent vertices by w_R and w_L, such that X looks locally like



with w_R on the right of v and w_L on the left. Any $\sigma \in D_n$ must send w_R and w_L to the two neighboring vertices of $\sigma(v)$, so there are two possibilities after applying σ:



or



.

Once $\sigma(v)$ is chosen, and the relative positions of $\sigma(w_L)$ and $\sigma(w_R)$ are known, σ is completely determined. Thus there are at most $2n$ elements in D_n.

Conversely, we can construct $2n$ elements of D_n by picking $\sigma(v)$ to be any vertex, and any orientation for $\sigma(w_R)$ and $\sigma(w_L)$. □

Let's find some elements of D_n. There is a rotation r by $\frac{2\pi}{n}$ radians. And for a vertex v, we can reflect around some line containing v, giving a reflection s.

Proposition 4.9.7. *Let X be a regular n-gon, and $r \in D_n$ be rotation by $\frac{2\pi}{n}$ radians. Then r has order n.*

Proof. Labeling the vertices with $1, \ldots, n$, the rotation r is simply the cyclic permutation $(1 \cdots n)$. □

Theorem 4.9.8. *The group D_n has the $2n$ elements*

$$\mathrm{id}, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1},$$

where r is rotation by $\frac{2\pi}{n}$ radians and s is a reflection around a fixed vertex.

Proof. We show that all of these elements are distinct. Since r has order n, we see that $\mathrm{id}, r, \dots, r^{n-1}$ are all distinct. For any vertex v, and of $s, sr, \dots, sr^{n-1}$ changes the order of the vertices neighboring v, thus the elements $s, sr, \dots, sr^{n-1}$ are distinct from $\mathrm{id}, r, \dots, r^{n-1}$. Finally, suppose that $sr^i = sr^j$ for some $1 \leq i, j \leq n-1$. Then multiplying by s on the left yields $r^i = r^j$, and thus $i = j$, as r has order n. Hence all of the elements $s, sr, \dots, sr^{n-1}$ are also all distinct. □

Corollary 4.9.9. *The group of symmetries of any two regular n-gons are isomorphic.*

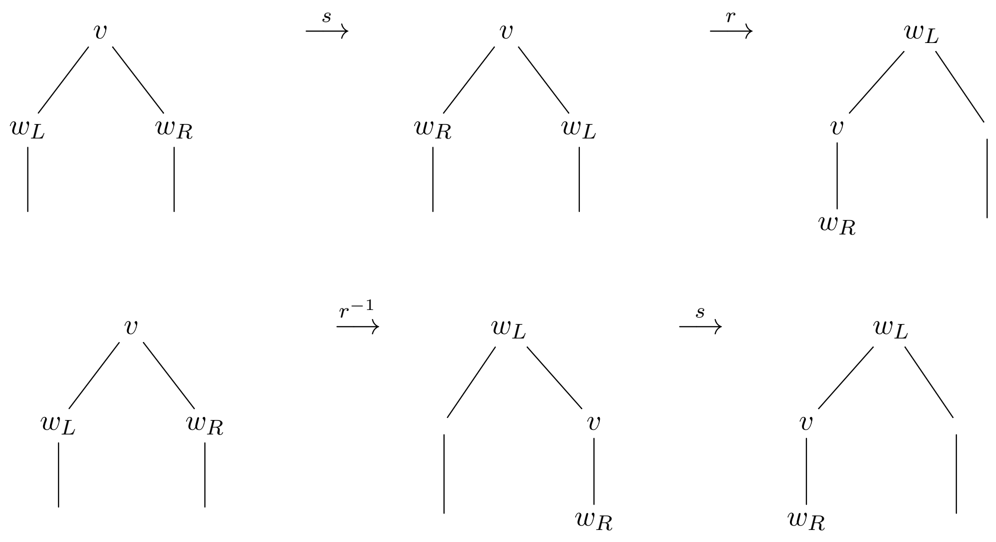
This, we can talk about *the* dihedral group D_n of order $2n$. Let's see how we can compute in D_n, in particular, how can we simplify expressions like rs, rs^{10}, or $rsrsr^3s^6$...

Lemma 4.9.10. *In D_n we have*

(1) $\mathrm{ord}(s) = \mathrm{ord}(sr^i) = 2$
(2) $r^is = sr^{-i}$

Proof. We have $s \neq \mathrm{id}$. However, we know that $s^2(v) = v$, and s^2 also fixes the neighbors of v (as it reflects across v twice). Thus $s^2 = \mathrm{id}$, and $\mathrm{ord}(s) = 2$.

We first prove that $r^is = sr^{-i}$ by induction. For the base case, we show that $rs = sr^{-1}$. We simply compute both sides:



and so $rs = sr^{-1}$.

Now assume that we have shown that $r^is = sr^{-i}$ for some $1 \leq i$. We compute

$$\begin{aligned} r^{i+1}s &= rr^is \\ &= rsr^{-i} \text{ by induction hypothesis} \\ &= sr^{-1}r^{-i} \text{ base case} \\ &= sr^{-i-1} = sr^{-(i+1)}. \end{aligned}$$

By induction, we have $r^is = sr^{-i}$ for all i.

Finally, we prove that $\operatorname{ord}(sr^i) = 2$. As sr^i does not fix v for $1 \leq i \leq n-1$, we have $sr^i \neq \mathrm{id}$. We compute

$$\begin{aligned}(sr^i)^2 &= sr^i sr^i \\ &= ssr^{-i}r^i \\ &= s^2 \\ &= \mathrm{id},\end{aligned}$$

and the result follows. □

4.10. Group Presentations.

We've seen that in the dihedral group D_n, we have some nice relations that help us simplify expressions, like $rs = sr^{-1}$, which can help us simplify any expression in D_n, noting that

$$s^2 = r^n = e.$$

Exercise 4.10.1. Let r and s be the rotation and reflection of D_7. Simplify the expression $s^6 r^{10} s r^2$ into the form $s^j r^i$ for some integers i, j.

So, we could describe the group D_n using generators and relations as

$$D_n = \left\langle \underbrace{r, s}_{\text{generators}} \mid \underbrace{s^2 = r^n = \mathrm{id},\ rs = sr^{-1}}_{\text{relations}} \right\rangle$$

which means that every element $g \in D_n$ can be written as a product of the elements r and s, and every equality between elements of D_n is a consequence of the relations.

Definition 4.10.2. Let G be a group, and $S \subseteq G$ such that $\langle S \rangle = G$. For some set of relations R, which are equations in the elements of $S \cup S^{-1} \cup \{\mathrm{id}\}$, we write $G = \langle S \mid R \rangle$ to mean that

- G is generated by S, and
- every relation in G is a consequence of the relations in R.

Remark 4.10.3. Warning! This is not really that precise, but is enough to work with.

Let's see some examples.

Example 4.10.4. The group $\mathbb{Z}^2$ has the presentation

$$\mathbb{Z}^2 = \langle a, b \mid ab = ba \rangle.$$

The relation just means that a and b commute. We could call $a = (1,0)$ and $b = (0,1)$, though there are also other elements that could generate $\mathbb{Z}^2$, for example $(3,1)$ and $(2,1)$.

Example 4.10.5. We've seen a few ways to write elements of the symmetric groups.

$$S_n = \langle \tau_1, \ldots, \tau_{n-1} \mid \tau_i^2 = e, \tau_i\tau_j = \tau_j\tau_i \text{ if } |i-j| > 1\ , \tau_i\tau_{i+1}\tau_i = \tau_{i+1}\tau_i\tau_{i+1} \rangle$$

which is just the presentation with generators transpositions $\tau_i = (i \;\; i+1)$. But showing that the group you get is actually just S_n would take some work!

More generally, we could ask:

Question 1. *Given a group presentation $G = \langle S \mid R \rangle$, what can we deduce about the group G?*

It turns out that even if S and R are finite sets of generators and relations, in which case G is called *finitely presented*, simple questions are still very difficult to answer.

Example 4.10.6.

- Any finite group is finitely presented.

- $\mathbb{Z}^n$ is finitely presented.
- $\mathbb{Q}$ is not finitely generated, and not finitely presented.
- The Lamplighter group
$$L = \langle a, t \mid a^2 = (at^nat^{-n})^2 = e, n \in \mathbb{Z}_{\geq 0}$$
is finitely generated, but has no finite presentation. Though we will not prove this.

Suppose we ask a *very* basic question. Let $G = \langle S \mid R \rangle$ be a group presentation, is G isomorphic to the trivial group?

Exercise 4.10.7. Let $G = \langle a, b \mid a^{-1}ba = b^2, b^{-1}ab = a^2 \rangle$. Show that G is the trivial group.

However, in general, there isn't really anything we can say.

Theorem 4.10.8 (Undecidability of group isomorphism problem)**.** *There is no algorithm that, given a group presentation $G = \langle S \mid R \rangle$, decides whether the group G is trivial or not.*

Group presentations are an active area of research, full of interesting open problems.

4.11. **Composition Series and Solvability.** So we may not be able to determine a group using generators and relations in a nice way. How else can we characterize groups? In this section we'll discuss the Hölder program for finite groups, and see how we could try to classify all groups.

Inspired by the fundamental theorem of arithmetic, that every positive integer factors into a product of primes in an essentially unique way, we want to find a way to "factor" groups.

Let G be a group, and suppose that $N \trianglelefteq G$ is a normal subgroup. We can think of N and G/N as "breaking up" G into smaller pieces. The "primes" should be the groups with no normal subgroups.

Definition 4.11.1. A non-trivial group G is called *simple* if the only normal subgroups are $\{e\}$ and G. That is, if G has no non-trivial normal subgroups.

Exercise 4.11.2. Show that if p is prime, then a cyclic group of order p is simple.

Exercise 4.11.3. Prove the following proposition.

Proposition 4.11.4. *Let G be a simple group, and $\varphi : G \to G'$ be a group homomorphism. Then φ is either injective or* $\mathrm{im}\varphi = \{e'\}$.

An interesting example of simple groups is given by the following theorem, which we will not prove.

Theorem 4.11.5. *The alternating groups A_n are simple for $n \geq 5$.*

The idea for our "factorization theorem" for groups is to break up a finite group G into pieces that are all simple.

Definition 4.11.6. Let G be a group. A *normal tower* is a sequence of subgroups
$$G_m \triangleleft \cdots \triangleleft G_2 \triangleleft G_1 \triangleleft G_0 = G$$
with $G_{i+1} \triangleleft G_i$ normal for every i. The groups G_i/G_{i+1} are called *factor groups.* A *refinement* of a normal tower
$$G_m \triangleleft \cdots \triangleleft G_2 \triangleleft G_1 = G$$

is a normal tower

$$H_n \lhd \cdots \lhd H_2 \lhd H_1 = G$$

obtained by inserting a finite number of subgroups into the first tower, each still normal in the next. We say two normal towers

$$\{e\} = G_m \lhd \cdots \lhd G_2 \lhd G_1 = G$$

and

$$\{e\} = H_n \lhd \cdots \lhd H_2 \lhd H_1 = G$$

are *equivalent* if $n = m$ and the sets of factor groups $\left\{G_i \big/ G_{i+1}\right\}_{i=1}^{n-1}$ and $\left\{H_i \big/ H_{i+1}\right\}_{i=1}^{n-1}$ are the same. A normal tower ending with $\{e\}$ with simple factor groups is called a *composition series* for G.

Exercise 4.11.7. Prove the following proposition.

Proposition 4.11.8. *Let G be a finite group. Then G has a normal tower*

$$\{e\} = G_r \lhd \cdots \lhd G_2 \lhd G_1 = G$$

such that each factor group $G_i \big/ G_{i+1}$ is simple. That is, every finite group has a composition series. Hint: use Theorem 4.7.21.

Example 4.11.9. Let's find some composition series for $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$. Starting with the normal subgroup $\mathbb{Z}/2\mathbb{Z}$, we can form the normal tower

$$\{e\} \lhd \mathbb{Z}/2\mathbb{Z} \lhd \mathbb{Z}/6\mathbb{Z},$$

and note that $\mathbb{Z}/6\mathbb{Z} \big/ \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z}$, thus all the factor groups are simple. Similarly, we have a normal subgroup $\mathbb{Z}/3\mathbb{Z}$ and can form the normal tower

$$\{e\} \lhd \mathbb{Z}/3\mathbb{Z} \lhd \mathbb{Z}/6\mathbb{Z},$$

and we note that $\mathbb{Z}/6\mathbb{Z} \big/ \mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z}$, thus all the factor groups are simple.

We observe that the factor groups appearing in the first composition series are

$$\mathbb{Z}/2\mathbb{Z} \text{ and } \mathbb{Z}/3\mathbb{Z}$$

and the factor groups of the second composition series are

$$\mathbb{Z}/3\mathbb{Z} \text{ and } \mathbb{Z}/2\mathbb{Z}.$$

This is exactly the "factorization theorem" for groups, that (for finite groups) we can find a composition series, and any normal tower can be refined to an equivalent one. Unfortunately, unlike for integers, the data of the simple factor groups does not determine the original group uniquely.

Exercise 4.11.10. Show that S_3 has a composition series with factor groups $\mathbb{Z}/3\mathbb{Z}$ and $\mathbb{Z}/2\mathbb{Z}$. Conclude that even if G and H have equivalent composition series (same length and with the same set of simple factor groups), G and H need not be isomorphic.

Remark 4.11.11. For finite groups, one could at least hope that we could classify all of the finite simple groups, and perhaps make some progress in classifying all finite groups. The classification of finite simple groups was a major program of finite group theory since the mid-20th century, and was completed around 2004.

To prove our “factorization theorem”, we’ll need a few nice facts.

Lemma 4.11.12 (Butterfly Lemma, Zassenhaus)**.** *Let U, V be subgroups of a group G. Let u, v be normal subgroups of U and V, respectively. Then*

$$u(U \cap v) \text{ is normal in } u(U \cap V)$$
$$(u \cap V)v \text{ is normal in } (U \cap V)v$$

and the factor groups are isomorphic, i.e.

$$\frac{u(U \cap V)}{u(U \cap v)} \cong \frac{(U \cap V)v}{(u \cap V)v}.$$

Proof. The statements about normality follow directly from the facts that $u \trianglelefteq U$ and $v \trianglelefteq V$.

Note that

$$(U \cap V) \cap u(U \cap v) = (u \cap V)(U \cap v),$$

and

$$(U \cap V) \cap (u \cap V)v = (u \cap V)(U \cap v).$$

Indeed, let us show the first equality. Let $x \in (U \cap V) \cap u(U \cap v)$, then $x = ab$ with $a \in u$ and $b \in U \cap v$. Since $x, b \in V$ we have $a = xb^{-1} \in V$, so $a \in u \cap V$. Thus $x = ab \in (u \cap V)(U \cap v)$, whence

$$(U \cap V) \cap u(U \cap v) \subseteq (u \cap V)(U \cap v)$$

Conversely, let $x \in (u \cap V)(U \cap v)$, then $x = ab$ with $a \in u \cap V$ and $b \in U \cap v$. Since $a, b \in U$, we have $x \in U$, and similarly $x \in V$, whereby $x \in U \cap V$. As $a \in u$ and $b \in U \cap v$, we have $x \in u(U \cap v)$, and so $x \in (U \cap V) \cap u(U \cap v)$, whence

$$(U \cap V) \cap u(U \cap v) \supseteq (u \cap V)(U \cap v).$$

The other equality is similar.

Hence $(u \cap V)(U \cap v)$ is the intersection (in $u(U \cap V)$ or in $(U \cap V)v$) of a normal subgroup and of $U \cap V$ and hence normal in $U \cap V$.

The lattice of relevant subgroups looks like

$$\begin{tikzcd}
 & U \arrow[d, dash] & & V \arrow[d, dash] & \\
 & u(U \cap V) \arrow[dd, dash, "\trianglelefteq"'] \arrow[dr, dash] & & (U \cap V)v \arrow[dd, dash, "\trianglelefteq"] \arrow[dl, dash] & \\
 & & U \cap V \arrow[dd, dash, "\trianglelefteq"'] & & \\
 & u(U \cap v) \arrow[dl, dash] \arrow[dr, dash] & & (u \cap V)v \arrow[dl, dash] \arrow[dr, dash] & \\
u \arrow[dr, dash] & & (u \cap V)(U \cap v) \arrow[dl, dash] \arrow[dr, dash] & & v \arrow[dl, dash] \\
 & u \cap V & & U \cap v &
\end{tikzcd}$$

from which the lemma gets its name.

We consider the two top parallelograms and show that there are isomorphisms of quotient groups

$$\frac{u(U \cap V)}{u(U \cap v)} \cong \frac{U \cap V}{(u \cap V)(U \cap v)} \cong \frac{(U \cap V)v}{(u \cap V)v}.$$

Let $H = U \cap V$ and $N = u(U \cap v)$. The second isomorphism theorem (Theorem 4.7.24) gives

$$\frac{H}{H \cap N} \cong \frac{HN}{N},$$

i.e.

$$\frac{U \cap V}{(u \cap V)(U \cap v)} \cong \frac{u(U \cap V)}{u(U \cap v)}.$$

The other isomorphism is given similarly. □

Theorem 4.11.13 (Schreier)**.** *Let G be a group. Two normal towers of subgroups ending with the trivial subgroup have equivalent refinements.*

Proof. Suppose G has two normal towers

$$\{e\} = G_r \lhd \cdots \lhd G_2 \lhd G_1 = G,$$
$$\{e\} = H_s \lhd \cdots \lhd H_2 \lhd H_1 = G.$$

For each $i = 1, \dots, r-1$ and $j = 1, \dots, s$ we define

$$G_{i,j} = G_{i+1}(H_j \cap G_i).$$

Then $G_{i,s} = G_{i+1}$ and $G_{i,1} = G_i$, and we have a refinement of the first tower by filling in $G_{i,j}$ in between G_{i+1} and G_i

$$\{e\} \lhd G_{r-1,s-1} \lhd \cdots \lhd G_{r-1,1} = G_{r-1} \lhd \cdots \lhd G_{2,2} \lhd G_{2,1} = G_2 \lhd G_{1,s-1} \lhd \cdots \lhd G_{1,2} \lhd G_{1,1} = G.$$

Similarly, we define $H_{j,i} = H_{j+1}(G_i \cap H_j)$ for $j = 1, \dots, s-1$ and $i = 1, \dots, r$. This gives a similar refinement of the second tower.

By the butterfly lemma (Lemma 4.11.12), for $i = 1, \dots, r-1$ and $j = 1, \dots, s-1$, we have isomorphisms

$$G_{i,j} \Big/ G_{i,j+1} \cong H_{j,i} \Big/ H_{j,i+1}.$$

We view each of the refined towers as having $(r-1)(s-1)+1$ elements, namely the $G_{i,j}$ and $\{e\}$ and the $H_{j,i}$ and $\{e\}$. The isomorphisms of factor groups shows that these two towers are equivalent, as desired. □

Remark 4.11.14. Hans Zassenhaus was a professor at OSU from 1963, and after retiring in 1982, maintained an active presence at OSU, mentoring students and published some 40 papers while retired.

We're now ready to prove our "factorization theorem" for groups.

Theorem 4.11.15 (Jordan–Hölder Theorem)**.** *Let G be a group, and let*

$$\{e\} = G_r \lhd \cdots \lhd G_2 \lhd G_1 = G$$

be a normal tower such that each factor group $G_i \Big/ G_{i+1}$ is simple and non-trivial. Then any normal tower of G having the same properties is equivalent to this one.

Proof. Given any other normal tower satisfying the same properties, Theorem 4.11.13 shows that they have equivalent refinements. However, as both towers have simple factor groups, they cannot be refined further, and thus are equivalent. □

Remark 4.11.16. Thus given a finite group G, we look at some composition series for G which gives us a set of simple factor groups, determined completely by G. The natural next question to ask is, can we go backward? That is, given some set of simple groups $\{Q_i\}$, can we reconstruct a group that gives these factor groups? If there was a unique way to go backwards, then we could reconstruct the group and classify all groups in this way. Unfortunately, there are *many* ways to go back, for example $\mathbb{Z}/6\mathbb{Z}$ and S_3 give the same set of simple factor groups, but $\mathbb{Z}/6\mathbb{Z} \ncong S_3$.

In general, reconstructing a group from simple factors is known as the "extension problem" and *very difficult*, and currently unknown and thought ot be unsolvable. The idea would be to find some additional data that allows us to reconstruct the group uniquely from its simple factors. For example if A and B are simple groups, how many groups can we find with simple factors A and B? It turns out that this depends on the groups A and B in an intricate way, which is beyond the scope of this course.

A similar notion, that of *solvability* of groups is crucial in understanding the solutions to polynomial equations, which you may see in a course covering Galois Theory. For now, we just give a definition and some results which you can prove if interested.

Definition 4.11.17. A group G is called *solvable* if G has a normal tower

$$\{e\} = G_r \lhd \cdots \lhd G_2 \lhd G_1 = G$$

such that the factor groups ${G_i}\big/{G_{i+1}}$ are all abelian. Such a normal tower is called an *abelian tower.*

Proposition 4.11.18. *Let G be a finite group. An abelian tower of G admits a refinement with cyclic factor groups.*

Proposition 4.11.19. *Let G be a group and $N \trianglelefteq G$ a normal subgroup. Then G is solvable if and only if N and G/N are solvable.*

Theorem 4.11.20. *If $n \geq 5$, then S_n is not solvable.*

Proof. We shall first show that if H, N are two subgroups of S_n such that $N \trianglelefteq H$, if H contains every 3-cycle, and if H/N is abelian, then N contains every 3-cycle. Let i, j, k, r, s be five distinct integers in $[n]$ (here we use that $n \geq 5$), and let

$$\sigma = (ijk) \text{ and } \tau = (krs).$$

Then

$$\begin{aligned} \sigma\tau\sigma^{-1}\tau^{-1} &= (ijk)(krs)(kji)(srk) \\ &= (rki). \end{aligned}$$

Thus, as H/N is abelian, the commutator subgroup of H must be contained in N (by Proposition 4.7.18), whereby all the cycles (rki) lie in N, as desired.

Now suppose that S_n is solvable and we have a normal tower

$$\{e\} \lhd H_m \lhd \cdots H_2 \lhd H_1 = S_n$$

with ${H_i}\big/{H_{i+1}}$ abelian. Since S_n contains every 3-cycle, H_2 must contain every 3-cycle. Then so must $H - 3, \ldots, H_m$. However $H_m = \{e\}$, which is a contradiction. Thus S_n is not solvable for $n \geq 5$. □

Remark 4.11.21. It turns out that Theorem 4.11.20 is the reason why there is no "quadratic formula" for solving polynomial equations of degree ≥ 5.

Exercise 4.11.22. If $n \geq 3$, show that A_n is generated by the 3-cycles.

Exercise 4.11.23. If $n \geq 5$ show that A_n is simple. Show that the subgroup
$$\{(1), (12)(34), (13)(24), (14)(23)\} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$$
is a normal subgroup of A_4.

Remark 4.11.24. The next theorem is very difficult to prove, and was one of the major achievements of finite group theory leading to the classification of finite simple groups, proven in 1963. Its proof is around 250 pages.

Theorem 4.11.25 (Feit–Thompson)**.** *If G is a simple group of odd order, then $G \cong \mathbb{Z}/p\mathbb{Z}$ for some prime p.*

5. Rings

So far, we've seen how rich the theory of groups, sets with just one nice operation, can be. Here, we'll begin the study of *rings*, which have two nice operations, usually called addition and multiplication.

Definition 5.0.1. A *ring* is a set R with two binary operations
$$+ : R \times R \to R,\ (a, b) \mapsto a + b$$
$$\cdot : R \times R \to R,\ (a, b) \mapsto a \cdot b,$$
called *addition* and *multiplication*, respectively, satisfying the following conditions:

(**RI 1**) $(R, +)$ is an abelian group with identity 0 ;
(**RI 2**) $\cdot$ is associative, that is for all $a, b, c \in R$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;
(**RI 3**) the *distributive law* holds on both sides, i.e. for all $a, b, c \in R$
$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c \text{ ; and}$$
(**RI 4**) there exists an element $1 \in R$ such that for all $a \in R$, $1 \cdot a = a \cdot 1 = a$.

Remark 5.0.2. To simplify notation, we will usually write $a \cdot b$ as ab.

Example 5.0.3. Let's see some examples

- The integers $\mathbb{Z}$ with the usual addition and multiplication, is a ring.
- The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$, with the usual addition and multiplication, are rings.
- The quotient group $\mathbb{Z}/n\mathbb{Z}$, with multiplication defined modulo n is a ring.
- Let $M_{n\times n}(\mathbb{R})$ be the set of $n \times n$ matrices with entries in $\mathbb{R}$. Then $M_{n\times n}(\mathbb{R})$ is a ring with addition given by matrix addition and multiplication given by matrix multiplication.
- Let R be the set of continuous real-valued functions on the interval $[1, 0]$. We define addition and multiplication of functions as usual, namely $(f + g)(x) = f(x) + g(x)$, and $(fg)(x) = f(x)g(x)$. Then R is a ring.
- More generally, let X be a non-empty set, and let R be a ring. Let
$$M(X, R) = \{ \text{ maps } f : X \to R\}$$
be the set of maps from X to R. Since R is a ring, we can define addition and multiplication of functions by the usual rules
$$(f + g)(x) = f(x) + g(x), \text{ and } (fg)(x) = f(x)g(x),$$
and 1 given by the map sending all of X to $1 \in R$. Then $M(X, R)$ is a ring.

- Let A be an abelian group, and let $\mathrm{End}(A) = \{\text{group homomorphisms } \varphi : A \to A\}$. Then $\mathrm{End}(A)$ is a ring, with addition given by addition of functions and multiplication given by function composition.

Remark 5.0.4. In some books, rings are not required to have a multiplicative identity 1, and those rings with a 1 are called *unital rings*. Though perhaps rings without a multiplicative identity should instead be called *rngs*, without the identity "i".

Definition 5.0.5. A ring R is called *commutative* if $xy = yx$ for all $x, y \in R$. That is, if multiplication is commutative.

Remark 5.0.6. The study of commutative rings goes by the name "commutative algebra", and is connected to many prominent areas of modern mathematics.

Let us prove some basic rules of arithmetic in rings.

Proposition 5.0.7. *Let R be a ring, then 1 is unique.*

Proof. Same as for groups. □

Proposition 5.0.8. *Let R be a ring. Then $0x = 0 = x0$ for all $x \in R$*

Proof. We compute

$$0x + x = 0x + 1x = (0 + 1)x = (1)x = x,$$

and subtracting x from both sides gives

$$0x = 0x + x - x = x - x = 0.$$

Similarly, $x0 = 0$. □

Proposition 5.0.9. *Let R be a ring, then $(-1)x = -x = x(-1)$, i.e. $(-1)x$ is the additive inverse of x.*

Proof. We compute

$$(-1)x + x = (-1)x + 1x = (-1 + 1)x = (0)x = 0.$$

Similarly, $x(-1) = -x$. □

Proposition 5.0.10. *Let R be a ring, then $(-1)(-1) = 1$.*

Proof. We multiply the equation

$$1 + (-1) = 0$$

by (-1) and find

$$-1 + (-1)(-1) = 0,$$

adding 1 to both sides gives the result. □

Exercise 5.0.11. Let R be a ring. Show that for all $x, y \in R$, we have

$$(-x)y = -xy \text{ and } (-x)(-y) = xy.$$

Example 5.0.12. If R is a ring such that $1 = 0$, then for all $x \in R$, $1x = 0x = 0$, so $R = \{0\}$. This is called the *zero ring.*

Notation 5.0.13. From now on, we will assume that in a ring R, $1 \neq 0$, unless we specify that R is the zero ring.

Definition 5.0.14. Let R and S be rings. The *product of the rings R and S* is the Cartesian product $R \times S$ with addition and multiplication defined component wise.

As with groups, for a ring we have subsets that are themselves rings.

Definition 5.0.15. Let R be a ring. A subset $R' \subseteq R$ is called a *subring* if

(**SR 1**) $R' \leq R$ is a subgroup under addition,
(**SR 2**) $1 \in R'$, and
(**SR 3**) and for all $x, y \in R'$,

$$-x, x+y, \text{ and } xy \in R'.$$

That is, $1 \in R'$, and with the operations $+$ and $\cdot$, R' is a ring.

Example 5.0.16.

- $\mathbb{Z} \subset \mathbb{Q}$ is a subring.
- $2\mathbb{Z} \subset \mathbb{Z}$ is not a subring, since $1 \notin 2\mathbb{Z}$.
- The differentiable real-valued functions on $\mathbb{R}$ form a subring of the continuous functions.

There are many adjectives that describe some properties of rings, let us name a few.

Definition 5.0.17. Let R be a ring, we say that

- two elements $x, y \in R$ with $x, y \neq 0$ are called *zero divisors* if $xy = 0$.
- R is an *(integral) domain* if R is commutative and has no non-zero zero divisors. That is, if $xy = 0$, then $x = 0$ or $y = 0$.
- R is a *division ring* if the subset $R \setminus \{0\}$ forms a group under multiplication. That is, if every non-zero element has a multiplicative inverse.
- R is a *field* if R is a commutative division ring.

Example 5.0.18.

- The integers $\mathbb{Z}$ are a domain.
- $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ are fields.
- If F is a field, then the polynomials over F form a domain. For example $\mathbb{R}[x]$, the ring of polynomials in one variable with real coefficients.

Example 5.0.19. Let

$$\mathbb{H} := \{a + bi + cj + dk \mid a, b, c, d \in \mathbb{R}\},$$

called the *Hamiltonian quaternions*, where i, j, k satisfy the rules

$$i^2 = j^2 = k^2 = ijk = -1,$$

which were carved into Broom Bridge in Dublin by William Hamilton on October 16, 1843 when he realized these relations would make $\mathbb{H}$ into a ring. In fact, $\mathbb{H}$ is a division ring, though not a field. Indeed, we can check that

$$(a + bi + cj + dk)(a - bi - cj - dk) = a^2 + b^2 + b^c + d^2$$

thus the inverse of $a + bi + cj + dk \neq 0$ is given by

$$\frac{a - bi - cj - dk}{a^2 + b^2 + c^2 + d^2}.$$

Exercise 5.0.20. Show that a field is a domain.

Definition 5.0.21. Let R be a ring, we denote by

$$R^\times := \{x \in R \mid \text{ there exits } y \in R \text{ such that } xy = yx = 1\},$$

called the *units* of R.

Exercise 5.0.22. Let R be a ring. Show that $R^\times$ is a group under multiplication.

Example 5.0.23.

- $\mathbb{Z}^\times = \{1, -1\}$
- $\mathbb{Q}^\times = \mathbb{Q} \setminus \{0\}$
- $\mathbb{Z}/8\mathbb{Z}^\times = \{\overline{1}, \overline{3}, \overline{5}, \overline{7}\}$
- $\mathbb{Z}/n\mathbb{Z}^\times = \{\overline{k} \mid \gcd(k,n) = 1\}$, as we showed in Proposition 4.2.25.

Definition 5.0.24. Let R be a commutative ring. A *monomial* in the (commuting) variables $x_1, \ldots, x_n$ is an expression of the form

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$ is a tuple of non-negative integers denoting the exponents. A *polynomial* in the variables $x_1, \ldots, x_n$ over R is an expression of the form

$$\sum_{\alpha \in \mathbb{Z}^n_{\geq 0}} c_\alpha x^\alpha,$$

where $c_\alpha \in R$ is the coefficient of the monomial x^α, and $c_\alpha = 0$ for all but finitely many α, so the sum is finite. We denote by

$$R[x_1, \ldots, x_n]$$

the ring of polynomials over R int he variables $x_1, \ldots, x_n$, with polynomial addition and multiplication defined as usual by

$$\left(\sum_\alpha c_\alpha x^\alpha\right) + \left(\sum_\alpha d_\alpha x^\alpha\right) = \sum_\alpha (c_\alpha + d_\alpha) x^\alpha$$

$$\left(\sum_\alpha c_\alpha x^\alpha\right) \left(\sum_\alpha d_\alpha x^\alpha\right) = \sum_\alpha \left(\sum_{\alpha_1 + \alpha_2 = \alpha} c_{\alpha_1} d_{\alpha_2}\right) x^\alpha.$$

The identity is the polynomial 1.

Remark 5.0.25. Polynomial rings are a very important class of rings, leading to very interesting geometry, algebra, and number theory.

Example 5.0.26. The polynomial ring $\mathbb{R}[x]$ of polynomials in one variable, with coefficients in $\mathbb{R}$ are a great example. Some of its elements are $3x^2 + 2x + 1$, and $x^5 + x^4 + x^3 + x^2 + x + 1$.

Exercise 5.0.27. Prove the following proposition.

Proposition 5.0.28. *Let R be a domain, and let $a, b, c \in R$ with $a \neq 0$. If $ab = ac$, then $b = c$.*

Exercise 5.0.29. Let R be a finite domain. Show that R is a field. Hint: show that for $a \in R$, the map of sets $L_a : R \to R,\ x \mapsto ax$ is injective.

Exercise 5.0.30. Let R be a ring such that for all $x \in R$, $x^2 = x$. Show that R is commutative.

Exercise 5.0.31. Let R be a ring, and let

$$Z(R) \coloneqq \{a \in R \mid ax = xa \text{ for all } x \in R\},$$

called the *center of* R. Show that $Z(R)$ is a subring of R.

Exercise 5.0.32. Let $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q},\ (\sqrt{2})^2 = 2\}$, with addition and multiplication defined by adding like terms and distributing. Show that $\mathbb{Q}(\sqrt{2})$ is a field.

Exercise 5.0.33. Let $\mathbb{Z}[i] \coloneqq \{a + bi \mid a, b \in \mathbb{Z},\ i^2 = -1\}$. Show that $\mathbb{Z}[i]$ is a ring. Find the units of $\mathbb{Z}[i]$.

Exercise 5.0.34. Let p be a prime number. Define

$$\mathbb{Z}_{(p)} \coloneqq \left\{ \frac{a}{b} \in \mathbb{Q} \mid \gcd(b, p) = 1 \right\}.$$

Show that $\mathbb{Z}_{(p)}$ is a ring. This is called the *ring of integers localized at the prime* p.

5.1. **Ideals.** We've seen the notion of ideals of $\mathbb{Z}$. The notion of ideals of rings is similar, but we add one condition.

Definition 5.1.1. Let R be a ring. A *left ideal* of R is a subset $J \subseteq R$ such that

- if $x, y \in J$, then $x + y \in J$
- $0 \in J$
- if $x \in J$ and $a \in R$, then $ax \in J$.

We could similarly define a *right ideal* by the rule $xa \in J$, and a *two-sided ideal* by the rule $ax, xa \in J$.

Remark 5.1.2. Since $(-1)x = -x$, we see that and ideal J forms a subgroup of $(R, +)$. An alternative definition of an ideal is: a subgroup $J \leq R$ such that for all $a \in R$ and $x \in J$, $ax \in J$.

Example 5.1.3.

- $d\mathbb{Z} \subseteq \mathbb{Z}$ is an ideal.
- Let R be a ring, then R is an ideal, called the *unit ideal.*
- Let R be a ring, and $a \in R$ the set of elements of the form xa with $x \in R$ is an ideal, denoted by (a), or Ra, called the *principal left ideal generated by* a.

Exercise 5.1.4. Let R be a ring, and $a \in R$, show that $(a) = \{xa \mid x \in R\}$ is a left ideal of R. If R is commutative, show that (a) is a two-sided ideal.

Definition 5.1.5. Let R be a ring and $a_1, \ldots, a_n \in R$. The set

$$(a_1, \ldots, a_n) \coloneqq \{x_1 a_1 + \cdots x_n a_n \mid x_i \in R\}$$

is a left ideal. The elements $a_1, \ldots, a_n$ are called *generators* for this ideal, and $(a_1, \ldots, a_n)$ is called the *ideal generated by* $a_1, \ldots, a_n$. More generally, if $A \subset R$ is a subset, then the *(left) ideal generated* *by* A is the set

$$(A) := \left\{ \sum_{i=1}^{n} r_i a_i \;\middle|\; n \in \mathbb{Z}_{\geq 0}, r_1, \ldots, r_n \in R, a_1 \ldots, a_n \in A \right\}.$$

Proposition 5.1.6. *Let R be a ring and $a_1, \ldots, a_n \in R$, then $(a_1, \ldots, a_n)$ is a left ideal.*

Proof. Let $x_1, \ldots, x_n, y_1 \ldots, y_n, z \in R$. We see that

$$\begin{aligned}(x_1a_1 + \cdots + x_na_n) + (y_1a_1 + \cdots + y_na_n) &= x_1a_1 + y_1a_1 + \cdots + x_na_n + y_na_n \\ &= (x_1 + y_1)a_1 + \cdots (x_n + y_n)a_n,\end{aligned}$$

thus $(a_1, \ldots, a_n)$ is closed under sums. We also see that

$$z(x_1a_1 + \cdots + x_na_n) = (zx_1)a_1 + \cdots + (zx_n)a_n \in (a_1, \ldots, a_n),$$

so $(a_1, \ldots, a_n)$ is closed under multiplication be elements of R. Finally,

$$0 = 0a_1 + \cdots 0a_n,$$

so $0 \in (a_1, \ldots, a_n)$.

Therefore, $(a_1, \ldots, a_n)$ is a left ideal of R. □

Remark 5.1.7. The proof above also shows that for any subset $A \subseteq R$, (A) is a left ideal of R. If R is commutative, all of the operations above commute, and (A) is a two-sided ideal.

Example 5.1.8. In the ring $\mathbb{Z}$, we have $(a, b) = (\gcd(a, b))$.

In $\mathbb{Z}$, we've seen that every ideal is generated by one element. This is quite special.

Definition 5.1.9. Let R be a ring. We say an ideal $J \subseteq R$ is *principal* if $J = (a)$ for some $a \in R$.

Definition 5.1.10. We say a ring R is a *principal ideal domain* if R is commutative, a domain, and if for every ideal of R is principal. We abbreviate this by saying R is a PID.

Theorem 5.1.11. *$\mathbb{Z}$ is a PID.*

However, not every ring is a PID.

Example 5.1.12. The ideal $(x, y) \subset R[x, y]$ is not principal.

There are some operations we can do with ideals.

Definition 5.1.13. Let R be a ring and I, J be left ideals. We denote by IJ the ideal generated by products of elements ij with $i \in I$ and $j \in J$. That is,

$$IJ = \{i_1j_1 + \cdots + i_nj_n \;|\; n \in \mathbb{Z}_{\geq 0}, i_k \in I, j_k \in J\}.$$

You can check that if I, J, K are three left ideals, then $(IJ)K = I(JK)$.

Definition 5.1.14. Let R be a ring and I, J be left ideals. We define

$$I + J := \{i + j \;|\; i \in I, j \in J\}.$$

Then $I + J$ is a left ideal of R. You can check that if I, J, K are three left ideals, then $(I + J) + K = I + (J + K)$.

Exercise 5.1.15. Prove the following proposition.

Proposition 5.1.16. *Let R be a ring. If I, J are left ideals, then $I \cap J$ is a left ideal. If I, J are right ideals, then $I \cap J$ is a right ideal. If I, J are two-sided ideals, then $I \cap J$ is a two-sided ideal.*

Exercise 5.1.17. Let $I, J \subseteq R$ be two-sided ideals. Prove the following:

- Then $I + J$ is the smallest ideal containing both I and J.
- $IJ \subseteq I \cap J$.

Exercise 5.1.18. Prove the following proposition.

Proposition 5.1.19. *Show that if F is a field, then the only ideals of F are $\{0\}$ or F.*

5.2. **Ring Homomorphisms.** Just like with groups, we want to define maps between rings that play nicely with the ring structure.

Definition 5.2.1. Let R, S be rings with multiplicative identities 1 and $1'$, respectively. A map

$$f : R \to S$$

is called a *ring homomorphism* if

- $f(1) = 1$, and

for all $x, y \in R$,

- $f(x + y) = f(x) + f(y)$, and
- $f(xy) = f(x)f(y)$.

Example 5.2.2. The map $\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}, k \mapsto \overline{k}$ is a ring homomorphism. All this says is that $1 \equiv 1 \mod n$ and if $a \equiv a' \mod n$ and $b \equiv b' \mod n$, then $a + b \equiv a' + b' \mod n$ and $ab \equiv a'b' \mod n$.

Remark 5.2.3. If $f : R \to S$ is a ring homomorphism, since $f(x + y) = f(x) + f(y)$, f is a group homomorphism of the abelian groups $(R, +)$ and $(S, +)$. In particular, $f(0) = 0$, and $f(-x) = -f(x)$.

Proposition 5.2.4. *Let $R, S_1, \dots, S_n$ be rings. Then $f : R \mapsto S_1 \times \cdots \times S_n$ is a ring homomorphism if and only if each coordinate map $f_i : R \to S_i$ is a ring homomorphism.*

Proposition 5.2.5. *Let R be a ring. There is a unique ring homomorphism $\mathbb{Z} \to R$*

Proof. Let $a \in R$. Since R is an abelian group under addition, we can define na for every positive integer n as usual by

$$na = \underbrace{a + \cdots + a}_{n \text{ times}},$$

and if $n < 0$, we define

$$na = -(-na).$$

Thus we get a map

$$f : \mathbb{Z} \to R, n \mapsto n1.$$

This is a homomorphism of abelian groups, and we note that clearly $f(1) = 1$ and

$$f(nm) = (nm)1 = n(m1) = (n1)(m1) = f(n)f(m),$$

hence $f : \mathbb{Z} \to R$ is a ring homomorphism.

Let $g : \mathbb{Z} \to R$ be a ring homomorphism. Then since $g(1) = 1$, for $n \geq 0$ we have

$$g(n) = g(\underbrace{1 + 1 + \cdots + 1}_{n \text{ times}}) = \underbrace{g(1) + g(1) + \cdots + g(1)}_{n \text{ times}} = ng(1),$$

and for $n < 0$, letting $k = -n$ we have

$$g(n) = g(-(-n)) = -g(k) = -kg(1) = ng(1).$$

Thus the ring homomorphism $g : \mathbb{Z} \to R$ is completely determined by $g(n) = ng(1) = n1$, and so $g = f$. Therefore the only ring homomorphism $f : \mathbb{Z} \to R$ is $f(n) = n1$. □

Definition 5.2.6. Let $f : R \to S$ be a ring homomorphism, the *kernel of f* is

$$\ker f := \{r \in R \mid f(r) = 0\}.$$

Example 5.2.7. The ring homomorphism $f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ has kernel $n\mathbb{Z}$.

Proposition 5.2.8. *Let $f : R \to S$ be a ring homomorphism. Then $\ker f \subset R$ is a two-sided ideal.*

Proof. Suppose $x, y \in \ker f$. Then $f(x+y) = f(x)+f(y) = 0+0 = 0$, thus $x+y \in \ker f$. We have $f(0) = 0$, thus $0 \in \ker f$. And if $a \in R$, then

$$f(ax) = f(a)f(x) = f(a)0 = 0 = 0f(a) = f(x)f(a) = f(xa),$$

whereby $ax, xa \in \ker f$. Therefore $\ker f$ is a two-sided ideal of R. □

Proposition 5.2.9. *If $fR \to S$ is a ring homomorphism, them $\mathrm{im} f$ is a subring of S.*

Proof. We check the conditions to be a subring. As $f : (R, +) \to (S, +)$ is a group homomorphism, $\mathrm{im} f$ is a subgroup of S. Since $f(1) = 1'$, $1' \in \mathrm{im} f$. And if $x' = f(x), y' = f(y) \in \mathrm{im} f$, then

$$-x' = f(-x) \in \mathrm{im} f,$$

$$x' + y' = f(x) + f(y) = f(x + y) \in \mathrm{im} f,$$

and

$$x'y' = f(x)f(y) = f(xy) \in \mathrm{im} f.$$

Thus $\mathrm{im} f \subseteq S$ is a subring of S. □

Exercise 5.2.10. Prove the following proposition.

Proposition 5.2.11. *Let $f : R \to S$ be a ring homomorphism. Show that f is injective if and only if* $\ker f = 0$.

Exercise 5.2.12. Prove the following proposition.

Proposition 5.2.13. *Let K be a field. Then every ring homomorphism $f : K \to R$ is injective.*

Just like with groups, a ring homomorphism with an inverse that is a ring homomorphism is called an *isomorphism*.

Definition 5.2.14. A ring homomorphism $f : R \to S$ is called an *isomorphism* if there is a ring homomorphism $g : S \to R$ such that

$$f \circ g = \mathrm{id}_S \text{ and } g \circ f = \mathrm{id}_R.$$

Exercise 5.2.15. Prove the following proposition.

Proposition 5.2.16. *Let $f : R \to S$ be a ring homomorphism. Then f is an isomorphism if and only if f is bijective.*

Exercise 5.2.17. Prove the following proposition.

Proposition 5.2.18. *Let $f : R \to S$ be a ring homomorphism. If I is an ideal of R, then $f(I)$ is not necessarily an ideal of S. If f is surjective, show that $f(I)$ is an ideal of S.*

Remark 5.2.19. The notion of *isomorphism* for groups and rings followed the same recipe, an isomorphism is a map that preserves the structure of the objects we care about and has an inverse that also preserves that structure. In general, if X and Y are two objects with some structure we care about, we call a map $f : X \to Y$ a *morphism* if f preserves that structure, and an *isomorphism* if there is a morphism $g : Y \to X$ such that

$$f \circ g = \mathrm{id}_X \text{ and } g \circ f = \mathrm{id}_Y.$$

For this to make sense, all we really need is composition of morphisms to be associative, and each object has an *identity morphism.*

In general, if we consider all objects with some structure, and all the morphisms between them, this gives a *category.* We've seen a few categories already

- The category of sets, **Set** where the objects are sets and the morphisms are just maps of sets.
- The category **Gps** of groups, where the objects are groups and the morphisms are group homomorphisms.
- The category **AbGps** of abelian groups, where the objects are abelian groups and the morphisms are group homomorphisms.
- The category **Ring** of rings, where the objects are rings and the morphisms are ring homomorphisms.

Many ideas in abstract algebra and modern mathematics can be phrased in terms of categories, abstracting away everything but the maps between objects, providing a clean and very general formulation of properties of the objects without getting too preoccupied with the intricacies of the objects themselves.

For example, an *automorphism* is an isomorphism of an object with itself, i.e. an isomorphism $f : X \to X$. It then follows directly from the definition that for an object X in some category, the set of automorphisms of X, $\mathrm{Aut}(X)$, forms a group. For instance, the automorphism group of the set $\{1, \dots, n\}$ is S_n. In Math 4581, you'll see Galois Theory, which studies the automorphism groups of fields.

5.3. **Quotient Rings, Isomorphism Theorems.** Just like for groups, we can quotient by kernels, and everything we can quotient by is a kernel of some (ring) homomorphism. We'll make this precise in this section. Just like with groups, we need to make sure that the operations are well-defined when we quotient.

Recall that if $I \subseteq R$ is an ideal, then I is a subgroup of R considered as an additive group. Since addition is commutative, $I \leq (R, +)$ is in particular a *normal subgroup*, and so we can form the quotient group R/I.

Example 5.3.1. The group $\mathbb{Z}/n\mathbb{Z}$ is also a ring! Let

$$\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z},\ a \mapsto \overline{a}$$

be the quotient map. Addition is defined as usual $\overline{a} + \overline{b} = \overline{a + b}$, and multiplication is defined by

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

The fact that multiplication is well-defined comes from the fact that multiplication plays nicely with the operation $\mod n$, which really only needs the fact that $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

Example 5.3.2. Let R be a ring and $I \subset R$ a two-sided ideal. If $x, y \in R$, we say that $x \equiv y \mod I$ if $x - y \in I$. Then we have

- $x \equiv x \mod I$
- If $x \equiv y \mod I$ and $y \equiv z \mod I$, then $x \equiv z \mod I$.
- If $x \equiv y \mod I$, then $y \equiv x \mod I$.

So being equivalent $\mod I$ is an equivalence relation! Furthermore, we have

- If $x \equiv y \mod I$ and $z \in R$, then $xz \equiv yz \mod I$ and $zx \equiv zy \mod I$.
- If $x \equiv y \mod I$ and $x' \equiv y' \mod I$, then $xx' \equiv yy' \mod I$ and $x + x' \equiv y + y' \mod I$.

  Proof. We give a proof of the last property, and leave the proofs of the remaining ones to the interested reader.

  So suppose that $x \equiv x' \mod I$ and $y \equiv y' \mod I$, then we can write $x = y + z$ and $x' = y' + z'$ for some $z, z' \in I$. We compute

$$xx' = (y+z)(y'+z') = yy' + \underbrace{\underbrace{zy'}_{\in I} + \underbrace{yz'}_{\in I} + \underbrace{zz'}_{\in I}}_{\in I}$$

  where $zy', yz', zz' \in I$ because I is a two-sided ideal. Thus $xx' - yy' \in I$ and so $xx' \equiv yy' \mod I$. Similarly,

$$x + x' = y + z + y' + z' = y + y' + \underbrace{z + z'}_{\in I},$$

  hence $x + x' \equiv y + y' \mod I$. □

Proposition 5.3.3. *Let R be a ring and $I \subset R$ a two-sided ideal. The quotient group R/I is a ring when equipped with the multiplication*

$$(r + I)(s + I) = rs + I.$$

Moreover, the map

$$\pi : R \to R/I$$
$$r \mapsto r + I$$

is a ring homomorphism with kernel I.

Proof. We must check that these operations are well-defined and they make R/I into a ring.

The fact that these operations are well-defined is the content of Example 5.3.2.

It remains to show that R/I satisfies the ring axioms.

For **RI 1**, since I is a normal subgroup of $(R, +)$, $(R/I, +)$ is an abelian group. The remaining ring axioms all follow from the corresponding axioms of R.

Let $x, y, z \in R$ be representatives of the cosets $x + I, y + I, z + I \in R/I$.

We leave checking **RI 2** to the reader.

For **RI 3**, observe that $x(y + z)$ is a representative of $(x + I)((y + I) + (z + I))$, but since

$$x(y + z) = xy + xz,$$

and xy is a representative of $xy + I$ and xz is a representative of $xz + I$, by definition we have

$$(x + I)((y + I) + (z + I)) = (xy + I) + (xz + I).$$

Similarly one proves that

$$((x + I) + (y + I))(z + I) = (xz + I) + (yz + I).$$

For **RI 4**, if 1 denotes the multiplicative identity of R, then $1x = x = x1$ and so $(1 + I)(x + I) = (x + I) = (x + I)(1 + I)$.

It remains to show that $\pi : R \to R/I$ is a ring homomorphism with kernel I. Since π is just the quotient map for groups, and $\pi(1) = 1 + I$, it remains to show that π respects multiplication. Let $r, s \in R$, then

$$\begin{aligned}\pi(rs) &= rs + I \\ &= (r+I)(s+I) \quad \text{by definition} \\ &= \pi(r)\pi(s).\end{aligned}$$

Finally, since π is the quotient map, we clearly have $I \subseteq \ker \pi$. For the other containment, let $r \in \ker \pi$, then $r + I = 0 + I = I$, hence $r \in I$, thus $\ker \pi \subseteq I$. Therefore $\ker \pi = I$, as claimed. □

Definition 5.3.4. Let R be a ring and $I \subset R$ be a two-sided ideal. The ring R/I is called the *quotient ring* of R by I, or the *quotient ring of R modulo I*, and $\pi : R \to R/I$ the *canonical quotient ring homomorphism.*

Example 5.3.5. $\mathbb{Z}/n\mathbb{Z}$ is just the quotient ring of $\mathbb{Z}$ by the ideal $n\mathbb{Z}$.

Exercise 5.3.6. Let $n \geq 2$ be an integer. Prove the following facts about the rings $\mathbb{Z}/n\mathbb{Z}$.
(1) Show that $\mathbb{Z}/n\mathbb{Z}$ is a domain if and only if n is prime.
(2) Let p be a prime number. Show that $\mathbb{Z}/p\mathbb{Z}$ is a field.
(3) If a is an integer such that $a \neq 0 \mod p$, show that $a^{p-1} \equiv 1 \mod p$.

Just as with groups and normal subgroups, every two-sided ideal is the kernel of a ring homomorphism, namely π, and every kernel of a ring homomorphism is a two-sided ideal.

Theorem 5.3.7 (Maps from quotients)**.** *Let $I \subset R$ be a twp-sided ideal with quotient map $\pi : R \to R/I$. Let $\varphi : R \to S$ be a ring homomorphism such that $I \subseteq \ker \varphi$. Then there is a unique ring homomorphism $\overline{\varphi} : R/I \to S$ such that $\overline{\varphi} \circ \pi = \varphi$.*

We symbolize this as saying that the diagram

$$\begin{tikzcd} R \arrow[d, "\pi"'] \arrow[dr, "\varphi"] & \\ R/I \arrow[r, dotted, "\exists!\ \overline{\varphi}"'] & S \end{tikzcd}$$

commutes, i.e. $\overline{\varphi} \circ \pi = \varphi$. We say that $\overline{\varphi}$ is *induced by* φ.

Proof of Theorem 5.3.7. The map $\overline{\varphi} : R/I \to S$ is the map induced by φ as a map of abelian groups, from Theorem 4.6.15.

Recall that $\overline{\varphi} : R/I \to S$ is defined by

$$\overline{\varphi}(r + I) = \varphi(r).$$

It remains to check that $\overline{\varphi}$ is in fact a ring homomorphism.

As $\varphi(1) = 1$, we have $\overline{\varphi}(1 + I) = \varphi(1) = 1$.

Let $r + I, s + I \in R/I$, then

$$\overline{\varphi}((r+I)(s+I)) = \overline{\varphi}(rs + I) = \varphi(rs) = \varphi(r)\varphi(s) = \overline{\varphi}(r+I)\overline{\varphi}(s+I),$$

and so $\overline{\varphi} : R/I \to S$ is a ring homomorphism. □

Theorem 5.3.8 (First Isomorphism Theorem for Rings)**.** *Let $\varphi : R \to S$ be a ring homomorphism, and let $\pi : R \to R/\ker\varphi$ be the quotient homomorphism. Then $\overline{\varphi} : R/\ker\varphi \to S$ gives a ring isomorphism*

$$R/\ker\varphi \cong \mathrm{im}\varphi.$$

The proof is virtually the same as for groups and once could give a direct proof. We'll instead give one using the First Isomorphism Theorem for groups (Theorem 4.7.1).

Proof. Let $I = \ker\varphi$. By Theorem 5.3.7, there is a unique ring homomorphism $\overline{\varphi} : R/I \to S$ defined by $\overline{\varphi}(r+I) = \varphi(r)$. By the first isomorphism theorem for groups we have $\overline{\varphi} : R/I \cong \mathrm{im}\varphi$ as groups. Since $\overline{\varphi}$ is also a ring homomorphism, this is also an isomorphism of rings. □

Example 5.3.9. Let $\mathbb{C}[x]$ be the polynomial ring in one variable. Then for $\alpha \in \mathbb{C}$, we have a ring homomorphism

$$\mathrm{ev}_\alpha : \mathbb{C}[x] \to \mathbb{C}, \quad f(x) \mapsto f(\alpha).$$

That this is a ring homomorphism is just the statement that $(f+g)(\alpha) = f(\alpha)+g(\alpha)$, the constant polynomial 1 has the value 1 at α and $(fg)(\alpha) = f(\alpha)g(\alpha)$.

The evaluation map ev_α is clearly surjective, as for $z \in \mathbb{C}$, the constant polynomial $z \mapsto z$. Thus, by the first isomorphism theorem $\mathbb{C}[x]/\ker\mathrm{ev}_\alpha \cong \mathbb{C}$.

In fact, although this is slightly more difficult to prove, $\ker\mathrm{ev}_\alpha$ is the ideal $(x-\alpha)$ generated by the polynomial $x - \alpha$.

We also have a correspondence theorem, for ideals of the quotient R/I.

Theorem 5.3.10. *Let R be a ring and I a two-sided ideal of R. Then the map*

$$\{\textit{subrings } I \subseteq S \subseteq R\} \to \{\textit{subrings } S' \subseteq R/I\}$$
$$S \mapsto S/I$$

is a bijection. Moreover, it restricts to a bijection between two-sided ideals of R that contain I, and ideals of R/I.

Proof. The proof is virtually the same as for subgroups of G/N, and is left to the interested reader. □

Exercise 5.3.11. Prove the following proposition.

Proposition 5.3.12. *Let $I \subset R$ be an ideal, and let $\pi : R \to R/I$ be the quotient map. Suppose that $J \subset R/I$ is an ideal. Then $\pi^{-1}(J) = \{j \in R \mid \pi(j) \in J\}$ is an ideal of R containing I.*

Theorem 5.3.13 (Third Isomorphism Theorem for Rings)**.** *Let R be a ring, and $J \subset I \subset R$ be ideals. Then*

$$R/I \cong \frac{R/J}{I/J}.$$

Proof. The proof is virtually the same as for groups, one defines a surjective ring homomorphism $R \mapsto \frac{R/I}{I/J}$ by composing the natural quotient ring homomorphisms, and the shows that the kernel is I. □

For the second isomorphism theorem, we will need some notation.

Proposition 5.3.14. *Let $I \subseteq R$ be a subring of R, and let $J \subseteq R$ be a two-sided ideal. Then*

$$I + J := \{a + b \mid a \in I, b \in J\}$$

is a subring of R, J is an ideal of $I + J$, and $I \cap J$ is an ideal of I. If I is a two-sided ideal of R, then $I + J$ is a two-sided ideal of R.

Proof. The proof is just verifying the definitions.

We see that $I + J$ is a subgroup of R, as both I and J are subgroups. Since $1 \in I$ and $0 \in J$, $1 + 0 \in I + J$. And for all $(a + b), (a' + b') \in I + J$, we have

$$(a + b) + (a' + b') = (a + a') + (b + b') \in I + J,$$

and

$$(a + b)(a' + b') = \underbrace{aa'}_{\in I} + \underbrace{ba' + ab' + bb'}_{\in J} \in I + J.$$

Thus $I + J$ is a subring of R.

For $b \in J$ and $(a' + b') \in I + J$, $(a' + b')b = a'b + b'b \in J$ as J is an ideal of R. Thus J is an ideal of $I + J$.

If $b \in I \cap J$, then for all $a \in I$, $ab \in I$ as I is a subring, and $ab \in J$ as J is an ideal. Thus $I \cap J$ is an ideal of I.

Finally, if I is a two-sided ideal of R, then for any $x \in R$, we have $x(a + b) = xa + xb \in I + J$, and similarly $(a + b)x = ax + bx \in I + J$. Thus $I + J$ is a two-sided ideal of R. □

Theorem 5.3.15 (Second Isomorphism Theorem for Rings)**.** *Let R be a ring, $I \subseteq R$ a subring, and $J \subseteq R$ a two-sided ideal of R. Then*

$$(I + J)\big/J \cong I\Big/(I \cap J).$$

Proof. The proof is practically the same as for groups. In fact, there is an isomorphism of groups

$$(I + J)\big/J \cong I\Big/(I \cap J)$$

from the Second Isomorphism Theorem for groups (Theorem 4.7.24). Since the quotient maps are all ring homomorphisms, the map is also a ring isomorphism. □

Definition 5.3.16. Let R be a commutative ring. We say two ideals $I, J \subseteq R$ are *relatively prime* if $I + J = R$.

Exercise 5.3.17. Let R be a commutative ring. If I, J are two relatively prime ideals, then $IJ = I \cap J$.

Exercise 5.3.18. Prove the following Proposition.

Proposition 5.3.19. *Let R be a commutative ring. If I, J are relatively prime ideals of R, then given $a, b \in R$ there exists an $x \in R$ such that $x - a \in I$ and $x - b \in J$. That is, there is some $x \in R$ such that*

$$x \equiv a \pmod{I} \text{ and } x \equiv b \pmod{J}.$$

More generally, the map

$$\varphi : R \to R/I \times R/J$$

is a surjective ring homomorphism with kernel $I \cap J$, hence

$$R/(I \cap J) \cong R/I \times R/J.$$

Even more generally, if $I_1, \dots, I_n$ are ideals that are pairwise relatively prime (I_k and I_m are coprime for any $k \neq m$) then there is a natural ring isomorphism

$$R/(I_1 \cdots I_n) \cong R/I_1 \times \cdots \times R/I_n.$$

5.4. **Prime and Maximal Ideals.**

Remark 5.4.1. From now on, we will assume all the rings we work with are commutative, and have a 1.

There are a few kinds of ideals with very nice properties, which are very useful in the abstract study of rings.

Definition 5.4.2. Let R be a (commutative) ring. An ideal $\mathfrak{m} \subset R$ is called *maximal* if $\mathfrak{m} \neq R$ and there is no ideal J such that $\mathfrak{m} \subset J \subset R$ and $J \neq R$ and $J \neq \mathfrak{m}$. That is, $\mathfrak{m}$ is a maximal ideal if whenever J is any ideal such that $\mathfrak{m} \subseteq J \subseteq R$, then either $J = \mathfrak{m}$ or $J = R$.

The first main result about maximal ideals is that every ideal is contained in some maximal ideal. However, this requires Zorn's lemma, and we briefly give some background before proving the result.

Definition 5.4.3. A set $\mathcal{P}$ is called *partially ordered* if there is a relation $\leq$ on $\mathcal{P}$ such that for all $x, y, z \in \mathcal{P}$ we have

- $x \leq x$,
- if $x \leq y$ and $y \leq x$, then $x = y$, and
- if $x \leq y$ and $y \leq z$, then $x \leq z$.

The relation $\leq$ is called a *partial order* as it need not be the case that any two elements of $\mathcal{P}$ are comparable. A partially ordered set is sometimes called a "poset".

Example 5.4.4. Let X be a set. Let $\mathcal{P}$ be the set of subsets of X, and define a partial order by inclusion. That is, for subsets $A, B \subset X$ we say $A \leq B$ if $A \subseteq B$.

Definition 5.4.5. Let $\mathcal{P}$ be a partially ordered set. A *chain* is a totally ordered subset $\mathcal{C} \subset \mathcal{P}$, that is, for any $x, y \in \mathcal{C}$ we have $x \leq y$ or $y \leq x$.

Remark 5.4.6. Posets are fundamental, sadly we don't have time to delve into properties or examples of posets. One example is the natural numbers ordered by size. Another example is subsets of $\mathbb{R}$ ordered by inclusion.

Lemma 5.4.7 (Zorn's Lemma)**.** *Let $\mathcal{P}$ be a partially ordered set. Suppose $\mathcal{P}$ is nonempty and every chain in $\mathcal{P}$ has an upper bound, i.e., for a chain $\mathcal{C} \subset \mathcal{P}$ there is an element $x \in \mathcal{P}$ such that $a \leq x$ for all $a \in \mathcal{C}$. Then $\mathcal{P}$ has at least one maximal element, i.e., an element $m \in \mathcal{P}$ such that $x \leq m$ for all $x \in \mathcal{P}$.*

Remark 5.4.8. Lemma 5.4.7 is not really a lemma, it is an axiom of set theory, which is equivalent to the Axiom of Choice (which says that there is a way to pick an element out of any set). In fact, the Axiom of Choice and Zorn's Lemma are both equivalent to the Well-Ordering Principle, tot he statement "every vector space has a basis", and to the statement "every non-zero ring with a 1 has a maximal ideal". So in some sense, we are assuming what we want to prove. On the other hand, if we want to prove maximal ideals exist, then we have to assume it.

For some nice rings, the existence of maximal ideals does not need Zorn's lemma, in particular for *Noetherian rings*, named in honor of Emmy Noether, which are rings where every ideal can be generated by a finite set.

Theorem 5.4.9. *Let R be a (commutative) ring. Then every proper ideal is contained in some maximal ideal.*

Proof. Let $\mathcal{S}$ be the set of proper ideals of R that contain I. Since $I \subseteq I$, $\mathcal{S}$ is non-empty, and $\mathcal{S}$ is a poset under inclusion. Moreover, for any chain $\mathcal{C} \subseteq \mathcal{S}$, let

$$J = \bigcup_{A \in \mathcal{C}} A.$$

We claim that J is an ideal. Certainly, J is non-empty and $0 \in J$. For any $a, b \in J$, there are ideals A, B in C such that $a \in A$ and $b \in B$. Since $\mathcal{C}$ is a chain, we can assume $A \subseteq B$, and so $a + b \in J$. Since each A is an ideal, then for any $r \in R$ and $a \in J$, $ra \in J$. Thus J is an ideal of R.

Since each $A \in \mathcal{C}$ is a proper ideal, $1 \notin A$ for all $A \in \mathcal{C}$, and so $1 \notin J$, so J is a proper ideal. And since $I \subseteq A$ for all $A \in \mathcal{C}$, $I \subseteq J$.

Thus every chain $\mathcal{C} \subseteq \mathcal{S}$ has an upper bound, and by Zorn's Lemma, $\mathcal{S}$ has a maximal element, which is therefore a maximal (proper) ideal containing I. □

Lemma 5.4.10. *A (commutative) ring R is a field if and only if R has no ideals other than* 0 *and R.*

Proof. If R is a field, then the only ideals are 0 and R, as proven in Proposition 5.1.19. Conversely, suppose that R has only the ideals 0 and R. Let $a \in R \setminus \{0\}$ and consider the ideal (a). Since $a \neq 0$, $(a) \neq 0$, this $(a) = R$. Hence $1 = ab \in (a)$, so $b = a^{-1}$. Thus every non-zero element of R is invertible, and R is a field. □

Theorem 5.4.11. *Let R be a (commutative) ring. An ideal $\mathfrak{m}$ of R is maximal if and only if $R/\mathfrak{m}$ is a field.*

Proof. Let $\mathfrak{m} \subset R$ be an ideal and let $\pi : R \to R/\mathfrak{m}$ be the quotient map.

Suppose that $\mathfrak{m}$ is maximal. Let $J \subset R/\mathfrak{m}$ be an ideal. Then by Proposition 5.3.12, $\pi^{-1}(\mathfrak{m})$ is an ideal of R containing $\mathfrak{m}$, hence must be either $\mathfrak{m}$ or R. Thus $J = 0$ or $J = R/\mathfrak{m}$, respectively. Hence $R/\mathfrak{m}$ is a field.

Conversely, suppose that $R/\mathfrak{m}$ is a field. Thus $\mathfrak{m} \neq R$. Suppose that J is an ideal of R containing $\mathfrak{m}$. We show that if $J \neq \mathfrak{m}$, then $J = R$. Let $a \in J \setminus \mathfrak{m}$. Then $\pi(a) \neq 0$, hence as $R/\mathfrak{m}$ is a field and π is surjective, there is a multiplicative inverse $\pi(a)^{-1} = \pi(b) \in R/\mathfrak{m}$. Thus $1 = \pi(ab)$, hence $ab \in 1 + \mathfrak{m}$. So there is $r \in \mathfrak{m}$ such that $ab = 1 + r$. Note that $a, b, r \in J$, hence $1 = ab - r \in J$. Thus $J = R$, as desired. Therefore $\mathfrak{m}$ is maximal. □

Example 5.4.12. The maximal ideals of $\mathbb{Z}$ are $p\mathbb{Z}$ where p is prime (as $n\mathbb{Z} \leq d\mathbb{Z}$ if and only if $d \mid n$.) Thus $\mathbb{Z}/p\mathbb{Z}$ is a field if and only if p is prime.

Quotienting by a maximal ideal gives a field. Domains share many properties of fields, so we may ask if there are kinds of ideals that correspond to domains. It turns out there are!

Suppose that I is an ideal of R and that R/I is a domain. Then for any $a + I, b + I \in R/I$, if $ab + I = 0 + I$, then $a + I = 0 + I$ or $b + I = 0 + I$. That is, if R/I is a domain, then I satisfies the property that if $ab \in I$, then $a \in I$ or $b \in I$.

Definition 5.4.13. Let R be a ring. An ideal I of R is *prime* if for all $a, b \in R$, if $ab \in I$, then $a \in I$ or $b \in I$.

Theorem 5.4.14. *Let R be a ring and I an ideal of R. Then I is prime if and only if R/I is a domain.*

Proof. Let $\pi : R \to R/I$ be the quotient map.

Suppose that I is prime, and let $\pi(a), \pi(b) \in R/I$, since π is surjective. Suppose that $\pi(a)\pi(b) = 0$. Then $\pi(a)\pi(b) = \pi(ab)$, thus $ab \in I$. As I is prime, $a \in I$ or $b \in I$. Thus $\pi(a) = 0$ or $\pi(b) = 0$. Therefore R/I is a domain.

Conversely, suppose that R/I is a domain, and let $a, b \in R$ such that $ab \in I$. Then $0 = \pi(ab) = \pi(a)\pi(b)$, thus $\pi(a) = 0$ or $\pi(b) = 0$. Therefore $a \in I$ or $b \in I$, whereby I is prime. □

Corollary 5.4.15. *R is a domain if and only if $\{0\}$ is a prime ideal.*

Example 5.4.16. Consider the ring $\mathbb{C}[x]$. The ideal (x) is prime, in fact it is maximal. However, the ideal (x^2) is not prime.

Example 5.4.17. The ideal (xy) of $\mathbb{C}[x, y]$ is not prime. Indeed, $xy \in (xy)$, but $x \notin (xy)$ and $y \notin (xy)$.

Since a field is a domain, we obtain the following corollary.

Corollary 5.4.18. *Any maximal ideal is prime.*

Proof. Let $\mathfrak{m}$ be a maximal ideal of R. Then $R/\mathfrak{m}$ is a field, in particular a domain. Then $\mathfrak{m}$ is prime. □

Example 5.4.19. Let R be a commutative ring. The ring of *dual numbers* is $R[x]/(x^2)$. The elements of $R[x]/(x^2)$ are of the form $a + b\varepsilon$ with $a, b \in R$ and $\varepsilon^2 = 0$. This is not a domain, since (x^2) is not prime. The ring of dual numbers is useful in defining the notion of derivatives in algebraic geometry.

Prime ideals play nicely with ring homomorphisms.

Proposition 5.4.20. *Let $\varphi : R \to S$ be a ring homomorphism, and let $I \subset S$ be a prime ideal. Then $\varphi^{-1}(I)$ is a prime ideal of R.*

Proof. Let $I \subset S$ be a prime ideal. By Proposition 5.3.12, $\varphi^{-1}(I) \subset R$ is an ideal. Let $a, b \in R$ such that $ab \in \varphi^{-1}(I)$. Then $\varphi(ab) = \varphi(a)\varphi(b) \in I$, and as I is prime, $\varphi(a) \in I$ or $\varphi(b) \in I$. Thus $a \in \varphi^{-1}(I)$ or $b \in \varphi^{-1}(I)$. □

Definition 5.4.21. Let R be a domain. The *characteristic* of R, $\operatorname{char}(R)$, is the smallest integer $p \in \mathbb{Z}_{>0}$ such that $p = 0$ in R, if such a p exists, and we say R has *characteristic* p. If no such p exists, we say R has *characteristic zero.*

Example 5.4.22. The rings $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ all have characteristic zero. The field $\mathbb{Z}/p\mathbb{Z}$ has characteristic p.

Proposition 5.4.23. *Let R be a domain. Then* $\operatorname{char}(R)$ *is zero or a prime.*

Proof. We show that if R has positive characteristic p, then p is a prime number. Indeed, by Proposition 5.2.5, there is a unique ring homomorphism $f : \mathbb{Z} \to R$. Since R is a domain, $\{0\}$ is a prime ideal, and $\ker f = f^{-1}(\{0\}) \subseteq \mathbb{Z}$ is a prime ideal, thus $\ker f = p\mathbb{Z}$ for some prime number p. □

Remark 5.4.24. Let R be a commutative ring. The set of prime ideals of R is usually denoted by $\operatorname{Spec}(R)$, called the *spectrum of* R. Note that if $\varphi : R \to S$ is a ring homomorphism of commutative rings, then by Proposition 5.4.20 we have an induced map

$$\varphi^* : \operatorname{Spec}(S) \to \operatorname{Spec}(R), \quad \mathfrak{p} \mapsto \varphi^{-1}(\mathfrak{p}).$$

One can define a *topology* on $\operatorname{Spec}(R)$, called the *Zariski topology*, turning $\operatorname{Spec}(R)$ into a space, and one can define *functions* on $\operatorname{Spec}(R)$ turning it into an *affine scheme*, and then define maps between affine schemes. In fact the *category* of affine schemes is equivalent to the category of commutative rings!

The study of the geometric spaces arising from affine schemes is called *algebraic geometry*, and is a rich field of math. In particular, by studying polynomial rings, algebraic geometry studies solutions to polynomial equations. A friendly introduction can also be found in *Ideals, Varieties, and Algorithms* by Cox–Little–O'Shea. For an introduction to algebraic geometry, I highly recommend Fulton's book *Algebraic Curves*, available for free on the author's website. Some additional references are Atiyah and MacDonald's book *Introduction to Commutative Algebra*, where much of the algebra background for algebraic geometry is developed. Another nice reference is Eisenbud's *Commutative Algebra: With a View Toward Algebraic Geometry.*

We'll end with a few observations about prime and maximal ideals.

Exercise 5.4.25. Let R and S be commutative rings and $f : R \to S$ be a ring homomorphism. If $\mathfrak{m} \subset S$ is a maximal ideal, is $f^{-1}(\mathfrak{m})$ a maximal ideal? Show that $f^{-1}(\mathfrak{m})$ need not be maximal. If f is surjective, show that $f^{-1}(\mathfrak{m})$ is a maximal ideal.

Proposition 5.4.26. *Let R be a PID. Then every non-zero prime ideal is maximal.*

Proof. Let (x) be a non-zero prime ideal, and suppose I is an ideal containing (x) with $I \neq (x)$. Since R is a PID, $I = (y)$ for some $y \in R$. Thus as $(x) \subset (y)$, $x \in (y)$ so $x = yz$ for some $xz \in R$. Since $yz = x \in (x)$, but $y \notin (x)$, (x) being prime implies that $z \in (x)$. So $z = tx$ for some $t \in R$. But then

$$x = yz = ytx,$$

hence as R is a domain, $yt = 1$. Thus $(y) = (1) = R$, and (x) is maximal. □

Example 5.4.27. It is not true that prime ideals are always maximal. Indeed, in the ring $\mathbb{C}[x, y]$, the ideal (x) is prime, but not maximal, as $\mathbb{C}[x, y]/(x) \cong \mathbb{C}[y]$, which is not a field.

Corollary 5.4.28. *Let R be a commutative ring. Suppose that $R[x]$ is a PID. Then R is a field.*

Proof. We have an injective ring homomorphism $R \to R[x], r \mapsto r \times 1$. Thus as $R[x]$ is a domain, R must also be a domain. Furthermore, $R[x]/(x) \cong R$. Since R is a domain, (x) is prime. Therefore, as $R[x]$ is a PID, (x) is maximal, and so $R[x]/(x) \cong R$ is a field. □

Our next topic will be about factorization in rings. After all, we have *prime* ideals...

5.5. Factorization Domains.

We will build up the notion of factorization in rings. We first set some terminology, which may differ from your intuition. Throughout this section, R will be a domain.

Definition 5.5.1. Let R be a domain.

- We say $u \in R$ is a *unit* if $u \in R^\times = \{a \in R \mid \text{ there exists } b \in R \text{ such that } ab = 1\}$
- Let $r \in R$ be a non-zero non-unit. We say r is *irreducible* if whenever $r = ab$ with $a, b \in R$, then a is a unit or b is a unit. Otherwise, r is called *reducible.*
- A non-zero element $p \in R \setminus 0$ is called *prime* if (p) is a prime ideal. That is, if whenever $p \mid ab$, then $p \mid a$ or $p \mid b$.
- We say that two elements $a, b \in R$ are *associate* if $a = ub$ for some unit $u \in R^\times$.

Proposition 5.5.2. *Let R be a domain and $p \in R$ be prime. Then p is irreducible.*

Proof. Suppose (p) is a non-zero prime ideal, and let $p = ab$. Then $ab = p \in (p)$, so by the definition of prime ideal, $a \in (p)$ or $b \in (p)$. Without loss of generality, we may assume $a \in (p)$. Thus $a = pr$ for some $r \in R$, hence

$$p = ab = prb,$$

whereby $rb = 1$. Thus r and b are units. In particular, $p = ab$ implied that b is a unit. Hence p is irreducible. □

Example 5.5.3. The converse of Proposition 5.5.2 is not always true. For example, consider the element $3 \in \mathbb{Z}[\sqrt{-5}]$. Then 3 is irreducible, but not prime since

$$(2 + \sqrt{-5})(2 - \sqrt{5}) = 9 \in (3),$$

but

$$(2 + \sqrt{-5}), (2 - \sqrt{5}) \notin (3).$$

Proposition 5.5.4. *Let R be a PID. Then a nonzero eleent is prime if and only if it is irreducible.*

Proof. We've seen that prime implies irreducible in Proposition 5.5.2. For the converse, suppose that $p \in R$ is irreducible. We show that (p) is a maximal ideal, hence prime. So suppose that I is an ideal of R containing (p). Since R is a PID, $I = (a)$. Thus as $p \in (p) \subseteq (a)$, $p = ra$ for some $r \in R$. Since p is irreducible, $r \in R^\times$ or $a \in R^\times$.In the former case, $(p) = (a)$, and in the latter case $(a) = R$. Thus (p) is a maximal ideal. □

More generally, the two notions of prime and irreducible coincide when we have unique factorization.

Definition 5.5.5. Let R be a domain.

- We call R a *factorization domain* (FD) if for every nonzero, nonunit $a \in R$, there exist irreducible elements $p_1 \dots, p_\ell \in R$ such that $a = p_1 \cdots p_\ell$.
- We call R a *unique factorization domain* (UFD) if R is a FD and whenever
$$p_1 \cdots p_\ell = q_1 \dots q_m$$
for irreducible elements $p_1, \dots, p_\ell, q_1, \dots, q_m \in R$, then $\ell = m$ and up to re-ordering, p_i and q_i are associates.

Example 5.5.6. The integers $\mathbb{Z}$ are a UFD. The reason we have to add the "associates" in the definition of UFD is because of factorization like

$$6 = 2 \cdot 3 = (-2) \cdot (-3).$$

Proposition 5.5.7. *Let R be a factorization domain. Then R is a UFD if and only if every irreducible element of R is prime.*

Proof. Suppose that R is a UFD, and let $p \in R$ be irreducible. We want to show that p is prime. So let $ab \in (p)$. Thus $ab = pc$ for some $c \in R$. Since R is a factorization domain, there are factorizations into irreducible elements

$$a = p_1 \cdots p_k, \quad b = q_1 \cdots q_\ell, \quad c = r_1 \cdots r_m.$$

Thus we have two irreducible factorizations of $ab = pc$, namely

$$p_1 \cdots p_k \cdot q_1 \cdots q_m = p \cdot r_1 \cdots r_m.$$

Since R is a UFD, there is a unit $u \in R^\times$ such that $p = up_i$ or $p = uq_j$ for some i or j. Thus $a \in (p)$ or $b \in (p)$, so p is prime.

Conversely, suppose that every irreducible element of R is prime. We want to show that R is a UFD. So let

$$p_1 \cdots p_\ell = q_1 \cdots q_m$$

be two factorizations into irreducible elements. Without loss of generality, we may assume that $\ell \leq m$. Since p_1 is irreducible, hence prime by assumption, and $q_1 \cdots q_m \in (p_1)$, there is some $q_j \in (p_1)$. Rearranging, we may assume $q_1 \in (p_1)$. Thus $q_1 = u_1 p_1$ for some $u_1 \in R$, and since q_1 is irreducible, $u \in R^\times$ is a unit. Thus we have

$$p_1 \cdots p_\ell = u_1 p_1 q_2 \cdots q_m,$$

and canceling p_1 on both sides (as R is a domain) we have

$$p_2 \cdots p_\ell = u_1 q_2 \cdots q_m.$$

Since $u_1 \notin (p_2)$, we can continue and find a unit u_2 such that up to reordering $q_2 = u_2 p_2$, and find

$$p_3 \cdots p_\ell = u_1 u_2 q_3 \cdots q_m.$$

After ℓ steps, we find units $u_1, \dots, u_\ell$ such that $q_i = u_i p_i$ and

$$1 = u_1 \cdots u_\ell q_{\ell+1} \cdots q_m.$$

Since each q_j is irreducible, and not a unit, we must have $\ell = m$. Thus R is a UFD. □

Theorem 5.5.8. *Let R be a PID. Then R is a UFD.*

Proof. We first show that if R is a PID, then R is a factorization domain.

Suppose for contradiction that there exists a nonzero, nonunit $a_0 \in R$ that doe snot factor as a product of irreducibles. Thus a_0 itself is not irreducible, thus we can factor $a_0 = a_1 b_1$ for some $a_1, b_1 \in R$ with neither a_1 nor b_1 a unit. If both a_1 and b_1 factored as a product of irreducibles, then so would their product. Therefore, without loss of generality, we may assume a_1 does not factor as a product of irreducibles. And we have $(a_0) \subsetneq (a_1)$. Continuing, since a_2 cannot be irreducible, there are nonunits $a_2, b_2 \in R$ such that $a_1 = a_2 b_2$ and we may again assume without loss of generality that a_2 does not factor as a product of irreducibles. Continuing like this, we find an infinite sequence $a_0, a_1, \ldots$ with $a_i = a_{i+1} b_{i+1} \in R$ for some nonunits b_{i+1}, and we have an infinite chain of strict containments of ideals

$$(a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \cdots .$$

The union of these ideals is an ideal, and since R is a PID, we have

$$\bigcup_{i=0}^{\infty} (a_i) = (c)$$

for some $c \in R$. Thus $c \in (a_n)$ for some n, hence $(c) = (a_k)$ for all $k \geq n$, which is a contradiction to the infinite chain of ideals being strict containments.

We've seen in Proposition 5.5.4 that if R is a PID, then every irreducible element is prime. Thus, as R is a factorization domain, Proposition 5.5.7 implies that R is a UFD. □

5.5.1. *Noetherian Rings.*

Remark 5.5.9. The argument in Theorem 5.5.8 really hinged on the fact that the infinite chain of ideals *stabilized* at some point. This is such an important property that it is useful to give it a name.

Definition 5.5.10. Let R be a (commutative) ring. We say that R satisfies the *ascending chain condition* if for any chain of ideals

$$I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$$

there is some N such that for all $k \geq N$, $I_k = I_N$. That is, every increasing chain of ideals stabilizes.

When we have the ascending chain condition, we don't really need Zorn's Lemma anymore.

Proposition 5.5.11. *Let $\mathcal{P}$ be a partially ordered set. The following are equivalent:*

(i) *Every increasing sequence $x_1 \leq x_2 \leq \cdots$ in $\mathcal{P}$ stabilizes, i.e., there is some N such that $x_k = x_N$ for all $k \geq N$.*
(ii) *Every non-empty subset of $\mathcal{P}$ has a maximal element.*

Proof. (i) $\Longrightarrow$ (ii): Suppose for contradiction that (ii) is false. Then there is a non-empty subset $T \subseteq \mathcal{P}$ with no maximal element, and we can construct inductively an infinite strictly increasing sequence of elements of T, contradicting (i).

(ii) $\Longrightarrow$ (i): The set $\{x_i\}_{i \geq 1}$ has a maximal element, x_N, and the sequence stabilizes. □

Theorem 5.5.12. *Let R be a commutative ring. Then the following are equivalent.*

(i) *R satisfies the ascending chain condition*
(ii) *every increasing chain of ideals has a maximal element*
(iii) *every ideal of R is finitely generated, i.e., if I is an ideal fo R then there exist $f_1, \ldots, f_m \in R$ such that $I = (f_1, \ldots, f_m)$.*

Proof. We've seen that (i) and (ii) are equivalent. It remains to show that (iii) is equivalent.

(ii) $\implies$ (iii): Let I be an ideal of R, and let $\mathcal{P}$ be the set of all finitely generated ideals contained in I. Since $(0) \in \mathcal{P}$, $\mathcal{P}$ is non-empty, and thus has a maximal element, say $J = (f_1, \dots, f_m)$. If $I = J$, we are done. Else, suppose that $I \neq J$ and let $x \in I \setminus J$. Then $J + (x) = (f_1, \dots, f_m, x) \subseteq I$ is a finitely generated ideal contained in I. Since $x \notin J$, $J \subsetneq J + (x)$, contradicting the fact that J is maximal. Hence $J = I$, as desired.

(iii) $\implies$ (i): Let $I_1 \subseteq I_2 \subseteq \cdots$ be an ascending chain of ideals of R. Then $I = \bigcup_{i=1}^{\infty} I_i$ is an ideal of R, hence finitely generated by $f_1, \dots, f_m$. Say $f_i \in I_{n_i}$ for some $n_1, \dots, n_m$. Then for $N = \max\{n_1, \dots, n_m\}$, $I = I_N$, and thus the chain stabilizes. □

Definition 5.5.13. We say a ring R is *Noetherian* if R satisfies any of the equivalent conditions in Theorem 5.5.12.

Remark 5.5.14. Noetherian rings are named in honor of Emmy Noether, who revolutionized the study of rings and other areas of abstract algebra.

Example 5.5.15. A field is Noetherian, since the only ideals are 0 and (1). More generally, a PID is Noetherian, every ideal is finitely generated (by one element).

An important fact about Noetherian rings is the following, the proof is taken from Fulton's *Algebraic Curves*.

Theorem 5.5.16 (Hilbert Basis Theorem). *Let R be a Noetherian ring, then $R[x]$ is Noetherian.*

Proof. Suppose that R is Noetherian, and let $I \subset R[x]$ be an ideal. We must find a finite set of generators for I.

If $F = a_d x^d + \cdots + a_1 x + a_0 \in R[x]$, and $a_d \neq 0$, we call a_d the leading coefficient of F. Let J be the set of leading coefficients of all polynomials in I. Then $J \subseteq R$ is an ideal, as if $F, G \in I$, then $F + G \in I$, and if $a \in R$, then $aF \in I$, thus the leading coefficients of polynomials in I are closed under addition and scaling by elements of R.

Since R is Noetherian, J is finitely generated, so there are some polynomials $F_1, \dots, F_r \in I$ whose leading coefficients generate J.

Now fix some $N \geq \max\{\deg(F_1), \dots, \deg(F_r)\}$. For each $m \leq N$, let J_m be the ideal in R consisting of all polynomials $F \in I$ such that $\deg(F) \leq m$. As for J, let $\{F_{m_j}\}$ be a finite set of polynomials whose leading coefficients generate J_m.

Let I' be the ideal generated by the finite set of the $\{F_1, \dots, F_r\}$ and all the $\{F_{m_j}\}$. So I' is finitely generated. We claim that $I' = I$.

Since each generator of I' is an element of I, we clearly have $I' \subseteq I$. Suppose for contradiction that $I' \neq I$, and let $G \in I \setminus I'$ be an element of lowest degree (which exists by well-ordering). If $\deg(G) > N$, then since the leading coefficient of G is in J, we can find polynomials Q_i such that $\sum F_i Q_i$ and G have the same leading coefficient. But then $\deg(G - \sum Q_i F_i) < \deg(G)$, so $G - \sum Q_i F_i \in I'$, whereby $G \in I'$, which is a contradiction. Similarly, if $\deg(G) = m \leq N$, then the leading coefficient of G is in J_m, and we can find polynomials Q_j such that G and $\sum Q_j F_{m_j}$ have the same leading coefficient, again leading to a contradiction. Thus $I' = I$, and I is finitely generated. □

Corollary 5.5.17. *Let K be a field, then $K[x_1, \dots, x_n]$ is Noetherian.*

Exercise 5.5.18. Prove the following proposition.

Proposition 5.5.19. *Let R be a ring, I an ideal of R, and $\pi : R \to R/I$ the natural quotient map. Show that if $J \subset R$ is a finitely generated ideal, then $\pi(J)$ is a finitely generated ideal of R/I. Therefore, if R is Noetherian, then R/I is Noetherian for any ideal I of R.*

One nice thing about Noetherian domains is that you can factor, though perhaps not uniquely. In particular, in the proof that a PID was a UFD, the crucial part in showing that a PID was a factorization domain was showing that the infinite chain of ideals

$$(a_0) \subsetneq (a_1) \subsetneq \cdots$$

eventually stabilized, which is a special case of the ascending chain condition.

Theorem 5.5.20. *Let R be a Noetherian domain, i.e. a domain that is Noetherian. Then R is a factorization domain.*

More generally, a domain which satisfies an "ascending chain condition on principal ideals" is a factorization domain, the proof being the same as for a PID or a Noetherian domain.

5.6. **Polynomial Rings.**

We'll develop some of the theory of polynomial rings $F[x_1, \dots, x_n]$ over a field F. You'll see much more of this in Math 4581 when you cover Galois theory.

Notation. Throughout this subsection, F will denote a field unless otherwise specified. We may sometimes emphasize that F is a field.

Definition 5.6.1. The ring of polynomials with coefficients in F is denoted by $F[x]$.

Slogan. $F[x]$ is like $\mathbb{Z}$.

First, $F[x]$ is a ring, we can add and multiply polynomials in the usual way. In $\mathbb{Z}$ we may not always be able to divide, but we can always divide with remainder.

Recall 5.6.2. We'll recall some facts about $\mathbb{Z}$.

Theorem 5.6.3 (Division algorithm in $\mathbb{Z}$)**.** *Given $a, b \in \mathbb{Z}$, there are unique $q, r \in \mathbb{Z}$ such that $b = ra + r$ with $|r| < |a|$ or $r = 0$.*

This gives the Euclidean algorithm for the greatest common divisor.

Definition 5.6.4. The greatest common divisor of a and b is $\gcd(a, b) = d$ if $d \mid a$, $d \mid b$, and if $e \in \mathbb{Z}$ also divides both a and b, then $e \mid d$.

The Euclidean algorithm for computing $\gcd(a, b)$ also shows that $\mathbb{Z}$ is a *principal ideal domain*. That is,

- $\mathbb{Z}$ is a domain (if $ab = 0$, then $a = 0$ or $b = 0$), and
- ideals of $\mathbb{Z}$ are principal (if $I \subseteq \mathbb{Z}$ is an ideal, then $I = (d)$ for some $d \in \mathbb{Z}$. In fact, one can take d to be the smallest non-negative element of I.)

Example 5.6.5. For $a, b \in \mathbb{Z}$, the ideal generated by a and b,

$$(a, b) \coloneqq \{n \in \mathbb{Z} \mid n = ax + by \text{ for some } x, y \in \mathbb{Z}\}$$

is principally generated by $\gcd(a, b)$, i.e., $(a, b) = (\gcd(a, b))$.

The ring of polynomials over a field enjoys similar properties.

Definition 5.6.6. Let $p(x) = a_n x^n + \cdots + a_1 x + a_0 \in F[x]$. The *degree of* $p(x)$ is the largest d such that $a_d \neq 0$, and we write $\deg p = d$.

Exercise 5.6.7. Let F be a domain. Then $F[x]$ is a domain.

Lemma 5.6.8. *Let F be a domain, and $f, g \in F[x]$. We have $\deg(fg) = \deg(f) + \deg(g)$ and $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$.*

Remark 5.6.9. This is no longer true if F is replaced with a ring that is not a domain! Namely in $\mathbb{Z}/6\mathbb{Z}[x]$ we have

$$(2x+1)(3x) = 6x^2 + 3x = 3x.$$

Theorem 5.6.10 (Division algorithm for polynomials)**.** *Given $a, b \in F[x]$, with $a \neq 0$, there are unique $q, r \in F[x]$ such that*

$$b = qa + r$$

with $\deg(r) < \deg(a)$ or $r = 0$.

Proof. We first show that q and r exist.

We'll fix some notation, let

$$a = a_n x^n + \cdots + a_0, \ a_n \neq 0,$$

and

$$b = b_k x^k + \cdots + b_0, \ b_k \neq 0.$$

If $k < n$, then

$$b = 0 \cdot a + b$$

is of the required form. So we may suppose $k \geq n$. We proceed by induction on k.

Note that

$$b - \frac{b_k}{a_n} x^{k-n} a$$

has degree $< k = \deg(b)$. So by induction, there are $\widetilde{q}, r \in F[x]$ such that $\deg(r) < \deg(a)$ and

$$b = \underbrace{(\frac{b_k}{a_n} x^{k-n} + \widetilde{q})a}_{q} + r,$$

and letting $q = (\frac{b_k}{a_n} x^{k-n} + \widetilde{q})a$, we have

$$b = qa + r$$

with $\deg(r) < \deg(a)$, as desired.

To prove that q and r are unique, suppose that

$$b = qa + r + \widetilde{q}a + \widetilde{r}$$

with $\deg(r), \deg(\widetilde{r}) < \deg(a)$. Then

$$a(q - \widetilde{q}) = \widetilde{r} - r,$$

and note that $\deg(\widetilde{r} - r) < \deg(a)$.

If $q - \widetilde{q} \neq 0$, then $\deg\left(a(q - \widetilde{q})\right) \geq \deg(a)$, contradicting the fact that

$$\deg\left(a(q - \widetilde{q})\right) = \deg(\widetilde{r} - r) < a.$$

Thus $q - \widetilde{q} = 0$, hence $\widetilde{r} - r = 0$, and thus

$$q = \widetilde{q}, \ r = \widetilde{r},$$

as desired. □

Definition 5.6.11. For $a, d \in F[x]$, we write $d \mid a$ and say "d divides a" if $a = qd$ for some $q \in F[x]$.

As in $\mathbb{Z}$, the division algorithm gives us a Euclidean algorithm.

Theorem 5.6.12. *For $a, b \in F[x]$, a and b have a greatest common divisor $d \in F[x]$ such that*

- *$d \mid a$ and $d \mid b$, and*
- *if $e \mid a$ and $e \mid b$, then $e \mid d$.*

Moreover, $\gcd(a,b) = fa + gb$ for some $f, g \in F[x]$.

Proof. Observe first that

$$\{\text{common divisors } e \text{ of } a \text{ and } b\} = \{\text{common divisors } e \text{ of } a \text{ and } b + ka \text{ for any } k \in F[x]\}.$$

Indeed, if e is a common divisor of a and b, then certainly e divides a and $b + ka$. Conversely, if e divides a and $b + ka$, then e divides a and $b = (b + ka) - ka$.

Now we apply the division algorithm to find $d = \gcd(a, b)$, as

$$\{\text{common divisors of } a \text{ and } b\} = \{\text{common divisors of } a \text{ and } \underbrace{b - qa}_{=r}\}.$$

If $\deg(a) \leq \deg(b)$, then

$$\min\{\deg(r), \deg(a)\} \leq \min\{\deg(a), \deg(b)\},$$

so using the division algorithm to find smaller degree common divisors must terminate, and it terminates when $r = 0$. So we transform the pair (a, b) to $(d, 0)$ with $d = \gcd(a, b)$.

As we found d by replacing a and b with $b - ka$ or vice versa, it is clear that d is obtained in the form $d = fa + gb$ for some $f, g \in F[x]$. □

Theorem 5.6.13. *Let F be a field. Then $F[x]$ is a PID.*

Proof. Since F is a field, if $f, g \in F[x]$ are non-zero, then fg is also non-zero as the top degree coefficients cannot cancel.

It remains to show that all ideals of $F[x]$ are principal. This is analogous to the proof that $\mathbb{Z}$ is a PID, replacing d with a polynomial of smallest degree in I. We give the details for completeness.

Let $I \subseteq F[x]$ be an ideal. If $I = 0$, then I is principal. So we may assume $I \neq 0$, and let $d \in I$ be a non-zero polynomial of lowest degree. We claim that $I = (d)$. Since $d \in I$, it is clear that $(d) \subseteq I$. Conversely, suppose that $f \in I$, and write $f = qd + r$ with $\deg(r) < \deg(d)$ or $r = 0$. Then $f - qd = r \in I$ and so $r = 0$. Thus $f = qd$, and so $I = (d)$. □

Corollary 5.6.14. *Let F be a field. Then $F[x]$ is a UFD.*

Proof. Every PID is a UFD (Theorem 5.5.8). □

Remark 5.6.15. Once we know that $F[x]$ is a PID, we can define greatest common divisors just like in $\mathbb{Z}$, we define $\gcd(f, g) = d$ where d is the generator of the ideal $(f, g) = (d)$.

Remark 5.6.16. Note that if F is a field, then $F[x, y]$ is no longer a PID. Indeed, the ideal (x, y) is not generated by one element. However, one can prove that if R is a UFD, then $R[x]$ is a UFD, though we will not develop the required theory here.

Definition 5.6.17. Let $a \in F$, then there is a ring homomorphism $\mathrm{ev}_a : F[x] \to F, p(x) \mapsto p(a)$, obtained by evaluating polynomials at a.

Definition 5.6.18. For $p(x) \in F[x]$, we call $a \in F$ a *root of* $p(x)$ or a *zero of* $p(x)$ if $\mathrm{ev}_a(p) = p(a) = 0$.

Proposition 5.6.19. *Let F be a field. Then $a \in F$ is a root of $p(x) \in F[x]$ if and only if $x - a$ divides $p(x)$.*

Proof. Let a be a root of p. By the division algorithm, we have

$$p = (x - a)q + r$$

with $\deg(r) < \deg(x - a) = 1$. So $r = b \in F$. Evaluating at a, we have

$$0 = p(a) = (a - a)q(a) + r = r,$$

and so $r = 0$ and $x - a$ divides p.

Conversely, if $p = (x-a)q$, then evaluating at a shows that $p(a) = 0$. □

Corollary 5.6.20. *Let F be a field. A non-zero polynomial of degree n has at most n roots in F.*

Proof. Let $p \in F[x]$ with $\deg(p) = n$. if $a_1, \dots, a_k$ are all the roots of p in F, then by Proposition 5.6.19

$$p = (x - a_1) \cdots (x - a_k) q,$$

for some non-zero $q \in F[x]$ of degree $\deg(q) \geq 0$. Taking degrees gives

$$n = \deg(p) = \deg\left((x-a_1)\cdots(x-a_k)q\right) = k + \deg(q) \geq k,$$

thus $n \geq k$. □

Remark 5.6.21. It is not true that a polynomial of degree n has to have n roots! For example, the polynomial $x^2 + 1 \in \mathbb{R}[x]$ has no roots!

Remark 5.6.22. Worse still, a polynomial $f \in F[x]$ can have every element of F as a root and still be non-zero. For example, let $F = \mathbb{Z}/2\mathbb{Z}$ and consider the polynomial $x^2 - x \in F[x]$.

The above examples only exist for polynomial rings over finite fields.

Theorem 5.6.23. *Let F be an infinite field. Let $f \in F[x_1, \dots, x_n]$ be a polynomial in n variables. If the corresponding function*

$$F^n \to F, \quad (a_1, \dots, a_n) \mapsto f(a_1, \dots, a_n)$$

is the zero function, then f is the zero polynomial.

Proof. We will use induction.

Let $g(x) \in F[x]$ be a polynomial such that $g(a) = 0$ for all $a \in F$. Suppose for contradiction that $g \neq 0$. Since F is infinite, $g(x)$ has infinitely many zeros, which is a contradiction. Thus $g = 0$.

For the inductive step, let $f \in F[x_1, \dots, x_n]$ be such that for every $(a-1\dots, a_n) \in F^n$, $f(a_1, \dots, a_n) = 0$. Write

$$f(x_1, \dots, x_n) = \sum_{j=0}^{d} f_j(x_1, \dots, x_{n-1}) x_n^j \in F[x_1, \dots, x_n] = F[x_1, \dots, x_{n-1}][x_n].$$

Let $(a_1, \dots, a_{n-1}) \in F^{n-1}$ be arbitrary. By assumption,

$$g(x_n) = f(a_1, \dots, a_{n-1}, x_n) = \sum_{j=0}^{d} f_j(a_1, \dots, a_{n-1}) x_n^j$$

is a polynomial in one variable with infinitely many roots, hence $g(x_n) = 0$, so $f_j(a_1, \dots, a_{n-1}) = 0$ for all $(a_1, \dots, a_{n-1}) \in F^{n-1}$ and all $1 \leq j \leq d$. Thus by induction $f_j = 0$ for all j, and so f is the zero polynomial. □

One nice thing about polynomial rings is that we can use them to construct bigger fields.

Proposition 5.6.24. *Let F be a field and $p \in F[x]$ be irreducible. Then $F[x]/(p)$ is a field.*

Proof. Since $F[x]$ is a PID, p is prime, by Proposition 5.5.4. Thus (p) is a prime ideal. Furthermore, by Proposition 5.4.26, (p) is a maximal ideal. Hence $F[x]/(p)$ is a field. □

Example 5.6.25. The polynomial $x^2 + 1 \in \mathbb{R}[x]$ is irreducible. Indeed, suppose that $x^2 + 1 = fg$. Then $0 \leq \deg(f), \deg(g) \leq 2$. Suppose for contradiction that $\deg(f) = 1 = \deg(g)$. But then f and g must have roots, hence $x^2 + 1$ has a root, which is a contradiction. Thus $\deg(f) = 0$ or $\deg(g) = 0$, and one of f or g is a unit. Therefore $x^2 + 1$ is irreducible in $\mathbb{R}[x]$.

In particular, $\mathbb{R}[x]/(x^2+1)$ is a field. It turns out that

$$\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}.$$

Let us prove this. Let $i \in \mathbb{C}$ be an element such that $i^2 = -1$. Consider the map

$$\mathrm{ev}_i : \mathbb{R}[x] \to \mathbb{C}, \quad f(x) \mapsto f(i).$$

Then ev_i is surjective, as $a + bi = \mathrm{ev}_i(a+bx)$. Moreover, since $\mathbb{R}[x]$ is a PID, $\ker \mathrm{ev}_i = (p)$ for some $p \in \mathbb{R}[x]$. Since $\mathrm{ev}_i(x^2+1) = (i)^2 + 1 = 0$, $(x^2+1) \subseteq \ker \mathrm{ev}_i$. And since x^2+1 is irreducible, (x^2+1) is maximal, thus $\ker \mathrm{ev}_i = (x^2+1)$. By the First Isomorphism Theorem (Theorem 5.3.8),

$$\mathbb{R}[x]/(x^2+1) \cong \mathbb{C}.$$

Let's see another example.

Proposition 5.6.26. *Let $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$. Then $x^2 - x - 1 \in \mathbb{F}_2[x]$ is irreducible, and the quotient $\mathbb{F}_2[x]/(x^2-x-1)$ is a field with four elements.*

Proof. Suppose for contradiction that $x^2 - x - 1$ is reducible, and write

$$pq = x^2 - x - 1.$$

If $\deg(p) = 1 = \deg(q)$, then p and q are of the form $ax+b$ and $\frac{-b}{a}$ is a root. Thus x^2-x-1 would have a root. However, evaluating at 0 or 1, we find that x^2-x-1 has no roots! Thus $\deg(p) = 0$ or $\deg(g) = 0$, in which case one of p or q is a unit, and so x^2-x-1 is irreducible.

Hence (x^2-x-1) is a maximal ideal (as $\mathbb{F}_2[x]$ is a PID, and x^2-x-1 is irreducible hence prime, and every prime ideal of a PID is maximal). Thus $\mathbb{F}_2[x]/(x^2-x-1)$ is a field.

We now show that $\mathbb{F}_2[x]/(x^2-x-1)$ has four elements. The cosets of $0, 1, x$, and $x+1$ are all distinct, since the difference of any two would have degree ≤ 2, and thus cannot be in the ideal (x^2-x-1). Hence $\mathbb{F}_2[x]/(x^2-x-1)$ has at least the four elements $0, 1, x, x+1$. Moreover, since $x^2 = x+1 \in \mathbb{F}_2[x]/(x^2-x-1)$, any class can be represented by an element of degree 1 or less. Hence the cosets of (x^2-x-1) are exactly represented by $0, 1, x, x+1$, thus $\mathbb{F}_2[x]/(x^2-x-1)$ has four elements. □

More generally, one can show that any finite field $\mathbb{F}_q$ has p^n elements where $p = \mathrm{char}(\mathbb{F}_q)$, and using similar constructions there is a finite field $\mathbb{F}_{p^n}$ for every prime p and $n \in \mathbb{Z}_{>0}$.

We'll end with some nice problems involving polynomial rings.

Exercise 5.6.27. Let K be a field, and let $F \in k[x_1, \ldots, x_n]$, and $a_1, \ldots, a_n \in K$.

(1) Show that there are some $\lambda_{(i)} \in K$ such that

$$F = \sum \lambda_{(i)} (x_1 - a_1)^{i_1} \cdots (x_n - a_n)^{i_n}.$$

(Hint: consider trying to eliminate the term of F with the largest degree using only products of the monomials $(x_i - a_i)$.)

(2) If $F(a_1, \ldots, a_n) = 0$, show that

$$F = \sum_{i=1}^{n} (x_i - a_i) G_i$$

for some (potentially not unique) $G_i \in K[x_1, \ldots, x_n]$.

Exercise 5.6.28. Let K be a field, and $a_1, \ldots, a_n \in K$. Show that

$$I = (x_1 - a_1, \ldots, x_n - a_n) \subset K[x_1, \ldots, x_n]$$

is a maximal ideal, and that $K[x_1, \dots, x_n]/I \cong K$.

Exercise 5.6.29. Let K be any field. Show that there are an infinite number of irreducible polynomials in $K[x]$ with leading coefficient 1. (Hint: Suppose $F_1, \dots, F_n$ were all of them, and factor $F_1 \cdots F_n + 1$ into irreducible factors.)

6. Further Topics in Algebra

We highlight some potential future topics that might be of interest after this first course in abstract algebra.

6.1. **Galois Theory.**

Galois theory, named in honor of Évariste Galois, arose in the study of the question "When can you find a formula for the solutions to a polynomial equation in one variable?" In particular, we know the famous quadratic formula that says $ax^2 + bx + c$ has solutions

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

It turns out that there are similar formulas for polynomials of degree 3 and 4, but there is *no* formula for the solutions of a polynomials of degree ≥ 5 using just radicals (this is known as Abel's Theorem).

The modern approach to Galois theory is to view finding roots of polynomials as finding larger fields, like $\mathbb{R}[x]/(x^2 + 1)$ is a larger field than $\mathbb{R}$ and has the new root $i \in \mathbb{C}$. We say a field L is an *extension* of a field K if $K \subseteq L$, and Galois theory studies "nice" field extensions. In particular, we consider the groups $\mathrm{Aut}(L/K)$ of field automorphisms of L that are the identity on K (like complex conjugation on $\mathbb{C}$ that fixes the subfield $\mathbb{R}$), and we notice that $\sigma \in \mathrm{Aut}(L/K)$ sends roots in L of $f \in K[x]$ to different roots of f, since $\sigma(0) = 0$. The magic of Galois theory is that there is a deep connection between subfields $K \subset F \subset L$ and subgroups of $\mathrm{Aut}(L/K)$, and many properties of the intermediate field F are reflected in the subgroup, and one can understand roots of polynomials using these groups. You'll see this in Math 4581. A very nice treatment is Milne's book *Fields and Galois Theory*, available on the authors website at https://www.jmilne.org/math/CourseNotes/FT.pdf, which you are now prepared to read. Dummit and Foote's book also has a very nice exposition of Galois theory.

And don't be fooled into thinking that Galois theory is all done, there are still many interesting open questions relating to Galois theory, and many people working on those and related questions.

6.2. **Representation Theory.**

Another way to study groups is by viewing them as symmetries acting on some objects. In the case of Representation theory, the objects will be vector spaces, and you can study groups by how they act on vector spaces. A (finite dimensional) *representation* of a group G is a group homomorphism $\rho : G \to \mathrm{GL}_n(F)$ for some field F, usually the field is $\mathbb{C}$ in a first introduction to representation theory.

For example, the way we represented elements of S_3 by matrices in Example 4.1.3(5) is a representation $S_3 \to \mathrm{GL}_3(\mathbb{C})$.

Many properties of groups can be seen in terms of their representations, and representation theory has deep connections to areas of math including number theory, algebraic geometry, Fourier analysis, and many other branches of math, and representation theory is an active area of modern research. A nice introduction can be found at the end of Dummit and Foote's book, as well as the classics *Linear Representations of Finite Groups* by Serre, and *Representation Theory* by Fulton and Harris.

6.3. **Algebraic Geometry and Commutative Algebra.** At its core, Algebraic Geometry is the study of solutions to polynomial equations. For example, one can take the solutions to $x^2+y^2-1 \in \mathbb{R}[x,y]$ and view the solutions as a circle, namely $\{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=1\}$. And one can study both the algebraic aspects of the polynomial x^2+y^2-1 and the geometric aspects of the space of solutions. There is a deep interplay between the algebra side and the geometry side, and algebraic geometry is the study of this relationship, or perhaps using tools from one side to study the other.

For example, solutions to equations like $y^2 = x^3+ax+b$ are called *elliptic curves*, and pop up all over the place in math, including applications like cryptography (you might have been using an elliptic curve to connect to the internet where you downloaded these notes). Over the complex numbers, the space of solutions to $y^2 = x^3+ax+b$ looks like a doughnut or a bagel, and elliptic curves have deep connections to number theory and many other areas of math.

In studying the solutions to polynomial equations, it is very useful to understand the polynomial rings $F[x_1,\dots,x_n]$ better, and in fact much of the theory works for commutative rings in general. The study of commutative rings is a rich topic, with many connections to other areas of math, and is central to algebraic geometry. A nice introduction to algebraic geometry can be found in *Ideals, Varieties, and Algorithms* by Cox–Little–O'Shea. For an introduction to algebraic geometry with a focus on the geometry of curves, I highly recommend Fulton's book *Algebraic Curves*, available for free on the author's website (https://dept.math.lsa.umich.edu/~wfulton/CurveBook.pdf). Some additional references are Atiyah and MacDonald's book *Introduction to Commutative Algebra*, where much of the algebra background for algebraic geometry is developed. Another nice reference is Eisenbud's *Commutative Algebra: With a View Toward Algebraic Geometry.*

6.4. **Algebraic Number Theory.** If you liked thinking about ideals, then algebraic number theory might be right for you! The central questions in algebraic number theory are to understand things like factorization, ideals, and similar structures in different number systems that behave much like $\mathbb{Z}$ does in $\mathbb{Q}$. More concretely, for a field K, we define some analog of the integers, called a "ring of integers", usually denoted by $\mathcal{O}_K$, and one can ask questions about whether $\mathcal{O}_K$ is a UFD, or how badly $\mathcal{O}_K$ fails to be a UFD. Doing this involves many beautiful constructions that touch on many other branches of math. Algebraic Number theory is a very active field of modern research. There are many nice introductions, for example Ireland and Rosen's *A Classical Introduction to Modern Number Theory* and you're already familiar with most of the first chapter, or Milne's *Algebraic Number Theory*, available for free on the author's website https://www.jmilne.org/math/CourseNotes/ANT.pdf, though this might require a fair bit of Galois Theory at some points.

6.5. **Algebra in Geometry and Topology.** One of the main uses of abstract algebra in other branches of math is to define *invariants* of other objects. One of the central questions in math is to "classify" all the objects of some type, i.e. to write down a complete list and understand the structure of the objects completely. For example, in Math 4581 you'll see the classification of finite abelian groups, which are all of the form

$$A \cong \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \cdots \times \mathbb{Z}/d_r\mathbb{Z}$$

with $2 \leq d_1 \mid d_2 \mid \cdots \mid d_r$, and the r and $d_1, \dots d_r$ are uniquely determined by the abelian group A.

But more generally, say you have two spaces X and Y, like the real plane $X = \mathbb{R}^2$ and the real line $Y = \mathbb{R}$. How do we know that X and Y are different? Well, we can define their *dimension*, and see that $\dim(\mathbb{R}^2) = 2$ and $\dim(\mathbb{R}) = 1$, so they are different. Here dimension is an invariant that helped us tell apart $\mathbb{R}^2$ and $\mathbb{R}$. But some spaces are much more complicated, and have a finer structure, like the real line $\mathbb{R}$ and the circle S^1. How can we tell them apart now? They both have dimension 1!

A useful idea is to define invariants that are groups, or rings, or some other algebraic objects with nice properties, and use those to tell things apart. These algebraic structures have much more intricate information than just a number like the dimension. In topology, there is a whole field called *algebraic topology* that tries to study different algebraic invariants to tell shapes apart. As you can imagine, they can get pretty complicated, but they are also very beautiful. For example, to tell the line and circle apart, one can define a *fundamental group* $\pi_1(X, x_0)$ of a space X, which is the group of (homotopy classes of) loops through a fixed point $x_0 \in X$, which are just loops up to bending and stretching and deforming in nice ways. The group operation is just "do one loop and then do the next loop". One can use this to show that $\pi_1(\mathbb{R}) = 0$ while $\pi_1(S^1) \cong \mathbb{Z}$. And so $\mathbb{R}$ and S^1 are different! Even more complicated invariants exist, like homology or cohomology or K-theory, and aspects of these algebraic invariants in geometry, topology, and many other areas of math are active areas of research. The classic reference is Hatcher's wonderful book *Algebraic Topology*, available for free on the author's website at https://pi.math.cornell.edu/~hatcher/AT/AT.pdf. For a brief introduction to topology, Hatcher also has some wonderful notes available here https://pi.math.cornell.edu/~hatcher/Top/Topdownloads.html.

6.6. **Other areas.** There are many other areas of modern math that I have not mentioned, and practically all of them use groups or rings in one way or another. If you want to do some searching for topics you might find interesting that use some ideas from this class, you could look into algebraic combinatorics, Lie theory, algebraic method s in data analysis, cryptography,...