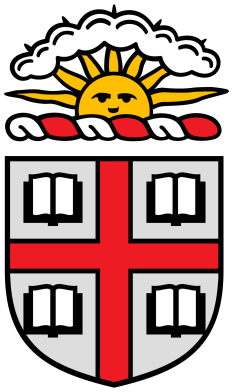
# Abstract Algebra

MATH1530

Professor Jordan Kostiuk

*Brown University*



Edited by

Richard Tang

# Contents

# Chapter 1

# Set Theory

Set theory forms a basis for all of higher mathematics. We begin with a brief introduction.

## §1.1 Sets

**Definition 1.1.1: Sets**

A **set** is a (possibly empty) collection of elements. If $S$ is a set and $a$ is some object, then $a$ is either an element of $S$ or not. We write:

- $a \in S$ if $a$ is an element of $S$.
- $a \notin S$ if $a$ is not an element of $S$.

The empty set is denoted $\varnothing$. We use $|S|$ or $\#S$ to denote the cardinality (number of elements) in a finite set.

**Definition 1.1.2: Natural Numbers**

The **natural numbers** are the set

$$\mathbb{N} = \{1, 2, \ldots\}.$$

Formally, we define $\mathbb{N}$ as follows:

1. $\mathbb{N}$ contains an initial element 1.
2. $\forall n \in \mathbb{N}$, there is an incremental rule that creates the next element $n + 1$.
3. We can reach every element of $\mathbb{N}$ by starting with 1 and repeatedly adding 1.

**Remark 1.** *$\mathbb{N}$ is totally ordered. We say $m$ is less than $n$ if $n$ appears before $n$ when we start from 1 and add repeatedly. In this case we write $m < n$ or $m \leq n$ if $m = n$.*

**Example 1.** *Let*

$$\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$$

*denote the set of integers, and*

$$Q = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}.$$

*the set of rationals.*

**Definition 1.1.3: Set Operations**

Let $S, T$ be sets.

1. $S$ is a **subset** of $T$ if every element of $S$ is an element of $T$, i.e. $a \in S \to a \in T$. We write
$$S \subset T.$$
2. The **union** of $S$ and $T$ is the set of elements that belong to S or belong to T, denoted
$$S \cup T = \{a \mid a \in S \text{ or } a \in T\}.$$
3. The **intersection** of $S$ and $T$ is the set of elements that belong to both S and T, denoted
$$S \cap T = \{a \mid a \in S \text{ and } a \in T\}.$$
4. If $S \subset T$, the **complement** of $S$ in T is the set of elements in $T$ *not* in $S$:
$$S^c = T - S = T\S = \{a \in T \mid a \notin S\}.$$
5. The **product** of $S$ and $T$ is the set of ordered pairs
$$S \times T = \{(a, b) \mid a \in S, b \in T\}.$$
We have projection maps
$$\begin{aligned} proj_1 : S \times T &\longrightarrow S \\ (a, b) &\longmapsto a. \end{aligned}$$
and
$$\begin{aligned} proj_2 : S \times T &\longrightarrow T \\ (a, b) &\longmapsto b. \end{aligned}$$

These definitions extend to sets $S_1, \ldots, S_n$:

$$S_1 \cup \ldots \cup S_n = \bigcup_{i \in I} S_i = \{a \mid a \in S_1 \text{ and } \ldots \text{ and } a \in S_n\} \tag{1.1}$$

### §1.1.1 The Well-Ordering Principle

**Theorem 1.1.1: Well-Ordering Principle**

Let $S \subset N$ be a non-empty subset of $\mathbb{N}$. Then $S$ has a *minimal element*. That is,

$\exists m \in S$ s.t. $n \geq m, \forall n \in S$. Informally, there exists a minimum element that is smaller than all other natural elements.

**Proof.** Since $S$ is non-empty, we can pick $k \in S$. By definition of $\mathbb{N}$, we can start with 1 and add 1 repeatedly to get $k$. So, there are only $k$ elements of $\mathbb{N}$ less than or equal to $k$:

$$1 < 2 < \ldots < k-1 < k.$$

So, we can keep moving down from $k$, until we find an element $j \notin S$; since there are no smaller elements than $j+1 \in S$, $j+1$ is the minimal element. □

## §1.2 Functions

**Definition 1.2.1: Functions**

A **function** from $S$ to $T$ is a rule that assigns some element of $T$ to each element of $S$ :

$$f : S \to T, s \mapsto f(s).$$

$S$ is the **domain** , and $T$ the **codomain**.

**Definition 1.2.2: Composition of Functions**

If $f : S \to T$ and $S : T \to U$, then the **composition** of $f$ and $g$ is

$$g \circ f = S \to U, a \mapsto g(f(a)).$$

**Definition 1.2.3: Bijectivity**

Let $f : S \to T$ be a function.

1. $f$ is **injective** or one-to-one if distinct elements of $S$ go to distinct elements of $T$. In other words,
$$f(a) = f(b) \to a = b.$$
2. $f$ is **surjective** or onto if every element of $T$ comes from some element in $S$:
$$\forall t \in T, \exists s \in S \text{ s.t. } f(s) = t.$$
3. $f$ is **bijective** if it is both injective and surjective.

**Definition 1.2.4: Invertibility**

Let $f : S \to T$ be a function. $f$ is **invertible** if

$$\exists g : T \to S, (g \circ f)(s) = s, s \in S \text{ and } (f \circ g)(t) = t, t \in T.$$

**Theorem 1.2.1: Bijective iff Invertible**

Let $f : S \to T$ be a function. Then $f$ is invertible $\iff$ $f$ is bijective.

**Proof.** Suppose first that $f$ is invertible. Let $g : T \to S$ denote the inverse. We need to prove that $f$ is bijective.
To prove injectivity, suppose $f(a) = f(b)$ for some $a, b \in S$. Applying $g$ to both sides and using the fact that $g$ is the inverse of $f$, we have

$$g(f(a)) = g(f(b)) \Rightarrow a = b.$$

Thus $f$ is injective.
To prove surjectivity, let $t \in T$; we need to find $s \in S$ such that $f(s) = t$. Using the inverse, let $s = g(t)$. Then

$$f(s) = f(g(t)) = t.$$

Thus $f$ is surjective.
Since $f$ is both injective and surjective, $f$ is bijective.

Now, suppose that $f$ is bijective. Then $\forall t \in T, !\exists s \in S$ s.t. $f(s) = t$. Define a new function $g : T \to S$

$$g(t) := \text{"the unique } s \in S \text{ s.t. } f(s) = t\text{"}.$$

We now show that $(g \circ f)(s) = s$ and $(f \circ g)(t) = t$ for $s \in S, t \in T$.
Given $t \in T$, $f(g(t))) = t$ by definition of $t$. Given $s \in S$, we know that $s$ maps to $f(s)$; so, by definition of $g$, $g(f(s)) = s$.
Thus, $g$ is the inverse of $f$. □

# Chapter 2

# Groups: Part I

Groups are a fundamental baseline for abstract algebra. We start with motivating examples, then move on to a concrete definition.

## §2.1 Motivation

### §2.1.1 Permutations

**Definition 2.1.1: Permutations**

Let $X$ be a set. A **permutation** of $X$ is a bijective function

$$\pi : X \to X$$

with the property: $\forall x \in X$, $!\exists x' \in X$ such that $\pi(x') = x$. This allows us to define an inverse $\pi^{-1}$ to be the permutation

$$\pi^{-1} : X \to X$$

with the rule that $\pi^{-1}(x) = x'$, where $x' \in X$ is the unique element such that $\pi(x') = x$.
The **identity permutation** of $X$ is the identity map

$$e : X \to X, e(x) = x, \ \forall x \in X.$$

In general, a *permutation* of a set $X$ is a rule that "mixes up" the elements of $X$.

**Example 2.** *Let $X = \{1, 2, 3, 4\}$. Then a permutation $\sigma : X \to X$ can be thought of as a shuffling of $X$ and visualized as follows:*

$$1 \Rightarrow 2$$
$$2 \Rightarrow 3$$
$$3 \Rightarrow 1$$
$$4 \Rightarrow 4$$

*$\sigma^{-1}$ would be defined as*

$$1 \Rightarrow 3$$
$$2 \Rightarrow 1$$
$$3 \Rightarrow 2$$
$$4 \Rightarrow 4$$

*Now, suppose $\tau$ is defined as $1 \Rightarrow 1, 2 \Rightarrow 3, 3 \Rightarrow 2, 4 \Rightarrow 4$. Then $\sigma \circ \tau$ is*

$$1 \Rightarrow 2$$
$$2 \Rightarrow 1$$
$$3 \Rightarrow 3$$
$$4 \Rightarrow 4$$

*and $\tau \circ \sigma$ is*

$$1 \Rightarrow 3$$
$$2 \Rightarrow 2$$
$$3 \Rightarrow 1$$
$$4 \Rightarrow 4$$

From this, we gather some observations.

- Given any 2 permutations, we can compose to get a new one.
- There was a permutation that didn't do anything ($\sigma \circ \sigma^{-1}$).
- We can invert any permutation.
- If $\sigma, \tau$ are two permutations, then we don't necessarily have $\tau \circ \sigma = \sigma \circ \tau$ (in other words, the group of permutations with composition is not commutative).

**Definition 2.1.2: Transformations**

Let $X$ be a figure in $\mathbb{R}^2$. Then $Trafo(X)$ is the set of transformations on $X$.

Consider the symmetries of a square (involving reflections/rotations on a square) as a motivating example of transformations; are they invertible? commutative?

**Remark 2.** *Each transformation gives a permutation of the vertices $\{A, B, C, D\}$.*

## §2.2 (Abstract) Groups

We now formally define the notion of a **group**.

**Definition 2.2.1: Groups**

A **group** $\{X, \cdot\}$ consists of a set $X$, together with a group rule/law

$$\cdot : G \times G \to G$$
$$(g_1, g_2) \mapsto g_1 \cdot g_2$$

satisfying the following axioms:

1. (identity) there is an element $e \in G$ such that
$$e \cdot g = g \cdot e = g.$$
for all $g \in G$.
2. (inverse) For all $g \in G$, there is an $h \in G$ such that
$$g \cdot h = h \cdot g = e.$$
The element $h$ is called $g^{-1}$, the inverse of $g$.
3. (associativity) Given $g_1, g_2, g_3$, we have
$$g_1(g_2 \cdot g_3) = (g_1 \cdot g_2)g_3.$$

If, in addition, the group satisfies

4. (commutative) Given $g_1, g_2 \in G$, we have
$$g_1 \cdot g_2 = g_2 \cdot g_1.$$

then $G$ is an **Abelian** group.

Now, we observe some interesting properties that follow from the group axioms.

**Proposition 2.2.1: Group Properties**

Let $G$ be a group.

1. The identity element is unique.
2. Each element of $G$ has only one inverse.
3. If $g, h \in G$, then $(gh)^{-1} = h^{-1}g^{-1}$.
4. Given $g \in G$, $(g^{-1})^{-1} = g$.

**Proof of (b).** Suppose $g \in G$ and that both $h_1, h_2$ satisfy the inverse axiom. Then
$$g \cdot h_1 = e = g \cdot h_2.$$
By the inverse axiom, we multiply on the left by an inverse of $g$:
$$e \cdot h_1 = e \cdot h_2$$
$$h_1 = h_2.$$
Thus the inverse is unique. □

**Definition 2.2.2: Order**

- The **order** of a group $G$ is denoted $\#G$ or $|G|$ is the number of elements in $G$ if finite, and $\infty$ if infinite.
- If $G$ is a group and $g \in G$, the smallest $n$ in which $g^n = e$ is called **the order**

**of** $g$. If no $n$ exists, we say $g$ has infinite order.

**Proposition 2.2.2: Individual Order and Group Order**

Suppose $G$ is a finite group and suppose $g^n = e$. Then the order of $g$ divides $n$.

**Proof.** Let $m$ be the order of $g \in G$; then $m$ is the smallest positive integer such that $g^m = e$. Dividing $n$ by $m$ yields

$$n = mq + r,\ q, r \in \mathbb{Z}, 0 \leq r < m.$$

In other words, dividing $n$ by $m$ leaves a quotient $q$ and a remainder $r$.
Using this equality together with $g^n = g^m = e$, we have

$$e = g^n = g^{mq+r} = (g^m)^q \cdot g^r = e^q \cdot g^r = g^r.$$

Hence $g^r = e$, and $r \in [0, m)$. But by definition, $m$ is the smallest integer such that $g^m = e$. Therefore $r = 0$, and $n = mq$, and so $m$, the order of $g$, divides $n$. □

**Proposition 2.2.3: Order of Inverse**

Let $G$ be a finite group, and $g \in G$. Then $|g| = \left|g^{-1}\right|$.

**Proof.** Let $|g| = n$; then $g^n = e$. From this, we get

$$e = (g \cdot g^{-1})^n = g^n \cdot (g^{-1})^n = e \cdot (g^{-1})^n,$$

and so $(g^{-1})^n = e$.
Now we show that $\left|g^{-1}\right| = n$. Suppose $\left|g^{-1}\right| = m$, and $m < n$. Then

$$e = g^n \cdot (g^{-1})^m = g^{n-m}.$$

But we know that $|g| = n$, or equivalently, $n$ is the smallest positive integer such that $g^n = e$; hence $g^{n-m} = e$ is a contradiction. Thus $m = n$, and so $\left|g^{-1}\right| = n$. □

### §2.2.1 Examples of Groups

**Example 3.** *$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are all Abelian groups with respect to addition. However, $\mathbb{Z}$ is not a group with respect to multiplication, as the multiplicative inverse does not exist. Additionally, $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ are not groups with respect to multiplication, due to zero; but $\mathbb{Q} \setminus \{0\}, \mathbb{R} \setminus \{0\}$, and $\mathbb{C} \setminus \{0\}$ are all groups under multiplication.*

**Example 4.** *Let $\mathbb{Z}/m\mathbb{Z}$ be the set of integers modulo $m$. Then $\mathbb{Z}/m\mathbb{Z}$ is a group under addition modulo $m$, $+_m$ ; $\mathbb{Z}/m\mathbb{Z}$ is finite with order $m$. We also observe that $\mathbb{Z}/m\mathbb{Z}$ is a cyclic group.*

**Example 5.** *Let the set of $n \times n$ matrices be $M_n$. Then $M_n$ is an Abelian group under addition, but not multiplication (since not all matrices have inverses).*

*Let*

$$GL_n(\mathbb{R}) = \{M \in M_n \mid \det(M) \neq 0\}$$

*denote the **general linear group**. Then $GL_n(\mathbb{R})$ is a non-Abelian group under matrix multiplication.*

### §2.2.2 Cyclic Groups

**Definition 2.2.3: Cyclic Groups**

A group $G$ is **cyclic** if there is a $g \in G$ such that

$$G = \{\ldots, g^{-2}, g^{-1}, e \text{ (or } g^0), g, g^2, g^3, \ldots\}.$$

We call $g$ a **generator**.
In general, for $n \geq 1$, the **abstract cyclic group order** $n$ is the set

$$\mathcal{C}_n = \{g_0, g_1, \ldots, g_{n-1}\}$$

together with the composition rule

$$g_i \cdot g_j = \begin{cases} g_{i+j}, & i+j < n \\ g_{i+j-n}, & i+j \geq n \end{cases}$$

The identity element of $\mathcal{C}_n$ is $g_0$, and the inverse of $g_i$ is $g_{n-i}$ (except $g_0$, whose inverse is $g_0$). Further, $\mathcal{C}_n$ is an Abelian group, as $g_{i+j} = g_{j+i}$.

Some examples of cyclic groups are $\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$; both have generators 1. Another one is the permutation group.

**Definition 2.2.4: Permutation Groups**

Given $X$ a set, let $S_X$ denote the **symmetric group of** $X$, or the group of permutations of $X$. If

$$X = \{1, \ldots, n\},$$

we use the notation $S_n$.
Let $P_n$ be a regular $n$-gon with vertices $1, \ldots, n$. The group of transformations of $D_n$ (e.g. rotations, reflections, and compositions of such) is called the **dihedral group** $D_n$. We will later prove that $D_n$ has order $2n$.

## §2.3 Group Homomorphisms

Suppose that $G, G'$ are groups, and suppose that $\phi$ is a function

$$\phi : G \longrightarrow G'$$

from elements of $G$ to elements of $G'$. Many functions exist, but we're interested in the ones that preserve the "structure", or *group-i-ness*, of $G$ and $G'$. But what makes a group a group? Specifically, groups are **associative**, and have **identity and inverse elements**. Thus, a function $\phi$ must preserve these qualities. We call such structure-preserving functions **homomorphisms**.

**Definition 2.3.1: Homomorphisms**

Let $G_1, G_2$ be groups. A **homomorphism** from $G_1$ to $G_2$ is a function

$$\phi : G_1 \to G_2$$

satisfying:

$$\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2).$$

In other words, the map $\phi$ preserves the group operations. (Note that the composition law is different on the left and right sides! The left is the composition law of $G_1$, while the right is the composition law of $G_2$.)

It turns out this property is enough to force the identities and inverses to exist under a function.

**Proposition 2.3.1**

Let $\phi : G \to G'$ be a homomorphism of groups.

1. Let $e \in G$ be the identity element of $G$. Then $\phi(e) \in G'$ is the identity element of $G'$.
2. Let $g \in G$, and let $g^{-1} \in G$ be its inverse. Then $\phi(g^{-1}) \in G$ is the inverse of $\phi(g)$.

**Proof.** 1. Observe that $e = e \cdot e$, and that $\phi$ is a homomorphism (and so $\phi(e) = \phi(e \cdot e) = \phi(e) \cdot \phi(e)$). Let $e' \in G'$ be the identity element of $G'$. Then

$$\begin{aligned} e' &= \phi(e) \cdot \phi(e)^{-1} \\ &= (\phi(e) \cdot \phi(e)) \cdot \phi(e)^{-1} \\ &= \phi(e) \cdot (\phi(e) \cdot \phi(e)^{-1}) \\ &= \phi(e) \cdot e' \\ &= \phi(e). \end{aligned}$$

Hence $e' = \phi(e)$.

2. We have

$$\begin{aligned} \phi(g^{-1}) \cdot \phi(g) &= \phi(g^{-1} \cdot \phi(g)) \\ &= \phi(e) \\ &= e'. \end{aligned}$$

The proof that $\phi(g) \cdot \phi(g^{-1}) = e'$ is similar. Hence $\phi(g^{-1})$ is the inverse of $\phi(g)$. □

**Example 6.** *Examples of homomorphisms:*

- *There exists a homomorphism from the dihedral group to the group* $\pm 1$*:*

$$\phi : D_n \to \{\pm 1\}$$

*, where* $\phi(\sigma) = 1$ *if rotation,* $\phi(\sigma) = -1$ *if flip.*

- *For* $n \geq m \geq 1$*, there is an injective homomorphism*

$$f : S_m \to S_n.$$

*Note that this homomorphism is not surjective. More generally, if* $X_1 \subseteq X_2$*, then there is an injective homomorphism* $f : S_{X_1} \to S_{X_2}$*.*

- *There is a homomorphism*

$$\log : (\mathbb{R}, \times) \to (\mathbb{R}, +)\,.$$

- *There is a homomorphism between the general linear group to the real numbers*

$$\begin{aligned} det : GL_n(\mathbb{R}) &\longrightarrow \mathbb{R} \\ AB &\longmapsto det(AB) = det(A) \cdot det(B). \end{aligned}$$

**Definition 2.3.2: Isomorphisms**

Groups $G_1, G_2$ are **isomorphic** if there exists a **bijective homomorphism** $f : G_1 \to G_2$. In this case, $f$ is called an **isomorphism**.

Interestingly, isomorphic groups are really the same group, but their elements are given different names.
We've now seen two examples of cyclic groups of order $n$: $\mathbb{Z}/n\mathbb{Z}$ and $\mathcal{C}_n$. Naturally, we wonder if these groups are actually different (from the perspective of group theory). Equivalently, **are these two groups isomorphic**?

**Example 7.** $\mathbb{Z}/n\mathbb{Z}$ *is isomorphic to* $\mathcal{C}_n$ *(*$\mathbb{Z}/n\mathbb{Z} \cong \mathcal{C}_n$*). Consider the map*

$$\begin{aligned} \phi : \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathcal{C}_n \\ a &\longmapsto \phi(a) = g_a. \end{aligned}$$

*Then* $\phi(a+b) = \phi(a) \cdot \phi(b)$ *by definition of group operations. So* $\phi$ *is a homomorphism.*
$\phi$ *is surjective since* $i \in \{0, \ldots, n-1\}$ *maps to* $g_i \in \{g_0, \ldots, g_{n-1}\}$*.*
*Since* $\mathbb{Z}/n\mathbb{Z}$ *and* $\mathcal{C}_n$ *both have* $n$ *elements,* $\phi$ *is injective as well.*
*So,* $\phi$ *is an isomorphism and* $\mathcal{C}_n \simeq \mathbb{Z}/n\mathbb{Z}$*.*

Note that if a group is isomorphic, there isn't necessarily a unique isomorphism. Consider the same isomorphism as above, except map $a \mapsto g_{a+1}$. This is also an isomorphism.

**Example 8.** *Given any group $G$, and an element $g \in G$, then multiplication by $g$ permutes the elements of $G$. This gives rise to an injective homomorphism $\phi : G \to S_G$.*

*This implies that by knowing every symmetric group, one knows much about every other group.*

## §2.4 Subgroups, Cosets, and Lagrange's Theorem

In all mathematics, a three-step process exists for studying complicated objects.

1. Deconstruction: Break your object into smaller and simpler pieces.
2. Analysis: Analyze the smaller, simpler pieces.
3. Fit the pieces back together.

For a group $G$, a natural way to form a smaller and simpler piece is by taking subsets $H \subseteq G$ that are themselves groups.

**Definition 2.4.1: Subgroups**

Let $G$ be a group. A **subgroup of** $G$ is a subset $H \subset G$ that is itself a group under $G$'s group law. Explicitly, $H$ needs to satisfy

1. (Closure Under Composition) For every $h_1, h_2 \in H$, $h_1 \cdot h2 \in H$
2. The identity element $e$ is in $H$.
3. For every $h \in H$, its inverse $h^{-1}$ is in $H$.

This is sometimes denoted $H < G$.

Note that since $H$ uses $G$'s composition law, associativity is automatically satisfied. If $H$ is finite, the **order** of $H$ is the number of elements in $H$.

**Proposition 2.4.1: Easier Subgroup Checking**

Let $G$ be a group, and $H \subseteq G$ a subset. If

- $H \neq \varnothing$
- For every $h_1, h_2 \in H$, the element $h_1 h_2^{-1}$ is in $H$

then $H$ is a subgroup of $G$.

**Proof.** Clearly, $H \neq \varnothing$ (otherwise the identity would not be in $H$).
To show that $e \in H$, let $h_2 = h_1$. Then

$$h_1 \cdot h_2^{-1} = h_1 \cdot h_1^{-1} = e \in H.$$

Thus the identity is in $H$.
To show that $\forall h \in H,\ h^{-1} \in H$, let $h_1 = e$. Then

$$h_1 \cdot h_2^{-1} = e \cdot h_2^{-1} = h_2^{-1} \in H.$$

Thus for any $h \in H$, its inverse $h^{-1}$ is in H.
To show closure, observe that for any $h \in H$, $h^{-1} \in H$ (from above), and that $(h^{-1})^{-1} = h$. Let $h_2 = h^{-1}$. Then

$$h_1 \cdot h_2^{-1} = h_1 \cdot (h^{-1})^{-1} = h_1 \cdot h \in H.$$

Thus for any $h1, h \in H$, we have $h_1 \cdot h \in H$. Thus $H$ is closed.

Hence $H$ is a subgroup of $G$. □

**Example 9.** *Every group $G$ has at least two subgroups, the **trivial subgroup** $\{e\}$ consisting of only the identity element, and the entire group $G$.*

**Example 10.** *Let $G$ be a group, and let $g \in G$ be an element of order $n$. The **cyclic subgroup of $G$ generated by** $g$, denoted $\langle g \rangle$, is the set*

$$\langle g \rangle = \{\ldots, g^{-2}, g^{-1}, e, g^1, g^2, g^3, \ldots\}.$$

*It is isomorphic to the cyclic group $\mathcal{C}_n$.*

*If $g$ has infinite order, then $\langle g \rangle \cong \mathbb{Z}$ ($\langle g \rangle$ is isomorphic to $\mathbb{Z}$).*

**Example 11.** *More examples of subgroups:*

- *Let $d \in \mathbb{Z}$; then we can form a subgroup of $\mathbb{Z}$ using multiples of $d$, or $d\mathbb{Z}$.*
- *The set of rotations in the dihedral group $\mathcal{D}_n$ is a subgroup of $\mathcal{D}_n$.*

Every group homomorphism has an associated subgroup, the **kernel**, which can be a convenient check to see if the homomorphism is injective.

**Definition 2.4.2: Kernel**

Let $\phi : G \to G'$ be a group homomorphism. The **kernel of** $\phi$, denoted $\ker(\phi)$, is the set of elements of $G$ that are sent to the identity element of $G'$,

$$\ker(\phi) = \{g \in G \mid \phi(g) = e'\}.$$

**Example 12.** *The kernel of the determinant homomorphism*

$$\det : \mathrm{GL}_n(\mathbb{R}) \longrightarrow \mathbb{R} \setminus \{0\}.$$

*is*

$$\ker(\det) = \{A \in \mathrm{GL}_n(\mathbb{R}) \mid \det(A) = 1\}.$$

We now observe two important properties of the kernel.

**Proposition 2.4.2: Kernel Properties**

Let $\phi : G \to G'$ be a group homomorphism.

1. $\ker(\phi)$ is a subgroup of $G$.
2. $\phi$ is injective if and only if $\ker(\phi) = \{e\}$.

**Proof.** We know that $\phi(e) = e'$, so $e \in \ker(\phi)$. Next, let $g_1, g2 \in \ker(\phi)$. By the homomorphism property,

$$\phi(g_1 \cdot g_2) = \phi(g_1) \cdot \phi(g_2) = e' \cdot e' \in G',$$

so $g_1, g_2 \in \ker(\phi)$. Finally, for $g \in \ker(\phi)$, we know $\phi(g^{-1}) = \phi(g)^{-1} = e'^{-1} = e$, so $g^{-1} \in \ker(\phi)$ Thus, $\ker(\phi)$ is a subgroup of $G$.

Now, we know again that $e \in \ker(\phi)$ (since $\phi(e) = e'$). If $\phi$ is injective, by definition $\ker(\phi) = \{e\}$ (at most one element $g \in G$ satisfies $\phi(g) = e'$).
Now, suppose $\ker(\phi) = \{e\}$. Let $\phi(g_1) = \phi(g_2)$ for some $g_1, g_2 \in G$. Observe that $g_2^{-1} \in G$, and $\phi(g_2^{-1}) = \phi(g_2)^{-1}$. Then

$$\phi(g_1) = \phi(g_2) \implies \phi(g_1) \cdot \phi(g_2)^{-1} = \phi(g_2) \cdot \phi(g_2)^{-1} = e',$$

and so $\phi(g_1) \cdot \phi(g_2)^{-1} = \phi(g_1 \cdot g_2^{-1}) = e'$, which means $g_1 \cdot g_2^{-1} \in \ker(\phi) = \{e\}$. Hence $g_1 \cdot g_2^{-1} = e \implies g_1 = g_2$, and so $\phi$ is injective. □

### §2.4.1 Cosets

We can use a subgroup $H$ of a group $G$ to break $G$ into pieces, called **cosets of** $H$.

**Definition 2.4.3: Cosets**

Let $G$ be a group, and let $H < G$ be a subgroup. For each $g \in G$, the (left) **coset of** $H$ **attached to** $g$ is the set

$$gH = \{gh \mid h \in H\}.$$

In other words, $gH$ is the resulting set we multiply $g$ by every element $h \in H$.

Note that $gH$ is **not** necessarily a subgroup of $H$; sometimes $e \notin gH$.

We now prove several properties of cosets that help explain their importance.

**Proposition 2.4.3: Properties of Cosets**

Let $G$ be a finite group, and let $H < G$.

1. Every element in $G$ is in some coset of $H$.
2. Every coset of $H$ has the same number of elements (namely, $|H|$).
3. Let $g_1, g_2 \in G$. Then the cosets $g_1H$ and $g_2H$ satisfy either
$$g_1H = g_2H \text{ or } g_1H \cap g_2H = \varnothing.$$
In other words, $g_1H$ and $g_2H$ are either equal or disjoint.

**Proof.**

1. Let $g \in G$. Since $e \in H$ for any subgroup $H < G$, the coset $gH$ contains $g \cdot e = g$.
2. Let $g \in G$. To prove that the cosets $gH$ and $H$ have the same number of elements, we show the map
$$\begin{aligned} F : H &\longrightarrow gH \\ h &\longmapsto F(h) = gh \end{aligned}$$
is a bijective map from $H$ to $gH$.
We first check that $F$ is injective. Suppose $h_1, h_2 \in H$ satisfy $F(h_1) = F(h_2)$. Then $gh_1 = gh_2$, and multiplying by $g^{-1}$, we get $h_1 = h_2$. Hence $F$ is injective.
For surjectivity, observe that every element of $gH$ looks like $gh$ for some $h \in H$, and $F(h) = gh$, so every element of $gH$ is the image of some element of $H$. Hence $F$ is surjective.
Thus $F$ is bijective, so $H$ and $gH$ have the same number of elements. Since this is true for any $g \in G$, every coset of $H$ has the same number of elements.
3. If $g_1H \cap g_2H = \varnothing$, we are done, so assume the two cosets are not disjoint. Then there are some elements $h_1, h_2 \in H$ such that $g_1h_1 = g_2h_2$. Since $h_1^{-1} \in H$, we rewrite this as $g_1 = g_2h_2h_1^{-1}$. Now, take any element $a \in g_1H$. $a$ is of the form $g_1h$ for some $h \in H$. Then
$$a = g_1h = g_2h_2h_1^{-1}h \in g_2H,$$
as $H$ is a subgroup, so $h_2h_1^{-1}h \in H$. Hence $g_1H \subseteq g_2H$; and from above, every coset has the same number of elements, so $g_1H \subseteq g_2H \implies g_1H = g_2H$.

□

These properties lead to a fundamental divisibility property for the orders of subgroups.

**Theorem 2.4.1: Lagranges Theorem**

Let $G$ be a finite group, and let $H < G$. Then the order of $H$ divides the order of $G$; or, $|G| = k\,|H|\,,\ k \in \mathbb{Z}$.

**Proof.** We start by choosing $g_1, \ldots, g_k \in G$ so that $g_1H, \ldots, g_kH$ is a list of every different coset of $H$. Since every element of $G$ is in some coset of $H$, we have that $G$ is equal to the union of the cosets of $H$, namely
$$G = g_1H \cup \ldots \cup g_kH.$$

Additionally, we know that distinct cosets share no elements, so if $i \neq j$, then $g_iH \cap g_jH = \varnothing$. Thus the union of cosets is a disjoint union, so the number of elements in $G$ is the sum of the number of elements in each coset:

$$|G| = |g_1H| + \ldots + |g_kH|.$$

But we know that every coset of $H$ has the same number of elements, so $|g_iH| = |H|$. Thus, we get

$$|G| = k\,|H|.$$

Thus the order of $G$ is a multiple of the order of $H$. □

**Definition 2.4.4: Index**

Let $G$ be a group, and $H < G$. The **index of $H$ in $G$**, denoted $(G : H)$, is the number of distinct cosets of $H$. In Lagrange's Theorem, the index $(G : H) = k$; so

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

**Corollary 2.4.1: Extension of Lagrange's Theorem to Finite Groups**

Let $G$ be a finite group, and let $g \in G$. Then the order of $g$ divides the order of $G$.

**Proof.** The order of the subgroup $\langle g \rangle$ generated by $G$ is equal to the order of the element $g$, and Lagrange's Theorem tells us that the order of $\langle g \rangle$ divides the order of $G$. □

We now give one application of Lagrange's Theorem, which marks the beginning of a long and ongoing mathematical journey that strives to classify finite groups according to their orders.

**Proposition 2.4.4: Prime-Ordered Groups**

Let $p$ be a prime, and let $G$ be a finite group of order $p$. Then $G$ is isomorphic to the cyclic group $\mathcal{C}_p$.

**Proof.** Since $p \geq 2$, we know that $G$ contains more than just the identity element, so we choose some non-identity element $g \in G$.
By Lagrange's Theorem, we know that the order of the subgroup $\langle g \rangle$ divides the order of $G$. But since $|G| = p$ is prime, the order of $\langle g \rangle$ is either 1 or $p$; and since $\langle g \rangle$ contains both $e$ and $g$ (and so $\langle g \rangle > 1$), we know $|\langle g \rangle| = p = |G|$. Thus the subgroup $\langle g \rangle$ has the same number of elements as the full group, so they are equal: $\langle g \rangle = G$.
Now, we denote the cyclic group $\mathcal{C}_p = \{g_0, g_1, \ldots, g_{p-1}\}$. We obtain an isomorphism

$$\begin{aligned} C_p &\longrightarrow G \\ g_i &\longmapsto g^i. \end{aligned}$$

Thus $G$ is isomorphic to $\mathcal{C}_p$. □

**Remark 3.** *The vast theory of finite groups has many fascinating (and frequently unexpected) results, with easy to understand statements, yet surprisingly intricate proofs. Two such theorems are stated.*

**Theorem 2.4.2**

Let $p$ be a prime number, and let $G$ be a group of order $p^2$. Then $G$ is an Abelian group.

On the other hand, we know that there exist non-Abelian groups of order $p^3$. For instance, $\mathcal{D}_4$ and the quaternion group $\mathcal{Q}$ are non-Abelian groups of order $8 = 2^3$.
The next result is a partial converse of Lagrange's Theorem.

**Theorem 2.4.3: Sylow's Theorem**

Let $G$ be a finite group, let $p$ be a prime, and suppose $p^n$ divides $|G|$ for some power $n \geq 1$. Then $G$ has a subgroup or order $p^n$.

One might hope, more generally, that if $d$ is any number that divides the order of $G$, then $G$ has a subgroup of order $d$. Unfortunately, this is not true; however, we have not yet seen a counterexample.
Both theorems will be proved later.

## §2.5 Products of Groups

Subgroups provide a way to break complicated objects (groups) down into smaller, simpler pieces. We now look at a way in which two smaller groups can be used to build a larger group.

**Definition 2.5.1: Products of Groups**

Let $G_1, G_2$ be groups. The **product** of $G_1$ and $G_2$ is the group whose elements consist of ordered pairs

$$G_1 \times G_2 = \{(a, b) \mid a \in G_1 \text{ and } b \in G_2\},$$

and whose group operation is performed separately on each component. In other words, if the group operation of $G_1 \times G_2$ is $*$, the group operation of $G_1$ is $\cdot$, and the group operation of $G_2$ is $\circ$, we have

$$(a_1, b_1) * (a_2, b_2) = (a_1 \cdot a_2, b_1 \circ b_2).$$

It is clear that the identity element of $G_1 \times G_2$ is $(e_1, e_2)$, and the inverse of an element $(a, b) \in G$ is given by

$$(a^{-1}, b^{-1}).$$

More generally, we can take any list of groups $G_1, \ldots, G_n$ and form the product

group

$$G_1 \times \ldots \times G_n.$$

**Remark 4.** *We observe that $G = G_1 \times G_2$ has order $|G_1| \cdot |G_2|$.*

**Example 13.** *For any non-zero numbers $m, n$, there is a homomorphism*

$$\begin{aligned} \mathbb{Z}/mn\mathbb{Z} &\longrightarrow (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \\ a \mod (mn) &\longmapsto (a \mod m, a \mod n). \end{aligned}$$

*We claim that if $\gcd(m,n) = 1$, then it is an isomorphism. To see this, suppose that $a \mod mn$ is in the kernel. Then*

$$a \equiv 0 \mod m \text{ and } a \equiv 0 \mod n.$$

*In other words, $a$ is divisible by both $m$ and $n$, and then the assumption that $\gcd(m,n) = 1$ implies that $a$ is divisible by $mn$. Thus $a \equiv 0 \mod mn$, which proves that the kernel of the homomorphism is $\{0\}$ (and thus the homomorphism is injective). Further, since the finite set $\mathbb{Z}/mn\mathbb{Z}$ has the same number of elements as $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ ($\mathbb{Z}/mn\mathbb{Z}$ has $mn$ elements, while $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ has $m \cdot n = mn$ elements), so it is surjective, and thus the homomorphism is an isomorphism.*

One interpretation of this example is that it tells us that if $\gcd(m,n) = 1$, then the large group $\mathbb{Z}/mn\mathbb{Z}$ may be broken down into the product of two smaller groups $\mathbb{Z}/m\mathbb{Z}$ and $\mathbb{Z}/n\mathbb{Z}$. Repeated applications demonstrate that any finite cyclic group is isomorphic to the product of cyclic groups of prime power order. The following theorem extends this to all finite Abelian groups.

**Theorem 2.5.1: Structure Theorem for Finite Abelian Groups**

Let $G$ be a finite Abelian group. Then there are integers $m_1, \ldots, m_r$) so that

$$G \cong (\mathbb{Z}/m_1\mathbb{Z}) \times (\mathbb{Z}/m_2\mathbb{Z}) \times \ldots \times (\mathbb{Z}/m_r\mathbb{Z}).$$

**Example 14.** *(**Projects and Inclusions**) Products of groups come with two natural **projection homomorphisms***

$$\begin{aligned} p_1 : G_1 \times G_2 &\longrightarrow G_1, & p_2 : G_1 \times G_2 &\longrightarrow G_2, \\ (a,b) &\longmapsto a, & (a,b) &\longmapsto b, \end{aligned}$$

*and two natural **inclusion homomorphisms***

$$\begin{aligned} \iota_1 : G_1 &\longrightarrow G_1 \times G_2 & \iota_2 : G_2 &\longrightarrow G_1 \times G_2 \\ a &\longmapsto (a, e_2) & b &\longmapsto (e_1, b). \end{aligned}$$

*The inclusion maps are clearly injective, but the projections have kernels*

$$\ker(p_1) = \{e_1\} \times G_2 \text{ and } \ker(p_2) = G_1 \times \{e_2\}.$$

# Chapter 3

# Rings: Part I

Unlike groups, which were completely new, the concept of a **ring** is mildly familiar! Some examples of rings include:

- $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are rings ($\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are actually **fields**, a special type of ring; such is a discussion for later)
- The set of integers modulo $m$ is a ring

These examples all share something in common: they each have two operations, "addition" and "multiplication", and each operation individually satisfies some axioms, along with the g̃reat and powerfuld̃istributive law.
In general, a ring is a set with two operations satisfying a bunch of axioms that are modeled after the properties of addition and muliplication of integers. We will later formalize this; but first, a little number theory.

## §3.1 Review of Number Theory

### §3.1.1 Equivalence Relations

We first introduce the notion of **equivalence relations**; while not strictly related to number theory, equivalence relations will be significant for modular arithmetic.

**Definition 3.1.1: Equivalence Relations**

An **equivalence relation** on a set $S$ is a relation "$\sim$" satisfying

1. **Reflexivity**: For $a \in S$, $a \sim a$
2. **Symmetry**: For $a, b \in S$, $a \sim b$ implies $b \sim a$
3. **Transitivity**: For $a, b, c \in S$, if $a \sim b$ and $b \sim c$, then $a \sim c$.

$a \sim b$ means $a$ is "related" to $b$; $a \not\sim b$ means $a$ is "not related" to $b$.
Given an $a \in S$, the **equivalence class** of $a$ is

$$S_a = \{b \in S \mid b \sim a\}.$$

Note that $S_a$ is never empty; it always contains $a$.

Some examples of equivalence relations are equality ($=$) and congruence $\mod m$; on the other hand, order (e.g. $\leq$) is **not** an equivalence relation (symmetry does not hold). We now look further into the congruence $\mod m$ equivalence relation.

**Example 15.** *Given $a \in \mathbb{Z}$, $b \equiv a \mod m$ iff*

$$n|b-a \iff b-a = kn,\ k \in \mathbb{Z} \iff b = a+kn.$$

*So $\mathbb{Z}_a$ actually forms a coset of $n\mathbb{Z}$:*

$$\begin{aligned}\mathbb{Z}_a &= \{b \in \mathbb{Z} \mid b \equiv a \mod m\} \\ &= \{a+kn \mid k \in \mathbb{Z}\} \\ &= a + n\mathbb{Z}.\end{aligned}$$

*That is, each equivalence class for congruence $\mod m$ is actually a coset of $m\mathbb{Z}$ in $\mathbb{Z}$:*

$$\mathbb{Z}/m\mathbb{Z} = \text{ set of cosets of } m\mathbb{Z} \text{ in } \mathbb{Z}\ .$$

**Theorem 3.1.1**

Let $S$ be a set with an equivalence relation $\sim$. Then

1. If $a, b \in S$, then either
$$S_a \cap S_b = \varnothing \text{ or } S_a = S_b.$$
2. Let $\{C_i\}_{i \in I}$ be the disjoint equivalence classes of $S$. Then
$$S = \coprod_{i \in I} C_i.$$
In particular, if $S$ is finite, then
$$|S| = |C_1| + \ldots + |C_n|\,.$$

## §3.1.2 Modular Arithmetic

We first observe an important characteristic of the gcd:

**Proposition 3.1.1**

Given integers $u, v$, there exists integers $x, y$ such that

$$ux + vy = \gcd(u, v).$$

$x, y$ can be found using the **Euclidean algorithm**.

This result leads us to an important proposition that facilitates equivalence relations in $\mathbb{Z}/m\mathbb{Z}$:

**Proposition 3.1.2**

$ax \equiv b \mod m$ is solvable if and only if

$$\gcd(a, m) | b.$$

**Proof.** First, suppose $ax \equiv b \mod m$ is solvable. It follows that

$$m | ax - b.$$

So, we can find $k \in \mathbb{Z}$ such that $ax - b = km$ is $b = ax - km$. From this, we get that any integer that divides both $a$ and $m$ must divide $b$; in particular, $\gcd(a, m) | b$. Conversely, if $\gcd(a, m) | b$, then $b = c \gcd(a, m)$ for some $c \in \mathbb{Z}$. Using the Euclidean algorithm to find $x, y$ with

$$ax + ym = \gcd(a, m),$$

and multiplying by $c$ to get

$$a(cx) + y(mc) = c \gcd(a, m) = b,$$

we have $a(cx) \equiv b \mod m$, so the congruence is solvable. □

One natural followup is:

$$\text{When does } ax \equiv 1 \mod m \text{ have a solution?}$$

From the proposition above, **only if** $\gcd(a, m) = 1$; that is, only if $a$ and $m$ are relatively prime. Rephrasing, if $a$ and $m$ are relatively prime, then $a$ has a unique multiplicative inverse in $\mathbb{Z}/m\mathbb{Z}$.

## §3.2 Abstract Rings and Ring Homomorphisms

**Definition 3.2.1: Rings**

A **ring** $R$ is a set with two operations, generally called **addition** and **multiplication** and written

$$\underbrace{a + b}_{\text{addition}} \qquad \text{and} \qquad \underbrace{a \cdot b \text{ or } ab}_{\text{multiplication}}$$

satisfying the following axioms:

1. The set $R$ with addition law $+$ is an Abelian group, with identity 0 (or $0_R$).
2. The set $R$ with multiplication law $\cdot$ is a **monoid** (associative, identity, **but no inverse**), with identity 1 (or $1_R$[a]).
3. **Distributive Law**: For all $a,\ b,\ c \in R$, we have
$$a \cdot (b + c) = a \cdot b + a \cdot c \qquad \text{and} \qquad (b + c) \cdot a = b \cdot a + c \cdot a.$$
4. If, in addition to these three properties, $a \cdot b = b \cdot a$ for all $a,\ b \in R$, then $R$ is a **commutative** ring.

[a] to avoid the trivial ring, we require $1_R \neq 0_R$; however, this is not strictly required

Experience with the integers seems to suggest that $0_R \cdot a = 0_R$, and $(-a) \cdot (-b) = a \cdot b$; yet why are these true? $0_R$ is the definition of the identity element for *addition*, so why should it say anything about *multiplication*? Similarly, $-a$ relates to the definition of *additive* inverse, but what does that tell us about its product with other elements in $R$? To show these intuitively obvious claims, we need the distributive law.

**Proposition 3.2.1**

1. $0_R \cdot a = 0_R$ for all $a \in R$.
2. $(-a) \cdot (-b) = a \cdot b$ for all $a,\ b \in R$. In particular, we have $(-1_R) \cdot a = -a$.

**Proof.** 1. Note that $1_R = 0_R + 0_R$. Then

$$\begin{aligned} a &= 1_R \cdot a && [1_R \text{ is the muliplicative identity }] \\ &= (1_R + 0_R) \cdot a && [\text{ from above }] \\ &= 1_R \cdot a + 0_R \cdot a && [\text{ from distributivity }] \\ &= a + 0_R \cdot a. \end{aligned}$$

Adding $-a$ to both sides, we get

$$0_R = 0_R + 0_R \cdot a,$$

and so $0_R \cdot a = 0_R$.

2. First, we show that $(-1_R) \cdot a = -a$:

$$\begin{aligned} a + (-1_R) \cdot a &= 1_R \cdot a + (-1_R) \cdot a \\ &= (1_R + -1_R) \cdot a \\ &= 0_R \cdot a \\ &= 0_R. \end{aligned}$$

Hence $(-1_R) \cdot a$ is the inverse of $a$, and so $-a = (-1_R) \cdot a$.

Now, observe that $-ab = (-a)b$ (this proof is left as an exercise for the reader). Then

$$\begin{aligned} (-a) \cdot (-b) + -ab &= (-a) \cdot (-b) + (-a) \cdot b \\ &= (-a) \cdot (-b + b) \\ &= (-a) \cdot 0_R \\ &= 0_R. \end{aligned}$$

Thus $(-a) \cdot (-b)$ is the inverse of $-ab$, and so $(-a) \cdot (-b) = ab$.

□

Just like groups, we want to investigate maps

$$\phi : R \to R'$$

from one ring to another that respect the *ring-i-ness* of $R$ and $R'$. Since rings are characterized by their addition and multiplication properties, we get the following definition.

**Definition 3.2.2: Ring Homomorphisms**

Let $R$, $R'$ be rings. A **ring homomorphism** from $R$ to $R'$ is a function $\phi : R \to R'$ satisfying[a]

1. $\phi(1_R) = 1_{R'}$
2. $\phi(a+b) = \phi(a) + \phi(b)$ for all $a, b \in R$
3. $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$ for all $a, b \in R$

The **kernel** of $\phi$ is the set of elements that are sent to 0:

$$\ker(\phi) = \{a \in R \mid \phi(a) = 0_{R'}\}.$$

As with groups, $R$ and $R'$ are **isomorphic** if here is a bijective ring homomorphism $\phi : R \to R'$, and we call such a map an **isomorphism**.

[a] we have the first axiom to disallow the boring and trivial zero map $\phi : R \to R'$, $\phi(a) = 0_{R'}$.

## §3.3 Interesting Examples of Rings

As described before, we have the four rings

$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C};$$

we say that $\mathbb{Z}$ is a **subring** of $\mathbb{Q}$, and similarly for the others. Another example is the ring of integers modulo $m$, $\mathbb{Z}/m\mathbb{Z}$.

**Example 16.** *(Integers Modulo $m$ $\mathbb{Z}/m\mathbb{Z}$) We construct a ring $\mathbb{Z}/m\mathbb{Z}$ with integers, and determining equality if $a-b$ is a multiple of $m$. Formally, we define an equivalence relation on $\mathbb{Z}$ by the rule*

$$a \equiv b \text{ if } a - b = km \text{ for some } k \in \mathbb{Z}. \text{ We say } a \text{ is } \textbf{\textit{congruent}} \text{ to } b \text{ modulo } m.$$

*While $\mathbb{Z}/m\mathbb{Z}$ is not a subring of $\mathbb{C}$, there is a natural homomorphism*

$$\phi : \mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z}, \ \phi(a) = a \mod m$$

*that assigns the integer $a$ to its equivalence class of all integers congruent to $a \mod m$. This homomorphism $\phi$ is called the **reduction** $\mod m$ **homomorphism**. The kernel of $\phi$ is the set of all multiples of $m$.*

**Example 17.** *Another subring of $\mathbb{C}$ is the **ring of Gaussian integers** $\mathbb{Z}[i]$, which consists of*

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}.$$

*The quantity $i$ represents the imaginary number (e.g. $\sqrt{-1} = i$), and addition and multiplication are defined following the usual rules of adding and multiplying complex numbers.*

*In general, we can define a ring*

$$\mathbb{F}[z] = \{a + bz \mid a, b \in \mathbb{F}\}$$

*for an arbitrary field and any number $z$.*

**Example 18.** *Polynomials offer another method of creating bigger rings from already known rings. For any commutative ring $R$, we use $R$ to build the* ***ring of polynomials over*** *$R$,*

$R[x] = \{$ *polynomials* $a_0 + a_1x + \ldots + a_dx^d$ *of all possible degrees with coefficients* $a_0, a_1, \ldots, a_d \in R$ $\}$.

*A common one we've seen before is $\mathbb{R}[x]$, but the rules of adding and multiplying polynomials hold in any commutative ring. Indeed, the rule for multiplying polynomials is a result of the distributive law!*
*Now, consider the polynomial*

$$f(x) = a_0 + a_1x + \ldots + a_dx^d \in R[x].$$

*Then for any element $c \in R$, we can* ***evaluate*** *$f$* ***at*** *$c$ simply by substituting $x$ with $c$:*

$$f(c) = a_0 + a_1c + \ldots + a_dc^d \in R.$$

*When we first studied polynomials, we viewed them as functions, i.e. $f(x)$ defined a function $f : R \to R$. While these polynomial functions are interesting, they are almost never ring homomorphisms!*
*Thus, we take another approach: using a particular $c \in R$, we define a function from the ring of polynomials $R[x]$ to the ring $R$:*

$$E_c : R[x] \longrightarrow R,\ E_c(f) = f(c).$$

*We call $E_c$ the* ***evaluation at*** *$c$* ***map****. If $R$ is commutative, then $E_c$ is a ring homomorphism, and its kernel is the set of polynomials that have a factor of $x - c$ (since $a_0 + a_1(c - c) + \ldots + a_d(c - c)^d = 0_R$).*

**Example 19.** *One famous non-commutative ring is the* ***ring of quaternions****,*

$$\mathbb{H} = \{a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} \mid a, b, c, d \in \mathbb{R}\}.$$

*$i$, $j$, and $k$ are the different square roots of $-1$, and although they commute with elements in $\mathbb{R}$, they don't commute with each other. To multiply two quaternions, we first use the distributive law, and then apply multiplication as specified in the quaternion group. Since $\mathcal{Q}$ is a non-commutative group, the ring of quaternions $\mathbb{H}$ is a non-commutative ring.*

The ring of quaternions played an important role in the development of math and physics because of the **cancellation law**:

if you know that $\alpha$ and $\beta$ are real (or complex) numbers satisfying $\alpha\beta = 0$, then either $\alpha = 0$ or $\beta = 0$. It turns out that the same is true with quaternions!

**Example 20.** *(Matrix Rings) There are also rings with matrix elements. Let*

$$M_2(\mathbb{R}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R}\}$$

*denote the set of $2 \times 2$ matrices with real entries. Matrices are added by their corresponding elements, and multiplied using matrix multiplication. With these operations, $M_2(\mathbb{R})$ is a non-commutative ring. However, $M_2(\mathbb{R})$ does not satisfy the cancellation law:*

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

*More generally, the set of $n \times n$ matrices over a ring $R$, $M_n(R)$, forms a non-commutative ring.*
*There are many interesting homomorphisms from rings to matrix rings. For example,*

$$\mathbb{C} \hookrightarrow M_2(\mathbb{R}),\ x + yi \mapsto \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$$

*is an injective ring homomorphism (proof left as an exercise).*

**Example 21.** *For every ring $R$, there is a unique homomorphism*

$$\phi : \mathbb{Z} \longrightarrow R.$$

*To understand why, note that by homomorphism requirements we must have $\phi(1) = 1_R$, and we must also have*

$$\phi(n) = \phi(1 + \ldots + 1) = \phi(1) + \ldots + \phi(1).$$

*But we also need $\phi(0) = 0_R$, and $\phi(-n) = -\phi(n)$, so there are really no other choices for $\phi$. In other words, the requirements that*

$$\phi(1) = 1_R \text{ and } \phi : \mathbb{Z} \longrightarrow R \text{ is a homomorphism}$$

*means that there is only one possibility for $\phi$. One still needs to check that $\phi$ is a homo morph ism; such a proof is left as an exercise.*

## §3.4 Important Properties of Rings

Some rings, such as $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ have the special property that every non-zero element has a multiplicative inverse. These rings are special, and we call them **fields.**

**Definition 3.4.1: Fields**

A **field** is a commutative ring $R$ with the property that every non-zero element of $R$ has a multiplicative inverse. In other words, for every non-zero $a \in R$, there is a $b \in R$ satisfying $ab = 1$.

**Example 22.** *In addition to* $\mathbb{Z}$, $\mathbb{R}$, *and* $\mathbb{C}$, *there are also* ***finite fields****. One important example of a finite field is the ring* $\mathbb{Z}/p\mathbb{Z}$, *where* $p$ *is a prime. This follows from number theory: if* $p \not| a$, *then* $\gcd(p, a) = 1$, *and so* $ax \equiv 1 \mod p$ *is solvable, e.g. there is a* $b \in \mathbb{Z}/p\mathbb{Z}$ *with* $ab \equiv 1 \mod p$. *We denote this field* $\mathbb{F}_p$.

Many other rings are not fields, such as $\mathbb{Z}$, $\mathbb{Z}[i]$, and $\mathbb{R}[x]$; however, they do have the nice property of cancellation:

**Definition 3.4.2: Cancellation Property**

Let $R$ be a commutative ring. $R$ has the **cancellation property** if for every $a, b, c \in R$,

$$ab = ac \text{ with } a \neq 0 \iff b = c.$$

Rings that maintain the cancellation property are called **integral domains**:

**Definition 3.4.3: Zero Divisors and Integral Domains**

Let $R$ be a ring. An element $a \in R$ is a **zero divisor** if $a \neq 0$ and there is some non-zero $b \in R$ such that $ab = 0$. The ring $R$ is an **integral domain** if it has no zero divisors. Equivalently, the ring $R$ is an integral domain if the only way to get $ab = 0$ is if either $a = 0$ or $b = 0$.

In fact, every field is an integral domain, and a ring $R$ is an integral domain if and only if it has the cancellation property. Moreover, every integral domain is a subring of a field (the reader should verify all of these statements). We will see later that the smallest such field is called the **field of fractions over** $R$.

## §3.5 Unit Groups and Product Rings

In groups, we saw that many interesting subgroups and larger groups could be formed from a group. Similarly, we get the notion in rings that every ring contains an interesting group, and smaller rings can form larger rings.

### §3.5.1 Unit Groups

**Definition 3.5.1: Unit Groups**

Let $R$ be a commutative ring[a]. The **group of units of** $R$ is the subset $R^*$ of $R$ defined by

$$R^* = \{a \in R \mid \text{ there exists some } b \in R \text{ satisfying } ab = 1\},$$

where group law is ring multiplication. Elements of $R^*$ are called **units**.

[a]For a non-commutative ring $R$, $a \in R$ is a **unit** if there are elements $b, c \in R$ such that $ab = ca = 1$, i.e. the element $a$ needs both a left- and right-inverse.

**Proposition 3.5.1**

The set of units $R^*$ of $R$ is a group, with group law being ring multiplication.

**Proof.** For each of the group axioms:

- Let $a_1, a_2 \in R^*$, and let $b_1, b_2 \in R^*$ be values such that $a_1 b_1 = 1$, $a_2 b_2 = 1$ (we can say $b_1, b_2 \in R^*$, not just $R$, since commutativity ensures that if $a \in R^*$, $b \in R^*$ as well). Then

$$\begin{aligned} 1 &= a_1 b_1 a_2 b_2 \\ &= a_1 a_2 b_1 b_2, \end{aligned}$$

  and since $b_1 b_2 \in R^*$ (multiplication is closed in monoids), we have $a_1 a_2 \in R^*$.
- $1 \in R$ is the identity element.
- By definition, units have inverses.
- Multiplicative associativity is guaranteed by the ring axioms.

Hence $R^*$ is a group. □

**Example 23.** *Some unit groups include*

$$\mathbb{Z}^* = \{\pm 1\}, \qquad \mathbb{Z}[i]^* = \{\pm 1, \pm i\}, \qquad \mathbb{R}[x]^* = \mathbb{R}^*.$$

*Another interesting example is the ring $\mathbb{Z}[\sqrt{2}]$, whose unit group has infinitely many elements. The proofs of these assertions is left as an exercise for the reader.*

**Example 24.** *A ring $R$ is a field if and only if*

$$R^* = \{a \in R \mid a \neq 0\} = R \setminus \{0\};$$

*in other words, every non-zero element has a multiplicative inverse.*

We now explore the unit group of the ring $\mathbb{Z}/m\mathbb{Z}$.

**Proposition 3.5.2**

Let $m \geq 1$ be an integer. Then

$$(\mathbb{Z}/m\mathbb{Z})^* = \{a \mod m \mid \gcd(a, m) = 1\}.$$

In particular if $m$ is a prime number, then $\mathbb{Z}/p\mathbb{Z}$ is a field (often denoted $\mathbb{F}_p$).

**Proof.** Suppose that $\gcd(a, m) = 1$. Then by the Euclidean algorithm, we can find $u, v \in \mathbb{Z}$ satisfying $au + mv = 1$. Hence

$$au = 1 - mv \equiv 1 \mod m.$$

Thus $u$ is a multiplicative inverse for $a$ in the ring $\mathbb{Z}/m\mathbb{Z}$, so $a \mod m$ is in $(\mathbb{Z}/m\mathbb{Z})^*$. In the other direction, suppose that $a \mod m \in (\mathbb{Z}/m\mathbb{Z})^*$. Then for any $a \in \mathbb{Z}/m\mathbb{Z}$, we can find some $b \mod m \in \mathbb{Z}/m\mathbb{Z}$ such that

$$(a \mod m)(b \mod m) = 1 \mod m.$$

In other words, $ab \equiv 1 \mod (m)$, so $ab - 1 = cm$ for some $c$. But then $ab - cm = 1$, and so $\gcd(a, m) = 1$, since any number dividing both $a, m$ divides 1, which is only true for 1. □

### §3.5.2 Product Rings

Now, we inspect building larger rings from smaller rings. Why make things more complicated? It turns out that reversing this process, breaking up complicated rings into smaller, easier rings, can be useful; one such example is the Chinese Remainder Theorem. The building procedure is analogous to building products of groups, as well as constructing vector spaces (e.g. $\mathbb{R}^n$, making $n$-tuples from $\mathbb{R}$).

**Definition 3.5.2: Products of Rings**

Let $R_1, \ldots, R_n$ be rings. The **product of** $R_1, \ldots, R_n$ is the ring

$$R_1 \times \ldots \times R_n = \{(a_1, \ldots, a_n) \mid a_1 \in R_1, \ldots, a_n \in R_n\}.$$

In other words, the product $R_1 \times \ldots \times R_n$ is the set of $n$-tuples, where the first entry is chosen from $R_1$, second from $R_2$, etc. $R_1 \times \ldots \times R_n$ becomes a ring using coordinate-wise addition and multiplication:

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$
$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 \cdot b_1, \ldots, a_n \cdot b_n).$$

Proving $R_1 \times \ldots \times R_n$ is a group is left as an exercise for the reader.

**Example 25.** *The product ring* $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ *has* 6 *elements,*

$$(0,0),\ (1,0),\ (0,1),\ (0,2),\ (1,1),\ (1,2).$$

*For example, addition and multiplication look like*

$$(1,1) + (1,2) = (0,0),\ (0,2) \cdot (1,2) = (0,1).$$

*It turns out that* $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}/6\mathbb{Z}$.

*However,* $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ *is* ***not*** *isomorphic to* $\mathbb{Z}/8\mathbb{Z}$. *To see this, if*

$$\phi : \mathbb{Z}/8\mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$$

*is a homomorphism, then by definition* $\phi(1) = (1,1)$*, so*

$$\phi(4) = \phi(1+1+1+1) = \phi(1)+\phi(1)+\phi(1)+\phi(1) = (1,1)+(1,1)+(1,1)+(1,1) = (0,0).$$

*Hence* $\ker(\phi)$ *is non-trivial, so* $\phi$ *cannot be injective.*

Now, we combine product rings and unit groups.

**Proposition 3.5.3**

Let $R_1, \ldots, R_n$ be rings. Then the unit groups of the product is isomorphic to the product of the unit groups:

$$(R_1 \times \ldots \times R_n)^* \cong R_1^* \times \ldots \times R_n^*.$$

**Proof.** If $(a_1, \ldots, a_n) \in (R_1 \times \ldots \times R_n)^*$, then by definition there is a $(b_1, \ldots, b_n) \in (R_1 \times \ldots \times R_n)^*$ satisfying

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (1, \ldots, 1).$$

However, this means that $a_i b_i = 1$, so $a_i \in R_i^*$. Hence $(a_1, \ldots, a_n) \in R_1^* \times \ldots \times R_n^*$.
Now, suppose $(a_1, \ldots, a_n) \in R_1^* \times \ldots \times R_n^*$. Then for $a_i \in R_i^*$, there exists some $b_i \in R_i^*$ such that $a_i b_i = 1$. Then

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (1, \ldots, 1).$$

Hence $(a_1, \ldots, a_n) \in (R_1 \times \ldots \times R_n)^*$. □

## §3.6 Ideals and Quotient Rings

Recall that in the ring $\mathbb{Z}/m\mathbb{Z}$, we pretend that $a, b \in \mathbb{Z}$ are "identical" if $a - b = km$ for some $k \in \mathbb{Z}$. In other words, we get an equivalence relation

$$a \equiv b \mod m \text{ if } ab = km \ (a - b \text{ is a multiple of } m).$$

We then defined $\mathbb{Z}/m\mathbb{Z}$ to be the set of equivalence classes.
Now, we attempt to generalize this construction to arbitrary (commutative) rings. We first start by generalizing the concept "being a multiple of $m$". Why would we want this?

- First, it provides us with a new ring to explore; when is $\mathbb{Z}/m\mathbb{Z}$ a field? An integral domain? Moreover, when exploring conjectures, $\mathbb{Z}/m\mathbb{Z}$ provides a testing ground.
- Second, sometimes $\mathbb{Z}/m\mathbb{Z}$ is easier to work with than $\mathbb{Z}$. Consider the Diophantine equation

  $$x^n + y^n = z^n.$$

  Solving this can be notoriously difficult (PepeLaugh). However, since $\mathbb{Z}/m\mathbb{Z}$ is finite, we can use the natural ring map $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z}$ for a variety of $m$'s to learn about the problem.

**Definition 3.6.1: Ideals**

Let $R$ be a commutative ring. An **ideal** of $R$ is a non-empty subset $I \subseteq R$ with the following two properties:

- If $a \in I$ and $b \in I$, then $a + b \in I$.
- If $a \in I$ and $r \in R$, then $ra \in I$.

Some examples of ideals include $m\mathbb{Z} \subseteq \mathbb{Z}$ (for any $m \in \mathbb{Z}$), $\{0\} \subseteq R$ for any ring $R$, and $R$ itself.
One way to create an ideal of $R$ is to start with one element in $R$ and take all of its multiples.

**Definition 3.6.2: Principal Ideals**

Let $R$ be a commutative ring, and let $c \in R$. The **principal ideal generated by** $c$, denoted $cR$ or $(c)$, is the set of all multiples of $c$,

$$cR = (c) = \{rc \mid r \in R\}.$$

Verifying that $cR$ is an ideal is straightforward, and left as an exercise.

The above examples still hold: $m\mathbb{Z}$, $\{0\}$, and $R$ are all principal ideals.
In some rings (e.g. $\mathbb{Z}$, $\mathbb{Z}[i]$, $\mathbb{R}[x]$), every ideal is a principal ideal (such rings are called **principal ideal domains**), although this is not immediately obvious. Moreover, this is not true for rings like $\mathbb{Z}[i]$; there exist non-principal ideals.

**Example 26.** *Every ring has at least two ideals:*

- *the* ***zero ideal***

  $$(0) = 0R = \{0\}$$

  *consisting of just the zero element, and the* ***unit ideal***

  $$(1) = 1R = R$$

  *consisting of the entire ring.*

*More generally, if $u \in R$ is a unit, then $uR = (u) = R$.*

Just like products of rings, we can formulate a sense of building multiple ideals. A principal ideal is generated by a single element in R; for a finite list of elements $c_1, \ldots, c_n \in R$, the **ideal generated by** $c_1, \ldots, c_n$ is

$$(c_1, \ldots, c_n) = c_1R + \ldots + c_nR = \{r_1c_1 + \ldots + r_nc_n \mid r_1, \ldots, r_n \in R\}.$$

Verifying that this is an ideal is left as an exercise.

**Remark 5.** *(Proof Technique) Let $I$ be an ideal of $R$. If $1 \in I$, then for every $r \in R$ we have $r \cdot 1 = r \in I$, so $I = R$ is the unit ideal. Conversely, if you can construct an ideal $I$ and prove $I = R$, we can often exploit this fact by using $1 \in I$.*

Now, we can start crafting quotient rings $R/I$ by identifying pairs of $R$ if their difference is in $I$, just like we did when defining $\mathbb{Z}/m\mathbb{Z}$:

For an element $a \in R$, the set of $b \in R$ that are equivalent to $a$ consists of the set of $b$ such that $a - b \in I$, or equivalently, such that $b$ is in the set $a + I$.

This parallel to the ring of integers modulo $m$ prompts the following definitions.

**Definition 3.6.3: Cosets of Ideals**

Let $R$ be a commutative ring, and let $I$ be an ideal of $R$. Then for every element $a \in R$, the **coset of** $a$ is the set

$$a + I = \{a + c \mid c \in I\}.$$

Note that $a$ is always an element of its coset, since $0 \in I$. If $a, b \in R$ satisfy $b - a \in I$, then people often write

$$b \equiv a \mod I$$

and say "$b$ is congruent to $a$ modulo $I$."
Given two cosets $a + I$ and $b + I$, we define their sum and product by the formulas

$$(a + I) + (b + I) = (a + b) + I, \qquad (a + I) \cdot (b + I) = (a \cdot b) + I,$$

and we denote the collection of distinct cosets by $R/I$.

Like $\mathbb{Z}/m\mathbb{Z}$, we wish to turn $R/I$ into a ring; it turns out that the above definitions successfully define a commutative ring.

**Proposition 3.6.1**

Let $R$ be a commutative ring, and let $I$ be an ideal of $R$.

1. Let $a + I, a' + I$ be two cosets. Then $a + I = a' + I$ if and only if $a' - a \in I$.
2. Addition and multiplication of cosets is well-defined, in that it doesn't matter which element of the coset we use in the definition.
3. Addition and multiplication of cosets in $R/I$ turn $R/I$ into a commutative ring.[a]

[a] To be precise, we must require that $I \neq R$, since if $I = R$, then $R/I$ only has one element, and we don't allow rings to have $1 = 0$.

**Proof.** 1. Suppose $a + I = a' + I$. Then $a' = a + c$ for some $c \in I$. But then $a' - a = c \in I$, and so $a' - a \in I$.

Now, suppose $a' - a \in I$. Then $a' - a = c \in I$ for some $c \in I$, so $a' = a + c$, $a = a' - c$. Thus any $\alpha \in a' + I$ can be written as $a' + b = a + (b + c) \in a + I$, and any $\beta \in a + I$

can be written as $\beta = a+b = a'+(b-c) \in a'+I$. Hence $a+I \subseteq a'+I$, $a'+I \subseteq a+I$, so $a+I = a'+I$.

2. For addition: let $a, b, a', b' \in R$ be elements that satisfy $a \equiv a' \mod I$, $b \equiv b' \mod I$; that is, $a'-a \in I$, $b'-b \in I$. By (1), $a+I = a'+I$, $b+I = b'+I$. Then

$$a' = a+s,\ b' = b+t$$

for some $s, t \in I$. But then $a'+b' = a+s+b+t = (a+b)+(s+t)$; and since $s+t \in I$, we have $a'+b' \equiv a+b \mod I$. Alternatively, $(a'+b')-(a+b) \in I$, so by (1) $(a+b)+I = (a'+b')+I$.

For multiplication: let $a, b, a', b' \in R$ be elements whose cosets satisfy

$$a'+I = a+I \text{ and } b'+I = b+I.$$

The assumption that $a+I = a'+I$ means that there is some $c \in I$ such that $a' = a+c$, and similarly the assumption that $b+I = b'+I$ means that there is some $d \in I$ such that $b' = b+d$ (since $a' \in a+I$ means that $a'$ is of the form $a+c$ for some $c \in I$). It follows that

$$a'b' = (a+c)(b+d) = ab + \underbrace{ad+cb+cd}_{\text{This is in } I, \text{ since } c,d \in I.}.$$

Since $c, d \in I$, $ad+cb+cd$ is also in $I$, and so $a'b'-ab = ad+cb+cd \in I$; and from (1), we see that $ab+I = a'b'+I$ are equal.

□

**Remark 6.** *One way to view the quotient ring $R/I$ is to divide a "pie" of $R$ into its individual "slices", where each slice represents an element of $R/I$ (for instance, $\mathbb{Z}/3\mathbb{Z}$ is divided into three slices: $3\mathbb{Z}$, $1+3\mathbb{Z}$, and $2+3\mathbb{Z}$). Alternatively, one can view $a$ and $b$ as "indistinguishable" if $a-b \in I$. As we think of $1 \sim 4$ in $\mathbb{Z}/3\mathbb{Z}$, we can abstractly partition $\mathbb{Z}/3\mathbb{Z}$ into colors; red might symbolize $3\mathbb{Z}$, green for $1+3\mathbb{Z}$, and blue for $2+3\mathbb{Z}$. Thus, $1, 4, 7$ would all be red, "the same element".*

Ideals and homomorphisms are closely related. Review the following proposition carefully; it will be used constantly.

**Proposition 3.6.2**

1. Let $R$ be a commutative ring. Then the map

$$\pi : R \longrightarrow R/I,\ a \longmapsto a+I$$

that sends an element to its coset is a surjective ring homomorphism with kernel $I$.

2. Let $\phi : R \to R'$ be a ring homomorphism.
   - The kernel of $\phi$ is an ideal of $R$ (recall that the kernel is $\ker(\phi) = \{a \in R \mid \phi(a) = 0_{R'}\}$).
   - The homomorphism is injective if and only if $\ker(\phi) = (0)$.

- Let $I_\phi = \ker(\phi)$. There is a well-defined injective ring homomorphism

$$\overline{\phi} : R/I_\phi \longrightarrow R,\ \overline{\phi}(a + I_\phi) = \phi(a).$$

Moreover, if $\phi$ is surjective, then $\overline{\phi}$ is an isomorphism.

**Proof.** To prove homomorphism:

- $\pi(0_R) = 0 + I = I = 0_{R/I}$.
- $\pi(a + b) = (a + b) + I = (a + I) + (b + I) = \pi(a) + \pi(b)$, by definition of addition of cosets of ideals.
- $\pi(a \cdot b) = (a \cdot b) + I = (a + I) \cdot (b + I) = \pi(a) \cdot \pi(b)$, by definition of multiplication of cosets of ideals.

Hence $\pi : R \to R/I$ is a ring homomorphism. Next, recall that $\ker(\pi) = \{a \in R \mid \pi(a) = 0_{R/I}\}$. But $0_{R/I} = 0 + I$, so we need $\pi(a) = I = 0 + I$. For any $a + I$, $a + I = 0 + I$ iff $a - 0 = a \in I$; thus $\ker(\pi) = I$. Surjectivity is trivial: for any $a + I \in R/I$, $a \in R$, so $\phi(a) = a + I$.

Now, let $\phi : R \to R'$ be a ring homomorphism. The kernel of $\phi$ is the set $\{a \in R \mid \phi(a) = 0_{R'}\}$. For any $a,\ b \in \ker(\phi)$, we have $\phi(a + b) = \phi(a) + \phi(b) = 0 + 0$, and so $a + b \in \ker(\phi)$. Next, let $c \in R$. For any $c \cdot a$, we have $\phi(c \cdot a) = \phi(c) \cdot \phi(a) = \phi(c) \cdot 0 = 0$, and so $c \cdot a \in \ker(\phi)$. Thus $\ker(\phi)$ forms an ideal of $R$.
To prove injectivity iff $\ker(\phi) = (0)$:

- Suppose $\phi$ is injective. Then for any $a \in \ker(\phi)$, $\phi(a) = 0 = \phi(0)$, so $a = 0$.
- Conversely, suppose $\ker(\phi) = (0)$. Suppose $\phi(a) = \phi(a')$. Then $\phi(a) - \phi(a') = \phi(a - a')$, so $a - a' \in \ker(\phi)$; but $\ker(\phi) = (0)$, so $a - a' = 0 \implies a = a'$. Hence $\phi$ is injective.

We finally prove well-definedness of $\overline{\phi}$. Suppose that $a + I_\phi = a' + I_\phi$ are two ways of writing a coset. We need to show $\overline{\phi}(a) = \overline{\phi}(a')$. The assumption of $a + I_\phi = a' + I_\phi$ allows us to write $a' = a + b$ for some $b \in I_\phi$. Thus

$$\overline{\phi}(a' + I) = \phi(a') = \phi(a + b) = \phi(a) + \phi(b) = \phi(a) + 0 = \phi(a),$$

and so $\overline{\phi}$ is well-defined. $\overline{\phi}$ being a ring homomorphism follows directly from $\phi$ being a ring homomorphism. To show injectivity,

$$\overline{\phi}(a+I_\phi) = \overline{\phi}(b+I) \iff \phi(a) = \phi(b) \iff \phi(a-b) = 0 \iff a-b \in I_\phi \iff a+I_\phi = b+I_\phi,$$

as desired (the second last step is accomplished because $\phi(a - b) = 0$ implies $a - b \in \ker(\phi) = I_\phi$).
Finally, if $\phi$ is surjective, then so is $\overline{\phi}$: $\phi$ surjective implies any $a' \in R'$ can be written as $\phi(a)$; but then any $\phi(a)$ can be written as $\overline{\phi}(a + I_\phi)$; thus $\phi(a) = \overline{\phi}(\pi(a)) = \overline{\phi}(a + I_\phi)$, and so $\overline{\phi}$ is surjective. We don't have to worry about $\pi$ "eating up" any $a \in R$: if $a \equiv b \mod I_\phi$, then $a = b + c$ for some $c \in I_\phi$, so

$$\phi(a) = \phi(b + c) = \phi(b) + \phi(c) = \phi(b) + 0 = \phi(b),$$

and so $\phi(a) = \phi(b)$. Hence any $\pi(a) = \pi(b)$ won't affect the surjectivity of $\overline{\phi}$.
Therefore $\overline{\phi}$ is an isomorphism. □

**Definition 3.6.4: Characteristics of Rings**

Let $R$ be a ring, and let

$$\phi : \mathbb{Z} \longrightarrow R$$

be the unique homomorphism determined by $\phi(1) = 1_R$ (see Example 21). The kernel of $\phi$ is an ideal of $\mathbb{Z}$, and recall that every ideal of $\mathbb{Z}$ is principal (therefore every ideal $I$ of $\mathbb{Z}$ is of the form $m\mathbb{Z}$); there is thus a unique integer $m \geq 0$ such that

$$\ker(\phi) = m\mathbb{Z}.$$

That integer $m$ is called the **characteristic of the ring** $R$.

Another way to describe $m$ is that $m$ is the smallest integer such that

$$\phi(m) = \underbrace{1_R + \ldots + 1_R}_{m \text{ terms}} = 0_R;$$

or, if no such $m$ exists, then $m = 0$ (the characteristic of $R$ is 0).

**Example 27.** *The ring $\mathbb{Z}/m\mathbb{Z}$ has characteristic $m$, while the rings $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ all have characteristic 0. It follows from Proposition 3.31 that if a ring has characteristic $m$, then there is an injective ring homomorphism*

$$\mathbb{Z}/m\mathbb{Z} \hookrightarrow R$$

*(this follows since characteristic $m$ implies $\ker(\phi) = m\mathbb{Z}$ for some $\phi : \mathbb{Z} \to R$, and every map $\mathbb{Z}/I_\phi = \mathbb{Z}/m\mathbb{Z} \to R$ is injective).*

Note that when $m = 0$, we aren't "dividing by 0." In general, the quotient of a ring $R$ by the zero ideal $(0)$ is just $R$ : $R/(0) = R$, since $R/(0)$ means we identify elements $a, b \in R$ as congruent if $a - b = 0$; or, $a = b$. Thus $R$ has infinitely many cosets $a + (0)$, and so $R/(0) = R$.

Now, we move on to the F̃reshman's Dream. Many freshmen hope for $(a+b)^n = a^n + b^n$. Clearly, this isn't true in $\mathbb{R}$; but is there a ring in which this is true? It turns out that yes, any ring with prime characteristic possesses that possibility!

**Theorem 3.6.1: Freshman's Dream**

Let $p$ be a prime, and let $R$ be a commutative ring with characteristic $p$. Then the map

$$f : R \longrightarrow R, \ f(a) = a^p$$

is a ring homomorphism, called the **Frobenius homomorphism of** $R$. In particular, for all $a, b \in R$ and $n \geq 0$, we have

$$(a + b)^{p^n} = a^{p^n} + b^{p^n}.$$

**Proof.** Clearly $f(1) = 1$, and preserving multiplication is clear:

$$f(ab) = (ab)^p = (a^p)(b^p) = f(a)f(b).$$

For addition, we use the Binomial Theorem:

$$\begin{aligned} f(a+b) &= (a+b)^p \\ &= \sum_{k=0}^{p} \binom{p}{k} a^{p-k} b^k \\ &= a^p + b^p + \sum_{k=1}^{p-1} \binom{p}{k} a^{p-k} b^k. \end{aligned}$$

For any $1 \leq k \leq p-1$, $\binom{p}{k}$ is a multiple of $p$:

$$\binom{p}{k} = \frac{p!}{k!(p-k)!} = \frac{(p-k+1)\dots(p-1)(p)}{k!}.$$

Moreover, since $p$ is prime and $1 \leq k \leq p-1$, none of $k!$ will cancel with $p$ (see Proposition 1.47 in the book). Since we're working in a ring with characteristic $p$, that means all the $\binom{p}{k}$ will be 0. Hence the entire sum disappears, and we are left with

$$f(a+b) = a^n + b^n = f(a) + f(b).$$

Thus $f$ is a ring homomorphism.
To prove the freshman's dream, we use induction: the case $n = 1$ is already proven, so suppose $n = k$ is true. Then

$$(a+b)^{p^{k+1}} = ((a+b)^{p^k})^p = (a^{p^k} + b^{p^k})^p = (a^{p^k})^p + (b^{p^k})^p = a^{p^{k+1}} + b^{p^{k+1}}.$$

Thus the freshman dream holds for any $(a+b)^{p^n}$. □

## §3.7 Prime Ideals and Maximal Ideals

Primes hold great significance in number theory. Recall that an integer $p$ is prime if and only if its only (positive) divisors are 1 and $p$. Moreover, if $p$ divides any $ab$, then either $p$ divides $a$ or $p$ divides $b$. We can rephrase this divisibility property for any arbitrary ideal: if a product $ab$ is in the ideal $p\mathbb{Z}$, then either $a \in p\mathbb{Z}$ or $b \in p\mathbb{Z}$. Additionally, we note that integral domains and fields are of particular interest; thus, we wish to find for which ideals $I$ is the quotient ring $R/I$ an integral domain, and for which is it a field.

**Definition 3.7.1: Prime Ideals**

Let $R$ be a commutative ring. An ideal $I$ of $R$ is a **prime ideal** if $I \neq R$ and if, whenever $ab \in I$, then either $a \in I$ or $b \in I$.

The contrapositive is also important: if $I$ is a prime ideal, and both $a \notin I$ and $b \notin I$, then we have $ab \notin I$.

**Example 28.** *Let $m \neq 0$. The ideal $m\mathbb{Z}$ is a prime ideal if and only if $|m|$ is a prime number.*

**Example 29.** *Let $F$ be a field. For every $a, b \in F$ with $a \neq 0$, the principal ideal $(a+bx)F[x]$ is a prime ideal. For every $a, b, c \in F$ with $a \neq 0$ and $b^2 - ac \neq d^2$ for any $d \in F$ ($b^2 - ac$ is not the square of any element), the principal ideal $(a^2 + bx + c)F[x]$ is a prime ideal.*

The largest possible ideal of $R$ that is not $R$ itself is also of interest.

**Definition 3.7.2: Maximal Ideals**

Let $R$ be a commutative ring. An ideal $I$ is a **maximal ideal** if $I \neq R$ and there is no ideal "larger" than $I$, or "contained" between $I$ and $R$; that is, for any ideal $J$, if $I \subseteq J \subseteq R$, then either $J = I$ or $J = R$.

**Example 30.** *Let $p \in \mathbb{Z}$ be a prime number. Then the ideal $p\mathbb{Z}$ is both a prime ideal **and** a maximal ideal. This follows by combining Proposition 3.5.2 (3.17 in the textbook), which says that $\mathbb{Z}/p\mathbb{Z}$ is a field, and the following Theorem 3.7.1 (3.40 in the textbook), which says that in general, $R/I$ is a field if and only if $I$ is a maximal ideal.*

**Example 31.** *In the ring $\mathbb{Z}[x]$ of polynomials with integer coefficients, the principal ideals $2\mathbb{Z}[x]$ and $x\mathbb{Z}[x]$ are prime ideals, but not maximal ideals, since they are properly contained in the ideal*

$$\{2a(x) + xb(x) \mid a(x),\ b(x) \in \mathbb{Z}[x]\}$$

*generated by 2 and $x$, and one can check that this ideal is not all of $\mathbb{Z}[x]$. Indeed, it is a non-principal maximal ideal.*

Just as prime numbers in $\mathbb{Z}$ form the basic building blocks for all numbers (recall that any integer can be represented uniquely as a product of prime numbers), the prime and maximal ideals of $R$ are, in some sense, the basic building blocks underlying the algebraic (and geometric) structure of $R$.

On the other hand, integral domains and fields are two particularly nice examples of rings. Thus, the next theorem is both interesting and important.

**Theorem 3.7.1**

Let $R$ be a commutative ring, and let $I$ be an ideal with $I \neq R$.

1. $I$ is a prime ideal if and only if the quotient ring $R/I$ is an integral domain.
2. $I$ is a maximal ideal if and only if the quotient ring $R/I$ is a field.

**Proof.** We first deal with the first statement:

- Suppose $I$ is a prime ideal. Let $a+I$ and $b+I$ be elements of $R/I$. If $(a+I)(b+I) = 0+I = I$, then either $a+I = I$ or $b+I = I$. Thus at least one of $a+I$ and $b+I$ is equal to $0+I$, and so $R/I$ is an integral domain (since for any $a, b \in R/I$, if $a \cdot b = (a+I)(b+I) = 0_{R/I} = 0+I$, then one of $a+I,\ b+I$ is $0_{R/I}$).
- Conversely, suppose $R/I$ is an integral domain. Then for any $a+I,\ b+I \in R/I$, if $(a+I)(b+I) = 0+I = I$, then either $a+I = I$ or $b+I = I$. This implies either $a \in I$ or $b \in I$ (since $a+I = 0+I$ implies $a-0 = a \in I$).

Now, we move on to the second:

- Suppose $I$ is a maximal ideal. Let $a+I$ be a non-zero element of $R/I$ (e.g. $a+I \neq I$). To exploit the fact that $I$ is a maximal ideal, construct an ideal $J$ where
$$J = \{ar + b \mid r \in R,\ b \in I\}$$
($J$ being an ideal is left as an exercise). Taking elements of $J$ with $r = 0$ shows that $I \subset J$, while taking $r = 1,\ b = 0$ shows that $a \in J$. Since $a \notin I$, $J$ is strictly larger than $I$; and since $I$ is a maximal ideal and $I \subsetneq J \subseteq R$, we have that $J = R$. In particular, $1 \in J$. Thus there exists some $c \in R,\ b \in I$ such that $1 = ac + b$. In terms of $R/I$, and using the fact that $b \in I$ implies $b+I = I$, we have
$$1 + I = (ac+b) + I = (ac+0) + I = ac + I = (a+I)(c+I).$$
Hence $a+I$ has a multiplicative inverse in $R/I$, and by definition $R$ is commutative, so $R/I$ is a field.
- Conversely, suppose $R/I$ is a field. Let $J$ be an ideal such that $I \subseteq J \subseteq R$. If $J = I$, we are done, so suppose $J \neq I$. Then for some $a \in J$, we have $a \notin I$. Then the coset $a+I \neq I = 0+I$, so $a+I$ is a non-zero element of $R/I$. Since $R/I$ is a field, $a+I$ has a multiplicative inverse $c+I$ for some $c \in R$; that is,
$$1 + I = (a+I) \cdot (c+I) = ac + I,$$
so there is an element $b \in I$ such that $1 = ac + b$. But $J$ is an ideal, so $a \in J$ implies $ac \in J$. Additionally, since $b \in I \subset J$, we have $ac + b = 1 \in J$. But since $1 \in J$, any $r \in R$ is in $J$: $r \cdot 1 = r \in J$. Hence $J = R$, and so $I$ is a maximal ideal.

□

This theorem leads to a very slick corollary:

**Corollary 3.7.1**

Every maximal ideal is a prime ideal.

**Proof.** If an ideal $I$ is a maximal ideal, then $R/I$ is a field. But $R/I$ is a field implies that $R/I$ is an integral domain, and thus $I$ is a prime ideal as well. □

Note that the converse does not always hold; Example 31 (3.39 in the textbook) illustrates prime non-maximal ideals.

**Remark 7.** *It would be nice to say that any ring has at least one maximal ideal. It turns out that this is yet another statement equivalent to the Axiom of Choice!*

Chapter 4

# Vector Spaces: Part 1

## §4.1 Introduction to Vector Spaces

We studied vectors on a real plane (e.g. 2D or 3D space) in the guise of arrows; for instance, $(1, 2)$ denotes an arrow with tail at the origin and head at the coordinate $(1, 2)$. Additionally, we could add vectors "tip-to-tail, " by concatenating one vector's tail to another's head; and we could multiply them by a scalar (e.g. $-2(1, 2)$ becomes $(-2, -4)$). These operations are represented as follows:

- $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$
- $c(a, b) = (ca, cb)$

Seem familiar? Abstractly, vectors are two types of operations: *scalar multiplication* and *vector addition*; and we can combine them using the all-powerful distributive law:

$$c(\vec{v}_1 + \vec{v}_2) = c\vec{v}_1 + c\vec{v}_2.$$

Now, we formalize this concept into a **vector space**.

## §4.2 Vector Spaces and Linear Transformations

The numbers we used before ($(1, 2)$) are real numbers; but that's not necessary. Generally, we can use any sort of "numbers" that allow us to add, subtract, multiply, and divide. As discussed before, these numbers live in **fields**.

**Definition 4.2.1: Fields**

A **field** $F$ is a commutative ring with the property that every non-zero $a \in F$ has a multiplicative inverse. In other words, for every $a \in F$ with $a \neq 0$, there is a $b \in F$ such that $ab = 1$; or, $F^* = F \setminus \{0\}$.

Familiar fields include $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$; additionally, for every prime $p$, there is a field $\mathbb{Z}/p\mathbb{Z}$ (see Proposition 3.17).

For vector spaces, we fix an underlying field $F$, and use it as a basic building block for vector spaces. Subsequently, we will use vector spaces to deeply study fields and field extensions.

**Definition 4.2.2: Vector Spaces**

Let $F$ be a field. A **vector space with a field F**, denoted $V_F$, or alternatively a **$F$-vector space**, is an Abelian group $V$ (with group law "vector addition"), together with a rule for multiplying a vector $\vec{v} \in V$ by a scalar $c \in F$ to obtain a new vector $c\vec{v} \in V$. Vector addition and scalar multiplication are required to

satisfy these axioms (in addition to the Abelian group axioms under addition and multiplication):

1. **Identity Law**: $1\vec{v} = \vec{v}$ for all $\vec{v} \in V$.
2. **Distributive Law 1**: $c(\vec{v}_1 + \vec{v}_2) = c\vec{v}_1 + c\vec{v}_2$ for all $c \in F$, $\vec{v}_1,\ \vec{v}_2 \in V$.
3. **Distributive Law 2**: $(c_1 + c_2)\vec{v} = c_1\vec{v} + c_2\vec{v}$ for all $c_1, c_2 \in F$, $\vec{v} \in V$.
4. **Associative Law**: $(c_1c_2)\vec{v} = c_1(c_2\vec{v})$ for all $c_1, c_2 \in F$, $\vec{v} \in V$.

The zero element of $V$ is the **zero vector**, denoted $\mathbf{0}$. It is **not** $0 \in F$, the zero element of $F$.

Just as with the axiomatic definition of groups and rings, some basic facts about vector spaces can be proven directly from their definitions.

**Proposition 4.2.1**

Let $V_F$ be a vector space.

- $0v = \mathbf{0}$ for all $v \in V$.
- $(-1)v = -v$ for all $v \in V$.
- Every vector $v \in V$ has a unique inverse $-v$.

**Proof.**

- $0v = (0 + 0)v = 0v + 0v$; adding $-0v$ to both sides yields $\mathbf{0} = 0v$.
- $v + (-1)v = 1v + (-1)v = (1 + -1)v = 0v = \mathbf{0}$. Hence $(-1)v = -v$.
- Suppose $v_1, v_2 \in V$ are two inverses of $v \in V$. Then

$$v_1 = v_1 + \mathbf{0} = v_1 + (v + v_2) = (v_1 + v) + v_2 = \mathbf{0} + v_2 = v_2.$$

Hence $v_1 = v_2$.

□

Vector spaces are characterized by vector addition and scalar multiplication, so we are interested in maps that preserve these properties:

**Definition 4.2.3: Linear Transformations**

Let $F$ be a field, and let $V, W$ be $F$-vector spaces. A **linear transformation** $L$ from $V$ to $W$ is a function

$$L : V \longrightarrow W$$

satisfying

$$L(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1L(\vec{v}_1) + c_2L(\vec{v}_2)$$

for all $c_1, c_2 \in F$, $\vec{v}_1, \vec{v}_2 \in V$. To specify a field, we call $L$ an **$F$-linear transformation**.

**Remark 8.** *An alternative, and perhaps better, name for a linear transformation would be a* ***vector space homomorphism****, given its similarities to group homomorphisms and ring homomorphisms. However, due to historical reasons linear transformations has stuck.*

## §4.3 Interesting Examples of Vector Spaces

We've already seen vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ (our world!... maybe?). More generally, we can form a vector space using a set of $n$-tuples with coordinates in any field.

**Example 32.** *Let $F$ be a field, with $n \geq 1$ an integer. Then $F^n$ is the $F$-vector space whose vectors are $n$-tuples of elements of $F$,*

$$F^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in F\}.$$

*Vector addition and scalar multiplication are done coordinate-wise. Among examples of $F^n$ vector spaces include $\mathbb{R}^n$, $\mathbb{C}^n$, and $\mathbb{F}_p^n$. Notice that $\mathbb{F}_p^n$ is finite; it has exactly $p^n$ different vectors.*

**Example 33.** *The maps*

$$L(a_1, a_2) = (3a_1 - 5a_2, 2a_1 + 3a_2)$$
$$L'(a_1, a_2, a_3) = (3a_1 - 5a_2 + 2a_3, 2a_1 + 3a_2 - 7a_3)$$

*are examples of linear transformations (as one should verify!).*

**Example 34.** *The set of polynomials $F[x]$ over a field $F$ is a vector space, with polynomial addition and scalar multiplication the usual way. For any $a \in F$, the* ***evaluation map***

$$E_a : F[x] \longrightarrow F,\ E_a(f(x)) = f(a),$$

*is a linear transformation. More generally, for any list of values $a_1, \ldots, a_n \in F$, we can define a linear transformation*

$$E_a : F[x] \longrightarrow F^n,\ E_a(f(x)) = (f(a_1), \ldots, f(a_n)).$$

*One should verify that this is a linear transformation.*

**Example 35.** *Linear algebra can be connected with calculus! Let*

$$V = \{\textit{functions } f : \mathbb{R} \to R\}$$
$$V^{cont} = \{\textit{continuous functions } f : \mathbb{R} \to \mathbb{R}\}$$
$$V^{diff} = \{\textit{differentiable functions } f : \mathbb{R} \to \mathbb{R}\}.$$

*These are $\mathbb{R}$-vector spaces, with function addition and scalar multiplication as usual:*

$$(f+g)(x) = f(x) + g(x),\ (cf)(x) = cf(x).$$

*Differentiation is a linear transformation*

$$D : V^{diff} \longrightarrow V,\ D(f(x)) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

*Similarly, for any $a \in \mathbb{R}$, integration is a linear transformation*

$$I_a : V^{cont} \longrightarrow V^{diff},\ I_a(f(x)) = \int_a^x f(t)dt.$$

*The Fundamental Theorem of Calculus can (almost; notice domain discrepancy) then be summarized by the two formulas*

$$I_a \circ D(f(x)) = f(x) - f(a) \text{ and } D \circ I_a(f(x)) = f(x).$$

## §4.4 Bases and Dimension

Conveniently, every vector in $\mathbb{R}^2$ can be uniquely expressed using the two vectors $e_1 = (1,0),\ e_2 = (0,1)$. More precisely, any vector $(a,b) \in \mathbb{R}^2$ can be written as

$$(a,b) = ae_1 + be_2,$$

and the coefficients $a, b$ uniquely identify the vector $v$. A similar construction works for $\mathbb{R}^n$ (and indeed $F^n$). We now axiomatize this notion to any vector space.

**Definition 4.4.1: Bases**

Let $V_F$ be a vector space. A **finite**[a] **basis for** $V$ is a finite set of vectors $\mathcal{B} = \{v_1, \ldots, v_n\} \subseteq V$ with the following property:

Every vector $\vec{v} \in V$ can be **uniquely** written in the form

$$\vec{v} = a_1v_1 + \ldots + a_nv_n$$

for $a_1, \ldots, a_n \in F$.

An expression of the form $a_1v_1 + \ldots + a_nv_n$ is called a **linear combination** of $v_1, \ldots, v_n$.

[a] not every basis need be finite!

**Example 36.** *Let $F$ be a field. The* ***standard basis for*** *$F^n$ is the collection of vectors $\{e_1, \ldots, e_n\}$ where*

$$e_k = (0, \ \ldots, \ 1, \ \ldots, \ 0)$$

*the $k$-th coordinate is 1, and all others 0. The vector $\vec{v} = (a_1, \ldots, a_n) \in F^n$ can be written uniquely by the sum*

$$\vec{v} = a_1 e_1 + \ldots + a_n e_n.$$

*The coefficients $a_1, \ldots, a_n$ are the* ***coordinates of $\vec{v}$ for the standard basis of*** *$F^n$.*

The uniqueness property is important. For instance, while every vector in $\mathbb{R}^2$ can be expressed by some linear combination of $(1,0), (0,1)$, and $(1,1)$, the representation is not unique. How can we tell if a given set of vectors is a basis? Two important concepts must thus be defined, that can be used to readily check for basis properties.

**Definition 4.4.2: Span and Linear Independence**

Let $V_F$ be a vector space, and let $\mathcal{A} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ be a set of vectors in $V$.

- The set $\mathcal{A}$ **spans** $V$ if every vector in $V$ is a linear combination of vectors in $\mathcal{A}$; that is, for any vector $\vec{v} \in V$, there exist scalars $a_1, \ldots, a_n$ such that

  $$\vec{v} = a_1 v_1 + \ldots + a_n v_n.$$

  In general, we denote the **span of** $\mathcal{A}$ by

  $$\text{span}(\mathcal{A}) = \{a_1 v_1 + \ldots + a_n v_n \mid a_1, \ldots, a_n \in F\}.$$

  If $\mathcal{A}$ spans $V$, we say $\text{span}(\mathcal{A}) = V$.

- The set $\mathcal{A}$ is **linearly independent** if the only scalars $a_1, \ldots, a_n \in F$ that satisfy

  $$a_1 v_1 + \ldots + a_n v_n = \mathbf{0}$$

  are $a_1 = \ldots = a_n = 0$. A set is **linearly dependent** if it is not linearly independent.

**Proposition 4.4.1: Basis Conditions**

Let $V_F$ be a vector space, and let $\mathcal{A} = \{v_1, \ldots, v_n\}$ be a set of vectors in $V$. Then $\mathcal{A}$ is a basis for $V$ if and only if $\mathcal{A}$ spans $V$ and is linearly independent.

**Proof.** Suppose $\mathcal{A}$ is a basis. By definition, $\mathcal{A}$ spans $V$, and since

$$a_1 v_1 + \ldots + a_n v_n = 0 v_1 + \ldots + 0 v_n = \mathbf{0},$$

uniqueness of coefficients necessarily forces $a_1 = \ldots = a_n = 0$. Hence $\mathcal{A}$ is linearly independent.
Conversely, suppose $\mathcal{A}$ spans $V$ and is linearly independent. $\mathcal{A}$ spanning $V$ means every vector $v \in V$ can be expressed by a linear combination of vectors in $\mathcal{A}$, so it remains

to show uniqueness. Suppose $a_1v_1 + \ldots + a_nv_n = b_1v_1 + \ldots + b_nv_n = \mathbf{0}$. Then linear independence tells us that each $a_i - b_i$ in

$$\mathbf{0} = v - v = (a_1 - b_1)v_1 + \ldots + (a_n - b_n)v_n$$

is 0; that is, $a_i - b - i = 0$, so $a_i = b_i$. Hence uniqueness is satisfied, and so $\mathcal{A}$ is a basis for $V$. □

A basis for a vector space provides the building block for the entire space, so we would like to know if a basis exists. The following theorem lets us find a basis within a spanning set.

**Theorem 4.4.1: Goldilock's Theorem**

Let $V_F$ be a vector space, let $\mathcal{S}$ be a finite set of vectors in $V$, and let $\mathcal{L} \subseteq \mathcal{S}$ be a linearly independent subset of $\mathcal{S}$ ($\mathcal{L}$ can be the empty set.) Then there is a basis $\mathcal{B}$ for $V$ satisfying

$$\mathcal{L} \subseteq \mathcal{B} \subseteq \mathcal{S}.$$

Rephrasing, every spanning set of $V$ contains a basis of $V$, and every linearly independent set in $V$ can be extended to a basis. We call it Goldilocks Theorem, since although $\mathcal{S}$ may span $V$, it could be too big, and although $\mathcal{L}$ may be linearly independent, it could be too small. Thus, Goldilocks builds a basis $\mathcal{B}$ that's "just right", i.e. simultaneously "big enough" to span $V$ and "small enough" to be linearly independent.

**Proof.** We look at the collections of subsets in $\mathcal{S}$ that contain $\mathcal{L}$ and are linearly independent:

$$\{\mathcal{A} \mid \mathcal{L} \subseteq \mathcal{A} \subseteq \mathcal{S},\ \mathcal{A} \text{ is linearly independent } \}.$$

This collection is non-empty, since it always has $\mathcal{L}$. Choose $\mathcal{B}$ = the subset $\mathcal{A}$ with the largest number of elements. $\mathcal{B}$ has the following properties:

- $\mathcal{B}$ is a linearly independent subset of $\mathcal{S}$ that contains $\mathcal{L}$.
- There are no linearly independent subsets of $\mathcal{S}$ that contain $\mathcal{L}$ and have more elements than $\mathcal{B}$.

We claim $\mathcal{B}$ is a basis. By construction, $\mathcal{B}$ is linearly independent, so it remains to show spanning.

- First, we show $\mathcal{S} \subset \text{span}(\mathcal{B})$. Take an arbitrary element $v \in \mathcal{S}$. If $v \in \mathcal{B}$, then $v$ is certainly in the span of $\mathcal{B}$. Otherwise, $\mathcal{B} \cup \{v\}$ is strictly larger than $\mathcal{B}$, so by the second property $\mathcal{B} \cup \{v\}$ cannot be linearly independent. Thus, with $\mathcal{B} = \{v_1, \ldots, v_n\}$, there exists non-zero $a_1, \ldots, a_m, b \in F$ such that

  $$a_1v_1 + \ldots + a_mv_m + bv = 0.$$

  $b \neq 0$, since $v_1, \ldots, v_m$ is linearly independent; so

  $$v = -\frac{a_1}{b}v_1 - \ldots - \frac{a_m}{b}v_m \in \text{span}(\mathcal{B}).$$

  Hence $\mathcal{S} \subset \text{span}(\mathcal{B})$.

- Next, we show that $\text{span}(\mathcal{S}) \subseteq \text{span}(\mathcal{B})$. This follows from the more general fact that if $\mathcal{A}_1$, $\mathcal{A}_2$ are any two finite subsets of $V$, then

$$\mathcal{A}_1 \subset \text{span}(\mathcal{A}_2) \implies \text{span}(\mathcal{A}_1) \subseteq \text{span}(\mathcal{A}_2).$$

  This is true because if any vector $v \in \mathcal{A}_1$ is in the span of $\mathcal{A}_2$, then we can rewrite any linear combination $v = a_1v_1 + \ldots + a_mv_m$ of vectors in $\mathcal{A}_1$ in terms of vectors of $\mathcal{A}_2$ (since each individual $v_i$ is in $\text{span}(\mathcal{A}_2)$); hence $\text{span}(\mathcal{A}_1) \subseteq \text{span}(\mathcal{A}_2)$.

  This thus shows that $\text{span}(\mathcal{S}) \subseteq \text{span}(\mathcal{B})$; but $\text{span}(\mathcal{S}) = V$, so $\mathcal{B}$ spans $V$. Since we've shown that $\mathcal{B}$ spans $V$ and is linearly independent (by construction), we have, by Proposition 4.4.1, that $\mathcal{B}$ is a basis for $V$.

□

**Remark 9.** *Does every vector space admit a basis? We must first understand what it means for an infinite set $\mathcal{B}$ to be a basis. We say such a set is a basis if every vector $v \in V$ can be written uniquely as a linear combination of some finite subset of $\mathcal{B}$. We require a finite subset, since in general there's no way to compute infinite sums. With this definition, the assertion that every vector space has a basis is another statement equivalent to the Axiom of Choice.*

We now move on to dimension, and a result of great importance in applying vector spaces to other areas of mathematics.

**Definition 4.4.3: Dimension**

Let $V$ be a vector space with a finite basis. The **dimension of** $V$, denoted $\dim(V)$, is the number of vectors in a basis of $V$. A vector space without a finite basis is called **infinite-dimensional**.

**Example 37.** *The dimension of $F^n$ is $n$; the standard basis has $n$ elements. For any $n \geq 0$, the set of polynomials*

$$\{f(x) \in F[x] \mid \deg(f) \leq n\}$$

*is an $F$-vector space with dimension $n+1$. The set $\{1, x, \ldots, x^n\}$ is a basis.*

We start with a lemma that builds up to an important result:

**Lemma 4.4.1: Swap Lemma**

Let $V_F$ be a vector space, let $\mathcal{S}$ be a spanning set of $V$, and let $\mathcal{L}$ be a linearly independent set of vectors in $V$. Then given any vector $v \in \mathcal{L} \setminus \mathcal{S}$, we can find a vector $w \in \mathcal{S} \setminus \mathcal{L}$ such that

$$(\mathcal{S} \setminus \{w\}) \cup \{v\}$$

is still a spanning set.

In other words, we can swap a vector in $\mathcal{L}$ but not in $\mathcal{S}$ with a vector in $\mathcal{S}$ but not in $\mathcal{L}$, while preserving the spanning set property.

**Proof.** We use a vector $v \in \mathcal{L} \setminus \mathcal{S}$ to build two sets of vectors:

$$(\mathcal{L} \cap \mathcal{S}) \cup \{v\} \subset \mathcal{S} \cup \{v\}.$$

The left set is linearly independent, since it is a subset of $\mathcal{L}$; the right set is clearly still a spanning set. Thus, by Goldilocks Theorem, there exists a basis $\mathcal{B}$ in between

$$(\mathcal{L} \cap \mathcal{S}) \cup \{v\} \subseteq \mathcal{B} \subseteq \mathcal{S} \cup \{v\}.$$

Further, since $\mathcal{S}$ is a spanning set, and $v \notin \mathcal{S}$, the larger set $\mathcal{S} \cup \{v\}$ cannot be linearly independent (since $v$ can be written as a linear combination of vectors in $\mathcal{S}$). Hence $\mathcal{B}$ is not equal to $\mathcal{S} \cup \{v\}$, so there exists some $w \in \mathcal{S} \cup \{v\}$ such that $w \notin \mathcal{B}$. In particular, $w \neq v$, since $v \in \mathcal{B}$, $w \notin \mathcal{B}$. It follows that

$$\mathcal{B} \subseteq (\mathcal{S} \setminus \{w\}) \cup \{v\},$$

since $w \notin \mathcal{B}$ and $\mathcal{B} \subseteq \mathcal{S}$ implies $\mathcal{B} \subseteq \mathcal{S} \setminus \{w\}$ also. In other words, the set $(\mathcal{S} \setminus \{w\}) \cup \{v\}$ contains a basis, so it certainly is a spanning set. □

We can then use the swap lemma to show that no linearly independent set is larger than a spanning set.

**Lemma 4.4.2**

Let $V_F$ be a vector space, let $\mathcal{S} \subset V$ be a finite spanning set of $V$, and let $\mathcal{L} \subset V$ be a linearly independent set. Then

$$|\mathcal{L}| \leq |\mathcal{S}|.$$

**Proof.** If $\mathcal{L} \subseteq \mathcal{S}$, this is clearly true. Otherwise, we can find a vector $v \in \mathcal{L} \setminus \mathcal{S}$, a vector in $\mathcal{L}$ but not $\mathcal{S}$. The Swap Lemma says we can find a $w \in \mathcal{S} \setminus \mathcal{L}$, a vector in $\mathcal{S}$ but not $\mathcal{L}$, so that $(\mathcal{S} \setminus \{w\}) \cup \{v\}$ is still a spanning set. Let

$$\mathcal{S}' = (\mathcal{S} \setminus \{w\}) \cup \{v\};$$

note that $|\mathcal{S}'| = |\mathcal{S}|$. Also, note that $\mathcal{S}'$ and $\mathcal{L}$ now share an extra vector in common; or,

$$|(\mathcal{S}' \cap \mathcal{L})| \geq 1 + |(\mathcal{S} \cap \mathcal{L})|.$$

If $\mathcal{L} \subseteq \mathcal{S}'$, we are done; otherwise, we repeat the swapping process, swapping an element in $\mathcal{L}$ but not in $\mathcal{S}'$ into $\mathcal{S}'$, and removing an element from $\mathcal{S}'$. With each repetition, the number of shared elements between $\mathcal{L}$ and $\mathcal{S}''$ increase, until eventually $\mathcal{L} \subseteq \widetilde{\mathcal{S}}$ and $\left|\widetilde{\mathcal{S}}\right| = |\mathcal{S}|$. Moreover, we can't repeat the process forever, since there are only finitely many elements in $\mathcal{L}$ that can be swapped (otherwise, we would have a spanning set $\widetilde{\mathcal{S}}$ that spans $V$ and contains only elements of $\mathcal{L}$; yet we would have another $v \in \mathcal{L}$ to swap. This would imply that $\widetilde{\mathcal{S}} \cup \{v\}$ is linearly independent; but that's not possible, since $\widetilde{\mathcal{S}}$ already spans $V$ using only elements in $\mathcal{L}$).
Thus $|\mathcal{L}| \leq |\mathcal{S}|$. □

Now, we arrive at a fundamental theorem regarding dimension of bases:

**Theorem 4.4.2: Invariance of Dimension**

Let $V$ be a vector space with a finite basis. Then every basis for $V$ has the same number of elements.

**Proof.** We note that $V$ has at least one finite basis, $\mathcal{B}$. Let $\mathcal{B}'$ be any other basis. Since they are bases, we know that $\mathcal{B}$ spans and $\mathcal{B}'$ is linearly independent; thus

$$|\mathcal{B}'| \leq |\mathcal{B}|\,.$$

Moreover, we also know that $\mathcal{B}$ is linearly independent and $\mathcal{B}'$ spans; thus

$$|\mathcal{B}| \leq |\mathcal{B}'|\,.$$

Hence $|\mathcal{B}| = |\mathcal{B}'|$. □

# Chapter 5

# Fields: Part I

## §5.1 Introduction to Fields

Recall that in Chapter 3, we introduced fields as commutative rings with an additional property: every non-zero element had a multiplicative inverse. We extended this in Chapter 4 by building vector spaces with fields.

**Definition 5.1.1: Fields**

A **field** is a commutative ring $F$ with the property that for every $a \in F$ with $a \neq 0$, there is a $b \in F$ such that $ab = 1$.

Now, we begin studying fields in their own right. First, we look at how fields fit into each other, how to construct new fields from old fields, and describe all fields with finitely many elements. Finite fields are not just a mathematical curiosity; they play important roles in pure and applied mathematics, as well as engineering; some applications include signal processing, error correcting codes, and cryptography.

**Remark 10.** *Field theory was originally developed to aid the study of polynomials, such as finding the roots of a certain polynomial. Interestingly, field theory and finite group theory both originated from the study of polynomials! For more history, Google Cardano's Formula.*

## §5.2 Abstract Fields and Homomorphisms

Recall the definition of a unit group:

**Definition 5.2.1: Unit Groups**

Let $R$ be a commutative ring. The **unit group** of $R$ is the group

$$R^* = \{a \in R \mid \text{ there is some } b \in R \text{ such that } ab = 1\},$$

where the group law is ring multiplication.

Thus, a succinct way to characterize a field is that it is a commutative ring $F$ satisfying

$$F^* = \{a \in F \mid a \neq 0\} = F \setminus \{0\}.$$

As for maps between fields, we obviously want them to preserve field properties. In particular, since fields are rings, we want our maps to *at least* be ring homomorphisms.

It turns out that it's enough to ensure that multiplicative inverses go to multiplicative inverses; moreover, somewhat surprisingly, maps between fields are always injective! Note that this statement is generally not true for rings; the ring homomorphism $\phi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ is very non-injective.

**Proposition 5.2.1**

Let $F, K$ be fields, and let $\phi : F \to K$ be a ring homomorphism.

- The map $\phi$ is injective.
- Let $a \in F^*$. Then $\phi(a^{-1}) = \phi(a)^{-1}$.

**Proof.**

- From Theorem 3.31(b)ii, we need to show that $\ker(\phi)$ is the zero ideal, so suppose there is a non-zero element $a \in \ker(\phi)$. $a \neq 0$ in a field $F$ implies that $ab = 1_F$ for some $b \in F$; thus

$$1_K = \phi(1_F) = \phi(ab) = \phi(a)\phi(b) = 0 \cdot \phi(b) = 0_K,$$

  a contradiction (since ring axioms require $1 \neq 0$). Thus $\ker(\phi)$ is the zero ideal, and so $\phi$ is injective.

- By definition of multiplicative inverse, and since $\phi$ is a homomorphism, we have

$$1_K = \phi(1_F) = \phi(a \cdot a^{-1}) = \phi(a) \cdot \phi(a^{-1}).$$

  Hence $\phi(a^{-1})$ is a multiplicative inverse for $\phi(a)$; i.e. $\phi(a^{-1}) = \phi(a)^{-1}$.

  Alternatively, from (a) we know the map $\phi : F^* \to K^*$ is well-defined (by injectivity, since if $\phi(a) = \phi(a')$, then $a = a'$), and by the homomorphism property, it is a group homomorphism from $F^*$ to $K^*$. Thus, the fact that $\phi$ sends multiplicative inverses to multiplicative inverses is a result of Proposition 2.20 (which states that for group homomorphisms, $\phi(a^{-1})$ is the inverse of $\phi(a)$).

□

## §5.3 Interesting Examples of Fields

**Example 38.** *Three fields are already familiar:* $\mathbb{Q}, \mathbb{R}$, *and* $\mathbb{C}$. *Moreover, they fit into each other:*

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

**Example 39.** *The following subset of* $\mathbb{C}$ *is a field:*

$$\mathbb{Q}(i) = \{a + bi \mid a, b \in \mathbb{Q}\}.$$

*The multiplicative inverse of a non-zero element* $a + bi$ *is obtained by "rationalizing the denominator:"*

$$\frac{1}{a+bi} = \frac{1}{a+bi} \cdot \frac{a-bi}{a-bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{bi}{a^2+b^2}.$$

*In similar fashion, we can use $\sqrt{2}$ to describe a subfield of $\mathbb{R}$:*

$$\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\},$$

*with a multiplicative inverse obtained again by rationalizing the denominator:*

$$\frac{1}{a+b\sqrt{2}} = \frac{1}{a+b\sqrt{2}} \cdot \frac{a-b\sqrt{2}}{a-b\sqrt{2}} = \frac{a-b\sqrt{2}}{a^2-2b^2} = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}.$$

*This is well-defined, since $\sqrt{2}$ is not a rational number, so $a^2 - 2b^2 \neq 0$ for any $a, b \in \mathbb{Q}$ as long as they are non-zero as well.*
*What about the set*

$$\{a + b\sqrt{2} + c\sqrt{3} \mid a, b, c \in \mathbb{Q}\}?$$

*This is not even a ring, since one can see $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$, which is not in the set. The larger set*

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\}$$

*is a field. It's easy to see ring properties, but proving that non-zero elements have an inverse is not quite so easy. We'll see tools to help with this later.*

**Example 40.** *In general, the ring $\mathbb{Z}/m\mathbb{Z}$ need not be a field. For example, $\mathbb{Z}/6\mathbb{Z}$ is not a field, since 2 does not have a multiplicative inverse. However, we've seen before that any ring $\mathbb{Z}/p\mathbb{Z}$, where $p$ is a prime, is a field; we denote this $\mathbb{F}_p$. $\mathbb{F}_p$ is an example of a **finite field**; it turns out that there are other finite fields. Indeed, for every prime power $p^k$, there is exactly one (up to isomorphism) finite field containing $p^k$ elements.*

**Example 41.** *A **skew field**, also called a **division ring**, is a ring in which every non-zero element has an inverse, but is no longer required to be commutative. Wedderburn's Theorem states that every finite skew field must be commutative (e.g. is a field), but there are also many interesting non-commutative infinite fields. One such example is $\mathbb{H}$, the ring of quaternions.*

## §5.4 Subfields and Extension Fields

**Definition 5.4.1: Subfields**

Let $K$ be a field. A **subfield of** $K$ is a subset $F$ of $K$ that is itself a field using the addition and multiplication operations of $K$.

**Definition 5.4.2: Extension Fields**

Let $F$ be a field. An **extension field of** $\mathbb{F}$ is a field $K$ such that $F$ is a subfield

of $K$. We write $K/F$ to indicate that $K$ is an extension field of $F$.[a]

[a]Note that $K/F$ is just a piece of convenient notation. It doesn't mean that we're taking the quotient of $K$ by $F$, despite its similarities to a quotient ring $R/I$.

**Example 42.** *The field $\mathbb{Q}$ is a subfield of $\mathbb{R}$, and thus $\mathbb{R}$ is an extension field of $\mathbb{Q}$; similarly with $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.*
*The fields $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt{2})$ are extension fields of $\mathbb{Q}$. The former is a subfield of $\mathbb{C}$ but not $\mathbb{R}$, while the latter is a subfield of $\mathbb{R}$. Neither are subfields of each other.*

The notation $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{2})$ is a special case of the following general construction.

**Proposition 5.4.1**

Let $L/F$ be an extension of fields, and let $\alpha_1, \ldots, \alpha_n \in L$. Then there is a unique field $K$ with the following properties:

1. $F \subseteq K \subseteq L$.
2. $\alpha_1, \ldots, \alpha_n \in K$.
3. If $K'$ is a field satisfying $F \subseteq K' \subseteq L$ and $\alpha_1, \ldots, \alpha_n \in K'$, then $K \subseteq K'$.

The field $K$ is denoted $F(\alpha_1, \ldots, \alpha_n)$ and is called the **extension field of $F$ generated by** $a_1, \ldots, a_n$. Intuitively, it is the smallest subfield of $L$ that contains both $F$ and $\alpha_1, \ldots, \alpha_n$.

**Proof.** Let $S$ be the set consisting of all subfields of $L$ that contain $F$ and $\alpha_1, \ldots, \alpha_n$. The set is not empty, since $L \in S$. Let $K$ be the intersection of all of the fields in $S$. Then $K$ is a field, since

$$\begin{aligned} \alpha, \beta \in K &\iff \alpha, \beta \in K' \text{ for every } K' \in S, \\ &\implies \alpha \pm \beta, \alpha\beta, \alpha^{-1} \in K' \text{ for every } K' \in S, \quad \text{since } K' \text{ is a field }, \\ &\implies \alpha \pm \beta, \alpha\beta, \alpha^{-1} \in K.. \end{aligned}$$

Clearly, $K$ contains $F$ and $\alpha_i$, since it is the intersection of fields that contain both; and if $K' \subseteq L$ contains both, then $K' \in S$, so $K'$ is one of the fields whose intersection forms $K$. Hence $K \subseteq K'$. □

Let $K/F$ be an extension of fields. Observe that we can add elements of $K$, and multiply elements in $K$ by elements in $F$; thus, the field $K$ becomes an $F$-vector space. Essentially, we discard most of the multiplication operation in $K$, and restrict it solely to multiplication by elements in $F$. This allows us to use tools from linear algebra to study field extensions.

**Definition 5.4.3: Degree of Field Extensions**

Let $K/F$ be an extension of fields. The **degree of $K$ over $F$**, denoted $[K : F]$, is the dimension of $K$ viewed as an $F$-vector space,

$$[K : F] = \dim_F(K).$$

If $[K : F]$ is finite, we say $K/F$ is a **finite extension**; otherwise, we say $K/F$ is an **infinite extension**.

**Example 43.** *The fields* $\mathbb{Q}(i)$, $\mathbb{Q}(\sqrt{2})$ *have degree* 2 *over* $\mathbb{Q}$*:*

- $[\mathbb{Q}(i) : \mathbb{Q}] = 2$*, since* $\{1, i\}$ *is a* $\mathbb{Q}$*-basis for* $\mathbb{Q}(i)$*.*
- $\left[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}\right] = 2$*, since* $\{1, \sqrt{2}\}$ *is a* $\mathbb{Q}$*-basis for* $\mathbb{Q}(\sqrt{2})$*.*

*Similarly, we have* $[\mathbb{C} : \mathbb{R}] = 2$*, since* $\{1, i\}$ *is an* $\mathbb{R}$*-basis for* $\mathbb{C}$*. On the other hand,* $[\mathbb{R} : \mathbb{Q}] = \infty$*:*

**Proof.** Suppose that $\{a_1, \ldots, a_n\} \subset \mathbb{R}$ is a finite $\mathbb{Q}$-basis for $\mathbb{R}$. Then

$$\mathbb{R} = \{c_1 a_1 + \ldots + c_n a_n \in \mathbb{Q}\}.$$

But the set on the right is countable (since its cardinality is the same as the cardinality of the set of $n$-tuples of rational numbers), while $\mathbb{R}$ is uncountable. □

The next theorem is similar to the index multiplication rule, which counted cosets in a chain of groups.

**Theorem 5.4.1**

Let $L/K/F$ be an extension of fields (that is, $L$ is a field extension of $K$, while $K$ is a field extension of $F$). Then

$$[L : F] = [L : K] \cdot [K : F],$$

in the sense that one of the following is true:

- All of the degrees $[L : F], [L : K], [K : F]$ are finite, and the above equation is true.
- $[L : F] = \infty$, and either $[L : K] = \infty$ or $[K : F] = \infty$.

**Proof.** We start with the case $L/K$ and $K/F$ are finite extensions. This means that we can choose bases

- $\mathcal{A} = \{a_1, \ldots, a_m\}$, a basis for $K$ as an $F$-vector space
- $\mathcal{B} = \{b_1, \ldots, b_n\}$, a basis for $L$ as a $K$-vector space

We claim that the $mn$ products in the set

$$\mathcal{C} = \{a_i b_j \mid 1 \leq i \leq m,\ 1 \leq j \leq n\}$$

are distinct and that they form a basis for $L$ as an $F$-vector space. Assuming this, it's easy to prove the equation:

$$[L : F] = \dim_F(L) = |\mathcal{C}| = mn = |\mathcal{A}|\,|\mathcal{B}| = \dim_F(K) \cdot \dim_K(L) = [K : F] \cdot [L : K].$$

We first prove that this is a linearly independent set. Suppose we have an $F$-linear combination of the elements of $\mathcal{C}$ that sum to 0:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} a_i b_j = 0,\ c_{ij} \in F.$$

We wish to show that every $c_{ij}$ is zero. We switch the order of the sums (due to commutativity) to get

$$\sum_{j=1}^{n} \left( \sum_{i=1}^{m} c_{ij} a_i \right) b_j = 0.$$

The inner sum is in $K$, since $c_{ij} \in F$, $a_i \in K$. This gives a linear combination of $b_j$ with coefficients in $K$; but $b_i \in \mathcal{B}$ is a basis for $L$ as a $K$-vector space, so it is $K$-linearly independent. That is, every $c_{ij} a_i = 0$, so we have

$$\sum_{i=1}^{m} c_{ij} a_i.$$

But we also know that $a_i \in \mathcal{A}$ is a basis for $K$ as an $F$-vector space, and each $c_{ij}$ is in $F$, so $F$-linear independence of $a_1, \ldots, a_m$ implies that every $c_{ij} \in F$ is 0, and so $\mathcal{C}$ is a linearly independent set.
Spanning is left as an exercise to the reader. □

## §5.5 Polynomial Rings

We will now briefly discuss some properties of polynomials with coefficients inside a field $F$.

**Definition 5.5.1: Degree of $f$**

Let $F$ be a field, and let $f(x) \in F[x]$ be a non-zero polynomial, where

$$f(x) = a_0 + a_1 x + \ldots + a_d x^d, a_d \neq 0;$$

in other words, $d$ is the highest power with a non-zero coefficient. The **degree of** $f$ is

$$\deg(f) = d.$$

By convention, we set $\deg(0) = -\infty$; further, if $a_d = 1$, then we say that $f$ is a **monic polynomial**.

One can easily check that for any two polynomials $f_1, f_2 \in F[x]$, we have

$$\deg(f_1 f_2) = \deg(f_1) + \deg(f_2).$$

A useful property of polynomials is the following algorithm, which allows one to extend division into the ring of polynomials. It is analogous to the Division Algorithm for integers.

**Proposition 5.5.1: Division-with-Remainder for Polynomials**

Let $F$ be a field, and let $f(x), g(x) \in F[x]$ be a polynomial with $g(x) \neq 0$. Then there are unique polynomials $q(x), r(x) \in F[x]$ satisfying

$$f(x) = g(x)q(x) + r(x) \text{ with } \deg(r) < \deg(g).$$

**Proof.** We prove existence by following the long-division with remainder algorithm for polynomials, and we use induction to generalize for any degree. So, if we assume that the degree of the polynomial $g(x)$ is fixed, say with

$$d = \deg(g) \geq 1,$$

our goal is to verify the statement $P_n$:

> If $f(x) \in F[x]$ is a polynomial with $\deg(f) = n$, then there are unique polynomials $q(x), r(x) \in F[x]$ satisfying
>
> $$f(x) = g(x)q(x) + r(x) \text{ and } \deg(r) < \deg(g).$$

We first prove that $P_0, \ldots, P_{d-1}$ are true. Let $f(x) \in F[x]$ with $\deg(f) < d$. Then we can take $q(x) = 0,\ r(x) = f(x)$, since

$$f(x) = 0 \cdot g(x) + f(x)$$

is clearly true, and $\deg(f) < d = \deg(g)$ by definition.
Now, let $n \geq d$, and assume $P_0, \ldots, P_{n-1}$ are true. Let $f(x) \in F[x]$ be a polynomial with $\deg(f) = n$; that is, we have

$$f(x) = ax^n + \ldots \text{ and } g(x) = bx^d + \ldots,\ a, b \in F,\ a, b \neq 0.$$

Then, if we multiply $g(x)$ by $\frac{a}{b}x^{n-d}$, we get a new polynomial

$$h(x) = f(x) - \frac{a}{b}x^{n-d}g(x).$$

Then $\deg(h) < n$, since we canceled out the $ax^n$ term in $f(x)$. Then, we can use the inductive hypothesis to find $q(x), r(x) \in F[x]$ such that

$$h(x) = g(x)q(x) + r(x),\ \deg(r) < \deg(g).$$

But then

$$\begin{aligned} f(x) &= h(x) + \frac{a}{b}x^{n-d}g(x) \\ &= g(x)q(x) + r(x) + \frac{a}{b}x^{n-d}g(x) \\ &= g(x)(q(x) + \frac{a}{b}a^{n-d}) + r(x). \end{aligned}$$

We have thus written $f(x)$ in the desired form:

$f(x) = g(x)\cdot$(polynomial)+(polynomial whose degree is strictly smaller than the degree of $g(x)$) .

(since $q(x) + \frac{a}{b}a^{n-d} \in F[x]$). Uniqueness is left as an exercise. □

Now, we move on to ideals of polynomial rings. Recall that an ideal is **principal** if it consists of multiples of a single element of $R$; that is, for $c \in R$,

$$I = \{cr \mid r \in R\}.$$

This ideal is called the **principal ideal generated by** $c$, and is often denoted $cR$ or $(c)$. Principal ideals are quite easy to work with, which makes the following theorem quite interesting and important.

**Theorem 5.5.1**

Let $F$ be a field, and let $I \subseteq F[x]$ be an ideal in the ring $F[x]$. Then $I$ is a principal ideal.

**Proof.** If $I$ is the zero ideal ($I = (0)$), then $I$ is clearly a principal ideal, so assume $I \neq (0)$; that is, $I$ contains at least one non-zero polynomial. Let $g(x) \in I$ be the non-zero polynomial with the lowest degree. We claim $I = (g)$.
Take an arbitrary $f \in I$. From the Division Algorithm for polynomials, we know that

$$f(x) = g(x)q(x) + r(x)$$

for some $q(x), r(x) \in F[x]$, $\deg(r) < \deg(g)$. Thus

$$r(x) = f(x) - g(x)q(x) \in I,$$

since $f(x), g(x)q(x) \in I$. But $g(x)$ has the smallest possible degree of any non-zero polynomial in $I$, while we also have $\deg(r) < \deg(g)$. Thus $r(x) = 0$ necessarily, and so

$$f(x) = g(x)q(x) \in (g) = g(x)F[x].$$

Thus $I \subseteq (g)$; and the fact that $g \in I$ forces $(g) \subseteq I$ by ideal properties. Hence $I = (g)$ is a principal ideal. □

## §5.6 Building Extension Fields

Now, we start building fields using roots of polynomials. However, we sometimes don't know where these roots might live. So instead, we take a field $F$ and a polynomial $f(x) \in F[x]$, and use them to construct a field extension $K/F$ that "magically" contains a root of $f(x)$. A prototype for this construction is the abstract construction of the field of complex numbers from the real numbers and the polynomial $x^2 + 1$ (recall how we can find a map $\phi : \mathbb{R}[x]/(x^2+1)\mathbb{R}[x] \to \mathbb{C}, \phi(f(x)) = f(i)$, by taking the evaluation homomorphism $E_i$).

**Definition 5.6.1: Irreducibility**

Let $F$ be a field. A non-constant polynomial $f(x) \in F[x]$ is **reducible** (over $F$) if there exist non-constant polynomials $g(x), h(x) \in F[x]$ such that $f(x) = g(x)h(x)$ ($f$ factors into them). An **irreducible polynomial** is a non-constant polynomial that is not reducible, i.e., a polynomial with no non-trivial factorizations in $F[x]$; that is, the only factorizations are constant polynomials and multiples of itself (we will later generalize this to arbitrary rings).

**Example 44.** *The polynomials $x^2 - 1$ and $x^2 - x - 2$ are reducible in $\mathbb{Q}[x]$:*

$$x^2 - 1 = (x+1)(x-1) \text{ and } x^2 - x - 2 = (x-1)(x-2).$$

*The polynomials $x^2 + 1$ and $x^2 - 2$ are irreducible in $\mathbb{Q}[x]$ (although the second one is equivalent to saying $\sqrt{2}$ is not rational). On the other hand, in the ring $\mathbb{Q}(i)[x]$, $x^2 + 1$ is reducible, since it can factor as*

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

*Reducibility depends on both the polynomial and the underlying field.*

**Example 45.** *If we take finite fields as coefficients, then there are only finitely many polynomials of any given degree. For instance, only four quadratic polynomials exist in* $\mathbb{F}_2[x]$*:*

$$x^2 = x \cdot x,\ x^2 + 1 = (x+1)^2,\ x^2 + x = x(x+1),\ x^2 + x + 1.$$

*Only the last one is irreducible. Similarly, there are eight cubic polynomials in* $\mathbb{F}_3[x]$*; only* $x^3 + x + 1$ *and* $x^3 + x^2 + 1$ *are irreducible.*

**Remark 11.** *Every polynomial of degree* 1 *is irreducible, and it's not hard to show that if* $f(x)$ *has degree* 2 *or* 3*, then it is irreducible if and only if it has no roots. However, this doesn't necessarily hold for higher degrees; for instance, given two irreducible polynomials of degree* 2*,* $g(x),\ h(x)$*, then* $g(x)h(x)$ *is a reducible polynomial of degree* 4*, yet it has no roots in* $F$*.*

**Theorem 5.6.1**

Let $F$ be a field, and let $f(x) \in F[x]$. The following are equivalent:

1. The polynomial $f(x)$ is irreducible.
2. The principal ideal $f(x)F[x]$, generated by $f(x)$, is a maximal ideal.
3. The quotient ring $F[x]/f(x)F[x]$ is a field.

**Proof.** The equivalence of the second and third statements is a result of Theorem 3.40 ($I$ is a maximal ideal if and only if $R/I$ is a field).
We start by proving $f(x)$ irreducible implies $f(x)F[x]$ is a maximal ideal. Suppose $I$ is an ideal satisfying

$$f(x)F[x] \subseteq I \subseteq F[x].$$

We wish to show $I = F[x]$ or $I = f(x)F[x]$.
From the Theorem above, we know that every ideal in $F[x]$ is principal, so we can find some $g(x) \in I$ such that $g(x)F[x] = I$. The inclusion of fields thus becomes

$$f(x)F[x] \subseteq g(x)F[x] \subseteq F[x].$$

In particular, since $f(x)$ is in the principal ideal generated by $g(x)$, we have

$$f(x) = g(x)h(x),\ h(x) \in F[x].$$

Now, we use the irreducibility of $f(x)$. Since $f(x)$ is irreducible, one of $g(x),\ h(x)$ is a non-zero constant polynomial. This gives two cases:

- $g(x) \in F^*$, which tells us $I = g(x)F[x] = F[x]$ (since $g(x) \in F^*$ and $g$ constant implies that its inverse is also in $F[x]$, so we can take $g(x)g^{-1}(x) = 1 \in I$. $1 \in I$ then forces $I = F[x]$.)

- $h(x) \in F^*$, which tells us $f(x)F[x] = g(x)h(x)F[x] = g(x)F[x] = I$ (since $h(x) \in F^*$ means we can find an inverse to cancel it out with.)

Thus $f(x)F[x]$ is a maximal ideal.
We then show the reverse: suppose $f(x)F[x]$ is a maximal ideal. Suppose $f(x)$ factors into $g(x)$ and $h(x)$ for some $g(x), h(x) \in F[x]$. We wish to show that either $g(x)$ or $h(x)$ is a constant polynomial. Consider the ideals

$$f(x)F[x] \subseteq g(x)F[x] \subseteq F[x].$$

$f(x)F[x]$ maximal means that either $f(x)F[x] = g(x)F[x]$ or $g(x)F[x] = F[x]$, so we consider both cases.

- Suppose $f(x)F[x] = g(x)F[x]$. Since $f(x) = g(x)h(x)$, we can replace $f(x)F[x]$ with $g(x)h(x)F[x]$. Then we have

$$\begin{aligned} g(x) \in f(x)F[x] = g(x)h(x)F[x] &\implies g(x) = g(x)h(x)s(x),\ s(x) \in F[x] \\ &\implies h(x)s(x) = 1 \qquad [\text{ by properties of fields: } g \cdot f = g \end{aligned}$$

  Thus $h(x)$ (and $s(x)$) must necessarily be a constant polynomial.

- Suppose $g(x)F[x] = F[x]$. Then $1 \in g(x)F[x]$, so $g(x)t(x) = 1$ for some $t(x) \in F[x]$, and so $g(x)$ must be a constant polynomial.

In both cases, either $g(x)$ or $h(x)$ is a constant polynomial; hence the only way to factor $f(x) = g(x)h(x)$ is for one of them to be constant. Thus $f(x)$ is definitionally irreducible. □

Now, given a field $F$ and a polynomial $f(x) \in F[x]$, we wish to use this theorem to construct an extension field of $F$ that contains a root of $f$. Later, we'll see a more general result; but even this one is spectacular: it says that if you're given any irreducible polynomial $f(x) \in F[x]$, if you're willing to make the field $F$ slightly larger, then you can find a root of $f(x)$.

**Theorem 5.6.2**

Let $F$ be a field, let $f(x) \in F[x]$ be an irreducible polynomial, let $I_f = f(x)F[x]$ be the principal (maximal) ideal generated by $f(x)$, and let $K_f = F[x]/I_f$ be the quotient ring.

1. The ring $K_f$ is a field.
2. The polynomial $f(x)$ has a root in the field $K_f$.
3. The field $K_f$ is a finite extension of the field $F$. Its degree is given by

$$[K_f : F] = \deg(f).$$

**Proof.** 1. The above theorem tells us that $I_f$ is a maximal ideal of the ring $F[x]$, and Theorem 3.40b (maximal ideal iff quotient ring is field) tells us that $K_f = F[x]/I_f$ is a field.

2. (It almost seems like cheating!) We claim that

$$\beta = x + I_f \in K_f$$

is a root of $f(x)$. To see this, write $f(x)$ as

$$f(x) = b_0 + b_1 x + \ldots + b - dx^d.$$

Then

$$\begin{aligned} f(\beta) &= b_0 + b_1\beta + \ldots + b_d\beta^d \\ &= b_0(1 + I_f) + b_1(x + I_f) + \ldots + b_d(x + I_f)^d \\ &= (b_0 + I_f) + (b_1 x + I_f) + \ldots + (b_d x^d + I_f) && [\text{ convince yourself why this is true, given the properti} \\ &= b_0 + b_1 x + \ldots + b_d x^d + I_f && [\text{ by coset addition }]. \end{aligned}$$

Thus $f(\beta) = 0 + I_f$, and so $f(\beta)$ is the zero element of the field $K_f$.

3. Let $d = \deg(f)$, and let

$$\beta = x + I_f = \text{ the coset of } x \text{ in the quotient ring } F[x]/I_f \ .$$

We claim that the set

$$\mathcal{B} = \{1, \beta, \ldots, \beta^{d-1}\}$$

forms an $F$-basis for $K_f$.

We first show that $\mathcal{B}$ spans. Let $g(x) + I_f$ be an arbitrary element of $K_f$. We divide the polynomial $g(x)$ by $f(x)$ to get a quotient and remainder:

$$g(x) = f(x)q(x) + r(x),\ q(x), r(x) \in F[x],\ \deg(r) < \deg(f) = d.$$

Since $f(x) \in I_f$, the cosets $g(x) + I_f$ and $r(x) + I_f$ are the same. On the other hand, if we write $r(x)$ as

$$r(x) = a_0 + a_1 x + \ldots + a_{d-1}x^{d-1} \text{ with } a_0, \ldots, a_{d-1} \in F,$$

then we can compute

$$\begin{aligned} g(x) + I_f &= r(x) + I_f \\ &= (a_0 + a_1 x + \ldots + a_{d-1}x^{d-1}) + I_f \\ &= a_0(1 + I_f) + a_1(x + I_f) + \ldots + a_2(x + I_f)^2 + \ldots + a_{d-1}(x + I_f)^{d-1} \\ &= a_0\beta + \ldots + a_{d-1}\beta^{d-1}. \end{aligned}$$

Thus, any element of $K_f = F[x]/I_f$ can be written as an $F$-linear combination of the elements in $\mathcal{B}$, and so $\mathcal{B}$ spans $K_f$.

Next, we show that $\mathcal{B}$ is linearly independent. Suppose we have

$$c_0 + c_1\beta_1 + \ldots + c_{d-1}\beta^{d-1} = 0,\ c_i \in F.$$

We wish to show that every $c_i$ vanishes. Using $\beta = x + I_f$, we have

$$\begin{aligned} 0 + I_f &= c_0 + c_1\beta + \ldots + c_{d-1}\beta^{d-1} \\ &= c_0(1 + I_f) + c_1(x + I_f) + \ldots + c_{d-1}(x + I_f)^{d-1} \\ &= (c_0 + I_f) + (c_1 x + I_f) + \ldots + (c_{d-1}x^{d-1} + I_f) \\ I_f &= (c_0 + c_1 x + \ldots + c_{d-1}x^{d-1}) + I_f. \end{aligned}$$

This tells us that the polynomial $c_0 + c_1 x + \ldots + c_{d-1}x^{d-1}$ is in the principal ideal $I_f$ generated by $f(x)$, so for some $g(x) \in F[x]$ we have

$$c_0 + c1x + \ldots + c_{d-1}x^{d-1} = f(x)g(x).$$

If $g(x) \neq 0$, we can take the degrees of both sides to obtain a contradiction:

$$d-1 \geq \deg(c_0+\ldots+c_{d-1}x^{d-1}) = \deg(f(x)g(x)) = \deg(f(x))+\deg(g(x)) = d+\deg(g(x)) = d.$$

Hence $g(x) = 0$, and so the above equation becomes the following equality

$$c_0 + \ldots + c_{d-1}x^{d-1} = 0$$

that is true in $F[x]$. Therefore $c_0 = c_1 = \ldots = c_{d-1} = 0$, which proves linear independence. Thus $\mathcal{B}$ is a basis for $K_f/F$, and so

$$[K_f : F] = |\mathcal{B}| = d = \deg(f).$$

□

## §5.7 Finite Fields

Now, we'll apply all of our tools that we've developed to explore finite fields, and construct finite fields of various prime power orders.
Before starting, we recall the definition of a **characteristic** of a ring $R$:

**Definition 5.7.1: Characteristic**

The **characteristic** of a ring $R$ is the integer $m \geq 0$ generating the kernel of the unique homomorphism

$$\phi : \mathbb{Z} \longrightarrow R.$$

Alternatively, $m$ is the smallest integer such that

$$\underbrace{\alpha + \ldots + \alpha}_{m \text{ terms}} = 0.$$

We know that $\phi$ is unique, since $\phi(1) = 1_R$, which determines $\phi(n)$ for all $n \geq 0$. Moreover, recall that every ideal of $\mathbb{Z}$ is principal, and the kernel of any ring homomorphism is an ideal of the starting ring. Thus, when we say $m$ generates the kernel, we mean

$$\ker(\phi) = m\mathbb{Z}.$$

**Proposition 5.7.1**

Let $F$ be a finite field.

1. The characteristic of $F$ is prime.
2. Let $p = \text{char}(F)$. Then the finite field $\mathbb{F}_p$ is a subfield of $F$, in the sense that there is a unique injective homomorphism

$$\mathbb{F}_p \hookrightarrow F.$$

3. The number of elements of $F$ is given by

$$|F| = p^{[F:\mathbb{F}_p]}.$$

In particular, the number of elements in a finite field is always the power of a prime.

**Proof.** 1. Let $m = \text{char}(F)$; by definition, $m\mathbb{Z}$ is the kernel of the unique homomorphism $\phi : \mathbb{Z} \to F$. It follows from Proposition 3.31 that $\phi$ induces an injective ring homomorphism

$$\overline{\phi} : \mathbb{Z}/m\mathbb{Z} \hookrightarrow F.$$

Moreover, $F$ finite tells us that $m \neq 0$ (since eventually, elements will have to cycle back again, or else it would be infinite). We need to show that $m$ is prime.

Suppose $m$ factors into $m = kn$ (we put lines above integers to distinguish elements in $\mathbb{Z}$ and $m\mathbb{Z}$). Now,

$$0 = \overline{\phi}(\overline{0}) = \overline{\phi}(\overline{m}) = \overline{\phi}(\overline{kn}) = \overline{\phi}(\overline{k}) \cdot \overline{\phi}(\overline{n}).$$

Since $F$ is an integral domain (it has no zero divisors), we conclude that either

$$\overline{\phi}(\overline{k}) = 0 \text{ or } \overline{\phi}(\overline{n}) = 0.$$

But $\overline{\phi}$ is injective; thus either $\overline{k} = 0$ or $\overline{n} = 0$. In other words, $k + m\mathbb{Z} = 0 + m\mathbb{Z} = m\mathbb{Z}$ or $n + m\mathbb{Z} = m\mathbb{Z}$; hence $m \mid k$ or $m \mid n$, and thus the factorization $m = kn$ is a trivial factorization. This shows that $m$ has no non-trivial factorizations, so it is prime.

2. We know from $a$ that $m$ is prime, so the injective map $\overline{\phi}$ gives this result.

3. From (b), we know that $F$ is a field extension of $\mathbb{F}_p$. And since $F$ has only finitely many elements, the dimension of $F$ as an $\mathbb{F}_p$-vector space must be finite, since any basis is a subset of $F$. Let $d = [F : \mathbb{F}_p] = \dim_{\mathbb{F}_p}(F)$, and let

$$\mathcal{B} = \{\beta_1, \ldots, \beta_d\}$$

be an $\mathbb{F}_p$-basis for $F$ (i.e. it has elements in $F$, and forms a basis with coefficients in $\mathbb{F}_p$). The definition of basis tells us that every element of $F$ looks like

$$c_1\beta_1 + \ldots + c_d\beta_d \text{ with } c_1, \ldots, c_d \in \mathbb{F}_p,$$

and that the different choices of $c_1, \ldots, c_d$ give distinct elements of $F$. In fancier terms, because every element in $F$ is spanned by unique $c_1, \ldots, c_d \in \mathbb{F}_p$, there is a bijection defined by

$$\mathbb{F}_p^d \longrightarrow F, \ (c_1, \ldots, c_d) \mapsto c_1\beta_1 + \ldots + c_d\beta_d.$$

Hence

$$|F| = (\mathbb{F}_p)^d = p^d.$$

Thus the number of elements in a finite field is always the power of a prime.

□

The full proof of the next result is highly involved, and would lead us too far afield (hehe), so we prove up to degree 3, and briefly indicate a way to prove it for higher degrees. This result is by no means obvious: for example, it would not be true if we replaced the field $\mathbb{F}_p$ with $\mathbb{R}$, since every irreducible polynomial in $\mathbb{R}[x]$ has degree 1 or 2.

**Theorem 5.7.1**

Let $p$ be a prime, and let $d \geq 1$. Then the ring $\mathbb{F}_p[x]$ contains an irreducible polynomial of degree $d$.

**Proof.** For $d = 1$, any degree 1 polynomial will work.
Since we can always multiply a polynomial by a non-zero scalar without affecting its irreducibility, it suffices to look at **monic polynomials**; that is, polynomials whose leading coefficient is 1. We let

- $\text{Poly}_d = \{\text{monic polynomials in } \mathbb{F}_p[x] \text{ having degree } d\}$
- $\text{Irred}_d = \{\text{monic irreducible polynomials in } \mathbb{F}_p[x] \text{ having degree } d\}$
- $\text{Red}_d = \{\text{monic reducible polynomials in } \mathbb{F}_p[x] \text{ having degree } d\}$

We note that $|\text{Poly}_d| = p^d$, since a monic polynomial of degree $d$ in $\mathbb{F}_p[x]$ has exactly $d$ coefficients, each of which may freely be chosen in $\mathbb{F}_p$ (and thus there are $p^d$ different choices).
We start with $d = 2$, so we wish to compute $|\text{Irred}_p|_2$, the number of irreducible monic quadratic polynomials. We won't compute this directly, but instead use a lesson from combinatorics: we'll count the number of reducible monic quadratic polynomials (the polynomials we don't want) and subtract this value from the total number of monic quadratic polynomials.
So, which quadratic polynomials are irreducible? Precisely the polynomials that factor as

$$x^2 + ax + b = (x - a)(x - b) \text{ for some } a, b \in \mathbb{F}_p.$$

It may look like there are $p^2$ different choices for $a, b$, but we must not double count polynomials. There are two cases:

- $a = b$: then we get $p$ different polynomials $(x - a)^2$ for some $a \in \mathbb{F}_p$.
- $a \neq b$: then there are $p(p-1)$ choices for the pair $(a, b)$, but since the order of $a$ and $b$ does not change the product $(x - a)(x - b)$, the number of different polynomials we get is $\binom{p}{2}$ ($p$ choose 2, where order doesn't matter).

Thus,

$$|\text{Red}_2| = p + \binom{p}{2} = p + \frac{p(p-1)}{2} = \frac{p^2 + p}{2}.$$

Removing these reducible polynomials from the totality of all monic polynomials gives us

$$|\text{Irred}_2| = |\text{Poly}_2| - |\text{Red}_2| = p^2 - \frac{p^2 + p}{2} = \frac{p^2 - p}{2}.$$

This is positive for all $p \geq 2$, and so $\text{Irred}_2$ is non-empty.
Now, we look at cubic polynomials. We wish to count how many are irreducible. There's quite a few ways to factor a monic cubic polynomial; we list them here, along with how many distinct ones there are:

- $(x - a)^3$; $a \in \mathbb{F}_p$, so $p$ distinct polynomials.
- $(x-a)^2(x-b)$; $a, b \in \mathbb{F}_p$, $a \neq b$, so $p(p-1)$ distinct polynomials. (Note that order **does** matter here, so we do not divide by 2).
- $(x - a)(x - b)(x - c)$; $a, b, c \in \mathbb{F}_p$, all distinct, so $\binom{p}{3}$

- $(x-a)(x^2+bx+c)$; $a,b,c \in \mathbb{F}_p$, $x^2+bx+c$ irreducible, so $p \cdot |\text{Irred}_2| = \frac{p^3-p^2}{2}$.

Thus the total number of reducible monic cubic polynomials is

$$\text{Red}_3 = p + p(p-1) + \frac{p(p-1)(p-2)}{6} + \frac{p^3-p^2}{2} = \frac{2p^3+p}{3}.$$

Hence

$$|\text{Irred}_3| = |\text{Poly}_3| - |\text{Red}_3| = p^3 - \frac{2p^3+p}{3} = \frac{p^3-p}{3}.$$

Thus $\text{Irred}_3$ is non-empty for all $p \geq 2$. □

**Example 46.** *Taking $p = 2$ in the formulas*

$$|\mathit{Irred}_2| = \frac{1}{2}(p^2-2) \text{ and } |\mathit{Irred}_3| = \frac{1}{3}(p^3-p)$$

*that we derived while proving the above theorem shows that $\mathbb{F}_2[x]$ has exactly one monic quadratic polynomials, and exactly two monic cubic polynomials. These values agree with an investigation in an earlier example.*

We're going to use the above theorem to prove the first part of the following fundamental theorem, but for completeness we include both parts. A later section will provide an alternative proof for the first part, and an even later section will provide both yet another alternative proof and a proof for the second part.

**Theorem 5.7.2**

Let $p$ be a prime, and let $d \geq 1$.

1. There exists a field $F$ containing exactly $p^d$ elements.
2. Any two fields containing $p^d$ elements are isomorphic.

**Proof.** 1. The above theorem tells us that there is an irreducible polynomial $f(x) \in \mathbb{F}_p[x]$ satisfying $\deg(f) = d$. We let $K_f = \mathbb{F}_p[x]/f(x)\mathbb{F}_p[x]$ be the field described in Theorem 5.6.2 (5.27 in the textbook). That theorem tells us that $[K_f : \mathbb{F}_p] = \deg(f) = d$, and then Proposition 5.7.1 (5.28 in the textbook) says that

$$|K_f| = p^{[K_f:\mathbb{F}_p]} = p^d.$$

2. We will see...

□

# Chapter 6

# Groups: Part II

We now continue our study of group theory. POGGIES!!!

## §6.1 Normal Subgroups and Quotient Groups

Let $G$ be a group, and let $H$ be a subgroup. Recall that each $g \in G$ gives a left coset of $H$, defined by

$$gH = \{gh \mid h \in H\}.$$

We now define the collection of cosets of $H$.

**Definition 6.1.1: Set of Cosets**

Let $G$ be a group, with $H < G$ a subgroup. We denote the set of (left) cosets of $H$ in $G$ by

$$G/H = \{ \text{ (left) cosets of } H\}.$$

Let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ be distinct cosets of $H$. Proposition 2.39 tells us that $G$ is equal to the disjoint union

$$G = \mathcal{C}_1 \cup \ldots \cup \mathcal{C}_k$$

We use the notation $\mathcal{C}_i$ to emphasize that for a given coset $\mathcal{C}$, many different elements $g \in G$ will satisfy $\mathcal{C} = gH$. Indeed, (as one should check) if $\mathcal{C}$ is a coset of $H$, then

$$\mathcal{C} = gH \iff g \in \mathcal{C}.$$

Now, is there a way to turn this collection of cosets $\{\mathcal{C}_1, \ldots, \ldots, \mathcal{C}_k\}$ into a group? If so, how should we define the product of two cosets $\mathcal{C}_i$ and $\mathcal{C}_j$? One intuitive method may be to take a product of two elements, one from $\mathcal{C}_i$ and one from $\mathcal{C}_j$, and declare the resulting set $\mathcal{C}_i \cdot \mathcal{C}_j$ to be the coset of the product.

**Definition 6.1.2: Proposed Coset Multiplication**

Let $G$ be a group, and $H < G$ a subgroup. Given cosets $\mathcal{C}_i, \mathcal{C}_j$ of $H$, select one element from each ($g_1 \in \mathcal{C}_i$, $g_2 \in \mathcal{C}_j$), and define the resultant coset $g_1 g_2 H$.

Is this a good definition? Where might this go wrong?
Well, if we have three cosets of a group $G$, and we select one element from each of two cosets, there's a possibility that if we select different elements from each of the cosets, the first product might not be in the same coset as the second product. Concretely, consider the dihedral group of three vertices, $\mathcal{D}_3$. The subgroup $\{e, \phi_1\}$ (where $\phi_i$ are flips, and $\rho_j$ are rotations) has three cosets,

$$\mathcal{C}_1 = \{e, \phi_1\},\ \mathcal{C}_2 = \{\rho_1, \phi_2\},\ \mathcal{C}_3 = \{\rho_2, \phi_3\}.$$

If we select $\phi_1 \in \mathcal{C}_2,\ \phi_2 \in \mathcal{C}_3$, then we get $e \in \mathcal{C}_1$; but if we select $\phi_2 \in \mathcal{C}_2,\ \phi_3 \in \mathcal{C}_3$, then we get $\rho_2 \in \mathcal{C}_3$. The above definition thus gives us two different results: either $\mathcal{C}_2 \cdot \mathcal{C}_3 = \mathcal{C}_1$, or $\mathcal{C}_2 \cdot \mathcal{C}_3 = \mathcal{C}_3$!
However, the above definition *does* work if we select our subgroup $H$ more carefully; for instance, if we chose $H = \{e, \rho_1, \rho_2\}$, then the definition seems to work! So the definition works for some, but not all, subgroups. How do we distinguish the "good" subgroups? We want the subgroups to have the following property:

For all cosets $\mathcal{C}_1, \mathcal{C}_2$ of $H$, for all pairs of elements $g_1, g_1' \in \mathcal{C}_1,\ g_2, g_2' \in \mathcal{C}_2$, we wish for

$$g_1 g_2 H = g_1' g_2' H.$$

Let's see where this takes us. $g_1$ and $g_1'$ in the same coset $\mathcal{C}_1$ of $H$ means that there is some $h_1 \in H$ with $g_1' = g_1 h_1$ (since all elements of a coset of $H$ can be represented as $g_1 h_1$, where $g_1$ is in the coset, and $h_1$ is in the subgroup), and similarly $g_2' = g_2 h_2$ for some $h_2 \in H$. So, we want

$$g_1' = g_1 h_1 \text{ and } g_2' = g_2 h_2 \implies g_2^{-1} g_1^{-1} g_1' g_2' \in H$$

(since we want $g_1' g_2' = g_1 g_2 h_1 h_2$ in order for $g_1' g_2' \in g_1 g_2 H$; or equivalently, $g_2^{-1} g_1^{-1} g_1' g_2' \in H$). Substituting values of $g_1', g_2'$, we want

$$g_2^{-1} g_1^{-1} g_1 h_1 g_2 h_2 \in H \text{ for all } g_1, g_2 \in G,\ h_1, h_2 \in H.$$

With cancellation, this becomes

$$g_2^{-1} h_1 g_2 h_2 \in H \text{ for all } g_2 \in G,\ h_1, h_2 \in H.$$

But we know that $g_2 h_2 \in H$ if and only if $g_2 \in H$, so we end up with

$$g_2^{-1} h_1 g_2 \in H \text{ for all } g_2 \in G,\ h_1 \in H.$$

Dropping subscripts, we now get the following definition:

**Definition 6.1.3: Normal Subgroups**

Let $G$ be a group, let $H < G$ be a subgroup, and let $g \in G$. The **$g$-conjugate of $H$** is the subgroup

$$g^{-1} H g = \{g^{-1} h g \mid g \in G\}.$$

We say that $H$ is a **normal subgroup** if it satisfies

$$g^{-1} H g = H \text{ for every } g \in G.$$

**Example 47.** *If $G$ is an Abelian group, then every subgroup $H \subseteq G$ is normal, since*

$$g^{-1} h g = g^{-1} g h = h.$$

**Definition 6.1.4: Simple Groups**

Every group $G$ has two normal subgroups, $\{e\}$ and $G$. If these are the only two normal subgroups, then $G$ is called a **simple group**.

**Example 48.** *As shown above, the subgroup $H = \{e, \phi_1\}$ is not a normal subgroup, since for example*

$$\phi_2^{-1}\{e, \phi_1\}\phi_2 = \{\phi_2^{-1}e\phi_2, \phi_2^{-1}\phi_1\phi_2\} = \{e, \phi_3\}.$$

*However, the other subgroup $H = \{e, \rho_1, \rho_2\}$ is a normal subgroup. One could tediously check all conjugates, or realize that three rotations is a rotation, or a flip, rotation, flip is still a rotation. Similar results hold for $\mathcal{D}_n$.*

Here is an important source of normal subgroups. Indeed, we will later show that every normal subgroup of a group $G$ arises in this way.

**Proposition 6.1.1**

et $\phi : G \to G'$ be a homomorphism of groups. Then $\ker(\phi)$ is a normal subgroup of $G$.

**Proof.** We know from before that $\ker(\phi)$ is a subgroup of $G$. Let $h \in \ker(\phi)$, $g \in G$. Then

$$\begin{aligned}
\phi(g^{-1} \cdot h \cdot g) &= \phi(g^{-1}) \cdot \phi(h) \cdot \phi(g) \\
&= \phi(g)^{-1} \cdot \phi(h) \cdot \phi(g) \\
&= \phi(g)^{-1} \cdot \phi(g) && [\text{ since } h \in \ker(\phi), \text{ so } \phi(h) = e' \,] \\
&= e'.
\end{aligned}$$

Hence $g^{-1} \cdot h \cdot g \in \ker(\phi)$; since the choice of $g, h$ were arbitrary, $\ker(\phi)$ is thus a normal subgroup of $G$. □

Before turning the collection of cosets $G/H$ into a group, we now give three elementary properties of normal subgroups.

**Proposition 6.1.2: Properties of Normal Subgroups**

Let $G$ be a group, and $H < G$ a subgroup.

1. If $g^{-1}Hg \subseteq H$, then $H$ is a normal subgroup of $G$; in other words, it is enough only to check one inclusion.
2. For all $g \in G$, the conjugate set $g^{-1}Hg$ is a subgroup of $G$.
3. For all $g \in G$, the map $H \to g^{-1}Hg$ defined by $h \mapsto g^{-1}hg$ is a group isomorphism. In particular, if $H$ is finite, then $H$ and its conjugates have the same number of elements.

**Proof.** Left as an exercise. □

We now go back to our goal of turning the set of cosets $G/H$ into a group via the multiplication rule

$$g_1H \cdot g_2H = g_1g_2H.$$

Unfortunately, as we've seen before, choosing different elements $g_1' \in g_1H$, $g_2' \in g_2H$ may yield a different coset $g_1'g_2'H$. However, if $H$ is a *normal* subgroup of $G$, then we do actually get the same coset. Yay! Let us formally verify this.

**Lemma 6.1.1**

Let $G$ be a group, and let $H < G$ be a normal subgroup. Let $g_1, g_1', g_2, g_2' \in G$ be elements of $G$ satisfying

$$g_1H = g_1'H \text{ and } g_2H = g_2'H.$$

Then

$$g_1g_2H = g_1'g_2'H.$$

**Proof.** The assumption that $g_1H = g_1'H$ implies in particular that

$$g_1' \in g_1'H = g_1H, \text{ so } g_1' = g_1h_1 \text{ for some } h_1 \in H.$$

Similarly,

$$g_2' \in g_2'H = g_2H \implies g_2' = g_2h_2 \text{ for some } h_2 \in H.$$

We now wish to prove the inclusion $g_1'g_2'H \subseteq g_1g_2H$, so let $g_1'g_2'h$ be an element of $g_1'g_2'H$. We want

$$g_1'g_2'h \subseteq g_1g_2H.$$

From above, we see that

$$g_1'g_2'h = g_1h_1g_2h_2h.$$

If we could flip $h_2$ and $g_2$, we are done (since we are left with $g_1g_2$ times an element of $H$); but alas, $G$ is not necessarily commutative, so we must look elsewhere.
We want the leftmost $g_1$ to be followed by $g_2$; we can't just insert a $g_2$ there, but we can use the standard mathematical trick of putting the quantity where we want, and multiplying it to cancel it out: we do this by multiplying by $e = g_2g_2^{-1}$. So, we have

$$g_1h_1g_2h_2h = g_1g_2g_2^{-1}h_1g_2h_2h = (g_1g_2)(g_2^{-1}h_1g_2)h_2h.$$

But $g_2^{-1}h_1g_2 \in H$, since $H$ is a normal subgroup; and closure implies $h_2h \in H$. Thus, we have

$$g_1'g_2'h = g_1h_1g_2h_2h = g_1g_2(\text{ three elements in } H\ ) \in g_1g_2H,$$

and hence that

$$g_1'g_2'H \subseteq g_1g_2H.$$

Reversing the roles gives an analogous argument for $g_1g_2H \subseteq g_1'g_2'H$, and thus we get

$$g_1g_2H = g_1'g_2'H.$$

□

Essentially, this lemma demonstrates that the multiplication rule for cosets, $g_1H \cdot g_2H = g_1g_2H$, is well-defined if $H$ is a normal subgroup of $G$. We are now able to turn $G/H$ into a group!

**Theorem 6.1.1: First Isomorphism Theorem For Groups**

Let $G$ be a group, and let $H$ be a normal subgroup of $G$.

1. The collection of cosets $G/H$ is a group via the well-defined group operation
$$g_1H \cdot g_2H = g_1g_2H.$$
2. The map
$$\phi : G \longrightarrow G/H,\ \phi(g) = gH,$$
is a homomorphism whose kernel is $\ker(\phi) = H$.
3. Let
$$\psi : G \longrightarrow G'$$
be a homomorphism with the property that $H \subseteq \ker(\psi)$. Then there is a unique homomorphism
$$\lambda : G/H \longrightarrow G' \text{ satisfying } \lambda(gH) = \psi(g).$$
4. If we take $H = \ker(\psi)$ from above, then the homomorphism
$$\lambda : G/\ker(\psi) \longrightarrow G'$$
is injective. In particular, we get an isomorphism onto the image of $\lambda$,
$$\lambda : G/\ker(\psi) \longrightarrow \lambda(G) \subseteq G'.$$

**Proof.** 1. The above lemma proves that coset multiplication is well-defined (with a normal subgroup $H$). The other group properties follow directly from the corresponding properties of the group operation on $G$:
$$eH \cdot gH = gH \cdot eH = gH \tag{6.1}$$
$$gH \cdot g^{-1}H = g^{-1}H \cdot gH = eH \tag{6.2}$$
$$(g_1H \cdot g_2H) \cdot g_3H = g_1H \cdot (g_2H \cdot g_3H). \tag{6.3}$$
Thus $G/H$ is a group under coset multiplication.

2. We first check that $\phi$ is a homomorphism. Note that
$$\phi(g_1)\phi(g_2) = g_1H \cdot g_2H = g_1g_2H = \phi(g_1g_2).$$
Thus $\phi$ is a homomorphism. The kernel of $\phi$ is
$$\ker(\phi) = \{g \in G \mid \phi(g) = H\} = \{g \in G \mid gH = H\} = H.$$

3. We want to define a homomorphism $\lambda : G/H \to G'$ by the following algorithm:
   (a) Let $\mathcal{C} \in G/H$ be a coset of $H$.
   (b) Choose some $g \in \mathcal{C}$ with $\mathcal{C} = gH$.
   (c) Define $\lambda(\mathcal{C})$ to be $\psi(g) = g'$; that is, $\psi(g) = \lambda(gH) = g'$.

   However, step 2 may pose a problem, since many possible $g \in \mathcal{C}$ can work. So, we need to prove that if $gH = g'H$, then $\psi(g) = \psi(g')$.
   $gH = g'H$ means that $g' = gh$ for some $h \in H$. Then
$$\begin{aligned} \psi(g') &= \psi(gh) \\ &= \psi(g)\psi(h) \\ &= \psi(g)e' && [\text{ since } H \subseteq \ker(\psi) \text{ }] \\ &= \psi(g), \end{aligned}$$

as required. Thus our algorithm gives a well-defined map

$$\lambda : G/H \longrightarrow G'.$$

It's also easy to check that $\lambda$ is a homomorphism: for $g_1, g_2 \in G$,

$$\lambda(g_1 g_2 H) = \psi(g_1 g_2) = \psi(g_1)\psi(g_2) = \lambda(g_1 H)\lambda(g_2 H).$$

Finally, for any given homomorphism $\psi : G \to G'$, there's only one map $\lambda$ satisfying $\psi(g) = \lambda(gH)$, since this equality completely determines the values of $\lambda$ in terms of values of $\psi$; that is, for any $\lambda$ that maps $G/H$ to $G'$, every $gH \in G/H$ must be mapped to the same $\psi$.

4. Let $H = \ker(\psi)$. From (c) we get that

$$\lambda : G/H \longrightarrow G',\ \lambda(gH) = \psi(g);$$

we just need to show that $\lambda$ is injective.

Let $gH \in ker(\lambda)$. Then $\lambda(gH) = \psi(g) = e'$, so $g \in \ker(\psi) = H$. Therefore $gH = H$, which is the identity element of $G/H$. Thus, the kernel $\ker(\lambda)$ is trivial and consists of only the identity element, so $\lambda$ is injective. Surjectivity follows by definition of a function; any $\lambda$ surjects onto its image. Thus $\lambda : G/H \to G'$ is an is an isomorphism.

□

## §6.2 Groups Acting On Sets

Among the first groups studied were the symmetric groups $\mathcal{S}_n$, and the dihedral groups $\mathcal{D}_n$. The elements of $\mathcal{S}_n$ were permutations of $\{1, \ldots, n\}$, so one could view an element of $\mathcal{S}_n$ as giving a rule that takes a number in $\{1, \ldots, n\}$ and assigns it to another number in $\{1, \ldots, n\}$. Similarly, the elements of $\mathcal{D}_n$ are rigid re-arrangements of vertices of an $n$-gon, so we can also view elements of $\mathcal{D}_n$ as assigning a rule that takes the vertex of an $n$-gon to another vertex of the $n$-gon (with additional restrictions). We can axiomatize these examples into arbitrary groups, by starting with a group $G$ and a set $X$ and having each element in $G$ re-arrange elements in $X$ (with some restrictions).

**Definition 6.2.1: Group Actions on Sets**

Let $G$ be a group, and let $X$ be a set. An **action of $G$ on $X$** is a rule $\cdot$ that assigns to each element $g \in G$ and each element $x \in X$ another element $g \cdot x \in X$, so that the following two axioms hold:

- **Identity Axiom**: $e \cdot x = x$ for all $x \in X$.
- **Associative Axiom**: $(g_1 g_2) \cdot x = g_1 \cdot (g_2 \cdot x)$ for all $g_1, g_2 \in G$ and all $x \in X$.

**Remark 12.** *We can choose a fancier definition by showing that defining an action of $G$ on $X$ is the same as giving a group homomorphism*

$$\alpha : G \longrightarrow \mathcal{S}_X$$

*from the group $G$ to the symmetric group of $X$. Thus $\alpha$ sends each $g \in G$ to a permutation $\alpha(g) : X \to X$ of the set $X$, and the group action is $g \cdot x = \alpha(g)(x)$.*

**Definition 6.2.2: Orbits and Stabilizers**

Given a group $G$ acting on a set $X$, each element $x \in X$ determines two natural objects of interest:

- What are the elements of $X$ to which $x$ is sent by the action of $G$? This set is the **orbit of** $x$, denoted
$$Gx = \{g \cdot x \mid g \in G\}.$$
In other words, it is the all possible values that an $x \in X$ can take—its "range"—when acted upon by the group $G$.
- What are the elements of $G$ that leave $x$ unchanged? This set is the **stabilizer of** $x$, denoted
$$G_x = \{g \in G \mid g \cdot x = x\}.$$
In other words, it is all values in $G$ that preserve $x$; similar to the "kernel" of the action.

**Example 49.** *Consider the action of $G = \mathcal{S}_n$ on $X = \{1, \ldots, n\}$. Given any two elements $x, y \in X$, there is a permutation $\pi \in \mathcal{S}_n$ that sends $x$ to $y$ (e.g. $\pi(x) = y$, $\pi(y) = x$), and fixes all other elements of $X$. Thus, the orbit of $x$ is all of $X$, i.e. $Gx = X$, since there exists a $\pi \in G$ that sends $x$ to any other element in $X$. The stabilizer $G_x$ of $x$ consists of all permutations of $X$ that fix $x$, so they are the permutations of the remaining $n-1$ elements of $X$. Thus $G_x$ is isomorphic to $\mathcal{S}_{n-1}$.*

**Example 50.** *Consider the action of $G = \mathcal{D}_n$ on an $n$-gon whose vertices are labeled $X = \{1, \ldots, n\}$ clockwise around the $n$-gon. The orbit of any vertex consists of every vertex ($X$), since there's a rotation that takes any vertex to any other vertex. With stabilizers, it's not quite as simple. Clearly, non-trivial rotations do not fix $x$; same with most flips. However, the flip about the axis that goes through $x$ fixes $x$ (it also fixes the opposite vertex if $n$ is even) Thus $G_x$ consists of two elements, the identity and the flip about the axis through $x$, so $G_x$ is a cyclic group of order 2.*

**Example 51.** *Let $G$ be the subgroup of $\mathcal{S}_5$ generated by the permutation*
$$\pi = (134)(25).$$
*Clearly, $\pi$ has order 6, so $G = \{e, \pi, \pi^2, \pi^3, \pi^4, \pi^5\}$. Then*
$$G \cdot 1 = G \cdot 3 = G \cdot 4 = \{1, 3, 4\}, \text{ and } G \cdot 2 = G \cdot 5 = \{2, 5\}.$$
*Thus $G$ has two distinct orbits.*

**Proposition 6.2.1**

Let $G$ be a group that acts on a set $X$.

1. Let $x \in X$. The stabilizer $G_x$ is a subgroup of $G$.
2. Define a relation $\sim$ on $X$ by the following rule:

$$x \sim y \text{ if } y = gx \text{ for some } g \in G.$$

Then $\sim$ is an equivalence relation, and

$$(\text{ the equivalence class of } x\ ) = (\text{ the orbit } Gx \text{ of } x\ ).$$

3. Let $x \in X$. There is a well-defined bijection

$$\alpha : G/G_x \longrightarrow Gx$$

defined by the following algorithm:

- **Input**: A coset $\mathcal{C}$ of the subgroup $G_x$.
- **Computation**: Choose an element $g \in \mathcal{C}$.
- **Output**: $\alpha(\mathcal{C})$ is the element $g \cdot x$ in the orbit $Gx$.

In particular, if $G$ is finite, then

$$|Gx| = \frac{|G|}{|G_x|}.$$

**Proof.** 1. First, we have $e \in G_x$, since the definition of a group action requires $ex = x$. Next, let $g, g' \in G_x$. Then since $g, g'$ both fix $x$ together with the associative law of group actions tell us that

$$(gg')x = g(g'x) = gx = x,$$

so $gg' \in G_x$. Finally, applying $g^{-1}$ to both sides of $x = gx$ yields

$$g^{-1}x = g^{-1}(gx) = (g^{-1}g)x = ex = x,$$

so $g^{-1} \in G_x$. Thus $G_x$ is a subgroup of $G$.

2. First we have

$$x = ex, \text{ which shows that } x \sim x\ ,$$

and so $\sim$ is reflexive. Second, we note that

$$\begin{aligned} x \sim y &\implies y = gx \text{ for some } g \in G \\ &\implies g^{-1}y = (g^{-1}g)x = x, \end{aligned}$$

thus $x = g^{-1}y$, and so $y \sim x$, which shows that $\sim$ is symmetric. Finally,

$$\begin{aligned} x \sim y \text{ and } y \sim z &\implies y = gx \text{ and } z = g'y \text{ for some } g, g' \in G \\ &\implies z = g'(gx) = (g'g)x \\ &\implies x \sim z. \end{aligned}$$

Thus $\sim$ is an equivalence relation.

The equivalence class of an element $x \in X$ is

$$\{y \in X \mid x \sim y\} = \{y \in X \mid y = gx \text{ for some } g \in G\} = \{gx \mid g \in G\} = Gx.$$

3. We first must show that the output is well-defined, in that it doesn't depend on the choice of the group element in the coset $\mathcal{C}$. Suppose that $g' \in \mathcal{C}$ is some other element in the coset. This means that

$$g' \in \mathcal{C} = gG_x, \text{ so } g' = gh \text{ for some } h \in G_x.$$

It follows that

$$g'x = (gh)x = g(hx) = gx,$$

since $h \in G_x$ means $hx = x$. This shows that $\alpha(\mathcal{C}) = gx$ is well-defined, since it doesn't depend on choice of $g \in \mathcal{C}$ (explicitly, $\alpha(g'G_x) = g'x = gx = \alpha(gG_x)$ for $g, g' \in \mathcal{C}$, a coset of $G_x$). It's easy to see that $\alpha$ is surjective, since every element of $Gx$ looks like $gx$ for some $g \in G$, and so we can simply choose $\alpha(gG_x) = gx$.

Finally, we prove that $\alpha$ is injective, so suppose we have cosets $\mathcal{C}_1, \mathcal{C}_2$ such that

$$\alpha(\mathcal{C}_1) = \alpha(\mathcal{C}_2).$$

We wish to show that $\mathcal{C}_1 = \mathcal{C}_2$. Write the cosets as

$$\mathcal{C}_1 = g_1G_x \text{ and } \mathcal{C}_2 = g_2G_x \text{ for some } g_1, g_2 \in G.$$

Then

$$\begin{aligned}
\alpha(\mathcal{C}_1) = \alpha(\mathcal{C}_2) &\implies \alpha(g_1G_x) = \alpha(g_2G_x) \\
&\implies g_1x = g_2x \\
&\implies x = g_1^{-1}g_2x \\
&\implies g_1^{-1}g_2 \in G_x && [\text{ since } g_1^{-1}g_2 \text{ fixes x }] \\
&\implies g_1^{-1}g_2G_x = G_x \\
&\implies g_2G_x = g_1G_x \\
&\implies \mathcal{C}_1 = \mathcal{C}_2.
\end{aligned}$$

Thus $\alpha$ is a well-defined bijection.

For the last part, we compute

$$|Gx| = |(G/G_x)| = |G| \,/\, |G_x|\,,$$

where the first equality is because they are isomorphic, and the second is due to Lagrange's Theorem.

□

**Definition 6.2.3: Transitivity**

We say that $G$ **acts transitively on** $X$ if $Gx = X$ for all $x \in X$.

The dihedral group example demonstrates that $\mathcal{D}_n$ acts transitively on the vertices of an $n$-gon, while the following example about the cyclic subgroup generated by $\pi$ gives an example of a group and a set on which the group does not act transitively.

## §6.3 The Orbit-Stabilizer Counting Theorem

Let $G$ be a finite group that acts on a finite set $X$. We start with an important proof relating the sizes of $G$, $X$, and the various orbits and stabilizers. It will later prove crucial in our proof of Bylaw's theorem; and it has many other applications.

**Theorem 6.3.1: Orbit-Stabilizer Counting Theorem**

Let $G$ be a finite group that acts on a finite set $X$. Choose $x_1, \ldots, x_k \in X$ so that

$$Gx_1, \ldots, Gx_k$$

are the distinct orbits of elements of $X$. Then

$$|X| = \sum_{i=1}^{n} Gx_i = \sum_{i=1}^{n} \frac{|G|}{|G_{x_i}|}.$$

**Proof.** From above, we see that there is an equivalence relation on $X$ such that the equivalence class of $x$ is exactly its orbit $Gx$. Then, from general properties of equivalence classes, we have that $X$ is the disjoint union of the distinct equivalence classes:

$$X = Gx_1 \cup \ldots \cup Gx_k,$$

so in particular

$$|X| = |Gx_1| + \ldots + |Gx_k|.$$

This proves the first equality. We've also shown that for any orbit and stabilizer, $|Gx| = \frac{|G|}{|G_x|}$; this thus proves the second equality. □

**Example 52.** *Let $\mathcal{D}_6$ be the deferral group acting on a hexagon whose vertices are labeled $\{A, B, C, D, E, F\}$. Let $r \in \mathcal{D}_6$ be a $60°$ counterclockwise rotation, $f$ a flip on the axis between $A$ and $B$, and $f'$ a flip that fixes $C$ and $F$.*
*We consider the following six subgroups of $\mathcal{D}_6$, and their actions on the hexagon:*

$$\{e, r, r^2, r^3, r^4, r^5\},\ \{e, r^2, r^4\},\ \{e, r^3\},\ \{e, f\},\ \{e, f'\}, \{e, r^3, f, f'\}.$$

*For each, we compute their orbits and stabilizers:*

- $\{e, r, r^2, r^3, r^4, r^5\}$*:*
  - *Orbit:* $\{A, B, C, D, E, F\}$
  - *Stabilizer:* $\{e\}$
- $\{e, r^2, r^4\}$*:*
  - *Orbits:* $\{A, C, E\}$, $\{B, D, F\}$
  - *Stabilizers:* $\{e\}, \{e\}$
- $\{e, r^3\}$*:*
  - *Orbits:* $\{A, D\}, \{B, E\}, \{C, F\}$
  - *Stabilizers:* $\{e\}, \{e\}, \{e\}$

- $\{e, f\}$:
  - *Orbits:* $\{A, B\}, \{C, F\}, \{D, E\}$
  - *Stabilizers:* $\{e\}, \{e\}, \{e\}$
- $\{e, f'\}$:
  - *Orbits:* $\{C\}, \{F\}, \{A, E\}, \{B, D\}$
  - *Stabilizers:* $\{e, f'\}, \{e, f'\}, \{e\}, \{e\}$
- $\{e, r^3, f, f'\}$:
  - *Orbits:* $\{A, B, D, E\}, \{C, F\}$
  - *Stabilizers:* $\{e\}, \{e, f'\}$

*In every case, the orbits are disjoint and satisfy* $|G_x| \cdot |Gx| = |G|$*, as we've illustrated before.*
*We can also check that the Orbit-Stabilizer Counting Theorem holds; the most interesting case is for* $\{e, r^3, f, f'\}$*. The two orbits,*

$$G \cdot A = \{A, B, D, E\} \text{ and } G \cdot C = \{C, F\},$$

*have respective stabilizers* $G_A = \{e\}$ *and* $G_C = \{e, f'\}$*. Then*

$$|X| = |(G \cdot A)| + |(G \cdot C)| = 4 + 2 = 6,$$

*as required.*

The Orbit-Stabilizer Counting Theorem is incredibly powerful for studying groups and the sets on which they act. We illustrate this by proving a fundamental property of groups with prime power order, and as an application show that every group with order $p^2$ is Abelian.
First, we start with a definition.

**Definition 6.3.1: Center**

Let $G$ be a group. The **center** of $G$, denoted $Z(G)$, is the set of elements of $G$ that commute with every element of $G$,

$$Z(G) = \{g \in G \mid gg' = g'g \text{ for every } g' \in\}.$$

We leave it as an exercise to prove that the center of $G$ is a normal subgroup of $G$.

**Theorem 6.3.2**

Let $p$ be a prime, and let $G$ be a finite group with $p^n$ elements for some $n \geq 1$. Then $Z(G) \neq \{e\}$, that is, $G$ contains a non-identity element that commutes with every element of $G$.

**Proof.** Take the set $X$ to be a copy of $G$, and let $G$ act on $X$ with the formula

$$g \in G \text{ sends } x \in X \text{ to } gxg^{-1} \in X.$$

This is called the **conjugation action** of $G$ on itself (check that this rule satisfies the group action axioms!). Let $x \in X$. What is the stabilizer $x \in X$? It is the set

$$G_x = \{g \in G \mid gxg^{-1} = x\} = \{g \in G \mid gx = xg\}.$$

In other words, the stabilizer of $x$ is the set of elements in $G$ that commute with $x$. In particular,

$$Gx = G \iff x \text{ commutes with every element of } G \iff x \in Z(G).$$

Choose elements $x_1, \ldots, x_k \in X$ that give the distinct orbits. Then the Orbit-Stabilizer Counting Theorem tells us that

$$|X| = \sum_{i=1}^{k} \frac{|G|}{|G_{x_i}|}.$$

We have $|X| = |G| = p^n$, and each $G_{x_i}$ is a subgroup of $G$, so Lagrange's Theorem tells us that each $|G_{x_i}|$ divides $|G|$. Thus, for each $i$ we have

$$|G_{x_i}| = p^{r_i},\ 0 \leq r_i \leq n,$$

so the summation above becomes

$$|X| = p^n = \sum_{i=1}^{k} \frac{p^n}{p^{r_i}} = \sum_{i=1}^{k} p^{n-r_i}.$$

Terms with $r_i = n$ are especially interesting, since

$$r_i = n \iff x_i \text{ commutes with every element of } G \iff x_i \in G.$$

So, if we separate out the summation above, we get the following formula:

$$p^n = \underbrace{\sum_{i=0}^{k} 1}_{\text{with } r_i = n} + \underbrace{\sum_{i=0}^{k} p^{n-r_i}}_{\text{with } r_i < n} = |Z(G)| + \sum_{i=0}^{k} p^{n-r_i}.$$

But we know that $|Z(G)| \geq 1$, since the identity element of $G$ is in $Z(G)$. Hence $|Z(G)| \geq p$ (indeed, the number of elements in $Z(G)$ is a power of $p$!). □

Using this, we can prove a surprising theorem about groups having $p^2$ elements.

**Corollary 6.3.1**

Let $p$ be prime, and let $G$ be a group with $p^2$ elements. Then $G$ is an Abelian group.

**Proof.** For notational purposes, let

$$Z = Z(G).$$

Lagrange tells us that the order of the center $Z$ divides $|G| = p^2$, so

$$|Z| = 1,\ p,\ \text{or } p^2.$$

From above, we see that $Z \neq \{e\}$, so $|Z| \neq 1$.
Now, assume $|Z| = p$. Since the center $Z$ is a normal subgroup of $G$, we can thus take the quotient group $G/Z$, and Lagrange tells us that

$$|G/Z| = \frac{|G|}{|Z|} = \frac{p^2}{p} = p.$$

Thus $G/Z$ has prime order, and so is cyclic. Let $hZ$ be a coset that generates $G/Z$,

$$G/Z = \{Z, hZ, h^2Z, \ldots, h^{p-1}Z\}.$$

In particular, this implies that

$$G = Z \cup hZ \cup \ldots \cup h^{p-1}Z,$$

since every element of $G$ is in some coset of $Z$.
Now, let $g_1, g_2 \in G$ be arbitrary elements. They are each in some coset of $Z$, so

$$g_1 = h^{i_1}z_1 \text{ and } g_2 = h^{i_2}z_2 \text{ for some } z_1, z_2 \in Z \text{ and } 0 \leq i_1, i_2 \leq p-1.$$

Since $z_1, z_2 \in Z$, they commute with every element of $G$; thus

$$\begin{aligned} g_1g_2 = (h^{i_1}z_1)(h^{i_2}z_2) = (h^{i_1}h^{i_2})(z_1z_2) = h^{i_1+i_2}z_2z_1 \\ = (h^{i_2}h^{i_1})(z_2z_1) = (h^{i_2}z_2)(h^{i_1}z_1) = g_2g_1. \end{aligned}$$

But then this shows that every element commutes with every other element; that is, $Z = G$; but we assumed that $|Z| = p \neq p^2 = |G|$, a contradiction.
Hence the order of $Z$ cannot be either 1 or $p$, and so must be $p^2$. In other words, $Z = G$—every element commutes with every other element—and so $G$ is Abelian. □

**Remark 13.** ***Some notes on Commutativity****: Commutativity is King! While that doesn't mean we should only study Abelian groups, whenever we're given a group $G$, we should exploit whatever commutative Ty we can find. For example, given an element $h \in G$, it's worthwhile to inspect elements of $G$ that commute with $h$. These elements satisfy*

$$gh = hg, \text{ or equivalently }, g^{-1}hg = h.$$

*More generally, we could take a subgroup $H \subset G$ and look at the elements $g \in G$ that commute with every element in $H$.*
*Alternatively, we can take a weaker commutativity condition: instead of insisting that $g^{-1}hg = h$ for every $h \in H$, we could require only that $g^{-1}hg \in H$, not that it equals $h$.*

The above remark provides the motivation for two important definitions.

**Definition 6.3.2: Centralizer**

Let $G$ be a group, and let $H \subseteq G$ be a subgroup of $G$. The **centralizer of** $H$, denoted $Z(H)$ or $Z_G(H)$, is the set of elements of $G$ that commute with every element of $H$,

$$Z_G(H) = \{g \in G \mid gh = hg \text{ for every } h \in H\}.$$

For $h \in G$, we write $Z_G(h)$ for the centralizer of the cyclic subgroup $\langle h \rangle$ generated

by $h$.

**Definition 6.3.3: Normalizer**

Let $G$ be a group, and $H \subseteq G$ a subgroup of $G$. We define the **normalizer of** $H$ to be

$$N_G(H) = \{g \in G \mid g^{-1}Hg = H\}.$$

For example, the definition of normal subgroup says that

$$N_G(H) = G \iff H \text{ is a normal subgroup of } G.$$

Verifying that $Z_G(H)$ and $N_G(H)$ are subgroups is left as an exercise.

## §6.4 Sylow's Theorem, Part I

We just proved that groups of $p$-power order have a non-trivial center. Now, we prove part of a fundamental theorem regarding subgroups of $p$-power order, Sylow's Theorem. Sylow's Theorem is incredibly important in the study of finite groups.
Let $G$ be a group, and $H \subseteq G$ a subgroup. Recall by Lagrange that the order of $H$ divides the order of $G$. It would be nice if the converse were true; that is, if $m$ divides $|G|$, then there exists a subgroup $H \subseteq G$ of order $m$. Unfortunately, this is not true in general; however, it turns out that this holds if $m$ is of prime power.

**Theorem 6.4.1: Sylow's Theorem: Part I**

Let $G$ be a finite group, let $p$ be prime, and let $p^n$ be the largest power of $p$ that divides $|G|$. Then $G$ has a subgroup of order $p^n$.

**Proof.** We're given that $p^n$ is the largest power of $p$ that divides $|G|$, which means we can factor

$$|G| = p^n m \text{ with } p \nmid m.$$

We proceed with induction on $m$. If $m = 1$, then $|G| = p^n$, so $G$ itself is the desired subgroup; so suppose $m \geq 2$, and that the theorem holds for groups of order $p^n m'$ with $m' < m$.
How might we create a subgroup of $G$ with $p^n$ elements? Suppose we take a subset $A \subseteq G$ with $p^n$ elements. How might we tell that $A$ is a subgroup? We first observe that

$$A \text{ is a subgroup} \iff aA = A \text{ for every } a \in A$$

(verify this!). This suggests that we look at the collection of $p^n$-element subsets of $G$ and let $G$ act on these subsets by left multiplication.
In other words, we start with the set

$$S = \{A \subseteq G \mid A \text{ is a subset of } G \text{ with } |A| = p^n\}.$$

Note the structure of $S$; $S$ is a **set of sets**. For example, given a group $G = \{g_1, g_2, g_3, g_4\}$ and $p^n = 2$, then

$$S = \{\{g_1, g_2\}, \{g_1, g_3\}, \{g_1, g_4\}, \{g_2, g_3\}, \{g_2, g_4\}, \{g_3, g_4\}\}.$$

How many subsets of $p^n$ are contained in $G$? Since we want to choose $p^n$ elements out of $p^n m$ total elements, the number of elements in $S$ is given by

$$|S| = \binom{p^n m}{p^n}.$$

Now, let $G$ act on $S$ by left-multiplication. Equivalently, for $g \in G$ and $A \in S$, we define

$$gA = \{g \cdot a \mid a \in A\} \in S$$

(since the resulting coset $gA$ also has $p^n$ elements, so must be in $S$; additionally, $gA$ is still a subset of $G$). Choose elements $A_1, \ldots, A_r$ of $S$ so that $GA_1, \ldots, GA_r$ form distinct orbits. With the Orbit-Stabilizer Counting Theorem, we thus have

$$|S| = \sum_{i=1}^{r} \frac{|G|}{|G_{A_i}|},$$

where $G_{A_i} = (\text{the stabilizer of } A_i) = \{g \in G \mid gA_i = A_i\}$. Substituting the sizes of $S$ and $G$, we then get

$$\binom{p^n m}{p^n} = \sum_{i=1}^{r} \frac{p^n m}{|G_{A_i}|}.$$

We'll now use a fact (proven later) that

$$\binom{p^n m}{p^n} \text{ is not divisible by } p.$$

It follows that the sum $\sum_{i=1}^{r} \frac{p^n m}{|G_{A_i}|}$ is not divisible by $p$, and since each $\frac{p^n m}{|G_{A_i}|}$ is an integer, it must be the case that at least one of these integers is not divisible by $p$ (why? Because otherwise, every $\frac{p^n m}{|G_{A_i}|}$ would still have a $p$ term, and thus the entire sum would still be divisible by $p$). Hence there exists at least one $A_j \in S$ with the property that $|G_{A_i}|$ is a multiple of $p^n$, since $|G_{A_i}|$ needs to cancel out the entire $p^n$ in the numerator.

In other words, we have shown that there exists a subset $A \subseteq G$ with $|A| = p^n$ (by construction of $A \in S$) such that the stabilizer $G_A$ of $A$ has order

$$|G_A| = p^n m' \text{ with } m' | m.$$

We consider two cases. If $m' < m$, then the induction hypothesis says that $G_A$ has a subgroup $H$ with order $p^n$. But $H$ is thus also a subgroup of $G$, so we're done.
So, suppose $m' = m$. This assumption means that $G_A$ has the same number of elements as $G$, so $G_A = G$. By definition of a stabilizer $G_A$, this means that

$$gA = A \text{ for every element } g \in G.$$

The set $A$ has $p^n$ elements, so it must be non-empty. Let $a \in A$ be some element of $A$. Then for every $g \in G$, we have

$$g = (ga^{-1})a \in (ga^{-1})A = A$$

(we use $ga^{-1}$ instead of just $g$). This shows that every element of $G$ is in $A$, so $G = A$, so $|G| = |A| = p^n$. But $G$ has order $p^n m$ with $m \geq 2$, a contradiction.
Therefore $m' < m$, and so $G$ has a subgroup of order $p^n$. □

We now prove a lemma about the divisibility of the binomial coefficient, that proved essential in Sylow's Theorem.

**Lemma 6.4.1**

Let $p$ be a prime, let $n \geq 0$, and let $m \geq 1$ with $p \nmid m$. Then the binomial coefficient $\binom{p^n m}{p^n}$ is not divisible by $p$.

**Proof.** The binomial coefficient is equal to

$$\binom{p^n m}{p^n} = \frac{(p^n m)!}{p^n!(p^n m - p^n)!} = \frac{p^n m(p^n m - 1)\dots(p^n m - p^n + 1)}{p^n(p^n - 1)\dots \cdot 3 \cdot 2 \cdot 1},$$

so it is equal to

$$\binom{p^n m}{p^n} = \prod_{r=0}^{p^n - 1} \frac{p^n m - r}{p^n - r}.$$

We claim that after simplifying any of the fractions $\frac{p^n m - r}{p^n - r}$, there are no $p$ factors left in the numerator. With $r = 0$, this is clear: since we're given that $p \nmid m$, and $\frac{p^n m}{p^n} = m$, that fraction is kosher.
For the others, take any $r$ between 1 and $p^n - 1$, and factor it into

$$r = p^i s \text{ with } 0 \leq i < n \text{ and } p \nmid s;$$

we can do this since if no power of $p$ divides $r$, then clearly for $r = s$, $p \nmid s$, and we simply choose $i = 0$. Otherwise, we divide $r$ by the highest power of $p$ possible, and thus the remaining $\frac{r}{p^i} = s$ is not divisible by $p$.
Then

$$\frac{p^n m - r}{p^n - r} = \frac{p^n m - p^i s}{p^n - p^i s} = \frac{p^{n-i} m - s}{p^{n-i} - s}.$$

Since $i < n$ and $p \nmid s$, we see that neither $p^{n-i} m - s$ nor $p^{n-i} - s$ is divisible by $p$. Thus, we can cancel out all of the $p$-factors in the numerator by all of the $p$-factors in the denominator, so for any $1 \leq r \leq p^n - 1$, the fraction $\frac{p^n m - r}{p^n - r}$ has no $p$-factors. Therefore, the product

$$\binom{p^n m}{p^n} = \prod_{r=0}^{p^n - 1} \frac{p^n m - r}{p^n - r}$$

has no factors of $p$, and thus is not divisible by $p$. □

Using Sylow's Theorem, we now classify subgroups further:

**Definition 6.4.1: Sylow Subgroups**

Let $G$ be a finite group, let $p$ be a prime, and let $p^n$ be the largest power of $p$ that divides $|G|$. A subgroup $H \subseteq G$ with $|H| = p^n$ is called a **$p$-Sylow subgroup of $G$**. Part I of Sylow's Theorem tells us that $G$ has at least one Sylow subgroup.

**Remark 14.** *We will see later that it is true, more generally, that if $p^r$ divides $|G|$, then $G$ has a subgroup of size $p^r$, even if $p^r$ is not the largest power of $p$ dividing $|G|$. [TODO: import exercise solution here]*

**Remark 15.** *A useful observation is that if $p$ and $q$ are two distinct primes dividing $|G|$, and if we take a $p$-Sylow subgroup $H_p$ and a $q$-Sylow subgroup $H_q$, then $H_p \cap H_q = \{e\}$. This actually follows immediately from Lagrange's Theorem, since if $H$ and $H'$ are any two subgroups of $G$, then $H \cap H'$ is a subgroup of both $H$ and $H'$, so Lagrange tells us that*

$$|H \cap H'| \text{ divides } \gcd(|H|, |H'|).$$

*However, in our situation $gcd(|H|, |H'|) = \gcd(p^n, q^m) = 1$, so $|H_p \cap H_q| = 1$.*

**Remark 16.** *In general, if $p_1, \ldots, p_r$ are the distinct primes that divide the order of $G$, then Sylow's Theorem tells us that the $p$-Sylow subgroups of $G$ are large; indeed, they satisfy*

$$|G| = |H_{p_1}| \cdot \ldots \cdot |H_{p_r}|.$$

*This suggests that one way to understand a finite group $G$ is to start with its various $p$-Sylow subgroups, and then see how they might fit together. However, if you start with Sylow subgroups, there may be multiple ways to combine them. For example, the cyclic subgroups $\mathcal{C}_2$ and $\mathcal{C}_3$ are the Sylow subgroups of both $\mathcal{C}_6$ and $\mathcal{S}_3$.*

We now provide the complete statement of Sylow's Theorem; we postpone the proof until later, when we've explored some additional results that will aid in our proof.

**Theorem 6.4.2: Sylow's Theorem**

Let $G$ be a finite group, and $p$ prime.

1. $G$ has at least one $p$-Sylow subgroup, i.e. there is a subgroup $H \subseteq G$ with $|H| = p^n$, where $p^n$ is the largest power of $p$ that divides $|G|$.
2. Let $H_1$ and $H_2$ be $p$-Sylow subgroups of $G$. Then $H_1$ and $H_2$ are conjugate, i.e. for some $g \in G$, $H_2 = g^{-1} H_1 g$.
3. Let $H$ be a $p$-Sylow subgroup of $G$, and let $k$ be the number of distinct $p$-Sylow subgroups of $G$. Then

$$k \,|\, |G| \text{ and } k \equiv 1 \mod p.$$

(In fact, the divisibility follows a more precise formula: $k\,|N_G(H)| = |G|$.)

Let's see how Sylow's Theorem can be used to analyze or classify groups based on their order. In particular, we can exploit the fact that the number of $p$-Sylow subgroups is simultaneously a divisor of $|G|$ and is congruent to 1 modulo $p$. This is quite a strict condition.

**Example 53.** *Let $G$ be a group of order 10, so $G$ has 2-Sylow subgroups and 5-Sylow subgroups. Let's look first at the case of 5-Sylow subgroups of $G$. Let $k$ be the number*

*of distinct* 5*-Sylow subgroups. Sylow's Theorem tells us that*

$$k|10 \text{ and } k \equiv 1 \mod 5.$$

*This forces* $k = 1$*; in other words,* $G$ *has a unique* 5*-Sylow subgroup* $H_5$*. It follows that* $H_5$ *is normal, since for any* $g \in G$*, the conjugate* $g^{-1}H_5g$ *is also a subgroup (see 6.1.2) of order* 5*, so equals* $H_5$*.*
*Next, we use the fact that there is also at least one* 2*-Sylow subgroup, say* $H_2$*. As noted before,* $H_2 \cap H_5 = \{e\}$*, so we can write*

$$H_2 = \{a, e\} \text{ and } H_5 = \{e, b, b^2, b^3, b^4\}$$

*(since prime order implies cyclic), where the only common element is* $e$*.*
*Let's look at the extent to which* $a$ *and* $b$ *commute:* $aba^{-1}$*. We have*

$$aba^{-1} \in aH_5a^{-1} = H_5, \text{ since } H_5 \text{ is normal .}$$

*Hence* $aba^{-1} = b_j$ *for some* $0 \leq j \leq 4$*.*
*In order to determine the value of* $j$*, we compute*

$$\begin{aligned} aba^{-1} = b^j \implies b &= a^{-1}b^ja \\ &= (a^{-1}ba)^j && [\text{since we can cancel out most of } a^{-1}a = e \text{ products}] \\ &= \left(a^{-1}(a^{-1}b^ja)a\right)^j \\ &= (a^{-2}b^ja^2)^j \\ &= a^{-2}b^{j^2}a^2 && [\text{again, we have a lot of } a^{-2}a^2 = e \text{ cancellation}] \\ &= b^{j^2} && [\text{since } a \text{ has order } 2, \text{ so } a^2 = e]. \end{aligned}$$

*In other words,* $b = b^{j^2}$*, so* $b^{j^2-1} = e$*. Since* $b$ *has order* 5*, this means*

$$j^2 - 1 \equiv 0 \mod 5 \iff j^2 \equiv 1 \mod 5.$$

*Inspection then reveals that either* $j = 1$ *or* $j = 4$*.*
*Suppose first that* $j = 1$*, and recall that* $aba^{-1} = b^j$*. This means that* $ab = ba$*, and since every element of* $G$ *is a power of* $a$ *times a power of* $b$*, this implies that* $G$ *is Abelian (to see why this is true, recall that* $a \neq b^i$ *for any power of* $b$*; that means* $ab^i \notin H_5$*; thus the two cosets of* $H_5$*, which are composed of* $b$ *to some power times* $a$ *to some power, is the disjoint union of* $G$*)! It is then easy to check that* $ab$ *has order* 10*:*

$$\begin{aligned} e = (ab)^k = a^kb^k &\implies a^k = b^{-k} \in H_2 \cap H_5 = \{e\} \\ &\implies a^k = b^k = e \\ &\implies 2|k \text{ and } 5|k \\ &\implies 10|k. \end{aligned}$$

*Hence if* $j = 1$*, then* $G$ *is a cyclic group of order* 10 *(we get the first inclusion statement since* $b^{-k} \in H_5$*, and if* $b^{-k} = a^k \in H_2$*, then* $b^{-k} \in H_2$ *as well).*
*For the other case of* $j = 4$*, that means that* $ab = b^4a$*. But* $b^5 = e$*, so* $b^4 = b^{-1}$*, which means that the group* $G$ *consists of the* 10 *elements* $a^ib^j$ *with* $0 \leq i \leq 1$*,* $0 \leq j \leq 4$ *(see above for why), and the group law is determined by the rules*

$$a^2 = e, b^5 = e, ba = ab^{-1}.$$

*Thus* $G$ *is simply the dihedral group* $\mathcal{D}_5$ *of order* 10*, i.e.* $G$ *is isomorphic to the symmetry group of a regular pentagon.*
*In summary, there are only two groups of order* 10*; specifically, if a group has order* 10*, then it is either isomorphic to the cyclic group* $\mathcal{C}_{10}$ *or the dihedral group* $\mathcal{D}_5$*.*

**Example 54.** *Let $G$ be a finite group of order $|G| = pq$, where $p$ and $q$ are distinct primes; say without loss of generality that $p > q$. Let $H$ be a $p$-Sylow subgroup of $G$. We claim that $H$ is a normal subgroup of $G$. We look at the normalizer $N_G(H)$ of $H$. It certainly contains $H$, which has $p$ elements, so $p|N_G(H)$ (by Lagrange). But $N_G(H)$ is a subgroup of $G$, so $|N_G(H)|$ divides $|G| = pq$. Thus $|N_G(H)|$ is an integer divisible by $p$ that divides $pq$, so either*

$$|N_G(H)| = p (so\ N_G(H) = H),\ or\ |N_G(H)| = pq (so\ N_G(H) = G).$$

*If $N_G(H) = G$, then $H$ is a normal subgroup, and we are done (by definition of normalizer). On the other hand, if $N_G(H) = H$, then from Sylow's Theorem (c), we see that*

$$1 \equiv \frac{|G|}{|N_G(H)|} = \frac{pq}{q} = q \mod p,$$

*so $p|q - 1$. But this contradicts the assertion that $p > q$, so this is not possible. Hence $H$ is a normal subgroup of $G$.*

# Chapter 7

# Rings: Part II

**Important**: in this chapter, *ring* means *commutative ring*.

## §7.1 Irreducible Elements and Unique Factorization Domains

You're likely familiar with the famous theorem that every non-zero integer can be factored uniquely as a product of primed, the **Fundamental Theorem of Arithmetic**. We now hope to extend this concept beyond the integers and into abstract rings, and to explore what unique factorization would mean for other rings; indeed, some rings have it, and others don't.

From the Fundamental Theorem of Arithmetic, in the ring of integers $\mathbb{Z}$, the primes are the basic building blocks, where an integer $p \in \mathbb{Z}$ is prime if it does not factor. But that's a lie, since we can always "factor" $p$; for example, we have

$$p = 1 \cdot p \text{ and } p = (-1) \cdot (-p).$$

Your first response is, "that's cheating", but how exactly is that cheating? If you say to ignore $\pm 1$, that's correct, but why exactly $\pm 1$, and not anything else?

Let's look at the analogous problem for the polynomial ring $F[x]$, where $F$ is a field. In Chapter 5, we saw the concept of irreducible polynomials, those $f(x) \in F[x]$ that "do not factor" in $F[x]$. But again, that's not exactly true, since for any non-zero constant $c \in F$, we can factor $f(x)$ into

$$f(x) = c^{-1} \cdot cf(x).$$

So when defining non-trivial factorizations in $F[x]$, we need to ignore all non-zero elements in $F$.

What do $\pm 1$ in $\mathbb{Z}$ and the non-zero constants in $F[x]$ have in common? **They're units in their respective rings**. Recall that

**Definition 7.1.1: Units**

Let $R$ be a ring. An element $a \in R$ is a **unit** if it has a multiplicative inverse, i.e. if there is an element $b \in R$ satisfying $ab = 1$. The set of units of $R$ is denoted $R^*$, which is indeed a group with multiplication.

In general, if $u \in R^*$ is a unit, then we can factor any element $a \in R$ as

$$a = u^{-1} \cdot u \cdot a.$$

We want to ignore these trivial factorizations; thus, we have our first important definition.

**Definition 7.1.2: Irreducible**

Let $R$ be a ring. A non-zero element $a \in R$ is **irreducible** if $a$ is not a unit and the only way to factor $a = bc$ is if either $b$ or $c$ is a unit.

As with the unique factorization of integers, it's necessary to carefully define what precisely uniqueness means. For example, ignoring even the trivial factorizations, the integer $60 \in \mathbb{Z}$ has "multiple different factorizations":

$$60 = 2 \cdot 2 \cdot 3 \cdot 5 = 2 \cdot 3 \cdot 5 \cdot 2 = \dots.$$

But these factorizations are intrinsically the same; only their orders have changed.
In order to discuss unique factorization in general rings, we thus need to account for potential ambiguities arising from re-ordering the factors and/or including extra unit factors.

**Definition 7.1.3: Unique Factorization Domain**

Let $R$ be an integral domain, i.e. a commutative ring with no zero divisors. Then $R$ is a **unique factorization domain** (UFD) if:

1. Let $a \in R$ be a non-zero element that is not a unit. Then $a$ can be written in the from
$$a = b_1 \cdot b_2 \cdot \dots \cdot b_n$$
using irreducible elements $b_1, \dots, b_n \in R$.

2. Suppose that $b_1, \dots, b_n \in R$, $c_1, \dots, c_m \in R$ are irreducible elements of $R$, and suppose further that
$$b_1 \cdot b_2 \cdot \dots \cdot b_n = c_1 \cdot c_2 \cdot \dots \cdot c_m.$$
Then $m = n$, and after rearranging $c_1, \dots, c_n$, there are units $u_1, \dots, u_n \in R^*$ so that
$$c_1 = u_1 b_1, \dots, c_n = u_n b_n.$$

Later, we'll show that $\mathbb{Z}$, $\mathbb{Z}[i]$, and $F[x]$ are UFDs. This proof will first show that every ideal in these rings are principal, then showing that in general, rings with this property are UFDs.
Many rings are not UFDs. For example, the ring

$$\mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} \in \mathbb{C} \mid a, b \in \mathbb{Z}\}$$

is not a UFD (TODO: prove this).
We finish this section with an important class of rings that are UFDs, even though they contain many non-principal ideals.

**Theorem 7.1.1**

Let $F$ be a field. Then the ring $F[x_1, \dots, x_n]$ of polynomials in $n$ variables with coefficients in $F$ is a unique factorization domain.

**Proof.** The proof is unfortunately beyond the scope of this chapter. The main key idea is proving that if $R$ is a UFD, then $R[x]$ is a UFD; then, we proceed by induction. □

## §7.2 Euclidean Domains and Principal Ideal Domains

We've seen before the importance of studying ideals of $R$. The simplest ideals are principal ideals; that is, ideals formed by taking all multiples of a $c \in R$:

$$cR = (c) = \{rc \mid r \in R\}.$$

Some rings have the nice property that all ideals are principal.

**Definition 7.2.1: Principal Ideal Domains**

A ring $R$ is a **principal ideal domain** (PID) if it is an integral domain in which every ideal of $R$ is principal.

We already know two examples of PIDs. The first is the ring of integers $\mathbb{Z}$ [TODO: insert proof]. The second is the ring of polynomials $F[x]$ with coefficients in a field $F$; see Theorem 5.21. In both cases, the key to proving the PID property is division-with-remainder, as described for $\mathbb{Z}$ in the prelude and $F[x]$ in Proposition 5.20. This motivates the next definition, which generalizes the idea of "division-with-remainder". The idea is that we want to "divide" $a$ by $b$ to get a quotient $q$ and a remainder $r$, where $r$ is "smaller" than $b$. The subtle part here is to define what exactly it means for an element of a ring to be "smaller" than another, and what properties "smallness" should obey.

**Definition 7.2.2: Euclidean Domains**

A ring $R$ is a **Euclidean domain** if it is an integral domain and if there is a **size function**

$$\sigma : R \longrightarrow \{0, 1, 2, 3, \ldots\}$$

with the following properties:

1. $\sigma(a) = 0$ if and only if $a = 0$;
2. For all $a, b \in R$ with $b \neq 0$, there exist elements $q, r \in R$ that satisfy
$$a = bq + r \text{ and } \sigma(r) < \sigma(b).$$
3. For all $a, b \in R$, we have $\sigma(ab) = \sigma(a)\sigma(b)$[a].

[a] Sometimes, only the weaker property $\sigma(a) \leq \sigma(a)\sigma(b)$ for all non-zero $a, b$ is required.

We now prove that every Euclidean domain is a principal ideal domain.

**Theorem 7.2.1**

Every Euclidean domain is a PID.

**Proof.** Let $R$ be a Euclidean domain with size function $\sigma$, and let $I$ be an ideal of $R$. If $I = (0)$, it is clearly principal, so consider $I \neq (0)$. We consider the following set of non-negative integers:

$$\{\sigma(c) \mid c \in I,\ c \neq 0\}.$$

This set is non-empty, since $I \neq (0)$. Therefore this set has a smallest element, say $\sigma(b)$. In other words, there is a non-zero element $b \in I$ satisfying

$$\sigma(b) \leq \sigma(c) \text{ for all } c \in I.$$

We claim that $I = (b)$. Let $a \in I$ be any element. Since $R$ is a Euclidean domain, we can find elements $q, r \in R$ such that

$$a = bq + r,\ \sigma(r) < \sigma(b).$$

Since $a, b \in I$ and $I$ is an ideal, we have

$$bq \in I \text{ and hence } r = a - bq \in I.$$

Thus $r \in I$ and $\sigma(r) < \sigma(b)$. Since $\sigma(b) \leq \sigma(c)$ for any non-zero element, we must have that $r = 0$. In other words, $a = bq$, so $a \in (b)$. Since we've proved that every $a \in I$ is in $(b)$, we have $I \subseteq (b)$. By the definition of an ideal, $(b) \subseteq I$, so we have $I = (b)$. Therefore every ideal in a Euclidean domain is principal. □

The ring $\mathbb{Z}$ is a Euclidean domain using the function $\sigma(b) = |b|$. The formula $a = bq + r$ just says that dividing $a$ by $b$ gives a quotient $q$ and remainder $r$ with $|r| < |b|$ (see Division Algorithm for integers). Thus $\mathbb{Z}$ is a PID.
Let $F$ be a field. The ring $F[x]$ of polynomials is a Euclidean domain using the size function

$$\sigma(p(x)) = \begin{cases} 2^{\deg p(x)} & \text{if } p(x) \neq 0, \\ 0 & \text{if } p(x) = 0 \end{cases}$$

Indeed, $\sigma$ satisfies the three properties of a size function:

- Clearly, $\sigma(p(x)) = 0$ if and only if $p(x) = 0$.
- Proposition 5.20 gives us the division algorithm for polynomials; that is, for polynomials $f(x), g(x) \in F[x]$, there exist $q(x), r(x) \in F[x]$ satisfying

$$f(x) = g(x)q(x) + r(x),\ \deg(r) < \deg(g).$$

- The multiplication property is immediate from the fact that the degree of a product of polynomials is the sum of their degrees, so

$$\sigma(p_1 p_2) = 2^{\deg p_1 + \deg p_2} = 2^{\deg p_1} \cdot 2^{\deg p_2} = \sigma(p_1)\sigma(p_2).$$

Thus $F[x]$ and $\mathbb{Z}$ are PIDs. Another interesting example is the Gaussian integers.

**Proposition 7.2.1**

The ring of Gaussian integers $\mathbb{Z}[i]$ is a Euclidean domain using the size function

$$\sigma(a + bi) = a^2 + b^2.$$

**Proof.** The ring $\mathbb{Z}[i]$ is a subring of the field $\mathbb{C}$ of complex numbers, and fields are always integral domains; thus $\mathbb{Z}[i]$ is automatically an integral domain. The hard part is proving that $\mathbb{Z}[i]$ has the Euclidean property using the size function $\sigma$. We'll leave it to you to check that

- $\sigma(\alpha) = 0$ if and only if $\alpha = 0$;

- $\sigma(\alpha\beta) = \sigma(\alpha)\sigma(\beta)$

We view the field of complex numbers $\mathbb{C}$ as forming the $xy$-plane, with the point $(x, y)$ corresponding to the complex number $x + yi$. Then the ring $\mathbb{Z}[i]$ consists of the points $(a, b)$ with integer coordinates, and the size $\sigma(a + bi)$ is the square of the distance from $(0, 0)$ to $(a, b)$.

We now proceed with a geometric argument. Let $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$. Our task is to divide $\alpha$ by $\beta$ to get a quotient $\gamma$ and remainder $\rho$ with $\sigma(\rho) < \sigma(\beta)$. The idea is to take the quotient $\alpha/\beta$ and round the coefficients to the nearest integer. $\alpha/\beta$ is a complex number, so we plop it into $\mathbb{C}$ and find its closest element in $\mathbb{Z}[i]$, as illustrated below. The star represents the quotient $\alpha/\beta$, while points in $\mathbb{Z}[i]$ are marked with dots. We then picked the closest corner to $\alpha/\beta$ and called it $\gamma$. Now, we use the fact that

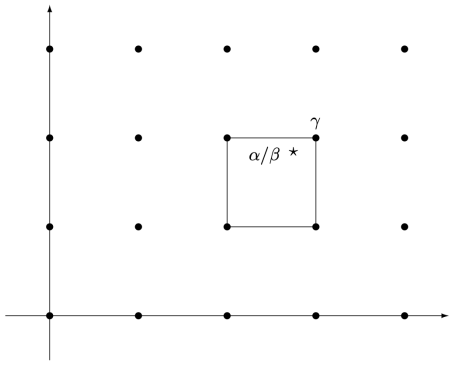



Figure 7.1: Choosing an element $\gamma$ of $\mathbb{Z}[i]$ close to $\alpha/\beta$

regardless of where $\alpha/\beta$ is, its distance to the closest corner of the square is at most half of the length of the diagonal of the square; that is, since the square's sides have length 1 and the diagonal thus has length $\sqrt{2}$, we have that for any $\alpha, \beta \in \mathbb{Z}[i]$ with $\beta \neq 0$, there exists a $\gamma \in \mathbb{C}$ so that

$$\text{Distance from } alpha/\beta \text{ to } \gamma \text{ is at most } \tfrac{\sqrt{2}}{2}.$$

Since $\sigma(x + yi) = x^2 + y^2$ is the square of the distance, we thus have

$$\sigma\left(\frac{\alpha}{\beta} - \gamma\right) \leq \frac{1}{2}.$$

Multiplying both sides by $\sigma(\beta)$ and using the fact that for any $\zeta_1, \zeta_2 \in \mathbb{C}$ we have $\sigma(\zeta_1\zeta_2) = \sigma(\zeta_1)\sigma(\zeta_2)$, we get

$$\sigma(\alpha - \beta\gamma) \leq \frac{1}{2}\sigma(\beta).$$

Let $\rho = \alpha - \beta\gamma$, so $\alpha = \beta\gamma + \rho$ is automatically true, and we further have

$$\sigma(\rho) \leq \frac{1}{2}\sigma(\beta) < \sigma(\beta),$$

since our assumption that $\beta \neq 0$ means $\sigma(\beta) > 0$. Thus $\mathbb{Z}[i]$ is a Euclidean domain. □

Finally, we explore some useful properties of PIDs.

**Theorem 7.2.2**

Let $R$ be a PID, and let $c \in R$ with $c \neq 0$. Then the following are equivalent:

1. $c$ is irreducible.
2. The principal ideal $cR$ is maximal.
3. The quotient ring $R/cR$ is a field.
4. The principal ideal $cR$ is prime.
5. The quotient ring $R/cR$ is an integral domain.

**Proof.** Note that

$$R/cR \text{ is a field} \iff cR \text{ maximal} \implies cR \text{ prime} \iff R/cR \text{ is an integral domain}$$

follow directly from Theorem 3.40 and Corollary 3.41. Thus, we only need to prove that $cR$ prime implies $c$ irreducible, and $c$ irreducible implies $cR$ maximal.
Suppose that $c$ factors into $ab$. We wish to show that either $a$ or $b$ is a unit. Since $ab = c \in cR$, the definition of a prime ideal tells us that either $a \in cR$ or $b \in cR$. Without loss of generality, suppose $a \in cR$. Then $a = cr$ for some $r \in R$, so

$$c = ab = crb.$$

Since $R$ is an integral domain and $c \neq 0$, we need $rb = 1$ (since $c(rb-1) = 0$, $c \neq 0$, and $R$ integral domain means $rb - 1 = 0$). Hence $b \in R^*$, so $c$ is irreducible.
Now, suppose $c$ irreducible and let $I \subseteq R$ be an ideal satisfying

$$cR \subseteq I \subseteq R.$$

We wish to show that $cR$ is maximal; that is, either $I = cR$ or $I = R$. Since $R$ is a PID, we can find some $a \in R$ such that $I = aR$. In particular,

$$c \in cR \subseteq I = aR,$$

so $c = ab$ for some $b \in R$. Since $c$ irreducible, either $a$ or $b$ is a unit:

- If $a$ unit, then $aR = R$, so $I = R$.
- If $b$ unit, then $I = aR = cb^{-1}R = cR$.

Hence $cR$ is a maximal ideal. □

## §7.3 Factorization in Principal Ideal Domains

Now, we'll prove that every Euclidean domain (and thus PID) is automatically a unique factorization domain (UFD). From Section 7.2, this will then show that $\mathbb{Z}$, $\mathbb{Z}[i]$, and $F[x]$ are unique factorization domains. Indeed, even more generally every PID is a UFD (since not all PIDs are Euclidean domains); make sure to prove this as an exercise!
We start by generalizing the notion of divisibility, then follow with a fundamental divisibility property in PIDs.

**Definition 7.3.1: Divisibility**

Let $R$ be an integral domain, and let $a, b \in R$ be elements of $R$. We say that $b$ **divides** $a$, or $b \mid a$, if there is some element $c \in R$ satisfying $a = bc$. We observe that $b \mid a$ is equivalent to the assertion that $a$ is in the ideal $bR$, which in turn is equivalent to the inclusion $aR \subseteq bR$.

**Proposition 7.3.1**

Let $R$ be a principal ideal domain, and let $a, b, c \in R$. Suppose that $a$ is irreducible and that $a \mid bc$. Then either $a \mid b$ or $a \mid c$, or both.

**Proof.** Consider the ideal generated by $a$ and $b$,

$$I = \{ar + bs \mid r, s \in R\}.$$

The assumption that $R$ is a PID tells us that $I$ is principal, so

$$I = dR$$

for some $d \in R$.
We know that $a \in I = dR$, so $a = de$ for some $e \in R$. But $a$ is irreducible, so either $d$ or $e$ is a unit. This gives us two cases:

- If $e \in R^*$, then $d = e^{-1}a$. But we also know that $b \in I = dR$, so $b = df$ for some $f \in R$. Hence $b = e^{-1}af = e^{-1}fa$, so $a \mid b$.
- If $d \in R^*$, then $1 = dd^{-1} \in dR = I$, so we can find some $r, s \in R$ such satisfying

  $$1 = ar + bs.$$

  Multiplying both sides by $c$, we get

  $$c = acr + bcs.$$

  Since we assumed $a \mid bc$, we can write $bc = ag$ for some $g \in R$. Substituting and factoring yields

  $$c = acr + ags = a(cr + gs),$$

  so $a \mid c$.

Thus if $a \mid bc$ and $a$ irreducible, then either $a \mid b$ or $a \mid c$. □

**Corollary 7.3.1**

Let $R$ be a principal ideal domain, and let $a, b_1, \ldots, b_n \in R$. Suppose $a$ is irreducible and $a$ divides the product $b_1 \cdot \ldots \cdot b_n$. Then there is at least one index $i$ such that $a$ divides $b_i$.

**Proof.** We proceed with induction on $n$. The statement is clearly true for $n = 1$, and we just proved the case when $n = 2$. Assume now that $n \geq 3$ and the statement is true for a product of $n - 1$ factors. We group the $n$ factors as

$$b_1 \cdot \ldots \cdot b_n = \underbrace{b_1}_{\text{b}} \cdot \underbrace{(b_2 \cdot \ldots \cdot b_n)}_{\text{c}}.$$

Then $a \mid bc$ and $a$ irreducible (given), so the above theorem shows that either $a \mid b$ or $a \mid c$. The first case is easy, since if $a \mid b$, then we're done (simply set $i = 1$). In the second case, we're given that $a \mid c$. Since $c$ is the product of $n - 1$ factors $b_2, \ldots, b_n$, the induction hypothesis tells us that some index $i$ satisfies $a \mid b_i$. □

Before approaching the main theorem of this section, we make an observation about size functions and units, whose proof is left as an exercise.

**Theorem 7.3.1**

Let $R$ be a Euclidean domain with size function $\sigma$, and let $u \in R$. Then $u \in R^*$ if and only if $\sigma(u) = 1$.

**Theorem 7.3.2: PIDs are UFDs**

Let $R$ be a principal ideal domain. Then $R$ is a unique factorization domain.

**Proof.** We first wish to show that every non-zero non-unit element of $R$ is a product of irreducible elements. We're going to make the stronger assumption that $R$ is a Euclidean domain; that is, $R$ has a size function $\sigma$. [TODO: Prove the more general case for any PID]
Let's consider the set of counterexamples to our goal:

$$S = \{ \text{ non-zero non-unit } a \in R \mid a \text{ is } \underline{\text{not}} \text{ a product of irreducibles} \}.$$

If $S$ is the empty set, we're done, so assume $S \neq \varnothing$.
Consider the sizes $\sigma(a)$ of all elements $a \in S$, and choose an element $a'$ of smallest size. In other words, we select an $a' \in S$ such that for every $a \in S$,

$$\sigma(a') \leq \sigma(a).$$

Since $a'$ is not irreducible (otherwise it wouldn't be in $S$), we can factor $a' = a_1 \cdot a_2$ such that neither $a_1$ or $a_2$ are units. But then

$$\sigma(a') = \sigma(a_1)\sigma(a_2) \text{ with } \sigma(a_1) \geq 2 \text{ and } \sigma(a_2) \geq 2,$$

since non-unit elements have size strictly bigger than 1. It follows that $\sigma(a_1) < \sigma(a')$ and $\sigma(a_2) < \sigma(a')$, so neither $a_1$ nor $a_2$ are in $S$ (since $a'$ is the smallest-sized element in $S$).
Thus, $a_1$ and $a_2$ are "not not" products of irreducibles; that is, they can be represented as products of irreducibles. But then $a' = a_1 a_2$ is a product of irreducibles, contradicting $a' \in S$. Thus $S$ is empty, and so every non-zero non-unit element in $R$ is a product of irreducibles.
Now, we must show that this factorization into irreducible elements is essentially unique. For this, we only assume that $R$ is a PID. Assume that $a \in R$ is written as

$$a = b_1 \cdot \ldots \cdot b_n = c_1 \cdot \ldots \cdot c_m$$

with irreducible elements $b_1, \ldots, b_n,\ c_1, \ldots, c_m$. In particular, this implies that

$$b_1 \mid c_1 \cdot \ldots \cdot c_m.$$

Since $b_1$ is irreducible and $R$ is a PID, by Corollary 7.3.1 (7.16) to deduce that for some index $i$, we have $b_1 \mid c_i$. Relabeling the other elements, we get $b_1 \mid c_1$.
This means that $c_1 = b_1 u_1$ for some $u_1 \in R$. However, since $c_1$ is also irreducible and $b_1$ is not a unit (since it's irreducible), we see that $u_1$ is a unit.
Cancelling $b_1$ from both sides (since integral domains have the cancellation property), we obtain

$$b_2 \cdot b_3 \cdot \ldots \cdot b_n = u_1 \cdot c_2 \cdot c_3 \cdot \ldots \cdot c_m.$$

We repeat the above argument with $b_2$; note that $b_2$ cannot divide $u_1$ since it's irreducible, so after relabeling, we find that $c_2 = b_2 u_2$ for some unit $u_2$. Canceling $b_2$ gives

$$b_3 \cdot \ldots \cdot b_n = u_1 \cdot u_2 \cdot c_3 \cdot \ldots \cdot c_m.$$

Continuing in this fashion, either all of the $b_i$ are gone, or else all of the $c_j$ are gone. But then one side is a unit, so the other side cannot have any irreducible elements left. In other words, we must have $n = m$, and with step-by-step relabeling, we have $c_i = b_i u_i$ with $u_i \in R^*$ for every $i$.
Thus every PID is a UFD. □

As an easy corollary, we get the following result (since each are PIDs).

**Corollary 7.3.2**

$\mathbb{Z}$, $\mathbb{Z}[i]$, and $F[x]$ are unique factorization domains.

**Example 55** (TODO: Finish proof)**.** *As a counterweight to the theorem, it's worthwhile to look at a ring that is* ***not*** *a UFD. Consider the ring*

$$R = \mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Z}\}.$$

*It is a subring of the field of complex numbers. The number* $4 \in R$ *can be factored into*

$$4 = 2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}).$$

## §7.4 The Chinese Remainder Theorem

The first instance of the Chinese Remainder Theorem appeared over 1500 years ago, in a Chinese mathematical work by Master Sun:

> We have a number of things, but we do not know exactly how many. If we count them by threes, we have two left over. If we count them by fives, we have three left over. If we count them by sevens, we have two left over. How many things are there?

We know that a congruence of the form

$$ax \equiv b \pmod{m}$$

has a solution provided that $\gcd(a, m) = 1$. The Euclidean algorithm even provides an efficient way to compute the solution. The simplest version of the Chinese Remainder Theorem (CRT) deals with the case of two simultaneous linear congruences with different moduli.

**Theorem 7.4.1: Chinese Remainder Theorem, v1**

Let $m_1, m_2 \in \mathbb{Z}$ be non-zero integers with $\gcd(m_1, m_2) = 1$ and let $c_1, c_2 \in \mathbb{Z}$.

1. There is a solution $x \in \mathbb{Z}$ to the simultaneous congruences

$$x \equiv c_1 \pmod{m_1} \text{ and } x \equiv c_2 \pmod{m_2}.$$

2. If $x, x'$ are two solutions, then

$$x_1 \equiv x_2 \pmod{m_1 m_2}.$$

Before providing a proof, we illustrate the Chinese Remainder Theorem using an example. Suppose we want to solve

$$x \equiv 8 \pmod{11} \text{ and } x \equiv 3 \pmod{17}.$$

From the first equivalence, we can have $x = 8$; but we could also have $x = -3, 19, \ldots$. Indeed, any solution to the first congruence looks like

$$x = 8 + 11y \text{ for some integer } y.$$

We explore this flexibility by substituting,

$$8 + 11y \equiv 3 \pmod{17} \implies 11y \equiv -5 \equiv 12 \pmod{17}.$$

To solve this, we simply multiply both sides by the inverse of 11 modulo 17. We could use the Euclidean algorithm, or simply trial-and-error, to reveal that $14 \cdot 11 \equiv 1 \pmod{17}$, so

$$y \equiv 14 \cdot 11y \equiv 14 \cdot 12 \equiv_1 68 \equiv 15 \pmod{17}.$$

So, we have $y = 15$, and substituting into $x = 8 + 11y$, we get $x = 173$ to solve the simultaneous congruences.

**Proof.** 1. We start with the first congruence, and observe (like above) that any solution in the form

$$x = c_1 + m_1 y \text{ for some } y \in \mathbb{Z}$$

solves the congruence. Now, we just need to choose a $y$ so that $x$ is also a solution to the second congruence $x \equiv c_2 \pmod{m_2}$. In other words, we want to find an integer $y$ that satisfies

$$c_1 + m_1 y \equiv c_2 \pmod{m_2},$$

or equivalently

$$m_1 y \equiv c_2 - c_1 \pmod{m_2}.$$

We know from Theorem 1.39 that this has a solution if and only if $\gcd(m_1, m_2) = 1$; but we know this is true by assumption. Thus, there exists a $y$ so that $x$ solves the simultaneous congruences, as desired.

2. If $x$ and $x'$ are both solutions to the given congruences, we have

$$x - x' \equiv c_1 - c_1 \equiv 0 \pmod{m_1} \text{ and } x - x' \equiv c_2 - c_2 \equiv 0 \pmod{m_2}.$$

Thus $x - x'$ is divisible by both $m_1$ and $m_2$. Since $\gcd(m_1, m_2) = 1$, we conclude that $x - x'$ is divisible by the product $m_1 m_2$; in other words,

$$x \equiv x' \pmod{m_1 m_2}.$$

□

**Remark 17.** *Here's a more modern way to state the Chinese Remainder Theorem, using product rings. Consider the natural homomorphism*

$$\phi : \mathbb{Z}/m_1m_2\mathbb{Z} \longrightarrow \mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z}$$
$$a \bmod m_1m_2 \longmapsto (a \bmod m_1, a \bmod m_2).$$

*The CRT is equivalent to the assertion that if* $\gcd(m_1, m_2) = 1$*, then the map* $\phi$ *is a ring isomorphism. More precisely, CRT(a) says that* $\phi$ *is surjective (since every simultaneous congruence has a solution, so there always exists an* $a \in \mathbb{Z}$ *that satisfies the requirements), and CRT(b) says that* $\phi$ *is injective (since if two solutions work, then they are congruent).*

There are many ways to generalize the CRT; for example, we could solve a longer list of simultaneous congruences, work with rings other than $\mathbb{Z}$, etc. Here, we provide a more general version, but [TODO: include exercise] we can prove a more general result.

**Theorem 7.4.2: Chinese Remainder Theorem, v2**

Let $R$ be a commutative ring, and let $c_1, \ldots, c_n \in R$ be elements having the property that[a]

$$c_iR + c_jR = R \text{ for all } i \neq j.$$

Let $c = c_1c_2 \cdots c_n$. Then there is an isomorphism

$$\phi : R/cR \longrightarrow R/c_1R \times R/c_2R \times \ldots \times R/c_nR$$
$$r \bmod c \longmapsto (r \bmod c_1, r \bmod c_2, \ldots, r \bmod c_n).$$

[a] Note that for $R = \mathbb{Z}$, this property is equivalent to $\gcd(c_i, c_j) = 1$. Keep this in mind when thinking about how this generalizes the CRT.

**Proof.** The proof is by induction on $n$. For $n = 1$, we have $c = c_1$, and $\phi$ is just the identity map $R/cR \to R/cR$. However, in order for the induction step to work, we also need to directly check the case of $n = 2$ (we will see later why this is the case):

**Lemma 7.4.1**

Let $R$ be an integral domain, and let $a, b \in R$ be non-zero elements satisfying $aR + bR = R$. Then there is an isomorphism

$$\phi : R/abR \longrightarrow R/aR \times R/bR, \ \phi(r \bmod ab) = (r \bmod a, r \bmod b).$$

**Proof.** We're given that $aR + bR = R$, so we can find elements $u, v \in R$ satisfying

$$au + bv = 1$$

(since $1 \in R$). We start with the map

$$\psi : R \longrightarrow R/aR \times R/bR, \ \psi(r) = (r \bmod a, r \bmod b).$$

We first compute the kernel of $R$. It's clear that $abR \subseteq \ker(\psi)$ (since any $c \in abR$ is a multiple of both $a$ and $b$, and so is equivalent to $(0, 0)$), so we wish to show the opposite

inclusion $\ker(\psi) \subseteq abR$. Let $r \in \ker(\psi)$. Then

$$r \equiv 0 \pmod{aR} \text{ and } r \equiv 0 \pmod{bR},$$

so we have

$$r = as = bt \text{ for some } s, t \in R.$$

Multiplying the relation $1 = au + bv$ by $s$, we get

$$s = s(au + bt) = asu + bsv = btu + bsv = b(tu + sv).$$

Thus, $b \mid s$, say $s = bw$. Then

$$r = as = abw \in abR.$$

Therefore, we get the opposite inclusion $\ker(\psi) \subseteq abR$, so equality holds, and by Proposition 3.31(c) (the first isomorphism theorem for rings), $\psi$ induces an injective homomorphism

$$\phi : R/abR \hookrightarrow R/aR \times R/bR.$$

We now only need to verify that $\psi$ is surjective, which shows that $\phi$ is surjective as well (and hence an isomorphism). Since $bv = 1 - au$, we have

$$\psi(1 - au) = \psi(bv) = (1 - au \bmod a, bv \bmod b) = (1, 0).$$

Similarly, we have $au = 1 - bv$, so

$$\psi(1 - bv) = \psi(au) = (au \bmod a, 1 - bv \bmod b) = (0, 1).$$

So for any desired $(c, d) \in R/aR \times R/bR$, we have

$$(c, d) = (c, 0) + (0, d) = c(1, 0) + d(0, 1) = c\psi(bv) + d\psi(au) = \psi(cbv + dau).$$

Thus every $(c, d)$ is in the image of $\psi$, and so $\psi$ and $\phi$ are surjective. □

We now resume the proof for the CRT. We've shown that it holds for $n = 1, 2$, so assume that the theorem holds for $n$. We then need to prove that it's true for $n + 1$.
Let $c_1, \ldots, c_{n+1} \in R$ as described in the statement (that is, each $c_i R + c_j R = R$ for $i \neq j$). In particular, we're given that

$$c_1 R + c_{n+1} R = R = (1), \ \ldots, \ c_n R + c_{n+1} R = 1.$$

Thus, we can find elements in $R$ that satisfy

$$c_1 u_1 + c_{n+1} v_1 = 1, \ldots, c_n u_n + c_{n+1} v_n = 1.$$

Multiplying all of these equations together and factoring out $c_{n+1}$, we get

$$\underbrace{c_1 c_2 \cdots c_n u_1 u_2 \cdots u_n}_{\text{terms that don't contain } c_{n+1}} + \underbrace{c_{n+1} \cdot (\text{a big mess})}_{\text{terms that do contain } c_{n+1}} = 1.$$

This shows that

$$c_1 \cdots c_n R + c_{n+1} R = R,$$

which means that we can use the above Lemma with $a = c_1 \cdots c_n$ and $b = c_{n+1}$ to deduce that there is an isomorphism

$$R/c_1 c_2 \cdots c_n R \xrightarrow{\sim} R/c_1 c_2 \cdots c_n R \times R/c_{n+1} R.$$

But our induction hypothesis tells us that there's an isomorphism

$$R/c_1c_2\cdots c_nR \xrightarrow{\sim} R/c_1R \times \ldots \times R/c_nR.$$

Combining these isomorphisms, we get

$$\begin{aligned} R/c_1c_2\cdots c_nR &\xrightarrow{\sim} R/c_1c_2\cdots c_nR \times R/c_{n+1}R \\ &\xrightarrow{\sim} R/c_1R \times \ldots \times R/c_nR \times R/c_{n+1}R, \end{aligned}$$

which completes the proof that the theorem holds for $n+1$. Hence by induction, the theorem is true for all $n$. □

### §7.4.1 An Application of the Chinese Remainder Theorem

[TODO]

## §7.5 Field of Fractions

Now, we wish to show that every integral domain $R$ is a subring of a field, and that there is a smallest such field $F$. The idea of the proof is to construct $F$ from $R$ just as one constructs $\mathbb{Q}$ from $\mathbb{Z}$. However, don't be fooled by the assumption that $\mathbb{Q}$ is just the set of fractions $\frac{a}{b}$; even in $\mathbb{Q}$, we must be careful with the fact that different looking fractions $\frac{1}{2}, \frac{2}{4}, \frac{137}{274}$ are really the same quantity.

**Theorem 7.5.1**

Let $R$ be an integral domain. There exists a field $F$, called the **field of fractions of** $R$, with the following properties:

1. The ring $R$ is a subring of the field $F$.
2. If $R$ is also a subring of some other field $K$, then there is a unique injective homomorphism $F \hookrightarrow K$ that takes $R$ to itself by the identity map.

**Proof.** We must first construct a fraction field $F$ from $R$, in the same way we constructed the field of rational numbers $\mathbb{Q}$ from the ring of integers $\mathbb{Z}$, but we must be careful. Let's start with the set of pairs

$$\{(a,b) \mid a,b \in R,\ b \neq 0\},$$

and we define an equivalence relation $\sim$ on this set by saying that

$$(a,b) \sim (a',b') \text{ if } ab' = a'b$$

(make sure to prove that this is an equivalence relation! This is equivalent to saying that $\frac{a}{b} = \frac{a'}{b'}$). We then define $F$ to be the set

$$F = \{\text{equivalence classes of pairs } (a,b)\}.$$

We can informally view the pair $(a,b)$ as the fraction $\frac{a}{b}$, but even for $\mathbb{Q}$, the fraction $\frac{a}{b}$ is the equivalence class of pairs of integers $(a,b)$ that all correspond to the same fraction. We want to make $F$ a field, so we need an addition rule and a multiplication rule. We define:

$$\begin{aligned} (a_1,b_1) + (a_2,b_2) &= (a_1b_2 + a_2b_1, b_1b_2) \\ (a_1,b_1) \cdot (a_2,b_2) &= (a_1a_2, b_1b_2). \end{aligned}$$

Addition may seem weird, but think about how we add $\frac{a_1}{b_1}$ and $\frac{a_2}{b_2}$; these are really the same!
There's a lot to check for well-definedness. First, we need to show that if

$$(a_1, b_1) \sim (a_1', b_1') \text{ and } (a_2, b_2) \sim (a_2', b_2'),$$

then the sum $(a_1, b_1) + (a_2, b_2)$ gives the same result as the sum $(a_1', b_1') + (a_2', b_2')$, and similarly for their products. [TODO: FINISH DA PROOF]
Now, we show that this addition and multiplication make $F$ a field. We skip most of the tedious proofs of the axioms here, but note that the additive identity is $(0, 1)$, and the multiplicative identity is $(1, 1)$. The additive inverse of any $(a, b)$ will be $(-a, b)$:

$$(a, b) + (-a, b) = (ab - ab, bb) = (0, bb) \sim (0, 1);$$

and for any $(a, b) \not\sim (0, 1)$, it has a multiplicative inverse $(b, a)$:

$$(a, b) \cdot (b, a) = (ab, ab) \sim (1, 1).$$

We leave the rest for you to check.

Now, we must show that the fraction field $F$ contains a copy of $R$. Consider the map

$$\phi : R \longrightarrow F, \ a \longmapsto (a, 1).$$

This is clearly injective, since $(a, 1) \sim (b, 1)$ if and only if $a = b$. $\phi$ is additionally a ring homomorphism, since

$$\phi(a + b) = (a + b, 1) = (a \cdot 1 + b \cdot 1, 1 \cdot 1) = (a, 1) + (b, 1) = \phi(a) + \phi(b),$$

and

$$\phi(a \cdot b) = (a \cdot b, 1) \sim (a, 1) \cdot (b, 1).$$

Hence $\phi$ is a ring isomorphism from $R$ onto its image (which is in $F$), so $F$ contains a copy of $R$.

Finally, we must show that $F$ is the smallest field containing a copy of $R$. What does this mean, exactly? If a field $K$ contains $R$, then $K$ also contains $F$; but from above, "$F$ contains $R$" should be interpreted as saying that there's an injective homomorphism from $R$ to $F$. So, we wish to prove:

If $\phi : R \hookrightarrow F$ and $\psi : R \hookrightarrow K$ are injective ring homomorphisms, then there is a unique field homomorphism $\lambda : F \hookrightarrow K$ satisfying $\lambda \circ \phi = \psi$.

If the map $\lambda$ is going to exist, then for every $a \in R$ we need

$$\psi(a) = \lambda \circ \phi(a) = \lambda(a, 1).$$

Therefore, we don't have any choice for the value of $\lambda(a, 1)$, since the map $\psi$ is already predetermined (from our assumption that $K$ contains a copy of $R$). Next, observe that any element $(a, b) \in F$ can be written as

$$(a, b) = (a, 1) \cdot (1, b) = (a, 1) \cdot (b, 1)^{-1}$$

(recall that in general, $(a, b)^{-1} = (b, a)$). Since $\lambda$ is required to be a homomorphism, we have

$$\begin{aligned} \lambda(a, b) &= \lambda((a, 1) \cdot (b, 1)^{-1}) \\ &= \lambda((a, 1)) \cdot \lambda((b, 1))^{-1} \\ &= \psi(a)\psi(b)^{-1}. \end{aligned}$$

(Since $b \neq 0$ and $\psi$ is injective, $\psi(b) \neq 0$, so $\psi(b)$ has an inverse in the field $K$). To reiterate, we've shown that if there is any map $\lambda : F \to K$ satisfying $\lambda \circ \phi = \psi$, then it must satisfy

$$\lambda(a, b) = \psi(a)\psi(b)^{-1}.$$

We must check that $\lambda$ is well-defined, and is actually a homomorphism. Assume that $(a, b) \sim (a', b')$, and inspect $\lambda(a, b)$ and $\lambda(a', b')$. $(a, b) \sim (a', b')$ means $ab' = a'b$, so $\psi$ homomorphism means

$$\psi(a)\psi(b') = \psi(a')\psi(b).$$

Hence

$$\lambda(a, b) = \psi(a)\psi(b)^{-1} = \psi(a')\psi(b')^{-1} = \lambda(a', b'),$$

and so $\lambda$ is well-defined.
[TODO: show $\lambda$ is a homomorphism] Hence $F$ is the smallest field containing $R$. □

Let's look at an important example involving the ring of polynomials $F[x]$. We know that $F[x]$ is an integral domain, so $F[x]$ has a field of fractions:

**Definition 7.5.1: Field of Rational Functions**

Let $F$ be a field. The **field of rational functions over** $F$ is the fraction field of $F[X]$, denoted $F(X)$.

There's nothing mysterious about $F(X)$; it is simply the field whose elements look like

$$\frac{\text{polynomial}}{\text{polynomial}}, \text{ where the denominator is not the zero polynomial },$$

with the usual requirement that $\frac{f_1}{g_2} = \frac{f_2}{g_1}$ if $f_1 g_2 = f_2 g_1$. We've already seen rational functions in calculus, since they're nice to integrate, differentiate, and graph. However, since we're calling them "functions", we must be careful. For instance, given some $\alpha \in F$, we might try to define the evaluation map

$$E_\alpha : F(X) \longrightarrow F, \; E_\alpha(\frac{f(x)}{g(x)}) = \frac{E_\alpha(f(x))}{E_\alpha(g(x))} = \frac{f(\alpha)}{g(\alpha)}.$$

However, this isn't always well-defined; what if $g(\alpha) = 0$, or both are 0? [TODO; Exercise 7.17 for a fix]

## §7.6 Multivariate and Symmetric Polynomials

[TODO: finish]

# Chapter 8

# Fields: Part II

More field theory. POGGIES!!!

## §8.1 Algebraic Numbers and Transcendental Numbers

We are vaguely familiar with algebraic and transcendental numbers from high-school algebra. We now formalize their definition.

**Definition 8.1.1: Algebraic and Transcendental Numbers**

Let $L/F$ be an extension of fields, and let $\alpha \in L$. $\alpha$ is **algebraic over** $F$ if $\alpha$ is the root of a non-zero polynomial in $F[x]$. Otherwise, $\alpha$ is **transcendental over** $F$.

The numbers $\sqrt{3}$ and $2+i$ are algebraic over $\mathbb{Q}$, since they are roots of, respectively, the polynomials $x^2-3$ and $x^2-4x+5$. Other numbers may not seem algebraic, but really are:

$$\sqrt{\sqrt{2}+1}+\sqrt[3]{5}$$

is the root of $x^{12}-6x^{10}-20x^9+9x^8+154x^6-360x^5+441x^4-180x^3-456x^2-1680x+274$.