Abstract Algebra

Muchang Bahng

Spring 2024

Contents

With set theory, we have established what sets, along with functions and relations are. Abstract algebra extends on this by studying *algebraic structures*, which are sets S with specific *operations* acting on their elements. This is a very natural extension and to be honest does not require much motivation. Let's precisely define what operations are.

Definition 0.1 (Operation)

A **p-ary operation**[a] $*$ on a set A is a map

$$* : A^p \longrightarrow A \tag{1}$$

where A^p is the p-fold Cartesian product of A. In specific cases,

1. If $p = 1$, then $*$ is said to be **unary**.
2. If $p = 2$, then $*$ is **binary**.

We can consider for $p > 2$ and even if p is infinite.

[a]or called an operation of arity p.

Definition 0.2 (Algebraic Structure)

An **algebraic structure** is a nonempty set A with a finite set of operations $*_1, \ldots, *_n$ and satisfying a finite set of axioms. It is written as $(A, *_1, \ldots, *_n)$.

If we consider functions between algebraic structures $f : A \to B$, there are some natural properties that we would like f to have.

Definition 0.3 (Preservation of Operation)

Given algebraic structures (A, μ_A), (B, μ_B), where μ_A and μ_B have the same arity p, a function $f : A \to B$ is said to **preserve the operation** if for all $x_1, \ldots, x_p \in A$,

$$f(\mu_A(x_1, \ldots, x_p)) = \mu_B(f(x_1), f(x_2), \ldots, f(x_p)) \tag{2}$$

Functions that preserve operations are generally called *homomorphisms*. However, given that preservation is defined with respect to each operation, a map may preserve one operation but not the other. Therefore, we will formally define homomorphisms for each class of algebraic structures we encounter.

Definition 0.4 (Commutative, Associative Operations)

A binary operation $\cdot : A \times A \to A$ is said to be

1. **associative** if for all $a, b, c \in A$, $(ab)c = a(bc)$.
2. **commutative** if for all $a, b \in A$, $ab = ba$.

Associativity is a particularly important property that we would like to have, and it is quite rare to work with algebraic structures that don't have associativity. It basically states that when doing an operation sequentially over 3 elements, it doesn't matter if we evaluate ab or bc first. Therefore, associativity allows us to throw the parentheses away since the evaluated result does not change.

Commutativity on the other hand is not as prevalent. It simply tells us that we can "swap" terms when evaluating. This usually is a another nice convenience, and in the theory of rings commutativity is very prevalent. Either way, in both of these scenarios we can extend to any finite sequence of operations.

Theorem 0.1 (Generalized Associativity)

Given that a binary operation $\cdot$ is associative on a set S, it is always the case that for any finite collection $a_1, \ldots, a_n$, the value $a_1 \ldots a_n$ is unique.

Proof.

We prove by strong induction on n from $n = 3$. Clearly $(a_1 a_2) a_3 = a_1 (a_2 a_3)$ by definition of associativity. The rest is a bit tedious but is mentioned in Jacobson's *Basic Algebra 1*.

Theorem 0.2 (Generalized Commutativity)

Given that a binary operation $\cdot$ is commutative and associative on a set S, with $\alpha = a_1 + \ldots + a_n$, we have

$$\alpha = a_{i_1} + \ldots + a_{i_n} \tag{3}$$

for any permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$.

Now that we've gotten these out of the way, we can start talking about algebraic structures. I've went through 4 main textbooks, plus Google and talking to friends/professors in creating these notes.

1. Vinberg's *A Course in Algebra*.
2. Nathan Jacobson's *Basic Algebra 1*, given to me by Marty.
3. Ted Shifrin's *Abstract Algebra, A Geometric Approach*, used in Duke Math 401.
4. Dummit and Foote's *Abstract Algebra, 3rd Edition*, used in Duke Math 501.

1 Group-Like Structures

1.1 Semigroups and Monoids

Now the endowment of some structures gives rise to the following. Usually, we will start with the most general algebraic structures and then as we endow them with more structure, we can prove more properties. Let's talk about the most basic type of algebraic structure. If you have a set S and some associative operation on it, we have a semigroup.

Definition 1.1 (Semigroup)

A **semigroup** $(S, \cdot)$ is a set S with an associative binary operation

$$\cdot : S \times S \to S \tag{4}$$

Example 1.1 (Continuous Time Markov Chain)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(S, \mathcal{S})$ a measurable space. Then, a homogeneous continuous-time Markov chain is a stochastic process $\{X_t\}_{t \geq 0}$ taking values in S (i.e. $X_t : \Omega \to S$) satisfying the **Markov property**: for every bounded measurable f and and $t, s \geq 0$,

$$\mathbb{E}[f(X_{t+s}) \mid \{X_r\}_{r \leq t}] = \mathbb{E}[f(X_{t+s}) \mid X_t] = (P_s f)(X_t) \tag{5}$$

The set $\{P_t\}_{t \geq 0}$ with the composition operation gives us the *Markov semigroup*.

To be honest the above example is the only time I have every seen a semigroup come up, so we proceed immediately to the next structure.

Definition 1.2 (Monoid)

A **monoid** $(M, \cdot, e)$ is a semigroup with an identity element $e \in M$ such that given a $m \in M$

$$e \cdot m = m \cdot e = m \tag{6}$$

We first should ask whether the identity is unique in a monoid. It turns out it is.

Lemma 1.1 (Uniqueness of Monoid Identity)

The identity e of a monoid M is unique.

Proof.

Assume not, i.e. there are 2 identities $e \neq e'$. But then

$$e = ee' = e' \implies e = e' \tag{7}$$

where the implication follows from transitivity of equivalence relations.

From set theory, we have directly worked with two examples of monoids.

Example 1.2 (Set Operations)

Let S be any nonempty set. Then $(2^S, \cup, \emptyset)$ and $(2^S, \cap, S)$ are monoids. So it seems that there are flavors of algebra that aren't really separable from set theory.

Definition 1.3 (Submonoid)

Given a monoid $(M, *)$, let $M' \subset M$. If the restriction of $*$ to $M' \times M'$ is closed in M', then we can define the **submonoid** $(M', *)$.

It may seem like the identity of a submonoid must be the identity of the monoid, but this is not always the case. We may take a subset $M' \subset M$ such that $\cdot$ is closed in M' and there may be some $e' \in M', e' \neq e$ such that it acts like an identity on M'.

Example 1.3 (Identities of Submonoids May Not be the Same)

Let $(M, \times, I)$ be the monoid of 2×2 matrices over $\mathbb{R}$ with the identity matrix I, and let M' be the set of matrices of form

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \text{ for } a \in \mathbb{R}, \qquad I' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{8}$$

Then $(M', \times, I')$ is a submonoid with a different identity element.

Example 1.4 ($\mathbb{N}$ is a Monoid)

The natural numbers, defined here are a monoid. More specifically,

1. $(\mathbb{N}, +, 0)$ is a monoid under addition.
2. $(\mathbb{N}, \times, 1)$ is a monoid under multiplication.
3. $(2\mathbb{N}, +, 0)$ is a monoid under addition, where $2\mathbb{N}$ is the set of all even numbers.
4. $2\mathbb{N}$ cannot be a monoid since $1 \notin 2\mathbb{N}$.

Definition 1.4 (Monoid of Transformations)

Given a set S, consider the set of all functions $S^S \coloneqq \{f : S \to S\}$. Then, with function composition $\circ$, $(S^S, \circ)$ is a monoid with the identity function $e : x \mapsto x$ as the identity element. This is called the **monoid of transformations** of S.

Theorem 1.1 (Cardinality of Monoid of Transformations)

If $|S| = n$, then the monoid of transformations has cardinality n^n.

1.2 Groups

Now we look at a specific case of monoids where invertibility is defined. The existence of inverses produces a whole suite of interesting properties, as we will see.

Definition 1.5 (Group)

A **group** $(G, \cdot)$ is a set with binary operation $x \cdot y$—also written as xy—having the following properties.

1. *Closure.* $a, b \in G \implies ab \in G$[a]
2. *Associativity.* $\forall a, b, c \in G, a(bc) = (ab)c$
3. *Identity.* $\exists e \in G$ s.t. $\forall a \in G, ae = ea = a$
4. *Inverses.* $\forall a \in G \; \exists a^{-1} \in G$ s.t. $aa^{-1} = a^{-1}a = e$

The **order** of a group is the cardinality $|G|$. An **abelian group** $(G, +)$ is a group where $+$ is commutative.[b]

[a] but not necessarily $ab = ba$

[b] Note that I switched the notation from $*$ to $+$. By convention and to avoid confusion, $+$ denotes commutative operations.

This is an extremely simple structure, and the first thing we should prove is the uniqueness of the identity and inverses.

Lemma 1.2 (Uniqueness of Identity and Inverse)

The identity and the inverse is unique, and for any a, b, the equation $xa = b$ has the unique solution $x = ba^{-1}$.

Proof.

Assume that there are two identities of group $(G, *)$, denoted e_1, e_2, where $e_1 \neq e_2$. According to the properties of identities, $e_1 = e_1 e_2 = e_2 \implies e_1 = e_2$.
As for uniqueness of a inverses, let a be an element of G, with its inverses a_1^{-1}, a_2^{-1}. Then,

$$\begin{aligned} aa_1^{-1} = e &\implies a_2^{-1}(aa_1^{-1}) = a_2^{-1}e && (9)\\ &\implies (a_2^{-1}a)a_1^{-1} = a_2^{-1} && (10)\\ &\implies ea_1^{-1} = a_2^{-1} && (11) \end{aligned}$$

Since the inverse is unique, we can operate on each side of the equation $xa = b$ to get $xaa^{-1} = ba^1 \implies xe = x = ba^{-1}$. Clearly, the derivation of this solution is unique since the elements that we have operated on are unique.

At this point, we can see that for each group there is a corresponding "multiplication table" defined by the operation. For example, we can create a set of 6 elements $\{r_0, r_1, r_2, s_0, s_1, s_2\}$ and define the operation $\times$ as the following.

$\times$	r_0	r_1	r_2	s_0	s_1	s_2
r_0	r_0	r_1	r_2	s_0	s_1	s_2
r_1	r_1	r_2	r_0	s_1	s_2	s_0
r_2	r_2	r_0	r_1	s_2	s_0	s_1
s_0	s_0	s_2	s_1	r_0	r_2	r_1
s_1	s_1	s_0	s_2	r_1	r_0	r_2
s_2	s_2	s_1	s_0	r_2	r_1	r_0

Figure 1: Multiplication table for some group. Note that we can only write such a table explicitly for a group of finite elements. But even for arbitrary groups, we should think of the operation completely defining a possibly "infinite" table.

It is clear that in an abelian group, the multiplication table must be symmetric across the diagonal.

Example 1.5 (Familiar Groups)

So what are some examples of groups?

1. $(\mathbb{N}, +)$ is not a group since $3 \in \mathbb{N}$ but $-3 \notin \mathbb{N}$. It is a commutative monoid.
2. $(\mathbb{N}, \times)$ is not a group but is a commutative monoid.
3. $(\mathbb{Z}, +)$ is an abelian group.
4. $(\mathbb{Z}, \times)$ is not a group.
5. $(\mathbb{Q}, +)$ and $(\mathbb{Q} \setminus \{0\}, \times)$ are both abelian groups.
6. $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \times)$ are both abelian groups.
7. The set of all invertible $n \times n$ matrices with matrix multiplication, denoted $(\mathrm{GL}(\mathbb{R}^n), \times)$ is a non-abelian group.
8. The set of all functions on a given interval $[a, b]$ is abelian with respect to addition, defined as $(f + g)(x) \equiv f(x) + g(x)$.

Example 1.6 (Group of Invertible Elements of a Monoid)

Given $x \in (M, \cdot, e)$, let x be **invertible** if there exists $x^{-1} \in M$ s.t. $xx^{-1} = x^{-1}x = e$. Then, the submonoid M' of invertible elements of M is a group. This must be proved.

1. *Closure.* If $x, y \in M'$, then $x^{-1}, y^{-1} \in M'$ since $(x^{-1})^{-1} = x$. Therefore $y^{-1}x^{-1} = (xy)^{-1} \in M'$, and so $xy \in M'$.
2. *Identity.* $e^{-1} = e$ so $e \in M'$.
3. *Inverses.* Exists by definition.
4. *Associativity.* Is inherited from associativity of $\cdot$ in M.

Let's prove a little more about groups so that we have more tools for manipulation.

Lemma 1.3 (Properties of Group Operation)

Given $a, b, c \in G$,

1. $ab = cb \implies a = c$.
2. $\forall a \in G$, $(a^{-1})^{-1} = a$.
3. $(ab)^{-1} = b^{-1}a^{-1}$.

Proof.

TBD.

Theorem 1.2 ()

Given group G, $(ab)^2 = a^2b^2$ for all $a, b \in G$ iff G is abelian.

Definition 1.6 (Subgroup)

Given group $(G, *)$, a **subgroup** $(H, *)$ is a group such that $H \subset G$. H is called a **proper subgroup** if $H \subsetneq G$.

Theorem 1.3 ()

If $H, K \subset G$ are subgroups, then $H \cap K$ is a subgroup.

Finally we end with an analogous result of the monoid of transformations. The problem with these transformations is that they may not be invertible, but if they are, i.e. bijective, then we can endow them with a group structure.

Definition 1.7 (Group of Transformations)

Given a set S, $\text{Sym}(S)$ is the group of bijective maps $f : S \to S$ with composition as the operator. This is also called the **symmetric group** of S.

Lemma 1.4 (Cardinality of Group of Transformations)

If S has cardinality n, then the order of $\text{Sym}(S)$ is $n!$.

1.3 Group Homomorphisms

At this point, we would like to try and classify groups (e.g. can we find *all* possible groups of a finite set?). But consider the two groups.

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

+	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

Figure 2: Two isomorphic groups.

These groups have different elements, but the operation behaves in exactly the same way between them (it may be a little harder if I relabeled the elements or permuted the rows/columns). Since we can trivially make arbitrary sets there really isn't much meaning to having two versions of the same group (at least in the algebraic sense). Therefore, these groups should be labeled "equivalent" in some way, and we will precisely define this notion now.

Definition 1.8 (Group Homomorphism)

Let $(G, \circ)$ and $(H, *)$ be two groups. The mapping $f : (G, \circ) \longrightarrow (H, *)$ is a **group homomorphism** if for all $a, b \in G$,

$$f(a \circ b) = f(a) * f(b) \tag{12}$$

Furthermore,

1. A **group isomorphism** is a bijective group homomorphism, and we call groups M, N **isomorphic**, denoted $M \simeq N$, if there exists an isomorphism between them.
2. An **endomorphism** is a homomorphism from a group to itself.
3. An **automorphism** is a isomorphism from a group to itself.

It turns out that from the simple property that $f(ab) = f(a)f(b)$, it also maps identities to identities, and inverses to inverses!

Lemma 1.5 (Homomorphisms Maps Identities/Inverses to Identities/Inverses)

Given a homomorphism $f : (G, *) \to (H, \times)$ and $a \in G$,

$$f(e_G) = e_H, \qquad f(a^{-1}) = f(a)^{-1} \tag{13}$$

Proof.

Let $a \in G$. Then

$$f(a) = f(ae_G) = f(a)f(e_G) \implies e_H = f(a)^{-1}f(a) = f(a)^{-1}f(a)f(e_G) = f(e_G) \tag{14}$$

To prove inverses, we see that

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(e_G) = e_H \tag{15}$$

from above, and this implies that $f(a^{-1}) = f(a)^{-1}$. We can also do this with right hand side multiplication.

Example 1.7 (Exponential Map)

The map $a \mapsto 2^a$ is an isomorphism between $(\mathbb{R}, +)$ and $(\mathbb{R}^+, \times)$ since

$$2^{a+b} = 2^a \times 2^b \tag{16}$$

which is proved in my real analysis notes when constructing the exponential map on the reals.

Example 1.8 (Determinant)

The determinant $\det : \mathrm{GL}_n(\mathbb{F}) \to \mathbb{F}^*$ is a homomorphism because of the product rule for determinants.

Example 1.9 (Projection onto Unit Circle)

Given $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ with $\times$ and $S^1 = \{x \in \mathbb{C} \mid |x| = 1\}$ (which is a group under multiplication), the map $f : \mathbb{C}^* \to S^1$ defined $f(z) = z/|z|$ is a group homomorphism since

$$f(z_1 z_2) = \frac{z_1 z_2}{|z_1 z_2|} = \frac{z_1 z_2}{|z_1||z_2|} = f(z_1) f(z_2) \tag{17}$$

Definition 1.9 (Kernel)

Given group homomorphism $f : G \to H$, the **kernel** of f is the preimage of the identity.

$$\ker(f) \coloneqq \{g \in G \mid f(g) = e_H\} \tag{18}$$

Theorem 1.4 (Images and Kernels of Group Homomorphisms)

If $f : G \to H$ is a group homomorphism, then
1. $\ker(f)$ is a subgroup of G.
2. $\mathrm{Im}(f)$ is a subgroup of H.
3. f is injective iff $\ker(f) = \{e_G\}$.

Proof.

Listed.
1. To show closed, consider $a, b \in \ker(f)$. Then $f(ab) = f(a)f(b) = e_H e_H = e_H \implies ab \in \ker(f)$. Since $f(e_G) = e_H$, $e_G \in \ker(f)$. If $a \in \ker(f)$, then $f(a^{-1}) = f(a)^{-1} = e_H^{-1} = e_H \implies a^{-1} \in \ker(f)$. Finally associativity follows from associativity of the supgroup.
2. TBD
3. We prove bidirectionally.
   (a) $(\rightarrow)$. Since f is injective, $f(a) = f(b) \implies a = b$. Let $a \in \ker(f)$. Then $f(a) = e_H$, and so $f(e_G) = e_H = f(a)$. By injectivity, $a = e_G$, and so $\ker(f) = \{e_G\}$.
   (b) $(\leftarrow)$. Let $a, b \in G$ s.t. $f(a) = f(b)$. Then $f(a)f(b)^{-1} = e_H \implies af(a)f(b^{-1}) = f(ab^{-1}) = e_H \implies ab^{-1} \in \ker(f)$. But by hypothesis $\ker(f) = \{e_G\} \implies ab^{-1} = e_G \implies a = b$.

Theorem 1.5 (Compositions of Group Homomorphisms)

Compositions of group homomorphisms are group homomorphisms.

Now let's focus more on isomorphisms, which we can interpret as a "renaming" of the elements. Not only does it rename the elements, but it preserves all the algebraic properties of the group and each element.

Theorem 1.6 (Properties of Group Isomorphisms)

If $f : G \to H$ is a group isomorphism, then
1. f^{-1} is also a group isomorphism.
2. G is abelian $\implies$ H is abelian.

Proof.

Listed.
1. Since f is bijective by definition, f^{-1} is well-defined and bijective as well. Now we show that f^{-1} is a group homomorphism. Given $c, d \in H$, take
$$f\big(f^{-1}(c), f^{-1}(d)\big) = f\big(f^{-1}(c)\big)\, f\big(f^{-1}(d)\big) = cd \tag{19}$$
where the first equality follows since f is a homomorphism, and the second since f^{-1} is the inverse mapping. Now mapping both sides through f^{-1}, we get
$$f^{-1}(c)f^{-1}(d) = f^{-1}(cd) \tag{20}$$
and so f^{-1} is a homomorphism.
2. Let $c, d \in H$. Then $c = f(a), d = f(b)$ for some $a, b \in G$, and so $cd = f(a)f(b) = f(ba) = f(b)f(a) = dc$.

A trivial example is the identity map, which is an automorphism. But can we generalize this a bit better?

Theorem 1.7 (Conjugate Shift is an Automorphism)

Let G be a group with $a \in G$. Then the following, called a **conjugate shift**, is an automorphism on G.
$$\phi : G \longrightarrow G,\ \phi(x) = axa^{-1} \tag{21}$$

Proof.

The map $\psi : G \longrightarrow G,\ \psi(x) = a^{-1}xa$ is clearly the inverse of ϕ, with $\phi\psi = \psi\phi = I$ for all $x \in G \implies \phi$ is bijective. Secondly, $\phi(x)\phi(y) = axa^{-1}aya^{-1} = a(xy)a^{-1} = \phi(xy) \implies \phi$ preserves the group structure.

1.4 Group Presentations

A group G may be very abstract and complicated, and so working with all its elements can be a bit painful. It would be more useful to work with a smaller subset S of G that can completely characterize G.[1] We would like to formalize this notion, which will be very useful later on. For now, let's start off with a simple element $a \in G$, and perhaps we can consider the elements
$$\ldots, a^{-2}, a^{-1}, a^0 = e, a^1, a^2, \ldots \tag{22}$$
However, there are two interpretations to a^{-2} is it the inverse of a^2 or $a^{-1}a^{-1}$? It turns out that these are equivalent.

Lemma 1.6 (Power to an Integer is Well-Defined)

For all $n \in \mathbb{N}$,
$$(a^{-1})^n = (a^n)^{-1} \tag{23}$$

[1] Note that this is similar to the basis that generates a topology.

Proof.

We prove by induction on n. It is trivially true for $n = 1$. Now given that it is true for some $n \in \mathbb{N}$, we have

$$(a^{-1})^{n+1} = (a^{-1})^n a^{-1} = (a^n)^{-1} a^{-1} = (aa^n)^{-1} = (a^{n+1})^{-1} \tag{24}$$

Therefore it makes sense to just write it as a^{-n}. It may or may not be the case that a may cycle back to itself for some n, i.e. $a = a^n$.

Definition 1.10 (Order of an Element)

The **order** of a group element $a \in G$ is the minimum number $n \in \mathbb{N}$ s.t. $a = a^n$, denoted $|a|$ or $\text{ord}(a)$.[a]

[a] Note that this is different from the order of a group. This is confusing but is the convention.

Let's pause for a bit and focus on the order. The order of an element a tells us how a "behaves" in the broader group. This means that when mapped through an isomorphism, it should behave similarly, i.e. the order shouldn't change. This gives us a nice way to check if two groups cannot be isomorphic, but the converse is not necessarily true in general!

Theorem 1.8 (Preservation of Order in a Group Homomorphism)

If f is a group isomorphism, then $\forall a \in G$ $\text{ord}(a) = \text{ord}(f(a))$.

Proof.

Now we come back to group presentations. The set of all multiples of a may or may not be the group, but if we take a certain subset of these elements and take all multiples of all combinations of them, we may have better coverage of the group.

Definition 1.11 (Word)

A **word** is any written product of group elements and inverses. They are generally in the form

$$s_1^{\epsilon_1} s_2^{\epsilon_2} s_3^{\epsilon_3} ... s_k^{\epsilon_k}, \text{ where } e_i \in \mathbb{Z} \tag{25}$$

e.g. given a set $\{x, y, z\}$, $xy, xz^{-1}yyx^{-2}, ...$ are words.

Definition 1.12 (Generating Set)

The **generating set** $\langle S \rangle$ of a group G is a subset of G such that every element of the group can be expressed as a word of finitely many elements under the group operations. The elements of the generating set are called **generators**.

Definition 1.13 (Group Presentations)

The **free group** F_S over a given set S consists of all words that can be built from elements of S. Often with this generating set S, we have a set of relations R that tell us which elements are equal. The **group presentation** writes both S and R in the form

$$\langle S \mid R \rangle \tag{26}$$

Theorem 1.9 ()

If every element other than the identity has order 2, then G is abelian.

With these group presentations we can start identifying specific groups. Let's start with the simplest group with one generator and zero/one relation: the cyclic group.

Definition 1.14 (Cyclic Group)

A **cyclic group** is a group generated by a single element.

1. In an infinite cyclic group, there is no relation and we write

$$Z \coloneqq \langle a \rangle = \{\ldots, a^{-2}, a^{-1}, e, a, a^2, \ldots\} \tag{27}$$

2. In a finite cyclic group, there exists a $n \in \mathbb{N}$ such that $a^n = e$ and we write

$$Z_n \coloneqq \langle a \mid a^n = e \rangle = \{e, a, a^2, \ldots, a^{n-1}\} \tag{28}$$

Example 1.10 (Cyclic Groups)

Here are some examples of cyclic groups.

1. $(\mathbb{Z}_n, +)$, the integers mod n, is a cyclic group of order n, generated by 1.[a]
2. The nth roots of unity in $\mathbb{C}$ is a cyclic group of order n, generated by the counterclockwise rotation $e^{2\pi/n}$.
3. The set of discrete angular rotations in $SO(2)$, in the form of

$$R = \left\{ \begin{pmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \;\middle|\; \theta \in \left\{ \frac{2\pi}{n} k \right\}_{k=0}^{n-1} \right\} \tag{29}$$

4. $(\mathbb{Z}, +)$ is an infinite cyclic group.

[a] In fact, the generator of $\mathbb{Z}_n$ can be any integer relatively prime to n and less than n.

That's really it for cyclic groups, and to make things simpler, there is a complete characterization of them.

Theorem 1.10 (Cyclic Groups are Unique up to Order)

Given a cyclic group, Z or Z_n

1. If it is finite, then $(Z_n, +) \simeq (\mathbb{Z}_n, +) \simeq \langle 1 \rangle$.
2. If it is infinite, then $(Z, +) \simeq (\mathbb{Z}, +) \simeq \langle 1 \rangle$.

Proof.

Therefore, we have completely characterized all cyclic groups! Furthermore, cyclic groups are contained in the sense that any subgroup is also a cyclic group. So you won't find any weird groups embedded in cyclic groups; you can safely assume that they are all cyclic. The proof for this is quite a useful technique, where we try to arrive at a contradiction between some minimally chosen k and the remainder r that must be less than k.

Theorem 1.11 (Subgroups of Cyclic Groups)

Any subgroup of a cyclic group is cyclic.

Proof.

Let $G = \langle a \rangle$ be a cyclic group. Then given a subgroup H, we must have $e \in H$. If there are no other elements we are done, and if there are extra elements then let $k \in \mathbb{N}$ be the smallest natural (which exists due to the well-ordering principle) such that $a^k \in H$. Now we claim that $H = \langle a^k \rangle$. Given any $a^n \in H$, we can use Euclidean algorithm on the integers to write $n = qk + r$ for $0 \leq r < k$. Therefore,

$$a^n = a^{qk+r} = (a^k)^q \cdot a^r \implies a^r = a^n (a^k)^{-q} \tag{30}$$
$$\implies a^r \in H \tag{31}$$

but this contradicts the fact that k is minimal, and so $r = 0$. This means that $a^n = (a^k)^q$ and so a^n is a multiple of a^k.

Example 1.11 (Integers to Even Integers)

Let $2\mathbb{Z}$ denote the set of all even integers with addition. Then we can verify that this is a group, and

$$\mathbb{Z} \simeq 2\mathbb{Z} \tag{32}$$

Theorem 1.12 (Homomorphisms between Cyclic Groups)

There are precisely $\gcd(n, m)$ homomorphisms $f : Z_n \to Z_m$.

Proof.

The next type of group we will focus on is the dihedral group. These are usually introduced as the symmetry group (group of rotations and flips you can do on a polygon) to preserve its symmetry. However, it seems a bit disconnected with cyclic groups and group presentations, so I introduce it in the following way. Once I define it, I connect to its geometric interpretations in the following examples.

Definition 1.15 (Dihedral Group)

The **Dihedral Group** of order $2n$ is the group

$$\mathrm{Dih}(n) := \langle r, f \mid r^n = f^2 = e, rfr = f \rangle \tag{33}$$

To parse this definition a bit, note that the relation $r^n = e$ behaves like a cyclic group of order n, and so we can interpret these as rotations of an object by $2\pi/n$. The second is that $f^2 = e$ is also a cyclic group of order 2, but it behaves more like a flip in that if you flip twice, you get back to the original. With these relations, we can think of the Dihedral group as having two "copies" of cyclic groups that have some extra properties.

Finally, the relation $rfr = f$ is a bit harder to parse, but it just means that a rotation, then flip, then rotation (which rotates backwards since we flipped), is equal to flipping once. Symbolically, this relation allows us to "push" all of the flips to the back.

$$fr = r^n fr = r^{n-1} f \tag{34}$$

Perhaps a slightly more complicated example for $n = 5$.

$$fr^3 f^3 r = fr^3 fr = fr^2 f = r^5 fr^2 f = r^4 frf = r^3 f^2 = r^3 \tag{35}$$

and after this the relation $r^n = f^2 = e$ allows us to cancel some out.

Example 1.12 (Dihedral Group of Order 4, aka Klein-4 Group)

We use the following group presentation to write the dihedral group of order 4. However, we can relabel them to get a simpler table.

	e	r	f	rf
e	e	r	f	rf
r	r	e	rf	f
f	f	rf	e	r
rf	rf	f	r	e

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Figure 3: Cayley multiplication table for the Klein 4-group.

It can be described as the symmetry group of a non-square rectangle. With the three non-identity elements being horizontal reflection, vertical reflection, and 180-degree rotation.

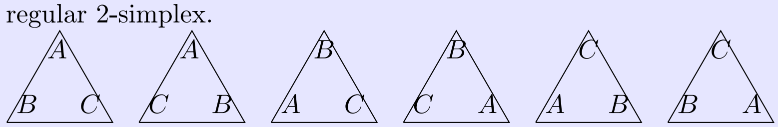
Example 1.13 (Dihedral Group of Order 6)

The group of rotations and flips you can do on a equilateral triangle is called the Dihedral Group Dih(3). It is not abelian.

	e	r	r^2	f	rf	r^2f
e	e	r	r^2	f	rf	r^2f
r	r	r^2	e	rf	r^2f	f
r^2	r^2	e	r	r^2f	f	rf
f	f	r^2f	rf	e	r^2	r
rf	rf	f	r^2f	r	e	r^2
r^2f	r^2f	rf	f	r^2	r	e

Figure 4: Multiplication table for D_3 using simplified notation.

Dih(3) is the group of all rotations and reflections that preserve the structure of the equilateral triangle in $\mathbb{R}^2$, a regular 2-simplex.



Example 1.14 (Dihedral Group of Order 8)

The group of rotations and reflections that preserve the structure of a square in $\mathbb{R}^2$. is called the Dihedral Group Dih(4).

	e	r	r^2	r^3	f	rf	r^2f	r^3f
e	e	r	r^2	r^3	f	rf	r^2f	r^3f
r	r	r^2	r^3	e	rf	r^2f	r^3f	f
r^2	r^2	r^3	e	r	r^2f	r^3f	f	rf
r^3	r^3	e	r	r^2	r^3f	f	rf	r^2f
f	f	r^3f	r^2f	rf	e	r^3	r^2	r
rf	rf	f	r^3f	r^2f	r	e	r^3	r^2
r^2f	r^2f	rf	f	r^3f	r^2	r	e	r^3
r^3f	r^3f	r^2f	rf	f	r^3	r^2	r	e

Figure 5: Multiplication table for D_4 using simplified notation.

Note that this is **not** the same as the symmetry group of the regular tetrahedron!

Following this pattern, we can extrapolate to find that the Dihedral group is a symmetry group.

Theorem 1.13 (Dihedral Groups as Symmetry Groups)

$\text{Dih}(n)$ is similarly the group of all rotations and reflections that preserve the structure of a regular n-gon in $\mathbb{R}^2$.

Example 1.15 (Groups of Order 3)

$\text{Dih}(3) \simeq S_3$, since permutations of the vertices of a triangle are isomorphic to a permutations of a 3-element set.

Theorem 1.14 (Tip)

To prove a group homomorphism, show that every element of G and H can be written as a word of certain g_i's in G and then h_i's in H, and map the g_i's to h_i's.

1.5 Symmetric and Alternating Groups

We have seen the natural construction of the symmetric group of a set as the set of bijective transformations. Now the reason that symmetric groups are nice is that we can embed a group into its symmetric group.

Theorem 1.15 (Cayley's Theorem)

This applies for both monoids and groups.

1. Any monoid is isomorphic to a monoid of transformations, i.e. there exists an injective monoid homomorphism
$$f : M \to M^M \tag{36}$$
2. Any group is isomorphic to a group of transformations, i.e. there exists an injective group homomorphism
$$f : G \to \text{Sym}(G) \tag{37}$$

Proof.

Let $(M, \cdot, 1)$ be a monoid. Then we will construct a homomorphism $f : M \to M^M$, the monoid of transformations from M to itself. For any $a \in M$, we define the *left translation* $a_L : x \mapsto ax$. We claim that the set $M' := \{a_L \in M^M \mid a \in M\}$ is indeed a monoid.

1. *Closure.* Given $a, b \in M$, $ab \in M$ and so $ab_L \in M'$. But $(ab_L)(x) = (ab)x = a(bx) = a_L(bx) = a_L(b_L(x)) = (a_L \circ b_L)(x)$, so $ab_L = a_L \circ b_L$.
2. *Identity.* $e \in M \implies e_L \in M$ where $e_L : x \mapsto x$.

Next we claim that is is an isomorphism.

1. This is a homomorphism due to the closure and identity properties proved above.
2. It is injective since given $a \neq b$ in M, a_L and b_L acts on the identity in different ways $a_L(e) = a \neq b = b_L(e)$, so $a_L \neq b_L$.
3. It is surjective by definition.

We have proved for monoids. For groups, we have the additional assumption that inverses exist in G, and we must prove that the set of left translations G' is indeed a group. It suffices to prove that inverses exist in G'. Given $a \in G$, $a_L \in G'$. But $a^{-1} \in G$ since G is a group, and so $a_L^{-1} \in G'$ as well. We can see that

$$(a_L^{-1} a_L)(x) = (a^{-1}a)x = ex = x \tag{38}$$
$$(a_L a_L^{-1})(x) = (aa^{-1})x = ex = x \tag{39}$$

and so indeed $(a^{-1})_L = (a_L)^{-1}$. From this additional fact all the rest follows exactly as for monoids.

Corollary 1.1 (Cayley)

Every group G is isomorphic to a subgroup of its symmetric group.

Now we limit our scope to only finite sets, i.e. finite symmetric groups, which are often called **permutation** groups. For such finite sets the labeling does not matter since such groups are always isomorphic, so we can say $S = \{1, 2, \ldots, n\}$.

Theorem 1.16 (Symmetric Group as a Symmetry Group)

The symmetric group S_n is isomorphic to the symmetry group of the n-simplex in $\mathbb{R}^{n-1}$.

Proof.

Now armed with group presentations and generating sets, let attempt to find a group presentation for a permutation group. Given set $S = \{1, 2, \ldots, n\}$, a permutation $\gamma \in \mathrm{Sym}(S)$ is denoted

$$\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & \ldots & n-2 & n-1 & n \\ i_1 & i_2 & i_3 & i_4 & \ldots & i_{n-2} & i_{n-1} & i_n \end{pmatrix} \in \mathrm{Sym}(S) \tag{40}$$

We begin by introducing a specific instance of a permutation.

Definition 1.16 (Cyclic Permutation)

A permutation is said to be **cyclic** if there exists some subset $A \subset S$ such that γ acts as

$$a_1 \mapsto a_2 \mapsto a_3 \ldots \mapsto a_k \mapsto a_1 \tag{41}$$

and leaves the rest unchanged. The notation for this is

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_k \end{pmatrix} \in \mathrm{Sym}(S) \tag{42}$$

A cycle acting on a subset of 2 elements, i.e. a swap of two elements, is called a **transposition**. Two cyclic rotations γ_1, γ_2 are **disjoint** if the subsets that they act on are disjoint: $A \cap B = \emptyset$.

Example 1.16 (Some Cyclic Permutations)

This notation can be a bit weird, so let's give some simple examples.

1. (12) is a mapping $1 \to 2,\ 2 \to 1$.
2. (123) is a mapping $1 \to 2,\ 2 \to 3,\ 3 \to 1$.
3. $(123)(45)$ is a mapping $1 \to 2,\ 2 \to 3,\ 3 \to 1,\ 4 \to 5,\ 5 \to 4$.

The reason that cyclic permutations are so important is that they are the building blocks of regular permutations.

Theorem 1.17 (Cycle Decomposition of Permutations)

Every element in S_n except the identity element can be written uniquely (up to order) as the product of disjoint cycles.

Proof.

We can compute $\gamma(1), \gamma^2(1), \ldots$. Since $S = \{1, \ldots, n\}$ is finite, there is some smallest positive natural k s.t.
$gamma^k(1) = 1$. This yields a k-cycle. Now remove the numbers $1, \pi(1), \ldots, \gamma^{k-1}(1)$ and continue the process. Since S is finite this must terminate, and we have such a decomposition. Proof of uniqueness omitted for now, but this whole theorem can be proved using proof by strong induction.

Example 1.17 (Cyclic Decompositions)

For the following permutation

$$\gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 5 & 4 & 8 & 2 & 7 & 1 \end{pmatrix} = (7)(4)(26)(1358) \tag{43}$$

One of the rewards of cycle decompositions is that we can easily compute the effect of conjugation in S_n.

Lemma 1.7 (Conjugation is Easy with Cycle Notation)

Given a k-cycle $\gamma = (i_1, i_2, \ldots, i_k) \in S_n$ and any permutation $\sigma \in S_n$, we have

$$\sigma \pi \sigma^{-1} = \begin{pmatrix} \sigma(i_1) & \sigma(i_2) & \ldots & \sigma(i_k) \end{pmatrix} \tag{44}$$

This shows that the cyclic permutations actually form a generating set of S_n. But can we do better? The answer is yes: there is a more minimal generating set.

Corollary 1.2 (Transposition Decomposition of Permutations)

The set of all transpositions forms a generating set of S_n.

Proof.

It suffices to prove that the cycles can be decomposed into transpositions. Indeed, we can just write out by hand

$$\begin{pmatrix} 1 & 2 & \ldots & k \end{pmatrix} = \begin{pmatrix} 1 & k \end{pmatrix} \begin{pmatrix} 1 & k-1 \end{pmatrix} \ldots \begin{pmatrix} 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix} \tag{45}$$

which by relabeling generalizes for those of form $(i_1 \ldots i_k)$.

Recognizing that the set of transpositions is the generating set of the permutation group, we must prove a few more statements before constructing the alternating group. One such fact is that transpositions allow us talk about the parity of an arbitrary permutation, through its signature.

Lemma 1.8 (Parity of Transpositions)

Every permutation can be written as the product of either an even number or an odd number of transpositions, but not both.

Proof.

Now that this is established, the following is well-defined.

Definition 1.17 (Signature)

The **signature** of a permutation is a homomorphism

$$\text{sgn} : S_n \longrightarrow \{1, -1\} \tag{46}$$

Lemma 1.9 ()

The signature of a permutation changes for every transposition that is applied to it.

Now we are ready to introduce another fundamental type of group.

Definition 1.18 (Alternating Group)

The **alternating group** is the kernel of the signature homomorphism $\text{sgn} : S_n \to \{\pm 1\}$.

$$A_n \coloneqq \ker \text{sgn} \tag{47}$$

It is the set of even permutations with order $n!/2$.[a]

[a] Note that the set of odd permutations do not form a group, since the composition of two odd permutations (each having signature -1 is an even permutation.

This construction might seem arbitrarily specific to study so early into algebra, but as we will see later, we will find that that for $n \geq 5$, they will be simple groups that can't be decomposed and therefore fundamental in a sense.

In lecture, we talked about the number of all finite set is e. Since $n!$ is the order of permutation groups, i.e. the order of automorphism groups, we can sum their inverses over all $n \in \mathbb{N}$ to get e.

1.6 Group Actions

We have studied the general properties of groups, but historically group theory arose from the study of transformation groups (which is why I also introduced is so early on). These transformation groups can be thought of as an abstract group it itself, but another way to interpret it is to see how it *acts* on a set.

Definition 1.19 (Group Action)

Let G be a group, S a set. Then, a (left) group action of G on S is a function

$$\sigma : G \times S \to S, \qquad \sigma(g, a) = g \cdot a \tag{48}$$

satisfying two axioms.

1. *Identity.* $\forall a \in S, \sigma(e, a) = a$.
2. *Compatibility.* $\forall g, h \in G$ and $\forall x \in X, \sigma(gh, x) = \sigma(g, \sigma(h, x))$.

The group G is said to **act on** S, and the evaluation $\sigma(g, a)$ can be interpreted as the result after transforming a through g.

Theorem 1.18 (Group Action as a Homomorphism onto the Symmetric Group)

We have the immediate facts.

1. For a fixed $g \in G$, the group action $\sigma_g(s) \coloneqq \sigma(g, s) : S \to S$ is a bijection, i.e. an element of $\mathrm{Sym}(S)$. The inverse is the function mapping $x \mapsto \sigma(g^{-1}, x)$.
2. The map from G to $\mathrm{Sym}(S)$ defined by $g \mapsto \sigma_g$ is a homomorphism.

Proof.

Example 1.18 (Permutations and Dihedral Groups as Group Actions)

In fact, we have seen two concrete examples of such group actions.

1. The permutation group acts on the set $S = \{1, \ldots, n\}$ by permuting its elements. It also acts on a set of n-simplexes by rotating/flipping them.
2. The dihedral group acts on the set of regular n-gons by rotating/flipping them.

2 Subgroups

We have seen a few examples of subgroups, but we will heavily elaborate on here. We know that given a set, we can define an equivalence relation on it to get a quotient set. Now if we have a group, defining any such equivalence relation may not be compatible with the group structure. Therefore, it would be nice to have some principles in which we can construct such compatible equivalence classes, i.e. through a **congruence relation** that preserves the operations.

We introduce some standard notation.

Definition 2.1 (Subgroup of Integer Multiples)

The set $k\mathbb{Z}$ is the set of all integer multiples of k. This is a group under addition.

2.1 Cosets

Fortunately, we can do such a thing by taking a subgroup $H \subset G$ and "shifting" it to form the cosets of G, which are the equivalence classes.

Definition 2.2 (Coset)

Given a group G, $a \in G$, and subgroup H,
1. A **left coset** is $aH := \{ah \mid h \in H\}$.
2. A **right coset** is $Ha := \{ha \mid h \in H\}$.
3. When G is abelian, the **coset** is denoted $a + H$.

With this, we can take arbitrary elements $a, b \in G$ and determine if they are in the same coset as such. Since $a \in aH$, $b \in aH$ iff $b = ah$ for some $h \in H$. Therefore, we have the equivalence relation.

$$a \equiv b \pmod{H} \iff a = bh \text{ for some } h \in H \tag{49}$$

Proof.

We show that this indeed forms an equivalence class.
1. *Reflexive.* $a \equiv a \pmod{H}$ since $e \in H \implies a = ae$.
2. *Symmetric.* Let $a \equiv b \pmod{H}$. Then $a = bh$ for some $h \in H$, but since H is a group, $h^{-1} \in H \implies ah^{-1} = b \implies b \equiv a \pmod{H}$.
3. *Transitive.* Let $a \equiv b \pmod{H}$ and $b \equiv c \pmod{H}$. Then $a = bh$ and $b = ch'$ for some $h, h' \in H$. But then
$$a = bh = (ch')h = c(h'h) \tag{50}$$
where $h'h \in H$ due to closure.

Note that a coset is *not* a subgroup. It is only the case that $eH = H$ is a subgroup, but for $a \neq e$, aH does not even contain the identity. We should think of a coset as a *translation* of the subgroup H.

Example 2.1 (Familiar Cosets)

Here are some examples. Note that all it takes is to find *some* subgroup, and the cosets will naturally pop up.
1. Let $H = 2\mathbb{Z} \subset (\mathbb{Z}, +)$ be the even integers. Then $0 + H$ and $1 + H$ are the even and odd integers, respectively.
2. Let $H = \{e, f\} \subset \mathrm{Dih}(3)$. Then
$$H = \{e, f\}, rH = \{r, rf\}, r^2H = \{r^2, r^2f\} \tag{51}$$
are the cosets.

With this partitioning scheme in mind, the following theorem on the order of such groups becomes very intuitive, and has a lot of consequences.

Theorem 2.1 (Lagrange's Theorem)

Let G be a finite group and H its subgroup. Then

$$|G| = [G:H]|H| \tag{52}$$

where $[G:H]$, called the **index of** H, is the number of cosets in G. Therefore, the order of a subgroup of a finite group divides the order of the group.

Proof.

The union of the $[G:H]$ disjoint cosets is all of G. On the other hand, every H is in one-to-one correspondence with each coset aH, so every coset has $|H|$ elements. Therefore, there are $[G:H]|H|$ elements altogether.

Therefore, Lagrange's theorem says that *given* that you find a subgroup, the order of the subgroup must divide the order of G. However, that doesn't mean that such a subgroup may even exist. For example, there is a group of order 12 having no subgroup of order 6.

Corollary 2.1 ()

The order of any element of a finite group divides the order of the group.

Proof.

Take any $a \in G$ and construct the cyclic subgroup $\langle a \rangle \subset G$. Then by Lagrange's theorem, $|a| = |\langle a \rangle|$ divides $|G|$.

Corollary 2.2 ()

Every finite group of a prime order is cyclic.

Proof.

Let $a \in G$ be any element other than the identity e, and consider $\langle a \rangle \subset G$. The order must divide $|G|$ which is prime, so $|a| = 1$ or $|G|$. But $|a| \neq 1$ since we did not choose the identity, so $|a| = |G| \implies \langle a \rangle = G$.

Corollary 2.3 ()

If $|G| = n$, then for every $a \in G$ $a^n = e$.

Proof.

Let $|a| = k$. Then $k \mid n$, and so $a^n = a^{kl} = (a^k)^l = e^l = e$.

Corollary 2.4 (Fermant's Little Theorem)

Let p be a prime number. The multiplicative group $\mathbb{Z}_p \setminus \{0\}$ of the field $\mathbb{Z}_p$ is an abelian group of order $p-1 \implies g^{p-1} = 1$ for all $g \in \mathbb{Z}_p \setminus \{0\}$. So,

$$a^{p-1} \equiv 1 \iff a^p \equiv a \pmod p \tag{53}$$

We can generalize this.

Definition 2.3 (Euler's Totient Function)

Euler's Totient Function, denoted $\varphi(n)$, consists of all the numbers less than or equal to n that are coprime to n.

Theorem 2.2 (Euler's Theorem)

For any n, the order of the group $\mathbb{Z}_n \setminus \{0\}$ of invertible elements of the ring $\mathbb{Z}_n$ equals $\varphi(n)$, where φ is Euler's totient function. In other words with $G = \mathbb{Z}_n \setminus \{0\}$,

$$a^{\varphi(n)} \equiv 1 \pmod n, \text{ where } a \text{ is coprime to } n \tag{54}$$

Example 2.2 ()

In $\mathbb{Z}_{125} \setminus \{0\}$, $\varphi(125) = 125 - 25 = 100 \implies 2^{100} \equiv 1 \pmod{125}$

2.2 Normal Subgroups

By introducing cosets, we have successfully constructed an equivalence relation on G. This set of cosets is indeed a partition of G, but we would like to endow it with a group structure that respects that of G. That is, let $a, b \in G$ and its corresponding cosets be aH, bH. Then, we would like to define an operation $\cdot$ on the cosets such that

$$(aH) \cdot (bH) \coloneqq (ab)H \tag{55}$$

That is, we would like to upgrade the equivalence relation to a *congruence relation*. If we try to show that this is indeed a well-defined operation, we run into some trouble. Suppose $aH = a'H$ and $bH = b'H$. Then with our definition, we should be able to derive that $(aH)(bH) = (a'H)(b'H)$ through the equation

$$(aH)(bH) = (ab)H = (a'b')H = (a'H)(b'H) \tag{56}$$

We have $a' = ah_1$, $b' = bh_2$, and $a'b' = abh$. Then,

$$\begin{aligned} (ab)H = (a'b')H &\implies a'b' = abh \text{ for some } h \in H &(57)\\ &\implies ah_1bh_2 = abh \text{ for some } h_1, h_2, h \in H &(58)\end{aligned}$$

But the final statement is not true in general. In an abelian group, we could just swap h_1 and b to derive it completely, but perhaps there is a weaker condition on just the subgroup H that allows us to "swap" the two.

Definition 2.4 (Normal Subgroups)

A subgroup $N \subset G$ is a **normal subgroup** iff the left cosets equal the right cosets. That is, $\forall g \in G, h \in H$.

$$g^{-1}hg \in H \tag{59}$$

We call $g^{-1}hg$ the **conjugate** of h by g.

Example 2.3 (Normal Subgroups)

For intuition, we provide some examples of normal subgroups.

1. If G is abelian, every subgroup is normal. So $(2\mathbb{Z}, +)$ is normal, and $(\mathbb{Q}, \times) \subset (\mathbb{R}, \times)$ is also normal.
2. Given $G = (\mathbb{R} \setminus \{0\}, \times)$, let $H = (\mathbb{R}^+, \times) \subset G$ be a subgroup. Then H is normal since for any $g \in \mathbb{R}$, g, g^{-1} are either both positive or both negative, and so $ghg^{-1} > 0 \implies ghg^{-1} \in H$. H and $(-1)H$ are two cosets of $\mathbb{R}$.
3. $\mathrm{SL}_n(\mathbb{F}) \subset \mathrm{GL}_n(\mathbb{F})$ is a normal subgroup since the determinant of the inverse is ihe inverse of the determinant, and so for any $g \in \mathrm{GL}_n(\mathbb{F})$,
$$\det(ghg^{-1}) = \det(g)\det(h) = \det(g^{-1}) = \det(g) \cdot 1 \cdot \frac{1}{\det(g)} = 1 \implies ghg^{-1} \in \mathrm{SL}_n(\mathbb{F}) \quad (60)$$
4. The subgroup $H = \{e, r^2\} \subset \mathrm{Dih}(4)$ is a normal subgroup. It is clearly a subgroup isomorphic to Z_2, and to see normality, note that r^2 commutes with any $g = r^n \in \mathrm{Dih}(4)$. If g contains a flip, then we can just check the 4 cases knowing that $fr = r^3 f$.
$$fr^2f^{-1} = fr^2f = (fr)(rf) = r^3frf = r^3r^3f^2 = r^2 \quad (61)$$
$$(rf)r^3(rf)^{-1} = \ldots = r^2 \quad (62)$$
Therefore $\mathrm{Dih}(4)/H$ has order 4, which means it must be isomorphic to either the cyclic group or the Klein 4 group. It turns out it's the Klein 4 group.
5. The subgroup $H = \{e, r, r^2, r^3\} \subset \mathrm{Dih}(4)$ is a normal subgroup because
$$\underbrace{(f^jr^i)}_{g}\underbrace{(r^l)}_{h}\underbrace{(r^{-i}f^{-j})}_{g^{-1}} = f^jri + l - if^{-j} \quad (63)$$
$$= f^jr^lf^{-j} \quad (64)$$
$$= f^jr^lf_j \quad (65)$$
$$= r^{l+3j} \quad (66)$$
where we used the fact that $frf = r^3 = r^{-1}$ in the penultimate step. So $|\mathrm{Dih}(4)/H| = 2 \implies \mathrm{Dih}(4)/H \simeq Z_2$ with generator fh.

Example 2.4 (Subgroups that are Not Normal)

Here are some subgroups that are not normal.

1. Given $G = \mathrm{Dih}(3)$, $H = \{e, f\}$ is not normal since $rfr^{-1} = rfr^2 = r^2 f \notin H$.
2. The subgroup
$$H = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \;\middle|\; ac \neq 0 \right\} \subset \mathrm{GL}_2(\mathbb{R}) \quad (67)$$
is not normal since
$$h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in H, a = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}) \implies aha^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix} \notin H \quad (68)$$

Finally, we present some relevent results of alternating subgroups.

Theorem 2.3 (Alternating Group is Normal in Symmetric)

A_n is a normal subgroup of S_n, of index[a] 2.

[a] i.e. the number of cosets

Proof.

Lemma 2.1 (Cycles in Alternating Group)

We have the following.

1. Every element of A_n can be written as the product of 3-cycles.
2. If $n \geq 4$, H is a normal subgroup of A_n, and H contains one 3-cycle, then $H = A_n$.

Proof.

Since we've proved that every permutation is the product of transpositions, it suffices to prove that the product of two transpositions can be written as the product of 3-cycles. We check this case by case, where distinct symbols represent distinct values.

1. $(\alpha\ \beta)(\gamma\ \delta) = (\alpha\ \beta\ \gamma)(\beta\ \gamma\ \delta)$
2. $(\alpha\ \beta)(\alpha\ \gamma) = (\alpha\ \gamma\ \beta)$
3. $(\alpha\ \beta)(\alpha\ \beta) = e$

Therefore every even permutation is the product of 3-cycles.

Definition 2.5 (Simple Group)

A **simple group** is a group with no proper normal subgroup. That is, the only normal subgroups are the trivial group and itself.

Theorem 2.4 (Alternating Groups are Simple)

For $n \geq 5$, A_n is a simple group.

Proof.

Let $H \subset A_n$ be a normal subgroup containing more than the identity. If we can find a single 3-cycle in H, then it follows from 2.2 that $H = A_n$. Let $\gamma \in H$, $\gamma \neq e$, and write $\gamma = \gamma_1 \ldots \gamma_m$ as a product of disjoint cycles. We have 4 cases.

1. Let $k \geq 4$ and suppose that some factor, say γ_1 is a k-cycle. WLOG let us assume that $\gamma_1 = (1 \ldots k)$. Since H is normal, $(1,2,3)\gamma(1,2,3)^{-1} \in H$ and $(1,2,3)$ commutes with all the factors of γ except γ_1 (since the cycles are disjoint and so γ_i for $i \neq 1$ does not contain $1,2,3$). Thus letting
$$\sigma = (1,2,3)\gamma(1,2,3)^{-1} = (2,3,1,4,\ldots,k)\gamma_2 \ldots \gamma_m \in H \tag{69}$$
since H is a group we have
$$\begin{aligned} \sigma\gamma^{-1} &= \begin{pmatrix} 2 & 3 & 1 & 4\ldots & k \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & \ldots & k \end{pmatrix}^{-1} && (70)\\ &= \begin{pmatrix} 2 & 3 & 1 & 4 & \ldots & k \end{pmatrix} \begin{pmatrix} k & \ldots & 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} && (71) \end{aligned}$$
2. Suppose γ has at least two 3-cycles as factors, say $\gamma_1 = (1,2,3), \gamma_2 = (4,5,6)$. Then
$$\sigma = (3,4,5)\gamma(3,4,5)^{-1} = (1,2,4)(3,6,5)\gamma_3 \ldots \gamma_m \in H \tag{72}$$
and again we have
$$\begin{aligned} \sigma\gamma^{-1} &= \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & 6 & 5 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}^{-1} && (73)\\ &= \begin{pmatrix} 1 & 6 & 3 & 4 & 5 \end{pmatrix} && (74) \end{aligned}$$
which is a 5-cycle, and we are done by case 1.

3. Suppose γ has precisely one 3-cycle factor and all others are transposisions. If the 3-cycle is $\gamma_1 = (1, 2, 3)$, then $\gamma^2 = (1, 2, 3)^2 = (1, 3, 2)$ is a 3-cycle.
4. Suppose γ is the product of disjoint transpositions. Say $\gamma_1 = (1, 2), \gamma_2 = (3, 4)$. Then as before

$$\sigma = \begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \gamma \begin{pmatrix} 1 & 2 & 4 \end{pmatrix}^{-1} \implies \sigma\gamma^{-1} = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \in H \tag{75}$$

Since $n \geq 5$ by our theorem hypothesis, the permutation $\tau = (2, 3, 5) \in A_n$, and so

$$\tau \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \tau^{-1} = \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 5 \end{pmatrix} \in H \tag{76}$$

$$\implies \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 3 \end{pmatrix} \in H \tag{77}$$

2.3 Quotient Groups

Now that we know about normal subgroups, this allows us to endow on the quotient set a group structure.

Definition 2.6 (Quotient Group)

Given a group G and a normal subgroup H, the **quotient group** G/H is the group of left cosets aH with
1. the operation $(aH) \cdot (bH) \coloneqq (ab)H$
2. the identity element eH.
3. inverses $(aH)^{-1}) = (a^{-1})H$.

and order $|G/H| = |G|/|H|$.

Proof.

We verify the properties of a group.
1. Suppose as above that $aH = a'H$ and $bH = b'H$. Then $a' = ah$ and $b' = bk$ for some $h, k \in H$. Since H is normal, $b^{-1}hb = h'$ for some $h' \in H$. Therefore,

$$a'b' = (ah)(bk) = a(hb)k = (abh')k = (ab)(h'k) \in (ab)H \tag{78}$$

and so $(ab)H = (a'b')H$.
2. eH is indeed the identity since $(aH)(eH) = (ae)H = aH$ and $(eH)(aH) = (ea)H = aH$.
3. Inverses are the same logic.
4. Associativity follows from associativity in G.

Finally, by Lagrange's theorem, the order is as stated.

Since the quotient defines a *congruence* class, this makes it a group homomorphism.

Theorem 2.5 (Quotient Maps are Homomorphisms)

The map $p : G \to G/H$ is a group homomorphism.

Proof.

Follows immediately from the definition.

It's a bit hard thinking of an intuitive picture of a normal subgroup. Unless you sit down and try to prove that a subgroup is normal, it's difficult to tell right away. The following lemma characterizes normal subgroups in a different manner.

Lemma 2.2 (Normal Subgroup as Kernel)

A subgroup $H \subset G$ is normal if and only if there exists a group homomorphism $\phi : G \to G'$ with $\ker \phi = H$.

Proof.

We prove bidirectionally.

1. $(\rightarrow)$. Since H is normal, we can form the quotient group G/H. Let $\phi : G \to G/H$ be defined $\phi(a) = aH$. Then,
$$\begin{aligned} \ker \phi = \phi^{-1}(eH) &= \{a \in G \mid aH = eH = H\} && (79) \\ &= \{a \in G \mid a \in H\} && (80) \end{aligned}$$
Therefore, ϕ is a homomorphism because $\phi(ab) = abH = (aH)(bH)$.
2. $(\leftarrow)$ Assume there is a group homomorphism ϕ. Then, $\ker \phi \subset G$ is a subgroup proven in 1.3. Now consider any $g \in G$. Then
$$\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1}) = \phi(g) \cdot e \cdot \phi(g)^{-1} = e \implies ghg^{-1} \in \ker \phi \tag{81}$$

Now that we can construct quotient groups, we would like to see if they are isomorphic to any current groups that we know. More specifically, if we have a normal subgroup $H \subset G$, we can cleverly think of some other group G' and construct a group homomorphism $f : G \to G'$ such that $H = \ker f$. If we can do this, then we can construct a nice isomrophism from G/H to G'. Recall a similar theorem in point set topology: given a topological space $(X, \mathscr{T})$ and its quotient space, if we can construct a map from X to a cleverly chosen space Z that agrees with the quotient, then this induces a homeomorphism $X \cong Z$.

Theorem 2.6 (Fundamental Group Homomorphism Theorem)

Let $f : G \to G'$ be a surjective homomorphism.[a] Then $G/\ker f \simeq G'$.[b]

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, "p"] & G' \\ G/\ker f \arrow[ur, "\bar{f}"'] & \end{tikzcd}$$

Figure 6: Given f and the projection map $p : G \to G/\ker f$, this induces an isomorphism $\bar{f}$ such that $f = \bar{f} \circ p$.

[a] Sometimes called an *epimorphism*.

[b] Note that if f is not surjective, we can just have it be surjective by restricting G' to be the image of f.

Proof.

Let $H = \ker f$, which is then a normal subgroup from 2.3. Now we define a homomorphism
$$\bar{f} : G/H \to G', \qquad \bar{f}(aH) = f(a) \tag{82}$$
We check the following.

1. $\bar{f}$ is well defined. If we have $a, a' \in G$ with $aH = a'H$, then $a' = ah$ for some $h \in H = \ker f$. So $f(a') = f(ah) = f(a)f(h) = f(a)$.

2. $\bar{f}$ is a homomorphism. We see that

$$\begin{aligned}\bar{f}((aH)(bH)) &= \bar{f}((ab)H) && (83)\\ &= f(ab) && (84)\\ &= f(a)f(b) && (85)\\ &= \bar{f}(aH)\bar{f}(bH) && (86)\end{aligned}$$

3. $\bar{f}$ is surjective. This is trivially true since if not, then $f = \bar{f} \circ p$ cannot be surjective.
4. $\bar{f}$ is injective. By 1.3, it suffices to show that $\ker \bar{f}$ is trivial. Suppose $aH \in \ker \bar{f}$. Then $\bar{f}(aH) = f(a) = e_{G'} \implies a \in H \implies aH = eH$.

Example 2.5 (Cyclic Groups)

$(k\mathbb{Z}, +) \subset (\mathbb{Z}, +)$ is a normal subgroup. Our intuition might tell us that the cosets of the form $k\mathbb{Z}, 1 + k\mathbb{Z}, \ldots, (k-1) + k\mathbb{Z}$ behave like integers modulo k, i.e. a cyclic group. Therefore, we can construct the map

$$f : \mathbb{Z} \to Z_k, \qquad f(x) = x \pmod{k} \tag{87}$$

This is a homomorphism and also $\ker f = k\mathbb{Z}$, and so by the fundamental homomorphism theorem

$$\frac{\mathbb{Z}}{k\mathbb{Z}} \simeq Z_k \tag{88}$$

By establishing the connection between the integers and cyclic groups, we establish the notation $Z_k = \mathbb{Z}_k$.

Example 2.6 (Quotient of Reals over Integers)

We can see that $(\mathbb{Z}, +) \subset (\mathbb{R}, +)$ is a normal subgroup. Our intuition might tell us that the cosets (which are disconnected sets consisting of isolated points $\{\ldots, x-1, x, x+1, \ldots\}$) behave sort of like the rotations on a circle S^1. Therefore, let us construct a map

$$f : \mathbb{R} \to S^1, \qquad f(x) = \cos 2\pi x + i \sin 2\pi x \subset \mathbb{C} \tag{89}$$

Since $f(x+y) = f(x)f(y)$, it follows that f is a homomorphism. On the other hand, $\ker f = \{x \in \mathbb{R} \mid \cos 2\pi x = 1, \sin 2\pi x = 0\} = \mathbb{Z}$. Therefore by the fundamental homomorphism theorem, we have

$$\mathbb{R}/\mathbb{Z} \simeq S^1 \tag{90}$$

Example 2.7 (Determinant)

The determinant $\det : \mathrm{GL}_n(\mathbb{F}) \to \mathbb{F}^* = \mathbb{F} \setminus \{0\}$ is a surjective group homomorphism (under multiplication on $\mathbb{F}$). Therefore,

$$\frac{\mathrm{GL}_n(\mathbb{F})}{\mathrm{SL}_n(\mathbb{F})} \simeq \mathbb{F}^* \tag{91}$$

2.4 Orbits and Stabilizers

Definition 2.7 (Orbits)

Let G be a transformation group on set X. Points $x, y \in X$ are equivalent with respect to G if there exists an element $g \in G$ such that $y = gx$. This has already been defined through the equivalence of figures before. This relation splits X into equivalence classes, called **orbits**. Note that cosets are the equivalence classes of the transformation group G; oribits are those of X. We denote it as

$$Gx \equiv \{gx \mid g \in G\} \tag{92}$$

By definition, transitive transformation groups have only one orbit.

Definition 2.8 ()

The subgroup $G_x \subset G$, where $G_x \equiv \{g \in G | gx = x\}$ is called the **stabilizer** of x.

Example 2.8 ()

The orbits of $O(2)$ are concentric circles around the origin, as well as the origin itself. The stabilizer of 0 is the entire $O(2)$.

Example 2.9 ()

The group S_n is transitive on the set $\{1, 2, ..., n\}$. The stabilizer of $k, (1 \leq k \leq n)$ is the subgroup $H_k \simeq S_{n-1}$, where H_k is the permutation group that does not move k at all.

Theorem 2.7 ()

There exists a 1-to-1 injective correspondence between an orbit G_x and the set G/G_x of cosets, which maps a point $y = gx \in Gx$ to the coset gG_x.

Corollary 2.5 ()

If G is a finite group, then

$$|G| = |G_x||Gx| \tag{93}$$

In fact, there exists a precise relation between the stabilizers of points of the same orbit, regardless of G being finite or infinite:

$$G_{gx} = gG_x g^{-1} \tag{94}$$

2.5 Centralizers and Normalizers

2.6 Lattice of Subgroups

3 Group Actions

3.1 Sylow Theorems

4 Classification of Groups

4.1 Direct Products

Definition 4.1 (Direct Product)

The **direct product** of two groups $(G, \cdot)$ and $(H, *)$ is the set

$$G \times H \equiv \{(g, h) \mid g \in G, h \in H\} \tag{95}$$

equipped with the operation

$$(g_1, h_1) \cdot (g_2, h_2) \coloneqq (g_1 \cdot g_2, h_1 * h_2) \tag{96}$$

Proof.

It is pretty trivial to see that this is a group.

Example 4.1 (General Affine Group)

The **general affine group** is defined

$$\mathrm{GA}(V) \equiv \mathrm{Tran}(V) \times \mathrm{GL}(V) \tag{97}$$

Example 4.2 (Galileo Group)

The **Galileo Group** is the transformation group of spacetime symmetries that are used to transform between two reference frames which differ only by constant relative motion within the constructs of Newtonian physics. It is denoted

$$\mathrm{Tran}\,\mathbb{R}^4 \times H \times \mathrm{O}(3) \tag{98}$$

where H is the group of transformations of the form

$$(x, y, z, t) \mapsto (x + at, y + bt, z + ct, t) \tag{99}$$

Example 4.3 (Poincaré Group)

The **Poincaré Group** is the symmetry group of spacetime within the principles of relativistic mechanics, denoted

$$G = \mathrm{Tran}\,\mathbb{R}^4 \times \mathrm{O}_{3,1} \tag{100}$$

where $\mathrm{O}_{3,1}$ is the group of linear transformations preserving the polynomial

$$x^2 + y^2 + z^2 - t^2 \tag{101}$$

4.2 Semidirect Products

4.3 Classification of Finite Abelian Groups

Theorem 4.1 (Groups of Order 1, 2, 3)

We have the following.

1. There is only one group of order 1.
$$Z_1 \simeq S_1 \simeq A_2 \tag{102}$$
2. There is only one group of order 2.
$$Z_2 \simeq S_2 \simeq D_2 \tag{103}$$
3. There is only one group of order 3.
$$Z_3 = A_3 \tag{104}$$

Theorem 4.2 (Groups of Order 4)

There are two groups of order 4.
$$Z_4, \qquad Z_2^2 \simeq D_4 \tag{105}$$

4.4 Group Extensions

4.5 Classification of Simple Groups of Small Order

Theorem 4.3 (Classification of Simple Groups of Small Order)

The following are the only groups of order n. You can notice that it is dominated by direct products of cyclic groups, since they exist for every order, while the other types increase in order very fast.

n	**Abelian Groups**	**Non-Abelian Groups**
1	$\{e\}$ (trivial group)	None
2	$\mathbb{Z}_2 = S_2 = \text{Dih}(1)$	None
3	$\mathbb{Z}_3 = A_3$	None
4	$\mathbb{Z}_4$, $\mathbb{Z}_2 \times \mathbb{Z}_2 = \text{Dih}(2)$	None
5	$\mathbb{Z}_5$	None
6	$\mathbb{Z}_6 = \mathbb{Z}_3 \times \mathbb{Z}_2$	$S_3 = \text{Dih}(3)$
7	$\mathbb{Z}_7$	None
8	$\mathbb{Z}_8$, $\mathbb{Z}_4 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$	$D_4 = \text{Dih}(4)$, Q_8 (quaternion)
9	$\mathbb{Z}_9$, $\mathbb{Z}_3 \times \mathbb{Z}_3$	None
10	$\mathbb{Z}_{10} = \mathbb{Z}_5 \times \mathbb{Z}_2$	$D_5 = \text{Dih}(5)$
11	$\mathbb{Z}_{11}$	None
12	$\mathbb{Z}_{12} = \mathbb{Z}_4 \times \mathbb{Z}_3$, $\mathbb{Z}_6 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3$	A_4, $D_6 = \text{Dih}(6)$, $\mathbb{Z}_3 \rtimes \mathbb{Z}_4$ (dicyclic)

Figure 7: Classification of groups up to order 12.

5 Ring-Like Structures

We have extensively talked about groups, and now we look at an algebraic structure called a ring that has two operations. As we introduce rings, we will use the integers as the primary structure to demonstrate our theorems, along with the ring of continuous functions and the ring of matrices.

Definition 5.1 (Ring)

A **ring** is a set $(R, +, \times)$ equipped with two operations, called addition and multiplication. It has properties:

1. R is an abelian group with respect to $+$, where we denote the additive identity as 0 and the additive inverse of x as $-x$.
2. R is a monoid with respect to $\times$, where we denote the multiplicative identity as 1, also known as the **unity**.
3. $\times$ is both left and right distributive with respect to addition $+$

$$a \times (b + c) = a \times b + a \times c \tag{106}$$
$$(a + b) \times c = a \times c + b \times c \tag{107}$$

for all $a, b, c \in \mathbb{R}$.

If $\times$ is associative, R is called an **associative ring**, and if $\times$ is commutative, R is called a **commutative ring**.

In fact, in some cases the existence of the multiplicative identity is not even assumed, though we will do it here.[2] Since a ring is a group with respect to addition, we know from 1.2 that additive inverses are unique. However, we can say a little more with rings because of the distributive property.

Lemma 5.1 (Additive Inverses)

For any $a \in R$, $-a = -1 \times a$.

Proof.

We can see that

$$-1 + 1 = 0 \implies (-1 + 1) \times a = 0 \times a \tag{108}$$
$$\implies -1 \times a + 1 \times a = 0 \tag{109}$$
$$\implies -1 \times a + a = 0 \tag{110}$$

and therefore by definition $-1 \times a$ must be the additive inverse.

It helps to see some familiar examples of rings first before examining their properties.

Example 5.1 (Integers, Rationals, Reals, Complexes)

$(\mathbb{Z}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$, $(\mathbb{C}, +, \times)$ are all commutative rings, with additive and multiplicative identities 0 and 1.

Example 5.2 (Matrices)

The set of matrices $\mathbb{R}^{n \times n\, a}$ forms a noncommutative ring under matrix addition $+$ and multiplication $\times$. It has the additive and multiplicative identities 0 and I_n. This forms a non-commutative ring for

[2] If a multiplicative identity is not assumed, then this is called an *rng*, or a *rung*.

$n > 1$, even when R is commutative.

[a] really over any field and even more generally a ring R

Example 5.3 (Continuous Functions)

The set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$ is a ring under point-wise addition and multiplication.

Example 5.4 (Power Set)

Given a set X, $(2^X, \triangle, \cap)$ is a commutative associative ring with respect to the operations of symmetric difference $M \triangle N := (M \setminus N) \cup (N \setminus M)$ and intersection. The additive identity is $\emptyset$ and the multiplicative identity is X. We can clearly see that both operations are commutative and $\cap$ is associative.

$$\begin{aligned} M \triangle N &= (M \setminus N) \cup (N \setminus M) \equiv N \triangle M \\ M \cap N &= N \cap M \\ M \cap N \cap P &= (M \cap N) \cap P = M \cap (N \cap P) \end{aligned}$$

Next, just like how we did for groups, we can talk about subrings.

Definition 5.2 (Subring)

Given ring $(R, +, \times)$ a **subgring** $(S, +, \times)$ is a ring such that $S \subset R$. S is called a **proper subring** if $S \subsetneq R$.

Theorem 5.1 ()

If S_1, S_2 are subrings of R, then $S_1 \cap S_2$ is a subring.

Note that we do not assume that there exists multiplicative inverses in a ring. However, there may be some elements for which multiplicative inverses do exist, i.e. $a, b \in R$ where $ab = 1$.

Definition 5.3 (Unit)

A **unit** of a ring R is an element $u \in R$ that has a multiplicative inverse in R. That is, there exists a $v \in R$ s.t. $uv = vu = 1$.

Another property that we would desire is some sort of decomposition of ring elements as other ring elements. More specifically, the existence of elements a, b such that $ab = 0$ will be of particular interest to us.

Definition 5.4 (Left, Right Divisor)

Let $a, b, r \in R$ a ring.

1. If $ab = r$, then a is said to be a **left divisor** of r and b a **right divisor** of r.
2. a is said to be a left divisor of r if it is a left divisor and a right divisor of r: $ax = ya = r$, but x does not necessarily equal y.
3. If $ab = 0$, then a and b are said to be a **left zero divisor** and **right zero divisor**, respectively.

If R is commutative, then we just call a a **divisor** of r or a **zero divisor**.[a]

[a] a is a right divisor of $b \iff \exists x(xa = b) \iff \exists x(ax = b) \iff a$ is a left divisor.

It turns out that the existence of units and zero divisors classify rings into subcategories, which we will elaborate on. That is, we will start with the most general theory on rings, and then shrink down into subcategories of rings.

$$\begin{tikzcd}
 & \text{Rings} \arrow[dl] \arrow[dr] & \\
\text{Commutative Rings} \arrow[d] & & \text{Domains} \arrow[dll] \arrow[ddd] \\
\text{Integral Domains} \arrow[d] & & \\
\text{PIDs} \arrow[d] & & \\
\text{Euclidean Domains} \arrow[dr] & & \text{Division Rings} \arrow[dl] \\
 & \text{Fields} &
\end{tikzcd}$$

Figure 8: Basic hierarchy of rings.

5.1 Ring Homomorphisms

So far, we have talked about many properties of rings but have not thoroughly gone over their classification. This is what we will do in this section, just like how we have classified groups. It turns out that classifying rings is significantly harder to do so, so we will talk about some low-order finite rings and provide some examples of isomorphisms between more complex rings.

Definition 5.5 (Ring Homomorphism, Isomorphism)

A **ring homomorphism** $f : R \to S$ is a function that satisfies for all $a, b \in R$

1. $f(a + b) = f(a) + f(b)$
2. $f(ab) = f(a)f(b)$
3. $f(1_R) = 1_S$ [a]

for all $a, b \in R$.[b] Furthermore,

1. A **ring isomorphism** is a bijective ring homomorphism, and we call rings R and S isomorphic, denoted $R \simeq S$ if there exists an isomorphism between them.
2. A **ring endomorphism** is a ring homomorphism onto itself.
3. A **ring automorphism** is an isomorphism from a ring to itself.

[a] The reason we need this third is that while f is a group homomorphism with respect to $+$, it automatically follows that $f(0) = 0$. However f is only a monoid homomorphism w.r.t. $\times$, and so we need this extra constraint.

[b] Note that the first is equivalent to it being a group homomorphism between $(R, +)$ and $(S, +)$. The second property may look like it is a group homomorphism between $(R, \times)$ and $(S, \times)$, but remember that neither are groups and it just states that closure distributes. Combined with the fact that the multiplicative identity matches, f is really a homomorphism of *monoids*.

Example 5.5 (Homomorphisms of Rings)

We provide some simple examples of ring homomorphisms.

1. The identity map $\iota : R \to R$ is a ring homomorphism.
2. If $R \subset S$ as rings, then the canonical injection map $\iota : R \to S$ is a ring homomorphism.
3. Complex conjugation $z \in \mathbb{C} \mapsto \bar{z} \in \mathbb{C}$ is a ring automorphism.

Definition 5.6 (Kernel)

The **kernel** of a ring homomorphism $f : R \to S$ is the preimage of $0 \in S$.[a]

[a] Note that this is the additive identity, not the multiplicative identity. We must specify which identity, unlike a group which has just one identity.

Lemma 5.2 (Images and Kernels of Ring Homomorphisms)

If $f : R \to S$ is a ring homomorphism, then

1. $\ker f$ is a subring of R.
2. $\operatorname{Im} f$ is a subring of S.
3. f is injective iff $\ker f = \{0\}$.

Proof.

Theorem 5.2 (Compositions of Ring Homomorphisms)

Compositions of ring homomorphisms are ring homomorphisms.

Proof.

Now let's focus a bit more on ring isomorphisms. The following should be intuitive.

Lemma 5.3 (Properties of Ring Isomorphisms)

If f is a ring isomorphism, then f^{-1} is a ring isomorphism.

5.2 Commutative Rings

Remember that for commutative rings, distinguishing left and right divisors are meaningless, and so we can talk about just *divisors*. Almost all rings that we will deal with are commutative, so let's try to find some properties of commutative rings.

Definition 5.7 (Prime and Compositive Elements)

In a commutative ring R, an element $p \in R$ is said to be **prime** if it is not 0, not a unit, and has only divisors 1 and p.

Lemma 5.4 (Euclid)

If p is prime, then $p \mid ab \implies p \mid a$ or $p \mid b$.

Lemma 5.5 ()

Let R be a commutative ring and $a, b, d \in R$. If $d \mid a$ and $d \mid b$, then $d \mid (ma + nb)$ for any $m, n \in R$.

Definition 5.8 (Greatest Common Divisor)

The **greatest common divisor** of elements a and b, denoted $\gcd(a, b)$ of an commutative ring R is a common divisor of a and b divisible by all their common divisors. That is, it is the element $d \in R$ satisfying

1. $d \mid a$ and $d \mid b$
2. if $k \mid a$ and $k \mid b$, then $k \mid d$.

If $\gcd(a, b) = 1$, then a and b are said to be **relatively prime**.

Note that in an arbitrary commutative ring, the gcd of two elements always exists since we can at least identify 1, but there may not be a *unique* gcd.

5.3 Domains

We can see that domains behave similarly to the integers, but with the missing property that $\times$ is commutative. This motivates the following definition of an integral domain, which can be seen as a generalization of the integers.

Definition 5.9 (Domain, Integral Domain)

A ring R with no zero divisors for every element is called a **domain**. An **integral domain** is a commutative domain R.[a]

[a]Almost always, we work with integral domains so we will default to this.

Example 5.6 (Domains vs Integral Domains)

We show some examples of domains and integral domains.

1. $(\mathbb{Z}, +, \times)$ is an integral domain
2. $(\mathbb{Q}, +, \times)$ is an integral domain.
3. $(\mathbb{R}, +, \times)$ is an integral domain.
4. Quaternions $\mathbb{H}$ are not commutative but are a domain.

Example 5.7 (Non-Domains)

Here are some examples of non-domains.

1. The ring of $n \times n$ matrices over any nonzero ring when $n \geq 2$ is not a domain. Given matrices A, B, if the image of B is in the kernel of A, then $AB = 0$.
2. The ring of continuous functions on the interval is not a domain. To see why, notice that given the piecewise functions
$$f(x) = \begin{cases} 1 - 2x & x \in [0, \frac{1}{2}] \\ 0 & x \in [\frac{1}{2}, 1] \end{cases}, \quad g(x) = \begin{cases} 0 & x \in [0, \frac{1}{2}] \\ 2x - 1 & x \in [\frac{1}{2}, 1] \end{cases} \tag{111}$$
$f, g \neq 0$, but $fg = gf = 0$.
3. A product of two nonzero commutative rings with unity $R \times S$ is not an integral domain since $(1, 0) \cdot (0, 1) = (0, 0) \in R \times S$.

Here is an alternative equivalent characterization of an integral domain.

Definition 5.10 (Regular Elements)

An element r of a ring R is **regular** if the mapping

$$\rho : R \longrightarrow R, \qquad x \mapsto xr \tag{112}$$

is injective for all $x \in R$.

Theorem 5.3 (Integral Domains w.r.t. Regularity)

An integral domain is a commutative associative ring where every element is regular.

Finally, we talk about the properties of integral domains. Namely, that the characteristic must be prime and that gcd's—while not yet unique—are now guaranteed to be *associated*.

Theorem 5.4 (Characteristic of an Integral Domain)

The characteristic of an integral domain is either 0 or a prime number.

While we have shown that gcd's exist in commutative rings, we can say a bit more when working in Euclidean domains.

Definition 5.11 (Associate Elements)

Elements a and b are **associated**, denoted $a \sim b$ if either of the following equivalent conditions holds

1. $a|b$ and $b|a$
2. $a = cb$, where c is invertible

The two conditions are equivalent because c and c^{-1} are both in A.

Theorem 5.5 (GCD's in a Euclidean Domain)

Any two distinct gcd's of a, b in a Euclidean domain must be associate elements.

5.4 Ideals

Now assuming that R and S are commutative rings, let's consider a special sort of subset of a commutative ring. Consider the kernel of the ring homomorphism. We can see that if $a, b \in \ker(f)$, then $f(a+b) = f(a) + f(b) = 0 + 0 = 0$, and so $\ker(f)$ is closed under addition. Furthermore, $a \in \ker(f)$ and *any* $b \in R$ gives $f(ab) = f(a)f(b) = 0f(b) = 0$, and so multiplying any element in the kernel by an arbitrary element in the rings keeps it in the kernel. We would like to generalize these properties into an *ideal*.

Definition 5.12 (Ideals)

For a commutative ring $(R, +, \times)$, a **two-sided ideal**—or **ideal**—is a subset $I \subset R$ satisfying

1. $(I, +)$ is a subgroup of $(R, +)$.
2. $a \in I, r \in R \implies ra = ar \in I$.[a]

If R is not necessarily commutative, then we $ra \neq ar$ in general, so we may distinguish between left and right ideals.

[a] Note that this property and closure under addition actually implies that it is a subgroup. Since we can see that $-1 \in R$ and $a \in I$ implies $-1 \cdot a = -a \in I$.

Therefore, we can see that it is an abelian group under $+$ and closed under $\times$. However, it is not guaranteed to have a multiplicative identity, which is why we can interpret I as a ring without a multiplicative identity, also known as a *rung*.

This seems like a pretty abstract definition, but a good intuition to have—though not completely accurate—is that ideals are a collection of *multiples* of a certain element.[3] They are analogous to normal subgroups, which were used to induce a congruence relation on a group to get its quotient. Ideals play a similar role.

Example 5.8 (Multiples of Elements Are an Ideal)

We give 2 ideals:
1. The set of even integers $2\mathbb{Z}$ is an ideal in the ring $\mathbb{Z}$, since the sum of any even integers is even and the product of any even integer with an integer is an even integer. However, the odd integers do not form an ideal.
2. The set of all polynomials with real coefficients which are divisible by the polynomial $x^2 + 1$ is an ideal in the ring of all polynomials.

Let's talk about a few more properties of ideals, namely their construction and behavior under set theoretic operations.

Theorem 5.6 (Sum and Intersection of Ideals are Ideals)

Given two ideals $I, J \subset R$,
1. $I \cap J$ is an ideal.
2. $I + J \coloneqq \{i + j \mid i \in I, j \in J\}$ is an ideal.

Proof.

Listed.
1. $I \cap J$ is an ideal. Given $a, b \in I \cap J$, then $a, b \in I \implies a + b \in I$, and $a, b \in J \implies a + b \in J$. So $a + b \in I \cap J$. Furthermore, for every $r \in R$, $a \in I \implies ra \in I$ and $a \in J \implies ra \in J$, so $a \in I \cap J \implies ra \in I \cap J$.
2. $I + J$ is an ideal. Given $x, y \in I + J$, then $x = a_x + b_x$ and $y = a_y + b_y$ for $a_x, a_y \in I, b_x, b_y \in J$. So
$$x + y = (a_x + b_x) + (a_y + b_y) = (a_x + a_y) + (b_x + b_y) \tag{113}$$
where $a_x + a_y \in I, b_x + b_y \in J$ by definition of an ideal, and so $x + y \in I + J$. Noe let $x = a_x + b_x \in I + J$. Then given $r \in R$,
$$rx = r(a_x + b_x) = ra_x + rb_x \tag{114}$$
where $ra_x \in I$ and $rb_x \in J$ since I, J are ideals. Therefore $rx \in I + J$.

Theorem 5.7 (Preimage of Ideals are Ideals)

If $f : R \to S$ is a ring homomorphism of commutative rings $J \subset S$ is an ideal, then $f^{-1}(J)$ is an ideal of R.

Proof.

[3] This is actually more accurate for a principal ideal.

Example 5.9 (Image of Ideal is Not Necessarily an Ideal)

It is not true in general that for an ideal $I \subset R$ and a ring homomorphism $f : R \to S$, the image $f(I)$ is an ideal of S.

Given the two examples above, let's formalize the idea of an ideal consisting of all multiples of a specific element a. This sounds pretty familiar to *generators* of groups.

Definition 5.13 (Generators of Ideals)

Given a commutative ring R, the **ideal generated by** $a \in R$ is denoted

$$\langle a \rangle := \{ra \mid r \in R\} \tag{115}$$

and more generally, we may have multiple generating elements.

$$\langle a_1, \ldots, a_n \rangle := \{r_1 a_1 + \ldots r_n a_n \mid r_1, \ldots, r_n \in R\} \tag{116}$$

Therefore, the ideals considered above can be written $\langle 2 \rangle \subset \mathbb{Z}$ and $\langle x - 2 \rangle \subset \mathbb{Q}[x]$. However, it may be the case that two elements generate the same ideal in a non-Euclidean domain, but constructing such an example is a bit challenging.

Example 5.10 (Matrix with Last Row of Zeros)

Let R be the set of all $n \times n$ matrices. Then
1. The set of all $n \times n$ matrices whose last row is zero forms a right ideal, but not a left ideal.
2. The set of all $n \times n$ matrices whose last column is zero is a left ideal, but not a right ideal.

5.5 Quotient Rings

What is nice about ideals is that they induce not just an equivalence relation—but a congruence relation—on a ring, which is a generalization of working in the integers modulo n.

Theorem 5.8 (Equivalence Relation Induced by an Ideal)

Given a commutative ring R and an ideal $I \subset R$, we say that two elements $a, b \in R$ are **congruent** (mod I), written $a \equiv b \pmod{I}$ iff $a - b \in I$. We claim two things:
1. $\equiv$ is an equivalence relation.
2. $\equiv$ is a congruence relation. Given that $a \equiv a' \pmod{I}$ and $b \equiv b' \pmod{I}$,

$$a + b \equiv a' + b' \pmod{I}, \qquad ab \equiv a'b' \pmod{I} \tag{117}$$

Occasionally, if the ideal I is clear from context, we will write $a \equiv b$.

Proof.

We first prove that $\equiv$ is indeed an equivalence relation.
1. *Reflexive.* $a \equiv a \pmod{I}$ is trivial since $a - a = 0 \in I$.
2. *Symmetric.* If $a \equiv b$, then $a - b \in I \implies -(a - b) = -a + b = b - a \in I \implies b \equiv a$.
3. *Transitive.* If $a \equiv b$ and $b \equiv c$, then $a - b \in I$ and $b - c \in I$. Since I is an additive group and so it is closed under addition, so $(a - b) + (b - c) = a - c \in I \implies a \equiv c$.

Note that so far, we have only used the group property of ideals to prove that is is an equivalence class. Now for congruence of multiplication, we need the ring properties.

1. $a \equiv a', b \equiv b' \implies (a - a'), (b - b') \in I$. By adding them together and distributivity, we have
$$a - a' + b - b' = (a + b) - (a' + b') \in I \implies a + b \equiv a' + b' \pmod{I} \tag{118}$$
2. We see that $a \in R, (b-b') \in I \implies a(b-b') \in I$. Similarly, $b' \in R, (a-a') \in I \implies (a-a')b' \in I$. Now adding the two, we have
$$a(b - b') + (a - a')b' = ab - ab' + ab' - a'b' = ab - a'b' \in I \implies ab = a'b' \pmod{I} \tag{119}$$

This quotient space maintains a lot of nice properties of the algebraic operations, and so we can form a new ring structure with this quotient space.

Definition 5.14 (Quotient Rings, Rings of Residue Class)

The quotient space R/I induced by the mapping $a \mapsto [a]$ is indeed a commutative ring, called the **quotient ring**, with addition and multiplication defined
$$[a] + [b] := [a + b], \qquad [ab] := [a]\,[b] \tag{120}$$

Proof.

Note that the properties of the operation in $\frac{M}{R}$ inherits all the properties of the addition operation on M that are expressed in the form of identities and inverses, along with the existence of the zero identity.
$$\begin{aligned} 0 \in M &\implies [0] \text{ is the additive identity in } \frac{M}{R} \\ a + (-a) = 0 &\implies [a] + [-a] = [0] \\ 1 \in M &\implies [1] \text{ is the multiplicative identity in } \frac{M}{R} \end{aligned}$$

Theorem 5.9 (Quotient Maps are Homomorphisms)

The map $p : R \to R/I$ is a ring homomorphism.

Proof.

This is true by definition since we have made $\equiv$ a congruence relation.

Example 5.11 (Quotient Rings of Integers)

The quotient set $\mathbb{Z}/\langle n \rangle$ by the relation of congruence modulo n is denoted $\mathbb{Z}_n$.
$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\} \tag{121}$$

Note that the quotient ring $(\mathbb{Z}/\langle n \rangle, +, \times)$ is precisely the cyclic quotient group $\mathbb{Z}_n = \mathbb{Z}/6\mathbb{Z}$ when considering only addition. We list some quotient rings of the integers.

1. In $\mathbb{Z}_5 = \mathbb{Z}/\langle 5 \rangle$, the elements $[2]$ and $[3]$ are multiplicative inverses of each other since $[2][3] = [6] = [1]$, and $[4]$ is its own inverse since $[4][4] = [16] = [1]$. The addition and multiplication tables for $\mathbb{Z}_5$ is shown below.
2. Consider the ideal $I = \langle 2 \rangle \subset \mathbb{Z}_6$. We have $0 \equiv 2 \equiv 4 \pmod{I}$ and $1 \equiv 3 \equiv 5 \pmod{I}$, and so the quotient ring $\mathbb{Z}_6/I$ consists of the two equivalence classes $[0]$ and $[1]$.

Example 5.12 (Quotient Rings of Polynomials)

We list some quotient rings of polynomials.

1. Consider $\mathbb{Q}[x]/\langle x^2-2\rangle$. We can see that any polynomial $f \in \mathbb{Q}[x]$ is equivalent $\pmod{I}$ to a linear polynomial, since $x^2 \equiv 2$. Alternatively we can apply the division algorithm to replace $f(x)$ by its remainder upon division by x^2-2, and thus in the quotient ring, $[x]$ plays the role of $\sqrt{2}$, which may indicate that $\mathbb{Q}[x]/\langle x^2-2\rangle = \mathbb{Q}[\sqrt{2}]$.
2. Consider $\mathbb{Z}_2[x]/\langle x^2+x+1\rangle$. As in the previous example, any polynomial in $\mathbb{Z}_2[x]$ is equivalent to a linear polynomial since $x^2 \equiv x+1 \pmod{I}$. Therefore the elements of the quotient ring are $[0],[1],[x],[x+1]$ with the addition and multiplication tables.

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

(a)

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

(b)

Note that just like how quotient topologies do not preserve topological properties, as shown here and here, quotient rings inherit some—but not all—algebraic properties.

Theorem 5.10 (Quotient Inherits Commutativity)

Let R be a commutative ring and $I \subsetneq R$ be an ideal. Then R/I is a commutative ring.

Example 5.13 (Quotient Does Not Inherit Integral Domain Property)

$\mathbb{Z}$ is an integral domain, but $\mathbb{Z}/\langle 6\rangle$ is not since $[2]\times[3]=[0]$.

Just like in group theory, we have a method of constructing isomorphisms between cleverly chosen rings S and a quotient ring R/I. This seems to be a common pattern here when considering groups, rings, and topological spaces... This will be investigated more in category theory.

Theorem 5.11 (Fundamental Ring Homomorphism Theorem)

Let R and S be commutative rings, and suppose $f : R \to S$ be a surjective ring homomorphism. Then this induces a ring isomorphism

$$R/\ker f \simeq S \tag{122}$$

satisfying $\phi = \bar{\phi}\circ\pi$.

$$\begin{tikzcd} R \arrow[r, "\phi"] \arrow[d, "\pi"] & S \\ R/\ker(\phi) \arrow[ur, "\bar{\phi}"] & \end{tikzcd}$$

Figure 10: The theorem states that the following diagram commutes.

Proof.

A direct application of this is the Chinese remainder theorem.

Corollary 5.1 (Chinese Remainder Theorem)

Given a commutative ring R, let $I, J \subset R$ be ideals such that $I + J = R$. Then,

$$\pi : R \to \frac{R}{I} \times \frac{R}{J}, \qquad r \mapsto ([r]_I, [r]_J) \tag{123}$$

with component-wise quotient mappings is a surjective ring homomorphism with $\ker \pi = I \cap J$. By the fundamental ring homomorphism theorem, it immediately follows that

$$\frac{R}{I \cap J} \simeq \frac{R}{I} \times \frac{R}{J} \tag{124}$$

Proof.

Since $I + J = R$, there exists $i \in I$ and $j \in j$ s.t. $i + j = 1$. Let $\bar{a} = a + I \in R/I$ and $\bar{b} = b + J \in R/J$ be any elements. Then

$$\pi(aj + bi) = ([aj + bi]_I, [a_j + bi]_J) = ([aj]_I, [bi]_J) \in \frac{R}{I} \times \frac{R}{J} \tag{125}$$

But we have

1. $a(j + i) = a \in R \implies aj = a(j + i) \in R/I$. Therefore $[aj]_I = [a]_I$
2. $b(j + i) = b \in R \implies bi = bj + bi \in R/J$. Therefore $[b]_J = [bi]_J$.

Therefore, we have $\pi(aj + bi) = ([a]_I, [b]_J)$, which proves surjectivity.

Example 5.14 (Chinese Remainder Theorem on Integers)

5.6 Principal Ideal Domains

A good intuition to have about ideals is that they are the set of multiples of a certain element. However, this may not be true for ideals in general, but if this intuition is true, then we call this a *principal ideal*.

Definition 5.15 (Principal Ideals)

Given commutative ring R and $I \subset R$, if $I = \langle a \rangle$ for some $a \in R$—i.e. it is generated by a single element—I is called a **principal ideal**.

Definition 5.16 (Principal Ideal Domain)

A **principal ideal domain**, also called a **PID**, is an integral domain in which every ideal is principal.

So a principal ideal domain is an integral domain by definition. It may seem that PIDs are an oddly specific structure to be studying separately, but this actually turns out to unlock a lot more nice properties that we are familiar with. The first is that GCDs are now unique, which is great. Second, we have Bezout's identity, saying that if x and y are elements of a PID without common divisors, then every element of the PID can be written in the form $ax + by$. Finally, and most importantly, any element of a PID has a unique decomposition into irreducible factors. We now introduce some examples of PIDs, which are not as trivial and should be introduced as theorems.

Theorem 5.12 (Integers and Polynomials over Fields are PIDs)

The following are all examples of principal ideal domains.
1. Any field $\mathbb{F}$.
2. The ring of integers $\mathbb{Z}$.
3. $\mathbb{F}[x]$, rings of polynomials in one variable with coefficients in a field $\mathbb{F}$.

Proof.

Listed.
1. It is quite easy to see that a field $\mathbb{F}$ is a PID since the only two possible ideals are $\{0\}$ and $\mathbb{F}$, both of which are principal.
2. If $I \subset \mathbb{Z}$ is an ideal, then if $I = \langle 0 \rangle$, then we're done. Otherwise, let $a \in I$ be the smallest positive integer in I. It is clear that $\langle a \rangle \subset I$. Now given an element $b \in I$, by the Euclidean algorithm we have $b = aq + r$ with $r < a$. Since $a, b \in I$, it follows that $r \in I$. But since $0 \leq r < a$ and a is the smallest positive integer, $r = 0$, and so $b = aq \implies b \in \langle a \rangle$.
3. The ring of polynomials $\mathbb{F}[x]$ is a PID since we can imagine a minimal polynomial p in each ideal I. Every element in I must be divisible by p, which means that the entire ideal I can be generated by the minimal polynomial p, making I principal.

Corollary 5.2 (Ideals Generated by Primes)

If $I \subsetneq \mathbb{Z}$ and a prime number $p \in I$, then $I = \langle p \rangle$. If $I \subset F[x]$ is an ideal and irreducible $f(x) \in I$, then $I = \langle f(x) \rangle$.

Proof.

Listed.
1. Since $\mathbb{Z}$ is a PID, $I = \langle a \rangle$ for some nonzero $a \in \mathbb{Z}$. We can assume a is positive, and if $a = 1$, then $I = \mathbb{Z}$, which contradicts the I is a proper subset. So $a \geq 2$. Now because $p \in I$, $p = ra$ for some $r \in \mathbb{Z}$, but since p is prime, $r = 1, a = p$.
2. Since $F[x]$ is a PID and $I = \langle g(x) \rangle$ for some $g(x) \in F[x]$, let us take $f(x) \in I$. Then it must be true that $f(x) = g(x)h(x)$ for some $h(x) \in R$. However, This means that $\deg(g)$ or $\deg(h)$ must be 0 since f is irreducible. But if $g(x)$ was a constant, then $I = R$, so $g(x) = f(x)$.

Corollary 5.3 (Kernel of Evaluation Homomorphism is Generated by Irreducible Factor)

Suppose $f(x) \in F[x]$ is irreducible in $F[x]$, and $K \supset F$ is a field containing a root α of $f(x)$. Then the ideal of all polynomials in $F[x]$ vanishing at α is generated by $f(x)$. That is, given the evaluation homomorphism

$$\mathrm{ev}_\alpha : F[x] \to K \tag{126}$$

we claim $\ker(\mathrm{ev}_\alpha) = \langle f(x) \rangle$.

Proof.

This is an immediate consequence of the previous corollary.

Theorem 5.13 (Greatest Common Divisor is Unique in PIDs)

Given $a, b \in R$ a PID, $\gcd(a, b)$ is unique.

Theorem 5.14 (Bezout's Theorem)

Given that one divides (with remainder) polynomial f by $g = x - c$, let the remainder be $r \in F$. That is,

$$f(x) = (x - c)q(x) + r, \; r \in F \tag{127}$$

This implies that the remainder equals the value of f at point c. That is,

$$f(c) = r \tag{128}$$

Note that a corollary of this is the single factorization theorem, but the single factorization holds for commutative rings in general.

Note that Bezout's does not hold in integral domains in general.

Example 5.15 (Counterexample in Integral Domains but not PIDs)

Theorem 5.15 (Unique Factorization Theorem)

Every element $x \in R$ of a PID can be uniquely factored (up to permutations and units) into irreducible elements in R.

5.7 Euclidean Domains

We have seen that PIDs unlock a lot of familiar properties that we see in integers. In fact, pretty much everything holds except for the existence of Euclidean algorithm for factorization, which turns out to be extremely powerful.

Definition 5.17 (Euclidean Domain)

Let R be an integral domain which is not a field. R is **Euclidean domain** if

1. there exists a *norm* $|\cdot| : R \setminus \mathbb{R}_0^+$, and
2. there exists a well-defined function, called **Euclidean division** $\mathcal{D} : R \times R \rightarrow R \times R$ that is defined
$$\mathcal{D}(a, b) = (q, r) \text{ where } a = bq + r \text{ and } 0 \leq r < |b| \tag{129}$$

The two prime examples are the integers and polynomials.

Example 5.16 (Integers)

$\mathbb{Z}$ is a Euclidean domain with Euclidean division, also called long division, defined

$$\begin{array}{r}
40 \\
13\overline{)521} \\
\underline{52} \\
01
\end{array}$$

Theorem 5.16 (Polynomials are Euclidean Domains)

Let $f(x), g(x) \in F[x]$ and $g(x) \neq 0$. Then, there exists polynomials $q(x), r(x)$ such that

$$f(x) = q(x)g(x) + r(x), \qquad 0 \leq \deg(r) < \deg(g) \tag{130}$$

where deg is the norm.

Example 5.17 (Gaussian Integers)

The subring of $\mathbb{C}$, defined
$$\mathbb{Z}[i] \equiv \{a + bi \mid a, b \in \mathbb{Z}\} \tag{131}$$
is a Euclidean integral domain with respect to the norm
$$N(c) \equiv a^2 + b^2 \tag{132}$$
since $N(cd) = N(c)N(d)$ and the invertible elements of $\mathbb{Z}[i]$ are $\pm 1, \pm i$.

Example 5.18 (Dyadic Rationals)

The ring of rational numbers of the form $2^{-n}m,\ n \in \mathbb{Z}_+, m \in \mathbb{Z}$, is a Euclidean domain. To define the norm, we can first assume that m can be prime factorized into the form
$$m = \pm \prod_i p_i^{k_i},\ p \text{ prime} \tag{133}$$
and the norm is defined
$$N(\frac{m}{2^n}) \equiv 1 + \sum_i k_i \tag{134}$$
We must further show that division with remainder is possible, but we will not show it here.

5.8 Characteristics

Note that given a ring R, we can pay attention to the subring $\langle 1 \rangle$. This must either be isomorphic to $\mathbb{Z}$ or $\mathbb{Z}_n$, so we can think of it being embedded in R.

Theorem 5.17 (Integer Ring Exists in Any Ring)

For every ring R, there exists a unique ring homomorphism $f : \mathbb{Z} \to R$.

Proof.

We know that $f(1_\mathbb{Z}) = 1_R$, and so for $n > 0$,
$$\begin{aligned} f(n_\mathbb{Z}) &= f(1_\mathbb{Z} + \ldots + 1_\mathbb{Z}) && (135)\\ &= f(1_\mathbb{Z}) + \ldots + f(1_\mathbb{Z}) && (136)\\ &= 1_R + \ldots + 1_R && (137)\\ &= n_R && (138)\end{aligned}$$
Similarly, we have
$$\begin{aligned} f(-n_\mathbb{Z}) &= f(-1_\mathbb{Z} - \ldots - 1_\mathbb{Z}) && (139)\\ &= f(1_\mathbb{Z}) - \ldots - f(1_\mathbb{Z}) && (140)\\ &= -1_R - \ldots - 1_R && (141)\\ &= -n_R && (142)\end{aligned}$$
Since $\mathbb{Z}$ is a PID, $\ker f$—which is an ideal—must be principal, and so $\ker f = \langle m \rangle$ for some $m \in \mathbb{Z}$.

Therefore, this motivates the following attribute of a ring, i.e. the smallest $\langle m \rangle$ that embeds (an injective homomorphism) into the ring.

Definition 5.18 (Characteristic Number)

The **characteristic** of ring R, denoted $\text{char}(R)$, is defined equivalently.

1. It is the smallest number of times one must successively add the multiplicative identity 1 to get the additive identity 0.
$$1 + 1 + ... + 1 = 0 \tag{143}$$
If no such number n exists, then $\text{char}(R) = 0$.
2. It is equal to m, where $\ker f = \langle m \rangle$ for the homomorphism defined above.[a]

[a] Note that m always exists since $\mathbb{Z}$ is a PID.

Often, it is not obvious whether two given rings R and S are isomorphic. The characteristic number is preserved across ring isomorphisms and therefore is a good sanity check.

Theorem 5.18 (Preservation of Characteristic Number in a Ring Homomorphism)

$R \simeq S \implies \text{char}(R) = \text{char}(S)$.

Proof.

However, the converse is not true! If so, we would have completely classified all rings just based on their characteristic number, and the study of rings would end pretty soon.

Example 5.19 (Same Characteristic does not Imply Isomorphic)

There exists no isomorphism from $\mathbb{Z}$ to $\mathbb{R}$.

Corollary 5.4 (Characteristic of Integral Domain)

If R is an integral domain, then

1. $\ker f = \langle 0 \rangle$ or $\langle p \rangle$ for p prime.
2. $\text{char}(R)$ is either 0 or a prime p.

Proof.

Let $m \in \mathbb{Z}$ be such that $\langle m \rangle = \ker f$. If $m = ab$, then $f(a)f(b) = f(m) = 0$. Since R is an integral domain, $f(a) = 0$ or $f(b) = 0$. Thus $d \in \ker f = \langle m \rangle$ or $e \in \ker f = \langle m \rangle \implies m$ is prime or 0.

Theorem 5.19 (Wilson's Theorem)

Let $p \in \mathbb{N}$ be prime. Then
$$(p-1)! \equiv -1 \pmod{p} \tag{144}$$

The following corollary isn't really worth stating in my opinion, but it has a popular name that might get mentioned a few times.

Corollary 5.5 (Freshman's Dream)

Given a ring R of characteristic p,
$$(a+b)^p = a^p + b^p \tag{145}$$

Proof.

We have

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^{p-k} b^k \tag{146}$$

It is clear that

$$\binom{p}{k} = \frac{p(p-1)...(p-k+1)}{k!} \tag{147}$$

is divisible by p for all $k \neq 0, p$, so all the middle terms must cancel out to 0.

5.9 Division Rings

Definition 5.19 (Division Ring)

A **division ring**, also called a **skew field**, is an associative ring where every nonzero element is invertible with respect to $\times$.[a]

[a] Division rings differ from fields in that multiplication is not required to be commutative.

Let's establish the hierarchy.

Lemma 5.6 (Division Rings are Domains)

Every division ring R is automatically a domain.

Proof.

Every nonzero element is invertible.

Example 5.20 (Invertible Matrices are a Division Ring)

At first, a division ring may not seem different from a field. However, a classic example is the ring of invertible matrices, which is not necessarily commutative, but is a ring in which "division" can be done by right and left multiplication of a matrix inverse.

$$aa^{-1} = a^{-1}a = I \tag{148}$$

This implies that every element in the division ring commutes with the identity, but again commutativity does not necessarily hold for arbitrary elements a, b.

5.10 Fields

Our final structure is field, which seems to add only a few more conditions to a ring, but again unlocks more structure. Field theory is usually pretty tame compared to groups and rings.

Definition 5.20 (Field)

A **field** $(F, +, \times)$ is a commutative, associative ring where every nonzero element is a unit.

Lemma 5.7 (Properties of Addition)

The properties of addition hold in a field.
1. If $x + y = x + z$, then $y = z$.
2. If $x + y = x$, then $y = 0$.
3. If $x + y = 0$, then $y = -x$.
4. $(-(-x)) = x$.

Proof.

For the first, we have

$$\begin{aligned} x + y = x + z &\implies -x + (x + y) = -x + (x + z) && \text{(addition is a function)} \\ &\implies (-x + x) + y = (-x + x) + z && \text{(+ is associative)} \\ &\implies 0 + y = 0 + z && \text{(definition of additive inverse)} \\ &\implies y = z && \text{(definition of identity)} \end{aligned}$$

For the second, we can set $z = 0$ and apply the first property. For the third, we have

$$\begin{aligned} x + y = 0 &\implies -x + (x + y) = -x + 0 && \text{(addition is a function)} \\ &\implies (-x + x) + y = -x + 0 && \text{(+ is associative)} \\ &\implies 0 + y = -x + 0 && \text{(definition of additive inverse)} \\ &\implies y = -x && \text{(definition of identity)} \end{aligned}$$

For the fourth, we simply follow that if y is an inverse of z, then z is an inverse of y. Therefore, $-x$ being an inverse of x implies that x is an inverse of $-x$. $-(-x)$ must also be an inverse of $-x$. Since inverses are unique[a], $x = -(-x)$.

[a] This is proved in algebra.

Lemma 5.8 (Properties of Multiplication)

The properties of multiplication hold in a field.
1. If $x \neq 0$ and $xy = xz$, then $y = z$.
2. If $x \neq 0$ and $xy = x$, then $y = 1$.
3. If $x \neq 0$ and $xy = 1$, then $y = x^{-1}$.
4. If $x \neq 0$, then $(x^{-1})^{-1} = x$.

Proof.

The proof is almost identical to the first. Since $x \neq 0$, we can always assume that x^{-1} exists. For the first, we have

$$\begin{aligned} xy = xz &\implies x^{-1}(xy) = x^{-1}(xz) && \text{(multiplication is a function)} \\ &\implies (x^{-1}x)y = (x^{-1}x)z && \text{($\times$ is associative)} \\ &\implies 1y = 1z && \text{(definition of multiplicative inverse)} \\ &\implies y = z && \text{(definition of identity)} \end{aligned}$$

For the second, we can set $z = 1$ and apply the first property. For the third, we have

$$\begin{aligned} xy = 1 &\implies x^{-1}(xy) = x^{-1}1 && \text{(multiplication is a function)} \\ &\implies (x^{-1}x)y = x^{-1}1 && (\times \text{ is associative}) \\ &\implies 1y = x^{-1}1 && \text{(definition of multiplicative inverse)} \\ &\implies y = x^{-1} && \text{(definition of identity)} \end{aligned}$$

For the fourth, we simply see that x^{-1} is a multiplicative inverse of both x and $(x^{-1})^{-1}$ in the group $(\mathbb{F} \setminus \{0\}, \times)$, and since inverses are unique, they must be equal.

Lemma 5.9 (Properties of Distribution)

For any $x, y, z \in \mathbb{F}$, the field axioms satisfy

1. $0 \cdot x = 0$.
2. If $x \neq 0$ and $y \neq 0$, then $xy \neq 0$.
3. $-1 \cdot x = -x$.
4. $(-x)y = -(xy) = x(-y)$.
5. $(-x)(-y) = xy$.

Proof.

For the first, note that

$$0x = (0 + 0) \cdot x = 0x + 0x \tag{149}$$

and subtracting $0x$ from both sides gives $0 = 0x$. For the second, we can claim that $xy \neq 0$ equivalently claiming that it will have an identity. Since $x, y \neq 0$, their inverses exists, and we claim that $(xy)^{-1} = y^{-1}x^{-1}$ is an inverse. We can see that by associativity,

$$(y^{-1}x^{-1})(xy) = y^{-1}(x^{-1}x)y = y^{-1}y = 1 \tag{150}$$

For the third, we see that

$$0 = 0 \cdot x = (1 + (-1)) \cdot x = 1 \cdot x + (-1) \cdot x = x + (-1) \cdot x \tag{151}$$

which implies that $-1 \cdot x$ is the additive inverse. The fourth follows immediately from the third by the associative property. For the fifth we can see that

$$\begin{aligned} (-x)(-y) &= (-1)x(-1)y && \text{(property 3)} \\ &= (-1)(-1)xy && (\times \text{ is commutative}) \\ &= -1 \cdot (-xy) && \text{(property 3)} \\ &= -(-xy) && \text{(property 3)} \\ &= xy && \text{(addition property 4)} \end{aligned}$$

Theorem 5.20 (Fields are Euclidean Domains)

Every field is a Euclidean domain.

Proof.

Given $x, y \in \mathbb{F}$, assume $xy = 0$ with $x \neq 0$. Since x is invertible,

$$0 = x^{-1}0 = x^{-1}(xy) = y \tag{152}$$

Now assuming that $y \neq 0$, since y is invertible,

$$0 = 0y^{-1} = (xy)y^{-1} = x \tag{153}$$

With this theorem, we have established the hierarchy in the beginning of this section. So as soon as we see a field, we can immediately apply everything we know, such as Euclidean division, unique factorization, GCDs, etc. The converse is not generally true, but with extra assumptions it is.

Theorem 5.21 (Wedderburn's little theorem)

Every finite Euclidean domain is a field.

Theorem 5.22 (Integral Domains are Embedded in Fields)

An integral domain is a ring that is isomorphic to a subring of a field.

Theorem 5.23 (Ideals of Fields)

The only ideals that exist in a field $\mathbb{F}$ is $\{0\}$ and $\mathbb{F}$ itself.

Proof.

Given a nonzero element $x \in \mathbb{F}$, every element of $\mathbb{F}$ can be expressed in the form of ax or xa for some $a \in \mathbb{F}$.

The ring $\mathbb{Z}_n$ has all the properties of a field except the property of having inverses for all of its nonzero elements. This leads to the following theorem.

Theorem 5.24 (Integer Quotient Rings as Finite Fields)

The ring $(\mathbb{Z}_n, +, \times)$ is a field if and only if n is a prime number.

Proof.

($\rightarrow$) Assume that n is composite $\implies n = kl$ for $k, n \in \mathbb{N} \implies k, n \neq 0$, but

$$[k]_n[l]_n = [kl]_n = [n]_n = 0 \tag{154}$$

meaning that $\mathbb{Z}_n$ contains 0 divisors and is not a field. The contrapositive of this states ($\rightarrow$).
($\leftarrow$) Given that n is prime, let $[a]_n \neq 0$, i.e. $[a]_n \neq [0]_n, [1]_n$. The set of n elements

$$[0]_n, [a]_n, [2a]_n, ..., [(n-1)a]_n \tag{155}$$

are all distinct. Indeed, if $[ka]_n = [la]_n$, then $[(k-l)a]_n = 0 \implies n = (k-l)a \iff n$ is not prime. Since the elements are distinct, exactly one of them must be $[1]_n$, say $[pa]_n \implies$ the inverse $[p]_n$ exists.

Corollary 5.6 (Invertibility in $\mathbb{Z}_n$)

For any n, $[k]_n$ is invertible in the ring $\mathbb{Z}_n$ if and only if n and k are relatively prime.

We will talk about finite fields again, which are extremely important in Galois theory and in practical applications in e.g. cryptography.

5.10.1 The Rational Numbers

Now that we've reviewed some fields, let's construct $\mathbb{Q}$ from $\mathbb{Z}$ and verify it's a field.

Definition 5.21 (Rationals)

Given the ordered ring of integers $(\mathbb{Z}, +_{\mathbb{Z}}, \times_{\mathbb{Z}}, \leq_{\mathbb{Z}})$ the **rational numbers** $(\mathbb{Q}, +_{\mathbb{Q}}, \times_{\mathbb{Q}})$ is the quotient space on $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ with the equivalence relation $\sim$

$$(a, b) \sim (c, d) \iff a \times_{\mathbb{Z}} d = b \times_{\mathbb{Z}} c \tag{156}$$

and the operation defined

1. The additive and multiplicative identities are
$$0_{\mathbb{Q}} := (0_{\mathbb{Z}}, a), \quad 1_{\mathbb{Q}} := (a, a) \tag{157}$$
2. Addition on $\mathbb{Q}$ is defined
$$(a, b) +_{\mathbb{Q}} (c, d) := \big((a \times_{\mathbb{Z}} d) +_{\mathbb{Z}} (b \times_{\mathbb{Z}} c), b \times_{\mathbb{Z}} d\big) \tag{158}$$
3. The additive inverse is defined
$$-(a, b) := (-a, b) \tag{159}$$
4. Multiplication on $\mathbb{Q}$ is defined
$$(a, b) \times_{\mathbb{Q}} (c, d) := \big(a \times_{\mathbb{Z}} c, b \times_{\mathbb{Z}} d\big) \tag{160}$$
5. The multiplicative inverse is defined
$$(a, b)^{-1} := (b, a) \tag{161}$$

Theorem 5.25 (Field Structure)

$\mathbb{Q}$ is a field.

Proof.

We do a few things.

1. Verify the additive identity.
$$(a, b) + (0, c) = (ac + 0b, bc) = (ac, bc) \sim (a, b) \tag{162}$$
2. Verify the multiplicative identity.
$$(a, b) \times (c, c) = (ac, bc) \sim (a, b) \tag{163}$$
3. Additive inverse is actually an inverse.
$$(a, b) + (-a, b) = (ab + (-ba), bb) = (0, bb) \sim (0, 1) \tag{164}$$

4. Multiplicative inverse is actually an inverse.
$$(a,b)\times(b,a)=(ab,ba)=(ab,ab)\sim(1,1) \tag{165}$$
5. Addition is commutative.
$$(a,b)+(c,d)=(ad+bc,bd)=(cb+ad,bd)=(c,d)+(a,b) \tag{166}$$
6. Addition is associative.
$$\begin{aligned}(a,b)+((c,d)+(e,f)) &= (a,b)+(cf+de,df) && (167)\\ &= (adf+bcf+bde,bdf) && (168)\\ &= (ad+bc,bd)+(e,f) && (169)\\ &= ((a,b)+(c,d))+(e,f) && (170)\end{aligned}$$
7. Multiplication is commutative.
$$(a,b)\times(c,d)=(ac,bd)=(ca,db)=(c,d)\times(a,b) \tag{171}$$
8. Multiplication is associative.
$$\begin{aligned}(a,b)\times((c,d)\times(e,f)) &= (a,b)\times(ce,df) && (172)\\ &= (ace,bdf) && (173)\\ &= (ac,bd)\times(e,f) && (174)\\ &= ((a,b)\times(c,d))\times(e,f) && (175)\end{aligned}$$
9. Multiplication distributes over addition.
$$\begin{aligned}(a,b)\times((c,d)+(e,f)) &= (a,b)\times(c,d)+(a,b)\times(e,f) && (176)\\ &= (ac,bd)+(ae,bf) && (177)\\ &= (abcf+abde,b^2df) && (178)\\ &= (acf+ade,bdf) \qquad = (a,b)\times(cf+de,df) && (179)\end{aligned}$$

Great, so we have established that $\mathbb{Q}$ is a field. The next property we want to formalize is order. There are countless ways to do it, but I just take the difference and claim that it is greater than 0. Note that given a set, we can really put whatever order we want on it. However, consider the field with the following order.
$$\mathbb{F}=\{0,1\},\ 0<1 \tag{180}$$
This does not behave well with respect to its operations because for example if we have $0<1$, then adding the same element to both sides should preserve the ordering. But this is not the case since $0+1=1>1+1=0$. While it may be easy to define an order, we would like it to be an ordered field.

Definition 5.22 (Ordered Field)

An **ordered field** is a field that has an order satisfying
1. $y<z \implies x+y<x+z$ for all $x\in\mathbb{F}$.
2. $x>0, y>0 \implies xy>0$.

Theorem 5.26 (Properties)

In an totally ordered field,

1. $x > 0 \implies -x < 0$.
2. $x \neq 0 \implies x^2 > 0$.
3. If $x > 0$, then $y < z \implies xy < xz$.

Proof.

The first property is a single-liner

$$0 < x \implies 0 + -x < x + -x \implies -x < 0 \tag{181}$$

For the second property, it must be the case that $x > 0$ or $x < 0$. If $x > 0$, then by definition $x^2 > 0$. If $x < 0$, then

$$x^2 = 1 \cdot x^2 = (-1)^2 \cdot x^2 = (-1 \cdot x)^2 = (-x)^2 \tag{182}$$

and since $-x > 0$ from the first property, we have $x^2 = (-x)^2 > 0$. For the third, we use the distributive property.

$$\begin{aligned} y < z &\implies 0 < z - y && (183)\\ &\implies 0 = x0 < x(z-y) = xz - xy && (184)\\ &\implies xy < xz && (185) \end{aligned}$$

Theorem 5.27 (Ordered Field Structure)

Second, $\mathbb{Q}$ is an ordered field. The order $\leq_{\mathbb{Q}}$ defined on the rationals as

$$(a,b) \leq_{\mathbb{Q}} (c,d) \iff ad \leq_{\mathbb{Z}} bc \tag{186}$$

is a total order. Remember that we have defined $b, d > 0$.

Proof.

For the order property, we have

1. Reflexive.
$$(a,b) \leq_{\mathbb{Q}} (a,b) \iff ab \leq_{\mathbb{Z}} ab \tag{187}$$
2. Antisymmetric.
$$(a,b) \leq_{\mathbb{Q}} (c,d) \implies ad \leq_{\mathbb{Z}} bc(c,d) \leq_{\mathbb{Q}} (a,b) \qquad \implies bc \leq_{\mathbb{Z}} ad \tag{188}$$
This implies that both $ad = bc$, which by definition means that they are in the same equivalence class.
3. Transitivity. Assume that $(a,b) \leq (c,d)$ and $(c,d) \leq (e,f)$. Then, we notice that $b, d, f > 0$ and therefore by the ordered ring property[a] of $\mathbb{Z}$, we have
$$(a,b) \leq_{\mathbb{Q}} (c,d) \implies ad \leq_{\mathbb{Z}} bc \implies adf \leq_{\mathbb{Z}} bcf \tag{189}$$
$$(c,d) \leq_{\mathbb{Q}} (e,f) \implies cf \leq_{\mathbb{Z}} de \implies bcf \leq_{\mathbb{Z}} bde \tag{190}$$
Therefore from transitivity of the ordering on $\mathbb{Z}$ we have $adf \leq bde$. By the ordered ring property[b] we have $0 \leq bde - adf = d(be - af)$. But notice that $d > 0$ from our definition of rationals, and therefore it must be the case that $0 \leq be - af \implies af \leq_{\mathbb{Z}} be$, which by definition means $(a,b) \leq_{\mathbb{Q}} (e,f)$.

For the ordered field property, we have

1. Assume that $y = (a, b) \leq (c, d) = z$. Let $x = (e, f)$. Then $x + y = (af + be, bf)$, $x + z = (cf + de, df)$. Therefore

$$\begin{aligned}(af + be)df &= adf^2 + bedf && (191)\\ &\leq bcf^2 + bedf && (192)\\ &= (cf + de)bf && (193)\end{aligned}$$

But $(af + be)df = (cf + de)bf$ is equivalent to saying $(af + be, bf) \leq_{\mathbb{Q}} (cf + de, df)$, i.e. $x + y \leq x + z$!
2. Let $x = (a, b), y = (c, d)$. Since $0 < x, 0 < y$, by construction this means that $0 < a, 0 < c$ (since $b, d > 0$ in the canonical rational form). By the ordered ring property of the integers, $0 < ac$. So

$$0 < ac \iff 0 \cdot bd < ac \cdot 1 \iff (0, 1) < (ac, bd) \iff 0_{\mathbb{Q}} < (a, c) \times_{\mathbb{Q}} (b, d) = xy \tag{194}$$

[a]If $a \leq b$ and $0 \leq c$, then $ac \leq bc$.
[b]If $a \leq b$, then $a + c \leq b + c$.

Theorem 5.28 (Vector Space Structure)

Third, it is an inner product space over $\mathbb{Q}$.
1. Addition and scalar multiplication is defined according to the field rules above.
2. The inner product, along with the induced norm and metric, are defined

$$\langle x, y\rangle := xy, \quad |x| := \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}, \quad d(x, y) := |x - y| \tag{195}$$

3. The topology induced by the metric is the set of open intervals (x, y), which also coincides with the order topology.

We have successfully defined the rationals, but now these are almost completely separate elements. We know that all integers are rational numbers, and so to show that the rationals are an extension of $\mathbb{Z}$ we want to identify a *canonical injection* $\iota : \mathbb{Z} \to \mathbb{Q}$. This can't just be any canonical injection; it must preserve the algebraic structure between the two sets and must therefore be a *ring homomorphism*.

Theorem 5.29 (Canonical Injection of $\mathbb{Z}$ to $\mathbb{Q}$ is an Ordered Ring Homomorphism)

Let us define the canonical injection $\iota : \mathbb{Z} \to \mathbb{Q}$ to be $\iota(a) = (a, 1)$. This is a ring homomorphism. Additionally, it preserves order: for $a, b \in \mathbb{Z}$,

$$a \leq_{\mathbb{Z}} b \iff \iota(a) \leq_{\mathbb{Q}} \iota(b) \tag{196}$$

Proof.

We show a few things.
1. Preservation of addition.

$$\begin{aligned}\iota(a) +_{\mathbb{Q}} \iota(b) &= (a, 1) +_{\mathbb{Q}} (b, 1) && (197)\\ &= (1a +_{\mathbb{Z}} 1b, 1^2) && (198)\\ &= (a +_{\mathbb{Z}} b, 1) && (199)\\ &= \iota(a +_{\mathbb{Z}} b) && (200)\end{aligned}$$

2. Preservation of multiplication.

$$\begin{aligned}\iota(a) \times_{\mathbb{Q}} \iota(b) &= (a, 1) \times_{\mathbb{Q}} (b, 1) && (201)\\ &= (a \times_{\mathbb{Z}} b, 1^2) && (202)\\ &= (a \times_{\mathbb{Z}} b, 1) && (203)\\ &= \iota(a \times_{\mathbb{Z}} b, 1) && (204)\end{aligned}$$

3. Preservation of multiplicative identity.

$$\iota(1_{\mathbb{Z}}) = (1, 1) = 1_{\mathbb{Q}} \tag{205}$$

To show that it preserves the order, we have

$$\begin{aligned}a \leq_{\mathbb{Z}} b &\iff a \cdot 1 \leq_{\mathbb{Z}} b \cdot 1 && (206)\\ &\iff (a, 1) \leq_{\mathbb{Q}} (b, 1) && (207)\\ &\iff \iota(a) \leq_{\mathbb{Q}} \iota(b) && (208)\end{aligned}$$

Example 5.21 (Numbers)

The rationals, reals, and complex numbers are all fields.[a]

[a] Quaternions are not!

Note the subfield structure $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$. However, we will find that there are tons of other fields lurking in between $\mathbb{Q}$ and $\mathbb{C}$ other than $\mathbb{R}$. We can actually say that there are no subfields of $\mathbb{Q}$.

Lemma 5.10 (Rationals are a Minimal Field)

Every subfield of $\mathbb{C}$ contains $\mathbb{Q}$.

Proof.

Must contain 0 and 1. Keep adding 1 and inverting it to get $\mathbb{Z}$. Now $\mathbb{Z}$ must contain units so $1/n$ also contained. Then multiply the elements to get $\mathbb{Q}$.

Theorem 5.30 (Finite Fields)

There are no finite ordered fields.

Proof.

Assume $\mathbb{F}$ is such an ordered field. It must be the case that $0, 1 \in \mathbb{F}$, with $0 < 1$. Therefore, we also have $0 + 1 < 1 + 1 \implies 1 < 1 + 1$. Repeating this we get

$$0 < 1 < 1 + 1 < 1 + 1 + 1 < \ldots \tag{209}$$

where these elements must be distinct (since only one of $>, <, =$ must be true for a totally ordered set). Since this can be done for a countably infinite number of times, $\mathbb{F}$ cannot be finite.

6 Polynomial Rings

In ring theory, the idea of *adjoining* an existing ring R with an arbitrary element x will appear frequently. That is, given a ring R and some x, can we try and construct a new ring S, denoted $R[x]$ that is the minimal ring containing both R and x? What kind of elements would be in $R[x]$?

1. Since $R \subset R[x]$, it must be the case that for $a \in R$, $a \in R[x]$.
2. Since $x \in R[x]$, it must be the case that $ax \in R[x]$ for all $a \in R$.
3. $x \times x = x^2 \in R[x]$, so it must be the case that $ax^2 \in R[x]$
4. In general, for any $n \in \mathbb{N}$, $x^n \in R[x]$, so it must be the case that $ax^n \in R[x]$.

If we add these terms up, we have elements of the general form

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \tag{210}$$

This is what we refer to as a polynomial, and it indeed does have a ring structure. More generally, given a set of elements $S = \{x_1, \ldots, x_n\}$, $R[S]$ can be defined accordingly. Now let's formally define them.

Definition 6.1 (Polynomial Ring)

Given ring R and a set of **indeterminates** $S = \{x_1, \ldots, x_n\}$, the **ring of polynomials** $R[S] = R[x_1, \ldots, x_n]$ is defined in the equivalent ways.

1. It is the minimal ring containing R as a subring and x.
2. It is the ring of formal expressions of the form

$$f(x_1, \ldots, x_n) = \sum_{0 \leq k_i \leq n} a_{k_1 \ldots k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}, \quad a_i \in R \tag{211}$$

   which for univariate polynomials simplifies to

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in R \tag{212}$$

From both definitions is becomes clear through the ring properties that

1. Addition is defined component-wise: $ax^i + bx^i = (a+b)x^i$.
2. Multiplication is defined component-wise: $x^i x^j = x^{i+j}$.
3. Additive and multiplicative identities is 0 and 1.

Note that x is just a formal symbol, whose powers x^i are just placeholders for the corresponding coefficients a_i so that the given formal expression is a way to encode the finitary sequence. $(a_0, a_1, a_2, ..., a_n)$. Two polynomials are equal if and only if the sequences of their corresponding coefficients are equal. Also note that unlike a function, which we write as f, for a polynomial we should write the indeterminate $f(x)$. We can however interpret $f(x)$ as a function as well (which may not be unique), and doing so allows us to determine special properties of $f(x) \in R[x]$. Before we move on, let's get some terms out of the way.

Definition 6.2 (Some Terms for Polynomials)

Given a univariate polynomial $f(x) \in R[x]$.

1. The **leading coefficient** is the last nonzero coefficient
2. The **degree** of f—denoted $\deg f$—is the index of the leading coefficient.
3. A **monomial** is a polynomial of a single term $a_j x^j$.
4. A **linear** polynomial is a polynomial of degree 1.
5. A **quadratic** polynomial is a polynomial of degree 2.
6. A **cubic** polynomial is a polynomial of degree 3.

We need to be very careful about the properties that hold for polynomials, as they may not be intuitive. For example, for certain finite fields (which are rings), some formally different polynomials may be indistinguishable in terms of mappings.[4] Second, a polynomial may have more roots than its degree. Therefore, we will work in different rings R and provide conditions where our intuition is true in $R[x]$. It is clear that if you have two polynomials of degree n and m, their sum may be degree $k < n, m$. This is not always true for multiplication.

Example 6.1 (Product of Two Linear Polynomials is 0)

Given $f, g \in \mathbb{Z}_6[x]$ with $f(x) = 2x + 4$ and $g(x) = 3x + 3$, we have

$$f(x) \cdot g(x) = (2x + 4)(3x + 3) = 6x^2 + 18x + 12 = 0 \tag{213}$$

There is a simple condition in which the degree is additive, however.

Theorem 6.1 (Bounds on Degrees From Operations)

Given that R is a ring and $f, g \in R[x]$,

$$\deg(f + g) \leq \max\{\deg f, \deg g\} \tag{214}$$

If R is a domain, then

$$\deg(fg) = \deg f + \deg g \tag{215}$$

Note that this automatically implies that $R[x]$ is a domain. Combined with the lemma above, we have: R is an integral domain $\implies$ $R[x]$ is an integral domain.

Proof.

The second may not be true if R has zero divisors.

Just working in domains do not make things all better. Sometimes, we may have two different polynomials but they may define the same function from R to R!

Example 6.2 (Polynomials as Same Function)

Given $f, g \in \mathbb{Z}_2[x]$,

$$f(x) = x \sim g(x) = x^2 \tag{216}$$

As shown in the example above, it is not so simple as to restrict which underlying set you are working on. Some rings R may or may not assert uniqueness of functions in $F[x]$, and vice versa. Therefore, here are some special theorems.

Theorem 6.2 (Uniqueness of Polynomials over Field)

If the field $\mathbb{F}$ is infinite, then different polynomials in $\mathbb{F}[x]$ determine different functions.

Unsurprisingly, the properties of the base ring R determines the properties on R that we see above.

Lemma 6.1 (Ring Properties Extending to Polynomials)

We have the following.

1. R is a commutative ring iff $R[x]$ is a commutative ring.

[4] x and x^2 are equivalent in $\mathbb{Z}_2[x]$.

2. R is an integral domain iff $R[x]$ is an integral domain.

Proof.

For the first claim,
1. $(\to)$ If R is commutative then given $a, b \in R$, we have $(ax^i)(bx^j) = a(x^i b)x^j = (ab)(x^i x^j) = (ba)(x^j x^i) = (bx^j)(ax^i)$, and so from distributive property $R[x]$ is commutative.
2. $(\leftarrow)$ This is trivial since given $R[x]$ commutative, take $a, b \in R \subset R[x]$, and so $ab = ba$ in $R[x]$ implies commutativity in R.

For the second claim, it suffices to prove the domain property since commutativity is proved above.
1. $(\to)$. Assume R is a domain. Now take any two nonzero polynomials $f(x), g(x) \in R[x]$. Then look at their leading term ax^n and bx^n. The leading coefficient of $(fg)(x)$ is $(ab)x^{n+m}$, and since R is a domain $ab \neq= 0 \implies (fg)(x) \neq 0$. So $R[x]$ is a domain.
2. $(\leftarrow)$. This is trivial since given $R[x]$ integral domain, take $a, b \in R \subset R[x]$ with $a, b \neq 0$, and so $ab \neq 0$ since $R[x]$ is a domain. Therefore R is a domain.

With this theorem we unlock all the properties that we have studied in general for the subclasses of rings. Almost always we will assume that R is at least commutative.

6.1 Euclidean Division

Just like how we can do Euclidean division with integers, there is an analogous result for polynomials. However, we require to work with a *field* F rather than an arbitrary ring R.

Theorem 6.3 (Polynomials as Euclidean Domain)

F is a field $\implies F[x]$ is a Euclidean domain. That is, given polynomials $f(x), g(x) \in F[x]$, there are unique polyomials $q(x), r(x) \in F[x]$ s.t.

$$f(x) = q(x)g(x) + r(x), \qquad \deg(r(x)) < \deg(g(x)) \tag{217}$$

Proof.

We first prove existence. If $\deg(f(x)) < \deg(g(x))$, then we can trivially set $q(x) = 0, r(x) = f(x)$. Therefore we can assume that $\deg(f(x)) \geq \deg(g(x))$. We can prove this by strong induction on $k = \deg(f(x))$. Assume that $\deg(f(x)) = 1$. Then if $\deg(g(x)) > 1$ it is trivial as before, so we show for $\deg(g(x)) = 1$. So let

$$f(x) = a_1 x + a_0, \qquad g(x) = b_1 x + b_0 \tag{218}$$

and we can find the solutions

$$f(x) = \frac{a_1}{b_1} g(x) + \left(a_0 - \frac{a_1 b_0}{b_1} \right) \tag{219}$$

Now suppose that the results is known for whenver $\deg(f(x)) \leq k$ and we have a polynomial $F(x) = a_{k+1}x^{k+1} + \ldots a_0$ of degree $k+1$. Then we must check that there exists a quotient and remainder for $0 \leq \deg(g(x)) = m \leq k+1$. Note that the coefficients of x^{k+1} in $F(x)$ and in the polynomial $\frac{a_{k+1}}{b_m} x^{k+1-m} g(x)$ are the same, so the polynomial

$$f(x) = F(x) - \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) \tag{220}$$

has degree at most k. Thus by our induction hypothesis we can write $f(x) = q(x)g(x) + r(x)$, and so

$$\begin{aligned} F(x) &= f(x) + \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) && (221) \\ &= q(x)g(x) + r(x) + \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) && (222) \\ &= \bigg(q(x) + \frac{a_{k+1}}{b_m} x^{k+1-m} \bigg) g(x) + r(x) && (223) \end{aligned}$$

which is indeed a decomposition. Now to prove uniqueness, suppose we had two different decompositions

$$f(x) = q(x)g(x) + r(x) = q'(x)g(x) + r'(x) \implies \big(q(x) - q'(x)\big)g(x) = r(x) - r'(x) \tag{224}$$

IF $q(x) \neq q'(x)$, then the degree of the LHS is at least $\deg(g(x))$, while the degree of the RHS must be strictly less, a contradiction.

Example 6.3 (Polynomial Long Division)

The algorithimc way to get such $q(x), r(x)$ is through *polynomial long division.*

$$\begin{array}{rl} & \quad\; x^2 \;\; + 6x + 11 \\ x-2) & x^3 + 4x^2 \quad - x \;\; + 7 \\ & -x^3 + 2x^2 \\ \hline & \qquad 6x^2 \quad - x \\ & \quad -6x^2 + 12x \\ \hline & \qquad\qquad 11x \;\; + 7 \\ & \qquad\quad -11x + 22 \\ \hline & \qquad\qquad\qquad 29 \end{array}$$

Given field $\mathbb{Z}_5$, $\mathbb{Z}_5[x]$ is a Euclidean domain, with Euclidean division.

In fact, it turns out that you don't necessarily require a polynomial to always come from a field in order to do long division. You can do polynomial long division over *any* commutative rings, as long as the leading coefficient of the divisor is a unit (and since all elements of a field are units, we can do so). This is because at each step, you only need to divide the leading coefficient of the divisor into the leading coefficient of the polynomial you have left. An immediate consequence of this theorem is the following.

Corollary 6.1 (Remainder Theorem)

Let $c \in F$ and $f(x) \in F[x]$. When we divide $f(x)$ by $g(x) = x - c$, the remainder is $f(c)$.

Proof.

By the Euclidean algorithm,

$$f(x) = (x-c)q(x) + r(x) \implies f(c) = (c-c)q(c) + r(c) = r(c) \tag{225}$$

6.2 Roots and Factorization

Next, we can define the all too familiar factors and roots of a polynomial.

Definition 6.3 (Factor)

Given a ring R and a polynomial $f(x) \in R[x]$, if there exists $g(x), h(x)$ of degree at least 1 such that

$$f(x) = g(x)h(x) \tag{226}$$

then g, h are said to be **factors**, or **divisors**, of f. If there are no such factors of f, then $f(x)$ is said to be **irreducible**.

Irreducible polynomials are analogous to prime numbers in $\mathbb{Z}$.

Definition 6.4 (Polynomial Root)

An element $r \in R$ is a **root** of polynomial $f \in R[x]$ if and only if

$$f(r) = 0 \tag{227}$$

Note that both factors and roots are intimately tied to Euclidean division, so the two are closely related.

Theorem 6.4 (Root-Factor Theorem)

Given a commutative ring R (usually R is a field) and $f(x) \in R[x]$, $(x - c)$ is a factor of $f(x)$, i.e. can be factored into

$$f(x) = (x - c)q(x) \tag{228}$$

for some $q(x) \in R[x]$ of degree $\deg(f) - 1$ if and only if $f(c) = 0$.[a]

[a]Note that this is not true for an arbitrary ring. R must be commutative at least.

Proof.

We prove for when R is a field F, but it turns out that the theorem also holds for commutative rings R.

1. $(\rightarrow)$. Given that $(x - c)$ is a factor of $f(x)$, this means that by the Euclidean algorithm $f(x) = (x - c)q(x)$ for some $q(x)$, and so $f(c) = (c - c)q(c) = 0$.
2. $(\leftarrow)$. Given that $f(c) = 0$. By the remainder theorem this means that when we divide $f(x)$ by $(x - c)$, the remainder is $f(c) = 0$, and so $f(x) = (x - c)q(x) + 0 = (x - c)q(x) \implies (x - c)$ is a factor of $f(x)$.

Notice how these polynomials mimick integers, and to drive this point even further, let's introduce the greatest common divisor.

Theorem 6.5 (GCD of Two Polynomials Exist)

Given nonzero polynomials $f(x), g(x) \in F[x]$, let

$$S = \{h(x) \in F[x] \mid h(x) = a(x)f(x) + b(x)g(x) \text{ for some } a(x), b(x) \in F[x]\} \tag{229}$$

Then there exists some polynomial $d(x) \in S$ of smallest degree, and every $h(x) \in S$ is divisible by $d(x)$.

Proof.

The existence is trivial since by the well-ordering principle on the degrees of polynomials in S, such a minimal degree must exist. Now we prove the second claim by proving $d(x) \mid f(x)$. We apply the division algorithm to write

$$f(x) = q(x)d(x) + r(x) \tag{230}$$

If $r(x) = 0$, then by root factor theorem we are done. If $r(x) \neq 0$, we then write

$$\begin{aligned} r(x) &= f(x) - q(x)d(x) && (231)\\ &= f(x) - \big(s(x)f(x) + t(x)g(x)\big)q(x) && (232)\\ &= \big(1 - s(x)q(x)\big)f(x) - \big(t(x)q(x)\big)g(x) \in S && (233) \end{aligned}$$

Since $r(x) \in S$ due to its form, the fact that $\deg(r(x)) < \deg(d(x))$ contradicts the way that $d(x)$ was chosen. Therefore $r(x) = 0$. It turns out that $d(x)$ is unique up to a constant factor.

Definition 6.5 (GCD)

$d(x)$ as above is called the **greatest common divisor** of $f(x), g(x)$, denoted $d(x) = \gcd(f(x), g(x))$ satisfying

1. $d(x) \mid f(x)$, $d(x) \mid g(x)$, and
2. $\forall e(x) \in F[x]$, if $e(x) \mid f(x)$ and $e(x) \mid g(x)$, then $e(x) \mid d(x)$.

$f(x), g(x)$ are said to be **relatively prime** if $\gcd(f(x), g(x)) = 1$.

The algorithmic way for computing the GCD is done the same way by performing Euclidean algorithm on two polynomials: dividing on by the other, taking the remainder, and dividing the lesser degree by the remainder again, until the remainder is 0.

Lemma 6.2 ()

Suppose $f(x)$ is irreducible and $f(x) \mid g(x)h(x)$. Then $f(x) \mid g(x)$ or $f(x) \mid h(x)$.

Now, we show an extremely important theorem. This should be intuitive since F a field implies $F[x]$ a Euclidean domain, which is a PID, which has the unique factorization theorem.

Theorem 6.6 (Unique Factorization of Polynomials over Fields)

Given field F and nonconstant polynomial $f(x) \in F[x]$ of degree n, we can always write $f(x)$ as a unique[a] product of at most n irreducible polynomials in $F[x]$.

[a] up to constant factors and rearrangement

Proof.

To prove the bound, the general idea is that by the root factor theorem, each root gives rise to a linear factor, and so inductively we cannot have more than n linear factors. Strong induction on degree of $f(x)$ by starting with linear.

Note that this is *not* true in arbitrary rings.

Example 6.4 (Linear Polynomial with 3 Roots)

Consider $f(x) = x^2 - 1 \in \mathbb{Z}_8[x]$, a commutative ring. Then $1, 3, 5, 7$ are all roots of $f(x)$, which is greater than its degree. Furthermore, it has two different factorizations

$$x^2 - 1 = (x+1)(x-1) = (x+3)(x-3) \tag{234}$$

Theorem 6.7 (Interpolation)

For any collection of given field values $y_1, y_2, ..., y_n \in F$ at given distinct points $x_1, x_2, ..., x_n \in F$, there exists a unique polynomial $f \in F[x]$ with $\deg f < n$ such that

$$f(x_i) = y_i, \quad i = 1, 2, ..., n \tag{235}$$

This is commonly known as the **interpolation problem**, and when $n = 2$, this is called **linear interpolation**.

6.3 Algebraically Closed Fields

Now that we have seen some examples of fields, what properties would we like it to have? Going back to polynomials, recall that if F is a field, then $F[x]$ as a Euclidean domain gave us a lot of nice properties, such as admitting a unique factorization of irreducible polynomials. However, we have only proved that the number of roots is *at most* the degree n, but not that it actually reaches n. In fact, in a more extreme case, a polynomial may not even factor *at all* in $F[x]$, since it could be irreducible. So while we have defined an upper bound for the number of roots for a polynomial, we have not determined whether a polynomial has any roots at all, i.e. a lower bound.

We don't have much *control* over what these irreducible polynomials can look like. We may have to check—either through theorems or manually—that a polynomial or arbitrary degree is irreducible. If we would like to assert that all irreducible polynomials must be of smallest degree—that is, linear—then such a field is called *algebraically closed*. This algebraic closed property asserts also that the lower bound on the number of (non-unique) factors is n.

Definition 6.6 (Algebraically Closed Field)

A field F is **algebraically closed** if every polynomial of positive degree (i.e. non-constant) in $F[x]$ has at least one root in F.

This is equivalent to saying that every polynomial can be expressed as a product of first degree polynomials. To extend our analysis more, we can talk about the multiplicity of these factors, which just tells us more about how many unique and non-unique factors a polynomial has.

Definition 6.7 (Multiplicity)

A root c of polynomial $f(x) \in F[x]$ is called simple if $f(x)$ is not divisible by $(x - c)^2$ and multiple otherwise. The **multiplicity** of a root c is the maximum k such that $(x - c)^k$ divides $f(x)$.

To restate the root-factor theorem for $R[x]$ with arbitrary commutative ring R, the number of roots of a polynomial—counted with multiplicity—does not exceed the degree of this polynomial. Furthermore, these numbers are equal if and only if the polynomial is a product of linear factors.

Example 6.5 (Reals are not Algebraically Closed)

$\mathbb{R}$ is not algebraically closed since we can identify the polynomial $f(x) = x^2 + 1 \in \mathbb{R}[x]$ which does not have any roots in $\mathbb{R}$. Consequently, any subfield of $\mathbb{R}$ (which contains 1) such as $\mathbb{Q}, \mathbb{Q}(\sqrt{2}), \ldots$ are not algebraically closed.

It turns out that the complex numbers are algebraically closed, which is presented with the following grand name. Ironically, this theorem cannot be proven with algebra alone. We need complex analysis.[5]

[5]Gauss proved this for the first time in 1799.

Theorem 6.8 (Fundamental Theorem of Algebra)

Suppose $f \in \mathbb{C}[x]$ is a polynomial of degree $n \geq 1$. Then $f(x)$ has a root in $\mathbb{C}$. It immediately follows from induction that it can be factored as a product of linear polynomials in $\mathbb{C}[x]$.

Proof.

WLOG we can assume that f is monic: $f(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1 z + a_0$. Since $\mathbb{C}$ is a field, we can set

$$f(z) = z^n \bigg(1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n} \bigg) \tag{236}$$

Since

$$\lim_{|z| \rightarrow \infty} \bigg(1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n} \bigg) = 0 \tag{237}$$

there exists a $R > 0$ s.t.

$$|z| > R \implies \bigg| 1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n} \bigg| < \frac{1}{2} \tag{238}$$

and hence

$$|z| > R \implies |f(z)| > |z|^n \cdot \bigg(1 - \frac{1}{2} \bigg) > \frac{R^n}{2} \tag{239}$$

So z cannot be a root if $|z| > R$. On the other hand, $f(z)$ is continuous (under the Euclidean topology) and so on the compact set $\{z \in \mathbb{C} \mid |z| \leq R\}$, $|f(z)|$ achieves a minimum value say at the point z_0. We claim that $\min_z f(z) = 0$.
For convenience, we let $z_0 = 0$ (we can do a change of basis on the polynomial) and assume that the minimum is some positive number, i.e. $f(0) = a_0 \neq 0$. Let j be the smallest positive integer such that $a_j = 0$. Let

$$g(z) = \frac{a_{j+1}}{a_j} z + \ldots + \frac{a_n}{a_j} z^{n-j} \implies f(z) = a_0 + a_j z^j \big(1 + g(z)\big) \tag{240}$$

We set $\gamma = \sqrt[j]{-a_0/a_j}$ and consider the values of

$$f(t\gamma) = a_0 + a_j (t\gamma)^j \big(1 + g(t\gamma)\big) \tag{241}$$
$$= a_0 - a_0 t^j \big(1 + g(t\gamma)\big) \tag{242}$$
$$= a_0 \big\{1 - t^j \big(1 + g(t\gamma)\big)\big\} \tag{243}$$

for $t > 0$. For t sufficiently small, we have

$$|g(t\gamma)| = \bigg| \frac{a_{j+1}}{a_j} (t\gamma) + \ldots + \frac{a_n}{a_j} (t\gamma)^{n-j} \bigg| < \frac{1}{2} \tag{244}$$

and for such t, this implies

$$|f(t\gamma)| = |a_0| |1 - t^j (1 + g(t\gamma))| \leq |a_0| |1 - t^j/2| < |a_0| \tag{245}$$

and so z_0 cannot have been the minimum of $|f(z)|$. Therefore, the minimum value must be 0.

Great, so through this theorem, we can work in any subfield of $\mathbb{C}$ and guarantee that will have all of its roots in $\mathbb{C}$.

Corollary 6.2 ($\mathbb{C}$ is algebraically closed)

$\mathbb{C}$ is algebraically closed, i.e. $\mathbb{C}$ is a splitting field of $\mathbb{C}[x]$.

Put more succinctly, the impossibility of defining division on the ring of integers motivates its extension into the field of rational numbers. Similarly, the inability to take square roots of negative real numbers forces us to extend the field of real numbers to the bigger field of complex numbers.

Theorem 6.9 (Eigenvector Conditions for Algebraic Closedness)

A field F is algebraically closed if and only if for each natural number n, every endomorphism of F^n (that is, ever linear map from F^n to itself) has at least one eigenvector.

Proof.

An endomorphism of F^n has an eigenvector if and only if its characteristic polynomial has some root. ($\rightarrow$) So, when F is algebraically closed, every characteristic polynomial, which is an element of $F[x]$, must have a root. ($\leftarrow$) Assume that every characteristic polynomial has some root, and let $p \in F[x]$. Dividing the polynomial by a scalar doesn't change its roots, so we can assume p to have leading coefficient 1. If $p(x) = a_0 + a_1 x + ... + x^n$, then we can identify matrix

$$A = \begin{pmatrix} 0 & 0 & ... & 0 & -a_0 \\ 1 & 0 & ... & 0 & -a_1 \\ 0 & 1 & ... & 0 & -a_2 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & -a_{n-1} \end{pmatrix} \tag{246}$$

such that the characteristic polynomial of A is p.

With this splitting condition, we can get a nice set of formulas often introduced in high-school math competitions.

Theorem 6.10 (Viete's Formulas)

Given that a polynomial f factors into linear terms, that is

$$f(x) = a_0 \prod_{i=1}^n (x - c_i), c_i \text{ roots of } f \tag{247}$$

Then the coefficients of f can be presented with the formulas

$$\begin{aligned} \sum_{i=1}^n c_i &= -\frac{a_1}{a_0} \\ \sum_{i_1 < i_2} c_{i_1} c_{i_2} &= \frac{a_2}{a_0} \\ \sum_{i_1 < ... < i_k} \prod_{j=1}^k c_{i_j} &= (-1)^k \frac{a_k}{a_0} \\ c_1 c_2 c_3 ... c_n &= (-1)^n \frac{a_n}{a_0} \end{aligned}$$

6.4 Reducibility of Real Polynomials

Theorem 6.11 ()

If c is a complex root of polynomial $f \in \mathbb{R}[x]$, then $\bar{c}$ is also a root of the polynomial. Moreover, $\bar{c}$ has the same multiplicity as c.

Corollary 6.3 ()

Every nonzero polynomial in $\mathbb{R}[x]$ factors into a product of linear terms and quadratic terms with negative discriminants.

Example 6.6 ()

$$\begin{aligned} x^5 - 1 &= (x-1)\bigg(x - \Big(\cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}\Big)\bigg)\bigg(x - \Big(\cos\frac{2\pi}{5} - i\sin\frac{2\pi}{5}\Big)\bigg) \\ &\times \bigg(x - \Big(\cos\frac{4\pi}{5} + i\sin\frac{4\pi}{5}\Big)\bigg)\bigg(x - \Big(\cos\frac{4\pi}{5} - i\sin\frac{4\pi}{5}\Big)\bigg) \\ &= (x-1)\bigg(x^2 - \frac{\sqrt{5}-1}{2}x + 1\bigg)\bigg(x^2 + \frac{\sqrt{5}+1}{2}x + 1\bigg) \end{aligned}$$

Corollary 6.4 ()

Every polynomial $f \in \mathbb{R}[x]$ of odd degree has at least one real root.

Proof.

This is a direct result of Theorem **. Alternatively, without loss of generality we can assume that the leading coefficient of f is positive. Then

$$\lim_{x\rightarrow +\infty} f(x) = +\infty, \; \lim_{x \rightarrow -\infty} f(x) = -\infty \tag{248}$$

By the intermediate value theorem, there must be some point where f equals 0.

Theorem 6.12 (Descartes' Rule of Signs)

Let $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{R}[x]$. Let C_+ be the number of times the coefficients of $f(x)$ change signs (here we ignore the zero coefficients); let Z_+ be the number of positive roots of $f(x)$, counting multiplicities. Then $Z_+ \leq C_+$ and $Z_+ \equiv C_+ \pmod 2$. Moreover, if we set $g(x) = f(-x)$, let C_- be the number of times the coefficients of $g(x)$ change signs, and Z_- the number of negative roots of $f(x)$. Then $Z_- \leq C_-$ and $Z_- \equiv C_- \pmod 2$.

Theorem 6.13 ()

The number of positive roots of $f(x)$ is the same as the number of negative roots of $f(-x)$.

Example 6.7 (Easy Way to Find Number of Positive Roots)

Given $f(x) = x^5 + x^4 - x^2 - 1$,

1. We have $C_+ = 1$. By Descartes' rule of signs, it must be the case that $Z_+ \leq 1$ and $Z_+ \equiv 1 \pmod 2 \implies Z_+ = 1$.
2. Since $f(-x) = -x^5 + x^4 - x^2 - 1$, we have $C_- = 2$, so $Z_- = 0$ or 2. This is the best that we can do, though it turns out that it actually has 0 negative roots.[a]

[a] On the other hand, $x^5 + 3x^3 - x^2 - 1$ has 2 negative roots.

Note that if a polynomial has a multiple root but its coefficients are known only approximately (but with any degree of precision), then it is impossible to prove that the multiple roots exists because under any perturbation of the coefficients, however small, it may separate into simple roots or simply cease to exist. This fact leads to the "instability" of the Jordan Normal form because under any perturbation of the elements of a matrix A, the change may drastically affect the characteristic polynomial, hence affecting the geometric multiplicities of its eigenvectors.

6.5 Reducibility of Integer Polynomials

Even though we have covered a more general theory of polynomials with rational coefficients, it is worthwhile to visit integer polynomials for two reasons. First, there are a few specialized theorems that allow us to easily determine reducibility in $\mathbb{Z}[x]$. Second, Gauss's lemma allows us to check for reducibility in $\mathbb{Q}[x]$ by checking for reducibility in $\mathbb{Z}[x]$, at which point we can abuse the specialized theorems we have developed.

Theorem 6.14 (Rational Root Theorem)

Let $a_n x^n + \ldots + a_0 \in \mathbb{Z}[x]$. If $r/s \in \mathbb{Q}$ with $\gcd(r, s) = 1$, then $r \mid a_0$ and $s \mid a_n$.

Proof.

Given that r/s is a root, we have

$$a_n (r/s)^n + \ldots + a_0 = 0 \tag{249}$$

Multiplying by s^n, we get

$$a_n r^n + a_{n-1} r^{n-1} s + \ldots + a_1 s^{n-1} r + a_0 s^n = 0 \tag{250}$$

and putting this equation on mod r and mod s implies that $r | a_0 s^n$ and $s | a_n r^n$, respectively. But since we assumed that $\gcd(r, s) = 1$, $r | a_0$ and $s | a_n$.

The next is quite a remarkable result, since it says that decompositions in $\mathbb{Q}[x]$ imply decompositions in $\mathbb{Z}[x]$! Therefore, to check irreducibility in $\mathbb{Q}[x]$, it suffices to check irreducibility in $\mathbb{Z}[x]$.

Lemma 6.3 (Gauss's Lemma)

Let $f \in \mathbb{Z}[x]$. If $\exists g, h \in \mathbb{Q}[x]$ s.t. $f(x) = g(x) h(x)$, then $\exists \bar{g}, \bar{h} \in \mathbb{Z}[x]$ s.t. $f(x) = \bar{g}(x) \bar{h}(x)$.

Proof.

We can find $k, l \in \mathbb{Z}$ s.t. $g_1(x) = k g(x)$ and $h_1(x) = l h(x)$ have integer coefficients, i.e. $g_1, h_1 \in \mathbb{Z}[x]$. Then, $kl f(x) = g_1(x) h_1(x) \in \mathbb{Z}[x]$. Let p be a prime factor of kl. We have

$$0 \equiv \bar{k} \bar{l} \bar{f}(x) \equiv \bar{g}_1(x) \bar{h}_1(x) \text{ in } \mathbb{Z}_p[x] \tag{251}$$

Since $\mathbb{Z}_p$ is an integral domain, $\mathbb{Z}_p[x]$ is an integral domain, and so $\bar{g}_1$ or $\bar{h}_1$ must be 0. WLOG let it be $\bar{g}_1$. Then every coefficient of $g_1(x)$ is divisible by p, and we can write it in the form $g_2(x) = pg_1(x)$. Therefore,

$$p(x) \cdot \frac{kl}{p} = \underbrace{\frac{g_1(x)}{p}}_{g_2(x)} \cdot \underbrace{h_1(x)}_{h_2(x)} \iff f(x)\frac{kl}{p} = g_2(x)h_2(x) \tag{252}$$

Since there are only finitely many prime divisors, we do this for all prime factors of kl, and we have

$$f(x) = g_n(x)h_n(x), \qquad g_n, h_n \in \mathbb{Z}[x] \tag{253}$$

Example 6.8 (Reducibility of Integer Polynomials)

Let $f(x) = x^4 - x^3 + 2$. The rational roots are in the set $S = \{\pm 1, \pm 2\}$, but none of them work since $f(\pm 1), f(\pm 2) \neq 0$. By degree considerations and Gauss's lemma, if $f(x)$ is reducible, then

$$f(x) = (x^2 + ax + b)(x^2 + cx + d), \qquad a, b, c, d \in \mathbb{Z} \tag{254}$$

We know that $bd \in S$, with $a + c = -1$, $d + b + ac = 0$, and so on for each coefficients. We can brute force this finite set of possibilities.

A great way to check irreducibility is to check in mod p.

Theorem 6.15 ()

Let $f(x) = a_n x^n + \ldots + a_0 \in \mathbb{Z}[x]$. If $p \nmid a_n$ and $f \in \mathbb{Z}_p[x]$ is irreducible, then f is irreducible in $\mathbb{Q}[x]$.[a]

[a]May need to verify this again.

Proof.

Suppose that $f(x) = g(x)h(x) \in \mathbb{Z}[x]$ with $\deg(g), \deg(h) > 0$. Then

$$f(x) \equiv g(x)h(x) \text{ in } \mathbb{Z}_p[x] \tag{255}$$

Since $f(x)$ is irreducible in $\mathbb{Z}_p[x]$, we must have that one of $g(x)$ or $h(x)$ has degree 0 in $\mathbb{Z}_p[x]$. WLOG let it be $g(x)$, but this means that the leading coefficient of $g(x)$ must be divisible by $p \implies$ leading coefficient of $f(x)$ is divisible by $p \iff p \mid a_n$.

Example 6.9 ()

x^4+x+1 is irreducible in $\mathbb{Z}_2[x]$. So we can extend this to $\mathbb{Z}[x]$ to see that *all* fourth degree polynomials of form $ax^4 + bx^3 + cx^2 + dx + e$, which a, d, e odd and b, c even is irreducible in $\mathbb{Q}[x]$.

This is a powerful theorem to quickly find a large class of polynomials that are irreducible. However, being reducible in $\mathbb{Z}_p[x]$ does not imply reducibility in $\mathbb{Q}$. In fact, there are polynomials $f(x) \in \mathbb{Z}[x]$ which are irreducible but reducible in $\mathbb{Z}_p$ for *every* prime p.

Theorem 6.16 (Eisenstein's Criterion)

Let $f(x) = a_n x^n + \ldots + a_0 \in \mathbb{Z}[x]$ and $p \in \mathbb{Z}$ a prime s.t. $p \nmid a_n$, $p \mid a_i$ for $i = 0, \ldots, a_{n-1}$, and $p^2 \nmid a_0$. Then $f(x)$ is irreducible in $\mathbb{Q}[x]$.

Proof.

Suppose that $f(x) = g(x)h(x) \in \mathbb{Q}[x]$ with $\deg(g), \deg(h) > 0$. Then, by Gauss's lemma, $g, h \in \mathbb{Z}[x]$. Reducing the equations mod p,

$$f(x) = g(x)h(x) \text{ in } \mathbb{Z}_p[x] \tag{256}$$

But $f(x) = a_n x^n$. By unique factorization theorem in $\mathbb{Z}_p[x]$, $g, h \in \mathbb{Z}_p[x]$ must be products of units and prime factors of $a_n x^n$, which are $\{x\}$. Therefore, let

$$g(x) = b_m x^m, h(x) = \frac{a_n}{b_m} x^{n-m} \in \mathbb{Z}_p[x] \tag{257}$$

with $\deg(g) = m > 0$ and $\deg(h) = n - m > 0$ in $\mathbb{Z}[x]$. This implies that the constant coefficients of $g(x), h(x)$ are divisible by p, which implies that the constant coefficients of $f(x) = g(x)h(x)$ are divisible by p^2, a contradiction.

Example 6.10 (Easy Checks for Irreducibility with Eisenstein)

Listed.

1. $x^{13} + 2x^{10} + 4x + 6$ is irreducible in $\mathbb{Q}[x]$ by Eisenstein for $p = 2$.
2. $x^3 + 9x^2 + 12x + 3$ is irreducible in $\mathbb{Q}[x]$ by Eisenstein for $p = 3$.
3. Let $f(x) = x^4 + x^3 + x^2 + x + 1$. Then, we know that $f(x) = \frac{x^5-1}{x-1}$ and so

$$\begin{aligned} f(x+1) &= \frac{(x+1)^5 - 1}{(x+1) - 1} && (258) \\ &= \frac{1}{x}\left(x^5 + \binom{5}{1}x^4 + \binom{5}{2}x^3 + \binom{5}{3}x^2 + \binom{5}{4}x + \binom{5}{5} - 1\right) && (259) \\ &= x^4 + 5x^3 + 10x^2 + 10x + 5 && (260) \end{aligned}$$

So all nonleading coefficients are divisible by 5 exactly once, which by Eisenstein implies that $f(x+1)$ is irreducible which implies that $f(x)$ is irreducible.

We have prod that for $\alpha \in \mathbb{C}$, subfield $F \subset \mathbb{C}$, and $f(x) \in F[x]$, with $f(\alpha) = 0$, then $B = \{1, \alpha, \ldots, \alpha^{\deg(f)-1}\}$ spans $F[\alpha]$ as a F-vector space. If $f(x)$ is irreducible then B is a basis.

6.6 Ring Extensions

We have already seen the idea of taking ring R and adjoining an element x to it to get a minimal ring containing both R (as a subring) and x. This was the polynomial ring. Now consider a field F and rather than adjoining an indeterminant x, we consider a bigger field $K \supset F$ with some $\alpha \in K$ (with $\alpha \notin F$). Then, we would like to similarly extend F to a bigger field $F[\alpha]$ containing both F (as a subring) and α. This can be naturally constructed when mapping α through polynomials $f(x) \in F[x]$. For ease of notation, we introduce the following.

Definition 6.8 (Ring Extension)

A pair of rings R, S where R is a subring of S is called a **ring extension**, denoted $R \hookrightarrow S$.

Now we introduce the ring R adjoined by α.

Definition 6.9 (Adjoining Ring of Polynomial Elements)

Given ring extension $R \hookrightarrow S$ with finite $\alpha = \{\alpha_1, \ldots, \alpha_n\} \subset S$, the **ring** R **adjoined by** α is defined

$$R[\alpha] = R[\alpha_1, \ldots, \alpha_n] := \{f(\alpha_1, \ldots, \alpha_n) \in K \mid f \in F[x_1, \ldots, x_n]\} \subset K \tag{261}$$

That is, we take the α_i's and map it through all polynomials in $F[x_1, \ldots, x_n]$. This induces a subring structure.

Most of the times, we will work with adjoining rings of univariate polynomial elements.

$$F[\alpha] := \{f(\alpha) \in F \mid f \in F[x]\} \subset K \tag{262}$$

Note that if $R \hookrightarrow S$ is a ring extension, then $R[\alpha]$ is a ring, and so it makes sense to write $R[\alpha][\beta] \subset S$. It had better be the case that this ring is consistent with equivalent constructions.

Lemma 6.4 (Consistency of Adjoning Rings)

It is the case that

$$R[\alpha, \beta] = R[\alpha][\beta] = R[\beta][\alpha] \tag{263}$$

It immediately follows for any finite set X as above, $F[\alpha_1, \ldots, \alpha_n]$ characterizes the $(n-1)$-fold adjoin operations.

Proof.

Since R being a ring implies $R[x]$ is also a ring, we can intuit some subring structure. Furthermore, $R[\alpha]$ is a minimal ring containing α.

Lemma 6.5 (Adjoining Fields are a Ring Extension)

Given an adjoining ring, we claim two things.

1. It is true that $F \subset F[\alpha] \subset K$. Furthermore, if $\alpha \notin F$, then $F \subsetneq F[\alpha]$.
2. $R[\alpha]$ is the minimal ring containing R and α.

Proof.

Note that $F \subset F[\alpha]$ since we can just take the constant polynomials, so this is not very interesting. Given two elements $\phi, \gamma \in F[\alpha]$, there exists polynomials $f, g \in F[x]$ s.t. $\phi = f(\alpha), \gamma = g(\alpha)$. Since $F[x]$ is a ring, we see that

$$\phi + \gamma = f(\alpha) + g(\alpha) = (f + g)(\alpha) \tag{264}$$
$$\phi \cdot \gamma = f(\alpha) \cdot g(\alpha) = (fg)(\alpha) \tag{265}$$

Furthermore, it is easy to check that 0 and 1 are the images of α through the 0 and 1 polynomials.

What allows us to make this inclusion proper is that the $\alpha \in K$, which does not necessarily have to be in F, *extends* this field a bit further, but since we can only map the one element α, it may not cover all of K. Let's go through some examples.

Example 6.11 (Radical Extensions of $\sqrt{2}$)

Let $F = \mathbb{Q}$ and $K = \mathbb{C}$. We claim $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$.

1. $\mathbb{Q}[\sqrt{2}] \subset \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. $\mathbb{Q}[\sqrt{2}]$ are elements of the form
$$f(\sqrt{2}) = a_n(\sqrt{2})^n + a_{n-1}(\sqrt{2})^{n-1} + \ldots + a_2(\sqrt{2})^2 + a_1\sqrt{2} + a_0 \tag{266}$$
This can be written by collecting terms, of the form $a + b\sqrt{2}$.
2. $\mathbb{Q}[\sqrt{2}] \supset \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$. Given an element $a + b\sqrt{2}$, this is clearly in $\mathbb{Q}[\sqrt{2}]$ since it is the image of $\sqrt{2}$ under the polynomial $f(x) = a + bx$.

Given this, we may extrapolate this pattern and claim that $\mathbb{Q}[\sqrt{2} + \sqrt{3}]$ consists of all numbers of form $a + (\sqrt{2} + \sqrt{3})b$. However, this is *not* the case.

Example 6.12 (Radical Extensions of $\sqrt{2} + \sqrt{3}$)

Given any element $\beta \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$, it is by definition of the form
$$\beta = \sum_{k=0}^{n} a_k(\sqrt{2} + \sqrt{3})^k \tag{267}$$
Clearly $1, \sqrt{2}+\sqrt{3} \in \mathbb{Q}[\sqrt{2}+\sqrt{3}]$ by mapping $\sqrt{2}+\sqrt{3}$ through the polynomials $f(x) = 1$ and $f(x) =$. However, we can see that $(\sqrt{2} + \sqrt{3})^2 = 5 + \sqrt{6}$,[a] and so $\sqrt{6} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$. Furthermore, we have $(\sqrt{2} + \sqrt{3})^3 = 11\sqrt{2} + 9\sqrt{3}$, and so with the ring properties we can conclude that
$$\frac{1}{2}\big[(11\sqrt{2} + 9\sqrt{3}) - 9(\sqrt{2} + \sqrt{3})\big] = \sqrt{2} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}] \tag{268}$$
$$-\frac{1}{2}\big[(11\sqrt{2} + 9\sqrt{3}) - 11(\sqrt{2} + \sqrt{3})\big] = \sqrt{3} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}] \tag{269}$$
$$\tag{270}$$
If we go a bit further, we can show that
$$\mathbb{Q}[\sqrt{2} + \sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\} \tag{271}$$

[a] where we use $\sqrt{6}$ as notation for $\sqrt{2} \cdot \sqrt{3}$

6.7 Rational Functions

Given a field F, we have constructed the Euclidean domain $F[x]$. However, this is one step away from being a field. We mimick the construction of the rational numbers $\mathbb{Q}$ as a quotient space over $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ by taking $F[x] \times (F[x] \setminus \{0\})$ and putting a quotient on it.

Definition 6.10 (Rational Functions)

The **rational functions** are defined to be the field of quotients (really just 2-tuples) of the form
$$F(x) := \left\{ \frac{f(x)}{g(x)} \;\middle|\; f(x), g(x) \in F[x], g(x) \neq 0 \right\} \tag{272}$$
where addition and multiplication is defined in the usual sense.

Theorem 6.17 (Partial Fractions Decomposition)

Let $f(x), g(x) \in F[x]$ where $\deg(f(x)) < \deg(g(x))$. If $g(x) = u(x)v(x)$ where u, v are relatively prime, then there are polynomials $a(x), b(x)$ with $\deg(a) < \deg(u), \deg(b) < \deg(v)$ s.t.

$$\frac{f(x)}{g(x)} = \frac{a(x)}{u(x)} + \frac{b(x)}{v(x)} \tag{273}$$

By induction, we can prove this for any finite set of irreducible polynomials.

Proof.

We describe an algorithm to get this decomposition. There are polynomials $s(x), t(x)$ s.t. $1 = s(x)u(x) + t(x)v(x)$. Therefore,

$$\frac{f(x)}{u(x)v(x)} = \frac{f(x)t(x)}{u(x)} + \frac{f(x)s(x)}{v(x)} \tag{274}$$

and we can use the Euclidean algorithm to write

$$\frac{f(x)t(x)}{u(x)} = q(x) + \frac{a(x)}{u(x)}, \qquad \deg(a) < \deg(u) \tag{275}$$

$$\frac{f(x)s(x)}{v(x)} = q(x) + \frac{a(x)}{u(x)}, \qquad \deg(b) < \deg(v) \tag{276}$$

which implies

$$\frac{f(x)}{u(x)v(x)} = \frac{a(x)}{u(x)} + \frac{b(x)}{v(x)} \tag{277}$$

Example 6.13 ()

Consider the rational function $\frac{x+3}{x^3(x-1)^2}$. Applying the Euclidean algorithm, we find that

$$1 = (3x^2 + 2x + 1)(x-1)^2 - (3x-4)x^3 \tag{278}$$

and so

$$\frac{x+3}{x^3(x-1)^2} = \frac{(x+3)(3x^2+2x+1)}{x^3} - \frac{(x+3)(3x-4)}{(x-1)^2} \tag{279}$$

$$= \frac{11x^2 + 7x + 3}{x^3} + \frac{-11x + 15}{(x-1)^2} \tag{280}$$

6.8 Symmetric Polynomials

Definition 6.11 ()

A polynomial $f \in \mathbb{F}[x_1, ..., x_n]$ is called **symmetric** if it is invariant under any permutation of the variables x_i.

Example 6.14 ()

Power sums are symmetric polynomials.

$$p(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i^k \tag{281}$$

Definition 6.12 ()

An **elementary symmetric polynomial** is a symmetric polynomial of one of these forms:

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + ... + x_n \\
\sigma_2 &= x_1 x_2 + x_1 x_3 + ... + x_{n-1} x_n \\
... &= ... \\
\sigma_k &= \sum_{i_1 < ... < i_k} x_{i_1} x_{i_2} ... x_{i_k} \\
... &= ... \\
\sigma_n &= x_1 x_2 ... x_n
\end{aligned}$$

The following theorem presents an extremely useful result about the decomposition of symmetric polynomials.

Theorem 6.18 ()

Every symmetric polynomial can be written as a polynomial of elementary symmetric polynomials σ_i.

Example 6.15 ()

The polynomial

$$f \equiv \sum_{i=1}^{n} x_i^3 \tag{282}$$

can be expressed as

$$f = \sigma_1^3 - 3\sigma_1 \sigma 2 + 3\sigma_3 \tag{283}$$

7 Modules

7.1 Modules over a PID

7.2 Rational Canonical Form

7.3 Jordan Canonical Form

8 Vector Space

Definition 8.1 (Vector Space)

A **vector space over a field** F consists of an abelian group $(V, +)$ and an operation called **scalar multiplication**

$$\cdot : F \times V \to V \tag{284}$$

such that for all $x, y \in V$ and $\lambda, \mu \in F$, we have

1. $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
2. $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
3. $(\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x)$, which equals $(\mu\lambda) \cdot x = \mu \cdot (\lambda \cdot x)$ since F is commutative
4. $1 \cdot x = x$, where 1 is the unity of F

Definition 8.2 ()

A **left R-module** M consists of an abelian group $(M, +)$ and an operation called **scalar multiplication**

$$\cdot : R \times M \longrightarrow M \tag{285}$$

such that for all $\lambda, \mu \in R$ and $x, y \in M$, we have

1. $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
2. $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
3. $(\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x)$, not necessarily equaling $(\mu\lambda) \cdot x = \mu \cdot (\lambda \cdot x)$
4. $1 \cdot x = x$, where 1 is the unity of R

Note that a left R-module is a vector space if and only if R is a field.

Definition 8.3 ()

A **right R-module** M is defined analogously to a left R-module, except that the scalar multiplication operation is defined

$$\cdot : M \times R \longrightarrow M \tag{286}$$

Definition 8.4 ()

Let A be a vector space over a field F equipped with an additional binary operation

$$\times : A \times A \longrightarrow A \tag{287}$$

A is an **algebra over** F if the following identities hold for all $x, y, z \in A$ and all $\lambda, \mu \in F$.

1. Right distributivity. $(x + y) \times z = x \times z + y \times z$
2. Left distributivity. $z \times (x + y) = z \times x + z \times y$
3. Compatibility with scalars. $(\lambda \cdot x) \times (\mu \cdot y) = (\lambda\mu) \cdot (x \times y)$

Note that vector multiplication of an algebra does not need to be commutative.

Example 8.1 ()

The set of all $n \times n$ matrices with matrix multiplication is a noncommutative, associative algebra. Similarly, the set of all linear endomorphisms of a vector space V with composition is a noncommutative, associative algebra.

Example 8.2 ()

$\mathbb{R}^3$ equipped with the cross product is an algebra, where the cross product is **anticommutative**, that is $x \times y = -y \times x$. $\times$ is also nonassociative, but rather satisfies an alternative identity called the **Jacobi Identity**.

Example 8.3 ()

The set of all polynomials defined on an interval $[a, b]$ is an infinite-dimensional subalgebra of the set of all functions $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined on $[a, b]$.

Definition 8.5 ()

Similar to division rings, a **division algebra** is an algebra where the operation of "division" defined as such: Given any $a \in A$, nonzero $b \in A$, there exists solutions to the equation

$$A = bx \tag{288}$$

that are unique. If we wish, we can distinguish left and right division to be the solutions of $A = bx$ and $A = xb$.

Definition 8.6 ()

Here are examples of division algebras.

1. $\mathbb{R}$ is a 1-dimensional algebra over itself.
2. $\mathbb{C}$ is a 2-dimensional algebra over $\mathbb{R}$.
3. There exists no 3-dimensional algebra.
4. Quaternions forms a 4-dimensional algebra over $\mathbb{R}$.

8.1 Modules

Vector space but over a ring.

8.2 Algebras

Vector space with bilinear product.

8.3 The Algebra of Quaternions

Definition 8.7 ()

The **quaternions** form an algebra of 4-dimensional vectors over $\mathbb{R}$, with elements of the form

$$(a, b, c, d) \equiv a + bi + cj + dk \tag{289}$$

where a is called the **scalar portion** and $bi + cj + dk$ is called the **vector/imaginary portion**. The algebra of quaternions is denoted $\mathbb{H}$, which stands for "Hamilton." $\mathbb{H}$ is a 4-dimensional associative normed division algebra over $\mathbb{R}$.

From looking at the multiplication table, we can see that multiplication in $\mathbb{H}$ is not commutative.

i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

Note the identity

$$i^2 = j^2 = k^2 = -1 \tag{290}$$

The algebra of quaternions are in fact the first noncommutative algebra to be discovered!

Proposition 8.1 ()

$\mathbb{H}$ and $\mathbb{C}$ are the only finite-dimensional divisions rings containing $\mathbb{R}$ as a proper subring.

Definition 8.8 ()

The **quaternion group**, denoted Q_8 is a nonabelian group of order 8, isomorphic to a certain 8-element subset in $\mathbb{H}$ under multiplication. It's group presentation is

$$Q_8 = \big\langle \bar{e}, i, j, k \mid \bar{e}^2 = e, i^2 = j^2 = k^2 = ijk = \bar{e} \big\rangle \tag{291}$$

Going back to the algebra, we can set $\{1, i, j, k\}$ as a basis and define addition and scalar multiplication component-wise, and multiplication (called the **Hamilton product**) with properties

1. The real quaternion 1 is the identity element.
2. All real quaternions commute with quaternions: $aq = qa$ for all $a \in \mathbb{R}, q \in \mathbb{H}$.
3. Every quaternion has an inverse with respect to the Hamilton product.

$$(a + bi + cj + dk)^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\big(a - bi - cj - dk\big) \tag{292}$$

Note that property 3 allows $\mathbb{H}$ to be a division algebra.

Proposition 8.2 (Scalar and Vector Components)

Let the quaternion be divided up into a scalar and vector part with the bjective mapping $a + bi + cj + dk \mapsto \big(a, (b, c, d)\big)$.

$$q = (r, v), r \in \mathbb{R}, v \in \mathbb{R}^3 \tag{293}$$

Then, the formulas for addition and multiplication are

$$\begin{aligned} q_1 + q_2 &= (r_1, v_1) + (r_2, v_2) = (r_1 + r_2, v_1 + v_2) \\ q_1 \cdot q_2 &= (r_1, v_1) \cdot (r_2, v_2) = (r_1 r_2 - v_1 \cdot v_2, r_1 v_2 + r_2 v_1 + v_1 \times v_2) \end{aligned}$$

where the $\cdot$ and $\times$ on the right hand side represnts the dot product and cross product, respectively.

Definition 8.9 ()

The conjugate of a quaternion $q = a + bi + cj + dk$ is defined

$$\bar{q}, q^* \equiv a - bi - cj - dk \tag{294}$$

It has properties

1. $q^{**} = q$
2. $(qp)^* = p^* q^*$

q^* can also be expressed in terms of addition and multiplication.

$$q^* = -\frac{1}{2}\big(q + iqi + jqj + kqk\big) \tag{295}$$

Definition 8.10 ()

The **norm** of q is defined

$$||q|| \equiv \sqrt{q^* q} = \sqrt{q q^*} = \sqrt{a^2 + b^2 + c^2 + d^2} \tag{296}$$

with properties

1. Scaling factor. $||\alpha q|| = |\alpha| ||q||$
2. Multiplicative. $||pq|| = ||p|| ||q||$

The norm allows us to define a metric

$$d(p, q) \equiv ||p - q|| \tag{297}$$

This makes $\mathbb{H}$ a metric space, with addition and multiplication continuous on the metric topology.

Definition 8.11 ()

The **unit quaternion** is defined to be

$$U_q = \frac{q}{||q||} \tag{298}$$

Corollary 8.1 ()

Every quaternion has a polar decomposition

$$q = U_q \cdot ||q|| \tag{299}$$

With this, we can redefine the inverse as

$$q^{-1} = \frac{q^*}{||q||^2} \tag{300}$$

8.3.1 Matrix Representations of Quaternions

We can represent q with 2×2 matrices over $\mathbb{C}$ or 4×4 matrices over $\mathbb{R}$.

Proposition 8.3 ()

The following representation is an injective homomorphism $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(2, \mathbb{C})$.

$$\rho : a + bi + cj + dk \mapsto \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} \tag{301}$$

It has properties

1. Constraining any two of b, c, d to 0 produces a representation of the complex numbers. When $c = d = 0$, this is called the **diagonal representation**.
$$\begin{pmatrix} a + bi & 0 \\ 0 & a - bi \end{pmatrix}, \begin{pmatrix} a & c \\ -c & a \end{pmatrix}, \begin{pmatrix} a & di \\ di & a \end{pmatrix}$$
2. The norm of a quaternion is the square root of the determinant of its corresponding matrix representation.

$$||q|| = \sqrt{\det \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}} = \sqrt{(a^2 + b^2) + (c^2 + d^2)} \tag{302}$$

3. The conjugate of a quaternion corresponds to the conjugate (Hermitian) transpose of its matrix representation.

$$\rho(q^*) = \rho(q)^H \iff a - bi - cj - dk \mapsto \begin{pmatrix} a - bi & -c - di \\ c - di & a + bi \end{pmatrix} \tag{303}$$

4. The restriction of this representation to only unit quaternions leads to an isomorphism between the subgroup of unit quaternions and their corresponding image in $\mathrm{SU}(2)$. Topologically, the unit quaternions is the 3-sphere, so the underlying space $\mathrm{SU}(2)$ is also a 3-sphere. More specifically,

$$\frac{\mathrm{SU}(2)}{2} \simeq \mathrm{SO}(3) \tag{304}$$

Proposition 8.4 ()

The following representation of $\mathbb{H}$ is an injective homomorphism $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(4, \mathbb{R})$.

$$\rho : a + bi + cj + dk \mapsto \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \tag{305}$$

or also as

$$a \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{306}$$

It has properties

1. $\rho(q^*) = \rho(q)^T$
2. The fourth power of the norm is the determinant of the matrix

$$||q||^4 = \det\big(\rho(q)\big) \tag{307}$$

3. Similarly, with the 2×2 representation, complex number representations can be produced by restricting 2 of b, c, d to 0.

Note that this representation in $\mathrm{GL}(4, \mathbb{R})$ is not unique. There are in fact 48 distinct representation of this form where one of the component matrices represents the scalar part and the other 3 are skew symmetric.

8.3.2 Square Roots of -1

In $\mathbb{C}$, there are two numbers, i and $-i$, whose square is -1. However, in $\mathbb{H}$, infinitely many square roots of -1 exist, forming the unit sphere in $\mathbb{R}^3$. To see this, let $q = a + bi + cj + dk$ be a quaternion, and assume that its square is -1. Then this implies that

$$a^2 - b^2 - c^2 - d^2 = -1, 2ab = 2ac = 2ad = 0 \tag{308}$$

To satisfy the second equation, either $a = 0$ or $b = c = d = 0$. The latter is impossible since then q would be real. Therefore,

$$b^2 + c^2 + d^2 = 1 \tag{309}$$

which forms the unit sphere in $\mathbb{R}^3$.

8.4 Tensor Algebras

Remember that an algebra is (loosely) a vector space V with a multiplication operation

$$\times : V \times V \longrightarrow V \tag{310}$$

Definition 8.12 ()

The **tensor algebra** of vector space V over field $\mathbb{F}$ is

$$\begin{aligned} T(V) \equiv \bigoplus_{n=0}^{\infty} V^{\otimes n} &= V^{\otimes 0} \oplus V^{\otimes 1} \oplus V^{\otimes 2} \oplus V^{\otimes 3} \oplus ... \\ &= \mathbb{F} \oplus V \oplus V^{\otimes 2} \oplus V^{\otimes 3} \oplus V^{\otimes 4} \oplus ... \end{aligned}$$

with elements being infinite-tuples

$$(a, B^{\mu}, C^{\nu\gamma}, D^{\alpha\beta\epsilon}, ...) \tag{311}$$

The addition operation is defined component-wise, and the multiplication operation is the tensor product

$$\otimes : T(V) \times T(V) \longrightarrow T(V) \tag{312}$$

and the identity element is

$$I = (1, 0, 0, ...) \tag{313}$$

Linearity can be easily shown.

The tensor algebra is often used to "add" differently ranked tensors together. But in order to do this rigorously, we must define the canonical injections

$$i_j : V^{\otimes j} \longrightarrow T(V),\ i_j(T^{\kappa_1, ..., \kappa_j}) = (0, ..., 0, T^{\kappa_1, ..., \kappa_j}, 0, ..., 0) \tag{314}$$

shown in the diagram

$$\begin{tikzcd} & & T(V) & & \\ \mathbb{F} \arrow[urr, "i_0"] & V \arrow[ur, "i_1"] & V^{\otimes 2} \arrow[u, "i_2"] & V^{\otimes 3} \arrow[ul, "i_3"'] & ... \arrow[ull] \end{tikzcd}$$

Therefore, with these i_j's, we can implicitly define the addition of arbitrary tensors $A \in V^{\otimes n}$ and $B \in V^{\otimes m}$ as

$$A + B \equiv i_n(A) + i_m(B) \in T(V) \tag{315}$$

along with multiplication of tensors as

$$A \otimes B \equiv i_n(A) \otimes i_m(B) \equiv i_{n+m}(A \otimes B) \tag{316}$$

We can also redefine the tensor product operation between two spaces to be an operation within $T(V)$ itself.

$$i_i(V^{\otimes i}) \otimes i_j(V^{\otimes j}) = i_{i+j}(V^{\otimes (i+j)}) \tag{317}$$

We can now proceed to define Exterior and Symmetric algebras as quotient algebras.

Definition 8.13 ()

The **exterior algebra** $\Lambda(V)$ of a vector space V over field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

$$\Lambda(V) \equiv \frac{T(V)}{I} \tag{318}$$

where I is the two-sided ideal generated by all elements of the form $x \otimes x$ for $x \in V$ (i.e. all tensors that can be expressed as the tensor product of a vector in V by itself).
The **exterior product** $\wedge$ of two elements of $\Lambda(V)$ is the product induced by the tensor product $\otimes$ of $T(V)$. That is, if

$$\pi : T(V) \longrightarrow \Lambda(V) \tag{319}$$

is the canonical projection/surjection and $a, b \in \Lambda(V)$,then there are $\alpha, \beta \in T(V)$ such that $a = \pi(\alpha), b = \pi(\beta)$, and

$$a \wedge b = \pi(\alpha \otimes \beta) \tag{320}$$

We can define this quotient space with the equivalence class

$$x \otimes y = -y \otimes x \pmod I \tag{321}$$

Definition 8.14 ()

The **symmetric algebra** $\mathrm{Sym}(V)$ of a vector space V over a field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

$$\Lambda(V) \equiv \frac{T(V)}{J} \tag{322}$$

where J is the two-sided ideal generated by all elements in the form

$$v \otimes w - w \otimes v \tag{323}$$

(i.e. commutators of all possible pairs of vectors).

9 Fields and Galois Theory

Now that we have established the theory of general rings and for polynomials, we will delve deeper into the theory of fields by talking about *field extensions*, which allows us to consider minimal fields in which polynomials can be decomposed into linear factors. This is called a *splitting field* and also turns out to have a vector space structure as well. This is similar to how we have constructed ring extensions, and we would like to find conditions in which an adjoining ring is a field.

9.1 Field Extensions and Vector Spaces

This method in which we have taken higher powers of α to reveal elements in $\mathbb{Q}$ reveals a deeper structure of a finite-dimensional vector space, which will be useful for analyzing certain fields in the examples below. Note that a vector space is only well-defined over a field F, so we will consider *field extensions* from now on. First, recall that a field is trivially a vector space.

Lemma 9.1 (Fields are a Vector Space)

A field F is a 1-dimensional vector space over itself.

It turns out that we can generalize this a bit more.

Theorem 9.1 (Fields are a Vector Space over Subfields)

Let F be a subfield of K. Then K is a F-vector space.

Proof.

A F-vector space has 0, addition, and multiplication by F. K indeed has 0, addition, and we can multiply any element of K by an element of F. The extra axioms follow but are too verbose to write a full proof.

Corollary 9.1 ()

$\mathbb{R}$ is an infinite-dimensional vector space over $\mathbb{Q}$.

Proof.

The fact that it is a vector space immediately follows from $\mathbb{Q} \hookrightarrow \mathbb{R}$. For dimensionality, the outline is to show that $\{\sqrt{p} \mid p \text{ prime }\}$ are linearly independent. This takes work to prove and won't do it.

Therefore, by constructing a subfield, we can model the original field as a vector space. This additional structure warrants a name.

Definition 9.1 (Field Extension)

If F, K are fields, then this is called a **field extension**. Its **degree** is the F-dimension of K, denoted

$$[K : F] \coloneqq \dim_F(K) \tag{324}$$

So when we are given a field, we can automatically treat it as a vector space. Furthermore, if we are given a field extension, we can treat the larger field as a vector space over the smaller field (though it may be finite or infinite-dimensional). As we create concatenated field extensions, sometimes called a *tower*, the dimensions behave nicely as well.

Theorem 9.2 ()

$F \hookrightarrow E$ and $F \hookrightarrow K$ are field extensions. Then $E \hookrightarrow K$ is a field extension with degree

$$[K : E] = [K : E][E : F] \tag{325}$$

Proof.

Let $\alpha_1, \ldots \alpha_m$ be a basis for E over F and $\beta_1, \ldots, \beta_n$ be a basis for K over E. We claim that $\{\alpha_i \beta_j\}$ is a basis for K over F, with multiplication done in the field K. We check linear indepdence. Let $\beta \in K$ be arbitrary. Then by the E-basis, we have

$$\beta = \sum_{j=1}^{n} x_j \beta_j \tag{326}$$

But since $x_j \in E$, there are elements

$$x_j = \sum_{i=1}^{m} y_{ij} \alpha_i \tag{327}$$

and so combining we get

$$\beta = \sum_{j=1}^{n} \sum_{i=1}^{m} y_{ij} (\alpha_i \beta_j) \tag{328}$$

To prove linear independence, suppose $\beta = 0$. Then we have

$$0 = \sum_{j=1}^{n} \sum_{i=1}^{m} y_{ij} (\alpha_i \beta_j) = \sum_{j=1}^{n} \bigg(\sum_{i=1}^{m} y_{ij} \alpha_i \bigg) \beta_j \tag{329}$$

Since $\beta_1, \ldots, \beta_n$ are linearly indpendent, we must have $\sum_{i=1}^{m} y_{ij} \alpha_i = 0$ for all j. But since α_i's are linear independent, this means $y_{ij} = 0$ for all i, j.

In fact, there is a nice classification algorithm that connects fields to vector spaces.

Theorem 9.3 ()

Once we know that field extensions are vector spaces, what consistutes linear maps? A first guess would be a ring homomorphism, but this may not be true.

Example 9.1 (Ring Homomorphisms are Not Always Linear Maps)

Therefore, we need some additional constraint.

Theorem 9.4 (Ring Homomorphisms as Linear Maps)

Let $F \subset K_1$ and $F \subset K_2$ be field extensions and $f : K_1 \rightarrow K_2$ be a ring-homomorphism. If $f(x) = x$ for all $x \in F$, then f is a linear map of F-vector spaces.

9.2 Adjoining Fields and Quotient Rings

We have examined the properties of field extensions in general. Now we consider the specific extension $F \subset F[\alpha]$. What we should think is that $F[\alpha]$ ends up becoming both a ring and a vector space, but not yet a field. It is only when α is the root of an irreducible polynomial in $F(x)$ that $F[\alpha]$ is a field.

Lemma 9.2 (Vector Space Structure)

If F is a field,

1. $F[\alpha]$ is a finite-dimensional vector space over F.
2. If $f(x) = a_n x^n + \ldots + a_0$ is any polynomial with root α, then $\{1, \alpha, \ldots, \alpha^{n-1}\}$ spans $F[\alpha]$.[a]

[a] Note that this does not mean that it is a basis.

Proof.

An element of $F[\alpha]$ is of the form

$$f(\alpha) = \sum_{k=0}^{n} a_k \alpha^k \tag{330}$$

for some $f \in F[x]$, and so it is immediate that $\{\alpha^k\}_{k \in \mathbb{N}_0}$ spans $F[\alpha]$. We claim that α^{n-1+i} is in S for all $i > 0$. By induction, if $i = 1$, then

$$\alpha^n = -\frac{1}{a_n}\big(a_{n-1}\alpha^{n-1} + \ldots + a_0\big) \tag{331}$$

which proves the claim. Now assume that $\alpha^n, \alpha^{n+1}, \ldots, \alpha^{n-1+i} \in \operatorname{span}\{1, \ldots, \alpha^{n-1}\}$. Then

$$\alpha^i f(\alpha) = 0 \implies a_n \alpha^{n+i} + \alpha_{n-1}\alpha^{n+i-1} + \ldots + a_0 \alpha^i = 0 \tag{332}$$

and so

$$\alpha^{n+i} = -\frac{1}{a_n}\big(a_{n-1}\alpha^{n+i-1} + \ldots + a_0\alpha^i\big) \tag{333}$$

which means that $\alpha^{n+i} \in \operatorname{span}\{1, \ldots, \alpha^{n-1}\}$, completing the proof.

Great, so we know that $F[\alpha]$ is a finite-dimensional vector space. What additional constrains do we need for it to be a field? With one more assumption, we can claim that it is a field, giving a subfield substructure $F \subset F[\alpha] \subset K$

Theorem 9.5 (Adjoining Fields)

Let $F \hookrightarrow K$ be a field extension.

1. If there exists a $f(x) \in F[x]$ s.t. $\alpha \in K$ is a root of f, then $F[\alpha] \subset K$ is a field. To emphasize that it is a field, we usually denote it as $F(\alpha)$ and refer it as the **adjoining field**.
2. If there exists a $f(x) \in F[x]$ is irreducible of degree n, then $[F[\alpha] : F] = n$.

Proof.

It is clear that $F[\alpha]$ is a commutative ring since F is a field. So it remains to show that every nonzero element of $\beta \in F[\alpha]$ is a unit. By definition $\beta = p(\alpha)$ for some polynomial $p \in F[x]$. Factor $f \in F[x]$ as the product of irreducible polynomials. Then α must be a root of one of those irreducible factors, say $g(x)$. Note that $g(x) \nmid p(x)$ since $p(\alpha) \neq 0$. Since g is irreducible, we know that $\gcd(g, p) = 1$ and so $\exists s, t \in F[x]$ s.t.

$$1 = sp + tg \implies 1 = s(\alpha)p(\alpha) + t(\alpha)g(\alpha) = s(\alpha)p(\alpha) \tag{334}$$

Therefore we have found a multiplicative inverse $s = p^{-1} \in F[\alpha]$.

Proof.

We can also prove it using the vector space structure, which requires the lemma below. Treating $F[\alpha]$as a finite-dimensional vector space over F, let us define the F-linear function[a]

$$m_b : F[\alpha] \to F[\alpha], \qquad m_b(\beta) = b\beta \tag{335}$$

Since $F[\alpha] \subset K$, $F[\alpha]$ is an integral domain. Thus $\nexists \beta \in F[\alpha] \setminus \{0\}$ s.t. $b\beta = 0$. This means that the kernel of m_b is 0, and so m_b is injective. By the rank-nullity theorem, it is bijective, and so there exists a $\beta \in F[\alpha]$ s.t. $b\beta = 1 \implies b$ is a unit.

[a] linearity is easy to check

Intuitively, the extra $\alpha \in K$ allows us to "expand" our field F into a bigger field of K. Since $F[\alpha]$ is the smallest ring containing both F and α, it immediately follow that it is the smallest *field* containing F and α. If we examine subfields of $\mathbb{C}$, this is equivalent to α being an *algebraic number* (i.e. α must be a root of some polynomial with rational—or equivalently, integer—coefficients).

Corollary 9.2 (Adjoining with Algebraic Numbers Creates Fields)

Let $\alpha \in \mathbb{C}$. Then $\mathbb{Q}[\alpha] \subset \mathbb{C}$ is a field if and only if α is an algebraic number.

Corollary 9.3 ()

Suppose $[K : F] = n$ and $\alpha \in K$ is the root of an irreducible polynomial $f(x) \in F[x]$. Then $\deg f(x) \mid n$.

Example 9.2 ($\mathbb{Q}[\sqrt{3}i]$ is a Field)

$\mathbb{Q}[\sqrt{3}i]$ is a field, hence denoted $\mathbb{Q}(\sqrt{3}i)$ since $\sqrt{3}i$ is a root of the polynomial $f(x) = x^2 + 3$.

Example 9.3 ($\mathbb{Q}[\pi]$ not a Field)

However, $\mathbb{Q}[\pi]$ is not a field.

Example 9.4 (Finding Multiplicative Inverses of elements in $\mathbb{Q}[\alpha]$)

Given $\beta = p(\alpha) = \alpha^2 + \alpha - 1 \in \mathbb{Q}[\alpha]$, where α is a root of $f(\alpha) = \alpha^3 + \alpha + 1$, we first know that β must have a multiplicative inverse since $\mathbb{Q}[\alpha]$ is a field. Applying the Euclidean algorithm, we have

$$1 = \frac{1}{3}\big\{(x+1)f(x) - (x^2+2)p(x)\big\} = -\frac{1}{3}(\alpha^2+2)p(\alpha) \tag{336}$$

and so $\beta^{-1} = (\alpha^2 + \alpha - 1)^{-1} = -\frac{1}{3}(\alpha^2 + 2)$. We can check that

$$\begin{aligned} -\frac{1}{3}(\alpha^2+2)(\alpha^2+\alpha-1) &= -\frac{1}{3}(\alpha^4+\alpha^3+\alpha^2+2\alpha-2) && (337)\\ &= -\frac{1}{3}(\alpha^3+\alpha-2) && (338)\\ &= -\frac{1}{3}(-3) = 1 && (339) \end{aligned}$$

Great, so we have seen examples of both ring and field extensions and seen a condition that makes $F[\alpha]$ a field as well. Now recall quotient rings, which do not necessarily preserve the properties of the original ring.

That is, if F is a field, then F/I may not be a field. Using the fundamental ring homomorphism theorem, we can precisely correlate certain quotient maps with adjoining fields.

Lemma 9.3 (Evaluation Homomorphism of Polynomials)

Given fields $F \subset K$, the **evaluation function** is a homomorphism.

$$\text{ev}_\alpha : F[x] \rightarrow K, \qquad f(x) \mapsto f(\alpha) \tag{340}$$

Proof.

Theorem 9.6 (Quotient Polynomial Ring Can be Splitting Field)

Let F be a field and $f(x) \in F[x]$, then

$$f(x) \text{ is irreducible in } F[x] \iff K = \frac{F[x]}{\langle f(x) \rangle} \text{ is a field} \tag{341}$$

and if either condition is satisfied, then K must contain a root α of $f(x)$, and

$$\frac{F[x]}{\langle f(x) \rangle} \simeq F[\alpha] \tag{342}$$

making $F \subset K$ a field extension. It immediately follows then as vector spaces,

$$\dim_F \frac{F[x]}{\langle f(x) \rangle} = \deg(f) \tag{343}$$

Proof.

Prove that a root α of $f(x)$ is contained in K.
If K is a field, then $F \subset K$ is a field extension. Then let $\alpha \in K$ be a root of $f(x)$. Then $F[\alpha]$ is a field.
We can see that $F \subset F[x]$. Therefore if $f(x)$ is irreducible, then $\alpha \in K$ By fundamental homomorphism theorem.
For dimension, we know that $\{1, \ldots, x^{n-1}\}$ is a basis.

Example 9.5 (Quotient Rings as Fields)

Since $x^2 + 1$ is irreducible in $\mathbb{Z}_7[x]$ and $\mathbb{R}$, the following are fields.

$$\frac{\mathbb{Z}_7[x]}{\langle x^2 + 1 \rangle}, \quad \frac{\mathbb{R}[x]}{\langle x^2 + 1 \rangle} \tag{344}$$

With this theorem, we can use it cleverly to prove that two rings are isomorphic to each other.

Example 9.6 ()

The evaluation map

$$\phi : \frac{\mathbb{R}[x]}{\langle x^2 + 1 \rangle} \rightarrow \mathbb{C}, \qquad \phi\big(f(x) \pmod{\langle x^2 + 1 \rangle}\big) = f(i) \tag{345}$$

is an isomorphism.[a] This is because we can think of the evaluation homomorphism $\text{ev}_i : f(x) \in \mathbb{R}[x] \mapsto f(i) \in \mathbb{R}[i]$. We know that $\mathbb{R}$ a field implies $\mathbb{R}[x]$ is a PID. Now take $\ker(\text{ev}_i)$. We can see that it contains the polynomial $x^2 + 1$, and since it is irreducible in $\mathbb{R}[x]$, it must be the case that $\ker(\text{ev}_i) = \langle x^2 + 1 \rangle$. Now it follows by the fundamental ring homomorphism theorem that

$$\frac{\mathbb{R}[x]}{\ker(\text{ev}_i)} = \frac{\mathbb{R}[x]}{\langle x^2 + 1 \rangle} \simeq \mathbb{R}[i] = \mathbb{C} \tag{346}$$

[a] Intuitively, we can see that the quotient ring can only consist up to linear polynomials since $x^2 \equiv -1$. This is a real vector space of dimension 2, and so is $\mathbb{C}$, so it makes sense that they may be isomorphic.

Example 9.7 ()

The evaluation map

$$\text{ev}_{\sqrt{2}} : \mathbb{Q}[x] \mapsto \mathbb{Q}[\sqrt{2}], \qquad \text{ev}_{\sqrt{2}}(f) = f(\sqrt{2}) \tag{347}$$

is a homomorphism. Furthermore, it has a kernel $\langle x^2 - 2 \rangle$ since $(x^2 - 2)$ is an irreducible polynomial in $\mathbb{Q}[x]$ containing the root $\sqrt{2}$. Therefore by the fundamental ring homomorphism theorem we have

$$\frac{\mathbb{Q}[x]}{\langle x^2 - 2 \rangle} \simeq \mathbb{Q}[\sqrt{2}] \tag{348}$$

Example 9.8 (Extensions of $\sqrt{2}$ and i)

We claim that

$$\mathbb{Q}[\sqrt{2}, i] = \{a + b\sqrt{2} + ci + d(\sqrt{2}i) \mid a, b, c, d \in \mathbb{Q}\} \tag{349}$$

From the previous example, we know that $\mathbb{Q}[\sqrt{2}]$ are all numbers of the form $a + b\sqrt{2}$. Now we take $i \in \mathbb{C}$ and map it through all polynomials with coefficients in $\mathbb{Z}[\sqrt{2}]$, which will be of form

$$f(i) = (a_n + b_n\sqrt{2})i^n + (a_{n-1} + b_{n-1}\sqrt{2})i^{n-1} + \ldots + (a_2 + b_2\sqrt{2})i^2 + (a_1 + b_1\sqrt{2})i + (a_0 + b_0\sqrt{2}) \tag{350}$$

However, we can see that since $i^2 = -1$, we only need to consider up to degree 1 polynomials of form

$$(a + b\sqrt{2}) + (c + d\sqrt{2})i \tag{351}$$

which is clearly of the desired form. For the other way around, this is trivial since we can construct a linear polynomial as before.

Example 9.9 ()

We claim $\mathbb{Q}[\sqrt{3} + i] = \mathbb{Q}[\sqrt{3}, i]$.

1. $\mathbb{Q}[\sqrt{3} + i] \subset \mathbb{Q}[\sqrt{3}, i]$
2. $\mathbb{Q}[\sqrt{3} + i] \supset \mathbb{Q}[\sqrt{3}, i]$. Note that

$$(\sqrt{3} + i)^3 = 8i \implies i \in \mathbb{Q}[\sqrt{3} + i] \tag{352}$$

$$\implies (\sqrt{3} + i) - i = \sqrt{3} \in \mathbb{Q}[\sqrt{3} + i] \tag{353}$$

Therefore, $\mathbb{Q}[\sqrt{3} + i]$ contains the elements $1, \sqrt{3}, i$, which form the basis of $\mathbb{Q}[\sqrt{3}, i]$.

Example 9.10 (Extensions of $\sqrt{3}i$ and $\sqrt{3}, i$)

We claim that $\mathbb{Q}[\sqrt{3}i] \subsetneq \mathbb{Q}[\sqrt{3}, i]$.

1. We can see that $\{1, \sqrt{3}i\}$ span $\mathbb{Q}[\sqrt{3}i]$ as a $\mathbb{Q}$-vector space. Therefore,
$$\sqrt{3}, i \in \mathbb{Q}[\sqrt{3}, i] \implies \sqrt{3}i \in \mathbb{Q}[\sqrt{3}, i] \tag{354}$$
implies that $\mathbb{Q}[\sqrt{3}i] \subset \mathbb{Q}[\sqrt{3}, i]$.
2. To prove proper inclusion, we claim that $i \not\in \mathbb{Q}[\sqrt{3}i]$. Assuming that it can, we represent it in the basis $i = b_0 + b_1\sqrt{3}i$, and so
$$-1 = (b_0 + b_1\sqrt{3}i)^2 = (b_0^2 - 3b_1^2) + 2b_0b_1\sqrt{3}i \tag{355}$$
Therefore we must have $2b_0b_1\sqrt{3} = 0 \implies b_0$ or b_1 should be 0. If $b_0 = 0$, then $b_0^2 - 3b_1^2 = -3b_1^2 \implies b_1^2 = 1/3$, which is not possible since $b_1^2 \in \mathbb{Q}$. If $b_1 = 0$, then $b_0 - 3b_1^2 = b_0^2 > 0$, and so it cannot be -1.

Theorem 9.7 ()

Given field $F = \mathbb{R}[x]/\langle x^2+1\rangle$[a]

1. F is a finite dimensional vector space over $\mathbb{R}$.
2. F is an infinite dimensional vector space over $\mathbb{Q}$.

[a]This is a field since x^2+1 is irreducible in $\mathbb{R}[x]$.

Proof.

Since $F \simeq \mathbb{R}[i] \supset \mathbb{R} \supset \mathbb{Q}$, F is a vector space over its subfields $\mathbb{R}$ and $\mathbb{Q}$. It suffices to prove dimensionality.

1. For (1), we claim that $F \simeq \mathbb{R}^2$.
2. For (2), we note that $F \supset \mathbb{R}$ and $\mathbb{R}$ is infinite dimensional over $\mathbb{Q}$, so it follows for F.

9.3 Splitting Fields

Remember that by previously establishing that $\mathbb{C}$ is algebraically closed, this gives us a "safe space" to work in, in the sense that if we take any subfield $F \subset \mathbb{C}$ and find a polynomial $f(x) \in F[x]$, we are *guaranteed* to find a linear factorization of f in $\mathbb{C}[x]$.

Therefore, if K is algebraically closed and $F \subset K$ is a field extension, $f(x) \in F[x]$ is guaranteed to *split* completely into linear factors. This is true for *all* $f(x) \in F[x]$, but now if we *fix* $f(x) \in F[x]$, perhaps we don't need the entire field K to split $f(x)$. Maybe we can work in a slightly larger field E—such that $F \subset E \subset K$—where $f(x)$ splits in E. This process of finding such a minimal field is important to understand the behavior of roots of such polynomials.

Definition 9.2 (Splitting Field)

Given a field extension $F \subset K$ and a polynomial $f \in F[x]$,

1. f **splits** in K if f can be written as the product of linear polynomials in $K[x]$.
2. If f splits in K and there exists no field E s.t. $F \subsetneq E \subsetneq K$, then K is called a **splitting field** of f.[a]

[a]i.e. the splitting field is the smallest field that splits f.

Example 9.11 (Don't Need Necessarily Complex Numbers to Split)

Consider the following.

1. Let $f(x) = x^2 - 1$. If $f(x) \in \mathbb{R}[x]$, it does split in $\mathbb{R}$. In fact, even if we consider it as an element of $\mathbb{Z}_2[x]$, it still splits into $(x+1)(x-1)$.
2. Let $f(x) = x^2 - 2$. If $f(x) \in \mathbb{Q}[x]$, it doesn't split in $\mathbb{Q}$ since the roots $\pm\sqrt{2} \not\in \mathbb{Q}$, but $\pm\sqrt{2}$ are real numbers, so $f(x)$ does in fact split in $\mathbb{R}$ since it splits into $(x+\sqrt{2})(x-\sqrt{2})$. However, maybe it is not the (smallest) splitting field.
3. Let $f(x) = x^2 + 1$. We can see that if we consider it as an element of $\mathbb{Q}[x]$ or $\mathbb{R}[x]$, neither fields split $f(x)$ since $\pm i$ are its roots and therefore are contained in the coefficients of its linear factors. We know that it definitely splits in $\mathbb{C}$, but can we find a smaller field that splits $f(x)$? Perhaps.

So how does one find a splitting field? Note that in the example above, we have found that there were some roots α of certain polynomials $f(x) \in F[x]$ are not contained in F. Therefore, what we want to do is find the smallest field F containing both F and α (plus any other α's). This is precisely the adjoining field $F[\alpha]$, which guarantees unique factorization since $F[\alpha]$ is a Euclidean domain.

Corollary 9.4 ()

Any polynomial $f(x) \in F[x]$ has a unique splitting field.

Example 9.12 (Simple Splitting Fields)

We provide some simple examples to gain intuition.

1. Let $f(x) = x^2 + 2x + 2 \in \mathbb{Q}[x]$. Then the roots of $f(x)$ are $-1 \pm i$, so
$$f(x) = (x - (-1+i))(x - (-1-i)) \tag{356}$$
and we can show that $\mathbb{Q}[-1-i, -1+i] = \mathbb{Q}[i]$ is the splitting field of f.
2. Let $f(x) = x^2 - 2x - 1 \in \mathbb{Q}[x]$. The roots are $1 \pm \sqrt{2}$, and so
$$f(x) = (x - (1+\sqrt{2}))(x - (1-\sqrt{2})) \tag{357}$$
and so $\mathbb{Q}[\sqrt{2}]$ is the splitting field of f.
3. Let $f(x) = x^6 - 1 \in \mathbb{Q}[x]$. We can factor
$$f(x) = (x-1)(x+1)(x^2+x+1)(x^2-x+1) \tag{358}$$
and the non-rational roots are $\frac{\pm 1 \pm \sqrt{3} i}{2}$. Thus the splitting field of f is $\mathbb{Q}[\sqrt{3}i]$.

Example 9.13 ()

Let $f(x) = x^4 - 2 \in \mathbb{Q}[x]$. It follows that the roots are

$$\{\sqrt[4]{2}, \sqrt[4]{2}, -\sqrt[4]{2}, -\sqrt[4]{2}i\} = \Big\{\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\Big\} \tag{359}$$

thus the splitting field of f is

$$\mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) \subset \mathbb{Q}(\sqrt[4]{2}, e^{\frac{2\pi i}{4}}) \tag{360}$$

since $\sqrt[4]{2}e^{\frac{m\pi i}{4}} \in \mathbb{Q}(\sqrt[4]{2}, e^{\frac{2\pi i}{4}})$. In fact, the two are equal, and to prove this we can see that since we are working in a field,

$$e^{2\pi i/4} = \frac{\sqrt[4]{2}e^{2\pi i/4}}{\sqrt[4]{2}} \in \mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) \tag{361}$$

which implies that $\sqrt[4]{2} \in \mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big)$. Therefore we can conclude that the splitting field is

$$\mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) = \mathbb{Q}\big(\sqrt[4]{2}, e^{\frac{2\pi i}{4}}\big) \tag{362}$$

Example 9.14 (Multivariate Splitting Fields)

$x^3 - 2 \in \mathbb{Q}[x]$ has a splitting field. By Eisenstein, $x^3 - 2$ is irreducible, and so

$$F = \frac{\mathbb{Q}[x]}{\langle x^3 - 2 \rangle} \tag{363}$$

is a field. Furthermore, there must be a root $\alpha \in F$. We factor this in the field F to get

$$x^3 - 2 = (x - \alpha)(x^2 + \alpha x + \alpha^2) \tag{364}$$

Then either $x^2 + \alpha x + \alpha^2$ has a root in F and F is the splitting field, or it is irreducible. It turns out that $x^2 + \alpha x + \alpha^2$ is irreducible in F, and so the splitting field is

$$E = \frac{F[y]}{\langle y^2 + \alpha y + \alpha^2 \langle} \simeq \frac{\mathbb{Q}[x,y]}{\langle x^3 - 2, y^2 + xy + x^2 \rangle} \tag{365}$$

9.4 Finite Fields

We know that a field—as an integral domain—has characteristic 0 or prime p. We also know that a field is a vector space, at least over itself. But given the characteristic of a field, we can model it as a vector space over two very specific fields.

Theorem 9.8 (Characteristic Determines Base Field of Vector Space)

Given a field F,

1. If $\text{char}(F) = p$, then F is a vector space over $\mathbb{Z}_p$.
2. If $\text{char}(F) = 0$, then F is a vector space over $\mathbb{Q}$.

Proof.

Therefore, just from the characteristic we can classify all fields as vector spaces over either $\mathbb{Q}$ or $\mathbb{Z}_p$. Now if we focus on finite fields, we can do a reverse classification.

Theorem 9.9 (Finite Fields Have Cardinality p^d)

Let F be a finite field. Then $|F| = p^n$ for some $n \in \mathbb{N}$.

Proof.

F is a vector space over $\mathbb{Z}_p$ from 9.4. Since F has finitely many elements, F has a finite spanning set, which implies $\dim_{\mathbb{Z}_p} F \leq +\infty$. Let d be the dimension and $\{b_1, \ldots, b_d\}$ be the basis. The elements of F are

$$a_1b_1 + \ldots + a_db_d \tag{366}$$

with $a_1, \ldots a_d \in \mathbb{Z}_p$. Thus there are p^d elements of F, so $F \simeq \mathbb{Z}_p^d$.

In fact, for *every* prime power there exists a unique field. Therefore we can create a bijection by proving the converse.

Theorem 9.10 (Field for Every p^d)

For every prime p and $n \in \mathbb{N}$, there exists a field with $q = p^d$ elements, unique up to isomorphism.

Proof.

Let $f(x) = x^q - x \in \mathbb{Z}_p[x]$. Then this polynomial has a splitting field $K \supset \mathbb{Z}$. Now we claim the roots of $f(x)$ in K are distinct and form a subfield $F_q \subset K$. This will complete the proof since $F_q \subset K$ and $K \subset F_q \implies K = F_q$. Assume $\alpha, \beta \in K$ are roots of $f(x)$, and so $\alpha^p = \alpha$ and $\beta^p = \beta$

1. $\alpha + \beta \in K$ since by a modification of Freshman's dream, $(\alpha+\beta)^p = \alpha^p + \beta^p = \alpha + \beta$.[a]
2. $(-\alpha)^q = (-1)^q \alpha^q = (-1)^q \alpha = -\alpha$ since $-1 = 1$ or q is odd.
3. $\alpha\beta \in K$ since $\mathbb{Z}_p$ is a field and so $(\alpha\beta)^p = \alpha^p \beta^p = \alpha\beta$.
4. For multiplicative inverses, let $\alpha \neq 0$. Then $(\alpha^{-1})^p = (\alpha^p)^{-1} = \alpha^{-1}$.
5. For all p, 0 and 1 are roots so $0, 1 \in K$.

Now we show that K consists of distinct roots. Certainly $0 \in K$ with multiplicity 1 since $f(x) = x(x^{q-1} - x)$. Now suppose nonzero $r \in K$ is a root with multiplicity m. The multiplicity of r is the multiplicity of 0 of

$$f(x+r) = (x+r)^q - (x+r) = x^q + r^q - x - r = x^q - x \tag{367}$$

where the final step follows from $0 = r^q - r$ since $r \in K$. Therefore r has multiplicity 1. Since $K[x]$ has unique factorization property, it follows that $m = 1$ and every r is a simple root.
To show that every field with p^n elements is unique, let F be such a field. We claim that $\text{char}(F) = p \implies \mathbb{Z}_p \subset F$. We claim that every element of F is a root of $f(x) = x^q - x \in \mathbb{Z}_p[x]$, where F is the splitting field. Let $G = F^*$ be the multiplicative group of units. Since F is a field, then $|F^*| = |F| - 1 = p^d - 1$, and by constructing the cyclic group $\langle g \rangle \subset G$ for any $g \in G$, we know by Lagrange's theorem that $g^{|G|} = 1_G$, which implies that for all $x \in F$,

1. If $x \neq 0$ then $x^{p^d - 1} = x \implies x^{p^d} = x$ and so $x \in K$.
2. If $x = 0$ then $x^{p^d} - x = 0$ and so $x \in K$.

Therefore $F \subset K$ with $|F| = |K|$ both finite, and so $F = K$.

[a] We induct on n for $q = p^n$. For $n = 1$, this is trivial by Freshmans dream. Now assume it holds for some $n \in \mathbb{N}$. Then $(x+y)^{p^{n+1}} = ((x+y)^{p^n})^p = (x^{p^n} + y^{p^n})^p = (x^{p^n})^p + (y^{p^n})^p = x^{p^{n+1}} + y^{p^{n+1}}$.

From this, we can write for every prime p and natural n the finite field of order p^n as $\mathbb{F}_{p^n}$. It is clear that if $n = 1$ then $\mathbb{F}_p \simeq \mathbb{Z}_p$. The final result we will show is a hierarchy of subfields.

Theorem 9.11 (Hierarchy of Fields)

For a given prime p, if $p^m < p^n$, then

$$F_{q^m} \subset F_{q^n} \iff m \mid n \tag{368}$$

9.5 Galois Theory

Definition 9.3 (Minimal Polynomial)

Now we unify the three ideas of groups, polynomials, and fields.

Definition 9.4 (Field Embedding)

Let $F \subset K_1, K_2$ be two field extensions. An F**-embedding** is a ring homomorphism $\phi : K_1 \to K_2$ such that $\phi(a) = a$ for all $a \in F$. An F**-automorphism of** K is an F-embedding $\phi : K \to K$.

Definition 9.5 (Galois Group)

The group of all F-automorphisms of K with composition is called the **Galois group of** K **over** F, denoted $G(K/F)$.

Proof.

This is indeed a group under composition. The identity map $\iota \in G(K/F)$. The composition is clearly closed. Now given that $\phi \in G(K/F)$, ϕ^{-1} is also an automorphism that is constant on F, so it is also in $G(K/F)$.

Essentially, the Galois group is a subgroup of the ring automorphism group of K that doesn't vary $F \subset K$. Let's provide a few examples to derive some of the Galois groups.

Example 9.15 (Galois Group of $\mathbb{Q}(\sqrt{2})$ over $\mathbb{Q}$)

Given the field extension $\mathbb{Q} \subset \mathbb{Q}(\sqrt{2})$, we have for any $a, b \in \mathbb{Q}$ and $\phi \in G(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$, we have

$$\phi(a + b\sqrt{2}) = \phi(a) + \phi(b\sqrt{2}) = a + b\phi(\sqrt{2}) \tag{369}$$

So ϕ is completely determined by the value of $\phi(\sqrt{2})$. Now let $\phi(\sqrt{2}) = \alpha + \beta\sqrt{2}$ for some $\alpha, \beta \in \mathbb{Q}$. We have

$$2 = \phi(2) = \phi(\sqrt{2}\sqrt{2}) = \phi(\sqrt{2})^2 = (\alpha + \beta\sqrt{2})^2 = (\alpha^2 + 2\beta^2) + 2\alpha\beta\sqrt{2} \tag{370}$$

Therefore $\alpha\beta = 0$ and $\alpha^2 + 2\beta^2 = 2$. With further casework, we must have $\alpha = 0, \beta = \pm 1$. In conclusion, there are exactly two $\mathbb{Q}$-automorphisms of $\mathbb{Q}(\sqrt{2})$.

1. The identity map $\iota(a + b\sqrt{2}) = a + b\sqrt{2}$, and
2. The conjugation map $\phi(a + b\sqrt{2}) = a - b\sqrt{2}$.

Example 9.16 (Galois Group of $\mathbb{Q}(2^{1/3})$ over $\mathbb{Q}$)

Given the field extension $\mathbb{Q} \subset \mathbb{Q}(\sqrt[3]{2})$, let us write $\xi = \sqrt[3]{2}$. Then since $\mathbb{Q}(\xi)$ is a vector space with basis $1, \xi, \xi^2$, we can write any element as $a + b\xi + c\xi^2$, and so by definition an element $\phi \in G(\mathbb{Q}(\xi)/\mathbb{Q})$ must satisfy

$$\phi(a + b\xi + c\xi^2) = a + b\phi(\xi) + c\phi(\xi)^2 \tag{371}$$

So the action is completely determined by the value of $\phi(\xi)$. Suppose $\phi(\xi) = \alpha + \beta\xi + \gamma\xi^2$ for some

$\alpha, \beta, \gamma \in \mathbb{Q}$. Through some derivation we have

$$\begin{aligned} 2 = \phi(2) = \phi(\xi^3) = (\phi(\xi))^3 &= (\alpha + \beta\xi + \gamma\xi^2)^3 && (372) \\ &= (\alpha^3 + 2\beta^3 + 4\gamma^3) + 3(\alpha^2\beta + \beta^2\gamma + \gamma^2\alpha)\xi + 3(\alpha\beta^2 + \beta\gamma^2 + \gamma\alpha^2)\xi^2 && (373) \end{aligned}$$

From the linear independence of $1, \xi, \xi^2$ we can see that

$$\alpha^3 + 2\beta^3 + 4\gamma^3 = 2 \quad (374)$$
$$\alpha^2\beta + \beta^2\gamma + \gamma^2\alpha = 0 \quad (375)$$
$$\alpha\beta^2 + \beta\gamma^2 + \gamma\alpha^2 = 0 \quad (376)$$

which turns out to have the only solution $\alpha = \gamma = 0, \beta = 1$. Therefore, the only $\mathbb{Q}$-automorphism of $\mathbb{Q}(\sqrt[3]{2})$ is the idnetity map.

Now let's look at how an F-automorphism acts on the roots of a polynomial. Let $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 \in F[x]$, $F \subset K$ a field extension, and suppose $f(x)$ has roots $\alpha_1, \ldots, \alpha_m$ lying in K (and perhaps other roots lying in a further extension). It turns out that any F-automorphism of K must permute $\alpha_1, \ldots, \alpha_m$, so the roots stay "within" the polynomial.

Lemma 9.4 (F-Automorphims Permute Roots)

Let $F \subset K$ and let $\alpha \in K$ be a root $f(x) \in F[x]$.
1. For any $\phi \in G(K/F)$, $\phi(\alpha)$ is also a root of $f(x)$ in K.[a]
2. If K is the splitting field of $f(x)$ and $f(x)$ has distinct roots $\alpha_1, \ldots, \alpha_n \in K$, then $G(F/K)$ is a subgroup of $\text{Perm}\{\alpha_1, \ldots, \alpha_n\}$.

[a]However it may not be a permutation!

Proof.

We know that $f(\alpha) = 0$. We claim first that $f(\phi(\alpha)) = 0$ and so $\phi(\alpha)$ is also a root. Now assume that $S = \{\alpha_1, \ldots, \alpha_n\}$ is contained in K. We know that $\phi(\alpha_j) \in S$. The map $\phi \mapsto \phi(\alpha_j)$ is indeed a group homomorphism $G(K/F) \to \text{Perm}(S)$. If K is the splitting field, then $K = F[S]$, and so if $\phi \in G(K/F)$ fixes all the α_j's, then it must fix all of K.

Example 9.17 ()

Therefore, we can get a much simply solution of the Galois group of $\mathbb{Q}(\sqrt[3]{2})$ over $\mathbb{Q}$. Since $\xi = \sqrt[3]{2}$ is a root of the irreducible polynomial $x^3 - 2$, any $\phi \in G(\mathbb{Q}[\xi]/\mathbb{Q})$ must carry ξ to some other root of $x^3 - 2$. The other roots are in the complex plane, and so ϕ must carry $\xi \mapsto \xi$. Thus, ϕ must be the identity.

Example 9.18 ()

Let $f(x) = x^2 - 2x - 1 \in \mathbb{Q}[x]$. One root of $f(x)$ is $1 + \sqrt{2} \in \mathbb{Q}(\sqrt{2})$. But we know that there exists a $\phi \in G(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$ that conjugates, and so $1 - \sqrt{2}$ must also be a root.

Example 9.19 ()

It follows that $G(\mathbb{C}/\mathbb{R})$ is a group of order 2 generated by complex conjugation.

By studying how F-automorphisms behave, we are able to understand a bit more about the Galois group.

Now we attempt to know more about the order of a Galois group by finding out how many possible F-automorphisms over K can exist. If we let $F \subset K$ be a field extension of degree $n > 1$ (i.e. $\dim_F K = n > 1$) and $\alpha \in K$, the set of $n+1$ vectors $\{1, \alpha, \alpha^2, \ldots, \alpha^n\}$ must be linearly dependent over a vector space of dimension n. Therefore, there exists some nontrivial linear combination

$$0 = a_n \alpha^n + a_{n-1}\alpha^{n-1} + \ldots + a_1 \alpha + a_0 \tag{377}$$

which corresponds to some nonzero $f(x) = a_n x^n + \ldots a_0$. Therefore α is a root of a polynomial $f(x) \in F[x]$ of degree at most n—therefore a root of an irreducible polynomial of degree at most n (since there could be a smaller linearly dependent set).

Now if we assume that $K = F[\alpha]$, let's try to see how many elements $G(K/F)$ could have.

1. $f(x)$ must have degree exactly n (why?). Since $1, \alpha, \ldots, \alpha^{n-1}$ form a basis of K over F, if $\phi \in G(K/F)$, the action of ϕ on $K = F[\alpha]$ is completely determined by $\phi(\alpha)$, which is what we saw for the first two examples above.
2. On the other hand, ϕ must map α to some root of $f(x)$ in K, and there are at most $n = [K : F]$ of these.

Thus we have $|G(F[\alpha]/F)| \leq [F[\alpha] : F]$, and it follows that if $F[\alpha] \subset K$, then $|G(F[\alpha]/F)| \leq [K : F]$. But our goal is to do so for arbitrary field extensions, i.e.

$$|G(K/F)| \leq [K : F] \tag{378}$$

To do this, we will need to talk about extensions of field isomorphisms. Recall that a field (i.e. a ring) isomorphism $\phi : F \to F'$ induces a ring isomorphism $\tilde{\phi} : F[x] \to F'[x]$.

Theorem 9.12 (Extending Isomorphisms of Fields)

Let $\phi : F \to F'$ be an isomorphism of fields.
1. Let $f(x) \in F[x]$ be an irreducible polynomial with root α in some field extension $K \supset F$.
2. Let $g(x) = \phi(f(x)) \in F'[x]$ and let β be a root of $g(x)$ in some field extension $K' \supset F'$.

Then there is a unique isomorphism

$$\bar{\phi} : F[\alpha] \to F'[\alpha'] \tag{379}$$

that is an extension of ϕ (i.e. behaves the same under F) and carries α to α.

We were interested in the Galois group of the symmetries of the field. Now we will extend this to bigger fields.

Theorem 9.13 (Bound on Number of Field Isomorphism Extensions)

Let $F \subset K_1, F \subset K_2$ be field extensions. Then there are at most $[K_1 : F]$ F-embeddings $\phi : K_1 \to K_2$. Moreoever, if there are exactly $[K_1 : F]$ F-embeddings, then for all $F \subset E \subset K$, there are $[E : F]$ F-embeddings $\phi : E \to K_2$.

Proof.

By induction on $[K_1 : F]$. When $[K_1 : F] = 1$. Now suppose the proposition holds for all F, K_1, K_2 with $[K_1 : F] \leq n-1$. Now choose F, K_1, K_2, as in the statement with $[K_1 : F] = n > 1$. Since $[K_1 : F] > 1$, we can choose $\alpha \in K_1$, $\alpha \notin F$. Let $f(x) \in F[x]$ be the minimal polynomial of α. We've seen that given the smallest subfield of K_1 containing α, i.e. $F[\alpha]$, we can use the evaluation homomorphism to state

$$F[\alpha] \simeq \frac{F[x]}{\langle f(x) \rangle}, \qquad F[\alpha] \xleftarrow{\mathrm{ev}_\alpha} \frac{F[x]}{\langle f(x) \rangle} \tag{380}$$

Any F-embedding of K_1 restricst to an F-embedding of $F[\alpha]$. By induction, there are at most

$[F[\alpha] : F] = m$ F-embeddings

$$\phi_1, \ldots, \phi_m : F[\alpha] \to K_2 \tag{381}$$

But by induction, we know that there are at most $[K_1 : F[\alpha]]$ $F[\alpha]$-embeddings $K_1 \to K_2$ where $F[\alpha] \subset K_2$ (which we can defined through multiple ways for each injection ϕ_i). This implies that there are at most

$$m\,[K_1 : F[\alpha]] = [F[\alpha] : F]\,[K_1 : F[\alpha] \tag{382}$$
$$= [K_1 : F] \tag{383}$$

F-embeddings. Suppose equality holds. Let $F \subset E \subset K_1$. We know

1. There are at most $[E : F]$ F-embeddings $\phi_1, \ldots, \phi_{[E:F]} : E \to K_2$
2. There are at most $[K_1 : E]$ E-embeddings $K_1 \to K_2$ where here $E \subset K_2$.

This gives at most $[K_2 : E][E : F]$ F-embeddings since $[K_2 : E][E : F] = [K_2 : F]$, and there are exactly $[K_2 : F]$ F-embeddings $K_1 \to K_2$, all the $\leq$'s must be $=$'s (otherwise it would be a strict inequality) in that there must be exactly $[E : F]$ F-embeddings $E \to K_2$ and exactly $[K_1 : E]$ E-embeddings $K_1 \to K_2$.
Correction. We need to see that there are at most $[F[\alpha] : F]$ F-embeddings $F[\alpha] \to K_2$. It's okay to use induction if $[F[\alpha] : F] \subset [K_1 : F]$. In general, we showed F-embeddings maps

If we reach this bound, then there are some special properties.

Definition 9.6 (Galois)

A field extension $F \subset K$ is **Glaois** if the number of F-embeddings $K \to K$ is $[K : F]$.

Example 9.20 ()

We review the derived Galois groups above and see if the field extensions are Galois.

1. $\mathbb{Q}(\sqrt{2})$ is a Galois extension of $\mathbb{Q}$.
2. $\mathbb{Q}(\sqrt[3]{2})$ is not a Galois extension of $\mathbb{Q}$ since $|G(\mathbb{Q}[\sqrt[3]{2}]/\mathbb{Q})| = 1$ but $[\mathbb{Q}[\sqrt[3]{2}] : \mathbb{Q}] = 3$.
3. Let $K = \mathbb{Q}[\sqrt[3]{2}, i\sqrt{3}]$ is the splitting field of $f(x) = x^3 - 2$. Then $[K : \mathbb{Q}] = 6$. $G(K/\mathbb{Q}) \simeq S_3$ has order 6 since we showed 6 $\mathbb{Q}$-automorphisms of K, but now we know that there can be no more.

Example 9.21 ()

Let $\alpha = \sqrt[7]{2}$. $\mathbb{Q} \subset \mathbb{Q}[\alpha]$ is not a Galois extension. $x^7 - 2$ is a polynomial with root α, and by Eisenstein $x^7 - 2$ is irreducible. So $f(x) = x^7 - 2$ is the minimal polynomial of α. This means that the number of $\mathbb{Q}$-embeddings $\mathbb{Q}[\alpha] \to \mathbb{Q}[\alpha]$ is in bijection with the number of roots of $f(x)$ in $\mathbb{Q}[\alpha]$. All 7 roots of $f(x)$ are $\sqrt[7]{2}e^{2\pi i j/7}$ for $j = 0, \ldots, 6$, which has one real root. So there is 1 $\mathbb{Q}$-embedding $\mathbb{Q}[\alpha] \to \mathbb{Q}[\alpha]$. But $[\mathbb{Q}[\sqrt[7]{2}] : \mathbb{Q}] = \deg f(x) = 7$. Since $\mathbb{Q}[\alpha] \simeq \mathbb{Q}[x]/\langle x^7 - 2\rangle$, which are polynomials of degree at most 6.

Theorem 9.14 (Splitting Fields of a Square Free Polynomial are Galois)

Suppose $F \subset K$ is the splitting field of a polynomial $f(x) \in F[x]$ such that no irreducible factor of $f(x)$ has repeated roots. Then $F \subset K$ is Galois.

Proof.

Since K is the splitting field of $f(x)$ over F, we have $F[r_1, \ldots, r_m]$ where the r_i are the distinct roots of $f(x)$. We show by induction that there are $[F[r_1, \ldots, r_j] : F]$ F-embeddings of $F[r_1, \ldots, r_j] \to K$. For $j = 1$, r_1 is a root of some irreducible factor $f_1(x)$ of $f(x)$.

$$F[r_1] \simeq \frac{F[x]}{\langle f_1(x) \rangle} \tag{384}$$

and the set of F-embeddings $F[r_1] \to K$ is in bijection with the set of roots $\alpha \in K$ of $f_1(x)$. By hypothesis , $f(x)$ has no repeated roots, which implies that the number of F-embeddings $F[r_1] \to K$ is $\deg f_1(x) = [F[r_1] : F]$ which gives the base case. For the inductive step, we know there are

$$[F[r_1, \ldots, r_{j-1}] : F] \tag{385}$$

F-embeddings $F[r_1, \ldots, r_{j-1}] \xrightarrow{\phi} K$. For each, we will show that there are exactly $[F[r_1, \ldots, r_j] : F[r_1, \ldots, r_{j-1}]]$ extensions of ϕ which completes the proof. Because

$$F[r_1, \ldots, r_j] : F] = F[r_1, \ldots, r_j] : F[r_1, \ldots, r_{j-1}]]F[r_1, \ldots, r_{j-1}] : F] \tag{386}$$

Let $g(x)$ be the minimal polynomial of r_j over $F[r_1, \ldots, r_{j-1}]$. Since $g(r_j) = 0$, g divides one of the irreducible factors of $f(x)$ in $F[r_1, \ldots, r_{j-1}][x]$ which implies it has no repeated roots. Then

$$F[r_1, \ldots, r_j] = \frac{E[x]}{\langle g(x) \rangle} \tag{387}$$

The number of E-embeddings is equal to the number of roots in in g which is $\deg g = F[r_1, \ldots, r_j : F]$.

Now we restrict our scope to $\mathbb{Q}$. Note the following.

Lemma 9.5 ()

Any irreducible polynomial in $\mathbb{Q}$ has no repeated roots.

With this, the following is immediately.

Theorem 9.15 ()

Let $\mathbb{Q} \subset K$ be the splitting field of $f(x) \in Q[x]$. Then $\mathbb{Q} \subset K$ is Galois.

What about the converse? There are two steps to proving that for any Galois field extension $F \subset K$, there exists a polynomial $f(x) \in F[x]$ that splits in K.

Lemma 9.6 (Irreducible Polynomial with Root in Galois Extension Splits)

Let $F \subset K$ be a Galois extension of fields. Let $f(x) \in F[x]$ be an irreducible polynomial with a root $\alpha \in K$. Then $f(x)$ splits in K, i.e. all other roots must be in K.

Theorem 9.16 (Every Galois Extension is a Splitting Field)

Let $F \subset K$ be a Galois extension. Then K is the splitting field of a polynomial $f(x) \in F[x]$.

Therefore, we have made a bijection of sets between set of F-embeddings and.

Theorem 9.17 (Fundamental Theorem of Galois Theory)

9.6 Cubic Equations

The well known discriminant of a quadratic equation

$$f(x) = ax^2 + bx + c \tag{388}$$

is known in the form $\nabla = b^2 - 4ac$. However, we will present it in a slightly different manner.

Definition 9.7 ()

The **discriminant** $D(\varphi)$ of a quadratic polynomial

$$\varphi = a_0 x^2 + a_1 x + a_2 \in \mathbb{C}[x] \tag{389}$$

with $c_1, c_2 \in \mathbb{C}$ as its roots is defined

$$D(\varphi) = a_1^2 - 4a_0a_2 = a_0^2\bigg(\Big(\frac{a_1}{a_0}\Big)^2 - \frac{4a_2}{a_0}\bigg) = a_0^2\big((c_1+c_2)^2 - 4c_1c_2\big) = a_0^2(c_1-c_2)^2 \tag{390}$$

Clearly, the value of $D(\varphi)$ can tell us three things

1. $c_1, c_2 \in \mathbb{R}, c_1 \neq c_2$. Then $c_1 - c_2$ is a nonzero real number and $D(\varphi) > 0$.
2. $c_1 = c_2 \in \mathbb{R}$. Then $c_1 - c_2 = 0$ and $D(\varphi) = 0$.
3. $c_1, c_2 \in \mathbb{C}, c_1 = \bar{c}_2$. Then, $c_1 - c_2$ is a nonzero strictly imaginary number and $D(\varphi) < 0$.

Definition 9.8 ()

We can generalize this notion of the discriminant to arbitrary polynomials

$$\varphi = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n \in \mathbb{F}[x],\ a_0 \neq 0 \tag{391}$$

The discriminant $D(\varphi)$ of the polynomial above is defined

$$D(\varphi) \equiv a_0^{2n-2}\prod_{i>j}(c_i - c_j)^2 \tag{392}$$

The a_0 term isn't very important in this formula, since it does not affect whether $D(\varphi)$ is positive, negative, or zero.

Definition 9.9 ()

A polynomial

$$\varphi = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n \in \mathbb{F}[x],\ a_0 \neq 0 \tag{393}$$

where $a_1 = 0$ is called **depressed**. A depressed cubic polynomial is of form

$$\varphi = x^3 + px + q \tag{394}$$

Proposition 9.1 ()

Every monic (leading coefficeint = 1) polynomial (and non-monic ones)

$$\varphi = x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n \in \mathbb{F}[x],\ a_0 \neq 0 \tag{395}$$

can be turned into a depressed polynomial with the change of variable

$$x = y - \frac{a_1}{n} \tag{396}$$

to get the polynomial

$$\psi = y^n + b_2 y^{n-2} + ... + b_{n-1} y + b_n \tag{397}$$

Lemma 9.7 ()

A cubic polynomial

$$\varphi = a_0 x^3 + a_1 x^2 + a_2 x + a_3 \in \mathbb{R}[x] \tag{398}$$

with roots $c_1, c_2, c_3 \in \mathbb{C}$ has discriminant

$$D(\varphi) \equiv a_0^4 (c_1 - c_2)^2 (c_1 - c_3)^2 (c_2 - c_3)^2 \tag{399}$$

With a bit of evaluation, it can also be expressed in terms of its coefficients as

$$D(\varphi) = a_1^2 a_2^2 - 4a_1^3 a_3 - 4a_0 a_2^3 + 18 a_0 a_1 a_2 a_3 - 27 a_0^2 a_3^2 \tag{400}$$

Again, three possibilities can occur (up to reordering of its roots).

1. c_1, c_2, c_3 are distinct real numbers. Then $D(\varphi) > 0$.
2. $c_1, c_2, c_3 \in \mathbb{R}, c_1 = c_2$. Then $D(\varphi) = 0$.
3. $c_1 \in \mathbb{R}, c_2 = \bar{c}_3 \notin \mathbb{R}$. Then $D(\varphi) < 0$.

Furthermore, the cubic formula used to find the roots of the polynomial is

$$c_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}} \tag{401}$$

known as **Cardano's formula**, after the mathematician Gerolamo Cardano.

10 Affine and Projective Spaces

10.1 Affine Spaces

Modeling the space of points as a vector space can be unsatisfactory for a number of reasons.

1. The origin 0 plays a special role, when it doesn't necessarily need to have one.
2. Certain notions, such as parallelism, are handled in an awkward manner.
3. The geometries of vector and affine spaces are intrinsically. That is,

$$\mathrm{GL}(V) \subset \mathrm{GA}(V) \tag{402}$$

In the ordinary Euclidean geometry, one can define the operation of the addition of a point and a vector. That is, the "sum" of a point p and a vector x is the endpoint of a vector that starts at p and equals x. We formalize it in the following definition.

Definition 10.1 ()

Let V be a vector space over field $\mathbb{F}$. The **affine space associated to** V is a set S with an operation of addition $+ : S \times V \longrightarrow S$ satisfying

1. $p + (x + y) = (p + x) + y$ for $p \in S, x, y \in V$
2. $p + 0 = p$ where $p \in S$, 0 is the zero vector
3. For any $p, q \in S$, there exists a unique vector x such that $p + x = q$

Elements of the set S are called **points**. The vector in condition 3 is called the **vector connecting points** p **and** q, denoted $\overline{pq}$. The dimension of an affine space is defined as the dimension of the corresponding vector space.

The first condition implies that

$$\overline{pq} + \overline{qr} = \overline{pr} \text{ for all } p, q, r \in S \tag{403}$$

Every vector space V can be regarded as an affine one if we view vectors both as points and as points and define the operation of addition of a vector to a point as addition of vectors. Under this interpretation, the vector $\overline{pq}$ is the difference between the vectors p and q.

Definition 10.2 ()

Conversely, if we fix a point o (the origin) in an affine space S, we can identify a point p with its **position vector** $\overline{op}$. Then, addition of a vector to a point just becomes the addition a vectors. This identification of points with vectors is called the **vectorization** of an affine space.

Definition 10.3 ()

A point o (the origin) together with a basis $\{e_1, ..., e_n\}$ of the space V is called a **frame** of the affine space S. Each frame is related to an **affine system of coordinates** in the space S. That is, a point p would get the coordinates equal to those of the vector $\overline{op}$ in the basis $\{e_1, ..., e_n\}$. It is easy to see that

1. Coordinates of the point $p + x$ are equal to the sums of respective coordinates of the point p and the vector x.
2. Coordinates of the vector $\overline{pq}$ are equal to the differences of respective coordinates of the points q and p.

Linear combinations of points are not defined in the affine space since the values of linear combinations are actually dependent on the choice of the origin. However, an analogous structure can be.

Definition 10.4 ()

The **barycentric linear combination** of points $p_1, ..., p_k \in S$ is a linear combination of the form

$$p = \sum_i \lambda_i p_i, \text{ where } \sum_i \lambda_i = 1 \tag{404}$$

This linear combination is equal to the point p such that

$$\overline{op} = \sum_i \lambda_i \overline{op_i} \tag{405}$$

where $o \in S$ is any origin point.

Definition 10.5 ()

In particular, the specific barycentric combination of points where $\lambda_1 = ... = \lambda_k = \frac{1}{k}$ is called the **center of mass** of the collection of points p_i.

Definition 10.6 ()

Let $p_0, p_1, ..., p_n$ be points of an n-dimensional affine space S such that the vectors $\overline{p_0p_1}, ..., \overline{p_0p_n}$ are linearly independent (that is, forms a basis). Then, every point $p \in S$ can be uniquely presented as

$$p = \sum_{i=0}^n x_i p_i, \text{ where } \sum_{i=0}^n x_i = 1 \tag{406}$$

This equality can be rewritten

$$\overline{p_0p} = \sum_{i=1}^n x_i \overline{p_0p_i} \tag{407}$$

implying that we can take the coordinates of the vector $\overline{p_0p}$ in the basis $\{\overline{p_0p_1}, ..., \overline{p_0p_n}\}$ as $x_1, ..., x_n$. Then, x_0 is determined as

$$x_0 = 1 - \sum_{i=1}^n x_i \tag{408}$$

The numbers $x_0, x_1, ..., x_n$ are called the **barycentric coordinates** of the point p with respect to $p_0, p_1, ..., p_n$.

Definition 10.7 ()

A **plane** in an affine space S is a subset of the form

$$p = p_0 + U \tag{409}$$

where p_0 is a point and U is a subspace of the space V. Note that we can choose any point p_0 in the plane in this representation. U is called the **direction subspace** for P.

Lemma 10.1 ()

If the intersection of two planes in an affine space is nonempty, then the intersection is also a plane.

Theorem 10.1 ()

Given any $k+1$ points of an affine space, there is a plane of dimension $\leq k$ passing through these points. If these points are not contained in a plane of dimension $< k$, then there exists a unique k-dimensional plane passing through them.

Definition 10.8 ()

Points $p_0, p_1, ..., p_k \in S$ are **affinely dependent** if they lie in a plane of dimension $< k$, and **affinely independent** otherwise. It is clear that the points $p_0, ..., p_k$ are affinely independent if and only if the vectors $\overline{p_0 p_1}, ..., \overline{p_0 p_k}$ are linearly independent.

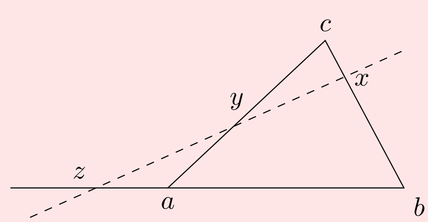
Theorem 10.2 ()

Points $p_0, ..., p_k \in S$ are affinely independent if and only if the rank of the matrix of their barycentric coordinates (with respect to some predetermined affinely independent points) equals $k+1$.

It is easy to see that the previous theorem is true, since the determinant represents the hypervolume of the parallelopiped spanned by the vectors $\overline{p_0 p_1}, ..., \overline{p_0 p_k}$, which must be nonzero if they are indeed affinely independent.

Corollary 10.1 (Menelaus' Theorem)

Let points x, y, z line on the sides bc, ca, ab of the triangle abc or their continuations.



Suppose that they divide these sides in the ratio

$$\lambda : 1, \mu : 1, \nu : 1$$

respectively. Then, the points x, y, z lie on the same line if and only if

$$\lambda \mu \nu = -1$$

Proof.

By the previous theorem, the points x, y, z are linearly dependent (i.e. lies on a line) if and only if the matrix of barycentric coordinates of x, y, z with respect to a, b, c, which is

$$\begin{pmatrix} 0 & \frac{1}{\lambda+1} & \frac{\lambda}{\lambda+1} \\ \frac{\mu}{\mu+1} & 0 & \frac{1}{\mu+1} \\ \frac{1}{\nu+1} & \frac{\nu}{\nu+1} & 0 \end{pmatrix} \tag{410}$$

has nonzero determinant. The determinant of the above matrix is 0 if and only if $\lambda \mu \nu = -1$.

Corollary 10.2 (Ceva's Theorem)

In the triangle above, the lines ax, by, cz intersect at one point if and only if

$$\lambda\mu\nu = 1 \tag{411}$$

Proof.

The proof can be done using barycentric coordinates.

Theorem 10.3 ()

A nonempty subset $P \subset S$ is a plane if and only if for any two distinct points $a, b \in P$, the line through a and b also lies in P.

Theorem 10.4 ()

Given an inhomogeneous system of linear equations of form

$$Ax = b \tag{412}$$

the set of solutions is an affine plane of dimension $n - r$, where n is the number of variables and r is the rank of the matrix A. More precisely, given that the plane is in the form $P = p_0 + U$, p_0 is one solution and U is the set of vectors that satisfy the homogeneous system

$$Ax = 0 \tag{413}$$

Let us observe the relative position of two planes.

Theorem 10.5 ()

Given two planes

$$P_1 = p_1 + U_1, P_2 = p_2 + U_2$$

P_1 and P_2 intersect if and only if

$$\overline{p_1 p_2} \subset U_1 + U_2 \tag{414}$$

where $U_1 + U_2$ is the set of all vectors of form $u_1 + u_2$, where $u_1 \in U_1, u_2 \in U_2$.

Now, consider the class of functions on an affine space corresponding to the class of linear functions on a vector space.

Definition 10.9 ()

An **affine-linear** function on an affine space S is a function $f : S \longrightarrow \mathbb{F}$ such that

$$f(p + x) = f(p) + \alpha(x), \;\; p \in S, x \in V \tag{415}$$

where α, called the **differential**, is a linear function on the vector space V. Let $o \in S$ be a fixed origin. By setting $p = o$, we can express an affine linear function in vectorized form as

$$f(x) = \alpha(x) + b, \;\; b \in \mathbb{F} \tag{416}$$

where $b = f(o)$. This implies the following coordinate form of f.

$$f(x) = b + \sum_i a_i x_i \tag{417}$$

A particular case of affine-linear functions are constant functions, where the defining characteristic is the zero differential.

Proposition 10.1 ()

Given that $\dim S = n$, affine-linear functions on S form a $(n+1)$-dimensional subspace on the space of all linear functions on S.

Proposition 10.2 ()

Barycentric coordinates are affine-linear functions.

Proposition 10.3 ()

Let f be an affine-linear function. Then

$$f\bigg(\sum_i \lambda_i p_i\bigg) = \sum_i \lambda_i f(p_i) \tag{418}$$

for any barycentric linear combination $\sum_i \lambda_i p_i$ of points $p_1, ..., p_k$.

Definition 10.10 ()

An affine space associated with a Euclidean vector space is called a **Euclidean affine space**. The **distance** ρ between two points in a Euclidean space is defined as

$$\rho(p, q) = ||\overline{pq}|| \tag{419}$$

This definition of ρ satisfies the axioms of a metric space.

10.2 Convex Sets

Let S be an affine space over the field of real numbers and V, the associated vector space.

Definition 10.11 ()

The **(closed) interval** connecting points $p, q \in S$ is the set

$$pq = \{\lambda p + (1 - \lambda) q \mid 0 \leq \lambda \leq 1\} \tag{420}$$

Geometrically, we can think of this as the straight line segment connecting point p with point q.

Definition 10.12 ()

A set $M \subset S$ is **convex** if for any two points $p, q \in S$, it contains the whole interval p, q.

Clearly, the intersection of convex sets is convex. However, the union of them is not.

Definition 10.13 ()

A **convex linear combination** of points in S is their barycentric linear combination with nonnegative coefficients.

It is clear to visualize the following proposition.

Proposition 10.4 ()

For any points $p_0, ..., p_k$ in a convex set $M \subset S$, the set M also contains every convex linear combination

$$p = \sum_i \lambda_i p_i \tag{421}$$

Furthermore, for any set $M \subset S$, the set $\operatorname{conv} M$ of all convex linear combinations of points in M is convex.

Definition 10.14 ()

Given $M \subset S$, the set $\operatorname{conv} M$ is the smallest convex set containing M. It is called the **convex hull** of M.

Definition 10.15 ()

The convex hull of a system of affinely independent points $p_0, p_1, ..., p_n$ in an n-dimensional affine space is called the **n-dimensional simplex** with vertices $p_0, ..., p_n$.

It is clear that the interior points of a simplex is precisely the set of all points whose barycentric coordinates with respect to the vertices are all positive.

Example 10.1 ()

Here are common examples of simplices.

1. A 0-dimensional simplex is a point.
2. A 1-dimensional simplex is a closed line interval.
3. A 2-dimensional simplex is a triangle.
4. A 3-dimensional simplex is a tetrahedron.

Proposition 10.5 ()

A convex set M has interior points if and only if $\operatorname{aff} M = S$.

Definition 10.16 ()

A convex set that has interior points is called a **convex body**. Clearly, every convex body in n-dimensional affine space S is n-dimensional.

The set of interior points of a convex body M, denoted M°, is an open convex body.

Definition 10.17 ()

For any nonconstant affine-linear function f on the set S, let

$$H_f \equiv \{p \in S \mid f(p) = 0\}$$
$$H_f^+ \equiv \{p \in S \mid f(p) \geq 0\}$$
$$H_f^- \equiv \{p \in S \mid f(p) \leq 0\}$$

The set H_f is a hyperplane, and H_f^+, H_f^- are called **closed half spaces**.

Definition 10.18 ()

A hyperplane H_f is a **supporting hyperplane** of a closed convex body M if $M \subset H_f^+$ and H_f contains at least one (boundary) point of M. The half space H_f^+ is then called the **supporting half-space** of M.

Proposition 10.6 ()

A hyperplane H that passes through a boundary point of a closed convex body M, is supporting if and only if $H \cap M^\circ = \emptyset$.

A key theorem of convex sets is the following separation theorem.

Theorem 10.6 (Separation Theorem)

For every boundary point of a closed convex body, there exists a supporting hyperplane passing through this point.

This theorem leads to the following one.

Theorem 10.7 ()

Every closed convex set M is an intersection of (perhaps infinitely many) half-spaces.

Definition 10.19 ()

A **polyhedron** is the intersection of a finite number of half-spaces. A convex polyhedron which is also a body is called a **convex solid**.

Example 10.2 ()

A simplex with vertices $p_0, p_1, ..., p_n$ is a convex polyhedron since it is determined by linear inequalities $x_i \geq 0$ for $i = 0, 1, ..., n$, where $x_0, x_1, ..., x_n$ are barycentric coordiantes with respect to $p_0, p_1, ..., p_n$.

Example 10.3 ()

A convex polyhedron determined by linear inequalities $0 \leq x_i \leq 1$ for $i = 1, ..., n$, where $x_1, ..., x_n$ are affine coordinates with respect ot some frame, is called an n-dimensional parallelopiped.

Definition 10.20 ()

A point p of a convex set M is **extreme** if it is not an interior point of any interval in M.

Theorem 10.8 ()

A bounded closed convex set M is the convex hull of the set $E(M)$ of its extreme points.

We can create a stronger statement with the following theorem.

Theorem 10.9 (Minkowski-Weyl Theorem)

The following properties of a bounded set $M \subset S$ is equivalent.
1. M is a convex polyhedron.
2. M is a convex hull of a finite number of points.

Definition 10.21 ()

A **face** of a convex polyhedron M is a nonempty intersection of M with some of its supporting hyperplanes. Given that $\dim \operatorname{aff} M = n$,
1. A 0-dimensional face is called a **vertex**.
2. A 1-dimensional face an **edge**.
3. ...
4. An $(n-1)$-dimensional face a **hyperface**.

Therefore, if a convex polyhedron is determined by a system of linear inequalities, we can obtain its faces by replacing some of these inequalities with equalities (in such a way that we do not get the empty set).

The following theorem demonstrates that in order to find its faces, it suffices to consider only the hyperplanes $H_{f_1}, ..., H_{f_m}$.

Theorem 10.10 ()

Every face Γ of the polyhedron M is of the form

$$\Gamma = M \cap \bigg(\bigcap_{j \in J} H_{f_j} \bigg) \tag{422}$$

where $J = \{1, 2, ..., m\}$

Proposition 10.7 ()

The extreme points of a convex polyhedron M are exactly its vertices.

The following theorem is used often in linear programming and in optimization.

Theorem 10.11 ()

The maximum of an affine-linear function on a bounded convex polyhedron M is attained at a vertex.

10.3 Affine Transformations and Motions

Let S and S' be affine spaces associated with vector spaces V and V', respectively, over the same field $\mathbb{F}$.

Definition 10.22 ()

An **affine map** from the space S to the space S' is a map $f : S \longrightarrow S'$ such that

$$f(p + x) = f(p) + \varphi(x), \quad p \in S, x \in V \tag{423}$$

for some linear map $\varphi : V \longrightarrow V'$. It follows that

$$\varphi(\overline{pq}) = \overline{f(p)f(q)}, \quad p, q \in S \tag{424}$$

Thus, f determines the linear map φ uniquely. Similarly, φ is called the **differential** of f, denoted df.

Proposition 10.8 ()

Let $f : S \longrightarrow S'$ and $g : S' \longrightarrow S''$ be two affine maps. Then the map

$$g \circ f : S \longrightarrow S'' \tag{425}$$

is also affine. Also

$$d(g \circ f) = dg \cdot df \tag{426}$$

where dg and df are the differentials of g and f, respectively.

For $\mathbb{F} = \mathbb{R}$, the differential of an affine map is a particular case of a differential of a smooth map in analysis. That is, the differential is the linear approximation of the function f.

Proposition 10.9 ()

An affine map is bijective if and only if its differential is bijective.

Definition 10.23 ()

Similar to linear transformations between vector spaces, bijective affine transformations are called **isomorphisms** of affine spaces. Affine spaces are **isomorphic** if there exists an isomorphism between them.

Corollary 10.3 ()

Finite-dimensional affine spaces over the same field are isomorphic if and only if they have the same dimension.

Definition 10.24 ()

An affine map from an affine space S to itself is called an **affine transformation**. Bijective affine transformations form a group called the **affine group of** S, denoted $\mathrm{GA}(S)$.

It follows that given affine space S with associated vector space V, the projection map

$$d : \mathrm{GA}(S) \longrightarrow \mathrm{GL}(V) \tag{427}$$

is a group homomorphism. It's kernel is the group of parallel translations, called $\mathrm{Tran}(S)$.

$$t_a : p \mapsto p + a, \quad a \in V \tag{428}$$

Proposition 10.10 ()

For any $f \in \mathrm{GA}(S)$ and $a \in V$,
$$f t_a f^{-1} = t_{df(a)} \tag{429}$$

Definition 10.25 ()

A **homothety** with the center o and coefficient λ is an affine transformation defined as
$$f(o + x) \equiv o + \lambda x \tag{430}$$
In its vectorized form, it is expressed
$$f(x) = \lambda x + b, \;\; b \in V \tag{431}$$
A homothety with coefficient -1 is called a **central symmetry**.

The group of affine transformations determines the **affine geometry** of the space. The following theorem shows that all simplices are equal in affine geometry.

Theorem 10.12 ()

Let $\{p_0, ..., p_n\}$ and $\{q_0, ..., q_n\}$ be two systems of affinely independent points in an n-dimensional affine space S. Then there exists a unique affine transformation f that maps p_i to q_i for $i = 0, 1, ..., n$.

Proof.

It is easy to see once we realize that there exists a unique linear map φ of the space V that maps the basis $\{\overline{p_0 p_1}, ..., \overline{p_0 p_n}\}$ to the basis $\{\overline{q_0 q_1}, ..., \overline{q_0 q_n}\}$. If we vectorize S by taking p_0 as the origin, the affine transformation in question has the form
$$f(x) = \varphi(x) + \overline{p_0 q_0} \tag{432}$$

Corollary 10.4 ()

In real affine geometry all parallelopipeds are equal.

Definition 10.26 ()

A **motion** of the space S is an affine transformation of S whose differential is an orthogonal operator (i.e. an origin preserving isometry). Every motion is bijective.

Motions of a Euclidean space S form a group denoted $\mathrm{Isom}\, S$. A motion is called **proper (orientation preserving)** if its differential belongs to $\mathrm{SO}(V)$ and improper otherwise.

Lemma 10.2 ()

The group $\mathrm{Isom}\, S$ is generated by reflections through hyperplanes.

Definition 10.27 ()

Let M be a solid convex polyhedron in an n-dimensional Euclidean space. A **flag of** M is a collection of its faces $\{F_0, F_1, ..., F_{n-1}\}$ where $\dim F_k = k$ and $F_0 \subset F_1 \subset ... \subset F_{n-1}$.

Definition 10.28 ()

A convex polyhedron M is **regular** if for any two of its flags, there exists a motion $f \in \operatorname{Sym} M$ mapping the first to the second, where

$$\operatorname{Sym} M \equiv \{f \in \operatorname{Isom} S \mid f(M) = M\} \tag{433}$$

Two dimensional regular polyhedra are the ordinary **regular polygons**. Their symmetry groups are known as the dihedral groups.

Three dimensional regular polyhedra are **Platonic solids**, which are the regular tetrahedron, cube, octahedron, dodecahedron, and icosahedron.

Definition 10.29 ()

A real vector space V with a fixed symmetric bilinear function α of signature (k, l), where $k, l > 0$ and $\dim V = k + l$, is called the **pseudo-Euclidean vector space** of signature (k, l). The group of α-preserving linear transformations of V is called the **pseudo-orthogonal group** and is denoted $\mathrm{O}(V, \alpha)$. In an orthonormal basis, the corresponding matrix group is denoted Ok, l.

10.4 Quadrics

Planes are the simplest objects of affine and Euclidean geometry, which are determined by systems of linear equations. The second simplest are quadratic functions. These types of objects are studied futher in algebraic geometry.

Definition 10.30 ()

An **affine-quadratic function** on an affine space S is a function $Q : S \longrightarrow \mathbb{F}$ such that its vectorized form is

$$Q(x) = q(x) + l(x) + c \tag{434}$$

for a quadratic function q, linear function l, and constant c.

10.5 Projective Spaces

Definition 10.31 ()

An n-dimensional **projective space** PV over a field $\mathbb{F}$ is the set of one-dimensional subspaces of an $(n+1)$-dimensional vector space V over $\mathbb{F}$. For every $(k+1)$-dimensional subspace $U \subset V$, the subset $PU \subset PV$ is called a k-dimensional **plane** of the space PV.

1. 0-dimensional planes are the points of PV.
2. 1-dimensional planes are called **lines**
3. ...
4. $(n-1)$-dimensional planes are called **hyperplanes**

Definition 10.32 ()

$\mathbb{RP}^1$ is called the real projective line, which is topologically equivalent to a circle.

Example 10.4 ()

The real projective space of $\mathbb{R}^2$ is the set of all lines that pass through the origin. It is denoted $\mathbb{RP}^2$ and called the **real projective plane**.

Example 10.5 ()

$\mathbb{RP}^3$ is diffeomorphic to SO(3).

Example 10.6 ()

The space $\mathbb{RP}^n$ is formed by taking the quotient of $\mathbb{R}^{n+1} \setminus \{0\}$ under the equivalence relation

$$x \sim \lambda x \text{ for all real numbers } \lambda \neq 0 \tag{435}$$

The set of these equivalence classes is isomorphic to $\mathbb{RP}^n$.

11 Representations

We will assume that V is a finite-dimensional vector space over field $\mathbb{C}$.

Definition 11.1 ()

The **general linear group** of vector space V, denoted $\mathrm{GL}(V)$, is the group of all automorphisms of V to itself. The **special linear group** of vector space V, denoted $\mathrm{SL}(V)$ is the subgroup of automorphisms of V with determinant 1.

When studying an abstract set, it is often useful to consider the set of all maps from this abstract set to a well known set (e.g. $\mathrm{GL}(V)$).

Definition 11.2 ()

A **representation** of an (algebraic) group $\mathcal{G}$ is a homomorphism

$$\rho : G \longrightarrow \mathrm{GL}(V) \tag{436}$$

for some vector space V. That is, given an element $g \in \mathcal{G}$, $\rho(g) \in \mathrm{GL}(V)$, meaning that $\rho(g)(v) \in V$. Additionally, since it is a homomorphism, the algebraic structure is preserved.

$$\rho(g_1 \cdot g_2) = \rho(g_1) \cdot \rho(g_2) \tag{437}$$

where $\cdot$ on the left hand side is the abstract group multiplication while the $\cdot$ on the right hand side is matrix multiplication. To shorten the notation, we will denote

$$gv = \rho(g)v, \; v \in V \tag{438}$$

Since ρ is a group morphism, we have

$$g_2(g_1 v) = (g_2 g_1)v \iff \rho(g_2)\big(\rho(g_1)(v)\big) = \big(\rho(g_2)\rho(g_1)\big)(v) \tag{439}$$

Additionally, since g (that is, $\rho(g)$) is a linear map,

$$g(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 g v_1 + \lambda_2 g v_2 \tag{440}$$

Usually, we refer to the map as the representation, but if the map is well-understood, we just call the vector space V the representation and say that the group acts on this vector space.

Example 11.1 ()

The group $\mathrm{GL}(2, \mathbb{C})$ can be represented a by the vector space $\mathbb{C}^2$, or explicitly, by the group of 2×2 matrices over $\mathbb{C}$ with nonzero determinant.

$$\mathrm{GL}(2, \mathbb{C}) \xmapsto{id} \mathrm{Mat}(2, \mathbb{C}) \tag{441}$$

This is a trivial representation.

We now show a nontrivial representation of $\mathrm{GL}(2, \mathbb{C})$.

Example 11.2 ()

We take $\mathrm{Sym}^2\mathbb{C}^2$, the second symmetric power of $\mathbb{C}^2$. Note that given a basis $x_1, x_2 \in \mathbb{C}^2$, the set

$$\{x_1 \odot x_1, x_1 \odot x_2, x_2 \odot x_2\} \tag{442}$$

forms a basis of $\mathrm{Sym}^2\mathbb{C}^2 \implies \dim \mathrm{Sym}^2\mathbb{C}^2 = 3$. So, we want to represent $\mathrm{GL}(2,\mathbb{C})$ by associating its element with elements of $\mathrm{GL}(Sym^2\mathbb{C}^2)$. More concretely, we are choosing to represent a 2×2 matrix over $\mathbb{C}$ with a 3×3 matrix group (since $\mathrm{GL}(Sym^2\mathbb{C}^2) \simeq \mathrm{GL}(3,\mathbb{C})$. Clearly,

$$\begin{aligned}
\rho(g)(x_1 \odot x_1) &= g(x_1) \odot g(x_1) \in Sym^2\mathbb{C}^2 \\
\rho(g)(x_1 \odot x_2) &= g(x_1) \odot g(x_2) \\
\rho(g)(x_2 \odot x_2) &= g(x_2) \odot g(x_2)
\end{aligned}$$

To present this in matrix form, let us have an element in $\mathrm{GL}(2,\mathbb{C})$

$$\mathcal{A} \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{443}$$

We evaluate the corresponding representation in $\mathrm{GL}(Sym^2\mathbb{C}^2)$. Using the identities above, we have

$$\begin{aligned}
\rho(g)(x_1 \odot x_1) &= g(x_1) \odot g(x_1) \\
&= (ax_1 + cx_2) \odot (ax_1 + cx_2) \\
&= a^2 x_1 \odot x_1 + 2ac x_1 \odot x_2 + c^2 x_2 \odot x_2 \\
\rho(g)(x_1 \odot x_2) &= g(x_1) \odot g(x_2) \\
&= (ax_1 + cx_2) \odot (bx_1 + dx_2) \\
&= ab x_1 \odot x_1 + (ad+bc) x_1 \odot x_2 + cd x_2 \odot x_2 \\
\rho(g)(x_2 \odot x_2) &= g(x_2) \odot g(x_2) \\
&= (bx_1 + dx_2) \odot (bx_1 + dx_2) \\
&= b^2 x_1 \odot x_1 + 2bd x_1 \odot x_2 + d^2 x_2 \odot x_2
\end{aligned}$$

And this completely determines the matrix. So,

$$\rho \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix} \tag{444}$$

is the 3×3 representation of $\mathcal{A}$ in $\mathrm{GL}(Sym^2\mathbb{C}^2)$.

We continue to define maps between two representations of $\mathcal{G}$.

Definition 11.3 ()

A **morphism** between 2 representations

$$\begin{aligned}
\rho_1 &: \mathcal{G} \longrightarrow \mathrm{GL}(V_1) \\
\rho_2 &: \mathcal{G} \longrightarrow \mathrm{GL}(V_2)
\end{aligned}$$

of some group but not necessarily the same vector space is a linear map $f: V_1 \longrightarrow V_2$ that is **compatible** with the group action. That is, f satisfies the property that for all $g \in \mathcal{G}$

$$f \circ g = g \circ f \tag{445}$$

Again, we use the shorthand notation that $g = \rho(g)$, meaning that the statement above really translates to $f \circ \rho(g) = \rho(g) \circ f$. This is equivalent to saying that the following diagram commutes.

$$\begin{tikzcd} V_1 \arrow[r, "\rho_1(g)"] \arrow[d, "f"] & V_1 \arrow[d, "f"] \\ V_2 \arrow[r, "\rho_2(g)"] & V_2 \end{tikzcd}$$

Definition 11.4 ()

Let V be a representation of $\mathcal{G}$. A **subrepresentation** is a subspace $W \subset V$ such that for all $g \in \mathcal{G}$ and for all $w \in W$,

$$\rho(g)(w) \in W \tag{446}$$

Example 11.3 ()

V and $\{0\}$ are always subrepresentations of V.

We now introduce the "building blocks" of all representations.

Definition 11.5 ()

A representation W is **irreducible representation** if $\{0\}$ and W are the only subrepresentations of W.

Lemma 11.1 (Schur's Lemma)

Let V_1, V_2 be irreducible representations and let $f : V_1 \longrightarrow V_2$ be a morphism (of representations). Then, either

1. f is an isomorphism.
2. $f = 0$

Furthermore, any 2 isomorphisms differ by a constant. That is,

$$f_1 = \lambda f_2 \tag{447}$$

Proof.

$\ker f$ is clearly a vector space. Furthermore, it is a subrepresentation (since it is a subspace of V_1) $\implies \ker f = V$ or $\ker f = 0$. If $\ker f = V$, then $f = 0$ and the theorem is satisfied. If $\ker f = 0$, then f is injective, and $\operatorname{Im} f$ is a subrepresentation of $V_2 \implies \operatorname{Im} f = 0$ or $\operatorname{Im} f = V_2$. But $\operatorname{Im} f \neq 0$ since f is injective, so $\operatorname{Im} f = V_2 \implies f$ is surjective $\implies f$ is bijective, that is, f is an isomorphism of vector spaces. So, the inverse f^{-1} exists, and this map f^{-1} satisfies

$$f^{-1} \circ \rho_2(g) = \rho_1(g) \circ f^{-1} \tag{448}$$

To prove the second part, without loss of generality, assume that the first isomorphism is the identity mapping. That is,

$$f_1 = id \tag{449}$$

Since we are working over the field $\mathbb{C}$, we can find an eigenvector of f_2. That is, there exists a $v \in V_1$ such that

$$f_2(v) = \lambda v \tag{450}$$

Now, we define the map

$$f : V_1 \longrightarrow V_2,\ f \equiv f_2 - \lambda f_1 \tag{451}$$

Clearly, $\ker f \neq 0$, since $v \in \ker f$. That is, we have a map f between 2 irreducible representations that has a nontrivial kernel. This means that $f = 0 \implies f_2 = \lambda f_1$.

Theorem 11.1 (Mache's Theorem)

Let V be finite dimensional, with $\mathcal{G}$ a finite group. Then, V can be decomposed as

$$V = \bigoplus_i V_i \tag{452}$$

where each V_i is an irreducible representation of $\mathcal{G}$.

Proof.

By induction on dimension, it suffices to prove that if W is a subrepresentation of V, then there exists a subrepresentation $W' \subset V$ such that $W \oplus W' = V$. So, if V isn't an irreducible representation, it can always be decomposed into smaller subrepresentations W and W' that direct sum to V. Now, we define the canonical (linear) projection

$$\pi : V \longrightarrow W \tag{453}$$

Then, we define the new map

$$\tilde{\pi} : V \longrightarrow W,\ \tilde{\pi}(v) \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \rho(g)\big|_W \circ \pi \circ \rho(g)^{-1} \tag{454}$$

This "averaging" of the group elements are done so that this mapping is a map of representations. This implies that

$$V = W \oplus \ker \tilde{\pi} \tag{455}$$

meaning that V can indeed be decomposed into direct sums of subrepresentations.

12 Lie Groups and Lie Algebras

Definition 12.1 ()

A **Lie group** is a group $\mathcal{G}$ that is also a finite-dimensional smooth manifold, in which the group operations of multiplication and inversion are smooth maps. Smoothness of the group multiplication

$$\mu : \mathcal{G} \times \mathcal{G} \to \mathcal{G},\ \mu(x,y) = xy \tag{456}$$

means that μ is a smooth mapping of the product manifold $\mathcal{G} \times \mathcal{G}$ into $\mathcal{G}$. These two requirements can be combined to the single requirement that tahe mapping

$$(x,y) \mapsto x^{-1}y \tag{457}$$

be a smooth mapping of the product manifold into $\mathcal{G}$.

Definition 12.2 ()

A **Lie Algebra** is a vector space $\mathfrak{g}$ with an operation called the **Lie Bracket**

$$[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{458}$$

Satisfying
1. Bilinearity: $[ax+by,z] = a[x,z] + b[y,z],\ [z,ax+by] = a[z,x] + b[z,y]$
2. Anticommutativity: $[x,y] = -[y,x]$
3. Jacobi Identity: $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$

Clearly, this implies that $\mathfrak{g}$ is a nonassociative algebra. Note that a Lie Algebra does not necessarily need to be an algebra in the sense that there needs to be multiplication operation that is closed in $\mathfrak{g}$.

Example 12.1 ()

A common example of a Lie Braket in the algebra of matrices is defined

$$[A,B] \equiv AB - BA \tag{459}$$

called the **commutator**. Note that in this case, the definition of the Lie bracket is dependent on the definition of the matrix multiplication. Without defining the multiplication operation, we wouldn't know what AB or BA means. Therefore, we see that the Lie algebra of $n \times n$ matrices has three operations: matrix addition, matrix multiplication, and the commutator (along with scalar multiplication). But in general, it is not necessary to have that multiplication operation for abstract Lie algebras. $\mathfrak{g}$ just needs to be a vector space with the bracket.

Example 12.2 ()

The set of all symmetric matrices is a vector space, but it is **not** a Lie algebra since the commutator $[A,B]$ is not symmetric unless $AB = BA$.

We will first talk about groups of matrices as a more concerete example before we get into abstract Lie groups. Recall that the matrix exponential map is defined

$$exp : \mathrm{Mat}(n,\mathbb{C}) \to \mathrm{mat}(n,\mathbb{C}),\ exp(A) = e^A = \sum_{p\geq 0} \frac{A^p}{p!} \tag{460}$$

Note that this value is always well defined. This lets us define

$$exp(tA) \equiv e^{tA} \equiv I + tA + \frac{1}{2}t^2A^2 + \frac{1}{3!}t^3A^3 + ... \tag{461}$$

where if t is small, we can expect a convergence. Note that exp maps addition to multiplication. That is, we can interpret it as a homomorphism from

$$exp : \mathfrak{g} \to \mathcal{G} \tag{462}$$

where $\mathfrak{g}$ is the Lie algebra and $\mathcal{G}$ is the Lie group (which we will treat just as a matrix group). To find the inverse of the exponential map, we can take the derivative of e^{tA} at $t = 0$. That is,

$$\bigg(\frac{d}{dt}e^{tA}\bigg)\bigg|_{t=0} = \bigg(\sum_{k=0}^{\infty} \frac{1}{k!}t^k A^{k+1}\bigg)\bigg|_{t=0} = A$$

So, the mapping

$$\frac{d}{dt}\bigg|_{t=0} : \mathcal{G} \to \mathfrak{g} \tag{463}$$

maps the Lie group back to the algebra. We can interpret this above mapping by visualizing the Lie Algebra as a tangent (vector) space of the abstract Lie group $\mathcal{G}$ at the identity element of the Lie group. The visualization below isn't the most abstract one, but it may help:

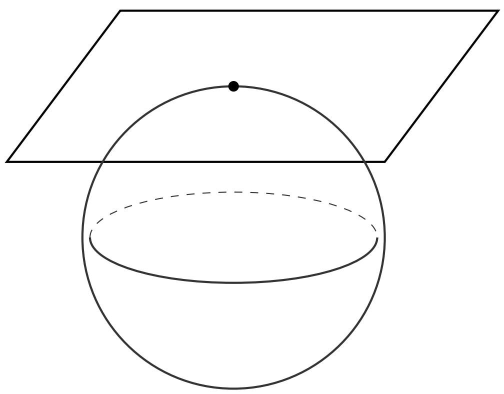
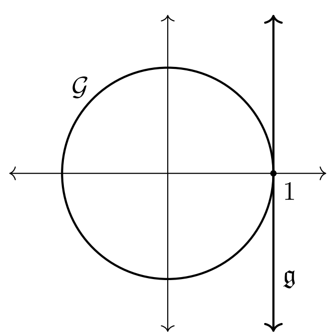


Figure 11: The Lie algebra can be visualized as the tangent space at the identity.

For example, say that the Lie group $\mathcal{G}$ is a unit circle in $\mathbb{C}$, then the Lie algebra of $\mathcal{G}$ is the tangent space at the identity 1, which can be identified as the imaginary line in the complex plane $\{it \mid t \in \mathbb{R}\}$, with

$$it \mapsto exp(it) \equiv e^{it} \equiv \cos t + i \sin t \tag{464}$$



So, analyzing the Lie group by looking at its Lie algebra turns a nonlinear problem to a linear one; this is called a **linearization** of the Lie group. The existence of this exponential map is one of the primary reasons that Lie algebras are useful for studying Lie groups.

Example 12.3 ()

The exponential map
$$exp : \mathbb{R} \rightarrow \mathbb{R}^+, \; x \mapsto e^x \tag{465}$$
is a group homomorphism that maps $(\mathbb{R}, +)$ to $(\mathbb{R}^+, \times)$. This means that $\mathbb{R}$ is the Lie algebra of the Lie group $\mathbb{R}^+$.

Theorem 12.1 ()

If A and B are commuting square matrices, then
$$e^{A+B} = e^A \, e^B \tag{466}$$
In general, the solution C to the equation
$$e^A \, e^B = e^C \tag{467}$$
is given by the **Baker-Campbell-Hausdorff formula**, defined
$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + ... \tag{468}$$
consisting of terms involving higher commutators of A and B. The full series is much too complicated to write, so we ask the reader to be satisfied with what is shown.

The BCH formula is messy, but it allows us to compute products in the Lie Group as long as we known the commutators in the Lie Algebra.

Therefore, we can describe the process of constructing a Lie group from a Lie Algebra (which a vector space) as such. We take a vector space V and endow it the additional bracket operation. We denote this as
$$\mathfrak{g} \equiv (V, [\cdot, \cdot]) \tag{469}$$
Then, we take every element of $\mathfrak{g}$ and apply the exponential map to them to get an another set $\mathcal{G}$. We then endow a group structure on $\mathcal{G}$ by defining the multiplication as
$$\cdot : \mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}, \; e^A \cdot e^B = e^{A*B} \tag{470}$$
where $A * B$ is defined by the BCH formula up to a certain kth order. Since the $*$ operation is completely defined by the bracket in the Lie algebra, it tells us how to multiply in the Lie group. This process can be made more abstractly, depending on what A, B and $[\cdot, \cdot]$ is, beyond matrices.

12.1 Lie Algebras of Classical Lie Groups

Definition 12.3 (General Linear Group)

The **general linear group**, denoted $\text{GL}(V)$, is the set of all bijective linear mappings from V to itself. Similarly, $\text{GL}_n(\mathbb{F})$, or $\text{GL}\,(n, \mathbb{F})$ is the set of all nonsingular $n \times n$ matrices over the field $\mathbb{F}$. Due to the same dimensionality of the following spaces, it is clear that $\text{GL}(V) \simeq \text{GL}(\mathbb{F}^n) \simeq \text{GL}_n(\mathbb{F})$. The **special linear group**, denoted $\text{SL}_n(\mathbb{F})$ or $\text{SL}(n, \mathbb{F})$, is the set of $n \times n$ matrices a with determinant 1. $\text{SL}_n(\mathbb{F})$ is a subgroup of $\text{GL}_n(\mathbb{F})$, which is a subset of the ring of all $n \times n$ matrices over field $\mathbb{F}$, denoted $\mathbb{L}_n(\mathbb{F})$.

Definition 12.4 (Translation Group)

The group of all translations in the space V is denoted $\text{Tran}\, V$. Its elements are usually denoted as t_u, where u is the vector that is being translated by. It can also be interpreted as shifting the origin by $-u$. It is clear that $\text{Tran}\, V \simeq V$.

Definition 12.5 (General Affine Group)

The **general affine group** is the pair of all transformations

$$\text{GA}(V) \equiv \text{Tran}(V) \times \text{GL}(V) \tag{471}$$

Definition 12.6 (Isometries)

The **Euclidean group** of **isometries** in the Euclidean space $\mathbb{E}^n$ (with the Euclidean norm), denoted $\text{Isom}\, \mathbb{E}^n$ or $\mathbb{E}(n)$, consists of all distance-preserving bijections from $\mathbb{E}^n$ to itself, called **motions** or **rigid transformations**. It consists of all combinations of rotations, reflections, and translations. The **special Euclidean group** of all isometries that preserve the **handedness** of figures is denoted $\mathbb{SE}(n)$, which is comprised of all combinations rotations and translations called **rigid motions** or **proper rigid transformations**.

Definition 12.7 (Orthogonal Group)

The **orthogonal group**, denoted $\text{O}(n)$, consists of all isometries that preserve the origin, i.e. consists of rotations and reflections. The **special orthogonal group**, denoted $\text{SO}(n)$, is a subgroup of $\text{O}(n)$ consisting of only rotations. We can see that

$$\text{O}(n) = \frac{\text{Isom}\, \mathbb{E}^n}{\text{Tran}\, V} \tag{472}$$

Definition 12.8 (Transitive)

A transformation group G is called **transitive** if for any $x, y \in X$, there exists a $\phi \in G$ such that $y = \phi(x)$.

Example 12.4 ()

$\text{Tran}(V)$ and $\text{GA}(V)$ are transitive groups.

Definition 12.9 (Congruence Classes)

Let X be a set and G its transformation group on X. The way we define G determines the **geometry** of X. More specifically, a figure $F_1 \subset X$ is **equivalent** or **congruent** to $F_2 \subset X$ iff there exists $\phi \in G$ such that $F_2 = \phi(F_1)$ (or equivalently, $F_1 = \phi(F_2)$). This is an equivalence relation since

1. $F \sim F$.
2. $F \sim H \implies H \sim F$.
3. $F \sim H, H \sim K \implies F \sim K$

Two figures that are in the same equivalence class are known to be **congruent** with respect to the geometry of X induced by G.

Clearly, if two figures are congruent in Euclidean geometry, then they are congruent in Affine geometry, since $\text{E}(n) \subset \text{GA}(n)$.

12.1.1 Lie Algebras of $\text{SL}(2, \mathbb{R})$ and $\text{SL}(2, \mathbb{C})$

Given the group $\text{SL}(2, \mathbb{R})$, there must be a corresponding Lie algebra of matrices such that $g = e^A \in \text{SL}(2, \mathbb{R})$. We attempt to find this Lie algebra. Let $g \in \text{SL}(2, \mathbb{R})$, with $g = e^A$. So, if $\det g = 1$, what is the corresponding restriction on A in the algebra? We use the following proposition.

Proposition 12.1 ()

$$\det(e^A) = e^{\text{Tr}\,(A)} \tag{473}$$

Proof.

Put A in Jordan Normal Form: $A = S^{-1}JS \implies A^n = S^{-1}J^nS \implies exp(A) = S^{-1}exp(A)S \implies \det(exp(A)) = \det e^J$. But since J is upper trianglar, J^n is upper triangular $\implies e^J$ is upper triangular, which implies that

$$\det e^J = \prod_i e^{\lambda_i} = e^{\text{Tr}\,(J)} = e^{\text{Tr}\,(A)} \tag{474}$$

since trace is invariant under a change of basis.

So, $\det(e^A) = 1 \implies \text{Tr}\,(A) = 2\pi in$ for $n \in \mathbb{Z}$. Since we want to component connected to the identity, we choose $n = 0$ meaning that $\text{Tr}\,(A) = 0$. And we are done. That is, the Lie algebra of $\text{SL}(2, \mathbb{R})$ consists of traceless 2×2 matrices, denoted $\mathfrak{sl}_2\mathbb{R}$. $\mathfrak{sl}_2\mathbb{R}$ has basis (chosen arbitrarily)

$$\left\{ H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \tag{475}$$

and the identity in the Lie algebra is the zero matrix, which translates to the 2×2 identity matrix in the Lie group.

$$exp \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = I \tag{476}$$

We must not forget to define the bracket structure in $\mathfrak{sl}_2\mathbb{R}$, so we define it as the commutator, which gives the identity

$$\begin{aligned} [H, X] &= HX - XH = 2X \\ [H, Y] &= HY - YH = -2Y \\ [X, Y] &= XY - YX = H \end{aligned}$$

Note that regular matrix multiplication is not closed within this Lie algebra. For example,

$$XY = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{477}$$

is clearly not traceless. However, the bracket operation keeps the matrices within this traceless condition (and thus, within this algebra), so you can't just stupidly multiply matrices together in a Lie algebra. Remember that regular matrix multiplication does not have anything to do with the Lie bracket and does not apply to this group. This algebra also simplifies the multiplicative inverse of a group to a simple additive inverse, making calculations easier.

Similarly, the Lie algebra of $\text{SL}(2, \mathbb{C})$ also has the same basis

$$\left\{ H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \tag{478}$$

but we choose the field to be $\mathbb{C}$, meaning that we take complex linear combinations rather than real linear ones.

12.1.2 Lie Algebra of SU(2)

$g \in \mathrm{SU}(2) \implies \det g = 1 \implies \operatorname{Tr} A = 0$. We also see that by definition e^A,

$$(e^A)^\dagger = e^{A^\dagger} \text{ and } (e^A)^{-1} = e^{-A} \tag{479}$$

which implies that $A^\dagger = -A$. That is, the unitary condition implies that the Lie algebra elements in $\mathfrak{su}(2)$ are traceless, anti-self adjoint 2×2 matrices over $\mathbb{C}$.

Definition 12.10 ()

The **Pauli matrices** are the three matrices

$$\left\{ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \tag{480}$$

Note that with some calculation,

$$\begin{aligned} [\sigma_x, \sigma_y] &= 2i\sigma_z \\ [\sigma_y, \sigma_z] &= 2i\sigma_x \\ [\sigma_z, \sigma_x] &= 2i\sigma_y \end{aligned}$$

To identify the basis of $\mathfrak{su}(2)$, we take the Pauli matrices and let

$$\begin{aligned} A_x &\equiv -\frac{i}{2}\sigma_x = \begin{pmatrix} 0 & -i/2 \\ -i/2 & 0 \end{pmatrix} \\ A_y &\equiv -\frac{i}{2}\sigma_y = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix} \\ A_z &\equiv -\frac{i}{2}\sigma_z = \begin{pmatrix} -i/2 & 0 \\ 0 & i/2 \end{pmatrix} \end{aligned}$$

be the basis of $\mathfrak{su}(2)$. Clearly, A_x, A_y, A_z are all traceless, anti-self adjoint 2×2 matrices. Moreover, they also satisfy

$$\begin{aligned} [A_x, A_y] &= A_z \\ [A_y, A_z] &= A_x \\ [A_z, A_x] &= A_y \end{aligned}$$

However, note that the algebra $\mathfrak{su}(2)$ consists of all **real** linear combinations of A_x, A_y, A_z. That is, $\mathfrak{su}(2)$ is a 3 dimensional **real** vector space, even though it has basis elements containing complex numbers.

However, we can always complexify this space by simply replacing real scalar multiplication in $\mathfrak{su}(2)$ with complex scalar multiplication. By complexifying $\mathfrak{su}(2)$, the Lie group SU(2) formed by taking the exponential map on this complexified space is actually identical to $\mathrm{SL}(2, \mathbb{C})$. Indeed, this is true because first, the basis $\{H, X, Y\}$ of $\mathfrak{sl}_2\mathbb{C}$ and the basis $\{A_x, A_y, A_z\}$ of $\mathfrak{su}(2)$ span precisely the same subspace in the vector space $\mathrm{Mat}(2, \mathbb{C})$, meaning that the two Lie algebras are the same vector space. Secondly, the bracket operation $[\cdot, \cdot]$ in both $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{su}(2)$ are equivalent since the operation defined to be the commutator in both cases, resulting in the similarities in the bracket behaviors.

$$\begin{aligned} [H, X] = 2X &\iff [A_x, A_y] = A_z \\ [H, Y] = -2Y &\iff [A_y, A_z] = A_x \\ [X, Y] = H &\iff [A_z, A_x] = A_y \end{aligned}$$

Therefore, the complexification of SU(2) and $\mathrm{SL}(2, \mathbb{R})$ both leads to the construction of $\mathrm{SL}(2, \mathbb{C})$.

$$\begin{tikzcd}
\mathrm{SL}(2,\mathbb{R}) \arrow[dr, hook] & \\
& \mathrm{SL}(2,\mathbb{C}) \\
\mathrm{SU}(2) \arrow[ur, hook, "\text{complexify}"'] &
\end{tikzcd}$$

We can interpret the "real forms" of $\mathrm{SL}(2,\mathbb{C})$ as "slices" of some complex group. However, this does not mean that the real version of these groups are equal. That is,

$$\mathrm{SL}(2,\mathbb{R}) \neq \mathrm{SU}(2) \tag{481}$$

12.1.3 Lie Algebra of SO(3)

It is easy to see that for SO(2), it is easy to see that its Lie algebra $\mathfrak{so}(2)$ has

$$\left\{ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\} \tag{482}$$

as its only basis, since

$$exp\bigg(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \theta \bigg) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{483}$$

meaning that the dimension of SO(2) is 1. By adding a component, we can get a rotation in $\mathbb{R}^3$.

$$\begin{aligned}
R_x &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \implies e^{R_x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \\
R_y &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \implies e^{R_y} = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \\
R_z &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies e^{R_z} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

That is, e^{R_x}, e^{R_y}, and e^{R_z} generates a rotation around the x, y, and z axis, respectively, which completely generates the group SO(3). Therefore, the Lie algebra $\mathfrak{so}(3)$ consists of the basis

$$\{R_x, R_y, R_z\} \tag{484}$$

The bracket structure (again, defined as the commutator) of this Lie algebra is

$$\begin{aligned}
[R_x, R_y] &= R_z \\
[R_y, R_z] &= R_x \\
[R_z, R_x] &= R_y
\end{aligned}$$

which is similar to the brakcet structure of $\mathfrak{su}(2)$. Therefore, SO(3) and SU(2) have the **same** Lie algebra, which is the algebra of dimension 3 with the same bracket structure. Note that Lie algebras are uniquely determined by the bracket structure and dimension. However, having the same Lie algebra does not imply that the groups are identical (obviously) nor isomorphic. For example,

$$exp(2\pi R_z) = \begin{pmatrix} \cos 2\pi & -\sin 2\pi & 0 \\ \sin 2\pi & \cos 2\pi & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \tag{485}$$

while

$$exp(2\pi A_z) = exp(-i\pi\sigma_z) = exp\bigg(-i\pi \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigg) = -I \tag{486}$$

There is discrepancy by a factor of -1. In fact, it turns out that

$$\mathrm{SO}(3) = \frac{\mathrm{SU}(2)}{\pm I} \tag{487}$$

We justify this in the following way. Let $v \in \mathbb{R}^3$ have components (x, y, z). Consider

$$M = x\sigma_x + y\sigma_y + z\sigma_z \tag{488}$$

M is clearly traceless and $M^\dagger = M$. Now, let $S \in \mathrm{SU}(2)$ and let $M' = S^{-1}MS$. Then, $\operatorname{Tr} M' = \operatorname{Tr} S^{-1}MS = \operatorname{Tr} M = 0$ and $(M')^\dagger = (S^{-1}MS)^\dagger = S^\dagger M^\dagger (S^{-1})^\dagger = S^{-1}MS = M'$. Therefore, since M' is self adjoint and traceless, it can be expressed in the form

$$x'\sigma_x + y'\sigma_y + z'\sigma_z \tag{489}$$

for some (x', y', z'). Now, since

$$M^2 = (-x^2 - y^2 - z^2)I \tag{490}$$

we have

$$\begin{aligned} (M')^2 = S^{-1}M^2S &= (-x^2 - y^2 - z^2)I \\ &= (-x'^2 - y'^2 - z'^2)I \end{aligned}$$

So, $x^2 + y^2 + z^2 = x'^2 + y'^2 + z'^2$, implying that the lengths of v stayed the same. (The proof of linearity of S is easy.) Therefore, the transformation $M \mapsto M'$, i.e. $(x, y, z) \mapsto (x', y', z')$ is a linear transformation preserving length in $\mathbb{R}^3$ (with respect to the usual inner product and norm) $\implies$ it is in SO(3). If we have

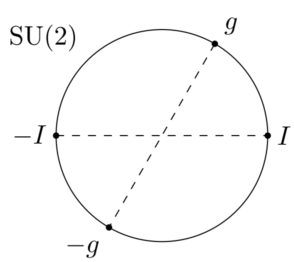
$$S = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{491}$$

then $M' = M$, which explains why SO(3) is a coset deviating by both I and $-I$. Visually, if we let SU(2) be a circle, points that are diametrically opposite of each other are "equivalent" in SO(3). That is, SU(2) is a three-dimensional sphere, and g and $-g$ are identified onto the same element in SO(3). This map

$$\rho : \mathrm{SU}(2) \to \mathrm{SO}(3) \tag{492}$$

in which 2 points are mapped to 1 point is a surjective map with

$$\ker \rho = \{I, -I\} \tag{493}$$



We can in fact explicitly describe exponential map from $\mathfrak{so}(3)$ to SO(3) with the following lemma.

Lemma 12.1 (Rodrigues' Formula)

The exponential map $exp : \mathfrak{so}(3) \to \mathrm{SO}(3)$ is defined by

$$e^A = \cos\theta I_3 + \frac{\sin\theta}{\theta}A + \frac{(1-\cos\theta)}{\theta^2}B \tag{494}$$

where

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}, B = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} \tag{495}$$

This formula has many applications in kinematics, robotics, and motion interpolation.

Theorem 12.2 ()

The Lie algebras for the following classical Lie groups are summarized as follows.

1. $\mathfrak{sl}_n\mathbb{R}$ is the real vector space of real $n \times n$ matrices with null trace.
2. $\mathfrak{so}(n)$ is the real vector space of real $n \times n$ skew-symmetric matrices.
3. $\mathfrak{gl}_n\mathbb{R}$ is the real vector space of all real $n \times n$ matrices.
4. $\mathfrak{o}(n) = \mathfrak{o}(n)$

Note that the corresponding groups $\mathrm{GL}(n, \mathbb{R}), \mathrm{SL}(n, \mathbb{R}), \mathfrak{gl}_n\mathbb{R}, \mathfrak{sl}_n\mathbb{R}$ are Lie groups, meaning that they are smooth real manifolds. We can view each of them as smooth real manifolds embedded in the n^2 dimensional vector space of real matrices, which is isomorphic to $\mathbb{R}^{n^2}$.

Theorem 12.3 ()

The Lie algebras $\mathfrak{gl}_{\mathbb{R}}, \mathfrak{sl}_n\mathbb{R}, \mathfrak{o}(n), \mathfrak{so}(n)$ are well-defined, but only

$$exp : \mathfrak{so}(n) \to \mathrm{SO}(n) \tag{496}$$

is surjective.

Theorem 12.4 ()

The Lie algebras for the following classical Lie groups are summarized as follows.

1. $\mathfrak{sl}_2\mathbb{C}$ is the real (or complex) vector space of traceless complex $n \times n$ matrices.
2. $\mathfrak{u}(n)$ is the real vector space of complex $n \times n$ skew-Hermitian matrices.
3. $\mathfrak{su}(n) = \mathfrak{u} \cap \mathfrak{sl}_2\mathbb{C}$. It is also a real vector space.
4. $\mathfrak{gl}_n\mathbb{C}$ is the real (or complex) vector space of complex $n \times n$ matrices.

Note that even though the matrices in these Lie algebras have complex coefficients, we have assigned them to be in a **real** vector space, which means that we are only allowed to take real linear combinations of these elements. That is, the field we are working over is $\mathbb{R}$ (this does not contradict any of the axioms for vector spaces). For example an element A in $\mathfrak{u}(n)$ or $\mathfrak{su}(n)$ must be anti-self adjoint, but iA is self adjoint.

Similarly, the Lie groups

$$\mathrm{GL}(n, \mathbb{C}), \mathrm{SL}(n, \mathbb{C}), \mathfrak{gl}_n\mathbb{C}, \mathfrak{sl}_n\mathbb{C} \tag{497}$$

are also smooth real manifolds embedded in $\mathrm{Mat}(n, \mathbb{C}) \simeq \mathbb{C}^{n^2} \simeq \mathbb{R}^{2n^2}$. So, we can view these four groups as manifolds embedded in $\mathbb{R}^{2n^2}$.

Note some of the similarities and differences between the real and complex counterparts of these Lie groups and algebras.

1. $\mathfrak{o}(n) = \mathfrak{so}(n)$, but $\mathfrak{u}(n) \neq \mathfrak{su}(n)$.
2. $exp : \mathfrak{gl}_n\mathbb{R} \to \mathrm{GL}(n, \mathbb{R})$ is not surjective, but $exp : \mathfrak{gl}_n\mathbb{C} \to \mathrm{GL}(n, \mathbb{C})$ is surjective due to the spectral theorem and surjectivity of $exp : \mathbb{C} \to \mathbb{C}^*$.
3. The exponential maps $exp : \mathfrak{u}(n) \to \mathrm{U}(n)$ and $exp : \mathfrak{su}(n) \to \mathrm{SU}(n)$ are surjective.
4. Still, $exp : \mathfrak{sl}_2\mathbb{C} \to \mathrm{SL}(2, \mathbb{C})$ is not surjective. This will be proved now.

Theorem 12.5 ()

$exp : \mathfrak{sl}_2\mathbb{C} \to \mathrm{SL}(2, \mathbb{C})$ is not surjective.

Proof.

Given $M \in \mathrm{SL}(n, \mathbb{C})$, assume that $M = e^A$ for some matrix $A \in \mathfrak{sl}_2\mathbb{C}$. Putting A into the Jordan Normal Form $J = NAN^{-1}$ means that J can either be of form

$$J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \implies e^J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} e^\lambda & 0 \\ 0 & e^{-\lambda} \end{pmatrix} \tag{498}$$

which is also in JNF in $\mathrm{SL}(2, \mathbb{C})$. But a matrix $P \in \mathrm{SL}(2, \mathbb{C})$ may exist with JNF of

$$K = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \tag{499}$$

which is not one of the 2 forms. So, $K \notin \mathrm{Im}\, exp \implies exp$ is not surjective.

Theorem 12.6 ()

The exponential maps

$$\begin{aligned} exp &: \mathfrak{u}(n) \to \mathrm{U}(n) \\ exp &: \mathfrak{su}(n) \to \mathrm{SU}(n) \end{aligned}$$

are surjective.

12.1.4 Lie Algebra of SE(n)

Recall that the group of affine rigid isometries is denoted $\mathrm{SE}(n)$. That is,

$$\mathrm{SE}(n) \equiv \mathrm{SO}(n) \ltimes \mathrm{Tran}\, \mathbb{R}^n \tag{500}$$

We can define the matrix representation of this affine transformation as such. Given an element $g \in \mathrm{SE}(n)$ such that

$$g(x) \equiv Rx + U,\ R \in \mathrm{SO}(n), U \in \mathrm{Tran}\, \mathbb{R}^n \tag{501}$$

we define the representation

$$\rho : \mathrm{SE}(n) \to \mathrm{GL}(n+1, \mathbb{R}), \rho(g) \equiv \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \tag{502}$$

where R is a real $n \times n$ matrix in $\mathrm{SO}(n)$ and U is a real n-vector in $\mathrm{Tran}\, \mathbb{R}^n \simeq \mathbb{R}^n$. We would then have

$$\rho(g) \begin{pmatrix} x \\ 1 \end{pmatrix} \equiv \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Rx + U \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1} \tag{503}$$

Clearly, $\mathrm{SE}(n)$ is a Lie group, and the matrix representation ϱ of its Lie algebra $\mathfrak{se}(n)$ can be defined as the vector space of $(n+1) \times (n+1)$ matrices of the block form

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix} \tag{504}$$

where Ω is an $n \times n$ skew-symmetric matrix and $U \in \mathbb{R}^n$. Note that there are two different exponential maps here: one belonging to the abstract Lie group $\mathrm{SE}(n)$ and another belonging to the concrete, matrix group

$\mathrm{GL}(n+1, \mathbb{R})$. This can be represented with the commutative diagram.

$$\begin{tikzcd} \mathfrak{se}(n) \arrow[r, "exp"] \arrow[d, "\varrho"] & SE(n) \arrow[d, "\rho"] \\ \mathfrak{gl}_{n+1}\mathbb{R} \arrow[r, "exp"] & \mathrm{GL}(n+1, \mathbb{R}) \end{tikzcd}$$

Lemma 12.2 ()

Given any $(n+1) \times (n+1)$ matrix of form

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix} \tag{505}$$

where Ω is any matrix and $U \in \mathbb{R}^n$,

$$A^k = \begin{pmatrix} \Omega^k & \Omega^{k-1}U \\ 0 & 0 \end{pmatrix} \tag{506}$$

where $\Omega^0 = I_n$, which implies that

$$e^A = \begin{pmatrix} e^\Omega & VU \\ 0 & 1 \end{pmatrix}, \; V = I_n + \sum_{k \geq 1} \frac{\Omega^k}{(k+1)!} \tag{507}$$

Theorem 12.7 ()

The exponential map

$$exp : \mathfrak{se}(n) \to SE(n) \tag{508}$$

is well-defined and surjective.

12.2 Representations of Lie Groups and Lie Algebras

Let $\mathcal{G}$ be an abstract group and let

$$\rho : \mathcal{G} \to \mathrm{GL}(V) \tag{509}$$

be the representation of $\mathcal{G}$. Then, let $\mathfrak{g}$ be the Lie algebra of $\mathcal{G}$, and $\mathfrak{gl}(V)$ be the Lie algebra of $\mathrm{GL}(V)$. Then, ρ induces another homomorphism

$$\varrho : \mathfrak{g} \to \mathfrak{gl}(V) \tag{510}$$

where the bracket structure (in this case, the comutator in the matrix algebra) is preserved.

$$\varrho([X, Y]) = [\varrho(X), \varrho(Y)] \tag{511}$$

We can visualize this induced homomorphism with the following commutative diagram, which states that $\rho \circ exp = exp \circ \varrho$.

$$\begin{tikzcd} \mathcal{G} \arrow[r, "\rho"] & \mathrm{GL}(V) \\ \mathfrak{g} \arrow[u, "exp"] \arrow[r, "\varrho"] & \mathfrak{gl}(V) \arrow[u, "exp"'] \end{tikzcd}$$

Note that there are very crucial differences between ρ and ϱ. First, ρ is a homomorphism between **groups**, while ϱ is a homomorphism between **vector spaces**. Additionally, $\mathrm{GL}(V)$ is a group, not a linear space, while

$\mathfrak{gl}(V)$ is a linear space. Finally, note that $\text{GL}(V)$ is restricted to only matrices with nonzero determinants, while the elements of $\mathfrak{gl}(V)$ can be any matrix.

Example 12.5 ()

The representation of $\text{SE}(n)$ to $\text{GL}(n+1\mathbb{R}$ and $\mathfrak{se}(n)$ to $\mathfrak{gl}_{n+1}\mathbb{R}$ induces the second homomorphism $\varrho : \mathfrak{gl}_{n+1}\mathbb{R} \rightarrow \text{GL}(n+1, \mathbb{R})$.

Definition 12.11 ()

The direct sum of representations is a representation. That is, if U is a representation and V is a representation, then $U \oplus V$ is a representation. That is, if

$$\rho_1 : \mathcal{G} \rightarrow U, \ \rho_1(g) = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} \tag{512}$$

and

$$\rho_2 : \mathcal{G} \rightarrow V, \ \rho_2(g) = \begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix} \tag{513}$$

are two representations of the same group element $g \in \mathcal{G}$, then

$$(\rho_1 \oplus \rho_2) : \mathcal{G} \rightarrow (U \oplus V), \ (\rho_1 \oplus \rho_2)(g) = \begin{pmatrix} u_1 & u_2 & 0 & 0 \\ u_3 & u_4 & 0 & 0 \\ 0 & 0 & v_1 & v_2 \\ 0 & 0 & v_3 & v_4 \end{pmatrix} \tag{514}$$

is a bigger representation of g in $U \oplus V$.

Definition 12.12 ()

V is irreducible if the only subspaces which are representations are only V and $\{0\}$.

For our case, we will consider that any representation can be written as a direct sum of irreducible representations. We will now proceed to find an irreducible representation of $\mathfrak{sl}_2\mathbb{C}$. This means that we want to find the smallest (lowest dimensional) vector space V such that there exists a representation

$$\varrho : \mathfrak{sl}_2\mathbb{C} \rightarrow \mathfrak{gl}(V) \tag{515}$$

We will write, as shorthand notation, that

$$H = \varrho(H), X = \varrho(X), Y = \varrho(Y) \tag{516}$$

Clearly, $H, X, Y \in \mathfrak{gl}(V) \simeq \mathfrak{gl}(\mathbb{C}^n)$. By the spectral theorem, we can find an orthonormal basis of eigenvectors $e_1, e_2, ..., e_n$ of the mapping H such that

$$He_i = \lambda_i e_i, \ \lambda_i \in \mathbb{C} \tag{517}$$

Since $[H, X] = 2X$, it follows that $HXe_i - XHe_i = 2Xe_i \implies H(Xe_i) = (\lambda_i + 2)(Xe_i) \implies Xe_i$ for all $i = 1, 2, ..., n$ are also eigenvectors of H with eigenvalue $(\lambda_i + 2)$, or $Xe_i = 0$. So, X is a "ladder operator" that maps each eigenvector e_i with eigenvalue λ_i to a different eigenvector e_j with eigenvalue $\lambda_j = \lambda_i + 2$. Having nowhere to be mapped to, the eigenvector with the largest eigenvalue (which must exist since V is finite dimensional) will get mapped to the 0 vector by X. Let us denote this eigenvector having the maximum eigenvalue m, as v_m.

Similarly, $[H, Y] = -2Y$ implies that

$$HYe_i - YHe_i = -2Ye_i \implies H(Ye_i) = (\lambda_i - 2)(Ye_i) \tag{518}$$

implying that Y maps each eigenvector e_i with eigenvaue λ_i to another eigenvector e_j with eigenvalue $\lambda_j = \lambda_i - 2$, except for the eigenvector with smallest eigenvalue, which gets mapped to 0. Since Y clearly maps each eigenvector to a different eigenvector that has a strictly decreasing eigenvalue, we can construct a basis of V to be

$$\{v_m, Yv_m, Y^2 v_m, Y^3 v_m, ..., Y^{n-1} v_m\} \tag{519}$$

(remember that $Y^n v_m = 0$). So, elements of $\mathfrak{sl}_2\mathbb{C}$ acts on the space V with basis above. To continue, we introduce the following proposition.

Proposition 12.2 ()

$$XY^j v_m = j(m - j + 1)Y^{j-1} v_m \tag{520}$$

Proof.

By induction on j using bracket relations.

V is n-dimensional. Since $Y^n v_m = 0$ and $Y^{n-1} v_m \neq 0$, we use the proposition above to get

$$0 = XY^n v_m = n(m - n + 1)Y^{n-1} v_m \implies m - n + 1 = 0 \tag{521}$$

So, $n = m + 1$, which means that the eigenvalues of H are

$$m, m - 2, m - 4, \dots, m - 2(n - 1) = -m \tag{522}$$

and we are done. We now classify the 1, 2, and 3 dimensional irreducible representations of $\mathfrak{sl}_2\mathbb{C}$.

1. When $n = 1$ (i.e. dimension is 1), $m = n - 1 = 0$, meaning that the greatest (and only) eigenvalue is 0. That is,
$$Hv_0 = 0,\ Xv_0 = 0,\ Yv_0 = 0 \tag{523}$$
which is the trivial representation of $\mathfrak{sl}_2\mathbb{C}$. Explicitly, we can completely define the representation (which is a linear homomorphism) with the three equations.
$$\varrho(H) = (0),\ \varrho(X) = (0),\ \varrho(Y) = (0) \tag{524}$$

2. When $n = 2$ and $m = 1$. We now look for a 2 dimensional irreducible representation. The eigenvalues are 1 and -1, with $\{v_1, v_{-1}\}$ as a basis of 2 dimensional space V. Then we have
$$\begin{aligned} &Hv_1 = v_1,\ Hv_{-1} = -v_{-1} \\ &Xv_1 = 0,\ Xv_{-1} = v_1 \\ &Yv_1 = v_{-1},\ Yv_{-1} = 0 \end{aligned}$$
which explicitly translates to the representation ϱ being defined
$$\varrho(H) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{525}$$

3. When $n = 3 \implies m = 2$, the basis is $\{v_{-2}, v_0, v_2\}$ with eigenvalues $2, 0, -2$, and the irreducible representation ϱ is defined
$$\varrho(H) = \begin{pmatrix} 2 & & \\ & 0 & \\ & & -2 \end{pmatrix}, \varrho(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \varrho(X) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{526}$$

4. The same process continues on for $n = 4, 5, ...$, and this entirely classifies the irreducible representations of $\mathfrak{sl}_2\mathbb{C}$.

12.2.1 Tensor Products of Group Representations

Definition 12.13 ()

If V and W are two different representations of a group $\mathcal{G}$, then we know that $V \oplus W$ is also a representation of $\mathcal{G}$. Furthermore, the tensor product space $V \otimes W$ also defines a representation of $\mathcal{G}$. That is, given representations

$$\begin{aligned} \rho_V &: \mathcal{G} \to \mathrm{GL}(V) \\ \rho_W &: \mathcal{G} \to \mathrm{GL}(W) \end{aligned}$$

The homomorphism $\rho_V \otimes \rho_W : \mathcal{G} \to \mathrm{GL}(V \otimes W)$ is also a representation of $\mathcal{G}$, which is defined

$$(\rho_V \otimes \rho_W)(g)(v \otimes w) \equiv \rho_V(g)(v) \otimes \rho_W(g)(w) \tag{527}$$

or represented in shorthand notation,

$$g(v \otimes w) \equiv (gv) \otimes (gw) \tag{528}$$

We know that $\exp(H)$ acts on V and W since it is an element of $\mathrm{GL}(V)$ and $\mathrm{GL}(W)$. This means that

$$exp(H)(v \otimes w) \equiv \big(exp(H)(v)\big) \otimes \big(exp(H)(w)\big) \tag{529}$$

If H $(= \rho_V(H)$ or $\rho_W(H))$ has an eigenvalue λ on v in V and eigenvalue μ on w in W, then

$$exp(H)(v \otimes w) = (e^\lambda v) \otimes (e^\mu w) = e^{\lambda+\mu} v \otimes w \tag{530}$$

That is, eigenvalues of H **add** on tensor products.

Example 12.6 ()

Recall that the 2 dimensional representation V of $\mathfrak{sl}_2\mathbb{C}$ has eigenvalues 1 and -1 (with corresponding eigenvectors e_1 and e_{-1}). So, $V \otimes V$ has eigenvalues

$$\begin{aligned} &(-1) + (-1) = -2, \;\; (-1) + 1 = 0 \\ &1 + (-1) = 0, \;\; 1 + 1 = 2 \end{aligned}$$

Therefore, the eigenvalues of $V \otimes V$ is -2 (geometric multiplicity of 1), 0 (geometric multiplicity of 2), and 2 (geometric multiplicity of 1), (Notation-wise, the n-dimensional irreducible representation of $\mathfrak{sl}_2\mathbb{C}$ is denoted $\mathbf{n}$.) which means that

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1} \tag{531}$$

We can decompose $V \otimes V$ into its symmetric and exterior power components. $\mathrm{Sym}^2 V$ has basis (of eigenvectors)

$$\{e_{-1} \odot e_{-1}, \; e_{-1} \odot e_1, \; e_1 \odot e_1\} \tag{532}$$

where the corresponding eigenvalues are -2, 0, and 2, respectively. So, $\dim Sym^2 V = 3$, which means that $Sym^2 V = \mathbf{3}$. As for the exterior power component of V, $\Lambda^2 V$ has basis $\{e_{-1} \wedge e_1\}$ with eigenvalue $= 0 \implies \dim \Lambda^2 V = 1$, meaning that $\Lambda^2 V = \mathbf{1}$. Therefore,

$$V \otimes V = Sym^2 V \oplus \Lambda^2 V = \mathbf{3} \oplus \mathbf{1} \tag{533}$$

12.3 Topological Decompositions of Lie Groups

Definition 12.14 ()

Let us define
1. $\mathrm{S}(n)$ is the vector space of real, symmetric $n \times n$ matrices.
2. $\mathrm{SP}(n)$ is the set of symmetric, positive semidefinite matrices.
3. $\mathrm{SPD}(n)$ is the set of symmetric, positive definite matrices.

Note that $\mathrm{SP}(n)$ and $\mathrm{SPD}(n)$ are not even vector spaces at all.

Lemma 12.3 ()

The exponential map
$$exp : S(n) \to SPD(n) \tag{534}$$
is a homeomorphism. One may be tempted to call $\mathrm{S}(n)$ the Lie algebra of $\mathrm{SPD}(n)$, but this is not the case. $\mathrm{S}(n)$ is not even a Lie algebra since the commutator is not algebraically closed. Furthermore, $\mathrm{SPD}(n)$ is not even a multiplicative group (since matrix multiplication is not closed).

Recall from linear algebra the Polar Decomposition. We express this result in a slightly modified way.

Theorem 12.8 (Polar Decomposition)

Given a Euclidean space $\mathbb{E}^n$ and any linear endomorphism f of $\mathbb{E}^n$, there are two positive definite self-adjoint linear maps $h_1, h_2 \in \mathrm{End}(\mathbb{E}^n)$ and $g \in \mathrm{O}(n)$ such that
$$f = g \circ h_1 = h_2 \circ g \tag{535}$$
That is, such that f can be decomposed into the following compositions of functions that commute.

$$\begin{tikzcd}
\mathbb{E}^n \arrow[r, "h_2"] & \mathbb{E}^n \\
\mathbb{E}^n \arrow[u, "g"] \arrow[r, "h_1"'] \arrow[ur, "f"] & \mathbb{E}^n \arrow[u, "g"']
\end{tikzcd}$$

This means that there is a bijection between $Mat(n, \mathbb{R})$ and $O(n) \times SP(n)$. If f is an automorphism, then this decomposition is unique.

Corollary 12.1 ()

The two topological groups are homeomorphic.
$$GL(n, \mathbb{R}) \cong O(n) \times SPD(n) \tag{536}$$

Corollary 12.2 ()

For every invertible real matrix $A \in GL(n, \mathbb{R})$, there exists a unique orthogonal matrix R and unique symmetric matrix S such that
$$A = Re^S \tag{537}$$
$\implies$ there is a bijection between $\mathrm{GL}(n, \mathbb{R})$ and $O(n) \times S(n) \simeq \mathbb{R}^{n(n+1)/2}$. Moreover, they are homeomorphic. That is,
$$\mathrm{GL}(n, \mathbb{R}) \simeq O(n) \times S(n) \simeq O(n) \times \mathbb{R}^{n(n+1)/2} \tag{538}$$

This essentially reduces the study of $\mathrm{GL}(n,\mathbb{R})$ to the study of $O(n)$, which is nice since $O(n)$ is compact.

Corollary 12.3 ()

Given a real matrix A, if $\det A > 0$, then we can decompose A as

$$A = Re^S \tag{539}$$

where $R \in SO(n)$ and $S \in S(n)$.

Corollary 12.4 ()

There exists a bijection between

$$\mathrm{SL}(n,\mathbb{R}) \text{ and } SO(n) \times (S(n) \cap \mathfrak{sl}_n\mathbb{R}) \tag{540}$$

Proof.

$A \in \mathrm{SL}(n,\mathbb{R}) \implies 1 = \det A = \det R \det e^S = \det e^S \implies \det e^S = e^{\mathrm{Tr}\, S} = 1 \implies \mathrm{Tr}\, S = 0 \implies S \in \mathrm{S}(n) \cap \mathfrak{sl}_n\mathbb{R}$.

Definition 12.15 ()

Let us define

1. $\mathrm{H}(n)$ is the real vector space of $n \times n$ Hermitian matrices.
2. $\mathrm{HP}(n)$ is the set of Hermitian, positive semidefinite $n \times n$ matrices.
3. $\mathrm{HPD}(n)$ is the set of Hermitian, positive definite $n \times n$ matrices.

Similarly, $\mathrm{HP}(n)$ and $\mathrm{HPD}(n)$ are not vector space. They are just sets.

Lemma 12.4 ()

The exponential mapping

$$exp : H(n) \rightarrow HPD(n) \tag{541}$$

is a homeomorphism.

However again, $\mathrm{HPD}(n)$ is not a Lie group (multiplication is not algebraically closed) nor is $\mathrm{H}(n)$ a Lie algebra (commutator is not algebraically closed). By the polar form theorem of complex $n \times n$ matrices, we have a (not necessarily unique) bijection between

$$\mathrm{Mat}(n,\mathbb{C}) \text{ and } U(n) \times HP(n) \tag{542}$$

which implies that

$$\mathrm{GL}(n,\mathbb{C}) \cong U(n) \times HPD(n) \tag{543}$$

Corollary 12.5 ()

For every complex invertible matrix A, there exists a unique decomposition

$$A = Ue^S \tag{544}$$

where $U \in U(n)$ and $S \in H(n)$, which implies that the following groups are homeomorphic.

$$\begin{aligned} \mathrm{GL}(n, \mathbb{C}) &\cong U(n) \times H(n) \\ &\cong U(n) \times \mathbb{R}^{n^2} \end{aligned}$$

This essentially reduces the study of $\mathrm{GL}(n, \mathbb{C})$ to that of $\mathrm{U}(n)$.

Corollary 12.6 ()

There exists a bijection between

$$\mathrm{SL}(n, \mathbb{C}) \text{ and } SU(n) \times (H(n) \cap \mathfrak{sl}_n\mathbb{C}) \tag{545}$$

Proof.

Similarly, when $A = Ue^S$, we know that $|\det U| = 1$ and $\operatorname{Tr} S$ is real (since by the Spectral theorem, every self adjoint matrix has a real spectral decomposition). Since S is Hermitian, this implies that $\det e^S > 0$. If $A \in \mathrm{SL}(n, \mathbb{C})$, then $\det A = 1 \implies \det e^S = 1 \implies S \in H(n) \cap \mathfrak{sl}_n\mathbb{C}$.

12.4 Linear Lie Groups

We will assume that the reader has the necessary background knowledge in manifolds, chart mappings, diffeomorphisms, tangent spaces, and transition mappings.

Recall that the algebra of real $n \times n$ matrices $\mathrm{Mat}(n, \mathbb{R})$ is bijective to $\mathbb{R}^{n^2}$, which is a topological space. Therefore, this bijection

$$i : (\mathbb{R}^{n^2}, \tau_E) \to \mathrm{Mat}(n, \mathbb{R}) \tag{546}$$

induces a topology on $\mathrm{Mat}(n, \mathbb{R})$, defined

$$\tau_M \equiv \{U \in \mathrm{Mat}(n, \mathbb{R}) \mid e^{-1}(U) \in \tau_E\} \tag{547}$$

With this, consider the subset

$$\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{Mat}(n, \mathbb{R}) \tag{548}$$

where

$$\mathrm{GL}(n, \mathbb{R}) \equiv \{x \in \mathrm{Mat}(n, \mathbb{R}) \mid \det x \neq 0\} \tag{549}$$

This set, as we expect, is a multiplicative group.

Definition 12.16 ()

The **general linear group**, denoted $\mathrm{GL}(n, \mathbb{R})$ is the set of $n \times n$ matrices with nonzero determinant. The more technical definition is that $\mathrm{GL}(n, \mathbb{R})$ is really just the automorphism group of $\mathbb{R}^n$,

$$\mathrm{GL}(n, \mathbb{R}) \equiv \mathrm{Aut}(\mathbb{R}^n) \tag{550}$$

but it is customary to assume a basis on $\mathbb{R}^n$ in order to realize $\mathrm{GL}(n, \mathbb{R})$ as a matrix group. Note that the procedure of assuming a basis on $\mathbb{R}^n$ is the same as defining a representation of the abstract group $\mathrm{GL}(n, \mathbb{R})$. Both assigns a real $n \times n$ matrix to each element of $\mathrm{GL}(n, \mathbb{R})$.

In this way, we can view $\mathrm{GL}(n, \mathbb{R})$ as a topological space in $\mathbb{R}^{n^2}$, and it is fine to interpret $\mathrm{GL}(n, \mathbb{R})$ as a matrix group rather than an abstract group.

Since the matrix representation of $\mathrm{GL}(n, \mathbb{R})$ is always well defined, the abstract subgroups of $\mathrm{GL}(n, \mathbb{R})$, which are $\mathrm{SL}(n, \mathbb{R}), O(n)$, and $SO(n)$, also have well defined matrix representations (that we are all familiar with). Additionally, since there exists a bijection

$$\mathrm{Mat}(n, \mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2} \tag{551}$$

we can view $\mathrm{GL}(n, \mathbb{C})$ as a subset of $\mathbb{R}^{2n^2}$, meaning that the subgroups $\mathrm{SL}(n, \mathbb{C}), U(n)$, and $SU(n)$ of $\mathrm{GL}(n, \mathbb{C})$ can also be viewed as subsets of $\mathbb{R}^{2n^2}$. This also applies to $SE(n)$ since it is a subgroup of $\mathrm{SL}(n+1, \mathbb{R})$. We formally state it now.

Proposition 12.3 ()

$\mathrm{SE}(n)$ is a linear Lie group.

Proof.

The matrix representation of elements $g \in SE(n)$ is

$$\rho(g) \equiv \begin{pmatrix} R_g & U_g \\ 0 & 1 \end{pmatrix}, \; R_g \in SO(n), U_g \in \mathbb{R}^n \tag{552}$$

But such matrices also belong to the bigger group $\mathrm{SL}(n+1, \mathbb{R}) \implies SE(n) \subset \mathrm{SL}(n+1, \mathbb{R})$. Moreover, this canonical embedding

$$i : SE(n) \to \mathrm{SL}(n+1, \mathbb{R}) \tag{553}$$

is a group homomorphism since

$$\begin{aligned} i\big(\rho(g_1 \cdot g_2)\big) &= \begin{pmatrix} RS & RV+U \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S & V \\ 0 & 1 \end{pmatrix} = \rho\big(i(g_1) \cdot i(g_2)\big) \end{aligned}$$

and the inverse is given by

$$\begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R^T & -R^T U \\ 0 & 1 \end{pmatrix} \tag{554}$$

is also consistent between the inverse operation in $\mathrm{SE}(n)$ and $\mathrm{SL}(n+1, \mathbb{R})$. Therefore, $\mathrm{SE}(n)$ is a subgroup of $\mathrm{SL}(n+1, \mathbb{R})$, which is a subgroup of $\mathrm{GL}(n+1, \mathbb{R})$.

Note that even though $\mathrm{SE}(n)$ is diffeomorphic (a topological relation) to $\mathrm{SO}(n) \times \mathbb{R}^n$, it is **not** isomorphic (an algebraic relation) sicne group operations are not preserved. Therefore, we write this "equality" as a semidirect product of groups.

$$SE(n) \equiv SO(n) \ltimes \mathbb{R}^n \tag{555}$$

Therefore, all of the classical Lie groups that we have mentioned can be viewed as subsets of $\mathbb{R}^N$ (with the subspace topology) and as subgroups of $\mathrm{GL}(N, \mathbb{R})$ for some big enough N. This defines a special family of Lie groups, called linear Lie groups.

Definition 12.17 ()

A **linear Lie group** is a subgroup of $\mathrm{GL}(n, \mathbb{R})$ for some $n \geq 1$ which is also a smooth manifold in $\mathbb{R}^{n^2}$.

Theorem 12.9 (Von Neumann, Cartan)

A closed subgroup $\mathcal{G}$ of $\text{GL}(n, \mathbb{R})$ is a linear Lie group. That is, a closed subgroup $\mathcal{G}$ of $\text{GL}(n, \mathbb{R})$ is a smooth manifold in $\mathbb{R}^{n^2}$.

Definition 12.18 ()

Since a linear Lie group $\mathcal{G}$ is a smooth submanifold in $\mathbb{R}^N$, we can take its tangent space at the identity element I, which is defined

$$T_I \mathcal{G} \equiv \{p'(0) \mid p : I \subset \mathbb{R} \to \mathcal{G}, p(0) = I\} \tag{556}$$

where p is a path function on $\mathcal{G}$.

Note that we haven't mentioned anything about the exponential map up to now. We mention the relationship between this map and the Lie algebra with the following theorem.

Theorem 12.10 ()

Let $\mathcal{G}$ be a linear Lie group. The set $\mathfrak{g}$ defined such that

$$\mathfrak{g} \equiv \{X \in \text{Mat}(n, \mathbb{R}) \mid e^{tX} \in \mathcal{G} \; \forall t \in \mathbb{R}\} \tag{557}$$

is equal to the tangent space of $\mathcal{G}$ at the identity element. That is,

$$\mathfrak{g} = T_I \mathcal{G} \tag{558}$$

Furthermore, $\mathfrak{g}$ is closed under the commutator

$$[A, B] \equiv AB - BA \tag{559}$$

This theorem ensures that given a linear Lie group $\mathcal{G}$, the tangent space $\mathfrak{g}$ exists and is closed under the commutator. We formally define this space.

Definition 12.19 ()

The Lie algebra of a linear Lie group is a real vector space (of matrices) together with a algebraically closed bilinear map

$$[A, B] \equiv AB - BA \tag{560}$$

called the **commutator**.

The definition of $\mathfrak{g}$ given in the previous theorem shows that

$$exp : \mathfrak{g} \to \mathcal{G} \tag{561}$$

is well defined. In general, exp is neither injective nor surjective. Visually, this exponential mapping is what connects the Lie algebra, i.e. the tangent space of manifold $\mathcal{G}$ to the actual Lie group $\mathcal{G}$. To define the inverse map that maps Lie group elements to Lie algebra ones, we can simply just compute the tangent vectors of the manifold $\mathcal{G}$ at the identity I by taking the derivative of arbitrary path functions in $\mathcal{G}$. That is, for every $X \in T_I \mathcal{G}$, we define the smooth curve

$$\gamma_X : t \mapsto e^{tX} \tag{562}$$

where $\gamma_X(0) = I$. If we take the derivative of this curve, with respect to t at $t = 0$, we will get the tangent vector X corresponding to that group element $g = e^X$. More visually, we just need to take the collection of all smooth path functions γ on manifold $\mathcal{G}$ such that $\gamma(0) = I$. Then, taking the derivative of all these paths at $t = 0$ will produce the collection of all tangent vectors at the identity element. We show this process in the following examples.

Theorem 12.11 ()

The matrix representation of $\mathfrak{sl}_n\mathbb{R}$ is precisely the set of traceless $n \times n$ matrices.

Proof.

Clearly, $\mathfrak{sl}_n\mathbb{R}$ is a vector space since it is a Lie algebra. So, $X \in \mathfrak{sl}_n\mathbb{R} \implies tX \in \mathfrak{sl}_n\mathbb{R}$ for all $t \in \mathbb{R} \implies \det e^{tX} = 1$ for all $t \in \mathbb{R}$, for all $X \in \mathfrak{sl}_n\mathbb{R}$. But we use the identity

$$\begin{aligned} \det e^{tX} = e^{\mathrm{Tr}\,(tX)} &\implies 1 = e^{\mathrm{Tr}\,(tX)} \\ &\implies \mathrm{Tr}\,(tX) = 0 \\ &\implies \mathrm{Tr}\,(X)t = 0 \implies \mathrm{Tr}\, X = 0 \end{aligned}$$

We now provide an alternative, better proof. We first need a lemma.

Lemma 12.5 ()

$\det'(I) = \mathrm{Tr}$. That is, the differential of the det operator, evaluated at the identity matrix, is equal to the trace. That is, given any matrix T in the vector space of matrices,

Proof.

$$\begin{aligned} \overset{\prime}{\det}(I)(T) = \nabla_T \det(I) & \\ &= \lim_{\varepsilon \to 0} \frac{\det\,(I + \varepsilon T) - \det I}{\varepsilon} \\ &= \lim_{\varepsilon \to 0} \frac{\det\,(I + \varepsilon T) - 1}{\varepsilon} \end{aligned}$$

Clearly, $\det(I + \varepsilon(T)) \in \mathbb{R}[\varepsilon]$, where the constant term of the polynomial approaches 1 and the linear term (coefficient of ε) is $\mathrm{Tr}\, T$. So,

$$\nabla_T \det I = \lim_{\varepsilon \to 0} ... + \mathrm{Tr}\, T = \mathrm{Tr}\, T \tag{563}$$

This means that the instantaneous rate at which det changes at I when traveling in direction T is directly proportional to $\mathrm{Tr}\, T$. Now, we provide an alternative proof of the theorem.

Proof.

Let $R : \mathbb{R} \to \mathrm{SL}(n, \mathbb{R})$ such that $R(0) = I$. Then, by definition, $\mathrm{Im}\, R \subset \mathrm{SL}(n, \mathbb{R}) \implies \det\,(R(t)) = 1$ for all $t \in (-\varepsilon, \varepsilon)$. Compute the derivative of the mapping $\det \circ R$.

$$\begin{aligned} (\det \circ R)(t) = 1 &\implies \overset{\prime}{\det}\big(R(t)\big) \cdot R'(t) \\ &\implies \overset{\prime}{\det}(I) = \overset{\prime}{\det}\big(R(t)\big) = 0 \end{aligned}$$

We now use the previous lemma get that

$$\overset{\prime}{\det}\big(R'(0)\big) = \overset{\prime}{\det}(I) = 0 \implies \mathrm{Tr}\, R'(0) = 0 \tag{564}$$

Theorem 12.12 ()

The matrix representation of $\mathfrak{so}(n)$ is precisely the set of antisymmetric matrices.

Proof.

Let $R : \mathbb{R} \to SO(n)$ be a arbitrary smooth curve in SL(n) such that $R(0) = I$. Then, for all $t \in (-\epsilon, \epsilon)$,

$$R(t)R(t)^T = I \tag{565}$$

Taking the derivative at $t = 0$, we get

$$R'(0)R(0)^T + R(0)R'(0)^T = 0 \implies R'(0) + R'(0)^T = 0 \tag{566}$$

which states that the tangent vector $X = R'(0)$ is skew symmetric. Since the diagonal elements of a skew symmetric matrix are 0, the trace is 0 and the condition that $\det R = 1$ yields nothing new. This shows that $\mathfrak{o}(n) = \mathfrak{so}(n)$.

We have only worked with linear Lie groups so far. The reason that linear Lie groups are so nice to work with is because they have well defined matrix representations. This allows us to have concrete structures on these groups and their Lie algebras.

1. A linear Lie group is concretely defined as a submanifold of $\mathbb{R}^N$, while a general one is an abstract manifold.
2. The Lie bracket with regards to a linear Lie group is defined to be the commutator
$$[A, B] \equiv AB - BA \tag{567}$$
but for elements that are not matrices this doesn't make sense.
3. The exponential map from the algebra to the group is defined
$$e^A \equiv \sum_{k=0}^{\infty} \frac{1}{k!} A^k \tag{568}$$
but if A is not a matrix, then exp cannot be defined this way.

We seek to generalize these concepts to abstract Lie groups, but we will do this in the next section.

12.4.1 Lie Algebras of SO(3) and SU(2), Revisited

Example 12.7 ()

The Lie algebra $\mathfrak{so}(3)$ is the real vector space of 3×3 skew symmetric matrices of form

$$\begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix} \tag{569}$$

where $b, c, d \in \mathbb{R}$. The Lie bracket $[A, B]$ of $\mathfrak{so}(3)$ is also just the usual commutator.
We can define an isomorphism of Lie algebras $\psi : (\mathbb{R}^3, \times) \to \mathfrak{so}(3)$ (where $\times$ is the cross product) by the formula

$$\psi(b, c, d) \equiv \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix} \tag{570}$$

where, by definition,

$$\psi(u \times v) = [\psi(u), \psi(v)] \tag{571}$$

It is also easily verified that for all $u, v \in \mathbb{R}^3$,

$$\psi(u)(v) = u \times v \tag{572}$$

Example 12.8 ()

Similarly, we can see that $\mathfrak{su}(2)$ is the real vector space consisting of all complex 2×2 skew Hermitian matrices of null trace, which is of form

$$i(d\sigma_1 + c\sigma_2 + b\sigma_3) = \begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} \tag{573}$$

where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli spin matrices. We can also define an isomorphism of Lie algebras $\varphi : (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$ by the formula

$$\varphi(b, c, d) = \frac{i}{2}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{2}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} \tag{574}$$

where, by definition of ismorphism, we have

$$\varphi(u \times v) = [\varphi(u), \varphi(v)] \tag{575}$$

We now restate the connection between the groups SO(3) and SU(2). Note that letting $\theta = \sqrt{b^2 + c^2 + d^2}$, we can write

$$A = \frac{1}{\theta}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{\theta}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} \tag{576}$$

such that $A^2 = I$. With this, we can rewrite the exponential map as

$$exp : \mathfrak{su}(2) \to SU(2),\ exp(i\theta A) = \cos\theta I + i \sin\theta A \tag{577}$$

As for the isomorphism $\varphi : (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$, we have

$$\varphi(b, c, d) \equiv \frac{1}{2}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} = i\frac{\theta}{2}A \tag{578}$$

Similarly, we can view the exponential map $exp : (\mathbb{R}^3, \times) \to SU(2)$ as

$$exp(\theta v) = \tag{579}$$

Example 12.9 ()

The lie algebra $\mathfrak{se}(n)$ is the set of all matrices of form

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} \tag{580}$$

where $B \in \mathfrak{so}(n)$ and $U \in \mathbb{R}^n$. The Lie bracket is given by

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix}\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} BC - CB & BV - CU \\ 0 & 0 \end{pmatrix} \tag{581}$$

12.5 Abstract Lie Groups

Definition 12.20 ()

A (real) **Lie group** $\mathcal{G}$ is a group $\mathcal{G}$ that is also a real, finite-dimensional smooth manifold where group multiplication and inversion are smooth maps.

Definition 12.21 ()

A (real) Lie algebra $\mathfrak{g}$ is a real vector space with a map

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{582}$$

called the Lie bracket satisfying bilinearity, antisymmetricity, and the Jacobi Identity.

To every Lie group $\mathcal{G}$ we can associate a Lie algebra $\mathfrak{g}$ whose underlying vector space is the tangent space of $\mathcal{G}$ at the identity element. Additionally, the exponential map allows us to map elements from the Lie algebra to the Lie group. These concrete definitions in the context of linear Lie groups is easy to work with, but has some minor problems: to use it we first need to represent a Lie group as a group of matrices, but not all Lie groups can be represented in this way.

To do this, we must introduce further definitions.

Definition 12.22 ()

Let M_1 (m_1-dimensional) and M_2 (m_2 dimensional) be manifolds in $\mathbb{R}^N$. For any smooth function $f : M_1 \to M_2$ and any $p \in M_1$, the function

$$f'_p : T_p M_1 \to T_{f(p)} M_2 \tag{583}$$

called the **tangent map, derivative, or differential** of f at p, is defined as follows. For every $v \in T_p M_1$ and every smooth curve $\gamma : I \to M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = v$,

$$f'_p(v) \equiv (f \circ \gamma)'(0) \tag{584}$$

The map f'_p is also denoted df_p and is a linear map.

Definition 12.23 ()

Given two Lie groups $\mathcal{G}_1$ and $\mathcal{G}_2$, a **homomorphism of Lie groups** is a function

$$f : \mathcal{G}_1 \to \mathcal{G}_2 \tag{585}$$

that is both a group homomorphism and a smooth map (between manifolds $\mathcal{G}_1$ and $\mathcal{G}_2$). An **isomorphism of Lie groups** is a bijective function f such that both f and f^{-1} are homomorphisms of Lie groups.

Definition 12.24 ()

Given two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, a **homomorphism of Lie algebras** is a function

$$f : \mathfrak{g}_1 \to \mathfrak{g}_2 \tag{586}$$

that is a linear homomorphism that preserves Lie brackets; that is,

$$f([A, B]) = [f(A), f(B)] \tag{587}$$

for all $A, B \in \mathfrak{g}$. An **isomorphism of Lie algebras** is a bijective function $f$such that both f and f^{-1} are homomorphisms of Lie algebras.

Proposition 12.4 ()

If $f : \mathcal{G}_1 \to \mathcal{G}_2$ is a homomorphism of Lie groups, then

$$f'_I : \mathfrak{g}_1 \to \mathfrak{g}_2 \tag{588}$$

is a homomorphism of Lie algebras.

We have explained how to construct the Lie bracket (as the commutator) of the Lie algebra of a linear Lie group, but we have not defined how to construct the Lie bracket for general Lie groups. There are several ways to do this, and we describe one such way through **adjoint representations**.

Definition 12.25 ()

Given a Lie group $\mathcal{G}$, we define a **left translation** as the map

$$L_a : \mathcal{G} \to \mathcal{G},\ L_a(b) \equiv ab \tag{589}$$

for all $b \in \mathcal{G}$. Similarly, the **right translation** is defined

$$R_a : \mathcal{G} \to \mathcal{G},\ R_a(b) \equiv ba \tag{590}$$

for all $b \in \mathcal{G}$.

Both L_a and R_a are diffeomorphisms. Additionally, given the automorphism

$$R_{a^{-1}} L_a \equiv R_{a^{-1}} \circ L_a,\ R_{a^{-1}} L_a(b) \equiv aba^{-1} \tag{591}$$

the derivative

$$(R_{a^{-1}} L_a)'_I : \mathfrak{g} \to \mathfrak{g} \tag{592}$$

is an ismorphism of Lie algebras, also denoted

$$\mathrm{Ad}_a : \mathfrak{g} \to \mathfrak{g} \tag{593}$$

Definition 12.26 ()

This induces another map $a \mapsto \mathrm{Ad}_a$, which is a map of Lie groups

$$Ad : \mathcal{G} \to \mathrm{GL}(\mathfrak{g}) \tag{594}$$

which is called the **adjoint representation of** $\mathcal{G}$. In the case of a linear map, we can verify that

$$\mathrm{Ad}(a)(X) \equiv \mathrm{Ad}_a(X) \equiv aXa^{-1} \tag{595}$$

for all $a \in \mathcal{G}$ and for all $X \in \mathfrak{g}$.

Definition 12.27 ()

Furthermore, the derivative of this map at the identity

$$\mathrm{Ad}'_I : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \tag{596}$$

is a map between Lie algebras, denoted simply as

$$\text{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \tag{597}$$

called the **adjoint representation** of $\mathfrak{g}$. It is easily visualized with the following commutative diagram.

$$\begin{tikzcd} \mathcal{G} \arrow[r, "Ad"] & \mathrm{GL}(\mathfrak{g}) \\ \mathfrak{g} \arrow[u, "exp"] \arrow[r, "ad"] & \mathfrak{gl}(\mathfrak{g}) \arrow[u, "exp"] \end{tikzcd}$$

We define the map ad to be

$$ad(A)(B) \equiv [A, B] \tag{598}$$

where $[A, B]$ is the Lie bracket (of $\mathfrak{g}$) of $A, B \in \mathfrak{g}$. We can actually conclude something stronger about this mapping. Since the Lie bracket of $\mathfrak{g}$ satisfies the properties of the bracket, the Jacobi identity of $[\cdot, \cdot]$ implies that ad is a Lie algebra homomorphism.

$$\begin{aligned} & [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \\ \implies & [x, ad(y)(z)] + [y, ad(z)(x)] + [z, ad(x)(y)] = 0 \\ \implies & ad(x)\big(ad(y)(z)\big) + ad(y)\big(ad(z)(x)\big) + ad(z)\big(ad(x)(y)\big) = 0 \\ \implies & ad(x)ad(y)(z) - ad(y)ad(x)z - ad\big(ad(x)(y)\big)(z) = 0 \\ \implies & \big(ad(x)ad(y) - ad(y)ad(x)\big)(z) = ad\big(ad(x)(y)\big)(z) \\ \implies & [ad(x), ad(y)](z) = ad([x, y])(z) \\ \implies & [ad(x), ad(y)] = ad([x, y]) \end{aligned}$$

Therefore, ad preserves brackets and thus ad is a Lie algebra homomorphism. That is,

$$\text{ad}([A, B]) = [\text{ad}(A), \text{ad}(B)] \tag{599}$$

Note that the bracket on the left side represents the bracket of $\mathfrak{g}$, while the bracket on the right represents the Lie bracket from the Lie algebra $\mathfrak{gl}(\mathfrak{g})$. The fact that ad is a Lie algebra homomorphism indicates that it is a representation of $\mathfrak{g}$, which is why it's called the adjoint representation.

Definition 12.28 ()

This construction finally allows us to define the Lie bracket in the case of a general Lie group. The Lie bracket on $\mathfrak{g}$ is defined as

$$[A, B] \equiv \text{ad}(A)(B) \tag{600}$$

We would also need to introduce a general exponential map for non-linear Lie groups, but we will not do it here.

13 Exercises

13.1 Group Like Structures

Exercise 13.1 (Math 401 Spring 2025 Midterm 2)

Listed.

1. Let G be a finite group with an even number of elements. Show that G contains an element of order 2.
2. Prove that a group of order 10 contains an element of order 5.

Solution 13.1

Listed.

1. We know that $e^{-1} = e$, and so remove it from G. Then G has an odd number of elements. Now as long as G is nonempty, we can remove a, a^{-1}, resulting in an odd cardinality. Since G is finite, this must terminate, and so there must be a case where $a = a^{-1} \implies \mathrm{ord}(a) = 2$.
2. Assume that there is no element of order 5. Then from above it must contain an element of order 2, and let us call it $a \in G$. $\mathrm{ord}(e) = 1$ obviously. If any $b \in G$ had order 10, then $G = Z_{10}$, which would mean that $\mathrm{ord}(b^5) = 2$. Therefore every element other than the identity must have order 2. But then given $a, b, ab \in G$, $ab \neq a, b$ since $ab = a \implies b = e$, and this is precisely the Klein 4 group. This subgroup has an order that doesn't divide 10, contradicting Lagrange's theorem.

Exercise 13.2 (Shifrin 6.1.1)

Which of the following are groups?

(a) $\{1, 3, 7, 9\} \subset \mathbb{Z}_{10}$, with operation multiplication
(b) $\{0, 2, 4, 6\} \subset \mathbb{Z}_{10}$, with operation addition
(c) $\{x \in \mathbb{Q} : 0 < x \leq 1\}$, with operation multiplication
(d) the set of all positive irrational real numbers, with operation multiplication
(e) the set of imaginary numbers $ix, x \in \mathbb{R}$, with operation addition
(f) the set of complex numbers of modulus 1, with operation multiplication
(g) $\mathbb{Z}$ with operation $a \bullet b = a + b + 1$
(h) $\mathbb{Z}$ with operation $a \bullet b = a - b$
(i) $\mathbb{Q} - \{1\}$ with operation $a \bullet b = a + b - ab$

Solution 13.2

Listed. We will denote the sets in question to be G.

(a) Is a group since product of 2 odds is odd, so is closed. Also we have 1 as the identity with $3^{-1} = 7, 7^{-1} = 3, 9^{-1} = 9$. It is associative since multiplication on $\mathbb{Z}_{10}$ is associative.
(b) Not a group since $4 + 4 = 8 \notin G$.
(c) Not a group since $1/2 \in G$ but $(1/2)^{-1} = 2 \notin G$.
(d) Not a group since $\sqrt{2} \times \sqrt{2}2 \notin G$.
(e) Is a group since identity is $0 = 0i$, $ix + iy = i(x + y)$ with $x + y \in \mathbb{R}$, and $-(ix) = i(-x)$ where $-x \in \mathbb{R}$.
(f) Is a group since this is a representation of $O(2)$.
(g) Is a group since this is obviously closed under $\mathbb{Z}$ since $+_{\mathbb{Z}}$ is closed. Now assume that i is the identity. Then $a \bullet i = a + i + 1 = a \implies i = -1$. Therefore $a \bullet a^{-1} = a + a^{-1} + 1 = -1 \implies a^{-1} = -a - 2$. This is associative since

$$(a \bullet b) \bullet c = (a + b + 1) \bullet c = a + b + c + 2 = a \bullet (b + c + 1) = a \bullet (b \bullet c) \tag{601}$$

(h) Not a group since it is not associative. Note $(a \bullet b) \bullet c = (a - b) \bullet c = a - b - c$, while $a \bullet (b \bullet c) = a \bullet (b - c) = a - b + c$.

(i) Is a group. We claim that it is closed. Assume not; given $a, b \neq 1$,

$$a \bullet b = a + b - ab = 1 \implies 0 = ab - a - b + 1 = (a-1)(b-1) \tag{602}$$

which means $a = 1$ or $b = 1$, which is a contradiction. As for the identity, $a \bullet i = a + i - ai = a \implies 0 = i - ai = i(1-a) \implies i = 0$ since $a \neq 1$. We can define the inverse by solving

$$0 = a \bullet a^{-1} = a + a^{-1} - aa^{-1} \implies a^{-1}(1-a) = -a \implies a^{-1} = \frac{a}{a-1} \tag{603}$$

which is well-defined since $a \neq 1$. Finally, it is associative since

$$\begin{aligned}
(a \bullet b) \bullet c &= (a + b - ab) \bullet c && (604)\\
&= a + b - ab + c - ac - bc - abc && (605)\\
&= a + b + c - bc - ab - ac - abc && (606)\\
&= a \bullet (b + c - bc) && (607)\\
&= a \bullet (b \bullet c) && (608)
\end{aligned}$$

Exercise 13.3 (Shifrin 6.1.10)

(a) Let G be a group. Prove that $(ab)^2 = a^2 b^2$ for all $a, b \in G$ if and only if G is abelian.

(b) Prove that if every element (other than the identity element) of a group G has order 2, then G is abelian.

Solution 13.3

For (a), if G is abelian, then

$$(ab)^2 = (ab)(ab) = a(ba)b = a(ab)b = (aa)(bb) = a^2 b^2 \tag{609}$$

If the identity holds, then

$$(ab)^2 = a^2 b^2 \implies (a^{-1}a)(ba)(bb^{-1}) = a^{-1}(ab)(ab)b^{-1} = a^{-1}a^2 b^2 b^{-1} \implies ba = ab \tag{610}$$

For (b), since we have $a^2 = e, b^2 = e$, and $(ab)^2 = e$, from (a) G is abelian.

Exercise 13.4 (Shifrin 6.1.17)

(a) A group has four elements a, b, c, and d, subject to the rules $ca = a$ and $d^2 = a$. Fill in the entire multiplication table at the left below.

$\cdot$	a	b	c	d
a				
b				
c	a			
d				a

(b) A group has six elements a, b, c, d, e, and f, subject to the rules $ae = a$, $bd = a$, $c^2 = a$, and $df = a$. Fill in the entire multiplication table at the right above.

$\cdot$	a	b	c	d	e	f
a					a	
b				a		
c			a			
d						a
e						
f						

Solution 13.4

We can see that $ca = a \implies c = caa^{-1} = aa^{-1} = i$, so c is the identity. We can fill in the row and column of c. Then, we can figure out what bd is. It cannot be b or d since c is the unique identity, so it must be either a or c. It cannot be a since then $bd = a = d^2$, and so $b = d$. So it must be c. By the same logic we can fill out the rest of the rows and columns.

$\cdot$	a	b	c	d
a	c	d	a	b
b	d	a	b	c
c	a	b	c	d
d	b	c	d	a

By the same logic as the previous, we can immediately see that $ae = a \implies e$ is the identity. The formal logic above can be simplified down to saying that there can be no two of the same elements in the same row or column, since if it were, then we are saying that $xy = xz \implies y = z$, which cannot be the case since y and z are distinct. So $fb = a$. We can also deduce that $da = ab$ and $ba = af$. At this point, we can recognize that this is the Dihedral group of order 6, and so we fill in the rest of the multiplication table.

$\cdot$	a	b	c	d	e	f
a	c	f	e	b	a	d
b	d	e	f	a	b	c
c	e	d	a	f	c	b
d	f	c	b	e	d	a
e	a	b	c	d	e	f
f	b	a	d	c	f	e

13.2 Subgroups and Quotient Groups

Exercise 13.5 (Shifrin 6.2.2)

Prove that $\mathbb{Z}_7^\times \cong \mathbb{Z}_6$. (It is crucial to remember that we multiply in $\mathbb{Z}_7^\times$ and add in $\mathbb{Z}_6$.)

Solution 13.5

Both groups are of order 6, and so $\mathbb{Z}_7^\times$—which is indeed a group (since it is the group of units of the ring $(\mathbb{Z}_7, +, \times)$)—must be isomorphic to either $\mathbb{Z}_6$ or S_3. However, S_3 is not abelian, while $\mathbb{Z}_7^\times$ is, so it must be the case that it is isomorphic to $\mathbb{Z}_6$.

Exercise 13.6 (Shifrin 6.2.15.a/b)

The **dihedral group** of order $2n$, denoted $\mathcal{D}_n$, is given by $\{\rho^i \psi^j : 0 \leq i < n, 0 \leq j \leq 1\}$ subject to the rules $\rho^n = e$, $\psi^2 = e$, and $\psi \rho \psi^{-1} = \rho^{-1}$.

1. Check this is really a group. That is, what is $(\rho^i \psi^j)^{-1}$, and what is the product $(\rho^i \psi^j)(\rho^k \psi^\ell)$?

2. Check that $\mathcal{T} \cong \mathcal{D}_3$ and $S_q \cong \mathcal{D}_4$.

Solution 13.6

We check the properties of a group. The following identity is useful:

$$(\psi\rho\psi^{-1})^{n-i} = (\rho^{-1})^{n-i} \implies \psi\rho^{n-i}\psi^{-1} = \rho^i \implies \psi\rho^{n-i} = \rho^i\psi \tag{611}$$

1. *Closure.* From simplifying according to the first two rules, we will automatically adjust the exponents to be $i, k < n$ (by subtracting out multiples of n) and $j \in \{0, 1\}$ (by subtracting out multiples of 2). Going case by case,
   (a) $j = 0, l = 0$. $\rho^i\rho^k = \rho^{i+k}$.
   (b) $j = 0, l = 1$. $\rho^i\rho^k\psi = \rho^{i+k}\psi$.
   (c) $j = 1, l = 0$. $\rho^i\psi\rho^k = \rho^i\rho^{n-k}\psi = \rho^{n-k+i}\psi$.
   (d) $j = 1, l = 1$. $\rho^i\psi\rho^k\psi = \rho^i\psi\psi\rho^{n-k} = \rho^i\rho^{n-k} = \rho^{n-k+i}$.
2. *Identity.* The identity is $e = \rho^0\psi^0$. We can see that $e\rho^i\psi^j = \rho^i\psi^j e = \rho^{i+0}\psi^j$.
3. *Inverse.* We have $\psi\rho\psi^{-1} = \psi\rho\psi = \rho^{-1} \implies \psi\rho = \rho^{-1}\psi^{-1} = (\psi\rho)^{-1}$. Therefore,

$$(\rho^i\psi^j)^{-1} = \begin{cases} \rho^{n-i} & \text{if } j = 0 \\ \rho^i\psi & \text{if } j = 1 \end{cases} \tag{612}$$

   which are both of the correct form and therefore in $\mathcal{D}_n$. To verify, we see that $\rho^i\rho^{n-i} = \rho^n = e$, and $(\rho^i\psi)(\rho^i\psi) = \rho^i\psi\psi\rho^{n-i} = \rho^i\rho^{n-i} = e$.
4. *Associativity.* Can also be proven tediously but problem only asked to state the product and inverse.

For (b) for $\mathcal{T}$, we can explicitly look at the multiplication tables and see that they are isomorphic. We denote r_1, r_2 as the 120 and 240 degree rotations, and f_1, f_2, f_3 as the flips across each axis.

	e	ρ	ρ^2	ψ	$\rho\psi$	$\rho^2\psi$
e	e	ρ	ρ^2	ψ	$\rho\psi$	$\rho^2\psi$
ρ	ρ	ρ^2	e	$\rho^2\psi$	ψ	$\rho\psi$
ρ^2	ρ^2	e	ρ	$\rho\psi$	$\rho^2\psi$	ψ
ψ	ψ	$\rho^2\psi$	$\rho\psi$	e	ρ^2	ρ
$\rho\psi$	$\rho\psi$	ψ	$\rho^2\psi$	ρ	e	ρ^2
$\rho^2\psi$	$\rho^2\psi$	$\rho\psi$	ψ	ρ^2	ρ	e

(a) $\mathcal{D}_3$

	e	r_1	r_2	f_1	f_2	f_3
e	e	r_1	r_2	f_1	f_2	f_3
r_1	r_1	r_2	e	f_3	f_1	f_2
r_2	r_2	e	r_1	f_2	f_3	f_1
f_1	f_1	f_2	f_3	e	r_2	r_1
f_2	f_2	f_3	f_1	r_1	e	r_2
f_3	f_3	f_1	f_2	r_2	r_1	e

(b) $\mathcal{T}$

For S_q, it is tedious to write the full table, so we construct the isrormorphisms using the generators. For S_q, the symmetry group of the square consists of 8 elements: the 4 rotations r_1, r_2, r_3, r_4 (of 90, 180, 270, and 360=0 degrees), and the flips f_1, f_2, f_3, f_4 (across each axis). Now we construct the function $g : \mathcal{D}_3 \to \mathcal{T}$ such that $f(\rho) = r_1$ and $f(\psi) = f_1$. Then we can see that

$$g(\rho^4) = g(e) = e = r_1^4 = g(\rho^4), \qquad g(\psi^2) = g(e) = e = f_1^2 = g(\psi)^2 \tag{613}$$

since 90 degrees rotated 4 times is 0 degrees, the identity, and two flips across the same axis is also the identity. Finally, we have

$$g(\psi\rho\psi) = g(\rho^{-1}) = r_1^{-1} = r_3 = f_1r_1f_1 = g(\psi)g(\rho)g(\psi) \tag{614}$$

Where $r_1^{-1} = r_3$ since a rotation of 270 after a 90 is the same as rotation by 360=0, and $r_3 = f_1r_1f_1$ is the change of basis symmetry observed in Shifrin Example 6.1.5. Therefore the rules match, making it a homomorphism, and since the order is the same ($\mathcal{D}_3$ has $4 \times 2 = 8$ elements from looking at the indices), this is an isomorphism.

Exercise 13.7 (Shifrin 6.3.8)

Let $H \subset G$ be a subgroup, and let $a \in G$ be given. Prove that $aHa^{-1} \subset G$ is a subgroup (called a **conjugate subgroup** of H). Prove, moreover, that it is isomorphic to H (cf. Exercise 6.2.12).

Solution 13.7

Let $x, y \in aHa^{-1}$. Then $x = ah_x a^{-1}, y = ah_y a^{-1}$ for some $h_x, h_y \in H$. Therefore,

1. It is closed. $xy = (ah_x a^{-1})(ah_y a^{-1}) = ah_x(a^{-1}a)h_y a^{-1} = ah_x h_y a^{-1} \in aHa^{-1}$ since $h_x h_y \in H$ by closure.
2. It has an identity since $e \in H \implies aea^{-1} = aa^{-1} = e \in aHa^{-1}$.
3. It has inverses since given $x \in H$ as above with inverses x^{-1}, we see that $(axa^{-1})^{-1} = (a^{-1})^{-1}x^{-1}a^{-1} = ax^{-1}a^{-1} \in aHa^{-1}$ since $x^{-1} \in H$ by H being a group.
4. Associativity is inherited from G.

It suffices to show that this is injective, since the map $\iota : H \to aHa^{-1}$ is surjective by definition. Given $x, y \in aHa^{-1}$ with $x = y$, we have $ah_x a^{-1} = ah_y a^{-1}$, and multiplying by a on the right and then a^{-1} on the left, we get $h_x = h_y$.

Exercise 13.8 (Shifrin 6.3.11)

Prove that a group of order n has a proper subgroup if and only if n is composite.

Solution 13.8

We prove bidirectionally. Call the group G and subgroup H.

1. $(\rightarrow)$. Assume n is prime. Then by Lagrange's theorem $|H|$ must divide n, and so $|H| = 1$ or n, neither of which results in a proper subgroup.
2. $(\leftarrow)$. Assume G has a proper subgroup H. Since it is proper, $|H| \neq 1, n$. Then by Lagrange's theorem, $|H|$ divides n, which implies that n is composite.

Exercise 13.9 (Shifrin 6.3.13)

Suppose $H, K \subset G$ are subgroups of orders 5 and 8, respectively. Prove that $H \cap K = \{e\}$.

Solution 13.9

Let us take an arbitrary element in $x \in H \cap K$ and consider the cyclic group $\langle x \rangle$. By Lagrange's Theorem, the order $|x|$ in H must be either 1 or 5, while the order in K must be $1, 2, 4, 8$. Therefore, $|x| = 1$ and so $x = e$.

Exercise 13.10 (Shifrin 6.3.17)

1. Prove that a group G of even order has an element of order 2. (Hint: If $a \neq e$, a has order 2 if and only if $a = a^{-1}$.)
2. Suppose m is odd, $|G| = 2m$, and G is abelian. Prove G has precisely one element of order 2. (Hint: If there were two, they would provide a Klein four-group.)
3. Prove that if G has exactly one element of order 2, then it must be in the center of G.

Solution 13.10

Listed.

1. Assume the contrary and take $H = G \setminus \{e\}$. Then $|H|$ is odd, and since no element has order 2, every element must be associated with a unique inverse a, a^{-1}. But this cannot happen since $|H|$ is odd. Therefore there must be at least one element of order 2.
2. It has at least 1 element of order 2 from (1). Now assume that there are two, call them a, b. Then $ab \neq a, b$ and ab also has order 2 since $(ab)(ab) = abba = aa = e$. Therefore, calling $c = ab$, we have $ac = ca = aab = b$ and $bc = cb = abb = a$. This fully defines the multiplication table for the Klein 4 group K of order 4. Therefore, by Lagrange's theorem, we have found a subgroup K and so $|K|$ must divide G. However, this would mean that m must be even, a contradiction. Therefore there is only one such unique a.
3. Given $a \in G$ with $|a| = 2$, we wish to show that it is an element of $Z = \{b \in G \mid bx = xb \forall x \in G\}$.[a] Consider $z = x^{-1}ax$. We have
$$z^2 = (x^{-1}ax)^2 = x^{-1}axx^{-1}ax = x^{-1}a^2x = x^{-1}x = e \tag{615}$$
which means that z also has order 2. But since this is unique, it must be that $z = a$. Therefore, by multiplying x on the left, we get
$$x^{-1}ax = a \implies ax = xa \tag{616}$$

[a] I am using the definition of center defined in Shifrin 6.3.7.

Exercise 13.11 (Assigned)

Find all group homomorphisms $\mathbb{Z}_n \to \mathbb{Z}_m$. (Your answer will depend on n and m.)

Solution 13.11

Given a homomorphism, f, we must have $f(0) = 0$. Let $f(1) = k$. Note that the value of $f(1) = k$ completely determines the homomorphism since the image of every other $l \in \mathbb{Z}_n$ is defined by
$$f(l) = f(\underbrace{1 + \ldots + 1}_{l \text{ times}}) = \underbrace{k + \ldots + k}_{l \text{ times}} \tag{617}$$
Since the image of f must be a cyclic subgroup of $\mathbb{Z}_m$, we must satisfy
$$\begin{aligned} 0 = f(0) &= f(\underbrace{1 + \ldots + 1}_{n \text{ times}}) && (618) \\ &= \underbrace{k + \ldots + k}_{n \text{ times}} && (619) \end{aligned}$$
and so $m \mid nk$. Therefore, k must be a multiple of $m/\gcd(n, m)$. So all homomorphisms are determined by the set
$$\left\{ k = \frac{am}{\gcd(n, m)} \;\middle|\; a \in \mathbb{N}, 0 \leq k \leq m - 1 \right\} \tag{620}$$
which we can see ranges from $0 \leq a < \gcd(n, m)$, and so the total number of homomorphisms is $\gcd(n, m)$. Note that there is always the trivial homomorphism when $a = 0$, i.e. everything maps to 0. For example, if we have $f : \mathbb{Z}_{14} \to \mathbb{Z}_{21}$, we have $k = 0, 3, 6, 9, 12, 15, 18$.

13.3 Group Actions

13.4 Product Groups

13.5 Ring Like Structures

Exercise 13.12 (Shifrin 1.2.1)

For each of the following pairs of numbers a and b, find $d = \gcd(a, b)$ and express d in the form $ma + nb$ for suitable integers m and n.

(a) $14, 35$
(b) $56, 77$
(c) $618, 336$
(d) $2873, 6643$
(e) $512, 360$
(f) $4432, 1080$

Solution 13.12

Listed.

1. $d = 7 = (-2) \cdot 14 + (1) \cdot 35.$
2. $d = 7 = (-4) \cdot 56 + 3 \cdot 77.$
3. $d = 6 = -25 \cdot 618 + 46 \cdot 336$
4. $d = 13 = 37 \cdot 2873 + (-16) \cdot 6643.$
5. $d = 8 = 19 \cdot 512 + (-27) \cdot 360.$
6. $d = 8 = 29 \cdot 4432 + (-119) \cdot 1080.$

Exercise 13.13 (Shifrin 1.2.2)

You have at your disposal arbitrarily many 4-cent stamps and 7-cent stamps. What are the postages you can pay? Show in particular that you can pay all postages greater than 17 cents.

Exercise 13.14 (Shifrin 1.2.3)

Prove that whenever $m \neq 0$, $\gcd(0, m) = |m|$.

Exercise 13.15 (Shifrin 1.2.4)

(a) Prove that if $a|x$ and $b|y$, then $ab|xy$.
(b) Prove that if $d = \gcd(a, b)$, then $\gcd(\frac{a}{d}, \frac{b}{d}) = 1$.

Exercise 13.16 (Shifrin 1.2.5)

Prove or give a counterexample: the integers q and r guaranteed by the division algorithm, Theorem 2.2, are unique.

Exercise 13.17 (Shifrin 1.2.6)

Prove or give a counterexample. Let $a, b \in \mathbb{Z}$. If there are integers m and n so that $d = am + bn$, then $d = \gcd(a, b)$.

Exercise 13.18 (Shifrin 1.2.7)

Generalize Proposition 2.5: if $\gcd(m, c) = 1$ and $m|cz$, then prove $m|z$.

Solution 13.13

Let $\gcd(m, c) = 1$ and $m|cz$. Then there exists $a, b \in \mathbb{Z}$ such that $am + bc = 1$. Multiply both sides of the equation by z to get by the distributive property

$$(am + bc)z = amz + bcz = z \tag{621}$$

$m|amz$ and $m|cz \implies m|bcz$. Therefore, the sum of the two, which is equal to z, must be divisible by m. Therefore $m|z$.

Exercise 13.19 (Shifrin 1.2.8)

Suppose $a, b, n \in \mathbb{N}$, $\gcd(a, n) = 1$, and $\gcd(b, n) = 1$. Prove or give a counterexample: $\gcd(ab, n) = 1$.

Exercise 13.20 (Shifrin 1.2.9)

Prove that if p is prime and $p|(a_1 a_2 \ldots a_n)$, then $p|a_j$ for some j, $1 \leq j \leq n$. (Hint: Use Proposition 2.5 and induction.)

Exercise 13.21 (Shifrin 1.2.10)

Given a positive integer n, find n consecutive composite numbers.

Exercise 13.22 (Shifrin 1.2.11)

Prove that there are no integers m, n so that $(\frac{m}{n})^2 = 2$. (Hint: You may start by assuming m and n are relatively prime. Why? Then use Exercise 1.1.3.)

Exercise 13.23 (Shifrin 1.2.12)

Find all rectangles whose sides have integral lengths and whose area and perimeter are equal.

Exercise 13.24 (Shifrin 1.2.13)

Given two nonzero integers a, b, in analogy with the definition of $\gcd(a, b)$, we define the **least common multiple** $\mathrm{lcm}(a, b)$ to be the positive number μ with the properties:

(i) $a|\mu$ and $b|\mu$, and
(ii) if $s \in \mathbb{Z}$, $a|s$ and $b|s \Rightarrow \mu|s$.

Prove that

(a) if $\gcd(a, b) = 1$, then $\mu = ab$. (Hint: If $\gcd(a, b) = 1$, then there are integers m and n so that $1 = ma + nb$; therefore, $s = mas + nbs$.)
(b) more generally, if $\gcd(a, b) = d$, then $\mu = ab/d$.

Solution 13.14

Listed.

1. We can simply verify the two properties. Since $\mu = ab$, $a|\mu$ and $b|\mu$ trivially by the existence of b and a, respectively. As for the second property, let $s \in \mathbb{Z}$ exist such that $a|s$ and $b|s$. Since $a|s$, $s = xa$ for some $x \in \mathbb{Z}$. But since $b|s$, $b|xa$. Since $\gcd(a,b) = 1$ by assumption, the result in [Shifrin 1.2.7] tells us that $b|x$, i.e. there exists some $k \in \mathbb{Z}$ such that $x = kb$. Therefore $s = xa = kba = kab = k\mu$. By existence of k, $\mu|s$, and we are done.
2. Given a, b with $\gcd(a,b) = d$, there exists some $a', b' \in \mathbb{Z}$ s.t. $a = da', b = db'$. We claim that $\mu = ab/d := da'b'$ is the lcm.[a] It is clear that $a|\mu$ and $b|\mu$ by the existence of integers b' and a', respectively. To prove the second property, let $s \in \mathbb{Z}$ with $a|s$ and $b|s$. Since $a|s \iff da'|s$, there must exist some $x \in \mathbb{Z}$ s.t. $s = da'x$. But since $b|s$, this means that $db'|s \iff db'|da'x \iff b'|a'x$. But $\gcd(a', b') = 1$ which follows from the definition of gcd, and so by [Shifrin 1.2.7] it must be the case that $b'|x$, i.e. there exists some $k \in \mathbb{Z}$ s.t. $x = b'k$. Substituting this back we have $s = da'b'k = \mu k$, and by existence of k it follows that $\mu|s$. Since it satisfies these 2 properties μ is the lcm.

[a] Since division isn't generally closed in the integers, I prefer to define ab/d this way.

Exercise 13.25 (Shifrin 1.2.14)

See Exercise 13 for the definition of $\text{lcm}(a,b)$. Given prime factorizations $a = p_1^{\mu_1} \cdots p_m^{\mu_m}$ and $b = p_1^{\nu_1} \cdots p_m^{\nu_m}$, with $\mu_i, \nu_i \geq 0$, express $\gcd(a,b)$ and $\text{lcm}(a,b)$ in terms of $p_1, \ldots, p_m$. Prove that your answers are correct.

Exercise 13.26 (Shifrin 1.3.8)

We see that in mod10,

$$3^{400} \equiv 9^{200} \equiv (-1)^{200} \equiv 1^{100} \equiv 1 \tag{622}$$

so the last digit is 1. To get the last 2 digits, we use the binomial expansion and focus on the last 2 terms.

$$3^{400} = 9^{200} = (10-1)^{200} = \ldots + \binom{200}{199} 10^1 (-1)^{199} + \binom{200}{200} (-1)^{200} \tag{623}$$

since every combination of the form $\binom{n}{k}$ is an integer and all the other terms have a factor of 10^2, the expansion mod100 becomes

$$3^{400} \equiv \binom{200}{199} 10^1 (-1)^{199} + \binom{200}{200} (-1)^{200} = 200 \cdot 10 \cdot (-1)^{199} + 1 \equiv 1 \pmod{100} \tag{624}$$

and so the last two digits is 01. To get the last digit of 7^{99}, we see that in mod10,

$$7^{99} \equiv 7^{96} \cdot 7^3 \equiv (7^4)^{24} \cdot 343 \equiv 2401^{24} \cdot 343 \equiv 1^{24} \cdot 3 \equiv 3 \tag{625}$$

Exercise 13.27 (Shifrin 1.3.10)

We must show that

$$n \equiv 0 \pmod{13} \iff n' = \sum_{i=1}^{k} a_i 10^{i-1} + 4a_0 \equiv 0 \pmod{13} \tag{626}$$

We see that $n \equiv n + 39a_0 \equiv 0 \pmod{13}$, and

$$n + 39a_0 = \sum_{i=0}^{k} 10^i a_i + 39a_0 \tag{627}$$
$$= \sum_{i=1}^{k} 10^i a_i + 40a_0 \tag{628}$$
$$= 10\bigg(\sum_{i=1}^{k} 10^{i-1} a_i + 4a_0\bigg) \tag{629}$$
$$= 10n' \tag{630}$$

and so we have $n \equiv 10n' \pmod{13}$, and so $n' \equiv 0 \pmod{13} \implies n \equiv 0 \pmod{13}$. Conversely, if $n \equiv 0 \pmod{13}$, then $4n \equiv 0 \pmod{13}$, but $4n \equiv 40n'$ and so $n' \equiv 40n' \equiv 4n \equiv 0 \pmod{13}$. Therefore both implications are proven.

Exercise 13.28 (Shifrin 1.3.12)

Suppose that p is prime. Prove that if $a^2 \equiv b^2 \pmod{p}$, then $a \equiv b \pmod{p}$ or $a \equiv -b \pmod{p}$.

Solution 13.15

We have

$$a^2 \equiv b^2 \pmod{p} \implies a^2 - b^2 \equiv 0 \pmod{p} \tag{631}$$
$$\implies (a+b)(a-b) \equiv 0 \pmod{p} \tag{632}$$

We claim that there are no zero divisors in $\mathbb{Z}_p$. If $mn \equiv 0 \pmod{p}$, then by definition this means $p|mn$, which implies that in the integers this must mean that $p|m$ or $p|n$.[a] But since $m, n \not\equiv 0$, $p \nmid n$ and $p \nmid m$, arriving at a contradiction. Going back to our main argument, it must be the case that $a + b \equiv 0 \implies a \equiv -b$ or $a - b \equiv 0 \implies a \equiv b$.

[a] Proposition 2.5

Exercise 13.29 (Shifrin 1.3.15)

Let us assume that $n = a^2 + b^2 + c^2$ for some $a, b, c \in \mathbb{Z}$. Let us consider for each integer z, all the possible values of $z^2 \pmod{8}$.

$$z \equiv 0 \implies z^2 \equiv 0 \pmod{8} \tag{633}$$
$$z \equiv 1 \implies z^2 \equiv 1 \pmod{8} \tag{634}$$
$$z \equiv 2 \implies z^2 \equiv 4 \pmod{8} \tag{635}$$
$$z \equiv 3 \implies z^2 \equiv 1 \pmod{8} \tag{636}$$
$$z \equiv 4 \implies z^2 \equiv 0 \pmod{8} \tag{637}$$
$$z \equiv 5 \implies z^2 \equiv 1 \pmod{8} \tag{638}$$
$$z \equiv 6 \implies z^2 \equiv 4 \pmod{8} \tag{639}$$
$$z \equiv 7 \implies z^2 \equiv 1 \pmod{8} \tag{640}$$

Therefore, $a^2 + b^2 + c^2 \pmod{8}$ can take any values of the form

$$x + y + z \pmod{8} \text{ for } x, y, z \in \{0, 1, 4\} \tag{641}$$

Since addition is commutative, WLOG let $x \leq y \leq z$. We can just brute force search this.

1. If $z = 0$, then $x = y = z = 0$ and $x + y + z = 0 \not\equiv 7$.
2. If $z = 1$, then we see

$$
\begin{aligned}
0 + 0 + 1 &\equiv 1 && (642)\\
0 + 1 + 1 &\equiv 2 && (643)\\
1 + 0 + 1 &\equiv 2 && (644)\\
1 + 1 + 1 &\equiv 3 && (645)
\end{aligned}
$$

3. If $z = 4$, then we see that

$$
\begin{aligned}
0 + 0 + 4 &\equiv 4 && (646)\\
0 + 1 + 4 &\equiv 5 && (647)\\
0 + 4 + 4 &\equiv 0 && (648)\\
1 + 1 + 4 &\equiv 6 && (649)\\
1 + 4 + 4 &\equiv 1 && (650)\\
4 + 4 + 4 &\equiv 4 && (651)
\end{aligned}
$$

And so $a^2 + b^2 + c^2 \not\equiv 7 \pmod 8$ for any $a, b, c \in \mathbb{Z}$.

Exercise 13.30 (Shifrin 1.3.20.a/b/g)

For (a),

$$3x \equiv 2 \pmod 5 \implies 6x \equiv 4 \pmod 5 \implies x \equiv 4 \pmod 5 \tag{652}$$

For (b),

$$
\begin{aligned}
6x + 3 \equiv 1 \pmod{10} &\implies 6x \equiv -2 \equiv 8 \pmod{10} && (653)\\
&\implies 10 | (6x - 8) && (654)\\
&\implies 5 | (3x - 4) && (655)\\
&\implies 3x \equiv 4 \pmod 5 && (656)\\
&\implies 3x \equiv 9 \pmod 5 && (657)\\
&\implies x \equiv 3 \pmod 5 && (658)
\end{aligned}
$$

For (g),

$$
\begin{aligned}
15x \equiv 25 \pmod{35} &\implies 35 | (15x - 25) && (659)\\
&\implies 7 | (3x - 5) && (660)\\
&\implies 3x \equiv 5 \pmod 7 && (661)\\
&\implies 3x \equiv 12 \pmod 7 && (662)\\
&\implies x \equiv 4 \pmod 7 && (663)
\end{aligned}
$$

Exercise 13.31 (Shifrin 1.3.21.b/c)

For (b), we see that 4 and 13 are coprime with $-3 \cdot 4 + 1 \cdot 13 = 1$. Therefore, by the Chinese remainder theorem

$$x \equiv 1 \cdot 1 \cdot 12 + (-3) \cdot 7 \cdot 4 \pmod{52} \implies x \equiv 33 \pmod{52} \tag{664}$$

For (c), we solve the first two congruences $x \equiv 3 \pmod 4$ and $x \equiv 4 \pmod 5$. 4 and 5 are coprime with $-1 \cdot 4 + 1 \cdot 5 = 1$. Therefore, by CRT

$$x \equiv -1 \cdot 4 \cdot 4 + 1 \cdot 5 \cdot 3 \pmod{20} \implies x \equiv -1 \pmod{20} \tag{665}$$

Then we solve $x \equiv -1 \pmod{20}$ with the final congruence $x \equiv 3 \pmod 7$. We see that 20 and 7 are coprime with $-1 \cdot 20 + 3 \cdot 7 = 1$. Therefore by CRT

$$x \equiv -1 \cdot 20 \cdot 3 + 3 \cdot 7 \cdot -1 \pmod{140} \implies x \equiv 59 \pmod{140} \tag{666}$$

Exercise 13.32 (Shifrin 1.3.25)

We prove bidirectionally.

1. Assume a solution exists for $cx \equiv b \pmod m$. Then $m|(cx - b)$, which means that there exists a $y \in \mathbb{Z}$ s.t. $my = cx - b \iff b = cx - my$. Since $d = \gcd(c, m)$, there exists $c', m' \in \mathbb{Z}$ s.t. $c = dc'$ and $m = dm'$. So

$$b = cx - my = d(c'x - m'y) \implies d|b \tag{667}$$

2. Assume that $d|b$. Then there exists a $b' \in \mathbb{Z}$ s.t. $b = db'$, and we have

$$\begin{aligned} cx \equiv b \pmod m &\iff m|(cx - b) && (668)\\ &\iff dm'|d(c'x - b') && (669)\\ &\iff m'|(c'x - b') && (670)\\ &\iff c'x \equiv b' \pmod{m'} && (671) \end{aligned}$$

   Since $\gcd(c', m') = 1$[a], by Shifrin Proposition 3.5 the equation $c'x \equiv b' \pmod{m'}$ is guaranteed to have a solution, and working backwards in the iff statements gives us the solution for $cx \equiv b \pmod m$.

We have proved existence of a solution in $\text{mod}(m/d) = m'$. Now we show uniqueness. Assume that there are two solutions $x \equiv \alpha$, $x \equiv \beta \pmod{m'}$ with $\alpha \not\equiv \beta \pmod{m'}$. Then, x can be written as $x = k_\alpha m' + \alpha$ and $x = k_\beta m' + \beta$. But we see that

$$\begin{aligned} 0 = x - x &= (k_\alpha m' + \alpha) - (k_\beta m' + \beta) && (672)\\ &= m'(k_\alpha - k_\beta) + (\alpha - \beta) && (673)\\ &\equiv \alpha - \beta \pmod{m'} && (674) \end{aligned}$$

which implies that $\alpha \equiv \beta \pmod{m'}$, contradicting our assumption that they are different in modulo. Therefore the solution must be unique.

[a] Since $\gcd(c, m) = d \implies$ that there exists a $y, z \in \mathbb{Z}$ s.t. $cy + mz = d$, and dividing both sides by d guarantees the existence of y, z satisfying $c'y + m'z = 1$, meaning that $\gcd(c', m') = 1$.

Exercise 13.33 (Shifrin 1.4.1)

For $\mathbb{Z}_7$. There are no zero divisors and the units are all elements.

$\times$	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

(675)

For $\mathbb{Z}_8$. The zero divisors are $2, 4, 6$. The units are $1, 3, 5, 7$.

$\times$	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	0	2	4	6
3	0	3	6	1	4	7	2	5
4	0	4	0	4	0	4	0	4
5	0	5	2	7	4	1	6	3
6	0	6	4	2	0	6	4	2
7	0	7	6	5	4	3	2	1

(676)

For $\mathbb{Z}_{12}$. The zero divisors are $2, 3, 4, 6, 8, 9, 10$. The units are $1, 5, 7, 11$.

$\times$	0	1	2	3	4	5	6	7	8	9	10	11
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7	8	9	10	11
2	0	2	4	6	8	10	0	2	4	6	8	10
3	0	3	6	9	0	3	6	9	0	3	6	9
4	0	4	8	0	4	8	0	4	8	0	4	8
5	0	5	10	3	8	1	6	11	4	9	2	7
6	0	6	0	6	0	6	0	6	0	6	0	6
7	0	7	2	9	4	11	6	1	8	3	10	5
8	0	8	4	0	8	4	0	8	4	0	8	4
9	0	9	6	3	0	9	6	3	0	9	6	3
10	0	10	8	6	4	2	0	10	8	6	4	2
11	0	11	10	9	8	7	6	5	4	3	2	1

(677)

Exercise 13.34 (Shifrin 1.4.5.a/b/c)

1. Prove that $\gcd(a, m) = 1 \iff \bar{a} \in \mathbb{Z}_m$ is a unit.
2. Prove that if $\bar{a} \in \mathbb{Z}_m$ is a zero-divisor, then $\gcd(a, m) > 1$, and conversely, provided $m \nmid a$.
3. Prove that every nonzero element of $\mathbb{Z}_m$ is either a unit or a zero-divisor.
4. Prove that in any commutative ring R, a zero-divisor cannot be a unit, and a unit cannot be a zero-divisor. Do you think c. holds in general?

Solution 13.16

For (a),

1. $(\rightarrow)$. If $\gcd(a, m) = 1$, then there exists $x, y \in \mathbb{Z}$ such that $ax + my = 1$. Taking the modulo

on both sides gives $ax \equiv 1 \pmod{m}$, and therefore we have established the existence of $x \in \mathbb{Z}$, which implies the existence of $\bar{x} \in \mathbb{Z}_m$.

2. $(\leftarrow)$. If we have $a \in \mathbb{Z}$ and $\bar{a}$ is a unit, then there exists a $\bar{x} \in \mathbb{Z}_m$ s.t. $\bar{a}\bar{x} = \bar{1} \iff ax \equiv 1 \pmod{m}$, which means that $m|(1-ax)$. So there exists an integer $y \in \mathbb{Z}$ s.t. $my = 1 - ax \iff ax + my = 1$. By Shifrin corollary 2.4 a, m must be coprime.

For (b),

1. $(\rightarrow)$ Let $\bar{a} \in \mathbb{Z}_m$ be a zero-divisor. Then there exists $\bar{x} \neq \bar{0}$ in $\mathbb{Z}_m$ such that $\bar{a}\bar{x} = \bar{0}$. This means: $ax \equiv 0 \pmod{m}$, so $m \mid ax$, and $m \nmid x$ (since $\bar{x} \neq \bar{0}$). Since $m \mid ax$ but $m \nmid x$, some prime factor of m must divide a. This prime factor is then a common divisor of a and m greater than 1, so $\gcd(a, m) > 1$.
2. $(\leftarrow)$ Let $a \in \mathbb{Z}$, $m \in \mathbb{N}$ where $\gcd(a, m) = d > 1$ and $m \nmid a$. Then $a = a'd$ and $m = m'd$ for some $a', m' \in \mathbb{Z}$. Therefore,

$$\bar{a}\bar{m'} = \overline{am'} = \overline{a'dm'} = \overline{a'm} = \bar{0} \tag{678}$$

Also since $m \nmid a$, we have $\bar{a} \neq \bar{0}$, and since $m = m'd$, we have $m \nmid m'$ (since $m \nmid a \implies d \neq m$), so $\bar{m'} \neq \bar{0}$. Therefore $\bar{a}$ is a zero-divisor in $\mathbb{Z}_m$.

For (c), let $a \in \mathbb{Z}_m$ be a nonzero element. Then it must be the case that $\gcd(a, m) = 1$ or $\gcd(a, m) > 1$. In the former case, a is a unit by (a), and in the latter case, $a \not\equiv 0 \implies m \nmid a$[a], and so by (b) a is a zero divisor.

[a] By contrapositive $m \mid a \implies a \equiv 0 \pmod{m}$ is trivial.

Exercise 13.35 (Shifrin 1.4.6.b/c/d)

Prove that in any ring R:

1. $0 \cdot a = 0$ for all $a \in R$ (cf. Lemma 1.1);
2. $(-1)a = -a$ for all $a \in R$ (cf. Lemma 1.2);
3. $(-a)(-b) = ab$ for all $a, b \in R$;
4. the multiplicative identity $1 \in R$ is unique.

Solution 13.17

For (a), note that $0a = (0 + 0) \cdot a = 0a + 0a$ and by subtracting $0a$ from both sides, we have $0 = 0a$. Similarly, $a0 = a(0 + 0) = a0 + a0 \implies 0 = a0$. For (b),

$$\begin{aligned} a + (-1) \cdot a &= 1 \cdot a + (-1) \cdot a && \text{(definition of 1)} \\ &= (1 + -1) \cdot a && \text{(left distributivity)} \\ &= 0 \cdot a && \text{(definition of add inverse)} \\ &= 0 && \text{(From (a))} \end{aligned}$$

For (c), note that by right distributivity,

$$\begin{aligned} (-1) \cdot a + a &= (-1) \cdot a + 1 \cdot a && \text{(definition of 1)} \\ &= (-1 + 1) \cdot a && \text{(right distributivity)} \\ &= a \cdot 0 && \text{(definition of add inverse)} \\ &= 0 && \text{(From (a))} \end{aligned}$$

Therefore,

$$\begin{aligned}
(-a)(-b) &= (-1 \cdot a)(-1 \cdot b) && \text{(from (b))} \\
&= -1 \cdot (a \cdot -1) \cdot b && \text{(associativity)} \\
&= -1 \cdot -a \cdot b && \text{(from (b))} \\
&= -1 \cdot -1 \cdot a \cdot b && \text{(from (b))} \\
&= (-1 \cdot -1) \cdot ab && \text{(associativity)} \\
&= 1ab && \text{(shown below)} \\
&= ab && \text{(definition of identity)}
\end{aligned}$$

where $(-1)(-1) = 1$ since by (b), $(-1)(-1) = -(-1)$. We know that $-(-1)$ is an additive inverse for -1 and so is 1. Since the multiplicative identity is unique in a ring, $-(-1) = 1$. We show uniqueness for (d). Let us have $1 \neq 1'$. Then by definition of identity,

$$1 = 11' = 1'1 = 1' \tag{679}$$

which is a contradiction.

Exercise 13.36 (Shifrin 1.4.10)

1. Prove that the multiplicative inverse of a unit a in a ring R is unique. That is, if $ab = ba = 1$ and $ac = ca = 1$, then $b = c$. (You will need to use associativity of multiplication in R.)
2. Indeed, more is true. If $a \in R$ and there exist $b, c \in R$ so that $ab = 1$ and $ca = 1$, prove that $b = c$ and thus that a is a unit.

Solution 13.18

For (a), we see that

$$c = 1c = (ab)c = (ba)c = b(ac) = b(ca) = b1 = b \tag{680}$$

For (b), we have

$$b = 1b = (ca)b = c(ab) = c1 = c \tag{681}$$

Exercise 13.37 (Shifrin 1.4.13)

Let p be a prime number. Use the fact that $\mathbb{Z}_p$ is a field to prove that $(p-1)! \equiv -1 \pmod{p}$. (Hint: Pair elements of $\mathbb{Z}_p$ with their multiplicative inverses; cf. Exercise 1.3.12.).

Solution 13.19

For $p = 2$, the result is trivial. Now let $p > 2$ be a prime. Then since $\mathbb{F}$ is a field, every element $a \in \mathbb{F}$ contains a multiplicative inverse a^{-1}. We claim that the only values for which $a = a^{-1}$ is $1, p-1$. Assume that $a = a^{-1}$. Then

$$a^2 = 1 \implies p|(a^2 - 1) \implies p|(a+1)(a-1) \tag{682}$$

and since p is prime, it must be the case that $p|a+1 \iff a \equiv -1 \pmod{p}$ or $p|a-1 \iff a \equiv 1 \pmod{p}$. Therefore, we are left to consider the $(p-3)$ elements: $2, \ldots, p-2$. Since inverses are unique and the inverses of inverses is the original element, we can partition these $p-2$ elements into $(p-3)/2$ pairs.[a] Let's call the set of pairs $K = \{(a, b)\}$ where $b = a^{-1}$. Therefore, by commutativity and associativity we have

$$(p-1)! \equiv (1)(p-1) \prod_{(a,b)\in K} ab \equiv -1 \cdot \prod_{(a,b)\in K} 1 \equiv -1 \pmod p. \tag{683}$$

[a] Since $p \neq 2$, p is odd and therefore $p-3$ is even.

Exercise 13.38 (Shifrin 2.3.2.a/b/c)

Recall that the conjugate of the complex number $z = a + bi$ is defined to be $\bar{z} = a - bi$. Prove the following properties of the conjugate:

1. $\overline{z+w} = \bar{z} + \bar{w}$
2. $\overline{zw} = \bar{z}\bar{w}$
3. $\bar{z} = z \iff z \in \mathbb{R}$ and $\bar{z} = -z \iff iz \in \mathbb{R}$
4. If $z = r(\cos\theta + i\sin\theta)$, then $\bar{z} = r(\cos\theta - i\sin\theta)$

Solution 13.20

Let $z = a + bi, w = c + di$. For (a),

$$\overline{z+w} = \overline{(a+c)+(b+d)i} = (a+c) - (b+d)i = a + c - bi - di = (a - bi) + (c - di) = \overline{z} + \overline{w} \tag{684}$$

For (b),

$$\overline{zw} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd)-(ad+bc)i = ac-bd-adi-bci = (a-bi)(c-di) = \bar{z}\bar{w} \tag{685}$$

For (c), consider

$$\begin{aligned} \overline{z} = z &\iff a + bi = a - bi && (686)\\ &\iff bi = -bi && (687)\\ &\iff 2bi = 0 && (688)\\ &\iff b = 0 && \text{(field has no 0 divisors)} \end{aligned}$$

Therefore, $z = a \in \mathbb{R}$.

$$\begin{aligned} \overline{z} = -z &\iff a - bi = -a - bi && (689)\\ &\iff a = -a && (690)\\ &\iff 2a = 0 && (691)\\ &\iff a = 0 && \text{(field has no 0 divisors.)} \end{aligned}$$

Therefore, $z = bi \implies iz = -b \in \mathbb{R}$.

Exercise 13.39 (Shifrin 2.3.3.a/b/c)

Recall that the modulus of the complex number $z = a+bi$ is defined to be $|z| = \sqrt{a^2+b^2}$. Prove the following properties of the modulus:

1. $|zw| = |z||w|$
2. $|\bar{z}| = |z|$
3. $|z|^2 = z\bar{z}$
4. $|z+w| \leq |z| + |w|$ (This is called the triangle inequality; why?)

Solution 13.21

Let $z = a + bi$ and $w = c + di$. For (a),

$$\begin{aligned} |zw| &= |(ac - bd) + (ad + bc)i| \\ &= \sqrt{(ac - bd)^2 + (ad + bc)^2} \\ &= \sqrt{a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2} \\ &= \sqrt{(a^2 + b^2)(c^2 + d^2)} \\ &= \sqrt{a^2 + b^2}\sqrt{c^2 + d^2} \\ &= |z||w| \end{aligned}$$

For (b), if $z = a + bi$, then $\bar{z} = a - bi$, so:

$$|\bar{z}| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z| \tag{692}$$

For (c),

$$\begin{aligned} z\bar{z} &= (a + bi)(a - bi) \\ &= a^2 + b^2 \\ &= |z|^2 \end{aligned}$$

Exercise 13.40 (Shifrin 3.1.2.c/d)

Find the greatest common divisors $d(x)$ of the following polynomials $f(x), g(x) \in F[x]$, and express $d(x)$ as $s(x)f(x) + t(x)g(x)$ for appropriate $s(x), t(x) \in F[x]$:

1. $f(x) = x^3 - 1$, $g(x) = x^4 + x^3 - x^2 - 2x - 2$, $F = \mathbb{Q}$
2. $f(x) = x^2 + (1 - \sqrt{2})x - \sqrt{2}$, $g(x) = x^2 - 2$, $F = \mathbb{R}$
3. $f(x) = x^2 + 1$, $g(x) = x^2 - i + 2$, $F = \mathbb{C}$
4. $f(x) = x^2 + 2x + 2$, $g(x) = x^2 + 1$, $F = \mathbb{Q}$
5. $f(x) = x^2 + 2x + 2$, $g(x) = x^2 + 1$, $F = \mathbb{C}$

Solution 13.22

For (c), the gcd is 1, with

$$-\frac{1}{1-i}(x^2 + 1) + \frac{1}{1-i}(x^2 - i + 2) = \frac{1}{1-i}(x^2 - i + 2 - x^2 - 1) = \frac{1}{1-i}(1 - i) = 1 \tag{693}$$

where $1/(1 - i) = (1 + i)/2$. For (d), the gcd is 1, with

$$\begin{aligned} &\frac{1}{5}(2x + 3)(x^2 + 1) + \frac{1}{5}(1 - 2x)(x^2 + 2x + 2) && (694) \\ &\quad = \frac{1}{5}(2x^3 + 3x^2 + 2x + 3) + \frac{1}{5}(-2x^3 - 3x^2 - 2x + 2) = 1 && (695) \end{aligned}$$

Exercise 13.41 (Shifrin 3.1.6)

Prove that if F is a field, $f(x) \in F[x]$, and $\deg(f(x)) = n$, then $f(x)$ has at most n roots in F.

Solution 13.23

We start when $n = 1$. Then $f(x) = mx + b$ and we claim that the only root is $x = -b/m$ since we can solve for $0 = mx + b$ with the field operations, which leads to a unique solution. This implies by corr 1.5 that $(x + b/m)$ is the only factor of f. Now suppose this holds true for some degree $n - 1$ and let us have a degree n polynomial f. Assume that some c is a root of f (if there exists no c, then we are trivially done), which means $(x - c)$ is a factor of f, and we can write

$$f(x) = (x - c)\, g(x) \tag{696}$$

for some polynomial $g(x)$ of degree $n-1$. By our inductive hypothesis, $g(x)$ must have at most $n-1$ roots, and so f has at most n roots.

Exercise 13.42 (Shifrin 3.1.8)

Let F be a field. Prove that if $f(x) \in F[x]$ is a polynomial of degree 2 or 3, then $f(x)$ is irreducible in $F[x]$ if and only if $f(x)$ has no root in F.

Solution 13.24

We prove bidirectionally.

1. $(\rightarrow)$. Let f be irreducible. Then it cannot be factored into polynomials $p(x)q(x)$ where $\deg(p) + \deg(q) = n$. Note that two positive integers adding up to 2 or 3 means that at least one of the integers must be 1, by the pigeonhole principle. This means that f irreducible is equivalent to saying that f does not have linear factors of form $(x - c)$, which by corollary 1.5 implies that there exists no root c for $f(x)$.
2. $(\leftarrow)$. Let f have no root in F. Then by corollary 1.5 there exists no linear factors $(x - c)$. By the same pigeonhole principle argument, we know that having a linear factor for degree 2 or 3 polynomials is equivalent to having (general) factors, and so f has no factors. Therefore f is irreducible.

Exercise 13.43 (Shifrin 3.1.13)

List all the irreducible polynomials in $\mathbb{Z}_2[x]$ of degree ≤ 4. Factor $f(x) = x^7 + 1$ as a product of irreducible polynomials in $\mathbb{Z}_2[x]$.

Solution 13.25

Listed by degree.

1. 1: $x, x + 1$.
2. 2: $x^2 + x + 1$.
3. 3: $x^3 + x^2 + 1, x^3 + x + 1$.
4. 4: $x^4 + x + 1, x^4 + x^3 + 1, x^4 + x^3 + x^2 + x + 1$.

We have

$$\begin{aligned} x^7 + 1 &= (x + 1)(x^6 + x^5 + x^4 + x^3 + x^2 + x + 1) && (697) \\ &= (x + 1)(x^3 + x + 1)(x^3 + x^2 + 1) && (698) \end{aligned}$$

Exercise 13.44 (Shifrin 3.2.2.b/c)

Prove that

1. $\mathbb{Q}[\sqrt{2}, i] = \mathbb{Q}[\sqrt{2} + i]$, but $\mathbb{Q}[\sqrt{2}i] \subsetneq \mathbb{Q}[\sqrt{2}, i]$

2. $\mathbb{Q}[\sqrt{2}, \sqrt{3}] = \mathbb{Q}[\sqrt{2} + \sqrt{3}]$, but $\mathbb{Q}[\sqrt{6}] \subsetneq \mathbb{Q}[\sqrt{2}, \sqrt{3}]$
3. $\mathbb{Q}[\sqrt[3]{2} + i] = \mathbb{Q}[\sqrt[3]{2}, i]$; what about $\mathbb{Q}[\sqrt[3]{2}i] \subset \mathbb{Q}[\sqrt[3]{2}, i]$?

Solution 13.26

From Shifrin, I use the fact that $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$, and the same proof immediately shows that $\mathbb{Q}[\sqrt{3}] = \{a + b\sqrt{3} \mid a, b \in \mathbb{Q}\}$ along with that for $\mathbb{Q}[\sqrt{6}]$. As for $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$, I also follow the same logic to show

$$\begin{aligned} \mathbb{Q}[\sqrt{2}, \sqrt{3}] &= \mathbb{Q}[\sqrt{2}][\sqrt{3}] && (699) \\ &= \{\alpha + \beta\sqrt{3} \mid a, b \in \mathbb{Q}[\sqrt{2}]\} && (700) \\ &= \{(a + b\sqrt{2}) + (c + d\sqrt{2})\sqrt{3} \mid a, b, c, d \in \mathbb{Q}\} && (701) \\ &= \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\} && (702) \end{aligned}$$

Where $\sqrt{2} \times \sqrt{3} = \sqrt{2 \times 3} = \sqrt{6}$ follows from the definition of nth roots plus associativity on the reals. For (b), we prove bidirectionally.

1. $\mathbb{Q}[\sqrt{2} + \sqrt{3}] \subset \mathbb{Q}[\sqrt{2}, \sqrt{3}]$. Consider $y \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$. Then there exists $p \in \mathbb{Q}[x]$ s.t.

$$y = p(\sqrt{2} + \sqrt{3}) = a_n(\sqrt{2} + \sqrt{3})^n + \ldots + a_1(\sqrt{2} + \sqrt{3}) + a_0 \tag{703}$$

where the terms can be expanded an rearranged to the form $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \in \mathbb{Q}[\sqrt{2}, \sqrt{3}]$.

2. $\mathbb{Q}[\sqrt{2}, \sqrt{3}] \subset \mathbb{Q}[\sqrt{2} + \sqrt{3}]$. Consider $\sqrt{2} + \sqrt{3} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$. Since it is a field and $\sqrt{2} + \sqrt{3}$ is a unit, by rationalizing the denominator, we can get

$$(\sqrt{2} + \sqrt{3})^{-1} = \frac{\sqrt{2} - \sqrt{3}}{2 - 3} = \sqrt{3} - \sqrt{2} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}] \tag{704}$$

Therefore by adding and subtracting the two elements, we have $\sqrt{2}, \sqrt{3} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}] \implies \sqrt{6} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$. Since $\mathbb{Q} \subset \mathbb{Q}[\sqrt{2} + \sqrt{3}]$, from the ring properties all elements of the form $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \in \mathbb{Q}[\sqrt{2} + \sqrt{3}]$.

For the second part, I claim that $\sqrt{2} \notin \mathbb{Q}[\sqrt{6}]$. Assuming it is, we have $\sqrt{2} = a + b\sqrt{6} \implies 2 = a^2 + 6b^2 + 2ab\sqrt{6}$. So $a = 0$ or $b = 0$. If $a = 0$, then $b^2 = 1/3 \implies b = 1/\sqrt{3}$ which contradicts that b is rational. If $b = 0$, then $a^2 = 2 \implies a = \sqrt{2}$ which contradicts that a is rational.

Solution 13.27

Note that $\mathbb{Q}[\sqrt[3]{2}] = \{a + b\sqrt[3]{2} + c\sqrt[3]{4}\}$, and so

$$\begin{aligned} \mathbb{Q}[\sqrt[3]{2}, i] &= \mathbb{Q}[\sqrt[3]{2}][i] && (705) \\ &= \{\alpha + \beta i \mid \alpha, \beta \in \mathbb{Q}[\sqrt[3]{2}]\} && (706) \\ &= \{(a + b\sqrt[3]{2} + c\sqrt[3]{4}) + (d + e\sqrt[3]{2} + f\sqrt[3]{4})i \mid a, b, c, d, e, f \in \mathbb{Q}\} && (707) \\ &= \{a + b\sqrt[3]{2} + c\sqrt[3]{4} + di + e\sqrt[3]{2}i + f\sqrt[3]{4}i \mid a, b, c, d, e, f \in \mathbb{Q}\} && (708) \end{aligned}$$

We prove bidirectionally.

1. $\mathbb{Q}[\sqrt[3]{2} + i] \subset \mathbb{Q}[\sqrt[3]{2}, i]$. Consider $y \in \mathbb{Q}[\sqrt[3]{2} + i]$. Then there exists a $p \in \mathbb{Q}[x]$ s.t.

$$y = p(\sqrt[3]{2} + i) = a_n(\sqrt[3]{2} + i)^n + \ldots + a_1(\sqrt[3]{2} + i) + a_0 \tag{709}$$

Then we can expand and rearrange the terms to be of the form

$$a + b\sqrt[3]{2} + c\sqrt[3]{4} + di + ei\sqrt[3]{2} + fi\sqrt[3]{4} \in \mathbb{Q}[\sqrt[3]{2}, i] \tag{710}$$

2. $\mathbb{Q}[\sqrt[3]{2}, i] \subset \mathbb{Q}[\sqrt[3]{2} + i]$. Consider $\alpha = \sqrt[3]{2} + i \in \mathbb{Q}[\sqrt[3]{2} + i]$. Then $(\alpha - i)^3 = 2$. Therefore

$$\alpha^3 - 3\alpha^2 i - 3\alpha + i = 2 \implies i(1 - 3\alpha^2) = 2 + 3\alpha - \alpha^3 \tag{711}$$

$$\implies i = \frac{2 + 3\alpha - \alpha^3}{1 - 3\alpha^2} \in \mathbb{Q}[\sqrt[3]{2} + i] \tag{712}$$

Therefore $\sqrt[3]{2} = \alpha - i \in \mathbb{Q}[\sqrt[3]{2} + i]$, which allows us add all combinations $\{1, \sqrt[3]{2}, \sqrt[3]{4}, i, \sqrt[3]{2}i, \sqrt[3]{4}i\}$ into our basis.

Exercise 13.45 (Shifrin 3.2.6.b/c/d/g)

Suppose $\alpha \in \mathbb{C}$ is a root of the given irreducible polynomial $f(x) \in \mathbb{Q}[x]$. Find the multiplicative inverse of $\beta \in \mathbb{Q}[\alpha]$.

1. $f(x) = x^2 + 3x - 3$, $\beta = \alpha - 1$
2. $f(x) = x^3 + x^2 - 2x - 1$, $\beta = \alpha + 1$
3. $f(x) = x^3 + x^2 + 2x + 1$, $\beta = \alpha^2 + 1$
4. $f(x) = x^3 - 2$, $\beta = \alpha + 1$
5. $f(x) = x^3 + x^2 - x + 1$, $\beta = \alpha + 2$
6. $f(x) = x^3 - 2$, $\beta = r + s\alpha + t\alpha^2$
7. $f(x) = x^4 + x^2 - 1$, $\beta = \alpha^3 + \alpha - 1$

Solution 13.28

For (b), using the Euclidean algorithm gives

$$(1)(x^3 + x^2 - 2x - 1) + (-x^2 + 2)(x + 1) = 1 \tag{713}$$

and substituting the root α gives $(-\alpha^2 + 2)(\alpha + 1) = 1$. So we have $\beta^{-1} = -\alpha^2 + 2$. For (c), doing the same thing gives

$$(-x)(x^3 + x^2 + 2x + 1) + (x^2 + x + 1)(x^2 + 1) = 1 \tag{714}$$

and substituting α gives $(\alpha^2 + \alpha + 1)(\alpha^2 + 1) = 1$, so $\beta^{-1} = \alpha^2 + \alpha + 1$. For (d), we have

$$(-\frac{1}{3})(x^3 - 2) + (\frac{1}{3}x^2 - \frac{1}{3}x + \frac{1}{3})(x + 1) = 1 \tag{715}$$

and so substituting α gives $(\frac{1}{3}\alpha^2 - \frac{1}{3}\alpha + \frac{1}{3})(\alpha + 1) = 1$, so $\beta^{-1} = \frac{1}{3}\alpha^2 - \frac{1}{3}\alpha + \frac{1}{3}$. For (g), we have

$$(-x^2 - x - 2)(x^4 + x^2 - 1) + (x^3 + x^2 + 2x + 1)(x^3 + x - 1) = 1 \tag{716}$$

and so substituting α gives $(\alpha^3 + \alpha^2 + 2\alpha + 1)(\alpha^3 + \alpha - 1) = 1$, and so $\beta^{-1} = \alpha^3 + \alpha^2 + 2\alpha + 1$.

Exercise 13.46 (Shifrin 3.2.7)

Let $f(x) \in \mathbb{R}[x]$.

1. Prove that the complex roots of $f(x)$ come in "conjugate pairs"; i.e., $\alpha \in \mathbb{C}$ is a root of $f(x)$ if and only if $\overline{\alpha}$ is also a root.
2. Prove that the only irreducible polynomials in $\mathbb{R}[x]$ are linear polynomials and quadratic polynomials $ax^2 + bx + c$ with $b^2 - 4ac < 0$.

Solution 13.29

Listed.

1. If $\alpha \in \mathbb{C}$ is a root of f, then

$$0 = f(\alpha) = a_n \alpha^n + \ldots + a_1 \alpha + a_0 \tag{717}$$

for $a_i \in \mathbb{R}$. Since

$$\begin{aligned} 0 = \overline{0} = \overline{f(\alpha)} & \quad (718) \\ = \overline{a_n \alpha^n + \ldots + a_1 \alpha + a_0} & \quad (719) \\ = \overline{a_n}\,\overline{\alpha^n} + \ldots + \overline{a_1 \alpha} + \overline{a_0} & \quad (720) \\ = a_n \overline{\alpha}^n + \ldots + a_1 \overline{\alpha} + a_0 & \quad (721) \\ = p(\overline{\alpha}) & \quad (722) \end{aligned}$$

we can see that $\overline{\alpha} \in \mathbb{C}$ is immediately a root as well. Since $\overline{\overline{\alpha}} = \alpha$, the converse is immediately proven.

2. Linear polynomials in $F[x]$ for a given field are trivially irreducible (since multiplying polynomials increases the degree of the product as there are no zero divisors in a field). Perhaps without Theorem 4.1, we can assume that a real quadratic polynomial $p(x) = ax^2 + bx + c$ is reducible, which is equivalent to

$$p(x) = (dx + e)(fx + g) = dfx^2 + (dg + ef)x + eg \tag{723}$$

For $d, e, f, g \in \mathbb{R}$, and evaluating $b^2 - 4ac = (dg + ef)^2 - 4dfeg = (dg - ef)^2 \geq 0$ since this is a squared term of a real number. So we have proved that if it is quadratic and reducible, then the discriminant ≥ 0. To prove the other way, we assume that it is not reducible, i.e. there exists some complex root α from the fundamental theorem of algebra. Then from (1), we know that $\overline{\alpha}$ must also be a complex conjugate. Then this is reducible in $\mathbb{C}$ as

$$p(x) = a(x - \alpha)(x - \overline{\alpha}) \tag{724}$$

for some constant factor a. Letting $\alpha = d + ei$ for $d, e \in \mathbb{R}$, expanding it gives us

$$\begin{aligned} p(x) &= a\big(x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha}\big) & \quad (725) \\ &= ax^2 + -2adx + a(d^2 + e^2) & \quad (726) \end{aligned}$$

and evaluating the discriminant gives

$$4a^2 d^2 - 4a^2(d^2 + e^2) = -4a^2 e^2 < 0 \tag{727}$$

and we are done. For higher degree polynomials, we can proceed by taking a complex root (which is guaranteed to exist by fundamental theorem of algebra). If it contains an imaginary term, then its conjugate is also a root, and we factor out the quadratic. If it is real, then we can factor out the linear term. We can keep going this until we hit our base cases of a quadratic or linear term.

Exercise 13.47 (Shifrin 3.2.13)

Let K be a field extension of F, and suppose $\alpha, \beta \in K$. Show that $(F[\alpha])[\beta] = (F[\beta])[\alpha]$, so that $F[\alpha, \beta]$ makes good sense.
(Remark: One way to do this is to think about the ring of polynomials in two variables. The other way is just to show directly that every element of one ring belongs to the other.)

Solution 13.30

Let $y \in (F[\alpha])[\beta]$. Then there exists a polynomial $p \in (F[\alpha])[x]$ s.t.

$$y = p(\beta) = b_n\beta^n + \ldots + b_1\beta + b_0 = \sum_{i=0}^{n} b_i\beta^i \tag{728}$$

for $b_i \in F[\alpha]$. But since $b_i \in F[\alpha]$, there exists a polynomial $q_i \in F[x]$ s.t. (omitting the subscript i for clarity)

$$b_i = q_i(\alpha) = a_{n_i}\alpha^n + \ldots + a_1\alpha + a_0 = \sum_{j=0}^{n_i} a_j\alpha^j \tag{729}$$

for $a_j \in F$. Substituting each b_i in gives

$$y = \sum_{i=0}^{n}\left(\sum_{j=0}^{n_i} a_j\alpha^j\right)\beta^i = \sum_{i=0}^{n}\sum_{j=0}^{n_i} a_j\alpha^j\beta^i \tag{730}$$

With the same logic, every element of $(F[\beta])[\alpha]$ can be written as

$$y = \sum_{i=0}^{n}\left(\sum_{j=0}^{n_i} a_j\beta^j\right)\alpha^i = \sum_{i=0}^{n}\sum_{j=0}^{n_i} a_j\alpha^i\beta^j \tag{731}$$

Note that since $F[\alpha]$ is a vector space spanned by $\{1, \ldots, \alpha^{n-1}\}$, and $F[\beta]$ is a also a vector space spanned by $\{1, \ldots, \beta^{m-1}\}$ for some m, the two spaces above are spanned by all products $\{\alpha^i\beta^j\}_{i<n, j<m}$, and they are the same set.

Exercise 13.48 (Shifrin 3.3.2.a/d/e/g)

Decide which of the following polynomials are irreducible in $\mathbb{Q}[x]$.

a $f(x) = x^3 + 4x^2 - 3x + 5$
1. $f(x) = 4x^4 - 6x^2 + 6x - 12$
2. $f(x) = x^3 + x^2 + x + 1$

d $f(x) = x^4 - 180$

e $f(x) = x^4 + x^2 - 6$

3. $f(x) = x^4 - 2x^3 + x^2 + 1$

g $f(x) = x^3 + 17x + 36$

4. $f(x) = x^4 + x + 1$
5. $f(x) = x^5 + x^3 + x^2 + 1$
6. $f(x) = x^5 + x^3 + x + 1$

Solution 13.31

For (a), by the rational root theorem the rational roots, if any, must be in the set $\{\pm 1, \pm 5\}$. Calculating them gives $f(x) = 7, 11, 215, -5$. Since this is third degree, no linear factors means that it is irreducible, so f is irreducible.
For (d), by the Eisenstein's criterion with $p = 5$ this polynomial is irreducible.
For (e), the rational root theorem states that the rational roots must be in $\{\pm 1, \pm 2, \pm 3, \pm 6\}$. This polynomial is clearly even, so it suffices to check the positive candidates. This gives $-4, 14, 84, 1326$. Therefore if it is reducible, by Gauss's lemma it must be of the form

$$(ax^2 + bx + c)(dx^2 + ex + f) \tag{732}$$

for integer coefficients. $a = d = 1$ is trivial ($-1, -1$ is also possible but constant factors don't matter).

Expanding this gives

$$x^4 + (b+e)x^3 + (c+f+be)x^2 + (bf+ce)x + cf = x^4 + x^2 - 6 \tag{733}$$

The coefficients of x^3 tell us that $e = -b$, which means that for the coefficents of x, $bf + ce = bf - bc = 0 \implies f = c$. So $c^2 = -6$, which has no solution. Therefore f is irreducible.
For (g), we must check rational roots of $\{\pm 1, \pm 2, \pm 3, \pm 4, \pm 6, \pm 9, \pm 12, \pm 18, \pm 36\}$. Since this polynomial is monotonically increasing, with $f(-2) = -6$ and $f(0) = 36$. It only suffices to check $x = -1$, which gives $f(-1) = 18$. Therefore there are no linear factors. Since this is third degree, no linear factors means that it is irreducible, so f is irreducible.

Exercise 13.49 (Shifrin 3.3.4)

Show that each of the following polynomials has no rational root:
1. $x^{200} - x^{41} + 4x + 1$
2. $x^8 - 54$
3. $x^{2k} + 3x^{k+1} - 12$, $k \geq 1$

Solution 13.32

Listed.
1. By the rational root theorem, the only possible rational roots are ± 1. Solving for both of these values gives
$$f(1) = 1 - 1 + 4 + 1 = 5 \tag{734}$$
$$f(-1) = 1 + 1 - 4 + 1 = -1 \tag{735}$$
Therefore there are no rational roots.
2. The only possible rational roots are $\pm 1, \pm 2, \pm 3, \pm 6, \pm 9, \pm 18, \pm 27, \pm 54$. But this polynomial is even, so it suffices to check the positive roots. $f(1) = -53$, $f(2) = 256 - 54 = 202$, and any greater inputs will increase the output since f is monotonic in $\mathbb{Z}^+$. Therefore f has no rational roots.
3. By Eisenstein's criterion with $p = 3$, this polynomial is irreducible and therefore has no rational roots.

Exercise 13.50 (Shifrin 3.3.6)

Listed.
1. Prove that $f(x) \in \mathbb{Z}_2[x]$ has $x + 1$ as a factor if and only if it has an even number of nonzero coefficients.
2. List the irreducible polynomials in $\mathbb{Z}_2[x]$ of degrees $2, 3, 4$, and 5.

Solution 13.33

Listed. Since $f(x)$ has $x + 1$ as a factor iff

$$f(1) = a_n 1^n + \ldots + a_1 1^1 + a_0 = a_n + \ldots + a_1 + a_0 = 0 \tag{736}$$

where each $a_i \in \{0, 1\}$. Therefore, this is equivalent to saying that there are an even number of 1's (nonzero coefficients), which sum to 0 mod 2. Therefore, the irreducible polynomials should at least have a constant coefficient of 1 (so we can't factor x) and should have odd number of terms (so that we can't factor $x + 1$). This will guarantee that $f(0) = f(1) = 1$.
1. Degree 2: $x^2 + x + 1$ is the only candidate and indeed is an irreducible polynomial.

2. Degree 3: $x^3 + x^2 + 1$, $x^3 + x + 1$ and indeed $f(0) = f(1) = 1$. Since it's only degree 3 we don't need to check irreducibility into 2 terms of both degree at least 2.
3. Degree 4: $x^4 + x^3 + x^2 + x + 1$, $x^4 + x^3 + 1$, $x^4 + x^2 + 1$, $x^4 + x + 1$ are candidates. However we need to check that they cannot be factored into two irreducible quadratic polynomials. The only possible such factorization is
$$(x^2 + x + 1)(x^2 + x + 1) = x^4 + x^2 + 1 \tag{737}$$
and so the irreducible polynomials are $x^4 + x^3 + x^2 + x + 1$, $x^4 + x^3 + 1$, $x^4 + x + 1$.
4. Degree 5: x^5+x^4+1, x^5+x^3+1, x^5+x^2+1, x^5+x+1, $x^5+x^4+x^3+x^2+1$, $x^5+x^4+x^3+x+1$, $x^5 + x^4 + x^2 + x + 1$, $x^5 + x^3 + x^2 + x + 1$ are the possible candidates. But we need to check that it is not factorable into an irreducible quadratic and cubic. The three candidates are
$$(x^2 + x + 1)(x^3 + x^2 + 1) = x^5 + x + 1 \tag{738}$$
$$(x^2 + x + 1)(x^3 + x + 1) = x^5 + x^4 + 1 \tag{739}$$
and so the irreducible polynomials are x^5+x^3+1, x^5+x^2+1, $x^5+x^4+x^3+x^2+1$, $x^5+x^4+x^3+x+1$, $x^5+x^4+x^2+x+1$, $x^5+x^3+x^2+x+1$.

Exercise 13.51 (Shifrin 3.3.7)

Prove that for any prime number p, $f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$ is irreducible in $\mathbb{Q}[x]$.

Solution 13.34

We can use the identity
$$f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1 = \frac{x^p - 1}{x - 1} \tag{740}$$
Therefore,
$$f(x+1) = \frac{(x+1)^p - 1}{(x+1) - 1} = \frac{1}{x}\left\{\left(\sum_{k=0}^{p} \binom{p}{k} x^k\right) - 1\right\} \tag{741}$$
$$= \frac{1}{x}\sum_{k=1}^{p} \binom{p}{k} x^k = \sum_{k=1}^{p} \binom{p}{k} x^{k-1} \tag{742}$$
Focusing on the coefficients, the leading coefficient is $\binom{p}{p} = 1$, and the rest of the coefficients are divisible by p. The constant coefficient is $\binom{p}{1} = p$, which is not divisible by p^2. By Eisenstein's criterion, $f(x+1)$ is irreducible $\implies f(x)$ is irreducible. To justify the final step, assume that $f(x)$ is reducible. Then $f(x) = g(x)h(x)$ for positive degree polynomials g, h. Then by substituting $x + 1$, we have that $f(x+1) = g(x+1)h(x+1)$, which means that $f(x+1)$ is irreducible.

Exercise 13.52 (Shifrin 4.1.3)

(a) Prove that if $I \subset R$ is an ideal and $1 \in I$, then $I = R$.
(b) Prove that $a \in R$ is a unit if and only if $\langle a \rangle = R$.
(c) Prove that the only ideals in a (commutative) ring R are $\langle 0 \rangle$ and R if and only if R is a field.

Solution 13.35

Listed.

(a) If $1 \in I$, then for every $r \in R$, we must have $r1 = r \in I$. Therefore $I = R$.

(b) If $a \in R$ is a unit, then $a^{-1} \in R$, and so for every $r \in R$, $ra^{-1} \in R$. Therefore, $\langle a \rangle$ must contain all elements of form $ra^{-1}a = r$, which is precisely R. Now assume that a is not a unit, and so there exists no $a^{-1} \in R$. Therefore, $\langle a \rangle$, which consists of all ra for $r \in R$, cannot contain 1 since $r \neq a^{-1}$, and so $\langle a \rangle \neq R$.

(c) For the forwards implication, assume that R is not a field. Then there exists some $a \neq 0$ that is not a unit, and taking $\langle a \rangle$ gives us an ideal that—from (b)—is not R. For the backward implication we know that $\langle 0 \rangle$ is an ideal. Now assume that there exists another ideal I containing $a \neq 0$. Since R is a field, a is a unit, and so by (b) $R = \langle a \rangle \subset I \subset R \implies I = R$.

Exercise 13.53 (Shifrin 4.1.4.a/b/c)

Find all the ideals in the following rings:

(a) $\mathbb{Z}$

(b) $\mathbb{Z}_7$

(c) $\mathbb{Z}_6$

(d) $\mathbb{Z}_{12}$

(e) $\mathbb{Z}_{36}$

(f) $\mathbb{Q}$

(g) $\mathbb{Z}[i]$ (see Exercise 2.3.18)

Solution 13.36

Listed.

(a) All sets of form $\{kz \in \mathbb{Z} \mid z \in \mathbb{Z}\}$ for all $k \in \mathbb{Z}$.

(b) Only $\{0\}$ and $\mathbb{Z}_7$ is an ideal.

(c) We have $\{0\}, \{0,2,4\}, \{0,3\}, \mathbb{Z}_6$.

Exercise 13.54 (Shifrin 4.1.5)

(a) Let $I = \langle f(x) \rangle$, $J = \langle g(x) \rangle$ be ideals in $F[x]$. Prove that $I \subset J \Leftrightarrow g(x) | f(x)$.

(b) List all the ideals of $\mathbb{Q}[x]$ containing the element $f(x) = (x^2 + x - 1)^3 (x-3)^2$.

Solution 13.37

For (a), we prove bidirectionally.

1. $(\rightarrow)$. Since $f(x) \in \langle f(x) \rangle \implies f(x) \in \langle g(x) \rangle$, this means that $f(x) = r(x) g(x)$ for some $r(x) \in F[x$. Therefore $g(x) \mid f(x)$.
2. $(\leftarrow)$. Given that $g(x) \mid f(x)$, let us take some $f_1(x) \in I$. Then it is of the form $f_1(x) = r(x) f(x)$ for some $r(x) \in F[x]$. But since $g(x) \mid f(x)$, $f(x) = h(x) g(x)$ for some $h(x) \in F[x]$. Therefore $f_1(x) = r(x) h(x) g(x) = (rh)(x) g(x)$, where $(rh)(x) \in F[x]$, and so $f_1(x) \in J$.

For (b), we can use the logic from (a) to find all the factors of $f(x)$, which generate all sup-ideals of $\langle f(x) \rangle$, which is the minimal ideal containing $f(x)$.

1. $g(x) = 1 \implies \langle 1 \rangle = F[x]$
2. $g(x) = x^2 + x - 1 \implies \langle x^2 + x - 1 \rangle$
3. $g(x) = (x^2 + x - 1)^2 \implies \langle (x^2 + x - 1)^2 \rangle$
4. $g(x) = (x^2 + x - 1)^3 \implies \langle (x^2 + x - 1)^3 \rangle$
5. $g(x) = x - 3 \implies \langle x - 3 \rangle$
6. $g(x) = (x^2 + x - 1)(x - 3) \implies \langle (x^2 + x - 1)(x - 3) \rangle$

7. $g(x) = (x^2+x-1)^2(x-3) \implies \langle (x^2+x-1)^2(x-3)\rangle$
8. $g(x) = (x^2+x-1)^3(x-3) \implies \langle (x^2+x-1)^3(x-3)\rangle$
9. $g(x) = (x-3)^2 \implies \langle (x-3)^2\rangle$
10. $g(x) = (x^2+x-1)(x-3)^2 \implies \langle (x^2+x-1)(x-3)^2\rangle$
11. $g(x) = (x^2+x-1)^2(x-3)^2 \implies \langle (x^2+x-1)^2(x-3)^2\rangle$
12. $g(x) = (x^2+x-1)^3(x-3)^2 \implies \langle (x^2+x-1)^3(x-3)^2\rangle$

Exercise 13.55 (Shifrin 4.1.14.a/b)

Mimicking Example 5(c), give the addition and multiplication tables of
(a) $\mathbb{Z}_2[x]/\langle x^2+x\rangle$
(b) $\mathbb{Z}_3[x]/\langle x^2+x-1\rangle$
(c) $\mathbb{Z}_2[x]/\langle x^3+x+1\rangle$
In each case, is the quotient ring an integral domain? a field?

Solution 13.38

For (a), note that the quotient allows us to state that $x^2 \equiv x \pmod{I}$, and therefore every polynomial in $\mathbb{Z}_2[x]/\langle x^2+x\rangle$ is equivalent to a linear polynomial. Therefore, the elements in this quotient are $0, 1, x, x+1$. As you can see, this is not an integral domain (and hence not a field) since $x, x+1$ are zero divisors.

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	x	0
$x+1$	0	$x+1$	0	$x+1$

Figure 13: Addition and multiplication tables for $\mathbb{Z}_2[x]/\langle x^2+x\rangle$.

For (b), note that the quotient allows us to state that $x^2 \equiv 2x+1 \pmod{I}$, and therefore every polynomial in $\mathbb{Z}_3[x]/\langle x^2+x-1\rangle$ is equivalent to a linear polynomial. Therefore, the elements in this quotient are $0, 1, 2, x, x+1, x+2, 2x, 2x+1, 2x+2$. This is indeed an integral domain since there are no zero divisors, and it is a field since every nonzero element is a unit (all rows/columns are filled with all elements of the set).

$+$	0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
0	0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
1	1	2	0	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$
2	2	0	1	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$
x	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$	0	1	2
$x+1$	$x+1$	$x+2$	x	$2x+1$	$2x+2$	$2x$	1	2	0
$x+2$	$x+2$	x	$x+1$	$2x+2$	$2x$	$2x+1$	2	0	1
$2x$	$2x$	$2x+1$	$2x+2$	0	1	2	x	$x+1$	$x+2$
$2x+1$	$2x+1$	$2x+2$	$2x$	1	2	0	$x+1$	$x+2$	x
$2x+2$	$2x+2$	$2x$	$2x+1$	2	0	1	$x+2$	x	$x+1$

Figure 14: Addition table for $\mathbb{Z}_3[x]/\langle x^2+x-1\rangle$.

$\times$	0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
0	0	0	0	0	0	0	0	0	0
1	0	1	2	x	$x+1$	$x+2$	$2x$	$2x+1$	$2x+2$
2	0	2	1	$2x$	$2x+2$	$2x+1$	x	$x+2$	$x+1$
x	0	x	$2x$	$2x+1$	1	$x+1$	$x+2$	$2x+2$	2
$x+1$	0	$x+1$	$2x+2$	1	$x+2$	$2x$	2	x	$2x+1$
$x+2$	0	$x+2$	$2x+1$	$x+1$	$2x$	2	$2x+2$	1	x
$2x$	0	$2x$	x	$x+2$	2	$2x+2$	$2x+1$	$x+1$	1
$2x+1$	0	$2x+1$	$x+2$	$2x+2$	x	1	$x+1$	2	$2x$
$2x+2$	0	$2x+2$	$x+1$	2	$2x+1$	x	1	$2x$	$x+2$

Figure 15: Multiplication table for $\mathbb{Z}_3[x]/\langle x^2+x-1\rangle$.

Exercise 13.56 (Shifrin 4.1.17)

Let R be a commutative ring and let $I, J \subset R$ be ideals. Define

$$\begin{aligned} I \cap J &= \{a \in R : a \in I \text{ and } a \in J\} \\ I + J &= \{a + b \in R : a \in I, b \in J\}. \end{aligned}$$

(a) Prove that $I \cap J$ and $I + J$ are ideals.
(b) Suppose $R = \mathbb{Z}$ or $F[x]$, $I = \langle a\rangle$, and $J = \langle b\rangle$. Identify $I \cap J$ and $I + J$.
(c) Let $a_1, \ldots, a_n \in R$. Prove that $\langle a_1, \ldots, a_n\rangle = \langle a_1\rangle + \cdots + \langle a_n\rangle$.

Solution 13.39

For (a), we prove it in 5.4. For (b), the argument is equivalent for $\mathbb{Z}$ and $F[x]$. $I \cap J$ consists of all elements that are divisible by both a and b, so $I \cap J = \langle \text{lcm}(a, b)\rangle$. $I + J$ consists of all elements that are of form $ra + sb$, but this are all multiples of $\gcd(a, b)$ and so $I + J = \langle \gcd(a, b)\rangle$.
For (c), it suffices to prove $\langle a, b\rangle = \langle a\rangle + \langle b\rangle$.

1. $\langle a, b\rangle \subset \langle a\rangle + \langle b\rangle$. $x \in \langle a, b\rangle \implies x = r_a a + r_b b$ for $r_a, r_b \in R$. But $a \in \langle a\rangle, b \in \langle b\rangle \implies r_a a \in \langle a\rangle, r_b b \in \langle b\rangle$, and so $x \in \langle a\rangle + \langle b\rangle$.
2. $\langle a, b\rangle \supset \langle a\rangle + \langle b\rangle$. $x \in \langle a\rangle + \langle b\rangle \implies x = a_x + b_x$ for $a_x \in \langle a\rangle, b_x \in \langle b\rangle$. But $a_x \in \langle a\rangle \implies a_x = r_a a$ for some $r_a \in R$, and $b_x \in \langle b\rangle \implies b_x = r_b b$ for some $r_b \in R$. So $x = r_a a + r_b b \iff x \in \langle a, b\rangle$.

We know that for $\langle a_1\rangle = \langle a_1\rangle$, and so by making this argument $n-1$ times we can build up by induction that $\langle a_1, \ldots a_{n-1}, a_n\rangle = \langle a_1, \ldots, a_{n-1}\rangle + \langle a_n\rangle$.

Exercise 13.57 (Shifrin 4.2.1)

(a) Prove that the function $\phi : \mathbb{Q}[\sqrt{2}] \to \mathbb{Q}[\sqrt{2}]$ defined by $\phi(a+b\sqrt{2}) = a - b\sqrt{2}$ is an isomorphism.
(b) Define $\phi : \mathbb{Q}[\sqrt{3}] \to \mathbb{Q}[\sqrt{7}]$ by $\phi(a + b\sqrt{3}) = a + b\sqrt{7}$. Is ϕ an isomorphism? Is there any isomorphism?

Solution 13.40

For (a), we first prove that it is a homomorphism.

$$\begin{aligned}
\phi((a+b\sqrt{2})+(c+d\sqrt{2})) &= \phi((a+c)+(b+d)\sqrt{2}) && (743)\\
&= (a+c)-(b+d)\sqrt{2} && (744)\\
&= (a-b\sqrt{2})+(c-d\sqrt{2}) && (745)\\
&= \phi(a+b\sqrt{2})+\phi(c+d\sqrt{2}) && (746)\\
\phi((a+b\sqrt{2})(c+d\sqrt{2})) &= \phi((ac+2bd)+(ad+bc)\sqrt{2}) && (747)\\
&= (ac+2bd)-(ad+bc)\sqrt{2} && (748)\\
&= (a-b\sqrt{2})(c-d\sqrt{2}) && (749)\\
&= \phi(a+b\sqrt{2})\times\phi(c+d\sqrt{2}) && (750)\\
\phi(1) &= 1 && (751)
\end{aligned}$$

This is injective since given that $a+b\sqrt{2}\neq c+d\sqrt{2}$, then at least $a\neq b$ or $c\neq d$, in which case $a-b\sqrt{2}\neq c-d\sqrt{2}$. Alternatively, we can see that the kernel is 0, so it must be injective. It is onto since given any $c+d\sqrt{2}$, the preimage is $c-d\sqrt{2}$. Therefore ϕ is an isomorphism.
For (b), no it is not an isomorphism since

$$\begin{aligned}
\phi((a+b\sqrt{3})(c+d\sqrt{3})) &= \phi((ac+3bd)+(ad+bc)\sqrt{3}) && (752)\\
&= (ac+3bd)+(ad+bc)\sqrt{7} && (753)\\
&\neq (ac+7bd)+(ad+bc)\sqrt{7} && (754)\\
&= (a+b\sqrt{7})(c+d\sqrt{7}) && (755)\\
&= \phi(a+b\sqrt{3})\phi(c+d\sqrt{3}) && (756)
\end{aligned}$$

We claim that there is no isomorphism. Assume that such ϕ exists. Then $\phi(1)=1$, and so $\phi(3)=\phi(1+1+1)=\phi(1)+\phi(1)+\phi(1)=1+1+1=3$. Now given $\sqrt{3}\in\mathbb{Q}[\sqrt{3}]$, we follows that

$$\phi(\sqrt{3})^2=\phi(3)=3 \qquad (757)$$

and so $\phi(\sqrt{3})$ must map to the square root of 3 which must live in $\mathbb{Q}[\sqrt{7}]$. Assume such a number is $a+b\sqrt{7} \implies (a^2+7b^2)+(2ab)\sqrt{7}=\sqrt{3}$. This implies that $2ab=0$, leaving the rational term, but we know that $\sqrt{3}$ does not exist in the rationals, and so $\sqrt{3}$ does not exist.

Exercise 13.58 (Shifrin 4.2.3.a/c/e)

Establish the following isomorphisms (preferably, using Theorem 2.2):
(a) $\mathbb{R}[x]/\langle x^2+6\rangle\cong\mathbb{C}$
(b) $\mathbb{Z}_{18}/\langle\bar{6}\rangle\cong\mathbb{Z}_6$
(c) $\mathbb{Q}[x]/\langle x^2+x+1\rangle\cong\mathbb{Q}[\sqrt{3}i]$
(d) $\mathbb{Z}[x]/\langle 2x-3\rangle\cong\mathbb{Z}[\frac{3}{2}]=\{\frac{a}{b}\in\mathbb{Q}: b=2^j \text{ for some } j\geq 0\}\subset\mathbb{Q}$
(e) $F[x]/\langle x\rangle\cong F$
(f) $\mathbb{Z}_3\times\mathbb{Z}_4\cong\mathbb{Z}_{12}$

Solution 13.41

For all, we construct the ring homomorphism $\phi: R\to S$ with the appropriate kernel, and the result is immediate from the theorem.

a) Given $f \in \mathbb{R}[x]$ which is a Euclidean domain, we claim that the map $\phi_1 : f(x) \mapsto r(x)$ where r is the remainder of f when divided by x^2+6, is a homomorphism. It is pretty easy to see that the map $\phi_2 : f(x) = \sum_{k=0}^n a_k x^k \mapsto a_0 + a_1 i$ is also a homomorphism, and thus $\phi = \phi_2 \circ \phi_1$ as the composition of homomorphisms is also a ring homomorphism. ϕ_1 is a homomorphism since given $f, g \in \mathbb{R}[x]$, we can write them as $f(x) = d_1(x)(x^2+6) + r_1(x)$ and $g(x) = d_2(x)(x^2+6)+r_2(x)$. Therefore,

$$(f+g)(x) = f(x) + g(x) = (d_1(x) + d_2(x))(x^2+6) + (r_1 + r_2)(x) \tag{758}$$

$$(fg)(x) = f(x) \cdot g(x) = (d_1(x)(x^2+6) + r_1(x))(d_2(x)(x^2+6) + r_2(x)) \tag{759}$$

$$= (\dots)(x^2+6) + (r_1 + r_2)(x) \tag{760}$$

$$1 = 0(x^2+6) + 1 \tag{761}$$

Therefore ϕ is a homomorphism, and the kernel is simply all polynomials divisible by x^2+6, which is $\langle x^2 + 6 \rangle$.

c) We define $\phi(f) = f(\frac{-1+\sqrt{3}i}{2})$, where $\frac{-1+\sqrt{3}i}{2}$ is a root of x^2+x+1. Therefore, since $f \in \mathbb{R}$, $\frac{-1-\sqrt{3}i}{2}$ must also be a root and so the kernel is $\langle x^2+x+1 \rangle$. Second, we will show that it is a homomorphism.

$$\phi(f+g) = (f+g)\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) = f\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) + g\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) = \phi(f) + \phi(g) \tag{762}$$

$$\phi(fg) = (fg)\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) = f\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) \cdot g\bigg(\frac{-1+\sqrt{3}i}{2}\bigg) = \phi(f) \cdot \phi(g) \tag{763}$$

$$\phi(1) = 1 \tag{764}$$

We are done.

e) Given $f(x) = \sum_{k=0}^n a_k x^k \in F[x]$, we show that $\phi : f \mapsto a_0$ is a homomorphism. Let f be as above and g have coeffients b_k from $k = 0 \dots m$.

$$\phi(f+g) = \phi\bigg(\sum_{k=0}^{\max\{n,m\}} (a_k + b_k) x^k\bigg) = a_0 + b_0 = \phi(f) + \phi(g) \tag{765}$$

$$\phi(fg) = \phi\bigg(\sum_{k=0}^{n+m} \Big(\sum_{i=0}^{k} a_i b_{k-i}\Big) x^k\bigg) = a_0 b_0 = \phi(f)\phi(g) \tag{766}$$

$$\phi(1) = 1 \tag{767}$$

So this is a homomorphism. Since $\langle x \rangle$ as all multiples of x consists of all polynomials with constant term $a_0 = 0$, we can see that $\ker(\phi) = 0$. Therefore we are done.

Exercise 13.59 (Shifrin 4.2.11.a/d)

True or false? (Give proofs or disproofs.)
(a) $\mathbb{Z}_2[x]/\langle x^2 \rangle \cong \mathbb{Z}_4$, or $\mathbb{Z}_2[x]/\langle x^2 \rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$?
(b) Same questions for $\mathbb{Z}_2[x]/\langle x^2 + x \rangle$.
(c) Same questions for $\mathbb{Z}_2[x]/\langle x^2 + 1 \rangle$.
(d) $\mathbb{Z}_3[x]/\langle x^2 - 1 \rangle \cong \mathbb{Z}_3 \times \mathbb{Z}_3$?
(e) $\mathbb{Q}[x]/\langle x^2 - 1 \rangle \cong \mathbb{Q} \times \mathbb{Q}$?

Solution 13.42

Listed.

(a) False for both. The characteristic of $\mathbb{Z}_2[x]/\langle x \rangle$ is 2 since $1+1=0$, but the characteristic of $\mathbb{Z}_4$ is 4 since $1+1+1+1=0$, so false. As for $\mathbb{Z}_2 \times \mathbb{Z}_2$, note that $(0,1)$ and $(1,0)$ are zero divisors of each other where $(0,1)\cdot(1,0)=(0,0)$. However, the two zero divisors in $\mathbb{Z}_2[x]/\langle x \rangle$ are x and $x+1$, where $x^2=(x+1)^2=0$. An isomomorphism $\phi : \mathbb{Z}_2[x]/\langle x \rangle \to \mathbb{Z}_2 \times \mathbb{Z}_2$ would have to preserve $0=\phi(0)=\phi(x^2)=\phi(x)\cdot\phi(x)$, but there are no nonzero elements $(a,b) \in \mathbb{Z}_2 \times \mathbb{Z}_2$ whose square is 0. Therefore, there cannot be an isomorphism.

(d) True. All elements of $\mathbb{Q}[x]/\langle x^2-1 \rangle$ are of form $a+bx$, with $a, b \in \mathbb{Q}$. We define the isomorphism $\phi(a+bx) = (a+b, a-b) \in \mathbb{Z}_3 \times \mathbb{Z}_3$. This is a homomorphism since

$$\begin{aligned}
\phi((a_1+b_1x)+(a_2+b_2x)) &= \phi((a_1+a_2)+(b_1+b_2)x) && (768)\\
&= (a_1+a_2+b_1+b_2, a_1+a_2-b_1-b_2) && (769)\\
&= (a_1+b_1, a_1-b_1)+(a_2+b_2, a_2-b_2) && (770)\\
&= \phi(a_1+b_1x)+\phi(a_2+b_2x) && (771)\\
\phi((a_1+b_1x)(a_2+b_2x)) &= \phi(a_1a_2+(a_1b_2+a_2b_1)x+b_1b_2x^2) && (772)\\
&= \phi((a_1a_2+b_1b_2)+(a_1b_2+a_2b_1)x) && (773)\\
&= (a_1a_2+b_1b_2+a_1b_2+a_2b_1, a_1a_2+b_1b_2-a_1b_2-a_2b_1) && (774)\\
&= (a_1+b_1, a_1-b_1)(a_2+b_2, a_2-b_2) && (775)\\
&= \phi(a_1+b_1x)\phi(a_2+b_2x) && (776)\\
\phi(1) &= 1 && (777)
\end{aligned}$$

This is also injective since given $a_1+b_1x \neq a_2+b_2x$, say that their images are the same. Then $a_1+b_1=a_2+b_2$ and $a_1-b_1=a_2-b_2$. Adding and subtracting the two equations, we have $2a_1=2a_2$ and $2b_1=2b_2$, which means the original elements were the same.

Exercise 13.60 (Shifrin 4.2.12)

Let R be a commutative ring, $I \subset R$ an ideal. Suppose $a \in R$, $a \notin I$, and $I+\langle a \rangle = R$ (see Exercise 4.1.17 for the notion of the sum of two ideals). Prove that $\bar{a} \in R/I$ is a unit.

Solution 13.43

Since $R = I+\langle a \rangle$, $1 \in R = I+\langle a \rangle$. So there exists $i \in I, ra \in \langle a \rangle$ s.t. $1 = i+ra \implies ra = 1-i$. Therefore, in the quotient ring, $\bar{i}=0$ and we have

$$\bar{r}\bar{a} = \bar{1}-\bar{0} = \bar{1} \tag{778}$$

and so $\bar{r}$ is a multiplicative inverse of $\bar{a}$. So $\bar{a}$ is a unit.

13.6 Polynomial rings

13.7 Modules

13.8 Vector Spaces

13.9 Field Theory and Galois Theory

Exercise 13.61 (Shifrin 5.3.3)

The polynomial $f(x) = x^2 + 1$ is irreducible in $\mathbb{Z}_3[x]$, and so $K = \mathbb{Z}_3[x]/(x^2 + 1)$ is a field with nine elements. Let $\alpha \in K$ be a root of $f(x)$. Find irreducible polynomials in $\mathbb{Z}_3[x]$ having as roots, respectively,

a. $\alpha + 1$

b. $\alpha - 1$.

Solution 13.44

Listed.

1. We can see that

$$\begin{aligned} (\alpha+1)^2 = \alpha^2 + 2\alpha + 1 = 2\alpha &\implies (\alpha+1)^2 - 2\alpha = 0 && (779)\\ &\implies (\alpha+1)^2 - 2\alpha - 2 + 2 = 0 && (780)\\ &\implies (\alpha+1)^2 - 2(\alpha+1) + 2 = 0 && (781)\end{aligned}$$

and so $f(x) = x^2 - 2x + 2 \in \mathbb{Z}_3[x]$ has $\alpha + 1$ as a root.

2. Similarly, we have

$$\begin{aligned} (\alpha-1)^2 = \alpha^2 - 2\alpha + 1 = -2\alpha &\implies (\alpha-1)^2 + 2\alpha = 0 && (782)\\ &\implies (\alpha-1)^2 + 2\alpha - 2 + 2 = 0 && (783)\\ &\implies (\alpha-1)^2 + 2(\alpha-1) + 2 = 0 && (784)\end{aligned}$$

and so $f(x) = x^2 + 2x + 2 \in \mathbb{Z}_3[x]$ has $\alpha - 1$ as a root.

Exercise 13.62 (Shifrin 5.3.4)

Construct explicitly an isomorphism

$$\mathbb{Z}_2[x]/(x^3 + x + 1) \to \mathbb{Z}_2[x]/(x^3 + x^2 + 1).$$

Solution 13.45

Both $x^3 + x + 1$ and $x^3 + x^2 + 1$ are irreducible in $\mathbb{Z}_2[x]$, so both are fields of order 8 (since the x^3 is equivalent to a lower order polynomial) consisting of all polynomials in $\mathbb{Z}_2[x]$ of degree ≤ 2. We can construct the isomorphism ϕ sending $\phi(f(x)) = f(x + 1)$. This is a homomorphism since it maps 1 to 1, and

$$\begin{aligned} \phi((f+g)(x)) = (f+g)(x+1) = f(x+1) + g(x+1) = \phi(f(x)) + \phi(g(x)) && (785)\\ \phi((fg)(x)) = (fg)(x+1) = f(x+1)g(x+1) = \phi(f(x))\phi(g(x)) && (786)\end{aligned}$$

It is also bijective since the inverse mapping $\phi(f(x)) = f(x-1) = f(x+1)$ is well-defined. Finally, we can see that that considering ϕ as an automorphism over $\mathbb{Z}_2[x]$, $\phi(x^3+x+1) = (x+1)^3+(x+1)+1 = x^3 + x^2 + 1$, so it maps the ideals to each other. This therefore induces an isomorphism between the quotient rings. We can explicitly write out the image of each element.

1. $\phi(0) = 0$.
2. $\phi(1) = 1$.
3. $\phi(x) = x + 1$
4. $\phi(x + 1) = x$.
5. $\phi(x^2) = x^2 + 1$.
6. $\phi(x^2 + 1) = x^2$.
7. $\phi(x^2 + x) = x^2 + x$.
8. $\phi(x^2 + x + 1) = x^2 + x + 1$.

Exercise 13.63 (Shifrin 5.3.5)

Let F be a finite field of characteristic p. Show that every element $a \in F$ can be written in the form $a = b^p$ for some $b \in F$. (Hint: Consider the Frobenius automorphism.)

Solution 13.46

Then F has $q = p^n$ elements for some $n \in \mathbb{N}$, and in Shifrin we have established through Frobenius automorphism $\sigma(a) = a^p$ that $\sigma^n(a)$ is the identity, i.e.

$$a = \sigma^n(a) = (a^p)^n = a^{pn} = (a^n)^p \tag{787}$$

Therefore, we have found $b = a^n \in F$ satisfying the condition.

Exercise 13.64 (Shifrin 5.3.7)

Let $q = p^n$, and let $f(x) = x^q - x$.

a. Prove that if $g(x)$ is an irreducible polynomial of degree d in $\mathbb{Z}_p[x]$, then $g(x)$ divides $f(x)$ if and only if $d|n$.
b. Prove that $f(x)$ is the product of all monic, irreducible polynomials in $\mathbb{Z}_p[x]$ whose degrees divide n.

Solution 13.47

For (a), we prove bidirectionally. Since $g(x)$ is irreducible, $F = \mathbb{Z}_p[x]/\langle g(x)\rangle$ is a field of p^d elements and $g(x)$ is the minimal polynomial of α over $\mathbb{Z}_p$. We also know that for any element a in a field of order p^d, it satisfies $a^{p^d} = a$. Additionally, the multiplicative group of units $(\mathbb{Z}_p[x]/\langle g(x)\rangle)^*$ is a cyclic group of order $p^d - 1$ generated by α. By Lagrange's theorem, the order of any element of this multiplicative group must divide $p^d - 1$. Choosing α, we have $\alpha^{p^d - 1} = 1 \implies \alpha^{p^d} = \alpha$.

1. $(\rightarrow)$. Let $g(x)$ divide $f(x) = x^{p^n} - x$. Then $g(\alpha) = 0 \implies f(\alpha) = \alpha^{p^n} - \alpha = 0 \implies \alpha^{p^n} = \alpha$. Therefore,
$$\alpha = \alpha^{p^d} = \alpha^{p^n} \tag{788}$$
The smallest positive integer m such that $\alpha^{p^m} = \alpha$ is $m = d$ as $g(x)$ is the minimal polynomial. Since $\alpha^{p^n} = \alpha$ and d is the smallest such exponent, we have $d \mid n$.
2. $(\leftarrow)$. Assume that $d \mid n$. Consider the field $F = \mathbb{Z}_p[x]/\langle g(x)\rangle$, which is a field of order p^d. We also know that for any element a in a field of order p^d, it satisfies $a^{p^d} = a$. Taking $x \in \mathbb{Z}_p[x]$, its image $\bar{x} \in F$ has the property that $\bar{x}^{p^d} - \bar{x} = 0$, and so this means that $x^{q^d} - x$ is in the kernel of this quotient map. Therefore $(x^{p^d} - x) \in \langle g(x)\rangle \implies g(x) \mid (x^{p^d} - x)$. To prove the final step, we prove that $\forall d, n$, $x^{p^d} - x \mid x^{p^n} - x$ iff $d \mid n$.
Then we have $n = kd$ for some $k \in \mathbb{N}$, and so
$$\alpha^{p^n} = \alpha^{p^{kd}} = \alpha \tag{789}$$

and so α is a root of $x^{p^n} - x$. Now assuming that $g(x) \nmid f(x)$, since $g(x)$ is irreducible the GCD is 1, and so there exists $a(x), b(x)$ s.t.

$$a(x)f(x) + b(x)g(x) = 1 \tag{790}$$

But by setting $x = \alpha$, we get $f(\alpha) = 0$ from above, and $g(\alpha) = 0$ by assumption, leading to $0 = 1$, which is a contradiction since $0 \neq 1$ always in fields. Therefore $g(x) \mid f(x)$.

For (b), we have shown in (a) that the irreducible factors of $f(x)$ are precisely all polynomials in $\mathbb{Z}_p[x]$ whose degree divides n. Since $\mathbb{Z}_p$ is a field, we can scalar multiply the polynomial—and hence the leading coefficient—by the multiplicative inverse of the leading coefficient to make it monic. This doesn't change the factorization since the leading coefficient of $f(x)$ is also 1. Since $\mathbb{Z}_p[x]$ is a Euclidean domain, by unique factorization theorem all such polynomials $g(x)$ must be contained within the product.

It now remains to show that $f(x)$ is square free, i.e. none of its factors have multiplicity greater than 1. Take f and its derivative (where $p = 0$ in $\mathbb{Z}_p$)

$$f(x) = x^{p^n} - x, \qquad f'(x) = p^n x^{p^n - 1} - 1 = -1 \tag{791}$$

It is clear that $\gcd(f, f') = 1$ since f' is constant. Now assume that there is some factor $a(x)$ of multiplicity at least 2. Then $f(x) = a(x)a(x)b(x)$ for some $b(x) \in \mathbb{Z}_p[x]$. Taking the derivative gives

$$\begin{aligned} f'(x) &= (a(x)a'(x) + a'(x)a(x))b(x) + a(x)^2 b'(x) && (792) \\ &= a(x)\big(a'(x)b(x) + a'(x)b(x) + a(x)b'(x)\big) && (793) \end{aligned}$$

which means that at least $a(x) \mid \gcd(f, f')$, contradicting that the gcd is 1. Therefore f is square free. Finally, since $\mathbb{Z}_p[x]$ is a Euclidean domain, by the unique factorization theorem all of its factors are precisely

13.10 Affine and Projective Spaces

13.11 Representations

13.12 Lie Groups and Lie Algebras