Abstract Algebra

Muchang Bahng

Spring 2024

Contents

1 Group-Like Structures **5**
1.1 Semigroups and Monoids . . . 5
1.2 Groups . . . 6
1.3 Group Homomorphisms . . . 8
1.4 Group Presentations . . . 10
1.5 Symmetric and Alternating Groups . . . 14
1.6 Group Actions . . . 16

2 Subgroups **18**
2.1 Centralizers and Normalizers . . . 21
2.2 Stabilizers and Orbits . . . 21
2.3 Subgroups Generated by Subsets of a Group . . . 21
2.4 Lattice of Subgroups . . . 21

3 Quotient Groups **22**
3.1 Normal Subgroups . . . 22
3.2 Quotients and Lagrange's Theorem . . . 22
3.3 Homomorphism and Isomorphism Theorems . . . 22

4 Group Actions **23**
4.1 Sylow Theorems . . . 23

5 Product Groups **24**
5.1 Direct Products . . . 24
5.2 Semidirect Products . . . 24
5.3 Classification of Finite Abelian Groups . . . 24
5.4 Group Extensions . . . 24

6 Rings **25**
6.1 Ring Homomorphisms and Characteristics . . . 28
6.2 Commutative Rings . . . 29
6.3 Domains . . . 29
6.4 Ideals . . . 31
6.5 Quotient Rings . . . 33
6.6 Homomorphism and Isomorphism Theorems . . . 35
6.7 Unique Factorization Domain . . . 35
6.8 Principal Ideal Domains . . . 35
6.9 Euclidean Domains . . . 37
6.10 Division Rings . . . 38
6.11 Fields . . . 39

7 Polynomial Rings **41**
7.1 Basic Properties of Polynomials . . . 42
7.2 Euclidean Division . . . 43
7.3 Roots and Factorization . . . 44
7.4 Algebraically Closed Fields . . . 46
7.5 Reducibility of Real Polynomials . . . 49
7.6 Reducibility of Integer Polynomials . . . 51
7.7 Rational Functions . . . 53
7.8 Exercises . . . 54

8 Modules **55**
8.1 Modules over a PID . . . 55
8.2 Rational Canonical Form . . . 55
8.3 Jordan Canonical Form . . . 55

9 Vector Space Structures **56**
9.1 Modules . . . 57
9.2 Algebras . . . 57
9.3 The Algebra of Quaternions . . . 57
9.3.1 Matrix Representations of Quaternions . . . 59
9.3.2 Square Roots of -1 . . . 60
9.4 Tensor Algebras . . . 61

10 Field and Galois Theory **63**
10.1 Extensions and Splitting Fields . . . 63
10.2 Finite Fields . . . 63

11 Affine and Projective Spaces **64**
11.1 Affine Spaces . . . 64
11.2 Convex Sets . . . 68
11.3 Affine Transformations and Motions . . . 71
11.4 Quadrics . . . 74
11.5 Projective Spaces . . . 74

12 Representations **76**

13 Lie Groups and Lie Algebras **80**
13.1 Lie Algebras of Classical Lie Groups . . . 82
13.1.1 Lie Algebras of SL(2, R) and SL(2, C) . . . 82
13.1.2 Lie Algebra of SU(2) . . . 83
13.1.3 Lie Algebra of SO(3) . . . 85
13.1.4 Lie Algebra of SE(n) . . . 88
13.2 Representations of Lie Groups and Lie Algebras . . . 89
13.2.1 Tensor Products of Group Representations . . . 92
13.3 Topological Decompositions of Lie Groups . . . 93
13.4 Linear Lie Groups . . . 95
13.4.1 Lie Algebras of SO(3) and SU(2), Revisited . . . 99
13.5 Abstract Lie Groups . . . 101

14 Fields **104**
14.1 Ring Extensions . . . 104
14.2 Field Extensions . . . 106
14.3 Splitting Fields . . . 108

15 TBD **110**

15.1 Abelian Groups . . . 110
15.2 Some Types of Groups . . . 112
15.2.1 Lie Groups . . . 113
15.3 Ring Homomorphisms in Polynomials . . . 114

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.

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.

Definition 1.1 (Semigroup)

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

Definition 1.2 (Monoid)

A **monoid** $(M, *)$ is a semigroup with an identity element $1 \in M$ such that given a $m \in M$

$$1 * m = m * 1 = m \tag{3}$$

Groupoids aren't necessarily that interesting, but there are cases in which semigroups and monoids come up.

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. $X_t : \Omega \rightarrow S$) satisfying the **Markov property**: for every bounded measurable f and and $t, s \geq 0$,

$$\mathbb{E}[f(X_{t+s}) \mid \{X_r\}_{r \leq t}] = \mathbb{E}[f(X_{t+s}) \mid X_t] = (P_s f)(X_t) \tag{4}$$

The set $\{P_t\}_{t \geq 0}$ is called the **Markov semigroup**.

Example 1.2 (Monoid of Transformations)

Given a set S, consider the set of all functions $f : S \rightarrow S$. This forms a monoid with the identity function $f(x) = x$ as the identity element. Proof of associativity is shown in my set theory notes.

Theorem 1.1 (Cardinality of Monoid of Transformations)

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

Example 1.3 ()

Let S be any nonempty set. Then $(2^S, \cup, \emptyset)$ and $(2^S, \cap, S)$ are monoids.

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 1 of a monoid M is unique.

Proof.

Assume not, i.e. there are 2 identities $1 \neq 1'$. But then

$$1 = 11' = 1' \implies 1 = 1' \tag{5}$$

where the implication follows from transitivity of equivalence relations.

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** $(M', *)$.

Theorem 1.2 (Identities of Submonoids)

If M' with identity $1'$ is a submonoid of M with identity 1, Then $1 = 1'$.

Proof.

Assume not. $1' \in M$, which means that

1.2 Groups

Definition 1.4 (Group)

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

1. *Closure.* $x, y \in G \implies xy \in G$[b]
2. *Associativity.* $\forall x, y, z \in G, x(yz) = (xy)z$
3. *Identity.* $\exists e \in G$ s.t. $\forall x \in G, xe = ex = x$
4. *Inverses.* $\forall x \in G \; \exists x^{-1} \in G$ s.t. $xx^{-1} = x^{-1}x = e$

The **order** of a group is the cardinality $|G|$.

[a] Note that this is a monoid with the additional property of inverses.
[b] but not necessarily $xy = yx$

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 $x * a = b$ has the unique solution $x = b * a^{-1}$.

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,

$$\begin{aligned} a * a_1^{-1} = e &\implies a_2^{-1} * \left(a * a_1^{-1}\right) = a_2^{-1} * e \\ &\implies \left(a_2^{-1} * a\right) * a_1^{-1} = a_2^{-1} \\ &\implies e * a_1^{-1} = a_2^{-1} \end{aligned}$$

Since the inverse is unique, we can operate on each side of the equation $x * a = b$ to get $x * a * a^{-1} = b * a^1 \implies x * e = x = b * a^{-1}$. 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 random (or is it?) 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.

Example 1.4 (Group of Invertible Elements of a Monoid)

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.

Definition 1.5 (Subgroup)

Given group $(G, *)$ and $(G', *)$ with the same operations, G' is a **subgroup** of G if $G' \subset G$.

Definition 1.6 (Abelian Group)

An **abelian group** $(A, +)$ is a group where $+$ is commutative.[a]

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

operations.

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

Example 1.5 (Abelian Groups)

Here are some examples.

1. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are all abelian groups with respect to addition. $\mathbb{Q}^* \equiv \mathbb{Q} \setminus \{0\}$ and $\mathbb{R}^* \equiv \mathbb{R} \setminus \{0\}$ are abelian groups with respect to multiplication.
2. The set of all functions on a given interval $[a, b]$ is abelian with respect to addition, defined as $(f + g)(x) \equiv f(x) + g(x)$.

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 1.7 (Group Homomorphism)

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

$$f(a \circ b) = f(a) * f(b) \tag{6}$$

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 1.4 (Homomorphisms Maps Identities/Inverses to Identities/Inverses)

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

$$f(e_G) = e_H, \qquad f(a^{-1}) = f(a)^{-1} \tag{7}$$

Proof.

Let $a \in G$. Then

$$f(a) = f(ae_G) = f(a)f(e_G) \implies e_H = f(a)^{-1}f(a) = f(a)^{-1}f(a)f(e_G) = f(e_G) \tag{8}$$

To prove inverses, we see that

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(e_G) = e_H \tag{9}$$

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

Example 1.6 (Exponential Map)

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

$$2^{a+b} = 2^a \times 2^b \tag{10}$$

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

Example 1.7 (Determinant)

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

Therefore, we can see that an isomorphism is really just a "renaming" of the elements, which aligns with our view of equivalence as above. Not only does it renamee the elements, but it preserves all the algebraic properties of the group and each element.

Theorem 1.3 (Preservation of Properties in Isomorphism)

If $f : G \to H$ is an isomorphism, then
1. f^{-1} is also an isomorphism.
2. $|G| = |H|$.
3. $\forall a \in G$, $\mathrm{ord}(a) = \mathrm{ord}(f(a))$.
4. 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
$$f\big(f^{-1}(c), f^{-1}(d)\big) = f\big(f^{-1}(c)\big)\, f\big(f^{-1}(d)\big) = cd \tag{11}$$
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
$$f^{-1}(c)f^{-1}(d) = f^{-1}(cd) \tag{12}$$
and so f^{-1} is a homomorphism.
2. This is trivial by bijectivity.
3. TBD.
4. 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 1.4 ()

Let G be a group with $a \in G$. Then the following is an automorphism on G.

$$\phi : G \longrightarrow G,\ \phi(x) = axa^{-1} \tag{13}$$

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.

Definition 1.8 (Kernel)

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

$$\ker(f) \coloneqq \{g \in G \mid f(g) = e_H\} \tag{14}$$

That is, it is the preimage of the identity.

Theorem 1.5 (Kernels are Subgroup)

Given a group homomorphism $f : G \to H$,

1. $\ker(f)$ is a subgroup of G.
2. f is injective $\iff \ker(f) = \{e_G\}$.

Proof.

For the first part, we prove the properties of a group. 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.

For the second part, we prove bidirectionally.

1. $(\rightarrow)$. 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 $\ker(f) = \{e_G\}$.
2. $(\leftarrow)$. 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 $\ker(f) = \{e_G\} \implies ab^{-1} = e_G \implies a = b$.

Example 1.8 (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

$$f(z_1 z_2) = \frac{z_1 z_2}{|z_1 z_2|} = \frac{z_1 z_2}{|z_1||z_2|} = f(z_1)f(z_2) \tag{15}$$

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

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

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

$$\ldots, a^{-2}, a^{-1}, a^0 = e, a^1, a^2, \ldots \tag{16}$$

It may or may not be the case that a may cycle back to itself for some n, i.e. $a = a^n$.

Definition 1.9 (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, I know.

Now 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 1.10 (Word)

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

$$s_1^{\epsilon_1} s_2^{\epsilon_2} s_3^{\epsilon_3} \ldots s_k^{\epsilon_k}, \text{ where} e_i \in \mathbb{Z} \tag{17}$$

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

Definition 1.11 (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 1.12 (Free Group)

The **free group** F_S over a given set S consists of all words that can be built from elements of S.

Now for notational convenience, one method of specifying a group is to put it in the form

$$\langle\, S \mid R \,\rangle \tag{18}$$

where S is the generating set and R is a set of relations. This is called the *group presentation.*

Example 1.9 (Group Presentations)

The cyclic group of order n could be presented as

$$\langle\, a \mid a^n = 1 \,\rangle \tag{19}$$

Dih (8), with r representing a rotation by 45 degrees in the direction of the orientation and f representing a flip over any axis, is presented by

$$\langle\, \{r, f\} \mid r^8 = 1, f^2 = 1, (rf)^2 = 1 \,\rangle \tag{20}$$

A lot of groups fall into one or more categories depending on what properties they have. We will proceed to just define these categories and introduce the groups as needed.

Theorem 1.6 (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.

Definition 1.13 (Cyclic Group)

A **cyclic group**, denoted Z_n, is a group generated by a single element. In a **finite cyclic group**, there exists a $k \in \mathbb{N}$ such that $g^k = g^0 = 1$ (or in additive notation, $kg = 0g = 0$), where g is the generator. A **finitely generated group** is a group generated by a finite number of elements. In **infinite cyclic groups**, all elements are distinct for distinct k.

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

$$R = \left\{ \begin{pmatrix} \sin\theta & \cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix} \,\middle|\, \theta \in \left\{ \frac{2\pi}{n} k \right\}_{k=0}^{n-1} \right\} \tag{21}$$

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 1.7 (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! However, note that there can be isomorphisms from a cyclic group to a subgroup.

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

$$\mathbb{Z} \simeq 2\mathbb{Z} \tag{22}$$

Definition 1.14 (Polytope)

A **polytope** in n-dimensions is a geometrical object with "flat sides, " called an n-polytope. It is a generalization of a polygon or a polyhedron to an arbitrary number of dimensions.

Definition 1.15 (Simplex)

A **n-simplex** is a n-polytope which is the n-dimensional convex hull of its $n+1$ vertices. Moreover, the $n+1$ vertices must be **affinely independent**, meaning that

$$\{u_1 - u_0, u_2 - u_0, ..., u_n - u_0 | \{u_i\}_{i=0}^n \text{ vertices}\} \tag{23}$$

are linearly independent vectors that span the n-dimensional space.

Definition 1.16 (Symmetry Group)

The **symmmetry group** of a geometrical object is the group of all transformations in which the object is invariant. Preserving all the relevant structure of the object.

Definition 1.17 (Dihedral Group)

A common example of such groups is the **dihedral group** of order $2n$, with the group presentation

$$\text{Dih}(n) \coloneqq \langle r, f \mid r^n = f^2 = e, rfr = f\rangle \tag{24}$$

, denoted $\text{Dih}(n)$ of order $2n$, which is the group of symmetries of a n-simplex, which includes rotations and reflections.

Example 1.12 (Dih(3) on Triangle)

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

	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 3: Multiplication table for D_3.

Example 1.13 (Groups of Order 3)

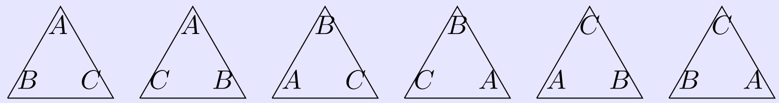
$\text{Dih}(3) \simeq S_3$, since permutations of the vertices of a triangle are isomorphic to a permutations of a 3-element set.

However, S_4 is not isomorphic to the symmetry group of a square. It is however, isomorphic to that of a tetrahedron, i.e. Dih(4).

Example 1.14 (Low Order Dihedral Group)

We introduce some low order Dihedral groups.

1. 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.



2. Dih(4) is the group of all rotations and reflections that preserve the structure of the regular tetrahedron in $\mathbb{R}^3$. An incorrect, yet somewhat useful, way of visualizing this group is to imagine a square in $\mathbb{R}^2$. However, the points are not pairwise equidistant and therefore does not preserve symmetry between all points.
3. Dih(n) is similarly the group of all rotations and reflections that preserve the structure of a regular $(n-1)$-simplex in $\mathbb{R}^{n-1}$.

Example 1.15 (Klein 4 Group)

The **Klein 4-Group** 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.

$\cdot$	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 4: Multiplication table for the Klein 4-group (V_4)

Example 1.16 (Groups of Order 4)

There are only 2 groups of order 4.

C_4	e	a	a^2	a^3
e	e	a	a^2	a^3
a	a	a^2	a^3	e
a^2	a^2	a^3	e	a
a^3	a^3	e	a	a^2

(a) Cyclic group C_4

V	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(b) Klein four-group V

Figure 5: Cayley tables for the two groups of order 4

1.5 Symmetric and Alternating Groups

Notice that given any set S, we can define the set of all functions $f : S \to S$ as a monoid. What if we consider the set of all invertible functions? This by definition means bijective functions, and so consider this subset.

Definition 1.18 (Symmetric/Transformation Group)

Given a set S, the **transformation group**, or **symmetric group**, of S is the group of all bijective maps from S to itself.

This exists for all sets S, and if S is finite, we call it a **permutation group**, since the set of bijective transformations of it is a permutation of its elements.

Definition 1.19 (Permutation Group)

The **permutation group** is the set of all bijective transformations from any set X to the same set, denoted either $\text{Sym}(X)$ or S_n. If $X = \{1, 2, 3, ..., n\}$, known as the set of all permutations of X, with cardinality $n!$.

Lemma 1.5 ()

Every element in finite S_n can be decomposed into a partition of cyclic rotations.

Example 1.17 ()

Listed.

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

Definition 1.20 ()

The **conjugacy class** of S_n correspond to the cycle structures of S_n. Two elements of S_n are conjugate in S_n if and only if they consist of the same number of disjoint cycles of the same lengths.

Example 1.18 ()

1. $(123)(45)$ is conjugate to $(143)(25)$.
2. $(12)(45)$ is not conjugate to $(143)(25)$.

Theorem 1.8 (Transpositions)

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

Definition 1.21 ()

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

$$\text{sgn} : S_n \longrightarrow \{1, -1\} \tag{25}$$

Lemma 1.6 ()

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

Now the reason that symmetric groups are nice is that we can embed a group into its symmetric group.

Theorem 1.9 (Cayley's Theorem)

Every group G is isomorphic to a subgroup of its symmetric group. If G is finite, then so is $\text{Sym}(G)$, so every finite group is a subgroup of S_n, for some n.

Proof.

Let $H = \text{Sym}(G)$. We define the map
$$\phi : G \longrightarrow H \tag{26}$$
by the following rule. For $a \in G$, map it to permutation $\sigma = \phi(a) \in H$ defined as $\sigma(g) = ag$ for all $g \in G$. Note that given an $a \in G$, ag must also be in G, meaning that a corresponding $\sigma \in H$ exists. It is sufficient to prove that ϕ is an isomorphism onto its image. We first prove injectivity. Given $a \neq b \in G$, $\phi(a) = \sigma, \phi(b) = \tau$. Assume $\sigma = \tau \implies a = ae = \sigma(e) = \tau(e) = be = b \implies a = b$, a contradiction. We now check that $\phi(ab) = \phi(a)\phi(b)$. Given $g \in G, \phi(a)\phi(b)(g) = \phi(a)(bg) = a(bg) = (ab)g = \phi(ab)(g)$.

Definition 1.22 (Alternating Group)

The **alternating group** of order n is the set of all **even permutations** (permutations that have signature 1) of $\{1, 2, ..., n\}$. It is denoted A_n or $\text{Alt}(n)$ and its cardinality is $\frac{1}{2}n!$. 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.

Example 1.19 (Low Order Symmetric Groups)

1. S_0 is the set of all permutations on the **null set**. S_1 is the set of all permutations on the **singleton set**. Both sets have cardinality 1 and the element is **trivial**. Note that $S_1 = A_1$.
2. S_2 is a cyclic, abelian group of order 2 consisting of the identity permutation and the transposition of two elements.
3. S_3 is the first cyclic, nonabelian group, with order $6.S \simeq \text{Dih}(3)$, which can be seen as the group of rotations and reflections on the equilateral triangle, and the elements of S_3 equate to permuting the vertices on the triangle.

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

Definition 1.23 (Group Action)

Let G be a group, X a set. Then, a (left) group action of G on X is a function:
$$\varphi : G \times X \longrightarrow X, \ (g, x) \longmapsto \varphi(g, x) \tag{27}$$
satisfying two axioms:

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

The group G is said to **act on** X. X is called a **G-set**. The two axioms, furthermore, imply that for every $g \in G$, the function that maps $x \in X$ to $\varphi(g, x) \in X$ is a bijective map, since the inverse is the function mapping $x \mapsto \varphi(g^{-1}, x)$.
(g, x) can be interpreted as the element g in the transformation group G acting on an element x in X.

Example 1.20 ()

$\text{Isom}\,\mathbb{R}^3$ acts on $\mathbb{R}^3$ since every element $g \in \text{Isom}\,\mathbb{R}^3$ acts on the entire space $\mathbb{R}^3$.

Example 1.21 ()

S_n acts on $\{1, 2, ..., n\}$by permuting its elements.

Example 1.22 ()

The $\text{GA}(V)$ acts transitively on the points of an affine space.

Equivalent Interpretation of Group Actions Note that this group action G on space X identifies a group homomorphism into the group of automorphisms of that space. Given an abstract group element $g \in G$, $\varphi(g, \cdot) : X \longrightarrow X$ is defined accordingly, where $\varphi(g, \cdot) \in \text{Aut}(X)$. So alternatively, we can interpret a group action as a homomorphism from G to $\text{Aut}(X)$.

$$\phi : G \longrightarrow \text{Aut}(X),\ g \mapsto \phi(g) = \varphi(g, \cdot) \tag{28}$$

Definition 1.24 (Representation)

A group action on a finite-dimensional vector space X is called a **representation** of that group.

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 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. 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 2.1 (Coset)

Given a group G, $g \in G$, and subgroup H,

1. A **left coset** is $gH \coloneqq \{gh \mid h \in H\}$.
2. A **right coset** is $Hg \coloneqq \{hg \mid h \in H\}$.
3. If G is abelian, then the **coset** is $gH \coloneqq \{g + h \mid h \in H\}$.

This divides the group into equivalence classes $g \mapsto [g] = gH$, and we write (for left cosets)

$$a \equiv b \pmod{H} \iff a = bh \text{ for some } h \in H \tag{29}$$

Proof.

We show that this indeed forms an equivalence class.

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

$$|G| = [G : H]|H| \tag{30}$$

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.

However, the converse is usually false, as 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 $|G| = n$ and $a \in G$ is arbitrary, then $a^n = e$.

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,

$$a^{p-1} \equiv 1 \iff a^p \equiv a \pmod{p} \tag{31}$$

Definition 2.2 (Normal Subgroups)

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

$$b^{-1}hb \in H \tag{32}$$

Every subgroup of an abelian group is normal.

The concept of normal subgroups allow us to endow on the quotient set a group structure.

Definition 2.3 (Quotient Group)

Given a group G and a normal subgroup H, the **quotient group** G/H is the set of left cosets aH with the operation

$$aH\, bH = abH \tag{33}$$

Lemma 2.1 ()

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,

$$\begin{aligned} \ker \phi = \phi^{-1}(eH) &= \{a \in G \mid aH = eH = H\} && (34) \\ &= \{a \in G \mid a \in H\} && (35) \end{aligned}$$

Therefore, ϕ is a homomorphism because $\phi(ab) = abH = (aH)(bH)$.

Theorem 2.2 (Quotient Maps are Homomorphisms)

The map $\pi : G \to G/H$ is a group homomorphism, and the **quotient group** is the set of left cosets with

Proof.

Corollary 2.5 ()

If $|G| = n$, then $g^n = e$ for all $g \in G$.

Definition 2.4 (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 2.3 (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 $G = \mathbb{Z}_n \setminus \{0\}$,

$$a^{\varphi(n)} \equiv 1 \pmod{n}, \quad \text{where } a \text{ is coprime to } n \tag{36}$$

Example 2.1 ()

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

Definition 2.5 ()

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

$$Gx \equiv \{gx \mid g \in G\} \tag{37}$$

By definition, transitive transformation groups have only one orbit.

Definition 2.6 ()

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

Example 2.2 ()

The orbits of $O(2)$ are concentric circles around the origin, as well as the origin itself. The stabilizer of the point $p \neq 0$ is the identity and the reflection across the line ??. The stabilizer of 0 is the entire $O(2)$.

Example 2.3 ()

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 2.4 ()

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.

Definition 2.7 ()

The **length of an orbit** is the number of elements in it.

Corollary 2.6 ()

If G is a finite group, then

$$|G| = |G_x||Gx| \tag{38}$$

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

$$G_{gx} = gG_x g^{-1} \tag{39}$$

2.1 Centralizers and Normalizers

2.2 Stabilizers and Orbits

2.3 Subgroups Generated by Subsets of a Group

Subgroups of cyclic group

2.4 Lattice of Subgroups

3 Quotient Groups

3.1 Normal Subgroups

3.2 Quotients and Lagrange's Theorem

3.3 Homomorphism and Isomorphism Theorems

4 Group Actions

4.1 Sylow Theorems

5 Product Groups

5.1 Direct Products

Definition 5.1 (Direct Product)

The **direct product** of two groups G and H is denoted

$$G \times H \equiv \{(g, h) \mid g \in G, h \in H\} \tag{40}$$

Note that the product need not be binary (nor must it be of finite arity).

Example 5.1 ()

The **general affine group** is defined

$$\mathrm{GA}(V) \equiv \mathrm{Tran}\, V \times \mathrm{GL}(V) \tag{41}$$

Example 5.2 ()

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

$$\mathrm{Tran}\, \mathbb{R}^4 \times H \times \mathrm{O}(3) \tag{42}$$

where H is the group of transformations of the form

$$(x, y, z, t) \longmapsto (x + at, y + bt, z + ct, t) \tag{43}$$

Example 5.3 ()

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

$$G = \mathrm{Tran}\, \mathbb{R}^4 \times \mathrm{O}_{3,1} \tag{44}$$

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

$$x^2 + y^2 + z^2 - t^2 \tag{45}$$

5.2 Semidirect Products

5.3 Classification of Finite Abelian Groups

5.4 Group Extensions

6 Rings

Definition 6.1 (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 $+$

$$a \times (b + c) = a \times b + a \times c \tag{46}$$
$$(a + b) \times c = a \times c + b \times c \tag{47}$$

   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]

Lemma 6.1 ()

Additive inverses are unique and $-1 \times a$ is the additive inverse of a.

Proof.

We can see that

$$-1 + 1 = 0 \implies (-1 + 1) \times a = 0 \times a \tag{48}$$
$$\implies -1 \times a + 1 \times a = 0 \tag{49}$$
$$\implies -1 \times a + a = 0 \tag{50}$$

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

Definition 6.2 (Characteristic Number)

The **characteristic** of ring R, denoted $\text{char}(R)$, is the smallest number of times one must successively add the multiplicative identity 1 to get the additive identity 0.

$$1 + 1 + ... + 1 = 0 \tag{51}$$

If no such number n exists, then $\text{char}(R) = 0$.

Theorem 6.1 (Freshman's Dream)

Given a field F with $\text{char}(F) = p$,

$$(a + b)^p = a^p + b^p \tag{52}$$

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

Proof.

We have

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^{p-k} b^k \tag{53}$$

It is clear that

$$\binom{p}{k} = \frac{p(p-1)...(p-k+1)}{k!} \tag{54}$$

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

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 6.3 (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. $uv = vu = 1$.

The next property that we would like to talk about is a zero divisor, which is the property that nonzero $a, b \in R$ satisfy $ab = 0$.

Definition 6.4 (Left, Right Zero Divisor)

An element a of a ring R is called a **left zero divisor** if there exists a nonzero x such that $ax = 0$ and a **right zero divisor** if there exists a nonzero x such that $xa = 0$.

Another property that we would desire is some sort of decomposition of ring elements as other ring elements.

Definition 6.5 (Left, Right Divisor)

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

1. If there exists an element $x \in R$ with $ax = b$, we say a is a **left divisor** of b.
2. If there exists an element $y \in R$ with $ya = b$, we say a is a **right divisor** of b.
3. We say a is a **two-sided divisor** if it is both a left divisor and a right divisor of b. Note that the x and y are not required to be equal.

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 6: Basic hierarchy of rings.

Example 6.1 (Integers, Rationals, Reals, Complexes)

The fields $\mathbb{Z}, \mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are rings with:

1. Sets:
   - $\mathbb{Q}$: rational numbers $\{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$
   - $\mathbb{R}$: real numbers
   - $\mathbb{C}$: complex numbers $\{a + bi : a, b \in \mathbb{R}\}$
2. Standard addition and multiplication
3. Additive identity 0
4. Multiplicative identity 1

These form commutative rings with unity where every non-zero element has a multiplicative inverse.

Example 6.2 (Continuous Functions)

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

Example 6.3 (Matrices)

The ring $M_n(R)$ of $n \times n$ matrices over a ring R consists of:

1. $n \times n$ arrays of elements from R
2. Matrix addition (entry-wise):
$$(A + B)_{ij} = A_{ij} + B_{ij} \tag{55}$$
3. Matrix multiplication:
$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \tag{56}$$
4. Zero matrix as additive identity
5. Identity matrix I_n as multiplicative identity

This forms a non-commutative ring for $n > 1$, even when R is commutative.

Example 6.4 (Power Set)

Given a set X, let 2^X be its power set, that is the set of all subsets of X. Then, 2^X is a commutative associative ring with respect to the operations of symmetric difference (i.e. the set of elements which is in exactly one of the sets)

$$M \triangle N \equiv (M \setminus N) \cup (N \setminus M) \tag{57}$$

and intersection $\cap$, taken for addition an multiplication, respectively. We will not prove all of the axioms of the ring, but we can state some important facts about this structure. The additive identity is $\emptyset$ and the multiplicative identity is X. Finally, it is clear that

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

6.1 Ring Homomorphisms and Characteristics

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. Recall that 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 $X \cong Z$.

Definition 6.6 (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. $f(1_R) = 1_S$

for all $a, b \in R$.[a] If f is a bijective ring homomorphism, it is called a **ring isomorphism**.

[a]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*.

Definition 6.7 (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 6.2 (Properties of Ring Homomorphisms)

It immediately follows that if $f : R \to S$ is a ring homomorphism, then

1. $f(0) = 0$
2. $\Im(f)$ is a subring of S.
3. A ring homomorphism is injective if and only if $\ker f = \langle 0 \rangle$.

Furthermore, if f is a ring isomorphism, then

1. f^{-1} is a ring isomorphism.

Theorem 6.2 (Compositions of Ring Homomorphisms)

Compositions of ring homomorphisms are ring homomorphisms.

6.2 Commutative Rings

Note that for commutative rings, distinguishing left and right divisors are meaningless, and so we can talk about just *divisors*.

Lemma 6.3 (Left=Right Divisors)

In a commutative ring R, a is a left divisor of b iff a is a right divisor of b. In this case, we just say that a is a **divisor** of b, written $a|b$.

Proof.

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

Definition 6.8 (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 6.4 (Euclid)

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

Lemma 6.5 ()

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

Definition 6.9 (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.

6.3 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 6.10 (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 6.5 (Domains vs Integral Domains)

We show some examples of integral domains.

1. The ring $\mathbb{Z}$ of integers.
2. The field $\mathbb{R}$.
3. The ring $\mathbb{Z}[x]$ of polynomials of one variable with integer coefficients.

We show examples of domains that are not integral domains.

1. Quaternions $\mathbb{H}$ are not commutative but are a domain.

Theorem 6.3 (Fields are Integral Domains)

Every field is an integral domain.

Proof.

Theorem 6.4 (Polynomial Integral Domains)

Rings of polynomials are an integral domain if the coefficients come from an integral domain.

Proof.

Factorization of polynomials over $\mathbb{C}$ into linear factors and polynomials over $\mathbb{R}$ into linear and quadratic factors is similar to the factoring of the integers to prime numbers. In fact, such a factorization exists for polynomials over any field F, but their factors can be of any degree. Moreover, there exists no general solution for the factoring of polynomials over any field.

Example 6.6 ()

$\mathbb{Z}$ and $F[x]$ over field F are integral domains. Any field F is also an integral domain.

Example 6.7 ()

The quotient ring $\mathbb{Z}_n$ is not an integral domain when n is composite.

Example 6.8 ()

A product of two nonzero commutative rings with unity $R \times S$ is not an integral domain since $(1,0) \cdot (0,1) = (0,0) \in R \times S$.

Example 6.9 ()

The ring of $n \times n$ matrices over any nonzero ring when $n \geq 2$ is not an integral domain. Given matrices A, B, if the image of B is in the kernel of A, then $AB = 0$.

Example 6.10 ()

The ring of continuous functions on the interval is not an integral domain. To see why, notice that given the piecewise functions

$$f(x) = \begin{cases} 1 - 2x & x \in [0, \frac{1}{2}] \\ 0 & x \in [\frac{1}{2}, 1] \end{cases}, \quad g(x) = \begin{cases} 0 & x \in [0, \frac{1}{2}] \\ 2x - 1 & x \in [\frac{1}{2}, 1] \end{cases} \tag{58}$$

$f, g \neq 0$, but $fg = gf = 0$.

Theorem 6.5 ()

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

Theorem 6.6 ()

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

Definition 6.11 (Regular Elements)

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

$$\rho : R \longrightarrow R, \qquad x \mapsto xr \tag{59}$$

is injective for all $x \in R$.

Theorem 6.7 ()

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

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

Definition 6.12 (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 6.8 (GCD's in a Euclidean Domain)

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

6.4 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 6.13 (Ideals)

For a commutative ring $(R, +, \times)$, a **two-sided ideal**—or **ideal**—is a subset $I \subset R$ satisfying
1. $a, b \in I \implies a + b \in I$.
2. $a \in I, r \in R \implies ra = ar \in I$.

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

Let's try to elaborate more on this interpretation by introducing immediate consequences.

Lemma 6.6 (Ideals are Groups Under +)

Given a commutative ring R and ideal $I \subset R$, $(I, +)$ is an abelian group.

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*.

Example 6.11 (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.

Given these two examples, we can think of an ideal consisting of all multiples of a specific element a that *generates* the ideal.

Definition 6.14 (Generators of Ideals)

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

$$\langle a \rangle \coloneqq \{ra \mid r \in R\} \tag{60}$$

and more generally, we may have multiple generating elements.

$$\langle a_1, \ldots, a_n \rangle \coloneqq \{r_1 a_1 + \ldots r_n a_n \mid r_1, \ldots, r_n \in R\} \tag{61}$$

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 6.12 (Matrix with Last Row of Zeros)

The set of all $n \times n$ matrices whose last row is zero forms a right ideal in the ring of all $n \times n$ matrices. However, it is not a left ideal.
The set of all $n \times n$ matrices whose last column is zero is a left ideal, but not a right ideal.

Theorem 6.9 (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 $a \in \mathbb{F}$.

6.5 Quotient Rings

What is nice about ideals is that they induce an equivalence relation defined on a ring, which reminds you of working in modulos on the integers.

Theorem 6.10 (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 equivalent $(\text{mod } I)$, written $a \equiv b \pmod{I}$ iff $a - b \in I$. We claim two things:

1. $\equiv$ is indeed an equivalence relation.
2. Given that $a \equiv a' \pmod{I}$ and $b = b' \pmod{I}$,

$$a + b \equiv a' + b' \pmod{I}, \qquad ab \equiv a'b' \pmod{I} \tag{62}$$

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. *Transitive.* If $a \equiv b$.

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 6.15 (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

$$[a] + [b] := [a + b], \qquad [ab] := [a]\,[b] \tag{63}$$

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.

$$\begin{aligned} 0 \in M &\implies [0] \text{ is the additive identity in } \frac{M}{R} \\ a + (-a) = 0 &\implies [a] + [-a] = [0] \\ 1 \in M &\implies [1] \text{ is the multiplicative identity in } \frac{M}{R} \end{aligned}$$

Example 6.13 (Quotient Rings of Integers)

The quotient set $\mathbb{Z}/\langle n \rangle$ by the relation of congruence modulo n is denoted $\mathbb{Z}_n$.

$$\mathbb{Z}_n = \{[0]_n, [1]_n, \ldots, [n-1]_n\} \tag{64}$$

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 6.14 (Quotient Rings of Polynomials)

We list some quotient rings of the integers.

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 6.11 (Quotient Inherits Commutativity)

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

Example 6.15 (Quotient Does Not Inherit Integral Domain Property)

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

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 6.12 (Integer Quotient Rings as Finite Fields)

The ring $\mathbb{Z}_n$ 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

$$[k]_n [l]_n = [kl]_n = [n]_n = 0 \tag{65}$$

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

$$[0]_n, [a]_n, [2a]_n, ..., [(n-1)a]_n \tag{66}$$

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 6.1 (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.

Theorem 6.13 (Wilson's Theorem)

Let n be a prime number. Then

$$(n-1)! \equiv -1 \pmod{n} \tag{67}$$

6.6 Homomorphism and Isomorphism Theorems

Theorem 6.14 (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

$$R/\ker f \simeq S \tag{68}$$

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 8: The theorem states that the following diagram commutes.

Proof.

6.7 Unique Factorization Domain

6.8 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 6.16 (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 6.17 (Principal Ideal Domain)

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

More generally, a **principal ideal ring** is a nonzero commutative ring in which every ideal is principal (i.e. can be generated by a single element). The distinction is that a principal ideal ring may have zero divisors whereas a principal ideal domain cannot. Principal ideal domains are thus mathematical objects that behave somewhat like the integers. That is,

1. Any element of a PID has a unique decomposition into prime elements.
2. Any two elements of a PID have a greatest common divisor.
3. If x and y are elements of a PID without common divisors, then every element of the PID can be written in the form
$$ax + by \tag{69}$$

We now introduce some examples of PIDs, which are not as trivial and should be introduced as theorems.

Theorem 6.15 (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 6.2 (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 6.3 (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
$$\mathrm{ev}_\alpha : F[x] \to K \tag{70}$$

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

Proof.

This is an immediate consequence of the previous corollary.

The great thing about PIDs is that they unlock a lot of the familiar properties that we see in the integers. In fact, pretty much everything holds except for the existence of Euclidean algorithm for factorization.

Theorem 6.16 (Greatest Common Divisor)

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

Theorem 6.17 (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.

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

Theorem 6.18 (Bezout's Theorem)

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

$$f(x) = (x - c)q(x) + r, \; r \in F \tag{71}$$

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

$$f(c) = r \tag{72}$$

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

6.9 Euclidean Domains

Definition 6.18 (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

$$\mathcal{D}(a, b) = (q, r) \text{ where } a = bq + r \text{ and } 0 \leq r < |b| \tag{73}$$

The two prime examples are the integers and polynomials.

Example 6.16 (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 6.19 (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

$$f(x) = q(x)g(x) + r(x), \qquad 0 \leq \deg(r) < \deg(g) \tag{74}$$

where deg is the norm.

Example 6.17 (Gaussian Integers)

The subring of $\mathbb{C}$, defined

$$\mathbb{Z}[i] \equiv \{a + bi \mid a, b \in \mathbb{Z}\} \tag{75}$$

is a Euclidean integral domain with respect to the norm

$$N(c) \equiv a^2 + b^2 \tag{76}$$

since $N(cd) = N(c)N(d)$ and the invertible elements of $\mathbb{Z}[i]$ are $\pm 1, \pm i$.

Example 6.18 (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

$$m = \pm \prod_i p_i^{k_i}, \; p \text{ prime} \tag{77}$$

and the norm is defined

$$N(\frac{m}{2^n}) \equiv 1 + \sum_i k_i \tag{78}$$

We must further show that division with remainder is possible, but we will not show it here.

Theorem 6.20 (Chinese Remainder Theorem)

6.10 Division Rings

Definition 6.19 (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 6.7 (Division Rings are Domains)

Every division ring R is automatically a domain.

Proof.

Every nonzero element is invertible.

Example 6.19 (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.

$$aa^{-1} = a^{-1}a = I \tag{79}$$

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.

6.11 Fields

Our final structure is field, which are usually pretty tame compared to groups and rings.

Definition 6.20 (Field)

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

Theorem 6.21 ()

Every field is a Euclidean domain.

Proof.

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

$$0 = x^{-1}0 = x^{-1}(xy) = y \tag{80}$$

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

$$0 = 0y^{-1} = (xy)y^{-1} = x \tag{81}$$

Let's give a few examples of fields.

Theorem 6.22 (Wedderburn's little theorem)

Every finite Euclidean domain is a field.

Example 6.20 (Finite Fields)

$\mathbb{Z}_p$ with p prime is a field.

Example 6.21 (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 6.8 (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 $\mathbb{Q}$.

7 Polynomial Rings

One of the most widely studied rings are the ring of polynomials. Let's reintroduce them.

Definition 7.1 (Univariate Polynomials)

For a ring R, the **univariate polynomial ring over** R, denoted $R[x]$ consists of elements called **polynomials** which are formal expressions of the form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 \text{ where } a_i \in R \tag{82}$$

with coefficients $a_i \in R$ and x is called a **variable**, or **indeterminant**.[a] Two polynomials are equal if and only if the sequences of their corresponding coefficients are equal. We can also see a polynomial as a function $f : R \to R$ as well.
Furthermore, $R[x]$ is a ring, with addition and multiplication defined

$$a_i x^i + b_i x^i = (a_i + b_i) x^i, \qquad x^i x^j = x^{i+j} \tag{83}$$

along with 0 as the additive identity and 1 as the multiplicative identity.

[a] 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. $(a_0, a_1, a_2, ..., a_n)$.

While we will mainly deal with univariate polynomials, we can also define multivariate polynomials similarly.

Definition 7.2 (Multivariate Polynomials)

For a ring R, the **multivariate polynomial ring over** R, denoted $R[x_1, \dots, x_n]$ consists of elements called **polynomials** which are formal expressions of the form

$$f(x_1, \dots, x_n) = \sum_{0 \leq k_i \leq n} a_{k_1 \dots k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \tag{84}$$

with coefficients $a \in R$ and x_i's the **variables**. We can treat an element $f \in R[x_1, \dots, x_n]$ as a function $f : R^n \to R$.
Furthermore, $R[x_1, \dots, x_n]$ is a ring, with addition and multiplication defined

$$a_{k_1 \dots k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} + b_{k_1 \dots k_n} x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} = (a_{k_1 \dots k_n} + b_{k_1 \dots k_n}) x_1^{k_1} x_2^{k_2} \dots x_n^{k_n} \tag{85}$$

$$x^{k_1 \dots k_n} x^{l_1 \dots l_n} = x^{k_1 + l_1, k_2 + l_2, \dots, k_n + l_n} \tag{86}$$

Usually, the properties of the base ring R determines the properties on R.

Lemma 7.1 (Commutativity Extends to Polynomials)

We have the following.

1. R is a commutative ring iff $R[x]$ is a commutative ring.
2. R is an integral domain iff $R[x]$ is an integral domain.
3. F is a field iff $F[x]$ is a Euclidean domain.

Proof.

TBD.

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, so let's get that out of the way. Before we move on, let's get some terms out of the way.

Definition 7.3 (Some Terms for Polynomials)

Given a univariate polynomial $f(x)$.
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.

7.1 Basic Properties of Polynomials

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.[3] 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 7.1 (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

$$f(x) \cdot g(x) = (2x+4)(3x+3) = 6x^2 + 18x + 12 = 0 \tag{87}$$

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

Theorem 7.1 (Bounds on Degrees From Operations)

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

$$\deg(f+g) \leq \max\{\deg f, \deg g\} \tag{88}$$

If R is a domain, then

$$\deg(fg) = \deg f + \deg g \tag{89}$$

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 7.2 (Polynomials as Same Function)

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

$$f(x) = x \sim g(x) = x^2 \tag{90}$$

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

[3] x and x^2 are equivalent in the polynomial algebra defined on the domain $\mathbb{Z}_2$.

some special theorems.

Theorem 7.2 (Uniqueness of Polynomials over Field)

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

7.2 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 7.3 (Polynomials as Euclidean Domain)

Given a field F, $F[x]$ 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.

$$f(x) = q(x)g(x) + r(x), \qquad \deg(r(x)) < \deg(g(x)) \tag{91}$$

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

$$f(x) = a_1 x + a_0, \qquad g(x) = b_1 x + b_0 \tag{92}$$

and we can find the solutions

$$f(x) = \frac{a_1}{b_1} g(x) + \left(a_0 - \frac{a_1 b_0}{b_1}\right) \tag{93}$$

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

$$f(x) = F(x) - \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) \tag{94}$$

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

$$\begin{aligned} F(x) &= f(x) + \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) && (95) \\ &= q(x)g(x) + r(x) + \frac{a_{k+1}}{b_m} x^{k+1-m} g(x) && (96) \\ &= \left(q(x) + \frac{a_{k+1}}{b_m} x^{k+1-m}\right) g(x) + r(x) && (97) \end{aligned}$$

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

$$f(x) = q(x)g(x) + r(x) = q'(x)g(x) + r'(x) \implies \big(q(x) - q'(x)\big)g(x) = r(x) - r'(x) \tag{98}$$

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 7.3 (Polynomials over Fields)

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

$$\begin{array}{r|rrrr}
 & x^2 & +6x & +11 & \\
\hline
x-2 & x^3+4x^2 & -x & +7 & \\
 & -x^3+2x^2 & & & \\
\hline
 & 6x^2 & -x & & \\
 & -6x^2+12x & & & \\
\hline
 & & 11x & +7 & \\
 & & -11x & +22 & \\
\hline
 & & & 29 &
\end{array}$$

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 7.1 (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,

$$f(x) = (x-c)q(x) + r(x) \implies f(c) = (c-c)q(c) + r(c) = r(c) \tag{99}$$

7.3 Roots and Factorization

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

Definition 7.4 (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

$$f(x) = g(x)h(x) \tag{100}$$

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 7.5 (Polynomial Root)

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

$$f(r) = 0 \tag{101}$$

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

Theorem 7.4 (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

$$f(x) = (x-c)q(x) \tag{102}$$

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 7.5 (GCD of Two Polynomials Exist)

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

$$S = \{h(x) \in F[x] \mid h(x) = a(x)f(x) + b(x)g(x) \text{ for some } a(x), b(x) \in F[x]\} \tag{103}$$

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

$$f(x) = q(x)d(x) + r(x) \tag{104}$$

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

$$\begin{aligned} r(x) &= f(x) - q(x)d(x) && (105)\\ &= f(x) - \big(s(x)f(x) + t(x)g(x)\big)q(x) && (106)\\ &= \big(1 - s(x)q(x)\big)f(x) - \big(t(x)q(x)\big)g(x) \in S && (107) \end{aligned}$$

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 7.6 (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 7.2 ()

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 7.6 (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 7.4 (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

$$x^2 - 1 = (x+1)(x-1) = (x+3)(x-3) \tag{108}$$

Theorem 7.7 (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

$$f(x_i) = y_i, \quad i = 1, 2, ..., n \tag{109}$$

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

7.4 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 7.7 (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 7.8 (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 7.5 (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.[4]

Theorem 7.8 (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

$$f(z) = z^n \left(1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n}\right) \tag{110}$$

Since

$$\lim_{|z| \to \infty} \left(1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n}\right) = 0 \tag{111}$$

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

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

$$|z| > R \implies \left|1 + \frac{a_{n-1}}{z} + \frac{a_{n-2}}{z^2} + \ldots + \frac{a_0}{z_n}\right| < \frac{1}{2} \tag{112}$$

and hence

$$|z| > R \implies |f(z)| > |z|^n \cdot \left(1 - \frac{1}{2}\right) > \frac{R^n}{2} \tag{113}$$

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

$$g(z) = \frac{a_{j+1}}{a_j} z + \ldots + \frac{a_n}{a_j} z^{n-j} \implies f(z) = a_0 + a_j z^j \big(1 + g(z)\big) \tag{114}$$

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

$$\begin{aligned} f(t\gamma) &= a_0 + a_j (t\gamma)^j \big(1 + g(t\gamma)\big) & (115)\\ &= a_0 - a_0 t^j \big(1 + g(t\gamma)\big) & (116)\\ &= a_0 \big\{1 - t^j \big(1 + g(t\gamma)\big)\big\} & (117) \end{aligned}$$

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

$$|g(t\gamma)| = \left|\frac{a_{j+1}}{a_j}(t\gamma) + \ldots + \frac{a_n}{a_j}(t\gamma)^{n-j}\right| < \frac{1}{2} \tag{118}$$

and for such t, this implies

$$|f(t\gamma)| = |a_0||1 - t^j(1 + g(t\gamma))| \leq |a_0||1 - t^j/2| < |a_0| \tag{119}$$

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 7.2 ($\mathbb{C}$ 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 7.9 (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

$$A = \begin{pmatrix} 0 & 0 & ... & 0 & -a_0 \\ 1 & 0 & ... & 0 & -a_1 \\ 0 & 1 & ... & 0 & -a_2 \\ ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & -a_{n-1} \end{pmatrix} \tag{120}$$

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 7.10 (Viete's Formulas)

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

$$f(x) = a_0 \prod_{i=1}^n (x - c_i), c_i \text{ roots of } f \tag{121}$$

Then the coefficients of f can be presented with the formulas

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

7.5 Reducibility of Real Polynomials

Theorem 7.11 ()

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 7.3 ()

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

Example 7.6 ()

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

Corollary 7.4 ()

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

$$\lim_{x \to +\infty} f(x) = +\infty, \;\; \lim_{x \to -\infty} f(x) = -\infty \tag{122}$$

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

Theorem 7.12 (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 7.13 ()

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

Example 7.7 (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.

7.6 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 7.14 (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

$$a_n (r/s)^n + \ldots + a_0 = 0 \tag{123}$$

Multiplying by s^n, we get

$$a_n r^n + a_{n-1} r^{n-1} s + \ldots + a_1 s^{n-1} r + a_0 s^n = 0 \tag{124}$$

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 7.3 (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

$$0 \equiv \bar{k}\bar{l}\bar{f}(x) \equiv \bar{g}_1(x) \bar{h}_1(x) \text{ in } \mathbb{Z}_p[x] \tag{125}$$

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 $g_2(x) = p g_1(x)$. Therefore,

$$p(x) \cdot \frac{kl}{p} = \underbrace{\frac{g_1(x)}{p}}_{g_2(x)} \cdot \underbrace{h_1(x)}_{h_2(x)} \iff f(x) \frac{kl}{p} = g_2(x) h_2(x) \tag{126}$$

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

$$f(x) = g_n(x) h_n(x), \qquad g_n, h_n \in \mathbb{Z}[x] \tag{127}$$

Example 7.8 (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

$$f(x) = (x^2 + ax + b)(x^2 + cx + d), \qquad a, b, c, d \in \mathbb{Z} \tag{128}$$

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 7.15 ()

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

$$f(x) \equiv g(x)h(x) \text{ in } \mathbb{Z}_p[x] \tag{129}$$

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 7.9 ()

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 7.16 (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,

$$f(x) = g(x)h(x) \text{ in } \mathbb{Z}_p[x] \tag{130}$$

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

$$g(x) = b_m x^m, h(x) = \frac{a_n}{b_m} x^{n-m} \in \mathbb{Z}_p[x] \tag{131}$$

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 7.10 (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

$$\begin{aligned} f(x+1) &= \frac{(x+1)^5 - 1}{(x+1) - 1} && (132) \\ &= \frac{1}{x}\left(x^5 + \binom{5}{1}x^4 + \binom{5}{2}x^3 + \binom{5}{3}x^2 + \binom{5}{4}x + \binom{5}{5} - 1\right) && (133) \\ &= x^4 + 5x^3 + 10x^2 + 10x + 5 && (134) \end{aligned}$$

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.

7.7 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 7.9 (Rational Functions)

The **rational functions** are defined to be the field of quotients (really just 2-tuples) of the form

$$F(x) := \left\{ \frac{f(x)}{g(x)} \;\middle|\; f(x), g(x) \in F[x], g(x) \neq 0 \right\} \tag{135}$$

where addition and multiplication is defined in the usual sense.

Theorem 7.17 (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.

$$\frac{f(x)}{g(x)} = \frac{a(x)}{u(x)} + \frac{b(x)}{v(x)} \tag{136}$$

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,

$$\frac{f(x)}{u(x)v(x)} = \frac{f(x)t(x)}{u(x)} + \frac{f(x)s(x)}{v(x)} \tag{137}$$

and we can use the Euclidean algorithm to write

$$\frac{f(x)t(x)}{u(x)} = q(x) + \frac{a(x)}{u(x)}, \qquad \deg(a) < \deg(u) \tag{138}$$

$$\frac{f(x)s(x)}{v(x)} = q(x) + \frac{a(x)}{u(x)}, \qquad \deg(b) < \deg(v) \tag{139}$$

which implies

$$\frac{f(x)}{u(x)v(x)} = \frac{a(x)}{u(x)} + \frac{b(x)}{v(x)} \tag{140}$$

Example 7.11 ()

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

$$1 = (3x^2 + 2x + 1)(x-1)^2 - (3x-4)x^3 \tag{141}$$

and so

$$\frac{x+3}{x^3(x-1)^2} = \frac{(x+3)(3x^2+2x+1)}{x^3} - \frac{(x+3)(3x-4)}{(x-1)^2} \tag{142}$$

$$= \frac{11x^2+7x+3}{x^3} + \frac{-11x+15}{(x-1)^2} \tag{143}$$

7.8 Exercises

8 Modules

8.1 Modules over a PID

8.2 Rational Canonical Form

8.3 Jordan Canonical Form

9 Vector Space Structures

Definition 9.1 (Vector Space)

A **vector space over a field** F consists of an abelian group $(V, +)$ and an operation called **scalar multiplication**

$$\cdot : F \times V \to V \tag{144}$$

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 9.2 ()

A **left R-module** M consists of an abelian group $(M, +)$ and an operation called **scalar multiplication**

$$\cdot : R \times M \longrightarrow M \tag{145}$$

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 9.3 ()

A **right R-module** M is defined analogously to a left R-module, except that the scalar multiplication operation is defined

$$\cdot : M \times R \longrightarrow M \tag{146}$$

Definition 9.4 ()

Let A be a vector space over a field F equipped with an additional binary operation

$$\times : A \times A \longrightarrow A \tag{147}$$

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 9.1 ()

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 9.2 ()

$\mathbb{R}^3$ 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 9.3 ()

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 9.5 ()

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

$$A = bx \tag{148}$$

that are unique. If we wish, we can distinguish left and right division to be the solutions of $A = bx$ and $A = xb$.

Definition 9.6 ()

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 $\mathbb{R}$.

9.1 Modules

Vector space but over a ring.

9.2 Algebras

Vector space with bilinear product.

9.3 The Algebra of Quaternions

Definition 9.7 ()

The **quaternions** form an algebra of 4-dimensional vectors over $\mathbb{R}$, with elements of the form

$$(a, b, c, d) \equiv a + bi + cj + dk \tag{149}$$

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

$$i^2 = j^2 = k^2 = -1 \tag{150}$$

The algebra of quaternions are in fact the first noncommutative algebra to be discovered!

Proposition 9.1 ()

$\mathbb{H}$ and $\mathbb{C}$ are the only finite-dimensional divisions rings containing $\mathbb{R}$ as a proper subring.

Definition 9.8 ()

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

$$Q_8 = \big\langle \bar{e}, i, j, k \mid \bar{e}^2 = e, i^2 = j^2 = k^2 = ijk = \bar{e} \big\rangle \tag{151}$$

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.

$$(a + bi + cj + dk)^{-1} = \frac{1}{a^2 + b^2 + c^2 + d^2}\big(a - bi - cj - dk\big) \tag{152}$$

Note that property 3 allows $\mathbb{H}$ to be a division algebra.

Proposition 9.2 (Scalar and Vector Components)

Let the quaternion be divided up into a scalar and vector part with the bjective mapping $a + bi + cj + dk \mapsto \big(a, (b, c, d)\big)$.

$$q = (r, v), r \in \mathbb{R}, v \in \mathbb{R}^3 \tag{153}$$

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 9.9 ()

The conjugate of a quaternion $q = a + bi + cj + dk$ is defined

$$\bar{q}, q^* \equiv a - bi - cj - dk \tag{154}$$

It has properties
1. $q^{**} = q$
2. $(qp)^* = p^* q^*$

q^* can also be expressed in terms of addition and multiplication.

$$q^* = -\frac{1}{2}\big(q + iqi + jqj + kqk\big) \tag{155}$$

Definition 9.10 ()

The **norm** of q is defined

$$||q|| \equiv \sqrt{q^* q} = \sqrt{q q^*} = \sqrt{a^2 + b^2 + c^2 + d^2} \tag{156}$$

with properties
1. Scaling factor. $||\alpha q|| = |\alpha| ||q||$
2. Multiplicative. $||pq|| = ||p|| ||q||$

The norm allows us to define a metric

$$d(p, q) \equiv ||p - q|| \tag{157}$$

This makes $\mathbb{H}$ a metric space, with addition and multiplication continuous on the metric topology.

Definition 9.11 ()

The **unit quaternion** is defined to be

$$U_q = \frac{q}{||q||} \tag{158}$$

Corollary 9.1 ()

Every quaternion has a polar decomposition

$$q = U_q \cdot ||q|| \tag{159}$$

With this, we can redefine the inverse as

$$q^{-1} = \frac{q^*}{||q||^2} \tag{160}$$

9.3.1 Matrix Representations of Quaternions

We can represent q with 2×2 matrices over $\mathbb{C}$ or 4×4 matrices over $\mathbb{R}$.

Proposition 9.3 ()

The following representation is an injective homomorphism $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(2, \mathbb{C})$.

$$\rho : a + bi + cj + dk \mapsto \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix} \tag{161}$$

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.

$$||q|| = \sqrt{\det \begin{pmatrix} a + bi & c + di \\ -c + di & a - bi \end{pmatrix}} = \sqrt{(a^2 + b^2) + (c^2 + d^2)} \tag{162}$$

3. The conjugate of a quaternion corresponds to the conjugate (Hermitian) transpose of its matrix representation.

$$\rho(q^*) = \rho(q)^H \iff a - bi - cj - dk \mapsto \begin{pmatrix} a - bi & -c - di \\ c - di & a + bi \end{pmatrix} \tag{163}$$

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 SU(2). Topologically, the unit quaternions is the 3-sphere, so the underlying space SU(2) is also a 3-sphere. More specifically,

$$\frac{\mathrm{SU}(2)}{2} \simeq \mathrm{SO}(3) \tag{164}$$

Proposition 9.4 ()

The following representation of $\mathbb{H}$ is an injective homomorphism $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(4, \mathbb{R})$.

$$\rho : a + bi + cj + dk \mapsto \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \tag{165}$$

or also as

$$a \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{166}$$

It has properties

1. $\rho(q^*) = \rho(q)^T$
2. The fourth power of the norm is the determinant of the matrix

$$||q||^4 = \det\big(\rho(q)\big) \tag{167}$$

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.

9.3.2 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

$$a^2 - b^2 - c^2 - d^2 = -1, 2ab = 2ac = 2ad = 0 \tag{168}$$

To satisfy the second equation, either $a = 0$ or $b = c = d = 0$. The latter is impossible since then q would be real. Therefore,

$$b^2 + c^2 + d^2 = 1 \tag{169}$$

which forms the unit sphere in $\mathbb{R}^3$.

9.4 Tensor Algebras

Remember that an algebra is (loosely) a vector space V with a multiplication operation

$$\times : V \times V \longrightarrow V \tag{170}$$

Definition 9.12 ()

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

$$(a, B^{\mu}, C^{\nu\gamma}, D^{\alpha\beta\epsilon}, ...) \tag{171}$$

The addition operation is defined component-wise, and the multiplication operation is the tensor product

$$\otimes : T(V) \times T(V) \longrightarrow T(V) \tag{172}$$

and the identity element is

$$I = (1, 0, 0, ...) \tag{173}$$

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

$$i_j : V^{\otimes j} \longrightarrow T(V),\ i_j(T^{\kappa_1,...,\kappa_j}) = (0, ..., 0, T^{\kappa_1,...,\kappa_j}, 0, ..., 0) \tag{174}$$

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

$$A + B \equiv i_n(A) + i_m(B) \in T(V) \tag{175}$$

along with multiplication of tensors as

$$A \otimes B \equiv i_n(A) \otimes i_m(B) \equiv i_{n+m}(A \otimes B) \tag{176}$$

We can also redefine the tensor product operation between two spaces to be an operation within $T(V)$ itself.

$$i_i(V^{\otimes i}) \otimes i_j(V^{\otimes j}) = i_{i+j}(V^{\otimes(i+j)}) \tag{177}$$

We can now proceed to define Exterior and Symmetric algebras as quotient algebras.

Definition 9.13 ()

The **exterior algebra** $\Lambda(V)$ of a vector space V over field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

$$\Lambda(V) \equiv \frac{T(V)}{I} \tag{178}$$

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

$$\pi : T(V) \longrightarrow \Lambda(V) \tag{179}$$

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

$$a \wedge b = \pi(\alpha \otimes \beta) \tag{180}$$

We can define this quotient space with the equivalence class

$$x \otimes y = -y \otimes x \pmod I \tag{181}$$

Definition 9.14 ()

The **symmetric algebra** $\mathrm{Sym}(V)$ of a vector space V over a field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

$$\Lambda(V) \equiv \frac{T(V)}{J} \tag{182}$$

where J is the two-sided ideal generated by all elements in the form

$$v \otimes w - w \otimes v \tag{183}$$

(i.e. commutators of all possible pairs of vectors).

10 Field and Galois Theory

10.1 Extensions and Splitting Fields

Great, so by establishing the fact 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 $\mathbb{C}[x]$. Let's define this a bit more generally for arbitrary fields $F \subset K$.

Definition 10.1 (Field Extension)

The pair of fields $F \subset K$ is called a **field extension**.

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 10.2 (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 10.1 (Don't Need(?) Complex)

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 $\pm\sqrt{2} \notin \mathbb{Q}$, 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 smallest such field is called an *adjoining field.*

10.2 Finite Fields

11 Affine and Projective Spaces

11.1 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,

$$\mathrm{GL}(V) \subset \mathrm{GA}(V) \tag{184}$$

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 11.1 ()

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

$$\overline{pq} + \overline{qr} = \overline{pr} \text{ for all } p, q, r \in S \tag{185}$$

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 11.2 ()

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 11.3 ()

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 11.4 ()

The **barycentric linear combination** of points $p_1, ..., p_k \in S$ is a linear combination of the form

$$p = \sum_i \lambda_i p_i, \text{ where } \sum_i \lambda_i = 1 \tag{186}$$

This linear combination is equal to the point p such that

$$\overline{op} = \sum_i \lambda_i \overline{op_i} \tag{187}$$

where $o \in S$ is any origin point.

Definition 11.5 ()

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 11.6 ()

Let $p_0, p_1, ..., p_n$ be points of an n-dimensional affine space S such that the vectors $\overline{p_0 p_1}, ..., \overline{p_0 p_n}$ are linearly independent (that is, forms a basis). Then, every point $p \in S$ can be uniquely presented as

$$p = \sum_{i=0}^n x_i p_i, \text{ where } \sum_{i=0}^n x_i = 1 \tag{188}$$

This equality can be rewritten

$$\overline{p_0 p} = \sum_{i=1}^n x_i \overline{p_0 p_i} \tag{189}$$

implying that we can take the coordinates of the vector $\overline{p_0 p}$ in the basis $\{\overline{p_0 p_1}, ..., \overline{p_0 p_n}\}$ as $x_1, ..., x_n$. Then, x_0 is determined as

$$x_0 = 1 - \sum_{i=1}^n x_i \tag{190}$$

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 11.7 ()

A **plane** in an affine space S is a subset of the form

$$p = p_0 + U \tag{191}$$

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 11.1 ()

If the intersection of two planes in an affine space is nonempty, then the intersection is also a plane.

Theorem 11.1 ()

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 11.8 ()

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.

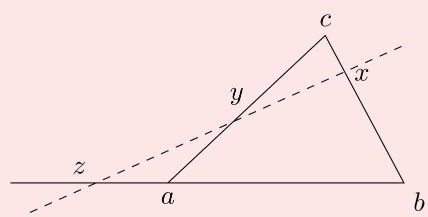
Theorem 11.2 ()

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 11.1 (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

$$\begin{pmatrix} 0 & \frac{1}{\lambda+1} & \frac{\lambda}{\lambda+1} \\ \frac{\mu}{\mu+1} & 0 & \frac{1}{\mu+1} \\ \frac{1}{\nu+1} & \frac{\nu}{\nu+1} & 0 \end{pmatrix} \tag{192}$$

has nonzero determinant. The determinant of the above matrix is 0 if and only if $\lambda \mu \nu = -1$.

Corollary 11.2 (Ceva's Theorem)

In the triangle above, the lines ax, by, cz intersect at one point if and only if

$$\lambda\mu\nu = 1 \tag{193}$$

Proof.

The proof can be done using barycentric coordinates.

Theorem 11.3 ()

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 11.4 ()

Given an inhomogeneous system of linear equations of form

$$Ax = b \tag{194}$$

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

$$Ax = 0 \tag{195}$$

Let us observe the relative position of two planes.

Theorem 11.5 ()

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

$$\overline{p_1 p_2} \subset U_1 + U_2 \tag{196}$$

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 11.9 ()

An **affine-linear** function on an affine space S is a function $f: S \longrightarrow \mathbb{F}$ such that

$$f(p + x) = f(p) + \alpha(x), \;\; p \in S, x \in V \tag{197}$$

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

$$f(x) = \alpha(x) + b, \;\; b \in \mathbb{F} \tag{198}$$

where $b = f(o)$. This implies the following coordinate form of f.

$$f(x) = b + \sum_i a_i x_i \tag{199}$$

A particular case of affine-linear functions are constant functions, where the defining characteristic is the zero differential.

Proposition 11.1 ()

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 11.2 ()

Barycentric coordinates are affine-linear functions.

Proposition 11.3 ()

Let f be an affine-linear function. Then

$$f\left(\sum_i \lambda_i p_i\right) = \sum_i \lambda_i f(p_i) \tag{200}$$

for any barycentric linear combination $\sum_i \lambda_i p_i$ of points $p_1, ..., p_k$.

Definition 11.10 ()

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

$$\rho(p, q) = ||\overline{pq}|| \tag{201}$$

This definition of ρ satisfies the axioms of a metric space.

11.2 Convex Sets

Let S be an affine space over the field of real numbers and V, the associated vector space.

Definition 11.11 ()

The **(closed) interval** connecting points $p, q \in S$ is the set

$$pq = \{\lambda p + (1-\lambda) q \mid 0 \leq \lambda \leq 1\} \tag{202}$$

Geometrically, we can think of this as the straight line segment connecting point p with point q.

Definition 11.12 ()

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 11.13 ()

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 11.4 ()

For any points $p_0, ..., p_k$ in a convex set $M \subset S$, the set M also contains every convex linear combination

$$p = \sum_i \lambda_i p_i \tag{203}$$

Furthermore, for any set $M \subset S$, the set $\operatorname{conv} M$ of all convex linear combinations of points in M is convex.

Definition 11.14 ()

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 11.15 ()

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 11.1 ()

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 11.5 ()

A convex set M has interior points if and only if $\operatorname{aff} M = S$.

Definition 11.16 ()

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 11.17 ()

For any nonconstant affine-linear function f on the set S, let

$$\begin{aligned} H_f &\equiv \{p \in S \mid f(p) = 0\} \\ H_f^+ &\equiv \{p \in S \mid f(p) \geq 0\} \\ H_f^- &\equiv \{p \in S \mid f(p) \leq 0\} \end{aligned}$$

The set H_f is a hyperplane, and H_f^+, H_f^- are called **closed half spaces**.

Definition 11.18 ()

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 11.6 ()

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 11.6 (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 11.7 ()

Every closed convex set M is an intersection of (perhaps infinitely many) half-spaces.

Definition 11.19 ()

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 11.2 ()

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 11.3 ()

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 11.20 ()

A point p of a convex set M is **extreme** if it is not an interior point of any interval in M.

Theorem 11.8 ()

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 11.9 (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 11.21 ()

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 11.10 ()

Every face Γ of the polyhedron M is of the form

$$\Gamma = M \cap \bigg(\bigcap_{j \in J} H_{f_j} \bigg) \tag{204}$$

where $J = \{1, 2, ..., m\}$

Proposition 11.7 ()

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 11.11 ()

The maximum of an affine-linear function on a bounded convex polyhedron M is attained at a vertex.

11.3 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 11.22 ()

An **affine map** from the space S to the space S' is a map $f : S \longrightarrow S'$ such that

$$f(p+x) = f(p) + \varphi(x), \quad p \in S, x \in V \tag{205}$$

for some linear map $\varphi : V \longrightarrow V'$. It follows that

$$\varphi(\overline{pq}) = \overline{f(p)f(q)}, \quad p, q \in S \tag{206}$$

Thus, f determines the linear map φ uniquely. Similarly, φ is called the **differential** of f, denoted df.

Proposition 11.8 ()

Let $f : S \longrightarrow S'$ and $g : S' \longrightarrow S''$ be two affine maps. Then the map

$$g \circ f : S \longrightarrow S'' \tag{207}$$

is also affine. Also

$$d(g \circ f) = dg \cdot df \tag{208}$$

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 11.9 ()

An affine map is bijective if and only if its differential is bijective.

Definition 11.23 ()

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 11.3 ()

Finite-dimensional affine spaces over the same field are isomorphic if and only if they have the same dimension.

Definition 11.24 ()

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

$$d : \mathrm{GA}(S) \longrightarrow \mathrm{GL}(V) \tag{209}$$

is a group homomorphism. It's kernel is the group of parallel translations, called $\mathrm{Tran}(S)$.

$$t_a : p \mapsto p + a, \quad a \in V \tag{210}$$

Proposition 11.10 ()

For any $f \in \text{GA}(S)$ and $a \in V$,

$$f t_a f^{-1} = t_{df(a)} \tag{211}$$

Definition 11.25 ()

A **homothety** with the center o and coefficient λ is an affine transformation defined as

$$f(o + x) \equiv o + \lambda x \tag{212}$$

In its vectorized form, it is expressed

$$f(x) = \lambda x + b, \;\; b \in V \tag{213}$$

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 11.12 ()

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

$$f(x) = \varphi(x) + \overline{p_0 q_0} \tag{214}$$

Corollary 11.4 ()

In real affine geometry all parallelopipeds are equal.

Definition 11.26 ()

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 $\text{Isom}\, S$. A motion is called **proper (orientation preserving)** if its differential belongs to $\text{SO}(V)$ and improper otherwise.

Lemma 11.2 ()

The group $\text{Isom}\, S$ is generated by reflections through hyperplanes.

Definition 11.27 ()

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 11.28 ()

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

$$\operatorname{Sym} M \equiv \{f \in \operatorname{Isom} S \mid f(M) = M\} \tag{215}$$

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 11.29 ()

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.

11.4 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 11.30 ()

An **affine-quadratic function** on an affine space S is a function $Q : S \longrightarrow \mathbb{F}$ such that its vectorized form is

$$Q(x) = q(x) + l(x) + c \tag{216}$$

for a quadratic function q, linear function l, and constant c.

11.5 Projective Spaces

Definition 11.31 ()

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 11.32 ()

$\mathbb{RP}^1$ is called the real projective line, which is topologically equivalent to a circle.

Example 11.4 ()

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 11.5 ()

$\mathbb{RP}^3$ is diffeomorphic to SO(3).

Example 11.6 ()

The space $\mathbb{RP}^n$ is formed by taking the quotient of $\mathbb{R}^{n+1} \setminus \{0\}$ under the equivalence relation

$$x \sim \lambda x \text{ for all real numbers } \lambda \neq 0 \tag{217}$$

The set of these equivalence classes is isomorphic to $\mathbb{RP}^n$.

12 Representations

We will assume that V is a finite-dimensional vector space over field $\mathbb{C}$.

Definition 12.1 ()

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 12.2 ()

A **representation** of an (algebraic) group $\mathcal{G}$ is a homomorphism

$$\rho : G \longrightarrow \mathrm{GL}(V) \tag{218}$$

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.

$$\rho(g_1 \cdot g_2) = \rho(g_1) \cdot \rho(g_2) \tag{219}$$

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

$$gv = \rho(g)v, \ v \in V \tag{220}$$

Since ρ is a group morphism, we have

$$g_2(g_1 v) = (g_2 g_1)v \iff \rho(g_2)\big(\rho(g_1)(v)\big) = \big(\rho(g_2)\rho(g_1)\big)(v) \tag{221}$$

Additionally, since g (that is, $\rho(g)$) is a linear map,

$$g(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 g v_1 + \lambda_2 g v_2 \tag{222}$$

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 12.1 ()

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.

$$\mathrm{GL}(2, \mathbb{C}) \xmapsto{id} \mathrm{Mat}(2, \mathbb{C}) \tag{223}$$

This is a trivial representation.

We now show a nontrivial representation of $\mathrm{GL}(2, \mathbb{C})$.

Example 12.2 ()

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

$$\{x_1 \odot x_1, x_1 \odot x_2, x_2 \odot x_2\} \tag{224}$$

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 $\mathrm{GL}(2,\mathbb{C})$

$$\mathcal{A} \equiv \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{225}$$

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,

$$\rho \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix} \tag{226}$$

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 12.3 ()

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 $g \in \mathcal{G}$

$$f \circ g = g \circ f \tag{227}$$

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.

$$\begin{array}{ccc} V_1 & \xrightarrow{\rho_1(g)} & V_1 \\ \big\downarrow f & & \big\downarrow f \\ V_2 & \xrightarrow{\rho_2(g)} & V_2 \end{array}$$

Definition 12.4 ()

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 $w \in W$,

$$\rho(g)(w) \in W \tag{228}$$

Example 12.3 ()

V and $\{0\}$ are always subrepresentations of V.

We now introduce the "building blocks" of all representations.

Definition 12.5 ()

A representation W is **irreducible representation** if $\{0\}$ and W are the only subrepresentations of W.

Lemma 12.1 (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,

$$f_1 = \lambda f_2 \tag{229}$$

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

$$f^{-1} \circ \rho_2(g) = \rho_1(g) \circ f^{-1} \tag{230}$$

To prove the second part, without loss of generality, assume that the first isomorphism is the identity mapping. That is,

$$f_1 = id \tag{231}$$

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

$$f_2(v) = \lambda v \tag{232}$$

Now, we define the map

$$f : V_1 \longrightarrow V_2,\ f \equiv f_2 - \lambda f_1 \tag{233}$$

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 12.1 (Mache's Theorem)

Let V be finite dimensional, with $\mathcal{G}$ a finite group. Then, V can be decomposed as

$$V = \bigoplus_i V_i \tag{234}$$

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

$$\pi : V \longrightarrow W \tag{235}$$

Then, we define the new map

$$\tilde{\pi} : V \longrightarrow W,\ \tilde{\pi}(v) \equiv \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \rho(g)\big|_W \circ \pi \circ \rho(g)^{-1} \tag{236}$$

This "averaging" of the group elements are done so that this mapping is a map of representations. This implies that

$$V = W \oplus \ker \tilde{\pi} \tag{237}$$

meaning that V can indeed be decomposed into direct sums of subrepresentations.

13 Lie Groups and Lie Algebras

Definition 13.1 ()

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

$$\mu : \mathcal{G} \times \mathcal{G} \to \mathcal{G}, \ \mu(x, y) = xy \tag{238}$$

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

$$(x, y) \mapsto x^{-1}y \tag{239}$$

be a smooth mapping of the product manifold into $\mathcal{G}$.

Definition 13.2 ()

A **Lie Algebra** is a vector space $\mathfrak{g}$ with an operation called the **Lie Bracket**

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{240}$$

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 13.1 ()

A common example of a Lie Braket in the algebra of matrices is defined

$$[A, B] \equiv AB - BA \tag{241}$$

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 13.2 ()

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

$$exp : \text{Mat}(n, \mathbb{C}) \to \text{mat}(n, \mathbb{C}), \ exp(A) = e^A = \sum_{p \geq 0} \frac{A^p}{p!} \tag{242}$$

Note that this value is always well defined. This lets us define

$$exp(tA) \equiv e^{tA} \equiv I + tA + \frac{1}{2}t^2A^2 + \frac{1}{3!}t^3A^3 + ... \tag{243}$$

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

$$exp : \mathfrak{g} \rightarrow \mathcal{G} \tag{244}$$

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,

$$\bigg(\frac{d}{dt}e^{tA}\bigg)\bigg|_{t=0} = \bigg(\sum_{k=0}^{\infty}\frac{1}{k!}t^kA^{k+1}\bigg)\bigg|_{t=0} = A$$

So, the mapping

$$\frac{d}{dt}\bigg|_{t=0} : \mathcal{G} \rightarrow \mathfrak{g} \tag{245}$$

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:

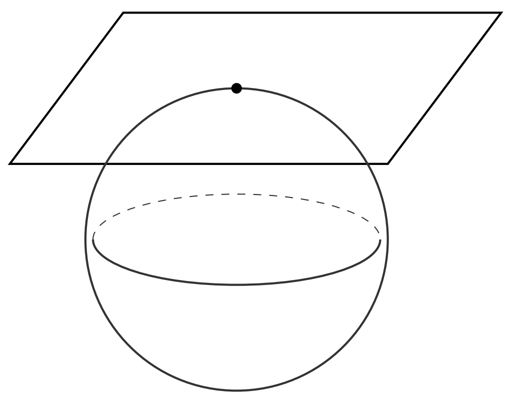
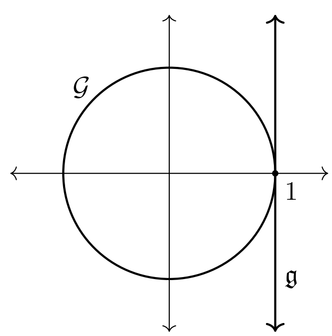


Figure 9: The Lie algebra can be visualized as the tangent space at the identity.

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

$$it \mapsto exp(it) \equiv e^{it} \equiv \cos t + i \sin t \tag{246}$$



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 13.3 ()

The exponential map
$$exp : \mathbb{R} \to \mathbb{R}^+, \ x \mapsto e^x \tag{247}$$
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 13.1 ()

If A and B are commuting square matrices, then
$$e^{A+B} = e^A \, e^B \tag{248}$$
In general, the solution C to the equation
$$e^A \, e^B = e^C \tag{249}$$
is given by the **Baker-Campbell-Hausdorff formula**, defined
$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + ... \tag{250}$$
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
$$\mathfrak{g} \equiv (V, [\cdot, \cdot]) \tag{251}$$
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
$$\cdot : \mathcal{G} \times \mathcal{G} \to \mathcal{G}, \ e^A \cdot e^B = e^{A*B} \tag{252}$$
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.

13.1 Lie Algebras of Classical Lie Groups

Definition 13.3 ()

The **general linear group** of vector space V is the group of all automorphisms of V, denoted $\mathrm{GL}(V)$. Additionally, $\mathrm{GL}(n, \mathbb{R})$ is the group of real $n \times n$ matrices with nonzero determinant, and $\mathrm{SL}(n, \mathbb{R})$ is the group of real $n \times n$ matrices with determinant $= 1$.

13.1.1 Lie Algebras of $\mathrm{SL}(2, \mathbb{R})$ and $\mathrm{SL}(2, \mathbb{C})$

Given the group $\mathrm{SL}(2, \mathbb{R})$, there must be a corresponding Lie algebra of matrices such that $g = e^A \in \mathrm{SL}(2, \mathbb{R})$. We attempt to find this Lie algebra. Let $g \in \mathrm{SL}(2, \mathbb{R})$, 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 13.1 ()

$$\det\left(e^A\right) = e^{\mathrm{Tr}\,(A)} \tag{253}$$

Proof.

Put A in Jordan Normal Form: $A = S^{-1}JS \implies A^n = S^{-1}J^nS \implies exp(A) = S^{-1}exp(A)S \implies \det\left(exp(A)\right) = \det e^J$. But since J is upper trianglar, J^n is upper triangular $\implies$ e^J is upper triangular, which implies that

$$\det e^J = \prod_i e^{\lambda_i} = e^{\mathrm{Tr}\,(J)} = e^{\mathrm{Tr}\,(A)} \tag{254}$$

since trace is invariant under a change of basis.

So, $\det\left(e^A\right) = 1 \implies \mathrm{Tr}\,(A) = 2\pi i n$ for $n \in \mathbb{Z}$. Since we want to component connected to the identity, we choose $n = 0$ meaning that $\mathrm{Tr}\,(A) = 0$. And we are done. That is, the Lie algebra of $\mathrm{SL}(2, \mathbb{R})$ consists of traceless 2×2 matrices, denoted $\mathfrak{sl}_2\mathbb{R}$. $\mathfrak{sl}_2\mathbb{R}$ has basis (chosen arbitrarily)

$$\left\{ H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \tag{255}$$

and the identity in the Lie algebra is the zero matrix, which translates to the 2×2 identity matrix in the Lie group.

$$exp \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = I \tag{256}$$

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,

$$XY = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{257}$$

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 $\mathrm{SL}(2, \mathbb{C})$ also has the same basis

$$\left\{ H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \tag{258}$$

but we choose the field to be $\mathbb{C}$, meaning that we take complex linear combinations rather than real linear ones.

13.1.2 Lie Algebra of $\mathrm{SU}(2)$

$g \in \mathrm{SU}(2) \implies \det g = 1 \implies \mathrm{Tr}\, A = 0$. We also see that by definition e^A,

$$(e^A)^\dagger = e^{A^\dagger} \text{ and } (e^A)^{-1} = e^{-A} \tag{259}$$

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 13.4 ()

The **Pauli matrices** are the three matrices

$$\left\{ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\} \tag{260}$$

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 $\mathrm{SU}(2)$ 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 $\mathrm{SU}(2)$ and $\mathrm{SL}(2, \mathbb{R})$ both leads to the construction of $\mathrm{SL}(2, \mathbb{C})$.

$$\begin{tikzcd} \mathrm{SL}(2, \mathbb{R}) \arrow[dr] & \\ & \mathrm{SL}(2, \mathbb{C}) \\ \mathrm{SU}(2) \arrow[ur, "\text{complexify}"'] & \end{tikzcd}$$

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,

$$\mathrm{SL}(2, \mathbb{R}) \neq \mathrm{SU}(2) \tag{261}$$

13.1.3 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

$$\left\{ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\} \tag{262}$$

as its only basis, since

$$exp\bigg(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \theta \bigg) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{263}$$

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

$$\{R_x, R_y, R_z\} \tag{264}$$

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,

$$exp(2\pi R_z) = \begin{pmatrix} \cos 2\pi & -\sin 2\pi & 0 \\ \sin 2\pi & \cos 2\pi & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \tag{265}$$

while

$$exp(2\pi A_z) = exp(-i\pi\sigma_z) = exp\bigg(-i\pi \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigg) = -I \tag{266}$$

There is discrepancy by a factor of -1. In fact, it turns out that

$$\mathrm{SO}(3) = \frac{\mathrm{SU}(2)}{\pm I} \tag{267}$$

We justify this in the following way. Let $v \in \mathbb{R}^3$ have components (x, y, z). Consider

$$M = x\sigma_x + y\sigma_y + z\sigma_z \tag{268}$$

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

$$x'\sigma_x + y'\sigma_y + z'\sigma_z \tag{269}$$

for some (x', y', z'). Now, since

$$M^2 = (-x^2 - y^2 - z^2)I \tag{270}$$

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

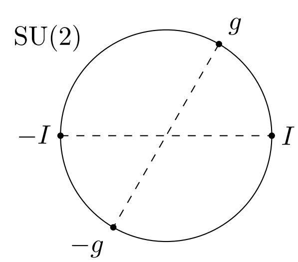
$$S = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \tag{271}$$

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

$$\rho : \text{SU}(2) \to \text{SO}(3) \tag{272}$$

in which 2 points are mapped to 1 point is a surjective map with

$$\ker \rho = \{I, -I\} \tag{273}$$



We can in fact explicitly describe exponential map from $\mathfrak{so}(3)$ to SO(3) with the following lemma.

Lemma 13.1 (Rodrigues' Formula)

The exponential map $exp : \mathfrak{so}(3) \to \text{SO}(3)$ is defined by

$$e^A = \cos\theta I_3 + \frac{\sin\theta}{\theta}A + \frac{(1-\cos\theta)}{\theta^2}B \tag{274}$$

where

$$A = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}, B = \begin{pmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{pmatrix} \tag{275}$$

This formula has many applications in kinematics, robotics, and motion interpolation.

Theorem 13.2 ()

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 13.3 ()

The Lie algebras $\mathfrak{gl}_{\mathbb{R}}, \mathfrak{sl}_n\mathbb{R}, \mathfrak{o}(n), \mathfrak{so}(n)$ are well-defined, but only

$$exp : \mathfrak{so}(n) \to \mathrm{SO}(n) \tag{276}$$

is surjective.

Theorem 13.4 ()

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

$$\mathrm{GL}(n,\mathbb{C}), \mathrm{SL}(n,\mathbb{C}), \mathfrak{gl}_n\mathbb{C}, \mathfrak{sl}_n\mathbb{C} \tag{277}$$

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 13.5 ()

$exp : \mathfrak{sl}_2\mathbb{C} \to \mathrm{SL}(2,\mathbb{C})$ 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

$$J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} \implies e^J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} e^\lambda & 0 \\ 0 & e^{-\lambda} \end{pmatrix} \tag{278}$$

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

$$K = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \tag{279}$$

which is not one of the 2 forms. So, $K \notin \operatorname{Im} exp \implies exp$ is not surjective.

Theorem 13.6 ()

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.

13.1.4 Lie Algebra of SE(n)

Recall that the group of affine rigid isometries is denoted $\mathrm{SE}(n)$. That is,

$$\mathrm{SE}(n) \equiv \mathrm{SO}(n) \ltimes \operatorname{Tran} \mathbb{R}^n \tag{280}$$

We can define the matrix representation of this affine transformation as such. Given an element $g \in \mathrm{SE}(n)$ such that

$$g(x) \equiv Rx + U, \; R \in \mathrm{SO}(n), U \in \operatorname{Tran} \mathbb{R}^n \tag{281}$$

we define the representation

$$\rho : \mathrm{SE}(n) \to \mathrm{GL}(n+1,\mathbb{R}), \rho(g) \equiv \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \tag{282}$$

where R is a real $n \times n$ matrix in $\mathrm{SO}(n)$ and U is a real n-vector in $\operatorname{Tran} \mathbb{R}^n \simeq \mathbb{R}^n$. We would then have

$$\rho(g) \begin{pmatrix} x \\ 1 \end{pmatrix} \equiv \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Rx + U \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1} \tag{283}$$

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

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix} \tag{284}$$

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 13.2 ()

Given any $(n+1)\times(n+1)$ matrix of form

$$A = \begin{pmatrix} \Omega & U \\ 0 & 0 \end{pmatrix} \tag{285}$$

where Ω is any matrix and $U \in \mathbb{R}^n$,

$$A^k = \begin{pmatrix} \Omega^k & \Omega^{k-1}U \\ 0 & 0 \end{pmatrix} \tag{286}$$

where $\Omega^0 = I_n$, which implies that

$$e^A = \begin{pmatrix} e^\Omega & VU \\ 0 & 1 \end{pmatrix},\ V = I_n + \sum_{k\geq 1} \frac{\Omega^k}{(k+1)!} \tag{287}$$

Theorem 13.7 ()

The exponential map

$$exp : \mathfrak{se}(n) \to SE(n) \tag{288}$$

is well-defined and surjective.

13.2 Representations of Lie Groups and Lie Algebras

Let $\mathcal{G}$ be an abstract group and let

$$\rho : \mathcal{G} \to \mathrm{GL}(V) \tag{289}$$

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

$$\varrho : \mathfrak{g} \to \mathfrak{gl}(V) \tag{290}$$

where the bracket structure (in this case, the comutator in the matrix algebra) is preserved.

$$\varrho([X,Y]) = [\varrho(X), \varrho(Y)] \tag{291}$$

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 $\mathrm{GL}(V)$ is restricted to only matrices with nonzero determinants, while the elements of $\mathfrak{gl}(V)$ can be any matrix.

Example 13.4 ()

The representation of $\mathrm{SE}(n)$ to $\mathrm{GL}(n+1\mathbb{R}$ and $\mathfrak{se}(n)$ to $\mathfrak{gl}_{n+1}\mathbb{R}$ induces the second homomorphism $\varrho : \mathfrak{gl}_{n+1}\mathbb{R} \to \mathrm{GL}(n+1,\mathbb{R})$.

Definition 13.5 ()

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

$$\rho_1 : \mathcal{G} \to U,\ \rho_1(g) = \begin{pmatrix} u_1 & u_2 \\ u_3 & u_4 \end{pmatrix} \tag{292}$$

and

$$\rho_2 : \mathcal{G} \to V,\ \rho_2(g) = \begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix} \tag{293}$$

are two representations of the same group element $g \in \mathcal{G}$, then

$$(\rho_1 \oplus \rho_2) : \mathcal{G} \to (U \oplus V),\ (\rho_1 \oplus \rho_2)(g) = \begin{pmatrix} u_1 & u_2 & 0 & 0 \\ u_3 & u_4 & 0 & 0 \\ 0 & 0 & v_1 & v_2 \\ 0 & 0 & v_3 & v_4 \end{pmatrix} \tag{294}$$

is a bigger representation of g in $U \oplus V$.

Definition 13.6 ()

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

$$\varrho : \mathfrak{sl}_2\mathbb{C} \to \mathfrak{gl}(V) \tag{295}$$

We will write, as shorthand notation, that

$$H = \varrho(H), X = \varrho(X), Y = \varrho(Y) \tag{296}$$

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

$$He_i = \lambda_i e_i,\ \lambda_i \in \mathbb{C} \tag{297}$$

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

$$HYe_i - YHe_i = -2Ye_i \implies H(Ye_i) = (\lambda_i - 2)(Ye_i) \tag{298}$$

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

$$\{v_m, Yv_m, Y^2 v_m, Y^3 v_m, ..., Y^{n-1} v_m\} \tag{299}$$

(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 13.2 ()

$$XY^j v_m = j(m-j+1)Y^{j-1}v_m \tag{300}$$

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

$$0 = XY^n v_m = n(m-n+1)Y^{n-1}v_m \implies m-n+1 = 0 \tag{301}$$

So, $n = m+1$, which means that the eigenvalues of H are

$$m, m-2, m-4, \ldots, m-2(n-1) = -m \tag{302}$$

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,
$$Hv_0 = 0,\ Xv_0 = 0,\ Yv_0 = 0 \tag{303}$$
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.
$$\varrho(H) = (0),\ \varrho(X) = (0),\ \varrho(Y) = (0) \tag{304}$$

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
$$\varrho(H) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{305}$$

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
$$\varrho(H) = \begin{pmatrix} 2 & & \\ & 0 & \\ & & -2 \end{pmatrix}, \varrho(Y) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \varrho(X) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{306}$$

4. The same process continues on for $n = 4, 5, \ldots$, and this entirely classifies the irreducible representations of $\mathfrak{sl}_2\mathbb{C}$.

13.2.1 Tensor Products of Group Representations

Definition 13.7 ()

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

$$\begin{aligned} \rho_V &: \mathcal{G} \to \text{GL}(V) \\ \rho_W &: \mathcal{G} \to \text{GL}(W) \end{aligned}$$

The homomorphism $\rho_V \otimes \rho_W : \mathcal{G} \to \text{GL}(V \otimes W)$ is also a representation of $\mathcal{G}$, which is defined

$$(\rho_V \otimes \rho_W)(g)(v \otimes w) \equiv \rho_V(g)(v) \otimes \rho_W(g)(w) \tag{307}$$

or represented in shorthand notation,

$$g(v \otimes w) \equiv (gv) \otimes (gw) \tag{308}$$

We know that $\exp(H)$ acts on V and W since it is an element of $\text{GL}(V)$ and $\text{GL}(W)$. This means that

$$exp(H)(v \otimes w) \equiv \big(exp(H)(v)\big) \otimes \big(exp(H)(w)\big) \tag{309}$$

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

$$exp(H)(v \otimes w) = (e^{\lambda} v) \otimes (e^{\mu} w) = e^{\lambda + \mu} v \otimes w \tag{310}$$

That is, eigenvalues of H **add** on tensor products.

Example 13.5 ()

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

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1} \tag{311}$$

We can decompose $V \otimes V$ into its symmetric and exterior power components. $\text{Sym}^2 V$ has basis (of eigenvectors)

$$\{e_{-1} \odot e_{-1},\; e_{-1} \odot e_1,\; e_1 \odot e_1\} \tag{312}$$

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,

$$V \otimes V = Sym^2 V \oplus \Lambda^2 V = \mathbf{3} \oplus \mathbf{1} \tag{313}$$

13.3 Topological Decompositions of Lie Groups

Definition 13.8 ()

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 13.3 ()

The exponential map

$$exp : S(n) \rightarrow SPD(n) \tag{314}$$

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 13.8 (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

$$f = g \circ h_1 = h_2 \circ g \tag{315}$$

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 13.1 ()

The two topological groups are homeomorphic.

$$GL(n, \mathbb{R}) \cong O(n) \times SPD(n) \tag{316}$$

Corollary 13.2 ()

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

$$A = Re^S \tag{317}$$

$\implies$ 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,

$$\mathrm{GL}(n, \mathbb{R}) \simeq O(n) \times S(n) \simeq O(n) \times \mathbb{R}^{n(n+1)/2} \tag{318}$$

This essentially reduces the study of $\text{GL}(n, \mathbb{R})$ to the study of $O(n)$, which is nice since $O(n)$ is compact.

Corollary 13.3 ()

Given a real matrix A, if $\det A > 0$, then we can decompose A as

$$A = Re^S \tag{319}$$

where $R \in SO(n)$ and $S \in S(n)$.

Corollary 13.4 ()

There exists a bijection between

$$\text{SL}(n, \mathbb{R}) \text{ and } SO(n) \times (S(n) \cap \mathfrak{sl}_n\mathbb{R}) \tag{320}$$

Proof.

$A \in \text{SL}(n, \mathbb{R}) \implies 1 = \det A = \det R \det e^S = \det e^S \implies \det e^S = e^{\text{Tr}\, S} = 1 \implies \text{Tr}\, S = 0 \implies S \in \text{S}(n) \cap \mathfrak{sl}_n\mathbb{R}$.

Definition 13.9 ()

Let us define

1. $\text{H}(n)$ is the real vector space of $n \times n$ Hermitian matrices.
2. $\text{HP}(n)$ is the set of Hermitian, positive semidefinite $n \times n$ matrices.
3. $\text{HPD}(n)$ is the set of Hermitian, positive definite $n \times n$ matrices.

Similarly, $\text{HP}(n)$ and $\text{HPD}(n)$ are not vector space. They are just sets.

Lemma 13.4 ()

The exponential mapping

$$exp : H(n) \to HPD(n) \tag{321}$$

is a homeomorphism.

However again, $\text{HPD}(n)$ is not a Lie group (multiplication is not algebraically closed) nor is $\text{H}(n)$ 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

$$\text{Mat}(n, \mathbb{C}) \text{ and } U(n) \times HP(n) \tag{322}$$

which implies that

$$\text{GL}(n, \mathbb{C}) \cong U(n) \times HPD(n) \tag{323}$$

Corollary 13.5 ()

For every complex invertible matrix A, there exists a unique decomposition

$$A = Ue^S \tag{324}$$

where $U \in U(n)$ and $S \in H(n)$, which implies that the following groups are homeomorphic.

$$\begin{aligned} \text{GL}(n, \mathbb{C}) &\cong U(n) \times H(n) \\ &\cong U(n) \times \mathbb{R}^{n^2} \end{aligned}$$

This essentially reduces the study of $\text{GL}(n, \mathbb{C})$ to that of $\text{U}(n)$.

Corollary 13.6 ()

There exists a bijection between

$$\text{SL}(n, \mathbb{C}) \text{ and } SU(n) \times (H(n) \cap \mathfrak{sl}_n \mathbb{C}) \tag{325}$$

Proof.

Similarly, when $A = Ue^S$, we know that $|\det U| = 1$ and $\text{Tr}\, S$ 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 $A \in \text{SL}(n, \mathbb{C})$, then $\det A = 1 \implies \det e^S = 1 \implies S \in H(n) \cap \mathfrak{sl}_n \mathbb{C}$.

13.4 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 $\text{Mat}(n, \mathbb{R})$ is bijective to $\mathbb{R}^{n^2}$, which is a topological space. Therefore, this bijection

$$i : (\mathbb{R}^{n^2}, \tau_E) \to \text{Mat}(n, \mathbb{R}) \tag{326}$$

induces a topology on $\text{Mat}(n, \mathbb{R})$, defined

$$\tau_M \equiv \{U \in \text{Mat}(n, \mathbb{R}) \mid e^{-1}(U) \in \tau_E\} \tag{327}$$

With this, consider the subset

$$\text{GL}(n, \mathbb{R}) \subset \text{Mat}(n, \mathbb{R}) \tag{328}$$

where

$$\text{GL}(n, \mathbb{R}) \equiv \{x \in \text{Mat}(n, \mathbb{R}) \mid \det x \neq 0\} \tag{329}$$

This set, as we expect, is a multiplicative group.

Definition 13.10 ()

The **general linear group**, denoted $\text{GL}(n, \mathbb{R})$ is the set of $n \times n$ matrices with nonzero determinant. The more technical definition is that $\text{GL}(n, \mathbb{R})$ is really just the automorphism group of $\mathbb{R}^n$,

$$\text{GL}(n, \mathbb{R}) \equiv \text{Aut}(\mathbb{R}^n) \tag{330}$$

but it is customary to assume a basis on $\mathbb{R}^n$ in order to realize $\text{GL}(n, \mathbb{R})$ 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 $\text{GL}(n, \mathbb{R})$. Both assigns a real $n \times n$ matrix to each element of $\text{GL}(n, \mathbb{R})$.

In this way, we can view $\text{GL}(n, \mathbb{R})$ as a topological space in $\mathbb{R}^{n^2}$, and it is fine to interpret $\text{GL}(n, \mathbb{R})$ 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

$$\mathrm{Mat}(n,\mathbb{C}) \cong \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2} \tag{331}$$

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 13.3 ()

SE(n) is a linear Lie group.

Proof.

The matrix representation of elements $g \in SE(n)$ is

$$\rho(g) \equiv \begin{pmatrix} R_g & U_g \\ 0 & 1 \end{pmatrix}, \; R_g \in SO(n), U_g \in \mathbb{R}^n \tag{332}$$

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

$$i : SE(n) \to \mathrm{SL}(n+1,\mathbb{R}) \tag{333}$$

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

$$\begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} R^{-1} & -R^{-1}U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R^T & -R^T U \\ 0 & 1 \end{pmatrix} \tag{334}$$

is also consistent between the inverse operation in SE(n) and $\mathrm{SL}(n+1,\mathbb{R})$. Therefore, 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 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.

$$SE(n) \equiv SO(n) \ltimes \mathbb{R}^n \tag{335}$$

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 13.11 ()

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 13.9 (Von Neumann, Cartan)

A closed subgroup $\mathcal{G}$ of $\mathrm{GL}(n,\mathbb{R})$ is a linear Lie group. That is, a closed subgroup $\mathcal{G}$ of $\mathrm{GL}(n,\mathbb{R})$ is a smooth manifold in $\mathbb{R}^{n^2}$.

Definition 13.12 ()

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

$$T_I\mathcal{G} \equiv \{p'(0) \mid p : I \subset \mathbb{R} \to \mathcal{G}, p(0) = I\} \tag{336}$$

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 13.10 ()

Let $\mathcal{G}$ be a linear Lie group. The set $\mathfrak{g}$ defined such that

$$\mathfrak{g} \equiv \{X \in \mathrm{Mat}(n,\mathbb{R}) \mid e^{tX} \in \mathcal{G} \; \forall t \in \mathbb{R}\} \tag{337}$$

is equal to the tangent space of $\mathcal{G}$ at the identity element. That is,

$$\mathfrak{g} = T_I\mathcal{G} \tag{338}$$

Furthermore, $\mathfrak{g}$ is closed under the commutator

$$[A,B] \equiv AB - BA \tag{339}$$

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 13.13 ()

The Lie algebra of a linear Lie group is a real vector space (of matrices) together with a algebraically closed bilinear map

$$[A,B] \equiv AB - BA \tag{340}$$

called the **commutator**.

The definition of $\mathfrak{g}$ given in the previous theorem shows that

$$exp : \mathfrak{g} \to \mathcal{G} \tag{341}$$

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

$$\gamma_X : t \mapsto e^{tX} \tag{342}$$

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 13.11 ()

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 13.5 ()

$\det'(I) = \mathrm{Tr}$. 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,

$$\nabla_T \det I = \lim_{\varepsilon \to 0} ... + \mathrm{Tr}\, T = \mathrm{Tr}\, T \tag{343}$$

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

$$\overset{\prime}{\det}\big(R'(0)\big) = \overset{\prime}{\det}(I) = 0 \implies \mathrm{Tr}\, R'(0) = 0 \tag{344}$$

Theorem 13.12 ()

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 $\text{SL}(n)$ such that $R(0) = I$. Then, for all $t \in (-\epsilon, \epsilon)$,

$$R(t)R(t)^T = I \tag{345}$$

Taking the derivative at $t = 0$, we get

$$R'(0)R(0)^T + R(0)R'(0)^T = 0 \implies R'(0) + R'(0)^T = 0 \tag{346}$$

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
$$[A, B] \equiv AB - BA \tag{347}$$
but for elements that are not matrices this doesn't make sense.
3. The exponential map from the algebra to the group is defined
$$e^A \equiv \sum_{k=0}^{\infty} \frac{1}{k!} A^k \tag{348}$$
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.

13.4.1 Lie Algebras of SO(3) and SU(2), Revisited

Example 13.6 ()

The Lie algebra $\mathfrak{so}(3)$ is the real vector space of 3×3 skew symmetric matrices of form

$$\begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix} \tag{349}$$

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

$$\psi(b, c, d) \equiv \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix} \tag{350}$$

where, by definition,

$$\psi(u \times v) = [\psi(u), \psi(v)] \tag{351}$$

It is also easily verified that for all $u, v \in \mathbb{R}^3$,

$$\psi(u)(v) = u \times v \tag{352}$$

Example 13.7 ()

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

$$i(d\sigma_1 + c\sigma_2 + b\sigma_3) = \begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} \tag{353}$$

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

$$\varphi(b, c, d) = \frac{i}{2}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{2}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} \tag{354}$$

where, by definition of ismorphism, we have

$$\varphi(u \times v) = [\varphi(u), \varphi(v)] \tag{355}$$

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

$$A = \frac{1}{\theta}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{\theta}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} \tag{356}$$

such that $A^2 = I$. With this, we can rewrite the exponential map as

$$exp : \mathfrak{su}(2) \to SU(2),\ exp(i\theta A) = \cos\theta I + i \sin\theta A \tag{357}$$

As for the isomorphism $\varphi : (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$, we have

$$\varphi(b, c, d) \equiv \frac{1}{2}\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} = i\frac{\theta}{2}A \tag{358}$$

Similarly, we can view the exponential map $exp : (\mathbb{R}^3, \times) \to SU(2)$ as

$$exp(\theta v) = \tag{359}$$

Example 13.8 ()

The lie algebra $\mathfrak{se}(n)$ is the set of all matrices of form

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} \tag{360}$$

where $B \in \mathfrak{so}(n)$ and $U \in \mathbb{R}^n$. The Lie bracket is given by

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix}\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} BC - CB & BV - CU \\ 0 & 0 \end{pmatrix} \tag{361}$$

13.5 Abstract Lie Groups

Definition 13.14 ()

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 13.15 ()

A (real) Lie algebra $\mathfrak{g}$ is a real vector space with a map

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{362}$$

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 13.16 ()

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

$$f'_p : T_p M_1 \to T_{f(p)} M_2 \tag{363}$$

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 $\gamma'(0) = v$,

$$f'_p(v) \equiv (f \circ \gamma)'(0) \tag{364}$$

The map f'_p is also denoted df_p and is a linear map.

Definition 13.17 ()

Given two Lie groups $\mathcal{G}_1$ and $\mathcal{G}_2$, a **homomorphism of Lie groups** is a function

$$f : \mathcal{G}_1 \to \mathcal{G}_2 \tag{365}$$

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 13.18 ()

Given two Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, a **homomorphism of Lie algebras** is a function

$$f : \mathfrak{g}_1 \to \mathfrak{g}_2 \tag{366}$$

that is a linear homomorphism that preserves Lie brackets; that is,

$$f([A, B]) = [f(A), f(B)] \tag{367}$$

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 13.4 ()

If $f : \mathcal{G}_1 \to \mathcal{G}_2$ is a homomorphism of Lie groups, then

$$f'_I : \mathfrak{g}_1 \to \mathfrak{g}_2 \tag{368}$$

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 13.19 ()

Given a Lie group $\mathcal{G}$, we define a **left translation** as the map

$$L_a : \mathcal{G} \to \mathcal{G},\ L_a(b) \equiv ab \tag{369}$$

for all $b \in \mathcal{G}$. Similarly, the **right translation** is defined

$$R_a : \mathcal{G} \to \mathcal{G},\ R_a(b) \equiv ba \tag{370}$$

for all $b \in \mathcal{G}$.

Both L_a and R_a are diffeomorphisms. Additionally, given the automorphism

$$R_{a^{-1}} L_a \equiv R_{a^{-1}} \circ L_a,\ R_{a^{-1}} L_a(b) \equiv aba^{-1} \tag{371}$$

the derivative

$$(R_{a^{-1}} L_a)'_I : \mathfrak{g} \to \mathfrak{g} \tag{372}$$

is an ismorphism of Lie algebras, also denoted

$$\mathrm{Ad}_a : \mathfrak{g} \to \mathfrak{g} \tag{373}$$

Definition 13.20 ()

This induces another map $a \mapsto \mathrm{Ad}_a$, which is a map of Lie groups

$$Ad : \mathcal{G} \to \mathrm{GL}(\mathfrak{g}) \tag{374}$$

which is called the **adjoint representation of $\mathcal{G}$**. In the case of a linear map, we can verify that

$$\mathrm{Ad}(a)(X) \equiv \mathrm{Ad}_a(X) \equiv aXa^{-1} \tag{375}$$

for all $a \in \mathcal{G}$ and for all $X \in \mathfrak{g}$.

Definition 13.21 ()

Furthermore, the derivative of this map at the identity

$$\mathrm{Ad}'_I : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \tag{376}$$

is a map between Lie algebras, denoted simply as

$$\text{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) \tag{377}$$

called the **adjoint representation** of $\mathfrak{g}$. It is easily visualized with the following commutative diagram.

$$\begin{tikzcd} \mathcal{G} \arrow[r, "Ad"] & \text{GL}(\mathfrak{g}) \\ \mathfrak{g} \arrow[u, "exp"] \arrow[r, "ad"] & \mathfrak{gl}(\mathfrak{g}) \arrow[u, "exp"'] \end{tikzcd}$$

We define the map ad to be

$$ad(A)(B) \equiv [A, B] \tag{378}$$

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,

$$\text{ad}([A, B]) = [\text{ad}(A), \text{ad}(B)] \tag{379}$$

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 13.22 ()

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

$$[A, B] \equiv \text{ad}(A)(B) \tag{380}$$

We would also need to introduce a general exponential map for non-linear Lie groups, but we will not do it here.

14 Fields

14.1 Ring Extensions

We will introduce this in a slightly different way, but by building up some theorems, we will unify these two soon enough.

Definition 14.1 (Ring of Univariate Polynomial Elements)

Let $F \subset K$ be fields, $F[x]$ a polynomial ring, and a constant $\alpha \in K$,

$$F[\alpha] \coloneqq \{f(\alpha) \in F \mid f \in F[x]\} \subset K \tag{381}$$

Lemma 14.1 (Ring Extension)

We have the following subring structure.

$$F \subset F[\alpha] \subset K \tag{382}$$

Furthermore, if $\alpha \notin F$, then $F \subsetneq F[\alpha]$.

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

$$\phi + \gamma = f(\alpha) + g(\alpha) = (f+g)(\alpha) \tag{383}$$
$$\phi \cdot \gamma = f(\alpha) \cdot g(\alpha) = (fg)(\alpha) \tag{384}$$

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 14.1 (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

$$f(\sqrt{2}) = a_n(\sqrt{2})^n + a_{n-1}(\sqrt{2})^{n-1} + \ldots + a_2(\sqrt{2})^2 + a_1\sqrt{2} + a_0 \tag{385}$$

   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 14.2 ()

Given any element $\beta \in \mathbb{Q}[\sqrt{2}+\sqrt{3}]$, it is by definition of the form

$$\beta = \sum_{k=0}^{n} a_k (\sqrt{2}+\sqrt{3})^k \tag{386}$$

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

$$\frac{1}{2}\big[(11\sqrt{2}+9\sqrt{3}) - 9(\sqrt{2}+\sqrt{3})\big] = \sqrt{2} \in \mathbb{Q}[\sqrt{2}+\sqrt{3}] \tag{387}$$

$$-\frac{1}{2}\big[(11\sqrt{2}+9\sqrt{3}) - 11(\sqrt{2}+\sqrt{3})\big] = \sqrt{3} \in \mathbb{Q}[\sqrt{2}+\sqrt{3}] \tag{388}$$

$$\tag{389}$$

If we go a bit further, we can show that

$$\mathbb{Q}[\sqrt{2}+\sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} \mid a, b, c, d \in \mathbb{Q}\} \tag{390}$$

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

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.

Lemma 14.2 (Vector Space Structure)

$F[\alpha]$ is a finite-dimensional vector space over F. If $f(x) = a_n x^n + \ldots a_0$, then $S = \{1, \alpha, \ldots, \alpha^{n-1}\}$ 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

$$f(\alpha) = \sum_{k=0}^{n} a_k \alpha^k \tag{391}$$

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

$$\alpha^n = -\frac{1}{a_n}\big(a_{n-1}\alpha^{n-1} + \ldots + a_0\big) \tag{392}$$

which proves the claim. Now assume that $\alpha^n, \alpha^{n+1}, \ldots, \alpha^{n-1+i} \in \text{span}\{1, \ldots, \alpha^{n-1}\}$. Then

$$\alpha^i f(\alpha) = 0 \implies a_n \alpha^{n+i} + \alpha_{n-1}\alpha^{n+i-1} + \ldots + a_0 \alpha^i = 0 \tag{393}$$

and so

$$\alpha^{n+i} = -\frac{1}{a_n}\big(a_{n-1}\alpha^{n+i-1} + \ldots + a_0 \alpha^i\big) \tag{394}$$

which means that $\alpha^{n+i} \in \text{span}\{1, \ldots, \alpha^{n-1}\}$, completing the proof.

14.2 Field Extensions

Great, so we automatically have the ring and vector space structures on $F[\alpha]$. However, what we would really like is a field structure since that was our original goal. Remember that $F[\alpha]$ is a ring that contains both F and α. With one more assumption, we can claim that it is a field.

Theorem 14.1 (Adjoining Fields)

Given fields $F \subset K$, if there exists a $f \in F[x]$ 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 field obtained by **adjoining** α to F.

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.

$$1 = sp + tg \implies 1 = s(\alpha)p(\alpha) + t(\alpha)g(\alpha) = s(\alpha)p(\alpha) \tag{395}$$

Therefore we have found a multiplicative inverse $s = p^{-1} \in F[\alpha]$.

Proof.

We can prove it using the vector space structure. Treating $F[\alpha]$as a finite-dimensional vector space over F, let us define the F-linear function[a]

$$m_b : F[\alpha] \to F[\alpha], \qquad m_b(\beta) = b\beta \tag{396}$$

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

Corollary 14.1 (Adjoining Field is Minimal)

$F[\alpha]$ is the smallest field containing F and α.

Example 14.3 ($\mathbb{Q}[\sqrt{3}i]$ 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 14.4 ($\mathbb{Q}[\pi]$ not a Field)

However, $\mathbb{Q}[\pi]$ is not a field.

Example 14.5 (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

$$1 = \frac{1}{3}\big\{(x+1)f(x) - (x^2+2)p(x)\big\} = -\frac{1}{3}(\alpha^2+2)p(\alpha) \tag{397}$$

and so $\beta^{-1} = (\alpha^2 + \alpha - 1)^{-1} = -\frac{1}{3}(\alpha^2 + 2)$. We can check that

$$-\frac{1}{3}(\alpha^2+2)(\alpha^2+\alpha-1) = -\frac{1}{3}(\alpha^4+\alpha^3+\alpha^2+2\alpha-2) \tag{398}$$

$$= -\frac{1}{3}(\alpha^3+\alpha-2) \tag{399}$$

$$= -\frac{1}{3}(-3) = 1 \tag{400}$$

Intuitively, the extra $\alpha \in K$ allows us to "expand" our field F into a bigger field of K. We can also define this for multivariate polynomials.

Definition 14.2 (Ring of Multivariate Polynomial Elements)

Given a polynomial ring $F[x, y]$ over a field F and constants $\alpha, \beta \in F$, the following definitions are equivalent.

$$F[\alpha, \beta] \coloneqq \{f(\alpha, \beta) \in F \mid f \in F[x, y]\} \tag{401}$$

$$= (F[\alpha])[\beta] \tag{402}$$

$$= (F[\beta])[\alpha] \tag{403}$$

Proof.

Example 14.6 (Extensions of $\sqrt{2}$ and i)

We claim that

$$\mathbb{Q}[\sqrt{2}, i] = \{a + b\sqrt{2} + ci + d(\sqrt{2}i) \mid a, b, c, d \in \mathbb{Q}\} \tag{404}$$

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

$$f(i) = (a_n + b_n\sqrt{2})i^n + (a_{n-1} + b_{n-1}\sqrt{2})i^{n-1} + \ldots + (a_2 + b_2\sqrt{2})i^2 + (a_1 + b_1\sqrt{2})i + (a_0 + b_0\sqrt{2}) \tag{405}$$

However, we can see that since $i^2 = -1$, we only need to consider up to degree 1 polynomials of form

$$(a + b\sqrt{2}) + (c + d\sqrt{2})i \tag{406}$$

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 14.7 ()

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

$$(\sqrt{3}+i)^3 = 8i \implies i \in \mathbb{Q}[\sqrt{3}+i] \tag{407}$$

$$\implies (\sqrt{3}+i) - i = \sqrt{3} \in \mathbb{Q}[\sqrt{3}+i] \tag{408}$$

Therefore, $\mathbb{Q}[\sqrt{3}+i]$ contains the elements $1, \sqrt{3}, i$, which form the basis of $\mathbb{Q}[\sqrt{3}, i]$.

Example 14.8 (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,

$$\sqrt{3}, i \in \mathbb{Q}[\sqrt{3}, i] \implies \sqrt{3}i \in \mathbb{Q}[\sqrt{3}, i] \tag{409}$$

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

$$-1 = (b_0 + b_1\sqrt{3}i)^2 = (b_0^2 - 3b_1^2) + 2b_0b_1\sqrt{3}i \tag{410}$$

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.

14.3 Splitting Fields

Now we return to the problem of taking a polynomial $f \in \mathbb{Q}[x]$ and finding the *smallest* possible field $K \subset \mathbb{C}$ s.t. f can be factored as a product of linear polynomials in $K[x]$.

Example 14.9 (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

$$f(x) = (x - (-1+i))(x - (-1-i)) \tag{411}$$

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

$$f(x) = (x - (1+\sqrt{2}))(x - (1-\sqrt{2})) \tag{412}$$

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

$$f(x) = (x-1)(x+1)(x^2+x+1)(x^2-x+1) \tag{413}$$

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 14.10 ()

Let $f(x) = x^4 - 2 \in \mathbb{Q}[x]$. It follows that the roots are

$$\{\sqrt[4]{2}, \sqrt[4]{2}, -\sqrt[4]{2}, -\sqrt[4]{2}i\} = \Big\{\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\Big\} \tag{414}$$

thus the splitting field of f is

$$\mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) \subset \mathbb{Q}(\sqrt[4]{2}, e^{\frac{2\pi i}{4}}) \tag{415}$$

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,

$$e^{2\pi i/4} = \frac{\sqrt[4]{2}e^{2\pi i/4}}{\sqrt[4]{2}} \in \mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) \tag{416}$$

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

$$\mathbb{Q}\big(\sqrt[4]{2}, \sqrt[4]{2}e^{\frac{2\pi i}{4}}, \sqrt[4]{2}e^{\frac{4\pi i}{4}}, \sqrt[4]{2}e^{\frac{6\pi i}{4}}\big) = \mathbb{Q}(\sqrt[4]{2}, e^{\frac{2\pi i}{4}}) \tag{417}$$

15 TBD

15.1 Abelian Groups

First, note that the successive addition of elements of an additive abelian group can be represented by integer multiplication.

$$x + x + ... + x = nx, \ n \in \mathbb{Z} \tag{418}$$

Similarly, we can take the integer power of an element to represent successive multiplication in a multiplicative abelian group.

Lemma 15.1 ()

It is easy to check that in an additive abelian group A, with $a, b \in A$ and $k, l \in \mathbb{Z}$,

$$k(a+b) = ka + kb \tag{419}$$
$$(k+l)a = ka + la \tag{420}$$
$$(kl)a = k(la) \tag{421}$$

which implies

$$k(a-b) = ka - kb, \ (k-l)a = ka - la \tag{422}$$

Definition 15.1 ()

For any subset $S \subset A$, the collection of all linear combinations

$$k_1 a_1 + k_2 a_2 + ... + k_n a_n, \ k_i \in \mathbb{Z}, a_i \in S \tag{423}$$

is the smallest subgroup of A containing S, called the **subgroup generated by** S and denoted $\langle S \rangle$. If $\langle S \rangle = A$, then we say that A is **generated** by S, or that S is a **generating set** of A.

Definition 15.2 ()

An abelian group that has a finite generating set is called **finitely generated**. Finitely generated abelian groups are similar to finite dimensional vector spaces.

Definition 15.3 ()

A system $\{a_1, a_2, ..., a_n\}$ of elements of a group A is called **linearly independent** if $k_1 a_1 + k_2 a_2 + ... + k_n a_n = 0 \implies k_1, k_2, ..., k_n = 0$. A system of linear independent elements that generates A is called a **basis**.

Note that every finite dimensional vector has a basis, but not every finitely generated abelian group has one. For example, $(\mathbb{Z}_n, +)$ is generated by one element, but it has no basis since every element $a \in \mathbb{Z}_n$ satisfies the nontrivial relation $na = 0$.

Definition 15.4 ()

A finitely generated abelian group is **free** if it has a basis.

Theorem 15.1 ()

All bases of a free abelian group L contain the same number of elements.

Definition 15.5 ()

The **rank** of a free abelian group L is the number of elements in its basis. It is denoted $\text{rk}L$. The zero group is regarded as a free abelian group of rank 0.

Theorem 15.2 ()

Every free abelian group L of rank n is isomorphic to the group $\mathbb{Z}^n$ of integer rows of length n.

Theorem 15.3 ()

Every subgroup n of a free abelian group l of rank n is a free abelian group of rank $\leq n$.

Note that unlike a vector space, a free abelian group of positive rank contains subgroups of the same rank that do not conside with the whole group. For example, the subgroup $m\mathbb{Z} \subset \mathbb{Z}, m > 0$ has rank 1, just as the whole group.

Moreover, a free abelian group of rank n can be embedded as a subgroup into an n-dimensional Euclidean vector space E^n. To do this, let $\{e_1, e_2, ..., e_n\}$ be a basis of E^n. Then, the subgroup generated by these basis vectors is the set of vectors with integer components, which is a free abelian group of rank n. This subgroup obtained as such is called a **lattice** in E^n.

Definition 15.6 ()

A subgroup $L \subset E^n$ is **discrete** if every bounded subset of E^n contains a finite number of elements in L. Clearly, every lattice is discrete, and a subgroup generated by a linearly independent system of vectors (i.e. a lattice in a subspace of E^n) is discrete.

Lemma 15.2 ()

A subgroup $L \subset E^n$ is discrete if and only if its intersection with any neighborhood of 0 consists of 0 itself.

Theorem 15.4 ()

Every discrete subgroup $L \subset E^n$ is generated by a linearly independent system of vectors of E^n.

Corollary 15.1 ()

A discrete subgroup $L \subset E^n$ whose linear span coincides with E^n is a lattice in E^n.

Lattices in E^3 play an important role in crystallography since the defining feature of a crystal structure is the periodic repetition of the configuration of atoms in all three dimensions. More explicitly, let Γ be the symmetry group of the crystal structure and let $\mathcal{L}$ be the group of all vectors a such that the parallel translation $t_a \in \Gamma$. Then, $\mathcal{L}$ is a discrete subgroup of E^n and thus, is a lattice in E^3. More specifically, we can present

$$\Gamma \equiv \text{Dih}\, C \times \mathcal{L} \tag{424}$$

where $\text{Dih}\, C$ is the Dihedral group of the crystal structure that preserves its lattices.

Definition 15.7 ()

An **integral elementary row transformation** of a matrix is a transformation of one of the following three types:

1. adding a row multiplied by an integer to another row
2. interchanging two rows
3. multiplying a row by -1

An **integral elementary column transformation** is defined similarly.

Lemma 15.3 ()

Every integral rectangular matrix $C = (c_{ij})$ can be reduced by integral elementary row transformations to the diagonal matrix $\text{diag}(u_1, ..., u_p)$, where $u_1, u_2, ..., u_p \geq 0$ and $u_i | u_{i+1}$ for $i = 1, 2, ..., p-1$.

Example 15.1 ()

The following matrix can be reduced (with a few steps now shown) to the stated form.

$$\begin{pmatrix} 2 & 6 & 2 \\ 2 & 3 & 4 \\ 4 & 2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & 4 \\ 0 & -3 & 2 \\ 4 & 2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 14 \\ 0 & 8 & 12 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 20 \end{pmatrix} \tag{425}$$

where $1|2$ and $2|20$.

Note that for $n \times 1$ or $1 \times n$ matrices, this procedure is precisely the Euclidean algorithm that produces the GCD of n integers.

Lemma 15.4 ()

Given square integral matrix C with reduced form $\text{diag}(u_1, ..., u_p)$,

$$u_i = \frac{d_i}{d_{i-1}} \tag{426}$$

where d_i is the GCD of the minors of order i of the original matrix C. Recall that a minor of a matrix is the determinant of the matrix with one of its rows and columns removed. d_0 is assumed to equal 1. This implies that the numbers $u_1, u_2, ..., u_p$, along with the reduced form, are uniquely determined by C.

Theorem 15.5 ()

For any subgroup N of a free abelian group L of rank n, there exists a basis $\{e_1, ..., e_n\}$ of L and natural numbers $u_1, ..., u_m$, $(m \leq n)$, such that $\{u_1 e_1, ..., u_m e_m\}$ is a basis fo the group N and $u_i | u_{i+1}$ for $i = 1, 2, ..., m-1$.

15.2 Some Types of Groups

At this point we are ready to start identifying some types of groups. We will introduce the following (non-exclusive) categories: cyclic groups, symmetric groups, symmetry groups[5], and Lie groups. There is one other category called the dicyclic group but I omit it. Just know that the group of quaterions of order 8 is one such dicyclic group that doesn't fit into any other categories. Here is a table of all known groups of small orders. We will prove this theorem as we go along.

[5] confusing name, I know

Theorem 15.6 (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 10: Classification of groups up to order 12.

15.2.1 Lie Groups

Definition 15.8 (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 15.9 (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 15.10 (General Affine Group)

The **general affine group** is the pair of all transformations

$$\text{GA}(V) \equiv \text{Tran}(V) \times \text{GL}(V) \tag{427}$$

Definition 15.11 (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 15.12 (Orthogonal Group)

The **orthogonal group**, denoted $\mathrm{O}(n)$, consists of all isometries that preserve the origin, i.e. consists of rotations and reflections. The **special orthogonal group**, denoted $\mathrm{SO}(n)$, is a subgroup of $\mathrm{O}(n)$ consisting of only rotations. We can see that

$$\mathrm{O}(n) = \frac{\operatorname{Isom} \mathbb{E}^n}{\operatorname{Tran} V} \tag{428}$$

Definition 15.13 (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 15.2 ()

$\mathrm{Tran}(V)$ and $\mathrm{GA}(V)$ are transitive groups.

Definition 15.14 (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 $\mathrm{E}(n) \subset \mathrm{GA}(n)$.

15.3 Ring Homomorphisms in Polynomials

Obviously, we can prove that things like the identity map are homomorphisms. However, the following will be used quite often.

Example 15.3 (Evaluation Homomorphism of Polynomials)

Given fields $F \subset K$, the **evaluation function**

$$\mathrm{ev}_\alpha : F[x] \to K \tag{429}$$

mapping $f(x) \mapsto f(\alpha)$ is a homomorphism.

Overall, we must use this theorem cleverly in order to prove that two rings are isomorphic to each other.

Example 15.4 ()

The evaluation map

$$\phi : \frac{\mathbb{R}[x]}{\langle x^2+1\rangle} \to \mathbb{C}, \qquad \phi\big(f(x) \pmod{\langle x^2+1\rangle}\big) = f(i) \tag{430}$$

is an isomorphism.[a] This is because we can think of the evaluation homomorphism $\mathrm{ev}_i : f(x) \in \mathbb{R}[x] \mapsto f(i) \in \mathbb{R}[i]$. We know that $\mathbb{R}$ a field implies $\mathbb{R}[x]$ is a PID. Now take $\ker(\mathrm{ev}_i)$. 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 $\ker(\mathrm{ev}_i) = \langle x^2+1\rangle$. Now it follows by the fundamental ring homomorphism theorem that

$$\frac{\mathbb{R}[x]}{\ker(\mathrm{ev}_i)} = \frac{\mathbb{R}[x]}{\langle x^2+1\rangle} \simeq \mathbb{R}[i] = \mathbb{C} \tag{431}$$

[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 15.5 ()

The evaluation map

$$\mathrm{ev}_{\sqrt{2}} : \mathbb{Q}[x] \mapsto \mathbb{Q}[\sqrt{2}], \qquad \mathrm{ev}_{\sqrt{2}}(f) = f(\sqrt{2}) \tag{432}$$

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

$$\frac{\mathbb{Q}[x]}{\langle x^2-2\rangle} \simeq \mathbb{Q}[\sqrt{2}] \tag{433}$$

Theorem 15.7 (Quotient Polynomial Ring Can be Splitting Field)

Let F be a field with $f(x) \in F[x]$.

1. Then $K = F[x]/\langle f(x)\rangle$ is a field iff $f(x)$ is irreducible in $F[x]$.
2. If $f(x)$ is irreducible, then K contains a root α of $f(x)$, and $K \simeq F[\alpha]$.

Corollary 15.2 ()

Any polynomial $f(x) \in F[x]$ has a splitting field.

Corollary 15.3 ()

Let $c \in \mathbb{C}$. Then $\mathbb{Q}[c] \subset \mathbb{C}$ is a field if and only if c is an algebraic number.