Abstract Algebra

Muchang Bahng

November 6, 2022

An introduction to a two-semester course in abstract algebra.

1 Algebraic Structures

Definition 1.1. An *operation* $*$ on a set M is a map

$$* : M \times M \longrightarrow M$$

Example 1.1. *Regular addition* $+$ *and multiplication* $\times$ *are operations in* $\mathbb{R}$, $\mathbb{Q}$, *and* $\mathbb{N}$, *but multiplication is not defined in* $\mathbb{R}_-$, *since the product of 2 negative numbers is a positive number.*

Example 1.2. *The product of functions* $f : N \longrightarrow M$ *and* $g : P \longrightarrow N$ *is defined as the composition of them*

$$(f \circ g)(x) \equiv f(g(x)) \; \forall x \in P$$

Example 1.3. $\mathbb{R}^3$ *can have operations of vector addition and the cross product. The inner product is not an operation on* $\mathbb{R}^3$.

Definition 1.2. Let $(M, \circ)$ and $(N, *)$ be two sets with their respective operations. The mapping $f : (M, \circ) \longrightarrow (N, *)$ is a *homomorphism* if

$$f(a \circ b) = f(a) * f(b) \; \forall a, b \in M$$

A homomorphism is an *isomorphism* if and only if it is bijective. If an isomorphism f exists between two algebraic structures M and N, then M and N are said to be *isomorphic*, denoted $M \simeq N$. An homomorphism from structure G to itself is called an *endomorphism*, and an endomorphism that is also an isomorphism is called an automorphism.

Example 1.4. *The map* $a \mapsto 2^a$ *is an isomorphism between* $(\mathbb{R}, +)$ *and* $(\mathbb{R}^+, \times)$ *since*

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

We now define a crucial type of measure on a set, called a relation.

Definition 1.3. given a set M, any subset $R \subset M \times M$ is called a *relation* on the set M. If $(a, b) \in R$, then a and b are *related*, denoted aRb.

Definition 1.4. An *equivalence relation* R, also written $\sim$, is a relation which is:

1. Reflexive. aRa
2. Symmetric. $aRb \iff bRa$
3. Transitive. $aRb, bRc \implies aRc$

An equivalence relation R defines an *equivalence class* $R(a)$, defined

$$R(a) \equiv \{b \in M | a \sim b\}$$

which directly implies that the set of equivalence classes $\{R(a)\}$ form a partition of M.

Definition 1.5. The set of equivalence classes under relation R is called the *quotient set* of M by R, denoted $\frac{M}{R}$. The map

$$q : M \longrightarrow \frac{M}{R},\ a \mapsto R(a)$$

is called the *quotient map*. We can also define an operation $*$ on the quotient set $\frac{M}{R}$ to get $\left(\frac{M}{R}, *\right)$, defined as

$$\{a\} * \{b\} \equiv \{a * b\}$$

to turn this quotient set into an algebraic structure. In words, $*$ applied on two classes takes arbitrary representatives of each class, does the operation on each of them, and finally outputs the class of the resulting product.

Example 1.5. *M is the set of circles in $\mathbb{R}^2$. Given $a, b \in M$, $a \sim b$ iff the radii are equal in length. We can denote each equivalence class by $\{r\}$, where r is the length of the radius. We can define addition as*

$$\{a\} + \{b\} \equiv \{a + b\}$$

1.1 Group-like Structures

Definition 1.6. A *groupoid*, also called a *magma*, is a set with operation $(M, *)$ where the operation $*$ is closed. No other properties are imposed.

Definition 1.7. A *semigroup* $(M, *)$ is a groupoid where the binary operation $*$ must be associative.

Definition 1.8. A *monoid* $(M, *)$ is a semigroup with an identity element $I \in M$ such that given a $m \in M$

$$I * m = m * I = m$$

Definition 1.9. A *group* is a monoid where every element has an inverse element. That is, $(G, *)$ is a set with binary operation having the properties of closure, associativity, existence of an identity and existence of inverses, in the following order:

1. $x, y \in S \implies x * y,\ y * x \in G$ but not necessarily $x * y = y * x$
2. $a * (b * c) = (a * b) * c\ \forall a, b, c \in G$
3. $\exists I \in G : x * I = I * x = x\ \forall x \in G$
4. $\forall x \in G\ \exists x^{-1} \in G : x * x^{-1} = x^{-1} * x = I$

Proposition 1.1. *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 I_1, I_2, where $I_1 \neq I_2$. According to the properties of identities, $I_1 = I_1 * I_2 = I_2 \implies I_1 = I_2$.
As for uniqueness of a inverses, let a be an element of G, with its inverses a_1^{-1}, a_2^{-1}. Then,

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

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

Definition 1.10. An *abelian group* $(A, +)$ is a group where $+$ is commutative. That is,

$$x + y = y + x\ \forall x, y \in A$$

The abstract operation for an abelian group is usually called addition.

Example 1.6. *$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ 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.*

Example 1.7. *The set of all functions on a given interval $F[a, b]$ is abelian with respect to addition, defined as $(f + g)(x) \equiv f(x) + g(x)$.*

1.2 Ring-like Structures

Definition 1.11. 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 $+$.
2. Multiplication $\times$ is distributive with respect to addition $+$
$$a \times (b + c) = a \times b + a \times c, (a + b) \times c = a \times c + b \times c \; \forall a, b, c \in R$$
3. There is an absorbing element, denoted 0 such that
$$0 \times a = a \times 0 = 0 \; \forall a \in R$$
4. Equivalence of Additive Inverses
$$a \times (-b) = (-a) \times b = -(a \times b)$$

Definition 1.12. A ring R is a *commutative ring* if and only if multiplication is commutative, i.e. $a \times b = b \times a \; \forall a, b \in R$. It is an *associative ring* if and only if multiplication is associative, i.e. $a \times (b \times c) = (a \times b) \times c \; \forall a, b, c \in R$.

Definition 1.13. The *unity* of a ring R is the multiplicative identity, denoted as 1.
$$a \times 1 = 1 \times a = a \; \forall a \in R$$
Note that a ring cannot have more than 1 unity, but it may not exist at all. But usually, a ring has a unity.

This is not to be confused with the unit of a ring.

Definition 1.14. A *unit* of a ring R is an element $u \in R$ that has a multiplicative inverse in R.

Example 1.8. *$\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are commutative, associative rings with respect to ordinary addition and multiplication.*

Example 1.9. *The set of even integers $2\mathbb{Z}$ is a commutative, associative ring without unity.*

Proposition 1.2. *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)*
$$M \triangle N \equiv (M \setminus N) \cup (N \setminus M)$$
and intersection $\cap$, taken for addition an multiplication, respectively.

Proof. 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
$$M \triangle N \equiv (M \setminus N) \cup (N \setminus M) \equiv N \triangle M$$
$$M \cap N = N \cap M$$
$$M \cap N \cap P = (M \cap N) \cap P = M \cap (N \cap P)$$
■

Example 1.10. *A division ring, also called a skew field, is an associative ring with unity where every nonzero element is invertible with respect to $\times$. Division rings differ from fields in that multiplication is not required to be commutative.*

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.
$$aa^{-1} = a^{-1}a = I$$
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.

Definition 1.15. A *field* $(F, +, \times)$ is a commutative, associative ring with unity where every nonzero element is invertible (with respect to $\times$). It is usually denoted as $\mathbb{F}$. Note that F is now an abelian group with respect to $\times$.

Definition 1.16. 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 $xa = 0$.

Definition 1.17. A ring R with no zero divisors for every element is called a *domain*.

Proposition 1.3. *Every field is a domain.*

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

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

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

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

■

While the converse is not true, we can state the following result.

Theorem 1.4 (Wedderburn's little theorem)**.** *Every finite domain is a field.*

1.3 Vector Space Structures

Definition 1.18. A *vector space over a field* F consists of an abelian group $(V, +)$ and an operation called *scalar multiplication*

$$\cdot : F \times V \to V$$

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

$$\cdot : R \times M \longrightarrow M$$

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

$$\cdot : M \times R \longrightarrow M$$

Definition 1.21. Let A be a vector space over a field F equipped with an additional binary operation

$$\times : A \times A \longrightarrow A$$

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 1.11. *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 1.12. $\mathbb{R}^3$ *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 1.13. *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 1.22. 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

$$A = bx$$

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

Definition 1.23. 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 $\mathbb{R}$.

1.4 Subgroups, Subrings, Subfields

Definition 1.24. Given a set M and a subset $N \subseteq M$, the subset N is closed with respect to $*$ if $a, b \in N \implies a * b \in N$

Definition 1.25. A *subgroup of group G* is a group that is a subset G. The *trivial subgroups* of a group G are 0 and G.

Example 1.14. $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ *are all groups.*

Theorem 1.5. *The subgroup $(N, *)$ of every abelian group $(M, *)$ is also an abelian group.*

Corollary 1.5.1. *Any subspace within a vector space is a subgroup.*

Definition 1.26. A subset L of a ring R is a *subring* if and only if it is a ring.

Example 1.15. *For any $n \in \mathbb{Z}_+$, the set $n\mathbb{Z}$ is a subring of $\mathbb{Z}$.*

Definition 1.27. A *subfield of field F* is a field that is a subset of F.

Example 1.16. $\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2 Group Theory

2.1 Classes of Groups

2.1.1 Symmetric Group

Definition 2.1. The *symmetric group*, also called 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 $n!$.

Proposition 2.1. *Every element in finite S_n can be decomposed into a partition of cyclic rotations.*

Example 2.1. *1.* (12) *is a mapping* $1 \to 2,\ 2 \to 1$.

2. (123) *is a mapping* $1 \to 2,\ 2 \to 3,\ 3 \to 1$.

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

Definition 2.2. 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 2.2. *1.* $(123)(45)$ *is conjugate to* $(143)(25)$.

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

Definition 2.3. The *signature* of a permutation is a homomorphism

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

Proposition 2.2. *The signature of a permutation changes for every transposition that is applied to it.*

Definition 2.4. 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 2.3 (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.$S \simeq Dih(3)$, 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.

Definition 2.5. A *permutation group* is some subgroup of $\text{Sym}(X)$.

2.1.2 General Linear and Affine Groups

Definition 2.6. 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 2.7. The *general affine group* is the pair of all transformations

$$\text{GA}(V) \equiv \text{Tran}(V) \times \text{GL}(V)$$

2.1.3 Isometries

Definition 2.8. 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 2.9. 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 2.10. The *orthogonal group*, denoted $\text{O}(n)$ or O_n, consists of all isometries that preserve the origin, i.e. consists of rotations and reflections. The *special orthogonal group*, denoted $\text{SO}(n)$, is a subgroup of $\text{O}(n)$ consisting of only rotations. We can see that

$$\text{O}(n) = \frac{\text{Isom}\, \mathbb{E}^n}{\text{Tran}\, V}$$

2.1.4 Geometrical Groups

Definition 2.11. 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 2.12. 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

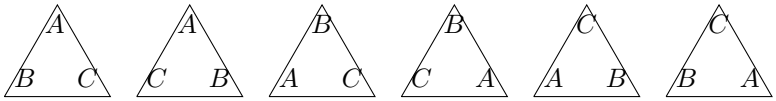
$$\{u_1 - u_0, u_2 - u_0, ..., u_n - u_0 | \{u_i\}_{i=0}^n \text{ vertices}\}$$

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

Definition 2.13. 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. A common example of such groups is the *dihedral group*, denoted D_n or $\mathrm{Dih}(n)$, which is the group of symmetries of a n-simplex, which includes rotations and reflections.

Example 2.4. *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* $\mathbb{R}^n$*.*

2.2 Direct Product of Groups

Definition 2.14. The *direct product* of two groups G and H is denoted

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

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

Definition 2.15. The *general affine group* is defined

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

Definition 2.16. 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

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

where H is the group of transformations of the form

$$(x, y, z, t) \longmapsto (x + at, y + bt, z + ct, t)$$

Definition 2.17. The *Poincaré Group* is the symmetry group of spacetime within the principles of relativistic mechanics, denoted

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

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

$$x^2 + y^2 + z^2 - t^2$$

2.3 Generating Sets and Group Presentations

2.3.1 Cyclic Groups

Definition 2.18. A *word* is any written product of group elements and inverses. They are generally in the form

$$s_1^{\epsilon_1} s_2^{\epsilon_2} s_3^{\epsilon_3} ... s_k^{\epsilon_k}$$

Example 2.5. *Given a set* $\{x, y, z\}$, $xy, xz^{-1}yyx^{-2}, ...$ *are words.*

Definition 2.19. 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 2.20. A *cyclic group*, denoted C_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 2.6. *A representation of a cyclic group of* n*th order is the* n*th roots of unity in* $\mathbb{C}$.

Example 2.7. *Another representation of a cyclic group of* n*th order is the set of discrete angular rotations in* $SO(2)$*, in the form of*

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

Example 2.8. $\mathbb{Z}$ *is an infinite cyclic group with generator* 1*. Furthermore,* $\mathbb{Z}/m\mathbb{Z}$ *is a finite cyclic group with generator* 1*. In fact, the generator of* $\mathbb{Z}/m\mathbb{Z}$ *can be any integer relatively prime to* m *(and less than* m*).*

Example 2.9. *The set of all transpositions forms a generating set of* S_n.

It is actually a fact that every finite cyclic group of order m is isomorphic to $\mathbb{Z}/m\mathbb{Z}$. Every infinite cyclic group is isomorphic to $\mathbb{Z}$. This implies that any two cyclic group of the same order are isomorphic, since we can define a mapping $f : a \longrightarrow b$, where a and b are generating elements of their respective groups.

Example 2.10. $Dih(3) \simeq S_3$*, since permutations of the vertices of a triangle are isomorphic to a permutations of a 3-element set.*

Definition 2.21. The *free group* F_S over a given set S consists of all words that can be built from elements of S. Clearly, S is the generating set of F_S.

2.3.2 Group Presentations

One method of specifying a group is to put it in the form

$$\langle\, S \mid R \,\rangle$$

where S is the generating set and R is a set of relations.

Example 2.11. *The cyclic group of order* n *could be presented as*

$$\langle\, a \mid a^n = 1 \,\rangle$$

Example 2.12. *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*

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

2.4 Cayley's Theorem

Lemma 2.3. *Let G be a group with $a \in G$. We define the map*

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

Then, ϕ is an automorphism of G.

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. ■

Theorem 2.4 (Cayley's Theorem)**.** *Every group G is isomorphic to a subgroup of its symmetric group. If G is finite, then so is 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

$$\phi : G \longrightarrow H$$

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 $g \in G, \phi(a)\phi(b)(g) = \phi(a)(bg) = a(bg) = (ab)g = \phi(ab)(g)$. ■

2.5 Group Actions

Definition 2.22. Let G be a group, X a set. Then, a (left) group action of G on X is a function:

$$\varphi : G \times X \longrightarrow X,\ (g, x) \longmapsto \varphi(g, x)$$

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

Example 2.14. *S_n acts on $\{1, 2, ..., n\}$ by permuting its elements.*

Example 2.15. *The $GA(V)$ 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 $\text{Aut}(X)$.

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

Definition 2.23. A group action on a finite-dimensional vector space X is called a *representation* of that group.

2.6 Equivalence and Congruence

Definition 2.24. 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 2.16. *Tran(V) and GA(V) are transitive groups.*

Definition 2.25. 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 $\mathrm{E}(n) \subset \mathrm{GA}(n)$.

2.7 Cosets and Lagrange's Theorem

Definition 2.26. Given a group G and a subgroup H, g_1 and g_2 are congruent modulo H, denoted $g_1 \equiv g_2 \pmod{H}$. The equivalence classes are known as *cosets*. A coset of somprised of all the products obtained by multiplying each element of H by a particular element in G. Since group multiplication is not necessarily commutative, we must distinguish between right and left cosets.

1. A *left coset* is
$$gH \equiv \{gh \mid h \in H\}$$
2. A *right coset* is
$$Hg \equiv \{hg \mid h \in H\}$$

It is easy to see that the cosets form a partition of the set X, with each coset of the same cardinality.

Definition 2.27. A subgroup $N \subset G$ is a *normal subgroup* iff the left cosets equal the right cosets. Every subgroup of an abelian group is normal.

Theorem 2.5 (Lagrange's Theorem)**.** *Let G be a finite group and H its subgroup. Then*

$$|G| = |G : H||H|$$

where $|G : H|$ is the number of cosets in G.

Corollary 2.5.1. *The order of a subgroup of a finite group divides the order of the group.*

Definition 2.28. The order of an element is the order of the cyclic subgroup that it generates.

Corollary 2.5.2. *The order of any element of a finite group divides the order of the group.*

Corollary 2.5.3. *Every finite group of a prime order is cyclic.*

Theorem 2.6 (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,*

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

Corollary 2.6.1. *If $|G| = n$, then $g^n = e$ for all $g \in G$.*

Definition 2.29. *Euler's Totient Function*, denoted $\varphi(n)$, consists of all the numbers less than or equal to n that are coprime to n.

Theorem 2.7 (Euler's Theorem (Generalization of Fermant's Little 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 $G = \mathbb{Z}_n \setminus \{0\}$,*

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

Example 2.17. *In $\mathbb{Z}_{125} \setminus \{0\}$, $\varphi(125) = 125 - 25 = 100 \implies 2^{100} \equiv 1 \pmod{125}$*

Definition 2.30. 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

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

By definition, transitive transformation groups have only one orbit.

Definition 2.31. The subgroup $G_x \subset G$, where $G_x \equiv \{g \in G | gx = x\}$ is called the *stabilizer* of x.

Example 2.18. *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 $\overrightarrow{0p}$. The stabilizer of 0 is the entire $O(2)$.*

Example 2.19. *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 2.8. *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 2.32. The *length of an orbit* is the number of elements in it.

Corollary 2.8.1. *If G is a finite group, then*

$$|G| = |G_x||Gx|$$

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

$$G_{gx} = gG_xg^{-1}$$

2.8 Abelian Groups

First, note that the successive addition of elements of an additive abelian group can be represented by integer multiplication.

$$x + x + ... + x = nx, \ n \in \mathbb{Z}$$

Similarly, we can take the integer power of an element to represent successive multiplication in a multiplicative abelian group.

Proposition 2.9. *It is easy to check that in an additive abelian group A, with $a, b \in A$ and $k, l \in \mathbb{Z}$,*

$$k(a + b) = ka + kb \tag{1}$$
$$(k + l)a = ka + la \tag{2}$$
$$(kl)a = k(la) \tag{3}$$

which implies

$$k(a - b) = ka - kb, \ (k - l)a = ka - la \tag{4}$$

Definition 2.33. For any subset $S \subset A$, the collection of all linear combinations

$$k_1a_1 + k_2a_2 + ... + k_na_n, \ k_i \in \mathbb{Z}, a_i \in S$$

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 2.34. 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 2.35. A system $\{a_1, a_2, ..., a_n\}$ of elements of a group A is called *linearly independent* if $k_1a_1 + k_2a_2 + ... + k_na_n = 0 \implies k_1, k_2, ..., k_n = 0$. 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 2.36. A finitely generated abelian group is *free* if it has a basis.

Theorem 2.10. *All bases of a free abelian group L contain the same number of elements.*

Definition 2.37. 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 2.11. *Every free abelian group L of rank n is isomorphic to the group $\mathbb{Z}^n$ of integer rows of length n.*

Theorem 2.12. *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 2.38. 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.

Proposition 2.13. *A subgroup $L \subset E^n$ is discrete if and only if its intersection with any neighborhood of 0 consists of 0 itself.*

Theorem 2.14. *Every discrete subgroup $L \subset E^n$ is generated by a linearly independent system of vectors of E^n.*

Corollary 2.14.1. *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

$$\Gamma \equiv \text{Dih}\, C \times \mathcal{L}$$

where $\text{Dih}\, C$ is the Dihedral group of the crystal structure that preserves its lattices.

Definition 2.39. 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.

Proposition 2.15. *Every integral rectangular matrix $C = (c_{ij})$ can be reduced by integral elementary row transformations to the diagonal matrix $diag(u_1, ..., u_p)$, where $u_1, u_2, ..., u_p \geq 0$ and $u_i | u_{i+1}$ for $i = 1, 2, ..., p - 1$.*

Example 2.20. *The following matrix can be reduced (with a few steps now shown) to the stated form.*

$$\begin{pmatrix} 2 & 6 & 2 \\ 2 & 3 & 4 \\ 4 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 2 & 3 & 4 \\ 0 & -3 & 2 \\ 4 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 6 & 14 \\ 0 & 8 & 12 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 20 \end{pmatrix}$$

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.

Proposition 2.16. *Given square integral matrix C with reduced form $diag(u_1, ..., u_p)$,*

$$u_i = \frac{d_i}{d_{i-1}}$$

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 2.17. *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 $i = 1, 2, ..., m-1$.*

3 Ring Theory

3.1 Field of Complex Numbers

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.

Definition 3.1. The *field of complex numbers* is a field $\mathbb{C}$ such that

1. It contains the field $\mathbb{R}$ as a subfield.
2. It contains an element i such that $i^2 = -1$.
3. It is minimal with respect to properties (i) and (ii). That is, if F is a subfield of $\mathbb{C}$ containing $\mathbb{R}$ and i, then $F = \mathbb{C}$.

Note that the identity $x^2 + 1 \equiv (x + i)(x - i)$ implies that the equation $x^2 = -1$ has exactly two solutions in $\mathbb{C}$, i and $-i$. Therefore, if a subfield of $\mathbb{C}$ contains one of these solutions, it must contain the other (since i and $-i$ are additive and multiplicative inverses).

Furthermore, since i is defined to be $\sqrt{-1}$, we could replace i with $-i$ and our calculations would still be consistent throughout the rest of mathematics. In fact, i and $-i$ behave *exactly* identically and cannot be distinguished in an abstract sense. Visually, the complex plane "flipped" across the real number axis produces the same complex plane.

Theorem 3.1. $\mathbb{C}$ *exists and is unique up to an isomorphism that maps all real numbers to themselves. Every complex number can be uniquely written as $a + bi$, where $a, b \in \mathbb{R}$ and i is a fixed element such that $i^2 = -1$.*

Proof. We first assume that $\mathbb{C}$ exists. Consider the subset of $\mathbb{C}$

$$K \equiv \{a + bi \mid a, b \in \mathbb{R}\}$$

By evaluating its operations, we can check for closure, identity, and invertibility of nonzero elements to conclude that K is a subfield of $\mathbb{C} \implies$ by prop. (iii), $K = \mathbb{C} \implies$ every element in $\mathbb{C}$ can be written in form $a + bi$. To prove uniqueness, we assume that $p \in \mathbb{C}$ can be written in distinct forms $p = a + bi = a' + b'i$. Then

$$\begin{aligned} a + bi = a' + b'i &\implies (a - a')^2 = (b'i - bi)^2 = -(b' - b)^2 \\ &\implies a - a' = b' - b = 0 \end{aligned}$$

To prove uniqueness of $\mathbb{C}$ up to ismorphism, we assume that $\mathbb{C}'$ exists with i' such that i'^2 containing elements $a + bi'$. Let $f : \mathbb{C} \longrightarrow \mathbb{C}'$ defined

$$f(a + bi) = a + bi'$$

Then,

$$\begin{aligned} f\big((a + bi) + (c + di)\big) &= f\big((a + c) + (b + d)i\big) \\ &= (a + c) + (b + d)i' \\ &= (a + bi') + (c + di') \\ &= f(a + bi) + f(c + di) \\ f\big(\kappa(a + bi)\big) &= f\big(\kappa a + \kappa bi\big) \\ &= \kappa a + \kappa bi' \\ &= \kappa(a + bi') \\ &= \kappa f(a + bi) \end{aligned}$$

So, f is an isomorphism, and $\mathbb{C} \simeq \mathbb{C}'$. From analysis, we can construct and prove the existence of $\mathbb{R}$. We then define the map

$$\rho : \mathbb{R}^2 \longrightarrow \mathbb{C},\ \rho(a, b) \equiv a + bi$$

with $\rho(1, 0)$ as the multiplicative identity and $\rho(0, 1) \equiv i$. Therefore, every element of $\mathbb{C}$ can be uniquely represented as an element of $\mathbb{R}^2$. ■

Definition 3.2. *Complex conjugation* is an automorphism of $\mathbb{C}$ defined

$$c = a + bi \mapsto \bar{c} = a - bi$$

This is identically defined by replacing i with $-i$. Clearly, $\bar{\bar{c}} = c$.

Definition 3.3. Real numbers are elements in $\mathbb{C}$ that are equal to their own conjugates.

Proposition 3.2. *For any $c \in \mathbb{C}$, $c + \bar{c}$ and $c\bar{c}$ are real.*

Proof. Using the fact that the complex conjugate is an isomorphism,

$$\begin{aligned} c \bar{+} \bar{c} &= \bar{c} + \bar{\bar{c}} = \bar{c} + c = c + \bar{c} \\ c\bar{c} &= \bar{c}\bar{\bar{c}} = \bar{c}c = c\bar{c} \end{aligned}$$

■

Note that we proved this abstractly using only the properties given above, and did not decompose c to its *algebraic form* $a + bi$.

If $c = a + bi,\ a, b \in \mathbb{R}$, then

$$c + \bar{c} = 2a,\ c\bar{c} = a^2 + b^2$$

In case the reader is unaware, it is common to interpret complex numbers $c = a + bi$ as points or vectors (a, b) on the complex plane.

3.1.1 Polar Representations of Complex Numbers

Definition 3.4. The *absolute value* of a complex number $c = a + bi$, denoted $|c|$, is the length of the vector representing c.

$$|c| \equiv \sqrt{a^2 + b^2}$$

Definition 3.5. The *argument* of a complex number $c = a + bi$, denoted arg$c$, is the angle formed by the corresponding vector with the polar axis. It is defined within the interval $[0, 2\pi)$.

$$\arg(c) \equiv \tan^{-1} \frac{b}{a}$$

Definition 3.6. The *polar representation*, or *trigonometric representation*, of a complex number $c = a + bi$ is defined using the equations

$$a = r\cos\varphi,\ b = r\sin\varphi \implies c = r(\cos\varphi + i\sin\varphi)$$

This mapping can be defined

$$\rho : \mathbb{R} \times \frac{\mathbb{R}}{2\pi} \longrightarrow \mathbb{C},\ \rho(r, \varphi) = r(\cos\varphi + i\sin\varphi)$$

Theorem 3.3. *ρ is "similar" to a homomorphism in the following way. By defining the domain and codomain as groups,*

$$\rho : (\mathbb{R}, \times) \times \left(\frac{\mathbb{R}}{2\pi}\right) \longrightarrow (\mathbb{C}, \times)$$

we can see that

$$\rho(r_1, \varphi_1) \times \rho(r_2, \varphi_2) = \rho(r_1 \times r_2, \varphi_1 + \varphi_2)$$

or equivalently,

$$r_1(\cos\varphi_1 + i\sin\varphi_1) \cdot r_2(\cos\varphi_2 + i\sin\varphi_2) = r_1r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2))$$

Corollary 3.3.1. *The formula for the ratio of complex numbers is defined*

$$\frac{r_1(\cos\varphi_1 + i\sin\varphi_1)}{r_2(\cos\varphi_2 + i\sin\varphi_2)} = \frac{r_1}{r_2}(\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2))$$

Corollary 3.3.2. *The positive integer power of a complex number can be written using De Moivre's formula.*

$$\big(r(\cos\varphi + i\sin\varphi)\big)^n = r^n(\cos n\varphi + i\sin n\varphi)$$

We can use this formula to extract a root of nth degree from a complex number $c = r(\cos\varphi + i\sin\varphi)$, which means to solve the equation $z^n = c$. Let $z = s(\cos\psi + i\sin\psi)$. Then by De Moivre's formula,

$$\begin{aligned} z^n &= s^n(\cos n\psi + i\sin n\psi) = r(\cos\varphi + i\sin\varphi) \\ &\implies s = \sqrt[n]{r},\ \psi = \frac{\varphi + 2\pi k}{n} \\ &\implies z = \sqrt[n]{r}\left(\cos\frac{\varphi + 2\pi k}{n} + i\sin\frac{\varphi + 2\pi k}{n}\right) \text{ for } k = 0, 1, ..., n-1 \end{aligned}$$

Geometrically, the n solutions lie at the vertices of a regular n-gon centered at the origin. When $c = 1$, the solutions are the nth roots of unity.

3.2 Rings of Residue Class

Definition 3.7. The quotient set $\mathbb{Z}$ by the relation of congruence modulo n is denoted $\mathbb{Z}_n$. It is called the *ring of residue class modulo n* or *residue ring modulo n.*

$$\mathbb{Z}_n = \{[0]_n, [1]_n, ..., [n-1]_n\}$$

By definition of the relation, congruence modulo n has properties:

1. $a \equiv a' \pmod n, b \equiv b' \pmod n \implies a + b \equiv a' + b' \pmod n$.
2. With same hypothesis as (i) $ab \equiv a'b \equiv ab' \equiv a'b' \pmod n$.

We can furthermore define operations of addition and multiplication on the ring $\mathbb{Z}_n$ as such

$$\begin{aligned} [a]_n + [b]_n &\equiv [a + b]_n \\ [a]_n[b]_n &\equiv [ab]_n \end{aligned}$$

making $\mathbb{Z}_n$ is a commutative, associative ring with unity.

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 3.1. *In $\mathbb{Z}_5$, 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.*

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 3.4. *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

$$[k]_n[l]_n = [kl]_n = [n]_n = 0$$

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

$$[0]_n, [a]_n, [2a]_n, ..., [(n-1)a]_n$$

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

Definition 3.8. The *characteristic* of ring R (or a field F), denoted char(R), is the smallest number of times one must successively add the multiplicative identity 1 to get the additive identity 0. That is char(R) is the smallest positive number n such that

$$1 + 1 + ... + 1 = 0$$

If no such number n exists, then char(R) = 0. The characteristic of $\mathbb{Z}_n = n$

Note that the characteristic of the field $\mathbb{Z}_n$ must be prime.

Theorem 3.5 (Freshman's Dream)**.** *Given a field F with char(F) = p,*

$$(a+b)^p = a^p + b^p$$

Proof.

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

It is clear that

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

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

3.3 Polynomial Algebra

3.3.1 Construction and Basic Properties

Definition 3.9. A *polynomial* f of x over a ring R is defined as a formal expression

$$f(x) = a_0 + a_1 x^1 + a_2 x^2 + ... + a_{n-1} x^{n-1} + a_n x^n \tag{5}$$

where n is a natural number, the coefficients $a_0, a_1, ..., a_n$ are elements of R, x is 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 infinite finitary sequence.

$$(a_0, a_1, a_2, ..., a_n, 0, 0, ...) \tag{6}$$

Two polynomials are equal if and only if the sequences of their corresponding coefficients are equal.

Note that this is really just a fancy way to write a finitary sequence.

Definition 3.10. The set of polynomials with coefficients in the ring R forms itself a ring, called the *ring of polynomials over* R, denoted $R[x]$. Addition on $R[x]$ is defined component-wise, and it suffices, by the distributive law, to define multiplication as

$$x^k x^l = x^{k+l}$$

given that we have chosen $\{x^i\}$ as a basis of $R[x]$. If R is a commutative associative ring (or a field), then $R[x]$ is called the *polynomial algebra*. From now, we will treat $R[x]$ and $F[x]$ as an algebra with R denoting a commutative associative ring and F denoting a field, respectively.

Note that the map from $R \longrightarrow R[x]$ sending $r \mapsto rx^0$ is an injective homomorphism of rings, by which R is viewed as a subring of $R[x]$.

The ring of polynomials over field $\mathbb{R}$ is denoted $\mathbb{R}[x]$. $R[x]$ is a subalgebra within the algebra of all function of $\mathbb{R}$.

However, for certain finite fields, some formally different polynomials may be indistinguishable in terms of mappings. For example, x and x^2 are equivalent in the polynomial algebra defined on the domain $\mathbb{Z}_2$.

Definition 3.11. The last nonzero coefficient is called the *leading coefficient*, and the degree of the polynomial f, denoted $\deg f$, is the index of the leading coefficient.

Theorem 3.6.

$$deg(f+g) \leq max\{deg\, f, deg\, g\} \tag{7}$$
$$deg\, fg = deg\, f + deg\, g \tag{8}$$

Proof. Simple when presenting polynomials if form (1). ■

Definition 3.12. The product of two finitary sequences $(a_0, a_1, a_2, ...)$ and $(b_0, b_1, b_2, ...)$ in the ring $F[x]$ is a sequence

$$(c_0, c_1, c_2, ...),\ c_k = \sum_{l=0}^{k} a_l b_{k-l}$$

This formula works for infinite (non-finitary) sequences too, allowing us to define a commutative, associative algebra with unity called the *algebra of formal power series over* F, denoted $F[[x]]$. The elements of $F[[x]]$ are written in the form

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 ...$$

Theorem 3.7. *If the field F is infinite, then different polynomials in $F[x]$ determine different functions.*

Theorem 3.8. *For any collection of given 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*

$$f(x_i) = y_i,\ i = 1, 2, ..., n$$

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

It is usually impossible to divide one polynomial by another in the algebra $F[x]$; the construction of it does not allow us to. However, division *with remainder* is possible, similarly to the procedure of division with remainder in the ring of integers.

Theorem 3.9. *Let $f, g \in F[x]$ and $g \neq 0$. Then, there exists polynomials q, r such that*

$$f = qg + r,\ deg\, r < deg\, g\ \ (or\ r = 0)$$

This procedure of finding such polynomials q, r is called division with a remainder. A polynomial f is divisible by g in $F[x]$ if and only if $r = 0$.

Theorem 3.10 (Bezout's Theorem)**.** *Given that one divides (with remainder) polynomial f by $g = x - c$, let the remainder be $r \in F$. That is,*

$$f(x) = (x - c)q(x) + r,\ r \in F$$

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

$$f(c) = r$$

3.3.2 Roots of Polynomials

Definition 3.13. An element $c \in F$ is a *root* of polynomial f if and only if

$$f(c) = 0$$

Corollary 3.10.1. *An element c of a field F is a root of polynomial f if and only if f is divisible by $x - c$.*

Definition 3.14. A root c of polynomial f is called *simple* if f 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.

Theorem 3.11. *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.*

Definition 3.15. A *monic polynomial* is a polynomial with leading coefficient equal to 1.

Theorem 3.12 (Viete's Formulas)**.** *Given that a polynomial f factors into linear terms, that is*

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

Then the coefficients of f can be presented with the formulas

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

Theorem 3.13 (Wilson's Theorem)**.** *Let n be a prime number. Then*

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

Definition 3.16. The *derivative* of a polynomial is a map $D : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ with the following properties:

1. It is linear.
2. $D(fg) = (Df)g + f(Dg)$.
3. $Dx = 1$.

In fact, there exists a unique map $D : F[x] \longrightarrow F[x]$ satisfying these properties for any field F.

Proposition 3.14. *If $charF = 0$, then the coefficients of $f \in F[x]$ regarded as a polynomial in $x - c$ can be expressed as*

$$b_k = \frac{f^{(k)}(c)}{k!} \tag{9}$$

where $f^{(k)}$ is the kth derivative of f.

Proof. We make the substitution $y = x - c$ in the polynomial $f \in F[x]$ and then express it as a polynomial in y

$$f = b_0 + b_1(x-c) + b_2(x-c)^2 + \ldots + b_n(x-c)^n \tag{10}$$

We differentiate this equation k times and substitute at $x = c$ to get the corresponding values of the coefficients. ■

3.3.3 Fundamental Theorem of Algebra of Complex Numbers

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. Fortunately, it is sufficient to extend the field to $\mathbb{C}$ in order to strongly define a lower limit, too.

Definition 3.17. 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.

Proposition 3.15. *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

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

such that the characteristic polynomial of A is p. ■

Proposition 3.16. *$\mathbb{R}$ is not algebraically closed.*

Proof. $x^2 + 1$ doesn't have any roots in $\mathbb{R}$. ■

Theorem 3.17. *Every polynomial of positive degree over field $\mathbb{C}$ has a root.*

Corollary 3.17.1. *In the algebra $\mathbb{C}[x]$, every polynomial splits into a product of linear factors.*

Corollary 3.17.2. *Every polynomial of degree n over $\mathbb{C}$ has n roots, counted with multiplicities.*

Corollary 3.17.3. *$\mathbb{C}$ is algebraically closed.*

3.3.4 Roots of Polynomials with Real Coefficients

Theorem 3.18. *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 3.18.1. *Every nonzero polynomial in $\mathbb{R}[x]$ factors into a product of linear terms and quadratic terms with negative discriminants.*

Example 3.2.

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

Corollary 3.18.2. *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

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

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

Theorem 3.19 (Descartes' Theorem)**.** *The number of positive roots (counted with multiplicities) of a polynomial $f \in \mathbb{R}[x]$ (denote this $N(f)$) does not exceed the number of changes of sign in the sequence of its coefficients (denote this $L(f)$). Additionally, $L(f) \equiv N(f) \pmod 2$. If all the complex roots of f are real, then $L(f) = N(f)$.*

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.

3.3.5 Factorization in Euclidean Domains

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.

Definition 3.18. A commutative associative ring with unity and without zero divisors is called an *integral domain*. That is, the product of any two nonzero elements $x, y \in A$ must be nonzero. Integral domains are generalizations of the ring of integers $\mathbb{Z}$ and provide a natural setting for studying divisibility.

Example 3.3. *$\mathbb{Z}$ and $F[x]$ over field F are integral domains. Any field F is also an integral domain.*

Example 3.4. *The quotient ring $\mathbb{Z}_n$ is not an integral domain when n is composite.*

Example 3.5. *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 3.6. *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 3.7. *The ring of continuous functions on the interval is not an integral domain. To see why, notice that given the piecewise functions*

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

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

We can classify the rings

$$\text{Integral Domains} \subset \text{Commutative Rings} \subset \text{Rings}$$

Proposition 3.20. *An integral domain is a ring that is isomorphic to a subring of a field.*

Proposition 3.21. *The characteristic of an integral domain is either 0 or a prime number.*

Definition 3.19. An element r of a ring R is *regular* if the mapping

$$\rho : R \longrightarrow R, \ x \mapsto xr$$

is injective for all $x \in R$.

Proposition 3.22. *An integral domain is a commutative associative ring where every element is regular.*

Definition 3.20. Let A be an integral domain. An element $a \in A$ is *divisible* by $b \in A$, denoted $b|a$ if there exists an element $q \in A$ such that $a = qb$. 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.

Definition 3.21. Let A be an integral domain which is not a field. A is *Euclidean* if there exists a function

$$N : A \setminus \{0\} \longrightarrow \mathbb{Z}_+$$

called a *norm* that satisfies the following conditions.

1. $N(ab) \geq N(a)$ and the equality holds if and only if b is invertible.
2. For any $a, b \in A,\ b \neq 0$, there exist $q, r \in A$ such that $a = qb + r$ with either $r = 0$ or $N(r) < N(b)$, known as division with remainder.

Uniqueness of q, r is not required in property 2.

Example 3.8. *The subring of $\mathbb{C}$, defined*

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

is a Euclidean integral domain with respect to the norm

$$N(c) \equiv a^2 + b^2$$

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

Example 3.9. *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*

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

and the norm is defined

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

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

Definition 3.22. The *greatest common divisor* of elements a and b of an integral domain is a common divisor of a and b divisible by all their common divisors. It is denoted $\text{GCD}(a, b)$.

Definition 3.23. A *Gaussian integer* is a complex number whose real part and imaginary part are both integers. That is,

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

3.3.6 Polynomials in Several Variables

Definition 3.24. A function of real variable $x_1, x_2, ..., x_n$ is called a *polynomial* if it can be represented as

$$f(x_1, ..., x_n) = \sum_{k_1,...,k_n} a_{k_1...k_n} x_1^{k_1} x_2^{k_2} ... x_n^{k_n}$$

where the summation is taken over a finite set of collections $(k_1, ..., k_n)$. The algebra of polynomials in $x_1, x_2, ..., x_n$ over $\mathbb{R}$ is denoted $\mathbb{R}[x_1, x_2, ..., x_n]$.

Definition 3.25. More generally, an infinite dimensional polynomial algebra of variables $x_1, ..., x_n$ over field $\mathbb{F}$ is denoted

$$\mathbb{F}[x_1, ..., x_n]$$

Like polynomials of one variable, it can be naturally identified with an abstract multi-dimensional "sequence." It has basis

$$\{e_{k_1 k_2 ... k_n} \mid k_1, k_2, ..., k_n \in \mathbb{Z}_+$$

with addition defined component-wise and the multiplication rule defined with the table

$$e_{k_1...k_n} e_{l_1...l_n} = e_{k_1+l_1, k_2+l_2, ..., k_n+l_n}$$

Clearly each polynomial in its usual presentation is gotten by the linear mapping

$$e_{k_1...k_n} \mapsto x_1^{k_1} x_2^{k_2} ... x_n^{k_n}$$

However, note that different polynomials may define the same functions if the field $\mathbb{F}$ is finite, similarly to polynomials with one variable. If $\mathbb{F}$ is infinite, then every polynomial will determine a different function.

Definition 3.26. A polynomial is called *homogeneous* if degree d if

$$a_{k_1 k_2 ... k_n} = 0 \text{ for } k_1 + k_2 + ... + k_n \neq d$$

The space of all homogeneous polynomials of fixed degree d forms a finite dimensional subspace in $\mathbb{F}[x_1, ..., x_n]$ with dimension

$$\frac{n(n+1)...(n+d-1)}{d!}$$

The dimension can be calculated by thinking of the combinatorics problem of having d indistinguishable balls to put into n distinguishable urns.

3.3.7 Symmetric Polynomials

Definition 3.27. A polynomial $f \in \mathbb{F}[x_1, ..., x_n]$ is called *symmetric* if it is invariant under any permutation of the variables x_i.

Example 3.10. *Power sums are symmetric polynomials.*

$$p(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i^k$$

Definition 3.28. An *elementary symmetric polynomial* is a symmetric polynomial of one of these forms:

$$\begin{aligned}
\sigma_1 &= x_1 + x_2 + ... + x_n \\
\sigma_2 &= x_1x_2 + x_1x_3 + ... + x_{n-1}x_n \\
... &= ... \\
\sigma_k &= \sum_{i_1 < ... < i_k} x_{i_1} x_{i_2} ... x_{i_k} \\
... &= ... \\
\sigma_n &= x_1x_2...x_n
\end{aligned}$$

The following theorem presents an extremely useful result about the decomposition of symmetric polynomials.

Theorem 3.23. *Every symmetric polynomial can be written as a polynomial of elementary symmetric polynomials σ_i.*

Example 3.11. *The polynomial*

$$f \equiv \sum_{i=1}^{n} x_i^3$$

can be expressed as

$$f = \sigma_1^3 - 3\sigma_1\sigma 2 + 3\sigma_3$$

3.3.8 Cubic Equations

The well known discriminant of a quadratic equation

$$f(x) = ax^2 + bx + c$$

is known in the form $\nabla = b^2 - 4ac$. However, we will present it in a slightly different manner.

Definition 3.29. The *discriminant* $D(\varphi)$ of a quadratic polynomial

$$\varphi = a_0x^2 + a_1x + a_2 \in \mathbb{C}[x]$$

with $c_1, c_2 \in \mathbb{C}$ as its roots is defined

$$D(\varphi) = a_1^2 - 4a_0a_2 = a_0^2\left(\left(\frac{a_1}{a_0}\right)^2 - \frac{4a_2}{a_0}\right) = a_0^2\big((c_1+c_2)^2 - 4c_1c_2\big) = a_0^2(c_1-c_2)^2$$

Clearly, the value of $D(\varphi)$ can tell us three things

1. $c_1, c_2 \in \mathbb{R}, c_1 \neq c_2$. Then $c_1 - c_2$ is a nonzero real number and $D(\varphi) > 0$.
2. $c_1 = c_2 \in \mathbb{R}$. Then $c_1 - c_2 = 0$ and $D(\varphi) = 0$.
3. $c_1, c_2 \in \mathbb{C}, c_1 = \bar{c}_2$. Then, $c_1 - c_2$ is a nonzero strictly imaginary number and $D(\varphi) < 0$.

Definition 3.30. We can generalize this notion of the discriminant to arbitrary polynomials

$$\varphi = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n \in \mathbb{F}[x],\ a_0 \neq 0$$

The discriminant $D(\varphi)$ of the polynomial above is defined

$$D(\varphi) \equiv a_0^{2n-2} \prod_{i>j} (c_i - c_j)^2$$

The a_0 term isn't very important in this formula, since it does not affect whether $D(\varphi)$ is positive, negative, or zero.

Definition 3.31. A polynomial

$$\varphi = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n \in \mathbb{F}[x],\ a_0 \neq 0$$

where $a_1 = 0$ is called *depressed*. A depressed cubic polynomial is of form

$$\varphi = x^3 + px + q$$

Proposition 3.24. *Every monic (leading coefficeint = 1) polynomial (and non-monic ones)*

$$\varphi = x^n + a_1 x^{n-1} + ... + a_{n-1} x + a_n \in \mathbb{F}[x],\ a_0 \neq 0$$

can be turned into a depressed polynomial with the change of variable

$$x = y - \frac{a_1}{n}$$

to get the polynomial

$$\psi = y^n + b_2 y^{n-2} + ... + b_{n-1} y + b_n$$

Lemma 3.25. *A cubic polynomial*

$$\varphi = a_0 x^3 + a_1 x^2 + a_2 x + a_3 \in \mathbb{R}[x]$$

with roots $c_1, c_2, c_3 \in \mathbb{C}$ has discriminant

$$D(\varphi) \equiv a_0^4 (c_1 - c_2)^2 (c_1 - c_3)^2 (c_2 - c_3)^2$$

With a bit of evaluation, it can also be expressed in terms of its coefficients as

$$D(\varphi) = a_1^2 a_2^2 - 4a_1^3 a_3 - 4a_0 a_2^3 + 18 a_0 a_1 a_2 a_3 - 27 a_0^2 a_3^2$$

Again, three possibilities can occur (up to reordering of its roots).

1. *c_1, c_2, c_3 are distinct real numbers. Then $D(\varphi) > 0$.*
2. *$c_1, c_2, c_3 \in \mathbb{R}, c_1 = c_2$. Then $D(\varphi) = 0$.*
3. *$c_1 \in \mathbb{R}, c_2 = \bar{c}_3 \notin \mathbb{R}$. Then $D(\varphi) < 0$.*

Furthermore, the cubic formula used to find the roots of the polynomial is

$$c_{1,2,3} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{p^3}{27} + \frac{q^2}{4}}}$$

known as Cardano's formula, after the mathematician Gerolamo Cardano.

3.4 Ideals and Quotient Rings

Definition 3.32. For an arbitrary ring $(R, +, \cdot)$, let $(R, +)$ be its additive group. A subset I is called a *left ideal* of R if it satisfies the two conditions.

1. $(I, +)$ is a subgroup of $(R, +)$.
2. For every $r \in R$ and every $x \in I$ the left product $r \cdot x \in I$.

Similarly, a *right ideal* I of R satisfies

1. $(I, +)$ is a subgroup of $(R, +)$.
2. For every $r \in R$ and every $x \in I$, the right product $r \cdot x \in I$.

Note that left and right modules are equivalence relations defined on a ring.

A left/right ideal can also be seen as a left/right R-submodule of R viewed as an R-module.

Definition 3.33. A *two-sided ideal*, or more simply an *ideal*, is a left ideal that is also a right idea.

Proposition 3.26. *Every right or left ideal of a commutative ring is a two sided ideal.*

Proof. Trivial. ■

Example 3.12. *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.*

Example 3.13. *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.*

Example 3.14. *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.

Proposition 3.27. *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 $a \in \mathbb{F}$. ■

Definition 3.34. A left ideal generated by a single element x is called the *principal left ideal generated by x* and is denoted Rx. Principal right ideals are denoted xR, and principal (two-sided) ideals are denoted RxR.

Definition 3.35. A *principal ideal domain*, also called a *PID*, is an integral domain in which every ideal is principal (i.e. can be generated by a single element).

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
$$ax + by$$

Proposition 3.28. *Every Euclidean domain is also a principal ideal domain.*

Example 3.15. *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* $\mathbb{F}$.

4. *Rings of formal power series* $\mathbb{F}[[x]]$.

5. *The ring of Gaussian integers* $\mathbb{Z}[i]$.

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. For the integers $\mathbb{Z}$, every ideal is of the form $n\mathbb{Z}$, which is principal since it is generated by the integer n. 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.

3.5 The Algebra of Quaternions

Definition 3.36. The *quaternions* form an algebra of 4-dimensional vectors over $\mathbb{R}$, with elements of the form

$$(a, b, c, d) \equiv a + bi + cj + dk$$

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.

$\times$	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

Note the identity

$$i^2 = j^2 = k^2 = -1$$

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

Proposition 3.29. $\mathbb{H}$ *and* $\mathbb{C}$ *are the only finite-dimensional divisions rings containing* $\mathbb{R}$ *as a proper subring.*

Definition 3.37. 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

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

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.

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

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

Proposition 3.30 (Scalar and Vector Components)**.** *Let the quaternion be divided up into a scalar and vector part with the bjective mapping* $a + bi + cj + dk \mapsto \big(a, (b, c, d)\big)$.

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

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

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

It has properties

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

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

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

Definition 3.39. The *norm* of q is defined

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

with properties

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

The norm allows us to define a metric

$$d(p, q) \equiv ||p - q||$$

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

Definition 3.40. The *unit quaternion* is defined to be

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

Corollary 3.30.1. *Every quaternion has a polar decomposition*

$$q = U_q \cdot ||q||$$

With this, we can redefine the inverse as

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

3.5.1 Matrix Representations of Quaternions

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

Proposition 3.31. *The following representation is an injective homomorphism* $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(2, \mathbb{C})$.

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

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.*

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

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

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

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,*

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

Proposition 3.32. *The following representation of $\mathbb{H}$ is an injective homomorphism $\rho : \mathbb{H} \longrightarrow \mathrm{GL}(4, \mathbb{R})$.*

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

or also as

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

It has properties

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

$$||q||^4 = \det\big(\rho(q)\big)$$

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.

3.5.2 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

$$a^2 - b^2 - c^2 - d^2 = -1, 2ab = 2ac = 2ad = 0$$

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

$$b^2 + c^2 + d^2 = 1$$

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

4 Affine and Projective Spaces

4.1 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,

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

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 4.1. 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

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

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 4.2. 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 4.3. 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 4.4. The *barycentric linear combination* of points $p_1, ..., p_k \in S$ is a linear combination of the form

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

This linear combination is equal to the point p such that

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

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

Definition 4.5. 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 4.6. Let $p_0, p_1, ..., p_n$ be points of an n-dimensional affine space S such that the vectors $\overline{p_0p_1}, ..., \overline{p_0p_n}$ are linearly independent (that is, forms a basis). Then, every point $p \in S$ can be uniquely presented as

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

This equality can be rewritten

$$\overline{p_0p} = \sum_{i=1}^{n} x_i \overline{p_0p_i}$$

implying that we can take the coordinates of the vector $\overline{p_0p}$ in the basis $\{\overline{p_0p_1}, ..., \overline{p_0p_n}\}$ as $x_1, ..., x_n$. Then, x_0 is determined as

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

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

$$p = p_0 + U$$

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

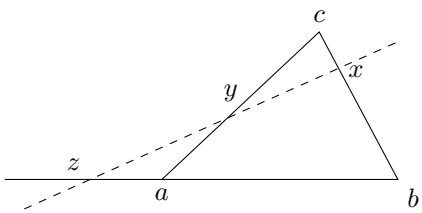
Theorem 4.2. *Given any $k+1$ points of an affine space, there is a plane of dimension $\leq k$ passing through these points. If these points are not contained in a plane of dimension $< k$, then there exists a unique k-dimensional plane passing through them.*

Definition 4.8. 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 $\overline{p_0p_1}, ..., \overline{p_0p_k}$ are linearly independent.

Theorem 4.3. *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 $\overline{p_0p_1}, ..., \overline{p_0p_k}$, which must be nonzero if they are indeed affinely independent.

Corollary 4.3.1 (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

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

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

Corollary 4.3.2 (Ceva's Theorem)**.** *In the triangle above, the lines ax, by, cz intersect at one point if and only if*

$$\lambda\mu\nu = 1$$

Proof. The proof can be done using barycentric coordinates. ■

Theorem 4.4. *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 4.5. *Given an inhomogeneous system of linear equations of form*

$$Ax = b$$

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

$$Ax = 0$$

Let us observe the relative position of two planes.

Theorem 4.6. *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

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

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

$$f(p + x) = f(p) + \alpha(x), \;\; p \in S, x \in V$$

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

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

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

$$f(x) = b + \sum_i a_i x_i$$

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

Proposition 4.7. *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 4.8. *Barycentric coordinates are affine-linear functions.*

Proposition 4.9. *Let f be an affine-linear function. Then*

$$f\left(\sum_i \lambda_i p_i\right) = \sum_i \lambda_i f(p_i)$$

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

Definition 4.10. 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

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

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

4.2 Convex Sets

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

Definition 4.11. The *(closed) interval* connecting points $p, q \in S$ is the set

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

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

Definition 4.12. 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 4.13. 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 4.10. *For any points $p_0, ..., p_k$ in a convex set $M \subset S$, the set M also contains every convex linear combination*

$$p = \sum_i \lambda_i p_i$$

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

Definition 4.14. 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 4.15. 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 4.1. *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 4.11. *A convex set M has interior points if and only if* $\operatorname{aff} M = S$.

Definition 4.16. 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 4.17. For any nonconstant affine-linear function f on the set S, let

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

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

Definition 4.18. 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 4.12. *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 4.13 (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 4.14. *Every closed convex set M is an intersection of (perhaps infinitely many) half-spaces.*

Definition 4.19. 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 4.2. *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 4.3. *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 4.20. A point p of a convex set M is *extreme* if it is not an interior point of any interval in M.

Theorem 4.15. *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 4.16 (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 4.21. 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 4.17. *Every face Γ of the polyhedron M is of the form*

$$\Gamma = M \cap \left(\bigcap_{j \in J} H_{f_j} \right)$$

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

Proposition 4.18. *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 4.19. *The maximum of an affine-linear function on a bounded convex polyhedron M is attained at a vertex.*

4.3 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 4.22. An *affine map* from the space S to the space S' is a map $f : S \longrightarrow S'$ such that

$$f(p + x) = f(p) + \varphi(x), \quad p \in S, x \in V$$

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

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

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

Proposition 4.20. *Let $f : S \longrightarrow S'$ and $g : S' \longrightarrow S''$ be two affine maps. Then the map*

$$g \circ f : S \longrightarrow S''$$

is also affine. Also

$$d(g \circ f) = dg \cdot df$$

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

Definition 4.23. 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 4.21.1. *Finite-dimensional affine spaces over the same field are isomorphic if and only if they have the same dimension.*

Definition 4.24. 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

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

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

$$t_a : p \mapsto p + a, \quad a \in V$$

Proposition 4.22. *For any $f \in \mathrm{GA}(S)$ and $a \in V$,*

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

Definition 4.25. A *homothety* with the center o and coefficient λ is an affine transformation defined as

$$f(o+x) \equiv o + \lambda x$$

In its vectorized form, it is expressed

$$f(x) = \lambda x + b, \quad b \in V$$

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 4.23. *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 $\{\overline{p_0p_1}, ..., \overline{p_0p_n}\}$ to the basis $\{\overline{q_0q_1}, ..., \overline{q_0q_n}\}$. If we vectorize S by taking p_0 as the origin, the affine transformation in question has the form

$$f(x) = \varphi(x) + \overline{p_0q_0}$$

■

Corollary 4.23.1. *In real affine geometry all parallelopipeds are equal.*

Definition 4.26. 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 $\operatorname{Isom} S$. A motion is called *proper (orientation preserving)* if its differential belongs to $\mathrm{SO}(V)$ and improper otherwise.

Lemma 4.24. *The group $\operatorname{Isom} S$ is generated by reflections through hyperplanes.*

Definition 4.27. 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 4.28. 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

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

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 4.29. 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.

4.4 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 4.30. An *affine-quadratic function* on an affine space S is a function $Q : S \longrightarrow \mathbb{F}$ such that its vectorized form is

$$Q(x) = q(x) + l(x) + c$$

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

4.5 Projective Spaces

Definition 4.31. 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 4.32. $\mathbb{RP}^1$ is called the real projective line, which is topologically equivalent to a circle.

Example 4.4. *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 4.5. $\mathbb{RP}^3$ *is diffeomorphic to* $SO(3)$.

Example 4.6. *The space* $\mathbb{RP}^n$ *is formed by taking the quotient of* $\mathbb{R}^{n+1} \setminus \{0\}$ *under the equivalence relation*

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

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

5 Tensor Algebras

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

$$\times : V \times V \longrightarrow V$$

Definition 5.1. 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

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

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

$$\otimes : T(V) \times T(V) \longrightarrow T(V)$$

and the identity element is

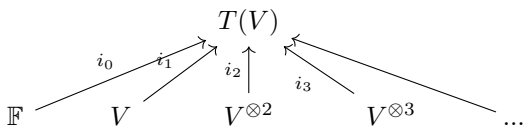
$$I = (1, 0, 0, ...)$$

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

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

shown in the diagram



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

$$A + B \equiv i_n(A) + i_m(B) \in T(V)$$

along with multiplication of tensors as

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

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

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

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

Definition 5.2. The *exterior algebra* $\Lambda(V)$ of a vector space V over field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

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

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

$$\pi : T(V) \longrightarrow \Lambda(V)$$

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

$$a \wedge b = \pi(\alpha \otimes \beta)$$

We can define this quotient space with the equivalence class

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

Definition 5.3. The *symmetric algebra* $\mathrm{Sym}(V)$ of a vector space V over a field $\mathbb{F}$ is the quotient algebra of the tensor algebra $T(V)$

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

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

$$v \otimes w - w \otimes v$$

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

6 Representation Theory

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

Definition 6.1. 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 6.2. A *representation* of an (algebraic) group $\mathcal{G}$ is a homomorphism

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

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.

$$\rho(g_1 \cdot g_2) = \rho(g_1) \cdot \rho(g_2)$$

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

$$gv = \rho(g)v, \ v \in V$$

Since ρ is a group morphism, we have

$$g_2(g_1 v) = (g_2 g_1)v \iff \rho(g_2)\big(\rho(g_1)(v)\big) = \big(\rho(g_2)\rho(g_1)\big)(v)$$

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

$$g(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 g v_1 + \lambda_2 g v_2$$

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 6.1. *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.*

$$\mathrm{GL}(2,\mathbb{C}) \overset{id}{\longmapsto} Mat(2,\mathbb{C})$$

This is a trivial representation.

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

Example 6.2. *We take* $Sym^2\mathbb{C}^2$*, the second symmetric power of* $\mathbb{C}^2$*. Note that given a basis* $x_1, x_2 \in \mathbb{C}^2$*, the set*

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

forms a basis of $Sym^2\mathbb{C}^2 \implies \dim Sym^2\mathbb{C}^2 = 3$*. 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 $\mathrm{GL}(2,\mathbb{C})$

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

We evaluate the corresponding representation in $\mathrm{GL}(Sym^2\mathbb{C}^2)$*. Using the identities above, we have*

$$\begin{aligned}
\rho(g)(x_1 \odot x_1) &= g(x_1) \odot g(x_1) \\
&= (ax_1 + cx_2) \odot (ax_1 + cx_2) \\
&= a^2 x_1 \odot x_1 + 2acx_1 \odot x_2 + c^2 x_2 \odot x_2 \\
\rho(g)(x_1 \odot x_2) &= g(x_1) \odot g(x_2) \\
&= (ax_1 + cx_2) \odot (bx_1 + dx_2) \\
&= abx_1 \odot x_1 + (ad + bc)x_1 \odot x_2 + cdx_2 \odot x_2 \\
\rho(g)(x_2 \odot x_2) &= g(x_2) \odot g(x_2) \\
&= (bx_1 + dx_2) \odot (bx_1 + dx_2) \\
&= b^2 x_1 \odot x_1 + 2bdx_1 \odot x_2 + d^2 x_2 \odot x_2
\end{aligned}$$

And this completely determines the matrix. So,

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

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 6.3. A *morphism* between 2 representations

$$\rho_1 : \mathcal{G} \longrightarrow \mathrm{GL}(V_1)$$
$$\rho_2 : \mathcal{G} \longrightarrow \mathrm{GL}(V_2)$$

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 $g \in \mathcal{G}$

$$f \circ g = g \circ f$$

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.

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

Definition 6.4. 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 $w \in W$,

$$\rho(g)(w) \in W$$

Example 6.3. *V and $\{0\}$ are always subrepresentations of V.*

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

Definition 6.5. A representation W is *irreducible representation* if $\{0\}$ and W are the only subrepresentations of W.

Lemma 6.1 (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,

$$f_1 = \lambda f_2$$

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 $\mathrm{Im}\, f$ is a subrepresentation of V_2 $\implies$ $\mathrm{Im}\, f = 0$ or $\mathrm{Im}\, f = V_2$. But $\mathrm{Im}\, f \neq 0$ since f is injective, so $\mathrm{Im}\, f = V_2$ $\implies$ 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

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

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

$$f_1 = id$$

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

$$f_2(v) = \lambda v$$

Now, we define the map

$$f : V_1 \longrightarrow V_2,\ f \equiv f_2 - \lambda f_1$$

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 6.2 (Mache's Theorem)**.** *Let V be finite dimensional, with $\mathcal{G}$ a finite group. Then, V can be decomposed as*

$$V = \bigoplus_i V_i$$

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

$$\pi : V \longrightarrow W$$

Then, we define the new map

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

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

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

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

7 Lie Groups and Lie Algebras

Definition 7.1. 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

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

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

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

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

Definition 7.2. A *Lie Algebra* is a vector space $\mathfrak{g}$ with an operation called the *Lie Bracket*

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$$

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

$$[A, B] \equiv AB - BA$$

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 7.2. *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

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

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

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

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

$$exp : \mathfrak{g} \longrightarrow \mathcal{G}$$

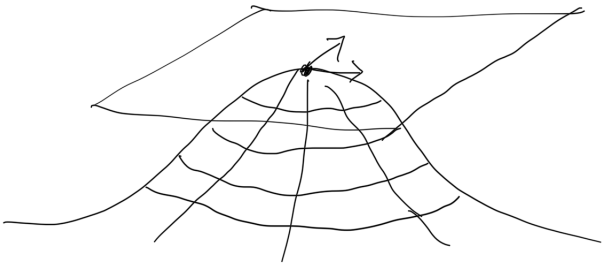
where $\mathfrak{g}$ is the Lie algebra and $\mathcal{G}$ is the Lie group (which we will treat just as a matrix group). To find the inverse of the exponential map, we can take the derivative of e^{tA} at $t = 0$. That is,

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

So, the mapping

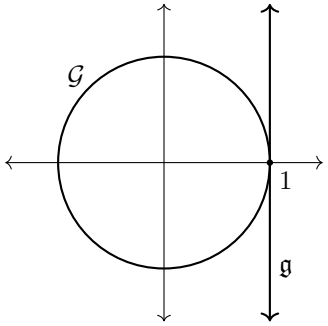
$$\frac{d}{dt}\Big|_{t=0} : \mathcal{G} \longrightarrow \mathfrak{g}$$

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:



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

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



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 7.3. *The exponential map*

$$exp : \mathbb{R} \longrightarrow \mathbb{R}^+,\ x \mapsto e^x$$

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 7.1. *If* A *and* B *are commuting square matrices, then*

$$e^{A+B} = e^A\, e^B$$

In general, the solution C *to the equation*

$$e^A\, e^B = e^C$$

is given by the Baker-Campbell-Hausdorff formula, defined

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}[A, [A, B]] - \frac{1}{12}[B, [A, B]] + ...$$

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

$$\mathfrak{g} \equiv (V, [\cdot, \cdot])$$

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

$$\cdot : \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G},\ e^A \cdot e^B = e^{A*B}$$

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.

7.1 Lie Algebras of Classical Lie Groups

Definition 7.3. 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 $= 1$.

7.1.1 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 7.2.

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

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

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

since trace is invariant under a change of basis. ∎

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

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

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

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

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,

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

is clearly not traceless. However, the bracket operation keeps the matrices within this traceless condition (and thus, within this algebra), so you can't just stupidly multiply matrices together in a Lie algebra. Remember that regular matrix multiplication does not have anything to do with the Lie bracket and does not apply to this group. This algebra also simplifies the multiplicative inverse of a group to a simple additive inverse, making calculations easier.

Similarly, the Lie algebra of $\text{SL}(2, \mathbb{C})$ also has the same basis

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

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

7.1.2 Lie Algebra of SU(2)

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

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

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 7.4. The *Pauli matrices* are the three matrices

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

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

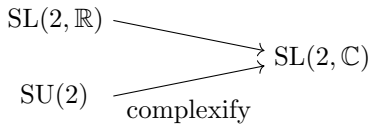
$$[A_x, A_y] = A_z$$
$$[A_y, A_z] = A_x$$
$$[A_z, A_x] = A_y$$

However, note that the algebra $\mathfrak{su}(2)$ consists of all *real* linear combinations of A_x, A_y, A_z. That is, $\mathfrak{su}(2)$ is a 3 dimensional *real* vector space, even though it has basis elements containing complex numbers.

However, we can always complexify this space by simply replacing real scalar multiplication in $\mathfrak{su}(2)$ with complex scalar multiplication. By complexifying $\mathfrak{su}(2)$, the Lie group SU(2) formed by taking the exponential map on this complexified space is actually identical to $\text{SL}(2, \mathbb{C})$. 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 $\text{Mat}(2, \mathbb{C})$, 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.

$$[H, X] = 2X \iff [A_x, A_y] = A_z$$
$$[H, Y] = -2Y \iff [A_y, A_z] = A_x$$
$$[X, Y] = H \iff [A_z, A_x] = A_y$$

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



We can interpret the "real forms" of $\text{SL}(2, \mathbb{C})$ as "slices" of some complex group. However, this does not mean that the real version of these groups are equal. That is,

$$\text{SL}(2, \mathbb{R}) \neq \text{SU}(2)$$

7.1.3 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

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

as its only basis, since

$$exp\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \theta \right) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

meaning that the dimension of SO(2) is 1. By adding a component, we can get a rotation in $\mathbb{R}^3$.

$$R_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \implies e^{R_x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$
$$R_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \implies e^{R_y} = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix}$$
$$R_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies e^{R_z} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

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

The bracket structure (again, defined as the commutator) of this Lie algebra is

$$[R_x, R_y] = R_z$$
$$[R_y, R_z] = R_x$$
$$[R_z, R_x] = R_y$$

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,

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

while

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

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

$$\text{SO}(3) = \frac{\text{SU}(2)}{\pm I}$$

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

$$M = x\sigma_x + y\sigma_y + z\sigma_z$$

M is clearly traceless and $M^\dagger = M$. Now, let $S \in \text{SU}(2)$ and let $M' = S^{-1}MS$. Then, $\text{Tr}\, M' = \text{Tr}\, S^{-1}MS = \text{Tr}\, M = 0$ 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

$$x'\sigma_x + y'\sigma_y + z'\sigma_z$$

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

$$M^2 = (-x^2 - y^2 - z^2)I$$

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

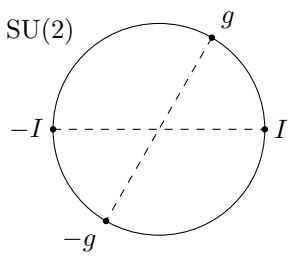
$$S = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

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

$$\rho : \text{SU}(2) \longrightarrow \text{SO}(3)$$

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

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



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

Lemma 7.3 (Rodrigues' Formula)**.** *The exponential map $exp : \mathfrak{so}(3) \longrightarrow SO(3)$ is defined by*

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

where

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

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

Theorem 7.4. *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 7.5. *The Lie algebras $\mathfrak{gl}_{\mathbb{R}}, \mathfrak{sl}_n\mathbb{R}, \mathfrak{o}(n), \mathfrak{so}(n)$ are well-defined, but only*

$$exp : \mathfrak{so}(n) \longrightarrow SO(n)$$

is surjective.

Theorem 7.6. *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 $\mathrm{GL}(n,\mathbb{C}), \mathrm{SL}(n,\mathbb{C}), \mathfrak{gl}_n\mathbb{C}, \mathfrak{sl}_n\mathbb{C}$ 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. $exp : \mathfrak{gl}_n\mathbb{R} \longrightarrow \mathrm{GL}(n,\mathbb{R})$ is not surjective, but $exp : \mathfrak{gl}_n\mathbb{C} \longrightarrow \mathrm{GL}(n,\mathbb{C})$ is surjective due to the spectral theorem and surjectivity of $exp : \mathbb{C} \longrightarrow \mathbb{C}^*$.
3. The exponential maps $exp : \mathfrak{u}(n) \longrightarrow \mathrm{U}(n)$ and $exp : \mathfrak{su}(n) \longrightarrow \mathrm{SU}(n)$ are surjective.
4. Still, $exp : \mathfrak{sl}_2\mathbb{C} \longrightarrow \mathrm{SL}(2,\mathbb{C})$ is not surjective. This will be proved now.

Theorem 7.7. *$exp : \mathfrak{sl}_2\mathbb{C} \longrightarrow \mathrm{SL}(2,\mathbb{C})$ 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

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

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

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

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

Theorem 7.8. *The exponential maps*

$$exp : \mathfrak{u}(n) \longrightarrow U(n)$$
$$exp : \mathfrak{su}(n) \longrightarrow SU(n)$$

are surjective.

7.1.4 Lie Algebra of SE(n)

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

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

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

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

we define the representation

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

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

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

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

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

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.

$$\begin{array}{ccc} \mathfrak{se}(n) & \xrightarrow{exp} & SE(n) \\ \downarrow{\scriptstyle \varrho} & & \downarrow{\scriptstyle \rho} \\ \mathfrak{gl}_{n+1}\mathbb{R} & \xrightarrow{exp} & \mathrm{GL}(n+1, \mathbb{R}) \end{array}$$

Lemma 7.9. *Given any $(n+1) \times (n+1)$ matrix of form*

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

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

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

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

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

Theorem 7.10. *The exponential map*

$$exp : \mathfrak{se}(n) \longrightarrow SE(n)$$

is well-defined and surjective.

7.2 Representations of Lie Groups and Lie Algebras

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

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

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

$$\varrho : \mathfrak{g} \longrightarrow \mathfrak{gl}(V)$$

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

$$\varrho([X, Y]) = [\varrho(X), \varrho(Y)]$$

We can visualize this induced homomorphism with the following commutative diagram, which states that $\rho \circ exp = exp \circ \varrho$.

$$\begin{array}{ccc} \mathcal{G} & \xrightarrow{\rho} & \mathrm{GL}(V) \\ {\scriptstyle exp}\uparrow & & \uparrow{\scriptstyle exp} \\ \mathfrak{g} & \xrightarrow{\varrho} & \mathfrak{gl}(V) \end{array}$$

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

Definition 7.5. 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

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

and

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

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

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

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

Definition 7.6. 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

$$\varrho : \mathfrak{sl}_2\mathbb{C} \longrightarrow \mathfrak{gl}(V)$$

We will write, as shorthand notation, that

$$H = \varrho(H), X = \varrho(X), Y = \varrho(Y)$$

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

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

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

$$HYe_i - YHe_i = -2Ye_i \implies H(Ye_i) = (\lambda_i - 2)(Ye_i)$$

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

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

(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 7.11.

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

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

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

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

$$m, m-2, m-4, ..., m-2(n-1) = -m$$

and we are done. We now classify the 1, 2, and 3 dimensional irreducible representations of $\mathfrak{sl}_2\mathbb{C}$.

When $n = 1$ (i.e. dimension is 1), $m = n - 1 = 0$, meaning that the greatest (and only) eigenvalue is 0. That is,

$$Hv_0 = 0,\ Xv_0 = 0,\ Yv_0 = 0$$

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.

$$\varrho(H) = (0),\ \varrho(X) = (0),\ \varrho(Y) = (0)$$

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

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

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

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

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

7.2.1 Tensor Products of Group Representations

Definition 7.7. 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

$$\rho_V : \mathcal{G} \longrightarrow \text{GL}(V)$$
$$\rho_W : \mathcal{G} \longrightarrow \text{GL}(W)$$

The homomorphism

$$\rho_V \otimes \rho_W : \mathcal{G} \longrightarrow \text{GL}(V \otimes W)$$

is also a representation of $\mathcal{G}$, which is defined

$$(\rho_V \otimes \rho_W)(g)(v \otimes w) \equiv \rho_V(g)(v) \otimes \rho_W(g)(w)$$

or represented in shorthand notation,

$$g(v \otimes w) \equiv (gv) \otimes (gw)$$

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

$$exp(H)(v \otimes w) \equiv \big(exp(H)(v)\big) \otimes \big(exp(H)(w)\big)$$

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

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

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

Example 7.5. *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*

$$(-1) + (-1) = -2, \;\; (-1) + 1 = 0$$
$$1 + (-1) = 0, \;\; 1 + 1 = 2$$

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*

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

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

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

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,*

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

7.3 Topological Decompositions of Lie Groups

Definition 7.8. 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 7.12. *The exponential map*

$$exp : S(n) \longrightarrow SPD(n)$$

is a homeomorphism. One may be tempted to call $S(n)$ the Lie algebra of $SPD(n)$, but this is not the case. $S(n)$ is not even a Lie algebra since the commutator is not algebraically closed. Furthermore, $SPD(n)$ 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 7.13 (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 $h_1, h_2 \in End(\mathbb{E}^n)$ and $g \in O(n)$ such that*

$$f = g \circ h_1 = h_2 \circ g$$

That is, such that f can be decomposed into the following compositions of functions that commute.

$$\begin{array}{ccc} \mathbb{E}^n & \xrightarrow{h_2} & \mathbb{E}^n \\ {\scriptstyle g}\uparrow & \nearrow_{f} & \uparrow{\scriptstyle g} \\ \mathbb{E}^n & \xrightarrow{h_1} & \mathbb{E}^n \end{array}$$

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 7.13.1. *The two topological groups are homeomorphic.*

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

Corollary 7.13.2. *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*

$$A = Re^S$$

$\implies$ *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,*

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

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

Corollary 7.13.3. *Given a real matrix A, if* $\det A > 0$*, then we can decompose A as*

$$A = Re^S$$

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

Corollary 7.13.4. *There exists a bijection between*

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

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

Definition 7.9. Let us define

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

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

Lemma 7.14. *The exponential mapping*

$$exp : H(n) \longrightarrow HPD(n)$$

is a homeomorphism.

However again, $\mathrm{HPD}(n)$ is not a Lie group (multiplication is not algebraically closed) nor is $\mathrm{H}(n)$ a Lie algebra (commutator is not algebraically closed). By the polar form theorem of complex $n \times n$ matrices, we have a (not necessarily unique) bijection between

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

which implies that

$$\mathrm{GL}(n, \mathbb{C}) \cong U(n) \times HPD(n)$$

Corollary 7.14.1. *For every complex invertible matrix A, there exists a unique decomposition*

$$A = Ue^S$$

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

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

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

Corollary 7.14.2. *There exists a bijection between*

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

Proof. Similarly, when $A = Ue^S$, we know that $|\det U| = 1$ and $\operatorname{Tr} S$ is real (since by the Spectral theorem, every self adjoint matrix has a real spectral decomposition). Since S is Hermitian, this implies that $\det e^S > 0$. If $A \in \mathrm{SL}(n, \mathbb{C})$, then $\det A = 1 \implies \det e^S = 1 \implies S \in H(n) \cap \mathfrak{sl}_n \mathbb{C}$. ■

7.4 Linear Lie Groups

We will assume that the reader has the necessary background knowledge in manifolds, chart mappings, diffeomorphisms, tangent spaces, and transition mappings.

Recall that the algebra of real $n \times n$ matrices $\mathrm{Mat}(n, \mathbb{R})$ is bijective to $\mathbb{R}^{n^2}$, which is a topological space. Therefore, this bijection

$$i : (\mathbb{R}^{n^2}, \tau_E) \longrightarrow \mathrm{Mat}(n, \mathbb{R})$$

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

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

With this, consider the subset

$$\mathrm{GL}(n, \mathbb{R}) \subset \mathrm{Mat}(n, \mathbb{R})$$

where

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

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

Definition 7.10. The *general linear group*, denoted $\mathrm{GL}(n, \mathbb{R})$ is the set of $n \times n$ matrices with nonzero determinant. The more technical definition is that $\mathrm{GL}(n, \mathbb{R})$ is really just the automorphism group of $\mathbb{R}^n$,

$$\mathrm{GL}(n, \mathbb{R}) \equiv \mathrm{Aut}(\mathbb{R}^n)$$

but it is customary to assume a basis on $\mathbb{R}^n$ in order to realize $\mathrm{GL}(n, \mathbb{R})$ as a matrix group. Note that the procedure of assuming a basis on $\mathbb{R}^n$ is the same as defining a representation of the abstract group $\mathrm{GL}(n, \mathbb{R})$. Both assigns a real $n \times n$ matrix to each element of $\mathrm{GL}(n, \mathbb{R})$.

In this way, we can view $\mathrm{GL}(n, \mathbb{R})$ as a topological space in $\mathbb{R}^{n^2}$, and it is fine to interpret $\mathrm{GL}(n, \mathbb{R})$ as a matrix group rather than an abstract group.

Since the matrix representation of $\mathrm{GL}(n, \mathbb{R})$ is always well defined, the abstract subgroups of $\mathrm{GL}(n, \mathbb{R})$, which are $\mathrm{SL}(n, \mathbb{R}), O(n)$, and $SO(n)$, also have well defined matrix representations (that we are all familiar with). Additionally, since there exists a bijection

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

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 7.15. *$SE(n)$ is a linear Lie group.*

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

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

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

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

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

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

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

Note that even though $\mathrm{SE}(n)$ is diffeomorphic (a topological relation) to $\mathrm{SO}(n) \times \mathbb{R}^n$, it is *not* isomorphic (an algebraic relation) sicne group operations are not preserved. Therefore, we write this "equality" as a semidirect product of groups.

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

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 7.11. 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 7.16 (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 7.12. 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

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

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

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

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

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

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

$$[A, B] \equiv AB - BA$$

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

$$[A, B] \equiv AB - BA$$

called the *commutator*.

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

$$exp : \mathfrak{g} \longrightarrow \mathcal{G}$$

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

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

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 7.18. *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 7.19. $\det'(I) = \mathrm{Tr}$. *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 $\operatorname{Tr} T$. So,

$$\nabla_T \det I = \lim_{\varepsilon \to 0} ... + \operatorname{Tr} T = \operatorname{Tr} T$$

■

This means that the instantaneous rate at which det changes at I when traveling in direction T is directly proportional to $\operatorname{Tr} T$. Now, we provide an alternative proof of the theorem.

Proof. Let $R : \mathbb{R} \longrightarrow \mathrm{SL}(n, \mathbb{R})$ such that $R(0) = I$. Then, by definition, $\operatorname{Im} R \subset \mathrm{SL}(n, \mathbb{R}) \implies \det(R(t)) = 1$ 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

$$\overset{\prime}{\det}\big(R'(0)\big) = \overset{\prime}{\det}(I) = 0 \implies \operatorname{Tr} R'(0) = 0$$

■

Theorem 7.20. *The matrix representation of $\mathfrak{so}(n)$ is precisely the set of antisymmetric matrices.*

Proof. Let $R : \mathbb{R} \longrightarrow SO(n)$ be a arbitrary smooth curve in $\mathrm{SL}(n)$ such that $R(0) = I$. Then, for all $t \in (-\epsilon, \epsilon)$,

$$R(t)R(t)^T = I$$

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

$$R'(0)R(0)^T + R(0)R'(0)^T = 0 \implies R'(0) + R'(0)^T = 0$$

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

   $$[A, B] \equiv AB - BA$$

   but for elements that are not matrices this doesn't make sense.
3. The exponential map from the algebra to the group is defined

   $$e^A \equiv \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

   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.

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

Example 7.6. *The Lie algebra $\mathfrak{so}(3)$ is the real vector space of 3×3 skew symmetric matrices of form*

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

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 $\psi : (\mathbb{R}^3, \times) \longrightarrow \mathfrak{so}(3)$ (where $\times$ is the cross product) by the formula

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

where, by definition,

$$\psi(u \times v) = [\psi(u), \psi(v)]$$

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

$$\psi(u)(v) = u \times v$$

Example 7.7. *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*

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

where $\sigma_1, \sigma_2, \sigma_3$ are the Pauli spin matrices. We can also define an isomorphism of Lie algebras $\varphi : (\mathbb{R}^3, \times) \longrightarrow \mathfrak{su}(2)$ by the formula

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

where, by definition of ismorphism, we have

$$\varphi(u \times v) = [\varphi(u), \varphi(v)]$$

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

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

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

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

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

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

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

$$exp(\theta v) =$$

Example 7.8. *The lie algebra $\mathfrak{se}(n)$ is the set of all matrices of form*

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

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

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

7.5 Abstract Lie Groups

Definition 7.14. 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 7.15. A (real) Lie algebra $\mathfrak{g}$ is a real vector space with a map

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \longrightarrow \mathfrak{g}$$

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 7.16. Let M_1 (m_1-dimensional) and M_2 (m_2 dimensional) be manifolds in $\mathbb{R}^N$. For any smooth function $f : M_1 \longrightarrow M_2$ and any $p \in M_1$, the function

$$f'_p : T_p M_1 \longrightarrow T_{f(p)} M_2$$

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 $\gamma : I \longrightarrow M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = v$,

$$f'_p(v) \equiv (f \circ \gamma)'(0)$$

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

Definition 7.17. Given two Lie groups $\mathcal{G}_1$ and $\mathcal{G}_2$, a *homomorphism of Lie groups* is a function

$$f : \mathcal{G}_1 \longrightarrow \mathcal{G}_2$$

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

$$f : \mathfrak{g}_1 \longrightarrow \mathfrak{g}_2$$

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

$$f([A, B]) = [f(A), f(B)]$$

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 7.21. *If $f : \mathcal{G}_1 \longrightarrow \mathcal{G}_2$ is a homomorphism of Lie groups, then*

$$f'_I : \mathfrak{g}_1 \longrightarrow \mathfrak{g}_2$$

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

$$L_a : \mathcal{G} \longrightarrow \mathcal{G},\ L_a(b) \equiv ab$$

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

$$R_a : \mathcal{G} \longrightarrow \mathcal{G},\ R_a(b) \equiv ba$$

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

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

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

the derivative

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

is an ismorphism of Lie algebras, also denoted

$$\text{Ad}_a : \mathfrak{g} \longrightarrow \mathfrak{g}$$

Definition 7.20. This induces another map $a \mapsto \text{Ad}_a$, which is a map of Lie groups

$$Ad : \mathcal{G} \longrightarrow \text{GL}(\mathfrak{g})$$

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

$$\text{Ad}(a)(X) \equiv \text{Ad}_a(X) \equiv aXa^{-1}$$

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

Definition 7.21. Furthermore, the derivative of this map at the identity

$$\text{Ad}'_I : \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$$

is a map between Lie algebras, denoted simply as

$$\text{ad} : \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$$

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

$$\begin{array}{ccc} \mathcal{G} & \xrightarrow{Ad} & \text{GL}(\mathfrak{g}) \\ {\scriptstyle exp}\uparrow & & \uparrow{\scriptstyle exp} \\ \mathfrak{g} & \xrightarrow{ad} & \mathfrak{gl}(\mathfrak{g}) \end{array}$$

We define the map ad to be

$$ad(A)(B) \equiv [A, B]$$

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,

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

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 7.22. 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

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

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