Willoughby Seago

Theoretical Physics

Symmetries of Quantum Mechanics

February 17, 2022

Course Notes

Symmetries of Quantum Mechanics

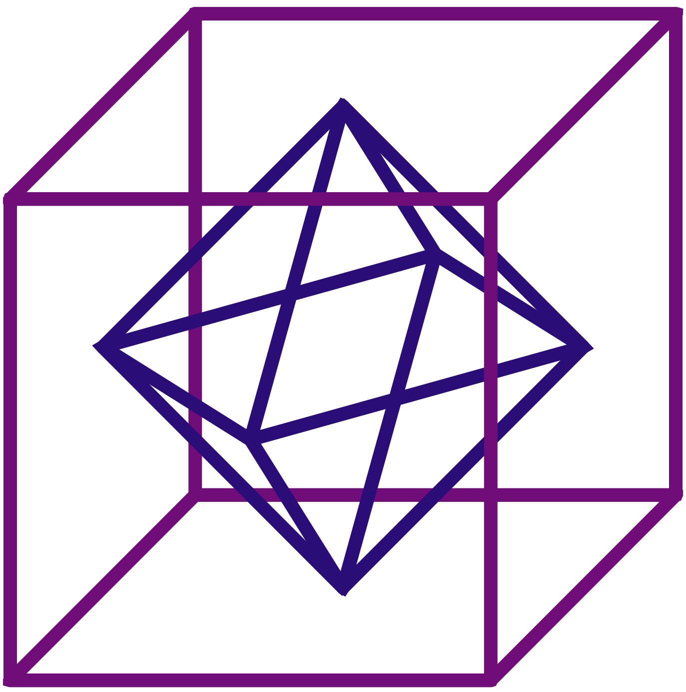
Willoughby Seago

February 17, 2022

Abstract

These are my notes from the course symmetries of quantum mechanics. I took this course as a part of the theoretical physics degree at the University of Edinburgh.

These notes were last updated at 15:04 on March 19, 2022. For notes on other topics see `https://github.com/WilloughbySeago/Uni-Notes`.



Chapters

Page

Contents

Page

List of Figures

Page

Part I

Group Theory

One

Introduction

1.1 Binary Operations

Definition 1.1.1 — Binary Operation A **binary operation** on a set X is a map, $f : X \times X \to X$.

R We say that X is **closed** under the binary operation since combining two elements of X gives another element of X.

Notation 1.1.2 Binary operations are usually written with infix notation, for example, the binary operation $\cdot : X \times X \to X$ maps $(x, x') \mapsto x \cdot x'$ whereas for a normal function, say, $f : X \times X \to X$, we usually use prefix notation: $(x, x') \mapsto f(x, x')$.
The other common notation when there is only one (obvious) choice of binary operation is juxtaposing the two elements, for example $(x, x') \mapsto xx'$. This is exactly what we do with multiplication most of the time rather than writing $x \cdot x$, or, $x \times x$. We will use this notation most of the time, particularly when the binary operation is denoted $\cdot$, and we will not comment on it further.

The notion of a binary operation is very general. We typically restrict ourselves to various classes of binary operations which are easier to work with due to possessing various properties.

1.1.1 Associativity

Definition 1.1.3 — Associativity We say that the binary operation $\cdot : X \times X \to X$ is **associative** if

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z. \tag{1.1.4}$$

From this it follows that for an associative binary operation and any number of elements in a product the answer will be the same no matter how we write the brackets, so we usually don't write any brackets at all. For example, with four elements two possible ways to write the product of four elements are

$$(x_1x_2)(x_3x_4) = x_1(x_2(x_3x_4)) = x_1x_2x_3x_4. \tag{1.1.5}$$

Writing $x_3x_4 = x$ it follows that $(x_1x_2)(x_3x_4) = (x_1x_2)x = x_1(x_2x) = x_1(x_2(x_3x_4))$ where the second equality is where we apply the associativity axiom.

■ **Example 1.1.6 — Function Composition** Denote by $\mathrm{Hom}(A, B)$ the set of functions from A to B. We define **function composition** to be the binary operation $\circ\colon \mathrm{Hom}(B, C) \times \mathrm{Hom}(A, B) \to \mathrm{Hom}(A, C)$ for sets A, B, and C, such that for $f \in \mathrm{Hom}(B, C)$ and $g \in \mathrm{Hom}(A, B)$ we have

$$(f \circ g)(a) = f(g(a)) \tag{1.1.7}$$

for all $a \in A$. An alternative way of saying this is that the following diagram commutes, meaning that the result is independent of the path taken:

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[dr, "f\circ g"'] & B \arrow[d, "f"] \\ & C. \end{tikzcd} \tag{1.1.8}$$

Function composition is associative. That is if $f \in \mathrm{Hom}(A, B)$, $g \in \mathrm{Hom}(B, C)$, and $h \in \mathrm{Hom}(C, D)$ then

$$(f\circ(g\circ h))(d) = f((g\circ h)(d)) = f(g(h(d))) = (f\circ g)(h(d)) = ((f\circ g)\circ h)(d),$$

or in other words

$$f \circ (g \circ h) = (f \circ g) \circ h \tag{1.1.9}$$

and so $\circ$ is associative.
The commutative diagram expressing this fact is

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] \arrow[rrr, bend right, "h\circ g\circ f"'] & B \arrow[r, "g"] \arrow[rr, bend left, "h\circ g"] & C \arrow[r, "h"] & D. \end{tikzcd} \tag{1.1.10}$$

One important corollary is that matrix multiplication is just composition of linear maps and so matrix multiplication is associative.

Many of the binary operations that we are familiar with are associative, such as addition, and multiplication, but not all, for example, subtraction isn't associative, consider $5 - (2 - 3) = 6$ and $(5 - 2) - 3 = 0$.

■ **Example 1.1.11 — Nonassociativity** An example of a binary operation that *isn't* associative is the vector cross product, $\times\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ (note that the first $\times$ is the cross product of vectors and the second one is the Cartesian product of sets). This can be shown by an example. Take $\boldsymbol{a} = (1, 1, 0)^\top$,

$\boldsymbol{b} = (0, 1, 0)^\top$, and $\boldsymbol{c} = (0, 0, 1)^\top$. Then

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \implies (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \tag{1.1.12}$$

whereas

$$\boldsymbol{b} \times \boldsymbol{c} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \implies \boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{1.1.13}$$

1.1.2 Identity

Definition 1.1.14 — Identity Given a set, X, and a binary operation on X, $\cdot\colon X \times X \to X$, we say that $e \in X$ is the **identity** if

$$x \cdot e = e \cdot x = x \tag{1.1.15}$$

for all $x \in X$.

Notation 1.1.16 — Identities There are many notations for identities since it is an idea that emerged in many areas before being unified by group theory and other algebraic concepts. The notation we will choose typically depends on both what the elements of X are and the nature of $\cdot$. For example,

- if the elements of X are matrices then the identity may be denoted I, or $\mathbb{1}$,
- if the elements of X are functions then the identity may be denoted id, or ι,
- if $\cdot$ can be thought of as multiplication then the identity is often denoted 1, and
- if $\cdot$ can be thought of as addition (in which case we are more likely to denote the operation $+$) then the identity is often denoted 0.

■ **Example 1.1.17 — Identities**

- The identity for multiplication in $\mathbb{R}$ is 1.
- The identity for addition in $\mathbb{Q}$ is 0.
- The identity for matrix multiplication is I, which has δ_{ij} as elements.
- The identity function is $\mathrm{id}_X\colon X \times X \to X$ defined by $\mathrm{id}_X(x) = x$ for all $x \in X$.

Not all binary operations have an identity, for example, there is no identity for the cross product. It is also important that the identity must be an element of X. For example if we set $X = \mathbb{Z}_{>0}$[1] and take our operation to be addition then there is no identity since $0 \notin \mathbb{Z}_{>0}$.

[1] $\mathbb{Z}_{>0} \coloneqq \{1, \dots\} = \mathbb{N} \setminus \{0\}$

1.1.3 Inverse

Definition 1.1.18 — Inverse Given a set, X, and a binary operation on X, $\cdot \colon X \times X \to X$, such that $e \in X$ acts as an identity element then we say that $x \in X$ has an **inverse** in X if there exists some $x^{-1} \in X$ such that

$$x \cdot x^{-1} = x^{-1} \cdot x = e. \tag{1.1.19}$$

Notation 1.1.20 — Inverses If we think of the binary operation, $\cdot$, as multiplicative then we write the inverse of x as x^{-1}, taking inspiration from division being the inverse of multiplication.
If we think of the binary operation, $+$, as additive then we write the inverse of x as $-x$, and we write $y - x$ as shorthand for $y + (-x)$, taking inspiration from subtraction being the inverse of multiplication.

As with the identity it is important that the inverse is an element of X. For example, taking $X = \mathbb{N}$[2] and our operation to be addition we have an identity, $0 \in \mathbb{N}$, but no inverses (apart from 0, which is its own inverse), since, for example, -3 is the inverse of 3, but $-3 \notin \mathbb{N}$.

[2] $\mathbb{N} \coloneqq \{0, 1, \dots\} = \mathbb{Z}_{>0} \cup \{0\}$.

1.2 Groups

Definition 1.2.1 — Group Formally a **group** is an ordered pair, $(G, \cdot)$, where G is a set and $\cdot$ is a binary operation on G satisfying the following **group axioms**:

1. **Associativity**: For all $g_1, g_2, g_3 \in G$
$$g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3. \tag{1.2.2}$$
2. **Identity**: There exists some $e \in G$ such that $e \cdot g = g \cdot e = g$.
3. **Inverse**: For all $g \in G$ there exists some g^{-1} such that
$$g \cdot g^{-1} = g^{-1} \cdot g = e \tag{1.2.3}$$
where e is the identity of the group.

R In practice, we don't really think of groups as an ordered pair, $(G, \cdot)$, but as a set and an operation on the set and rather than saying "the group $(G, \cdot)$" most of the time we will say "the group G under $\cdot$", or simply "the group G" when it is clear what the group operation.

Some sources include a fourth axiom:

4. **Closure**: The product of two elements of a group is another element of the group.

This is implicit however in the definition of a binary operation as a function $X \times X \to X$, and so we leave it out. It will be important to consider when we think about subgroups by restricting the binary operation to a subset.

Notation 1.2.4 — Multiple Groups Say G and H are two groups of interest. Then we will use G and H as subscripts to differentiate between the two groups. For example the product of two elements in G may be written as $g \cdot_G g'$, as opposed to $g \cdot_H g'$, which is the product of H applied to elements of G, as we may sometimes have reason to do if, say, $H \subseteq G$. The identity in H may be denoted e_H, and so $e_H \cdot_H h = h$, but we may not have $h \cdot_G e_H = h$, since a different operation means we can have a different identity.

Lemma 1.2.5 The identity of a group is unique.

Proof. Suppose that G is a group and $e, e' \in G$ both act as identities. That is $e'g = ge = g$ for all $g \in G$. Then $e = e'e = e'$ where the first equality holds by the identity property of e' and the second by the identity property of e. This means $e = e'$ and so the identity is unique. □

Lemma 1.2.6 The inverse of a group element is unique.

Proof. Suppose that G is a group and $g \in G$ is such that $h, h' \in G$ act as inverses to G. That is $hg = gh' = e$ where $e \in G$ is the identity of G. Right multiplying $hg = gh' = e$ by h' we have $hgh' = gh'h'$. Using the inverse property of h' on both sides we have $he = eh'$ which implies $h = h'$ and so the inverse is unique. □

Lemma 1.2.7 Let G be a group and $g, h \in G$. Then $(gh)^{-1} = h^{-1}g^{-1}$.

Proof. The defining property of the inverse is that $gg^{-1} = e$, so we simply need to show that $(gh)(h^{-1}g^{-1}) = e$ and we can then identify that $h^{-1}g^{-1} = (gh)^{-1}$. Using associativity we can rewrite the brackets in the expression however we like, so we have $(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = geg^{-1} = gg^{-1} = e$. □

Lemma 1.2.8 Let G be a group, and $g \in G$. Then $(g^{-1})^{-1} = g$.

Proof. If $(g^{-1})^{-1} = g$ then we expect that $(g^{-1})^{-1}g^{-1} = e$. We can identify that $(g^{-1})^{-1}g^{-1} = (gg^{-1})^{-1}$ using Lemma 1.2.7. We then have $(g^{-1})^{-1}g^{-1} = (gg^{-1})^{-1} = e^{-1} = e$, since $ee = e$, so clearly $e^{-1} = e$. □

1.2.1 Examples of Groups

■ **Example 1.2.9 — Additive Groups** $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all form groups under addition. In particular the identity is 0, and the inverse of x is $-x$.

These same sets don't form groups under multiplication. The identity of multiplication is 1, and there is no number which acts as an inverse for multiplication by zero, that is there are no solutions to $0x = 1$ in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$. Noticing that zero causes issues with division it's sensible to consider these sets with zero removed. Denote by $\mathbb{Z}^* := \mathbb{Z} \setminus \{0\}$ the set of integers with zero removed. This is not a group since, for example, the multiplicative inverse of 2 is $1/2$, and $1/2 \notin \mathbb{Z}^*$.

■ **Example 1.2.10 — Multiplicative Groups** $\mathbb{Q}^*$, $\mathbb{R}^*$, and $\mathbb{C}^*$ all form a group under multiplication, where we use the notation $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$, $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$, and $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$. The identity is 1 and the inverse of x is $1/x$.

Definition 1.2.11 — Permutation Formally a **permutation** on n objects is a bijection

$$\sigma\colon \{1, \dots, n\} \to \{1, \dots, n\}. \tag{1.2.12}$$

This idea can be extended to any set of size n, not just $\{1, \dots, n\}$. Informally we can think of a permutation as a way of ordering the n objects such that the mth object is in position $\sigma(m)$.

■ **Example 1.2.13 — Permutation Group** The **permutation group**, S_n, defined as the set of all permutations on n objects, is a group under function composition. What this means is that if we permute the objects then permute them again we will have a permutation of the objects (closure), we can always leave the objects in the order they are (identity), and we can always undo a permutation (inverse).

■ **Example 1.2.14 — Cyclic Group** Take some $n \in \mathbb{Z}_{>0}$. We define $\mathbb{Z}_n := \{e^{2i\pi m/n} \mid m = 0, \dots, n-1\}$. This is a group under multiplication, called the **cyclic group** of order n.
Instead, define $\mathbb{Z}_n := \{0, \dots, n-1\}$. This is a group under addition modulo n. This is actually the same group as the previous definition of $\mathbb{Z}_n$ (they are isomorphic, a term defined later in Definition 2.1.1).

■ **Example 1.2.15 — Rotation Group** Define $\mathrm{O}(3) := \{O \in \mathcal{M}_3(\mathbb{R}) \mid O^\top O = \mathbb{1}\}$. This is a group under matrix multiplication. The identity is the identity matrix and the inverse is the normal matrix inverse, which is guaranteed to exist since for $O \in \mathrm{O}(3)$ we have $\det O = \pm 1 \neq 0$. This is called the **rotation group**.

R Strictly this is the fundamental representation of $\mathrm{O}(3)$.

1.2.2 Basic Definitions

Definition 1.2.16 — Group Size Given some group G we classify it as **finite**, **discrete**, or **continuous**, depending on whether G has a finite number of elements, the same number of elements as $\mathbb{Z}$, or more elements than $\mathbb{Z}$.

Recall that two sets have the same cardinality if there is a bijection between them. For example, $\mathbb{Z}$, $\mathbb{Z}_{>0}$, $\mathbb{N}$, and $\mathbb{Q}$ all have the same cardinality. A set is larger than a second set if there is an injective function from the first set to the second, but not vice versa. For example, there are more real numbers than integers.

■ **Example 1.2.17** Of the groups mentioned so far S_n, and $\mathbb{Z}_n$ are finite. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Z}^*$, and $\mathbb{Q}^*$ are discrete. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^*$, $\mathbb{C}^*$, and $\mathrm{O}(3)$ are continuous.

Definition 1.2.18 — Order The **order** of a finite group, G, is the number of elements in G, denoted $|G|$.
Given some group, G, the **order** of $g \in G$ is the smallest $n \in \mathbb{Z}_{>0}$ such that $g^n = e$, where e is the group identity and g^n has the expected meaning of the product of g with itself n times.
Note that the order of the identity is always 1.

■ **Example 1.2.19 — Order** The order of S_n is $|S_n| = n!$. The order of $\mathbb{Z}_n$ is $|\mathbb{Z}_n| = n$.
The order of $\mathrm{e}^{2i\pi 3/9} = \mathrm{e}^{2i\pi/3} \in \mathbb{Z}_9$ is 3 since $(\mathrm{e}^{2i\pi/3})^1 = \mathrm{e}^{2i\pi/3}$, $(\mathrm{e}^{2i\pi/3})^2 = \mathrm{e}^{4i\pi/3}$, and $(\mathrm{e}^{2i\pi/3})^3 = \mathrm{e}^{6i\pi/3} = \mathrm{e}^{2i\pi} = 1$, which is the identity of $\mathbb{Z}_n$.

Definition 1.2.20 — Abelian A group, G, is **Abelian** if all of its elements commute. That is $gg' = g'g$ for all $g \in G$. If this is not the case we say that G is non-Abelian.

■ **Example 1.2.21 — Abelian** Of the groups mentioned so far $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}^*$, $\mathbb{R}^*$, $\mathbb{C}^*$, and $\mathbb{Z}^*$ are Abelian. S_n and $\mathrm{O}(3)$ are non-Abelian.

Definition 1.2.22 — Subgroup Let G be a group. We say that H is a **subgroup** of G, denoted $H \subseteq G$, if

- H is a subset of G, and
- H is a group under the group operation of G restricted to elements of H.

A subgroup is said to be a **proper subgroup** if the subgroup is not equal to the full group, or the **trivial group**, $\{e\}$.

■ **Example 1.2.23 — Subgroup** $\mathbb{Z}_3$ is a subgroup of $\mathbb{Z}_9$ since both are groups and $\mathbb{Z}_3 = \{1, \mathrm{e}^{2i\pi/3}, \mathrm{e}^{2i\pi 2/3}\} \subset \mathbb{Z}_9$.

Definition 1.2.24 — Conjugate Given a group, G, we say that $g_1, g_2 \in G$ are **conjugate** if there exists $g \in G$ such that $g_1 = g g_2 g^{-1}$.

Lemma 1.2.25 Let G be a group. Then the relation $\sim$ defined by $g \sim h$ if g and h are conjugate in G is an equivalence relation.

Proof. Let $a \in G$. Then $a = eae^{-1}$ where e is the identity of G. This shows that $a \sim a$ and so $\sim$ is reflexive.
Let $a, b \in G$ be such that $a \sim b$. Then $a = gbg^{-1}$ for some $g \in G$. Right multiplying by g and left multiplying by g^{-1} this becomes $g^{-1}ag = g^{-1}gbg^{-1}g = b$. Noticing that $g^{-1} = g' \in G$ and $g = g'^{-1}$ this becomes $b = g'ag'^{-1}$ which shows that $b \sim a$ and so $\sim$ is symmetric.
Let $a, b, c \in G$ be such that $a \sim b$ and $b \sim c$. Then there exists $g, g' \in G$ such that $a = gbg^{-1}$ and $b = g'cg'^{-1}$. Inserting the second equation into the first we see that $a = gg'cg'^{-1}g^{-1}$. Now we write $g'' = gg' \in G$ and notice from Lemma 1.2.7 that $g'^{-1}g^{-1} = (gg')^{-1} = g''^{-1}$ we can write $a = g''cg''^{-1}$ and so $a \sim c$, meaning that $\sim$ is transitive. Hence, $\sim$ is an equivalence relation. □

Definition 1.2.26 — Generators Given a set $\{g_i\} \subseteq G$ we say that $\{g_i\}$ **generate** G if all elements of G can be written as a product of g_i. We call g_i **generators**.
The **rank** of a group is the size of the smallest set of generators.
If the rank of a group is 1 then there is one generator, g, and all elements are of the form g^n. We call such a group **cyclic**.

Definition 1.2.27 — Centre The **centre** of the group G is the set

$$Z(G) := \{z \in G \mid gz = zg \text{ for all } g \in G\}. \tag{1.2.28}$$

That is the centre is the set of all elements that commute with all other elements.

R Notice that the identity is always in the centre.

R The Z comes from the German *Zentrum* for centre.

Lemma 1.2.29 Given a group G the centre, $Z(G)$, is a subgroup of G.

Proof. Clearly $e \in Z(G)$ since $eg = ge$ for all $g \in G$. Let $z, z' \in Z(G)$, then $zz' \in Z(G)$ since

$$(zz')g = z(z'g) = z(gz') = (zg)z' = (gz)z' = g(zz') \tag{1.2.30}$$

for all $g \in G$. Finally, let $z \in Z(G)$, then $z^{-1} \in Z(G)$ since if $gz = zg$ for all $g \in G$ then left and right multiplying by z^{-1} we get $z^{-1}gzz^{-1} = z^{-1}g = z^{-1}zgz^{-1} = gz^{-1}$, and so $z^{-1} \in Z(G)$. We have shown that $Z(G)$ is a group and by construction it is a subset of G so $Z(G)$ is a subgroup of G. □

■ **Example 1.2.31 — Centre** If G is Abelian then $Z(G) = G$ The centre of S_3 is the trivial group. The centre of $\mathrm{O}(3)$ is $Z(\mathrm{O}(3)) = \{\mathbb{1}, -\mathbb{1}\}$.

Theorem 1.2.32 — Subgroup Criteria. Let G be a group and let H be a nonempty subset of G. Then H is a subgroup of G if and only if $g_1 g_2^{-1} \in H$ for all $g_1, g_2 \in H$.

Proof. Suppose H is a subgroup of G and $g_1, g_2 \in H$. The group axioms require that $g_2^{-1} \in H$. In order for H to be closed we must have $g_1 g_2^{-1} \in H$. Hence, if H is a subgroup of G then $g_1 g_2^{-1} \in H$ for all $g_1, g_2 \in H$.
Now suppose that $g_1 g_2^{-1} \in H$ for all $g_1, g_2 \in H$. Take some $g \in H$ and the condition gives $gg^{-1} = e \in H$, so the identity is in H. Using this we have $eg^{-1} = g^{-1} \in H$, so all elements of H have inverses in H. Take some $g_1, g_2 \in H$. We now know that $g_2^{-1} \in H$ and so using Lemma 1.2.8 we get $g_1(g_2^{-1})^{-1} = g_1 g_2 \in H$, thus H is closed under the operation. Hence, H is a group, and by definition it is a subset of G so H is a subgroup of G. □

1.2.3 Cayley Tables

Definition 1.2.33 — Cayley Table Given a finite group, G, we can list all possible products of pairs of group elements in a table, called a **Cayley table**, or **multiplication table**. This is done by listing the elements along the edge in some chosen order, usually starting with the identity, and taking the value in the ith row and jth column as the product of the ith and jth element of the group in the chosen order. That is

$$
\begin{array}{c|ccccc}
G & e & a & b & c & \cdots \\
\hline
e & e & a & b & c & \cdots \\
a & a & a^2 & ab & ac & \cdots \\
b & b & ba & b^2 & bc & \cdots \\
c & c & ca & cb & c^2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
\tag{1.2.34}
$$

The Cayley table for $\mathbb{Z}_2$ is

$$
\begin{array}{c|cc}
\mathbb{Z}_2 & 1 & -1 \\
\hline
1 & 1 & -1 \\
-1 & -1 & 1
\end{array}
\tag{1.2.35}
$$

The Cayley table for $\mathbb{Z}_3$ is

$$
\begin{array}{c|ccc}
\mathbb{Z}_3 & 1 & e^{2i\pi/3} & e^{2i\pi 2/3} \\
\hline
1 & 1 & e^{2i\pi/3} & e^{2i\pi 2/3} \\
e^{2i\pi/3} & e^{2i\pi/3} & e^{2i\pi 2/3} & 1 \\
e^{2i\pi 2/3} & e^{2i\pi 2/3} & 1 & e^{2i\pi/3}
\end{array}
\tag{1.2.36}
$$

This isn't that easy to read. There is a perhaps simpler way to think of $\mathbb{Z}_n$, as the set $\{0, \dots, n-1\}$, with addition modulo n as an operation. Using this we get the Cayley table

$$\begin{array}{c|ccc} \mathbb{Z}_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \tag{1.2.37}$$

Notice that the structure of these two tables is the same if we make the identification $1 \leftrightarrow 0$, $e^{2i\pi/3} \leftrightarrow 2$, and $e^{2i\pi 2/3} \leftrightarrow 2$. This is made clearer by colouring the entries in to the tables matching the colours based on this correspondence:

$\mathbb{Z}_3$	1	$e^{2i\pi/3}$	$e^{2i\pi2/3}$
1	1	$e^{2i\pi/3}$	$e^{2i\pi2/3}$
$e^{2i\pi/3}$	$e^{2i\pi/3}$	$e^{2i\pi2/3}$	1
$e^{2i\pi2/3}$	$e^{2i\pi2/3}$	1	$e^{2i\pi/3}$

$\mathbb{Z}_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

(1.2.38)

We will see in a bit that what we really are saying here is that the multiplicative group $\{1, e^{2i\pi/3}, e^{2i\pi2/3}\}$ and the group $\{0, 1, 2\}$ under addition modulo 3 are isomorphic, and so have the same structure and all of their group theoretical properties are the same. For this reason we often simply think of them as being the same and just consider a single group $\mathbb{Z}_3$ using whichever of these groups is most useful at the moment.

We can make a similar identification between the group generated by multiplication of $e^{2i\pi/n}$ and the group of $\{0, \dots, n-1\}$ under addition modulo n, both of which can be thought of as $\mathbb{Z}_n$ by identifying $1 \leftrightarrow 0$ and $e^{2i\pi m/n} \leftrightarrow m$. For example, we can thin, of $\mathbb{Z}_2$ as addition modulo 2 on $\{0, 1\}$.

> **Notation 1.2.39 — Cycle Notation** A k-**cycle** is a way of writing a permutation down. For $a_i \in \{0, 1, \dots, n\}$ we write $(a_1\, a_2\, \dots\, a_m)$ to denote the permutation in S_n that sends a_1 to a_2, a_2 to a_3, and so on, sending a_{m-1} to a_m, and finally a_m to a_1.
> Using this notation the identity permutation is denoted $()$.

For example, consider the 2-cycle $(1\,2)$ acting on the objects tuple (a, b, c). This sends 1, which here is the first object, a, to 2, which here is the second object b, and sends the second object to the first object. We can write this as

$$(1\,2)(a, b, c) = (b, a, c). \tag{1.2.40}$$

Applying this 2-cycle a second time we get

$$(1\,2)^2(a, b, c) = (1\,2)(1\,2)(a, b, c) = (1\,2)(b, a, c) = (a, b, c), \tag{1.2.41}$$

and so we see that $(1\,2)^2 = ()$.

Now consider the 3-cycle $(1\,2\,3)$ acting on (a, b, c). We see that

$$(1\,2\,3)(a, b, c) = (c, a, b), \tag{1.2.42}$$

$$(1\,2\,3)^2 = (1\,2\,3)(c, a, b) = (b, c, a), \tag{1.2.43}$$

$$(1\,2\,3)^3 = (1\,2\,3)(b, c, a) = (a, b, c), \tag{1.2.44}$$

so $(1\,2\,3)^3 = ()$.

Carrying on like this we can build up the Cayley table for S_3, notice that all 2-cycles are self-inverses (i.e. they are order 2, so square to give the identity):

S_3	$()$	$(1\,2)$	$(2\,3)$	$(1\,3)$	$(1\,2\,3)$	$(3\,2\,1)$
$()$	$()$	$(1\,2)$	$(2\,3)$	$(1\,3)$	$(1\,2\,3)$	$(3\,2\,1)$
$(1\,2)$	$(1\,2)$	$()$	$(1\,2\,3)$	$(3\,2\,1)$	$(2\,3)$	$(1, 3)$
$(2\,3)$	$(1\,3)$	$(3\,2\,1)$	$()$	$(1\,2\,3)$	$(1\,3)$	$(1\,2)$
$(1\,3)$	$(1\,3)$	$(1\,2\,3)$	$(3\,2\,1)$	$()$	$(1\,2)$	$(2\,3)$
$(1\,2\,3)$	$(1\,2\,3)$	$(1\,3)$	$(1\,2)$	$(2\,3)$	$(3\,2\,1)$	$()$
$(3\,2\,1)$	$(3\,2\,1)$	$(2\,3)$	$(1\,3)$	$(1\,2)$	$()$	$(3\,2\,1)$

(1.2.45)

Cayley tables can be a useful way to visualise group operations for small groups. For example, we can see that S_3 has as a subgroup $\{(), (1\,2)\}$, which is the upper left-hand corner of the table, and that this is equivalent to $\mathbb{Z}_2$ after making the correspondence $1 \leftrightarrow ()$ and $-1 \leftrightarrow (1\,2)$, again this can be seen more easily by colouring in the relevant entries:

$\mathbb{Z}_2$	1	-1
1	1	-1
-1	-1	1

S_3	$()$	$(1\,2)$	$(2\,3)$	$\cdots$
$()$	$()$	$(1\,2)$	$(2\,3)$	$\cdots$
$(1\,2)$	$(1\,2)$	$()$	$(1\,2\,3)$	$\cdots$
$(2\,3)$	$(1\,3)$	$(3\,2\,1)$	$()$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

(1.2.46)

Theorem 1.2.47 — Rearrangement Theorem. The rows and columns of a multiplication table are permutations of the group. That is they contain each element of the group exactly once.

Proof. Suppose that there is a row of the Cayley table for G such that $g \in G$ appears more than once, say this is the row associated with $g' \in G$. That means that there exist two elements $g_1, g_2 \in G$ such that $g'g_1 = g'g_2 = g$. Applying the left inverse to g' we get $g_1 = g_2$, and so g cannot appear more than once.
Since all columns of the table must be filled and there are $|G|$ columns and $|G|$ elements in order to have no repeats each element must appear once. □

Identifying the permutations giving the rows with the element of the group that is associated with that row we get the next theorem. The statement of the theorem is in terms of isomorphisms which we will define shortly but for now think of "is isomorphic to" as meaning "is equivalent to in the sense of the Cayley tables above having the same structure after renaming elements". Skip the proof until we've covered isomorphisms and then come back and look at it.

Theorem 1.2.48 — Cayley's Theorem. Any finite group is isomorphic to a subgroup of the symmetric group.

Proof. Let G be a finite group and let $S_{|G|}$ be the permutation group of order $|G|$. The for each $g \in G$ we can define $\sigma_g \colon G \to G$ to be $\sigma_g(g') = gg'$. This

function is invertible since $\sigma_{g^{-1}}$ is its inverse, as can be seen by considering $\sigma_{g^{-1}}(\sigma_g(g')) = \sigma_{g^{-1}}(gg') = g^{-1}gg' = g'$. This means that σ_g is bijective and hence is a permutation on the set G.
Now define $\varphi\colon G \to S_{|G|}$ by $\varphi(g) = \sigma_g$. Then φ is a homomorphism since

$$\begin{aligned}(\varphi(gg'))(g'') &= \sigma_{gg'}(g'') = gg'g'' = \sigma_g(g'g'') \\ &= \sigma_g(\sigma_{g'}(g'')) = (\sigma_g \circ \sigma_{g'})(g'') = (\varphi(g)\varphi(g'))(g''). \end{aligned} \tag{1.2.49}$$

Now suppose $\varphi(g) = \varphi(g')$, then $\sigma_g = \sigma_{g'}$, meaning $gg'' = g'g''$ for all $g'' \in G$, which means that $g = g'$ since we can apply g''^{-1} to the right of this equation. This shows that φ is injective.
The function $\tilde{\varphi}\colon G \to \mathrm{Im}(\varphi)$ given by $\tilde{\varphi}(g) = \varphi(g)$ is a surjective. It remains only to show that $\mathrm{Im}\,\varphi$ is a subgroup of $S_{|G|}$. This will be proven in Lemma 2.1.21, and so we have proven the theorem. □

Cayley's theorem is similar in nature to the Whitney embedding theorem which states that any manifold can be embedded into Euclidean space, $\mathbb{R}^n$, for suitable n. We just swap "manifold" with "group", "embedding" with "isomorphism", and "Euclidean space, $\mathbb{R}^n$" with "a subgroup of the permutation group, S_n".

1.3 Almost Groups

In this course we are interested in the study of groups. But why? Well, it turns out that groups are the natural way to discuss symmetries, and symmetries occur all over the place in physics. Broadly if we have some sort of symmetry we can always do nothing (identity) and undo the symmetry (inverses). In the act of chaining symmetries one after another it also shouldn't matter how we combine them (associativity). This is why we use groups to define symmetries. It is worth wandering briefly what happens if we relax some of the group axioms. We aren't the first to think this and many of the resulting structures have already got names. Throughout the next few chapters we will prove many theorems and lemmas. Many of these hold for some relaxed version of a group, just check which of the group axioms are used in the proof.

The most relaxed form of a group is just a set. This isn't that interesting on its own, so we move on to the next most relaxed form of a group.

Definition 1.3.1 — Magma A **magma**, $(X, \cdot)$, is a set, X, and a binary operation, $\cdot\colon X \times X \to X$.

■ Example 1.3.2 — Magma

- $(\mathbb{R}^3, \times)$ is a magma where $\mathbb{R}^3$ is the space of three-tuples of real numbers, (x, y, z) and $\times$ is the usual vector cross product such that

$$(x, y, z) \times (a, b, c) := (yc - zb, za - xc, xb - ya). \tag{1.3.3}$$

- $(\mathbb{R}, f\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}; f(x, y) = x^3y + 2x)$ is a magma.

Definition 1.3.4 — Semigroup A **semigroup**, $(X, \cdot)$, is a magma such that $\cdot$ is associative.

■ Example 1.3.5 — Semigroup

- $(\mathbb{Z}_{>0}, +)$ is a semigroup.
- $(\{f^{n*}\}, *)$ is a semigroup where f is some (sufficiently smooth) function, $*$ is convolution, and $f^{n*} = f * f * \cdots * f$ for n copies of f with $n \in \mathbb{Z}_{>0}$.
- $(X, \cdot)$ is a semigroup called the null semigroup if there exists $0 \in X$ such that $xy = 0$ for all $x, y \in X$.

Definition 1.3.6 — Monoid A **monoid**, $(X, \cdot, 1)$, is a semigroup such that $1 \in X$ acts as an identity for $\cdot$.

■ Example 1.3.7 — Monoid

- $(\mathbb{N}, +, 0)$ is a monoid.
- $(\mathbb{Z}, \cdot, 1)$ is a monoid.
- More generally if $(R, \cdot, +, 1, 0)$ is a ring (with unity) then $(R, \cdot, 1)$ is a monoid (and $(R, +, 0)$ is an Abelian group).
- $(\mathcal{M}_n(\mathbb{F}), \cdot)$ is a monoid where $\mathcal{M}_n(\mathbb{F})$ is the set of $n \times n$ square matrices over the field $\mathbb{F}$ and $\cdot$ is matrix multiplication.
- If $(X, \cdot)$ is a semigroup then $(X \cup \{e\}, \cdot, e)$ is a monoid where we extend the definition of $\cdot$ to $X \cup \{e\}$ by defining $xe = ex = x$ for all $x \in X$ and $e^2 = e$.
- $(\{\mathrm{True}, \mathrm{False}\}, \mathrm{XOR}, \mathrm{False})$ is a monoid.
- $(\mathcal{P}(A), \cup, \emptyset)$ is a monoid where $\mathcal{P}(A)$ is the power set of some set A, $\cup$ is the union of sets, and $\emptyset$ is the empty set.
- $(\mathcal{P}(A), \cap, A)$ is a monoid where $\cap$ is the intersection of sets.
- $(\mathrm{Hom}_{\mathrm{Set}}(X, X), \circ, \mathrm{id}_X)$ is a monoid where $\mathrm{Hom}_{\mathrm{Set}}(X, X)$ is the set of all functions $X \to X$ for some set X, more generally $\mathrm{Hom}_{\mathbf{C}}(A, B)$ is the class of all morphisms from A to B in some category $\mathbf{C}$. Also, $\circ$ is composition of functions and id_X is the identity function, defined by $\mathrm{id}_X(x) = x$ for all $x \in X$. It turns out that all monoids are isomorphic to a submonoid of a monoid of this type, this is similar to all groups being isomorphic to a subgroup of a permutation group, in particular by requiring invertible functions for a group we get bijections instead of just normal functions, and for finite groups bijections can be interpreted as permutations.

- $(\mathrm{End}_{\mathbf{Grp}}(G), \circ, \mathrm{id}_G)$ is a monoid where G is a group, $\mathrm{End}_{\mathbf{Grp}}(G)$ is the set of all **endomorphisms** of G, which is to say homomorphisms (to be defined in the next section) of G with itself, $\mathrm{End}_{\mathbf{Grp}}(G) := \mathrm{Hom}_{\mathbf{Grp}}(G, G)$.

- More generally $(\mathrm{End}_{\mathbf{C}}(X), \circ, \mathrm{id}_X)$ is a monoid where $\mathbf{C}$ is a category, X is an object of $\mathbf{C}$, and $\mathrm{End}_{\mathbf{C}}(X)$ is the set of morphisms from X to X in $\mathbf{C}$, that is $\mathrm{End}_{\mathbf{C}}(X) = \mathrm{Hom}_{\mathbf{C}}(X, X)$. The previous example is subsumed by this one taking $\mathbf{C} = \mathbf{Grp}$, the category of groups with group homomorphisms as morphisms.

- $(\mathrm{Hom}_{\mathbf{C}}(\bullet, \bullet), \circ, \mathrm{id}_\bullet)$ is a monoid where $\mathbf{C}$ is a category with a single object, $\bullet \in \mathrm{Obj}(\mathbf{C})$. The elements of this monoid are the morphisms from this object to itself with morphism composition as a binary operation and the identity morphism, $\mathrm{id}_\bullet$, as the identity. This operation is associative, and the identity exists by the definition of a category. If all morphisms are isomorphisms then they are invertible and so this is a group.

We can now define a group having worked our way up to it one property at a time.

Definition 1.3.8 — Group A **group**, $(X, \cdot, 1, -^{-1})$, is a monoid, $(X, \cdot, 1)$, equipped with a function $-^{-1}\colon X \to X$ such that x^{-1} acts as the inverse of x for all $x \in X$.

Note that we previously just wrote $(G, \cdot)$, leaving 1 and $-^{-1}$ out of the notation. In fact, we usually don't even write the operations, we just say

- Let M be a magma with operation $\cdot$, for $(M, \cdot)$,
- Let M be a monoid with operation $\cdot$ and identity 1 for $(M, \cdot, 1)$, and
- Let G be a group with operation $\cdot$, identity 1, and inverses g^{-1} for $(G, \cdot, 1, -^{-1})$.

Two

Morphisms and Cosets

2.1 Morphisms

Definition 2.1.1 — Morphism A **homomorphism** between groups G and H is a map $\varphi\colon G \to H$ which preserves the group product. That is for all $g, g' \in G$ we have

$$\varphi(gg') = \varphi(g)\varphi(g'). \tag{2.1.2}$$

R The product gg' on the left is the group product of G whereas the product $\varphi(g)\varphi(g')$ on the right is the group product of H. We can emphasise this by writing $\varphi(g \cdot_G g') = \varphi(g) \cdot_H \varphi(g')$.

An **isomorphism** between groups G and H is a bijective homomorphism. If there exists an isomorphism between G and H then we say that G and H are **isomorphic** and denote this $G \cong H$.

An isomorphism preserves all group structure. That means we can think of isomorphic groups as being the same group, just with the labels of the elements and the group operation renamed. We've already seen one example of this, $\{\pm 1, \pm i\}$ with the group operation of multiplication is isomorphic to, and hence considered the same as, $\{0, 1, 2, 3\}$ with the group operation of addition modulo 4.

In group theory we are almost always only interested in properties holding "up to isomorphism". For example, we may say "there is one group up to isomorphism with some property", by which we actually mean that all groups with this property are isomorphic. Often the "up to isomorphism" is left implicit, and we just say "there is one group with some property".

Homomorphism comes from ὁμός (*homos*) meaning same, and μορφή (*morphe*) meaning shape or form. Isomorphism comes from ἴσος (*isos*) meaning equivalent or equal, and μορφή (*morphe*) meaning shape or form.

Note that the relation $\cong$ on the set of all groups defined by $G \cong H$ if G and H are isomorphic is an equivalence relation (see Example A.1.25). This is what justifies us saying that two isomorphic groups are the same. Isomorphism is exactly what we mean when we say two groups are equivalent, rather than the stricter meaning of being exactly equal.

■ **Example 2.1.3 — Trivial Examples** Consider the trivial group, $\{e\}$, consisting of a single element, which must act as an identity. Then $\varphi\colon G \to \{e\}$ for some group G, defined by $\varphi(g) = e$ for all $g \in G$ is a homomorphism since $\varphi(gg') = e = ee = \varphi(g)\varphi(g')$. This is not an isomorphism unless $G = \{e\}$.
Similarly, there exists a homomorphism between any two groups, G and H, by sending all elements of G to the identity of H.
Every group, G, is isomorphic to itself since the identity function, $\mathrm{id}_G\colon G \to G$, defined by $\mathrm{id}_G(g) = g$ for all $g \in G$ is an isomorphism. That is $\mathrm{id}_G(gg') = gg' = \mathrm{id}_G(g)\mathrm{id}_G(g')$, and id_G is a self inverse, so id_G is bijective.

■ **Example 2.1.4 — Groups of Order 2 and 3** $\mathbb{Z}_2$ and S_2 are isomorphic. First notice that there are two permutations on 2 objects, we either leave them as is, $()$, or swap them, $(1\,2)$. Then $() \mapsto 1$ and $(1\,2) \mapsto -1$ is an isomorphism. To see this note that $(1\,2)(1\,2) = ()$, that is swapping and swapping back has no net effect, and so

$$\varphi(1 \cdot 1) = \varphi(1) = () = ()() = \varphi(1)\varphi(1), \tag{2.1.5}$$
$$\varphi((-1) \cdot (-1)) = \varphi(1) = () = (1\,2)(1\,2) = \varphi(-1)\varphi(-1), \tag{2.1.6}$$
$$\varphi(1 \cdot (-1)) = \varphi(-1) = (1\,2) = ()(1\,2) = \varphi(1)\varphi(-1). \tag{2.1.7}$$

The final $\varphi((-1)\cdot 1)$ case is covered by the $\varphi(1\cdot(-1))$ case since both groups are Abelian.
In fact all groups of order two are isomorphic to $\mathbb{Z}_2$ under the isomorphism of sending the identity to 1 and the non-identity to the -1. Similarly, all groups of order three are isomorphic to $\mathbb{Z}_3$.

■ **Example 2.1.8 — Discrete Isomorphisms** The group $\mathbb{Z}$ under addition is isomorphic to the group $2\mathbb{Z}$ under addition. Here $n\mathbb{Z}$ is understood to be the set of integer multiples of n, so $2\mathbb{Z}$ is the set of even integers.
One isomorphism, $\varphi\colon \mathbb{Z} \to 2\mathbb{Z}$, is the obvious choice of $\varphi(n) = 2n$. First we check that this is a homomorphism, given some $m, n \in \mathbb{Z}$ we have

$$\varphi(n+m) = 2(n+m) = 2n + 2m = \varphi(n) + \varphi(m), \tag{2.1.9}$$

so this is indeed a homomorphism.
Next we check that this is injective. Suppose $\varphi(n) = \varphi(m)$ for two elements $n, m \in \mathbb{Z}$. Then $2n = 2m$, which readily implies $n = m$, and so φ is injective.
Finally, we check that this is surjective. Consider some $n \in 2\mathbb{Z}$, since this is even[a] $2|n$ and so $n/2$ is an integer. It follows that for each $n \in 2\mathbb{Z}$ we have $n/2 \in \mathbb{Z}$ as the element such that $\varphi(n/2) = n$, and so φ is surjective.
Note that we could also have identified that $\varphi^{-1}(n) = n/2$ is the inverse of φ. Either way φ is a bijective homomorphism and hence an isomorphism.

[a] Recall that $m|n$ means m divides n, meaning that n is an integer multiple of m and n/m is an integer.

■ **Example 2.1.10 — Continuous Isomorphisms** The groups $(\mathbb{R}, +)$ and $(\mathbb{R}_{>0}, \cdot)$ of positive real numbers under multiplication are isomorphic.
One isomorphism between these groups is $x \mapsto \mathrm{e}^x$, which is a homomorphism since

$$\mathrm{e}^{x+y} = \mathrm{e}^x \mathrm{e}^y \tag{2.1.11}$$

for all $x, y \in \mathbb{R}$ and is bijective since $x \mapsto \ln x$ is the inverse.

■ **Example 2.1.12 — Complex Numbers as Matrices** The multiplicative group of complex numbers is isomorphic to the subset of 2×2 real matrices

$$\left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in \mathbb{R} \right\}. \tag{2.1.13}$$

An isomorphism between these two groups is given by

$$\varphi(a + bi) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}. \tag{2.1.14}$$

We first check that this is a homomorphism:

$$\varphi((a+bi)(c+di)) = \varphi((ac-bd)+(ad+bc)i) = \begin{pmatrix} ac - bd & ad + bc \\ -(ad + bc) & ac - bd \end{pmatrix} \tag{2.1.15}$$

and

$$\varphi(a + bi)\varphi(c + di) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac - bc & ad + bc \\ -bc - ad & bd + ac \end{pmatrix}, \tag{2.1.16}$$

so this is indeed a homomorphism.
We can see that this is bijective by noticing that the inverse is simply

$$\varphi^{-1}\left(\begin{pmatrix} a & b \\ -b & a \end{pmatrix}\right) = a + bi. \tag{2.1.17}$$

We will see later that this is a two-dimensional real representation of $\mathbb{C}^*$.
Notice that if we restrict ourselves to complex numbers with unit modulus then we can write $a = \cos\vartheta$ and $b = \sin\vartheta$ which allows us to make an identification between complex numbers of unit modulus and two-dimensional rotations. Denoting the multiplicative group of complex numbers with unit modulus by $\mathbb{T}$, the group of 1×1 unitary matrices by $\mathrm{U}(1)$ and the group of two-dimensional rotations by $\mathrm{SO}(2)$ what we see here is that $\mathrm{U}(1) \cong \mathbb{T} \cong \mathrm{SO}(2)$, where the isomorphism between $\mathbb{T}$ and $\mathrm{U}(1)$ is the obvious one mapping $z \in \mathbb{T}$ to $(z) \in \mathrm{U}(1)$ and the isomorphism between $\mathbb{T}$ and $\mathrm{SO}(2)$ is the restriction of φ to $\mathbb{T}$.

It isn't until we get to groups of order 4 that we get two groups which *aren't* isomorphic. The two groups of order 4 are $\mathbb{Z}_2$ and the **Klein *Vierergruppe***, $\mathbb{Z}_2 \times \mathbb{Z}_2$[1], which has the unique property for a group of this order that all non-trivial (i.e. not

[1] This notation will make sense when we talk about direct products of groups in Definition 5.1.1

the identity) elements are of order 2. These two groups have the Cayley tables

$\mathbb{Z}_4$	1	-1	i	$-i$
1	1	-1	i	$-i$
-1	-1	1	$-i$	i
i	i	$-i$	-1	1
$-i$	$-i$	i	1	-1

$\mathbb{Z}_2 \times \mathbb{Z}_2$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(2.1.18)

It is possible to fill the second one of these in by starting with the first row and column, which are simple taking e as the identity, and the leading diagonal, which must be e in every row since we have declared all nontrivial elements to be of order 2. This leaves the last two slots on the a line open, since each element must appear exactly once in each row and column these slots must be c and b. Continuing on we can fill out the rest of the table. Notice that we don't need to ever discuss what the elements e, a, b, and c are. It is enough to know that they form a group with this property of squaring to the identity. We will see later one possible set of elements that naturally form a group of this structure in Example 5.1.12.

There are some immediate consequences of these definitions that are worth considering.

Lemma 2.1.19 — Homomorphisms Map Identities to Identities Let G and H be groups. Then if $\varphi\colon G \to H$ is a homomorphism $\varphi(e_G) = e_H$ where e_G and e_H are the identities of G and H respectively.

Proof. By definition $\varphi(gg') = \varphi(g)\varphi(g')$ for all $g, g' \in G$. In particular, we have $\varphi(e_G g) = \varphi(e_G)\varphi(g)$, and $\varphi(e_G g) = \varphi(g) = e_H\varphi(g)$. Right multiplying these two results by $\varphi(g)^{-1}$ we have $\varphi(e_G) = e_H$. □

Lemma 2.1.20 — Homomorphisms Map Inverses to Inverses Let G and H be groups and $\varphi\colon G \to H$ a homomorphism. Then $\varphi(g^{-1}) = \varphi(g)^{-1}$.

Proof. By definition $\varphi(gg') = \varphi(g)\varphi(g')$ for all $g, g' \in G$. By Lemma 2.1.19 we have $\varphi(e_G) = e_H$ where e_G and e_H are the identities of G and H. We then have $\varphi(gg^{-1}) = \varphi(e_G) = e_H$, and also $\varphi(gg^{-1}) = \varphi(g)\varphi(g^{-1})$. From this we see that $e_H = \varphi(g)\varphi(g^{-1})$, which is to say that $\varphi(g^{-1}) = \varphi(g)^{-1}$ since $\varphi(g)\varphi(g^{-1})$ gives the identity, which defines the inverse. □

Lemma 2.1.21 — The Image of a Homomorphism is a Subgroup Let G and H be groups and $\varphi\colon G \to H$ a homomorphism. Then $\operatorname{Im}\varphi = \varphi(G)$ is a subgroup of H.

Proof. Consider $h, h' \in \operatorname{Im}\varphi$. Then there exists $g, g' \in G$ such that $h = \varphi(g)$ and $h' = \varphi(g')$. Since G is a group $gg'^{-1} \in G$. We then have $\varphi(gg'^{-1}) = \varphi(g)\varphi(g'^{-1}) = \varphi(g)\varphi(g')^{-1} = hh'^{-1}$ by Lemma 2.1.20 and the defining property of a homomorphism. This means that $hh'^{-1} \in \operatorname{Im}\varphi$ and hence $\operatorname{Im}\varphi$ is a subgroup of H by the subgroup criterion of Theorem 1.2.32. □

Lemma 2.1.22 Every group of rank 1 is isomorphic to some cyclic group.

Proof. First suppose that G is a finite group of order n. Then elements of G are of the form g^i for some $i = 0, \dots, n-1$. In particular $g^0 = e$. Notice that $f^i = g^j$ if $i \equiv j \bmod n$.
The map $\varphi\colon G \to \mathbb{Z}_n$ defined by $\varphi(g^i) = i$ is then an isomorphism, using $\mathbb{Z}_n$ as the group of integers under addition modulo n. Clearly $\varphi(g^i g^j) = \varphi(g^{i+j}) = i + j = \varphi(g^i) + \varphi(g^j)$ where addition outside of the argument of φ occurs modulo n.
The inverse of this map is simply $\varphi^{-1}(i) = g^i$ and so this is a bijection. Meaning that φ is an isomorphism. □

Lemma 2.1.23 A homomorphism is injective if and only if its kernel is trivial.

Proof. Let G and H be groups and $\varphi\colon G \to H$ a homomorphism. Suppose $\ker\varphi = \{e_G\}$, where e_G is the identity of G. Suppose $\varphi(g) = \varphi(h)$ for some $g, h \in G$. Then $\varphi(g)^{-1} = \varphi(h)^{-1}$ so $\varphi(g)\varphi(h)^{-1} = e_H$, where e_H is the identity of H. Now using Lemma 2.1.20 we can write this as $\varphi(g)\varphi(h^{-1}) = e_H$, and using the defining property of the homomorphism this gives us $\varphi(gh^{-1}) = e_H$. Finally, using Lemma 2.1.19 we have $gh^{-1} = e_G$. Right multiplying by h we have $g = h$, and so φ is injective.
Suppose instead that φ is injective. Let $g \in \ker\varphi$. Then $\varphi(g) = e_H$. However, we know from Lemma 2.1.19 that $e_G \in \ker\varphi$, and so $\varphi(e_G) = e_H = \varphi(g)$, then the injective nature of φ means $e_G = g$, and so $\ker\varphi = \{e_G\}$. □

2.2 Group Presentations

Definition 2.2.1 — Group Presentations A **group presentation** is a way of defining a specific group. A generic group presentation is of the form

$$G = \langle S \mid C \rangle \tag{2.2.2}$$

which we read as "G is the group generated by the elements of S subject to the constraints C".

▪ **Example 2.2.3 — Cyclic Groups** The cyclic group, $\mathbb{Z}_n$, has the group presentation

$$\mathbb{Z}_n = \langle a \mid a^n = e \rangle. \tag{2.2.4}$$

We can identify $a = \mathrm{e}^{2i\pi/n}$ as one possible generator with the group operation of multiplication, but it need not be the only one. For example if $n = 3$ then $a = \mathrm{e}^{2i\pi 2/3}$ also works.
If instead we take the group operation to be addition modulo n then $a = 1$ is a generator. For the $n = 3$ case we can also choose $a = 2$ as a generator.

■ **Example 2.2.5 — Permutation Group** The following is group presentations of S_3:

$$S_3 = \langle a, b, c \mid a^2 = b^2 = c^3 = abc = e \rangle. \tag{2.2.6}$$

We have a fair amount of choice here about exactly which elements a, b, and c are, clearly $e = ()$ is the identity. We can then choose a and b to be any of the 2-cycles, $(1\,2)$, $(1\,3)$, and $(2\,3)$, and c as one of the 3-cycles, $(1\,2\,3)$ or $(3\,2\,1)$.
Group presentations aren't unique. For example, the following is another valid presentation of S_3:

$$S_3 = \langle A, B \mid A^2 = B^2 = (AB)^3 = e \rangle. \tag{2.2.7}$$

We can see from this that A and B must be 2-cycles and AB is a three cycle.

■ **Example 2.2.8 — Quaternion Group** The **quaternion group** is the group with the presentation

$$Q := \langle -e, i, j, k \mid (-e)^2 = e, i^2 = j^2 = k^2 = ijk = -e \rangle. \tag{2.2.9}$$

This group is of order $|Q| = 8$, with $Q = \{\pm e, \pm i, \pm j, \pm k\}$, where by $-i$ we mean $(-e)i$.
Making the identification of $e = 1$ and taking i, j, and k, as the quaternions $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ respectively it is clear that this is a subset of the quaternions, $\mathbb{H}$.
We can think Q being to $\mathbb{H}$ as $\mathbb{Z}_4 = \{\pm 1, \pm i\}$ is to $\mathbb{C}$, or $\mathbb{Z}_2 = \{\pm 1\}$ is to $\mathbb{R}$.
Another identification we can make is $-e = -\mathbb{1}$, $i = \sigma_1$, $j = \sigma_2$, and $k = \sigma_3$, where σ_i are the Pauli matrices. In fact the Pauli matrices are a representation of the quaternion group.

2.3 Cosets

Definition 2.3.1 — Coset Given some group G and subgroup H we define for each $g \in G$ the left (right) **coset** to be the set

$$gH := \{gh \mid h \in H\} \tag{2.3.2}$$

$$(Hg := \{hg \mid h \in H\}). \tag{2.3.3}$$

Typically, we will state and prove things for left cosets and then the equivalent statement about right cosets will hold and be proven in exactly the same way. We will often refer simply to cosets when we mean left cosets.

Definition 2.3.4 — Partition Given a nonempty set, X, we say that the collection of sets $\{P_i \subseteq X\}$ is a **partition** or **decomposition** of X if

- $\bigcup_i P_i = X$, and
- $P_i \cap P_j = \emptyset$ if $i \neq j$.

That is every element of X is in exactly one of P_i. We can assume that P_i are non-empty.

Lemma 2.3.5 Let G be a group with subgroup H. Then the set of all cosets, gH, partitions G. Further, all cosets are of the same size, meaning $|gH| = |H|$ for all $g \in G$.

Proof. In order for the cosets to be a partition we must show that two cosets are either equal or disjoint. Consider some element $g_1 \in G$ which is not in some coset g_2H. Clearly this means that the two cosets g_1H and g_2H are equal since $e \in H$ so $g_1e = g_1$ is in g_1H. Suppose then that $g_1H \cap g_2H \neq \emptyset$. Then it follows that there exist some $h_1, h_2 \in H$ such that $g_1h_1 = g_2h_2$. This then means that $g_1 = g_2h_2h_1^{-1}$, however, since H is a group $h_2h_1^{-1} \in H$, and so $g_1 = g_2h$ for some $h = h_2h_1^{-1} \in H$ meaning that $g_1 \in g_2H$, which contradicts our earlier assumption. Hence, $g_1H \cap g_2H = \emptyset$. Combining this with noticing that for $g \in G$ we have $g \in gH$ since $e \in H$ and so $g = ge$ means that $g \in gH$ proves the first part of the statement, that gH partition G.
Consider the map $g_1H \to g_2H$ defined by $g_1h \mapsto g_2h$ for $h \in H$. This is invertible since inverses are unique and hence $|g_1H| = |g_2H|$. In particular taking $g_1 = g$ and $g_2 = e$ we have $|gH| = |H|$. □

The fact that gH partition G into sets of equal size allows us to prove the next famous theorem. But first, a definition.

Definition 2.3.6 — Index Given a finite group G with subgroup H we define the **index** of H in G to be

$$[G : H] \coloneqq \frac{|G|}{|H|}. \tag{2.3.7}$$

Theorem 2.3.8 — Lagrange's Theorem. Given a finite group G with subgroup H the index $[G : H]$ is an integer.

Proof. The cosets partition G into sets of size $|H|$. Suppose that there are n distinct cosets. Then $|G| = n|H|$, meaning that $|G|/|H| = n$. □

Lagrange's theorem says that a subset can be a subgroup only if the cardinality of the subset divides the order of the group. Notice that just because this holds does not mean that the subset is a subgroup. There is also no requirement that just because a number divides the order of the group that there is a subgroup of that order. Lagrange's theorem is much better for ruling out possible subgroups than it is for actually finding them.

▪ **Application 2.3.9** In particle physics and statistical mechanics if a continuous global symmetry given by the group G is broken to some subgroup H then it is possible to formulate an effective field theory in terms of cosets. For example in the theory of strong interactions, quantum chromodynamics

(QCD), the breaking of left-right symmetry, known as chiral symmetry, gives rise to the effective theory of pions, known as chiral perturbation theory.

Three

Group Action

3.1 Group Action

Definition 3.1.1 Let G be a group and X a set. A **group action** is a map, $\varphi\colon G \times X \to X$, where we use the notation $\varphi(g, x) = g \,.\, x$. This map must be compatible with the group structure, by which we mean

- $e \,.\, x = x$ ($\varphi(e, x) = x$) for all $x \in X$, and
- $(gg') \,.\, x = g \,.\, (g' \,.\, x)$ ($\varphi(gg', x) = \varphi(g, \varphi(g', x))$) for all $g, g' \in G$ and $x \in X$.

! Be careful to distinguish "." used to the group action and "$\cdot$" used to denote a group product, in general $g \,.\, g' \neq g \cdot g'$. This is another good reason *not* to use a dot to denote the group product.

R Technically what we have defined here is a *left* group action. We can also define a right group action similarly. Let G be a group and X a set. The right group action is a map, $\varphi\colon X \times G \to X$, where we use the notation $\varphi(x, g) = x . g$. This map must be compatible with the group structure, by which we mean

- $x \,.\, e = x$ ($\varphi(x, e) = x$) for all $x \in X$, and
- $x \,.\, (gg') = (x \,.\, g) \,.\, g'$ ($\varphi(x, gg') = \varphi(\varphi(x, g), g')$) for all $g, g' \in G$ and $x \in X$.

The difference between left and right group actions is subtle. For a left group action if we act on x with the product gg' then g' acts first, for a right group action g acts first.

An alternative definition for finite groups is that a group action is a group homomorphism from G into $S_{|X|}$. That is we can think of the action as taking an element of G and then determining how to reorder the elements of X based on this choice.

Identifying $S_{|X|}$ with bijections from X to X we can further define a group action to be a homomorphism $\varphi\colon G \to \operatorname{Aut}(X)$, where $\operatorname{Aut}(X)$ is the group of automorphisms on X, which is to say exactly the set of bijections $X \to X$, with the group product being function composition.

■ **Example 3.1.2 — General Examples** For any group G and set X the trivial group action is

$$g \,.\, x = x \tag{3.1.3}$$

for all $x \in X$.
For the special case of $X = G$ we have a variety of choices for the group action of a group on itself:

- Left multiplication: $g \,.\, x = gx$ for all $g, x \in G$,
- Right multiplication: $g \,.\, x = xg$ for all $g, x \in G$ (strictly this is a *right* group action), and
- Conjugation: $g \,.\, x = gxg^{-1}$.

■ **Example 3.1.4 — Specific Examples** Let S_n be the permutation group on n objects and X the set of all tuples $(x_1, \dots, x_n)$ such that x_i are unique. Then S_n acts on $x = (1, \dots, n) \in X$ by permuting the elements. Notice that by acting on x with S_n we can get any element of X and that $|X| = n!$.
The group $(\mathbb{Z}, +)$ acts on the set $\mathbb{R}$ as $m \,.\, r = (-1)^m r$ for $m \in \mathbb{Z}$ and $r \in \mathbb{R}$.

■ **Example 3.1.5 — Representation** The group action of a group on a linear space is called a representation and will be the subject of study of much of the rest of this course.
The group action of a group on a nonlinear space is called a nonlinear representation, and is beyond the scope of this course.

■ **Application 3.1.6** Hilbert spaces describing wave functions are linear spaces. Particles correspond to representations of the Lorentz group, $\mathrm{O}(1, 3)$, which is the set of all Lorentz transformations with matrix multiplication as a group action, as well as various internal symmetry groups. That is group actions of the Lorentz group, on the Hilbert space of wave functions gives particles.

■ **Application 3.1.7 — Gauge Theory** In a gauge theory the gauge group acts on the gauge potential.
One family of groups that appears in this context are the **unitary groups**

$$\mathrm{U}(n) \coloneqq \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1}\}. \tag{3.1.8}$$

In particular, we often deal with

$$\mathrm{U}(1) = \{z \in \mathbb{C} \mid z^* z = zz^* = |z|^2 = 1\} = \{\mathrm{e}^{i\varphi} \mid \varphi \in \mathbb{R}\} \tag{3.1.9}$$

which is the group of complex numbers with unit modulus[a]. This particular group is the gauge group of (quantum) electrodynamics, it's action on the gauge potential is

$$\mathrm{e}^{i\varphi(x)} \,.\, A_\mu = A_\mu + \partial_\mu \varphi(x). \tag{3.1.10}$$

Recall that adding the derivative of a function to the electromagnetic potential doesn't change the electric or magnetic fields, $-\nabla A^0$ and $\nabla \times \boldsymbol{A}$.

R See the notes for quantum theory for more discussion of gauge invariance, and the notes from the particle physics part of relativity, nuclear, and particle physics for a discussion of QED.

[a]Strictly the group of complex numbers with unit modulus is the circle group, $\mathbb{T}$, and $\mathrm{U}(1)$ is the group of 1×1 unitary matrices over the complex numbers, however, under the obvious correspondence that $(z) \in \mathrm{U}(1)$ should correspond to $z \in \mathbb{T}$ these two groups are isomorphic, and therefore we don't distinguish between them.

3.2 Orbits and Stabilisers

Definition 3.2.1 — Orbit Given a group G which acts on the set X for each $x \in X$ we define the **orbit** of x to be the set, $g(x)$, containing all elements of X which are reached by acting on x with an element of G. That is

$$g(x) \coloneqq \{g \,.\, x \mid g \in G\} \subseteq X. \tag{3.2.2}$$

Notation 3.2.3 Sometimes the orbit is denoted $G(x)$ or using a notation similar to cosets $G \,.\, x$. Yet another notation is $\mathrm{Orb}_G(x)$, or $\mathrm{Orb}(x)$ when the group is clear.

■ **Example 3.2.4** Consider the group of rotations in the plane about some point $\boldsymbol{a}$. The orbit of some other point $\boldsymbol{b}$ is the circle of radius $|\boldsymbol{a} - \boldsymbol{b}|$ around the point $\boldsymbol{a}$.
Consider the group of integers, $\mathbb{Z}$, acting on $\mathbb{R}$ by $n \,.\, r = r + n$ for $n \in \mathbb{Z}$ and $r \in \mathbb{R}$. The orbit of $1 \in \mathbb{R}$ is the set $\mathbb{Z} \subset \mathbb{R}$.
Consider the group $\mathbb{R}^*$ acting on some vector space, V, by $r \,.\, \boldsymbol{v} = r\boldsymbol{v}$ for $r \in \mathbb{R}$ and $\boldsymbol{v} \in V$. Then the orbit of $\boldsymbol{v}$ is all vectors parallel to $\boldsymbol{v}$.

Definition 3.2.5 — Stabiliser Given a group G which acts on the set X for each $x \in X$ we define the **stabiliser** of x to be the subset, G_x, of G which leaves x invariant under the group action. That is

$$G_x \coloneqq \{g \in G \mid g \,.\, x = x\}. \tag{3.2.6}$$

Notation 3.2.7 Sometimes the stabiliser is denoted $\mathrm{Stab}_G(x)$, or $\mathrm{Stab}(x)$ where the group is clear.

■ Example 3.2.8 Consider the group S_4, which acts on the set of strings of four objects by permutation. The stabiliser of (a, b, c, c) is $\mathrm{Stab}_{S_4}((a, b, c, c)) = \{(), (3\,4)\}$.

■ Example 3.2.9 Consider the group $\mathrm{GL}(2, \mathbb{R})$ acting on $\mathbb{R}^2$ by matrix multiplication. Since

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \tag{3.2.10}$$

the stabiliser of $(1, 0)^\top$ is

$$\mathrm{Stab}_{\mathrm{GL}(2,\mathbb{R})}((1, 0)^\top) = \left\{ \begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix} \,\middle|\, b, d \in \mathbb{R} \text{ and } d \neq 0 \right\}. \tag{3.2.11}$$

Note that we require $d \neq 0$ so that $\det A = d \neq 0$ since $A \in \mathrm{GL}(2, \mathbb{R})$ must be invertible.

■ Example 3.2.12 Let G be a group which acts on itself by conjugation. Then the stabiliser of $a \in G$ is

$$\mathrm{Stab}_G(a) = \{g \in G \mid gag^{-1} = a\}, \tag{3.2.13}$$

which we can think of as the set of all elements, g, which commute with a, which follows by right multiplying the condition by g to get $gag^{-1}g = ga = ag$. This is sometimes called the **centraliser** of a, denoted $C_G(a)$. Compare this to the centre of G, $Z(G)$, which is the set of all commuting elements in G.

Lemma 3.2.14 If G is a group and acts on the set X then the stabiliser is a subgroup of G for any element of X.

Proof. Let $x \in X$. Suppose that $g, h \in \mathrm{Stab}(x)$, that is $g \,.\, x = h \,.\, x = x$. Then $(gh^{-1}) \,.\, x = g \,.\, (h^{-1} \,.\, x)$. Since $h^{-1}h = e$ we have that $(h^{-1}h) \,.\, x = e \,.\, x = x$, we also have that $(h^{-1}h) \,.\, x = h^{-1} \,.\, (h \,.\, x) = h^{-1} \,.\, x$ Hence $h^{-1} \,.\, x = x$. Using this we have $(gh^{-1}) \,.\, x = g \,.\, x = x$, and so $gh^{-1} \in \mathrm{Stab}(x)$. Thus, by the subgroup criterion (Theorem 1.2.32) $\mathrm{Stab}(x)$ is a subgroup of G. □

Theorem 3.2.15 — Orbit-Stabiliser Theorem. Let G be a group and $g \in G$. Define an action of G on some set X with $x \in X$. The map $\varphi g(x) \to G/\mathrm{Stab}(x)$, where $G/\mathrm{Stab}(x)$ is the set of all cosets of $\mathrm{Stab}(x)$, defined by $g \,.\, x \mapsto g\,\mathrm{Stab}(x)$ defines a bijection.

Proof. Since $\mathrm{Stab}(x)$ is a subgroup of G by Lemma 3.2.14 we can define cosets, gG_x, which partition G by Lemma 2.3.5. Hence, the map is surjective.
To demonstrate that this map is injective we need to prove that if $g_1\,\mathrm{Stab}(x) =$

$g_2 \operatorname{Stab}(x)$ then $g_1 . x = g_2 . x$. Notice that if $g_1 \operatorname{Stab}(x) = g_2 \operatorname{Stab}(x)$ then there exists some $g \in \operatorname{Stab}(x)$ such that $g_1 = g_2 g$ by the definition of a coset. We therefore have

$$g_1 . x = (g_2 g) . x = g_2 . g . x = g_2 . x \tag{3.2.16}$$

where in the last step we have used $g \in \operatorname{Stab}(x)$ and so by definition $g . x = x$. We have therefore demonstrated that this map is injective. Hence, the map is a bijection. □

Corollary 3.2.17 Given a finite group, G, which acts on a set X for all $x \in X$ we have

$$|G| = |\operatorname{Stab}(x)||\operatorname{Orb}(x)|. \tag{3.2.18}$$

Proof. By Lagrange's theorem (Theorem 2.3.8) $|G|/|G_x|$ is an integer. Further we can identify this as the number of sets in the partition of G by $\operatorname{Stab}(x)$, which has size $|G/\operatorname{Stab}(x)|$, which is exactly $|\operatorname{Orb}(x)|$, since by the orbit-stabiliser theorem (Theorem 3.2.15) the set of sets partitioning G by $\operatorname{Stab}(x)$, $G/\operatorname{Stab}(x)$ is in bijection with the set of orbits, $\operatorname{Orb}(x)$. □

Consider the limiting cases of this theorem. If $\operatorname{Stab}(x) = G$ then $|\operatorname{Orb}(x)| = 1$, since all elements of G leave x fixed. Clearly $|G| = |\operatorname{Stab}(x)|$, and so $|G| = |\operatorname{Stab}(x)||\operatorname{Orb}(x)|$ in accordance with the theorem.

If $\operatorname{Stab}(x) = \{e\}$ then $|\operatorname{Orb}(x)| = |G|$, since all elements of $G \setminus \{e\}$ must change x. Then $|G| = |\operatorname{Stab}(x)||\operatorname{Orb}(x)|$ in accordance with the theorem.

Theorem 3.2.19 — Cauchy's Theorem. Let G be a group and let p be prime. If $|G|/p$ is an integer then there exists an element of order p in G. That is there exists $g \in G$ with $g \neq e$ such that $g^p = e$ where e is the identity of G.

Proof. Let G be a group with $|G|/p$ an integer for some prime p. Define G^p to be the set of p-length strings of elements of G. That is $x \in G^p$ is of the form $(x_1, \dots, x_p)$ for $x_i \in G$. Define $X \subset G^p$ to be the set of elements $(x_1, \dots, x_p)$ such that $x_1 \cdots x_p = e$.

The size of X is $|X| = |G|^{p-1}$. This follows since for x_1 we can pick any element of G. For x_2 we can pick any element of G. So on until x_{p-1} for which we can pick any element of G. We then have to chose x_p such that $x_p = (x_1 \cdots x_{p-1})^{-1}$. We then have

$$x_1 \cdots x_p = (x_1 \cdots x_{p-1})x_p = (x_1 \cdots x_{p-1})(x_1 \cdots x_{p-1})^{-1} = e. \tag{3.2.20}$$

Hence, we make a choice from a set of size $|G|$ $p - 1$ times, and so we have $|G|^{p-1}$ possible elements in X.

Since $|X|$ is a multiple of $|G|$ and $|G|$ is divisible by p $|X|$ is also divisible by p.

Define the group action of $\mathbb{Z}_p$ on X by cyclic permutation. That is given

$m \in \mathbb{Z}_p$ and $x = (x_1, \ldots, x_p) \in X$ define

$$m \,.\, x = (x_{1+m}, \ldots, x_{p+m}) = (x_{m+1}, \ldots, x_p, x_1, \ldots, x_m) \tag{3.2.21}$$

where addition in the indices is done modulo m to so the indices remain in $\{1, \ldots, m\}$.
Clearly $|\mathbb{Z}_p| = p$. Then by Corollary 3.2.17 it follows that

$$p = |\mathbb{Z}_p| = |\mathrm{Stab}_{\mathbb{Z}_p}(x)||\mathrm{Orb}_{\mathbb{Z}_p}(x)|. \tag{3.2.22}$$

Since p is prime either $|\mathrm{Stab}_{\mathbb{Z}_p}(x)| = p$ and $|\mathrm{Orb}_{\mathbb{Z}_p}(x)| = 1$ or $|\mathrm{Stab}_{\mathbb{Z}_p}(x)| = 1$ and $|\mathrm{Orb}_{\mathbb{Z}_p}(x)| = p$.
If $|\mathrm{Orb}_{\mathbb{Z}_p}(x)| = 1$ then $x \in X$ must be of the form $(g, \ldots, g)$ for some $g \in G$ since if this wasn't the case then a cyclic permutation would not leave x invariant. An example of this case is $x = (e, \ldots, e)$. This cannot be the only example since then $|X|$ is not divisible by p, since we can write $|X|$ as $|X| = np + m1$, where n is the number of orbits of length p and m is the number of orbits of length 1. Clearly in order for $|X|$ to be divisible by p we need m to be divisible by p. Therefore, there must be some $g \in G$ such that $g \neq e$ but $(g, \ldots g) \in X$. By the definition of X this means that $g \cdots g = g^p = e$, and so g is some element of order p. □

■ **Example 3.2.23** S_3 is order 6 and hence has elements of order 2 and 3, for example $(1\,2)$ is of order 2 and $(1\,2\,3)$ is of order 3.
$\mathbb{Z}_6$ is of order 6 and hence has elements of order 2 and 3. Using addition modulo 6 as the group operation 3 is of order 2 and 2 is of order 3 in this group.

Corollary 3.2.24 If G is a group and $|G| = p$ for prime p then G is isomorphic to $\mathbb{Z}_p$.

Proof. By Cauchy's theorem (Theorem 3.2.19) G has an element of order p. That is there exists $g \in G$ such that $g^p = e$, further $g^m \neq e$ for $m < p$ since m doesn't divide p. Since there are p elements of G and each g^m must be distinct for $m < p$ it follows that all elements of G are of the form g^m for some $m \in [0, p) \cap \mathbb{Z}$. Therefore, G is cyclic and of order p and therefore G is isomorphic to $\mathbb{Z}_p$. □

Four

Normal Subgroups

4.1 Normal Subgroups

Definition 4.1.1 — Normal Subgroup Let G be a group and N be a subgroup of G. Then we say that N is a **normal subgroup** of G if N is invariant under the group action, $G \times N \longrightarrow N$, of conjugation. That is for all $n \in N$ and all $g \in G$ we have that $gng^{-1} \in N$. A normal subgroup is also referred to as an **invariant subgroup**.

Notation 4.1.2 If N is a normal subgroup of G we denote this $N \trianglelefteq G$. If N is a normal subgroup of G, and $N \neq G$, then we denote this $N \triangleleft G$.

We can think of normal subgroups as groups which are invariant under relabelling of the elements, this relabelling is done by conjugation. Compare this to matrix transformation. Suppose that M is a matrix in some basis, $\{\boldsymbol{e}_i\}$, and that $\{\boldsymbol{e}'_i\}$ is some other basis related to the first by $\boldsymbol{e}'_i = T_{ij}\boldsymbol{e}_j$. Then in this new basis M becomes TMT^{-1}. M describes the same transform in both bases but with different components.

■ **Example 4.1.3** For any group, G, both G and the trivial group, $\{e\}$, are normal subgroups.
If G is an Abelian group and H is a subgroup of G then H is a normal subgroup since $ghg^{-1} = gg^{-1}h = h \in H$ for all $g \in G$ and $h \in H$.

■ **Example 4.1.4** Recall that the quaternion group has the presentation

$$Q = \langle -e, i, j, k \mid (-e)^2 = e, i^2 = j^2 = k^2 = ijk = -e \rangle. \tag{4.1.5}$$

We can define $\mathbb{Z}_2$ to have the presentation $\mathbb{Z}_2 = \langle -e \mid (-e)^2 = e \rangle$, which has elements $\{e, -e\}$. This is a normal subgroup of Q. To see this note that

$i^{-1} = -i$, $j^{-1} = -j$, $k^{-1} = -k$, and $-e^{-1} = -e$ so

$$\begin{aligned} eee^{-1} &= ee(-e) &&= -e^3 = -e, && (4.1.6)\\ (-e)e(-e)^{-1} &= (-e)ee &&= -e^3 = -e, && (4.1.7)\\ e(-e)e^{-1} &= e(-e)(-e) = e^3 &&= e, && (4.1.8)\\ (-e)(-e)(-e)^{-1} &= (-e)(-e)e = e^3 &&= e, && (4.1.9)\\ iei^{-1} &= ie(-i) &&= -i^2 = e, && (4.1.10)\\ i(-e)i^{-1} &= i(-e)(-i) = i^2 &&= -e, && (4.1.11)\\ (-i)e(-i)^{-1} &= (-i)ei &&= -i^2 = e, && (4.1.12)\\ (-i)(-e)(-i)^{-1} &= (-i)(-e)i = i^2 &&= -e. && (4.1.13)\\ & && && (4.1.14) \end{aligned}$$

We also get the same results if we replace all is with j or k and so $\mathbb{Z}_2$ is invariant under conjugation and hence a normal subgroup of Q, $\mathbb{Z}_2 \lhd Q$.

In the previous example all elements of $\mathbb{Z}_2$ map to themselves under conjugation by elements of Q not in $\mathbb{Z}_2$. However, this need not be the case for a normal subgroup, as the next example shall show.

■ **Example 4.1.15** The permutations $A_3 = \{(), (1\,2\,3), (3\,2\,1)\}$ form a normal subgroup of S_3, this subgroup is called the alternating group, we will study it more in the future.
First we have to show that A_3 is a subgroup, which we can do via the subgroup criterion by considering $(1\,2\,3)(3\,2\,1)^{-1} = (1\,2\,3)(1\,2\,3) = (3\,2\,1) \in A_3$ and $(3\,2\,1)(1\,2\,3)^{-1} = (3\,2\,1)(3\,2\,1) = (1\,2\,3) \in A_3$, and clearly if we have eg^{-1} or ge^{-1} with $e = ()$ and $g = (1\,2\,3)$ or $g = (3\,2\,1)$ we will simply get g or g^{-1}, both of which are in A_3 since $(1\,2\,3)^{-1} = (3\,2\,1)$.

■ **Example 4.1.16** Invertible transformations of some vector space form a group, since the composition of two transformations is again a transformation, by definition for an *invertible* transformation the inverse exists, and the identity transformation is in this set. Given a basis for this vector space we can write these transformations as matrices with nonzero determinants, and in doing so define the **general linear group**:

$$\mathrm{GL}(n, \mathbb{F}) := \{A \in \mathcal{M}_n(\mathbb{F}) \mid \det A \neq 0\}. \tag{4.1.17}$$

Here $\mathbb{F}$ is a field, n is the dimension of the vector space, and $\mathcal{M}_n(\mathbb{F})$ are all square $n \times n$ matrices with entries from $\mathbb{F}$. Recall that $A \in \mathcal{M}_n(\mathbb{F})$ is invertible if and only if $\det A \neq 0$.
Another group we can define is the **special linear group**:

$$\mathrm{SL}(n, \mathbb{F}) := \{A \in \mathcal{M}_n(\mathbb{F}) \mid \det A = 1\} \subseteq \mathrm{GL}(n, \mathbb{F}). \tag{4.1.18}$$

$\mathrm{SL}(n, \mathbb{F})$ is a subgroup of $\mathrm{GL}(n, \mathbb{F})$ since if $A, B \in \mathrm{SL}(n, \mathbb{F})$ then $\det(AB^{-1}) = \det(A)\det(B^{-1}) = \det(A)/\det(B) = 1/1 = 1$, so $AB^{-1} \in \mathrm{SL}(n, \mathbb{F})$.

Further, $\mathrm{SL}(n, \mathbb{F})$ is a normal subgroup of $\mathrm{GL}(n, \mathbb{F})$ since if $A \in \mathrm{SL}(n, \mathbb{F})$ and $B \in \mathrm{GL}(n, \mathbb{F})$ we have

$$\det(BAB^{-1}) = \det(B)\det(A)\det(B^{-1}) = \det(B)\det(A)\frac{1}{\det(B)} = \det(A) = 1$$

so BAB^{-1} is again in $\mathrm{SL}(n, \mathbb{F})$.

There is an easier way to test if a subgroup is normal, and it is to consider the cosets.

Theorem 4.1.19. Let H be a subgroup of some group, G. Then H is a normal subgroup of G if and only if the left and right cosets are equal. That is $gH = Hg$ for all $g \in G$.

Proof. First, suppose that H is a normal subgroup. Consider some $g \in G$ and $h \in H$. Then $ghg^{-1} \in H$, call $h' = ghg^{-1}$. Similarly, since $g^{-1} \in G$ we have $g^{-1}hg = g^{-1}h(g^{-1})^{-1} \in H$, call $h'' = g^{-1}hg$. Then $gh = ghg^{-1}g = h'g \in Hg$ and $hg = gg^{-1}hg = gh'' \in gH$. Since $gh \in Hg$ for all $h \in H$ it follows that $gH \subseteq Hg$. Since $hg \in gH$ for all $h \in H$ it follows that $Hg \subseteq gH$. Hence, $gH = Hg$.
Now, suppose that $gH = Hg$ for all $g \in G$. Then for some $g \in G$ and $h \in H$ we have $gh \in gH = Hg$, which means there exists some $h' \in H$ such that $gh = h'g$. Left multiplying by g^{-1} we have $ghg^{-1} = h' \in H$, and so H is invariant under conjugation and hence normal in G. □

4.2 Coset Groups

Definition 4.2.1 — Coset Product Let G be a group and N be a normal subgroup of G. Then we can define G/N to be the set of cosets of N, that is

$$G/N := \{gN \mid g \in N\} = \{Ng \mid g \in G\}. \tag{4.2.2}$$

We can define a binary operation on G/N according to

$$(gN)(g'N) := (gg')N. \tag{4.2.3}$$

With this operation G/N forms a group which we call the **quotient group** or **factor group**.

Theorem 4.2.4. If G is a group and N is a normal subgroup of G then G/N with the operation defined above is a group.

Proof. We must first show that the operation is well-defined. That is that no matter which element of gN we choose to represent gN we get the same

result. Suppose $gN = g'N$ and $\tilde{g}N = \tilde{g}'N$, the operation is well-defined only if $(g\tilde{g})N = (g'\tilde{g}')N$. For this to be the case it is sufficient that $g'\tilde{g}' \in (g\tilde{g})N$. Since $g'N = gN$ we know that $g' = gn$ for some $n \in N$ and since $\tilde{g}'N = \tilde{g}N$ we know that $\tilde{g}' = \tilde{g}\tilde{n}$ for some $\tilde{n} \in N$. We then have $g'\tilde{g}' = gn\tilde{g}\tilde{n}$. We need to show that this is of the form $g\tilde{g}n''$ for some $n'' \in N$.
Now, by definition since N is normal we have $\tilde{g}n'\tilde{g}^{-1} = n$ for some $n' \in N$. From this it follows that $\tilde{g}n' = n\tilde{g}$ We then have

$$\begin{aligned} g'\tilde{g}' &= (gn)(\tilde{g}\tilde{n}) && (4.2.5)\\ &= g(n\tilde{g})\tilde{n} && (4.2.6)\\ &= g(\tilde{g}n')\tilde{n} && (4.2.7)\\ &= (g\tilde{g})(n'\tilde{n}) && (4.2.8)\\ &= (g\tilde{g})n'' && (4.2.9) \end{aligned}$$

where $n'' = n'\tilde{n} \in N$. Hence, this operation is well-defined.
It remains to show that G/N is a group. To do so notice that $eN = N$ acts as an identity: $(eN)(gN) = (eg)N = gN$ for all $g \in G$, and that $g^{-1}N$ is the inverse of gN, since $(gN)(g^{-1}N) = (gg^{-1})N = eN = N$. Associativity follows from associativity of the group product, that is if $g_1, g_2, g_3 \in G$ then

$$\begin{aligned} [(g_1N)(g_2N)](g_3N) &= [(g_1g_2)N](g_3N) && (4.2.10)\\ &= ((g_1g_2)g_3)N && (4.2.11)\\ &= (g_1(g_2g_3))N && (4.2.12)\\ &= (g_1N)[(g_2g_3)N] && (4.2.13)\\ &= (g_1N)[(g_2N)(g_3N)]. && (4.2.14) \end{aligned}$$

Hence, G/N is a group. □

■ **Example 4.2.15** Recall from Example 4.1.4 that the quaternion group, Q, has as a normal subgroup $\mathbb{Z}_2$.
The quotient group in this case is $Q/\mathbb{Z}_2$. To see this we first write out all cosets:

$$\mathbb{Z}_2 = \{e, -e\}, \quad i\mathbb{Z}_2 = \{i, -i\}, \quad j\mathbb{Z}_2 = \{j, -j\}, \quad \text{and} \quad k\mathbb{Z}_2 = \{k, -k\}.$$

We can then construct the group table:

$Q/\mathbb{Z}_2$	$\mathbb{Z}_2$	$i\mathbb{Z}_2$	$j\mathbb{Z}_2$	$k\mathbb{Z}_2$
$\mathbb{Z}_2$	$\mathbb{Z}_2$	$i\mathbb{Z}_2$	$j\mathbb{Z}_2$	$k\mathbb{Z}_2$
$i\mathbb{Z}_2$	$i\mathbb{Z}_2$	$\mathbb{Z}_2$	$k\mathbb{Z}_2$	$j\mathbb{Z}_2$
$j\mathbb{Z}_2$	$j\mathbb{Z}_2$	$k\mathbb{Z}_2$	$\mathbb{Z}_2$	$i\mathbb{Z}_2$
$k\mathbb{Z}_2$	$k\mathbb{Z}_2$	$j\mathbb{Z}_2$	$i\mathbb{Z}_2$	$\mathbb{Z}_2$

(4.2.16)

Here we have used $ij = k$, $jk = i$, and $ki = j$, $i^2 = j^2 = k^2 = -e$, and that permuting any two of i, j, and k adds a negative sign.
Comparing this Cayley table with those in Equation (2.1.18) we see that $Q/\mathbb{Z}_2 \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. In particular, we have the correspondence $\mathbb{Z}_2 \leftrightarrow e$, $i\mathbb{Z}_2 \leftrightarrow a$, $j\mathbb{Z}_2 \leftrightarrow b$, and $k\mathbb{Z}_2 \leftrightarrow c$.

▪ **Example 4.2.17** Let $n\mathbb{Z}$ denote the integer multiples of n, viewed as a subgroup of $\mathbb{Z}$. Then $n\mathbb{Z}$ is a normal subgroup of $\mathbb{Z}$, since $\mathbb{Z}$ is Abelian and so left and right cosets are equal. The cosets are of the form $m + n\mathbb{Z} = \{m+nk \mid k \in \mathbb{Z}\}$. There is then a natural isomorphism $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}_n$, namely $m + n\mathbb{Z} \mapsto \mathrm{e}^{2i\pi m/n}$. For this reason a lot of people denote the group we have been calling $\mathbb{Z}_n$ by $\mathbb{Z}/n\mathbb{Z}$.

Definition 4.2.18 — Simple Group If a group has no non-trivial proper normal subgroups then we say that it is a **simple group**.

Thinking of the quotient groups, G/N, as division of groups, as the notation suggests, it makes sense to view simple groups as the "primes" of groups, in that they can't be further divided.

Finite simple groups are classified up to automorphism. That is given any finite simple group it will be isomorphic to a group from a known set of groups. This classification of finite simple groups is not simple. The proof consists of tens of thousands of pages across several hundred journal articles written by approximately 100 authors over a period of more than 50 years, finally being completed in 2008.

Theorem 4.2.19. If G is a group and H is a subgroup of G and $[G : H] = 2$ then H is a normal subgroup of G.

Proof. Since $[G : H] = |G|/|H| = 2$ there are two cosets, namely H and gH for $g \in G \setminus H$. Considering right cosets instead we have H and Hg, again for $g \in G \setminus H$. By Lemma 2.3.5 we also know that the cosets partition G, so that $G = H \cup gH = H \cup Hg$. Since H is the same in both of these and cosets are disjoint we must have that $gH = Hg$ and so H is a normal subgroup of G by Theorem 4.1.19. □

Note that in this case $|G/H| = |G|/|H| = [G : H] = 2$ and so $G/H \cong \mathbb{Z}_2$.

As the name of the next theorem suggests there are multiple "isomorphism theorems", but we will discuss only on the first.

Theorem 4.2.20 — First Isomorphism Theorem. Let G and H be groups and $\varphi\colon G \to H$ a homomorphism. Then

1. The image of φ, $\operatorname{Im}\varphi$, is a subgroup of H.
2. The kernel of φ, $\ker\varphi$, is a normal subgroup of G.
3. $G/\ker\varphi \cong \operatorname{Im}\varphi$.

Proof. The first statement is simply Lemma 2.1.21.
Recall that $\ker\varphi = \{k \in G \mid \varphi(k) = e\}$ where e is the identity of H. For the proof of the second statement consider $k_1, k_2 \in \ker\varphi$. Then

$$\varphi(k_1k_2^{-1}) = \varphi(k_1)\varphi(k_2^{-1}) = \varphi(k_1)\varphi(k_2)^{-1} = ee = e \qquad (4.2.21)$$

so $k_1k_2^{-1} \in \ker\varphi$ and so by the subgroup criterion (Theorem 1.2.32) $\ker\varphi$ is a

subgroup of G. Here we used Lemma 2.1.20 to allow us to identify $\varphi(k_2^{-1}) = \varphi(k_2)^{-1}$.
It remains to show that $\ker\varphi$ is a *normal* subgroup of G. Take $k \in \ker\varphi$ and $g \in G$. Then

$$\begin{aligned}\varphi(gkg^{-1}) &= \varphi(g)\varphi(k)\varphi(g^{-1}) = \varphi(g)e\varphi(g^{-1}) \\ &= \varphi(g)\varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(e) = e \end{aligned} \tag{4.2.22}$$

where we have used Lemma 2.1.19 to identify $\varphi(e) = e$. This shows that $gkg^{-1} \in \ker\varphi$ and so $\ker\varphi$ is a normal subgroup of G. This proves the second point.
For the third point we need to show that there exists an isomorphism $\psi\colon G/\ker\varphi \longrightarrow \operatorname{Im}\varphi$. To do so we consider the obvious choice that $\psi(gK) = \varphi(g)$, where $K = \ker\varphi$. Since ψ is defined on representatives of the cosets we need to show that it is well-defined. Let $g_1, g_2 \in G$ be in the same coset, that is $g_1K = g_2K$. It follows that $(g_1^{-1}K)(g_2K) = (g_1g_2)K$ but also $(g_1^{-1}K)(g_2K) = (g_1^{-1}K)(g_1K) = (g_1^{-1}g_1)K = K$ and so $g_1^{-1}g_2 \in K$. Thus, $\varphi(g_1^{-1}g_2) = e$, since $K = \ker\varphi$. We then have $\varphi(g_1^{-1}g_2) = \varphi(g_1^{-1})\varphi(g_2) = \varphi(g_1)^{-1}\varphi(g_2)$ having used the definition of the homomorphism and Lemma 2.1.20. We then have $\varphi(g_1)^{-1}\varphi(g_2) = e$, and so $\varphi(g_1) = \varphi(g_2)$. This shows that ψ is well-defined.
Next, we verify that ψ is a homomorphism. To do so let $g_1, g_2 \in G$ and so $g_1K, g_2K \in G/K$. We then have

$$\psi((g_1K)(g_2K)) = \psi((g_1g_2)K) \tag{4.2.23}$$
$$= \varphi(g_1g_2) \tag{4.2.24}$$
$$= \varphi(g_1)\varphi(g_2) \tag{4.2.25}$$
$$= \psi(g_1K)\psi(g_2K), \tag{4.2.26}$$

so ψ is a homomorphism.
Finally, we show that ψ is an isomorphism, that is that it is bijective. The kernel of ψ consists of all cosets $gK \in G/K$ such that $\varphi(g) = e$, but these are exactly the elements $g \in G$ such that $g \in K = \ker\varphi$. Hence, the kernel of ψ is the trivial group $\{K\} \subset G/K$. This proves that ψ is injective by Lemma 2.1.23.
Finally, let $h \in \operatorname{Im}\varphi$. Then there exists $g \in G$ such that $\varphi(g) = h$. We then have $\psi(gK) = \varphi(g) = h$, and so ψ is surjective.
So ψ is a bijective homomorphism and hence an isomorphism. □

An alternative statement of the first isomorphism theorem is that the following diagram commutes:

$$\begin{tikzcd} G \arrow[r, "\varphi"] \arrow[d, "\pi"] & \operatorname{Im}\varphi \subseteq H \\ G/\ker\varphi. \arrow[ur, "\sim", "\psi"'] & \end{tikzcd} \tag{4.2.27}$$

Here $\pi\colon G \longrightarrow G/\ker\varphi$ is the **natural projection** defined by $\pi(g) = g\ker\varphi$. We use the notation $X \xrightarrow{\sim} Y$ to denote an isomorphism from X to Y.

One of the main uses of the first isomorphism theorem is to quickly check if a subgroup is normal. This works since if H is a subgroup of some group G and there exists some homomorphism $\varphi\colon G \to \tilde{G}$ for some group $\tilde{G}$ such that $H = \ker\varphi$ then by the first isomorphism theorem $\ker\varphi$, and hence H, is normal in G.

Five

Products of Groups

5.1 Direct Products

Definition 5.1.1 — Direct Product Let H and J be groups. Then we define a new group, $H \times J$, called the **direct product** of H and J such that

$$H \times J := \{(h, j) \mid h \in H \text{ and } j \in J\}. \tag{5.1.2}$$

We extend the products of the two groups to define a product

$$(h, j)(h', j') = (hh', jj') \tag{5.1.3}$$

for $h, h' \in H$ and $j, j' \in J$.

We can think of the direct product as extending the Cartesian product to groups. Notice that for finite groups $|H \times J| = |H||J|$.

Theorem 5.1.4. The direct product of two groups is a group.

Proof. Let H and J be groups. Consider $h_1, h_2, h_3 \in H$ and $j_1, j_2, j_3 \in J$. Then

$$\begin{aligned}
(h_1, j_1)[(h_2, j_2)(h_3, j_3)] &= (h_1, j_1)(h_2h_3, j_2j_3) && (5.1.5)\\
&= (h_1(h_2h_3), j_1(j_2j_3)) && (5.1.6)\\
&= ((h_1h_2)h_3, (j_1j_2)j_3) && (5.1.7)\\
&= (h_1h_2, j_1j_2)(h_3, j_3) && (5.1.8)\\
&= [(h_1, j_1)(h_2, j_2)](h_3, j_3) && (5.1.9)
\end{aligned}$$

so associativity in $H \times J$ follows from associativity in H and J.
Let e_H be the identity of H and e_J the identity of J. Then

$$(h, j)(e_H, e_J) = (he_H, je_J) = (h, j) \tag{5.1.10}$$

for all $h \in H$ and $j \in J$ and so $e_{H\times J} = (e_H, e_J)$ is the identity of $H \times J$.
Finally, notice that for all $h \in H$ and $j \in J$ we have

$$(h, j)(h^{-1}, j^{-1}) = (hh^{-1}, jj^{-1}) = (e_H, e_J) = e_{H\times J} \tag{5.1.11}$$

and since $h^{-1} \in H$ and $j^{-1} \in J$ we have that $(h^{-1}, j^{-1}) \in H \times J$ acts as the inverse for $(h, j) \in H \times J$.
Hence, $H \times J$ is a group. □

■ **Example 5.1.12 — Klein *Vierergruppe*** The **Klein *Vierergruppe*** is most simply defined as the direct product group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Note that

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{(1,1), (1,-1), (-1,1), (-1,-1)\} \tag{5.1.13}$$

and all elements square to $(1,1)$, which is the identity, for example, $(-1,1)^2 = ((-1)^2, 1^2) = (1,1)$. Hence, $\mathbb{Z}_2 \times \mathbb{Z}_2$ matches our earlier definition of the Klein *Vierergruppe*, namely the unique group of order 4 such that all elements square to the identity.

■ **Example 5.1.14** Let $\mathbb{R}$ be the group of real numbers under addition. Then $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ is the group of two-component vectors, (x, y), with addition defined by

$$(x, y) + (x', y') = (x + x', y + y'). \tag{5.1.15}$$

The direct product is both commutative and associative up to isomorphism. That is $G \times H \cong H \times G$ and $G \times (H \times J) \cong (G \times H) \times J$, with the obvious isomorphisms $(g, h) \mapsto (h, g)$ and $(g, (h, j)) \mapsto ((g, h), j)$, respectively. The latter means that we can define the direct product of multiple groups in a sensible way, so we typically write $G \times H \times J$ and (g, h, j) and so on.

The order of $(h, j) \in H \times J$ is the lowest common multiple of the orders of h and j. In particular if the orders of h and j are relatively prime then the order of (h, j) is the product of the orders of h and j. This means that if H and J are cyclic groups of relatively prime orders m and n their direct product is again cyclic:

$$\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{mn}. \tag{5.1.16}$$

Given two groups, G and H, it is common to view G and H as subgroups of $G \times H$, which can be done since we can identify $G \cong \{e\} \times H$ and $H \cong G \times \{e\}$ under the obvious isomorphisms $g \mapsto (g, e)$ and $h \mapsto (e, h)$, respectively. Making this identification both G and H are normal subgroups of $G \times H$. We can show this easily for G, and it can be shown similarly for H:

$$(g, h)(g', e)(g, h)^{-1} = (g, h)(g', e)(g^{-1}, h^{-1}) = (gg'g^{-1}, heh^{-1}) = (gg'^{-1}, e) \in G \times \{e\} \tag{5.1.17}$$

since $gg'g^{-1} \in G$ as G is a group.

5.2 Semidirect Product

Definition 5.2.1 — Semidirect Product Let H and J be groups and $\varphi\colon H \times J \to J$ a group action of H on J. Then we can define a new group, $H \ltimes J$, as the set

$$\{(h, j) \mid h \in H \text{ and } j \in J\} \tag{5.2.2}$$

and the group product

$$(h, j)(h', j') = (hh', j\varphi(h, j')) = (hh', j(h \,.\, j')) \tag{5.2.3}$$

for $h, h' \in H$ and $j, j' \in J$.

R The vertical line in the symbol $\ltimes$ goes with the group which acts on the other group.

For each group action we technically have two semidirect product groups, $H \ltimes J$ and $J \rtimes H$, but these differ only in the order of elements, (h, j) vs. (j, h), and are isomorphic. The order of $H \ltimes J$ is $|H \ltimes J| = |H||J|$.

■ **Example 5.2.4 — Isometries of Euclidean Space** The **isometries of Euclidean space** is the group which preserves Euclidean distance. This group is typically denoted $\mathrm{ISO}(n)$ or $\mathrm{E}(n)$, where n is the dimension of the space on which the group acts. This group consists of rotations, reflections and translations. Transformations in the first two categories, rotations and reflections, form $\mathrm{O}(n)$, that is the group of transformations which preserve Euclidean distance and leave the origin invariant. Translations can be viewed as $\mathbb{R}^n$.
A point in Euclidean space can be viewed as $\boldsymbol{x} \in \mathbb{R}^n$, which is somewhat confusing since $\mathbb{R}^n$ appears twice, as both the translations and the Euclidean space upon which they act.
We can write $\mathrm{ISO}(n)$ as a semidirect product, $\mathrm{ISO}(n) = \mathrm{O}(n) \ltimes \mathbb{R}^n$. We just need to work out the correct group action. First we need to state how $\mathrm{ISO}(n)$ acts on $\mathbb{R}^n$ (as Euclidean space). This is simple, given a rotation and/or reflection $R \in \mathrm{O}(n)$ and translation $\boldsymbol{a} \in \mathbb{R}^n$ this acts on $\boldsymbol{x} \in \mathbb{R}^n$ as $(R, \boldsymbol{a}).\boldsymbol{x} = R\boldsymbol{x}+\boldsymbol{a}$, that is we rotate $\boldsymbol{x}$ with R and then translate by $\boldsymbol{a}$, notice that the order is important, which we will see means that $\mathrm{ISO}(n)$ is *not* a direct product of $\mathrm{O}(n)$ and $\mathbb{R}^n$.
Consider what happens when we act on some $\boldsymbol{x} \in \mathbb{R}^n$ (as a point in Euclidean space) by two isometries, $(R, \boldsymbol{a})$ and $(R', \boldsymbol{a}')$, where $R, R' \in \mathrm{O}(n)$ and $\boldsymbol{a}, \boldsymbol{a}' \in \mathbb{R}^n$ (as the group of translations). We then have

$$\begin{aligned}
(R', \boldsymbol{a}')(R, \boldsymbol{a}) \,.\, \boldsymbol{x} &= (R', \boldsymbol{a}') \,.\, (R, \boldsymbol{a}) \,.\, \boldsymbol{x} && (5.2.5)\\
&= (R', \boldsymbol{a}) \,.\, (R\boldsymbol{x} + \boldsymbol{a}) && (5.2.6)\\
&= R'(R\boldsymbol{x} + \boldsymbol{a}) + \boldsymbol{a}' && (5.2.7)\\
&= R'R\boldsymbol{x} + R'\boldsymbol{a} + \boldsymbol{a}' && (5.2.8)\\
&= (R'R, R'\boldsymbol{a} + \boldsymbol{a}') \,.\, \boldsymbol{x}. && (5.2.9)
\end{aligned}$$

So we identify the group action associated with the semidirect product as

$$R' \,.\, \boldsymbol{a} = R'\boldsymbol{a}, \qquad \text{or} \qquad \varphi(R', \boldsymbol{a}) = R\boldsymbol{a}, \tag{5.2.10}$$

which is probably what most people would expect, the rotation (or reflection) acts by rotating (or reflecting).

We can further generalise the isometries of Euclidean space by dropping the requirement that lengths be preserved and allowing uniform scaling. In this case the group of symmetries is the affine group, $\mathrm{Aff}(V)$, which is the semidirect product $\mathrm{GL}(V) \ltimes V$ where $\mathrm{GL}(V)$ acts on the vector space V with the expected action, $M \,.\, \boldsymbol{v} = M\boldsymbol{v}$ for $M \in \mathrm{GL}(V)$ and $\boldsymbol{v} \in V$. This contains $\mathrm{ISO}(V)$ as a subgroup.

■ **Example 5.2.11 — Dihedral Group** The **dihedral group** of order $2n$ can be defined abstractly as

$$D_n \coloneqq \langle r, s \mid r^n = s^2 = e, s^{-1}rs = r^{-1} \rangle. \tag{5.2.12}$$

This can be identified as the group of symmetries of the regular n-gon. Here r is identified as a rotation by $2\pi/n$ and s as a mirror symmetry or inversion in a perpendicular bisector of one of the sides.

! Some sources denote the dihedral group of order $2n$ by D_{2n}, since it has $2n$ elements, whereas we denote it D_n, as it is the group of symmetries of the regular n-gon.

We can identify a subgroup generated by s as $\mathbb{Z}_2 = \{e, s\}$. We can identify a subgroup generated by r as $\mathbb{Z}_n = \{e, r, \dots, r^{n-1}\}$. It turns out that D_n can then be written as the semidirect product $D_n \cong \mathbb{Z}_2 \ltimes \mathbb{Z}_n$.
To see this it is best to just consider a few examples. First we identify $\mathbb{Z}_2 = \{\pm 1\}$, and $\mathbb{Z}_n = \{\mathrm{e}^{2i\pi m/n}\}$. A few examples of products in D_n are then

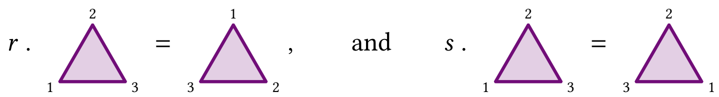
$$(+1, \mathrm{e}^{2i\pi m/n})(+1, \mathrm{e}^{2i\pi m'/n}) = (+1, \mathrm{e}^{2i\pi(m+m')/n}), \tag{5.2.13}$$

$$(-1, \mathrm{e}^{2i\pi m/m})(+1, \mathrm{e}^{2i\pi m'/n}) = (-1, \mathrm{e}^{2i\pi(m-m')/n}), \tag{5.2.14}$$

$$(+1, \mathrm{e}^{2i\pi m/n})(+1, \mathrm{e}^{2i\pi m'/n}) = (-1, \mathrm{e}^{2i\pi(m+m')/n}). \tag{5.2.15}$$

That is the group action associated with the semidirect product is $\pm 1 \,.\, \mathrm{e}^{2i\pi m/n} = \mathrm{e}^{\pm 2i\pi m/n}$. We can further identify this as $1 \,.\, z = z$ and $-1 \,.\, z = z^*$.

It is worth examining the dihedral group more, particularly as it comes up in geometry and chemistry. Starting with geometry we claimed that D_3 is the group of symmetries of an equilateral triangle. By this we mean that D_3 acts on an equilateral triangle such that there is no noticeable change. However, in order to keep track of what is happening we label the corners of the triangle, but these labels have no meaning besides keeping track of how D_3 is acting. Taking r to be a clockwise rotation by 120° and s to be a reflection in the vertical. Graphically the action of r and s on the triangle is



(5.2.16)

We should check that these symmetries really correspond to D_3 as defined by the presentation above. First, notice that three rotations correspond to a rotation by $3 \cdot 120° = 360°$, which is the same as no rotation at all. Second, notice that repeating

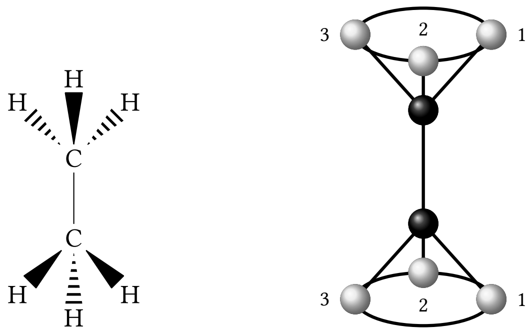
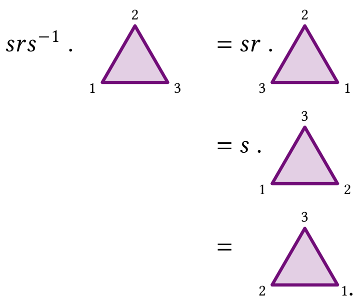


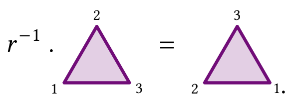
Figure 5.1: Ethane molecule.

a mirror symmetry undoes it so $s^2 = e$. Finally, consider the action of srs^{-1} on the triangle, noticing that $s^{-1} = s$, we have



(5.2.17)

Now compare this to the action of r^{-1}, which is an anticlockwise rotation by 120°:



(5.2.18)

Noticing that these are the same we see that D_3 does truly describe the symmetries of the equilateral triangle.

As well as the equilateral triangle D_3 is also the symmetry group of an ideal C_2H_6 molecule. This molecule is shown in Figure 5.1. Here we identify r as a rotation around the carbon-carbon bond and s as a rotation by 180° about the perpendicular bisector to this bond in the page, such that the numbers on the upper and lower hydrogens match. Shown here is staggered ethane, which is such that viewed end on the hydrogens don't line up. D_3 is also the symmetry group of eclipsed ethane, where, when viewed end on, the hydrogens line up, we just change s to be inversion about the centre of the carbon-carbon bond.

Six

Permutation Groups

6.1 Symmetric Group

Definition 6.1.1 — Permutation A **permutation**, σ, on n objects is a bijection $\sigma\colon X \to X$ where $|X| = n$.

Typically, we identify X as $\{1, \dots, n\}$.

Definition 6.1.2 — Symmetric Permutation Group The **symmetric group** on n objects is the set of all permutations on n objects with function composition as a group operation. This group is denoted S_n.

The order of S_n is $n!$, since we can choose to permute the first element to any of n possible options, the second to any of $n-1$ options, and so on giving $n(n-1)\cdots 1 = n!$ choices.

Lemma 6.1.3 The symmetric group on n objects is a group.

Proof. Let $\sigma, \rho \in S_n$. Then both of these are bijections on some set X with $|X| = n$. Their composition is defined by $(\sigma \circ \rho)(x) = \sigma(\rho(x))$ for all $x \in X$. Straight away we see that this is indeed a permutation, since $\sigma \circ \rho\colon X \to X$ and the inverse is $\rho^{-1} \circ \sigma^{-1}$, the existence of said inverses in S_n is guaranteed as $\sigma \in S_n$ is a bijection and so its inverse exists and is also a bijection. This shows that S_n is closed.
The identity function, $\mathrm{id}_X\colon X \to X$ defined by $\mathrm{id}_X(x) = x$ for all $x \in X$ is a permutation and $\mathrm{id}_X \circ \sigma = \sigma$ for all $\sigma \in S_n$. As previously discussed σ has the inverse σ^{-1}, which is such that $\sigma \circ \sigma^{-1} = \mathrm{id}_X$. Hence, S_n is a group. □

S_n acts on the set of all tuples $(x_1, \dots, x_n)$ where $x_i \in X$ are distinct in the obvious way, namely by permuting the elements:

$$\sigma \,.\, (x_1, \dots, x_n) = (\sigma(x_1), \dots, \sigma(x_n)). \tag{6.1.4}$$

Definition 6.1.5 — Cycle A k-**cycle** is a way of writing a certain permutation. Namely, $(a_1 \ \dots \ a_k)$ with $a_i \in X$ is the permutation that sends a_1 to a_2, a_2 to a_3, and so on until a_{k-1} to a_k and a_k to a_1. All $x \in X$ such that $x \neq a_i$

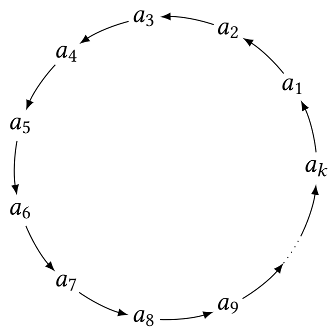


Figure 6.1: The k-cycle $(a_1\ \dots\ a_k)$.

are left unchanged.
A 2-cycle is also called a **transposition**.
The identity is usually written as () when using cycle notation, although we could also write it as a 1-cycle, (a) for any $a \in X$.
Given a k-cycle $(a_1\ \dots\ a_k)$ we can start on any element of this cycle, so this is equivalent to $(a_m\ \dots\ a_k\, a_1\ \dots\ a_{m-1})$.

■ **Example 6.1.6** Consider S_4, this contains the 3-cycles $(1\,4\,2)$ and $(1\,2\,3)$. We can work out their product by considering their action on some 4-tuple (a, b, c, d):

$$(1\,4\,2)(1\,2\,3)\,.\,(a\,b\,c\,d) = (1\,4\,2)\,.\,(c, a, b, d) \tag{6.1.7}$$
$$= (a, d, b, c) \tag{6.1.8}$$
$$= (2\,3\,4)\,.\,(a, b, c, d), \tag{6.1.9}$$

hence, we have $(1\,4\,2)(1\,2\,3) = (2\,3\,4)$.

Definition 6.1.10 — Disjoint Cycles Two cycles are disjoint if no element of X appears in both cycles.

Lemma 6.1.11 All permutations can be written as a product of disjoint cycles.

Proof. We proceed by induction on the size of $n = |X|$. Clearly if $n = 1$ then the only permutation is the identity, ().
Let $\sigma \in S_n$ and suppose that all cycles in S_{n-1} can be written as disjoint cycles. For simplicity, we will take $X = \{1, \dots, n\}$. If $\sigma(n) = n$ then we can consider σ as a permutation on $\{1, \dots, n-1\}$ leaving n fixed, and we are done since this can be written as a product of disjoint cycles. If $\sigma(n) = k \neq n$ then consider the permutation $\rho = (n\,k)\sigma$. We have that $\rho(n) = (n\,k)\sigma(n) = (n\,k)(k) = n$, here we are treating $(n\,k)$ as a function, $(n\,k)\colon \{1, \dots, n\} \to \{1, \dots, n\}$, defined by $(n\,k)(n) = k$, $(n\,k)(k) = n$ and $(n\,k)(a) = a$ for $a \neq n, k$. So

we can think of ρ as being a permutation on $\{1, \dots, n-1\}$, and hence can be written as a product of disjoint cycles, $\rho = \tau_1 \cdots \tau_r$. The cycles τ_i only contain the numbers $1, \dots, n-1$, and each appears in at most one of these cycles.

Clearly $(n\,k)(n\,k) = ()$, and so it follows that $\sigma = (n\,k)(n\,k)\sigma = (n\,k)\tau_1 \cdots \tau_r$. If k doesn't appear in any of the cycles τ_i then we are done as this is a product of disjoint cycles. Disjoint cycles commute, since by being disjoint they act on different elements of the tuple $(x_1, \dots, x_n)$, and so don't interact, meaning the order doesn't matter. Using this we are free to assume that the cycle in which k appears, if it appears, is τ_1.

We are free to start on any element of the cycle, so we write $\tau_1 = (k\,a_1\,\dots\,a_m)$. We then have

$$(n\,k)\tau_1 = (n\,k)(k\,a_1\,\dots\,a_m) = (n\,k\,a_1\,\dots\,a_m). \tag{6.1.12}$$

This follows by considering $(n\,k)\tau_1(k) = (n\,k)(a_1) = a_1$, $(n\,k)\tau_1(n) = (n\,k)(n) = k$, $(n\,k)\tau_1(a_m) = (n\,k)(k) = n$, and $(n\,k)\tau_1(a_i) = (n\,k)a_{i+1} = a_{i+1}$ for $i \neq m$. It follows then that we can write

$$\sigma = (n\,k\,a_1\,\dots\,a_m)\tau_2 \cdots \tau_r \tag{6.1.13}$$

which is a product of disjoint cycles.

Hence, by induction we can write any permutation in S_n as a product of disjoint cycles for all $n \in \mathbb{N}$. □

One question that we may reasonably ask is how many m-cycles are there in S_n for some fixed $m \in \{1, \dots, n\}$. If the order of a cycle didn't matter then there would be $\binom{n}{m}$ m-cycles in S_n. However, the order does matter. Suppose we have chosen our m terms to appear in the cycle. We can start with any of them, reducing the number of choices that give distinct cycles by a factor of $1/m$. There are then $(m-1)!$ choices for ordering the $m-1$ elements remaining, giving the number of m-cycles to be

$$\binom{n}{m}\frac{1}{m}(m-1)! = \frac{n!}{(n-m)!m!}\frac{1}{m}(m-1)! = \frac{n!}{(n-m)!} \tag{6.1.14}$$

where we have used $m(m-1)! = m!$.

Theorem 6.1.15. The transpositions generate S_n. That is, all permutations can be written as a product of 2-cycles.

Proof. All elements of S_n are simply k-cycles for some $k \in \{1, \dots, n\}$. An arbitrary k-cycle can be written as

$$(a_1\,a_2\,\dots\,a_k) = (a_1\,a_k)(a_1\,a_{k-1}) \cdots (a_1\,a_2). \tag{6.1.16}$$

To see why this works we consider three cases. First, if this acts on some a_i with $i \neq 1$, on the left-hand side we clearly see that a_i maps to a_{i+1}. On the right-hand side a_i commutes with cycles until a cycle with a_i occurs, this cycle will be $(a_1\,a_i)$, and so a_i will be sent to a_1. The next cycle is then $(a_1\,a_{i+1})$, and hence a_1 maps to a_{i+1} which then commutes with all of the remaining cycles. Hence, a_i will map to a_{i+1}.

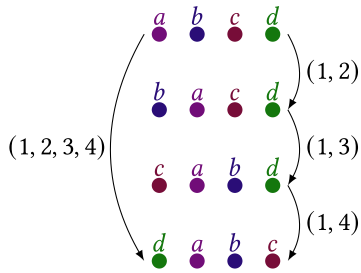


Figure 6.2: The permutation $(1\,2\,3\,4)$ acts on (a, b, c, d), which can be done in steps where each step is a transposition.

The second case is when this acts on a_1, in which case the first cycle sends a_1 to a_2, which then commutes with all remaining cycles and so a_1 maps to a_2, which is what we want.
The final case is trivial, it's where this acts on some $a \neq a_i$ for any i, in which case on both the left and right this element is not changed, and we are finished. □

The above theorem is fairly obvious. It states that we can do any permutation just by swapping two items at a time. This is demonstrated in Figure 6.2.

Notice that this theorem implies that the rank of S_n is at most

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{1}{2}n(n-1), \tag{6.1.17}$$

although we will see it is less than this.

Lemma 6.1.18 The transpositions, $\{(1\,2), (1\,3), \dots, (1\,n)\}$, generate S_n.

Proof. Notice that

$$(a_1\,a_2) = (1\,a_1)(1\,a_2)(1\,a_1). \tag{6.1.19}$$

To see this we can consider this acting on $(1\,\dots\,a_1\,\dots\,a_2\,\dots)$:

$$\begin{aligned} &(1\,a_1)(1\,a_2)(1\,a_1)(1\,\dots\,a_1\,\dots\,a_2\,\dots) && (6.1.20)\\ &= (1\,a_1)(1\,a_2)(a_1\,\dots\,1\,\dots\,a_2\,\dots) && (6.1.21)\\ &= (1\,a_1)(a_1\,\dots\,a_2\,\dots\,1\,\dots) && (6.1.22)\\ &= (1\,\dots\,a_2\,\dots\,a_1\,\dots) && (6.1.23)\end{aligned}$$

so the action of $(1\,a_1)(1\,a_2)(1\,a_1)$ is to swap a_1 and a_2, which is exactly the action of $(a_1\,a_2)$. Using this we can generate any transposition from transpositions of the form $(1\,a)$. Therefore, by Theorem 6.1.15 transpositions of this form generate S_n. □

Notice that this theorem implies that the rank of S_n is at most n, for $n > 3$ this is an improvement on our previous bound.

Lemma 6.1.24 The transpositions, $\{(1\,2), (2\,3), \dots, (n-1\,n)\}$, generate S_n.

Proof. Notice that

$$(1\,k) = g(1\,2)g^{-1} \tag{6.1.25}$$

where $g = (k-1\,k)\cdots(3\,4)(2\,3)$. To see this first notice that $g^{-1} = (2\,3)(3\,4)\cdots(k-1\;\;k)$. Then notice that the action of g^{-1} is to exchange $k-1$ and k, then swap k and $k-2$, and then k and $k-3$, and so on until k and 2 have been swapped. Then $(1\,2)$ swaps k and 1. We then use g to swap 2 and 3 back, then 3 and 4, and so on until we swap $k-2$ and $k-1$ back to their original positions. The result is that 1 and k swap, which is exactly what $(1\,k)$ does.
Using this we can generate any transposition of the form $(1\,a)$ from transpositions of the form $(k-1\;\;k)$, and so by Lemma 6.1.18 these transpositions generate S_n. □

We can think of this proof as consisting of a basis change from the $(1\,a)$ transpositions case, and similarly for the next proof.

Lemma 6.1.26 The cycles, $\{(1\,2), (1\,2\,\dots\,n)\}$, generate S_n.

Proof. Notice that

$$(k\;\;k+1) = g(1\,2)g^{-1} \tag{6.1.27}$$

where $g = (1\,2\,\dots\,n)^{k-1}$. To see this first notice that $g^{-1} = (n\,\dots\,2\,1)^{k-1}$. The action of g^{-1} is then to cycle backwards through the elements $k-1$ times. The result is that k and $k+1$ end up in the first two positions. These are then swapped by $1, 2$. Then g cycles forwards through the elements $k-1$ times, and we end up with the same tuple but with k and $k+1$ swapped, which is exactly what $(k\;\;k+1)$ does.
Using this we can generate any transposition of the form $(k\,k+1)$, and so by Lemma 6.1.24 these two cycles generate S_n. □

Corollary 6.1.28 The rank of S_n is 2.

Proof. S_n is generated by $\{(1\,2), (1\,\dots\,n)\}$ by Lemma 6.1.26, so S_n is of rank 2. □

6.2 Alternating Group

We have seen that any permutation can be written as a product of transpositions by Theorem 6.1.15. It turns out that the number of transpositions it takes to write any given permutation is unambiguously even or odd. We then term the permutation as even or odd based on the parity of the number of transpositions it takes to write it. This is the obvious definition of even and odd permutations, but it isn't that easy to work with, so we use an equivalent, but easier to work with, definition.

Definition 6.2.1 — Sign of a Permutation Let S_n be the symmetric group on n letters. Define the **Vandermonde polynomial** to be

$$\begin{aligned}
P(x_1, \ldots, x_n) &= P(\boldsymbol{x}) && (6.2.2)\\
&:= (x_1 - x_2)(x_1 - x_2) \cdots (x_1 - x_n) && (6.2.3)\\
&\quad \cdot (x_2 - x_3) \cdots (x_2 - x_n) \cdots (x_{n-1} - x_n) && (6.2.4)\\
&= \prod_{\substack{i,j \in \{1,\ldots,n\} \\ i<j}} (x_i - x_j). && (6.2.5)
\end{aligned}$$

This polynomial is such that exchanging any two variables, x_i and x_j, results in the sign of the polynomial changing.
Define X to be the set of all permutations of $(x_1, \ldots, x_n)$. We can define a group action, $\varphi\colon S_n \times X \longrightarrow X$, in the usual way as S_n acting on $\boldsymbol{x} \in X$ by permutation.
Now define $\operatorname{sgn}\colon S_n \longrightarrow \{\pm 1\}$ by $\operatorname{sgn}(\sigma) = P(\sigma . \boldsymbol{x})/|P(\boldsymbol{x})|$. Then if $\operatorname{sgn}(\sigma) = 1$ we say σ is an **even permutation** and if $\operatorname{sgn}(\sigma) = -1$ we say that σ is an **odd permutation**. This agrees with the "parity of the number of transpositions" definition since each transposition swaps two variables and so an even number of transpositions corresponds to an even number of swaps and hence no overall sign change, similarly an odd number of transpositions will result in a sign change.

Definition 6.2.6 — Alternating Group The **alternating group**, A_n, is the group of all even permutations on n letters.

Theorem 6.2.7. The alternating group, A_n, is a normal subgroup of the symmetric group, S_n.

Proof. We claim that $\psi\colon S_n \longrightarrow \mathbb{Z}_2$ as defined in Definition 6.2.1 is a group homomorphism. First notice that $\operatorname{sgn}(\sigma\rho) = P(\sigma\rho\,.\,\boldsymbol{x})/|P(\boldsymbol{x})|$. Now consider $\operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = P(\sigma\,.\,\boldsymbol{x})P(\rho\,.\,\boldsymbol{x})/|P(\boldsymbol{x})|^2$. Hence, $\operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = 1$ if σ and ρ have the same parity and $\operatorname{sgn}(\sigma)\operatorname{sgn}(\rho) = -1$ if σ and ρ have opposite parities. Suppose $\sigma = s_1 \cdots s_k$ and $\rho = r_1 \cdots r_m$ where s_i and r_i are transpositions. Then $\sigma\rho = s_1 \cdots s_k r_1 \cdots r_m$ is a product of $k+m$ transpositions. This is even if k and m are both even, or both odd, and odd if k and m have opposite parities. Therefore, $\operatorname{sgn}(\sigma\rho) = P(\sigma\rho\,.\,\boldsymbol{x})/|P(\boldsymbol{x})|$ is 1 if σ and ρ are the same parity and -1 if they are opposite parities. Hence, sgn is a homomorphism.
The kernel of sgn is A_n, since even permutations map to 1 by definition. Hence, by Theorem 4.2.20 A_n is a normal subgroup of S_n. □

Lemma 6.2.8 The three-cycles generate A_n.

Proof. Notice that $(1\,a_1)(1\,a_2) = (1\,a_1\,a_2)$. Since transpositions of the form $(1\,a)$ generate S_n and A_n consists of all permutations which can be written as a product of an even number of transpositions we can always pair up transpositions like this to write elements of A_n as a product of three cycles. Hence, the three-cycles generate A_n. □

Notice that the order of A_n is $|A_n| = |S_n|/|\mathbb{Z}_2| = n!/2$.

Seven

Applications

7.1 Platonic Solids

Definition 7.1.1 — Platonic Solid A **Platonic solid** is a regular convex polyhedron.

There are five Platonic solids, the tetrahedron, cube, octahedron, dodecahedron, and icosahedron. These have 4, 6, 8, 12, and 20 faces, respectively, and are formed from equilateral triangles, squares, equilateral triangles, regular pentagons, and equilateral triangles, respectively.

Each Platonic solid has an associated symmetry group, which acts on the solid leaving it invariant. These symmetry groups are all permutation groups, both symmetric and alternating. We can think of them as acting by permuting the vertices of the solid. The reason why the full symmetry group is not necessarily all of S_n is because certain permutations aren't allowed, for example, if two vertices are connected by an edge then this must be the case after permuting vertices.

The dual of a Platonic solid is the Platonic solid you get if you swap the vertices and faces, that is if you join the centre of two faces if they have a common edge. This is demonstrated in Figure 7.1 for the cube and octahedron. The tetrahedron is its own dual, the cube and octahedron are dual and the dodecahedron and icosahedron are dual. Duals have the same symmetry group.

Consider the tetrahedron, with the symmetry group A_4. This group is of order $4!/2 = 12$. A_4 is generated by two symmetries shown in Figure 7.2.

Table 7.1: The Platonic solids, along with the number of faces, F, vertices, V, and edges, E, which are related by Euler's formula, $V - E + F = 2$, the regular polygons that make up their faces, and their symmetry groups. Notice that duals have the same symmetry groups, the same number of edges and that the number of faces and vertices are swapped.

Sold	F	V	E	Polygon	Symmetry Group
Tetrahedron	4	4	6	Triangle	A_4
Cube	6	8	12	Square	S_4
Octahedron	8	6	12	Triangle	S_4
Dodecahedron	12	20	30	Pentagon	A_5
Icosahedron	20	12	30	Triangle	A_5

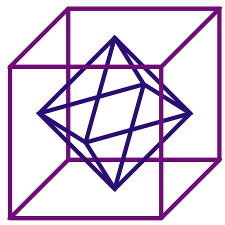


Figure 7.1: The octahedron is the dual of the cube.

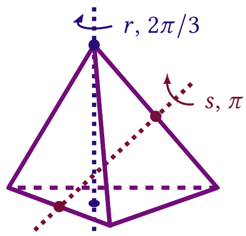


Figure 7.2: Two possible symmetries of the tetrahedron

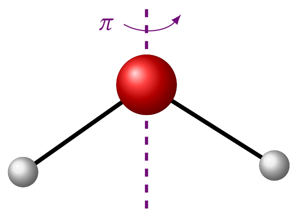


Figure 7.3: Water, with symmetry group $\mathbb{Z}_2$, has a permanent dipole.

For the case of the cube we can view S_4 as acting by permuting the diagonals of the cube.

7.2 Molecules

For a molecule to have a permanent electric dipole it must necessarily have some level of asymmetry, since a spherically symmetric molecule cannot have a preferred direction for a dipole to lie along. For example, water, has a permanent dipole, and it has as a symmetry group $\mathbb{Z}_2$, as shown in Figure 7.3.

Any molecule with symmetry group $\mathbb{Z}_n$ with $n \in \{2, 3, \dots\}$ cannot have a permanent electric dipole perpendicular to the symmetry axis. For example, water's electric dipole is aligned with its symmetry axis.

Consider again the case of ethene, C_2H_6, shown in Figure 5.1. This has symmetry group $D_3 \cong \mathbb{Z}_2 \ltimes \mathbb{Z}_3$. With $\mathbb{Z}_3$ corresponding to rotations by $2\pi/3$ about the carbon-carbon bond and $\mathbb{Z}_2$ corresponding to either inversion about the middle of the carbon-carbon bond or rotations about the perpendicular bisector to the carbon-carbon bond, depending on whether the molecule is eclipsed or staggered. Either way C_2H_6 cannot have a permanent dipole since the two axis of symmetry are orthogonal. For example, suppose there was an electric dipole aligned with the $\mathbb{Z}_3$ symmetry axis. Then the action of $\mathbb{Z}_2$ would be to reverse the dipole, and hence this isn't a valid permanent dipole.

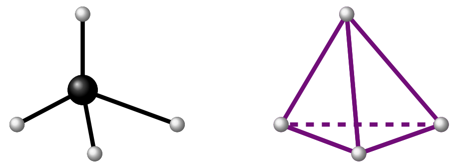
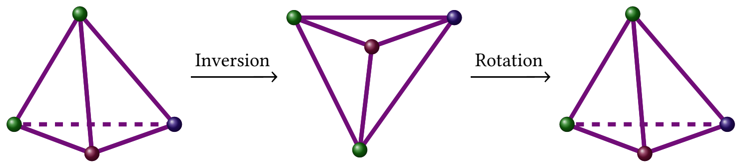


Figure 7.4: Methane is a tetragonal molecule, meaning that it has a central atom, here carbon, and four atoms arranged around it making the points of a tetrahedron.

It can be shown that only molecules with a cyclic symmetry group, $\mathbb{Z}_n$, can have permanent dipoles.

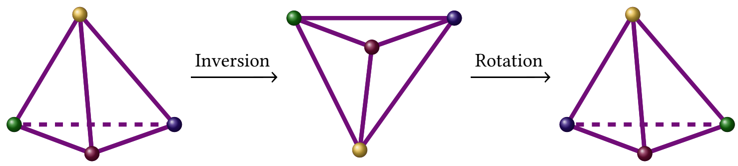
Another example of symmetry applications to molecules is chirality. A molecule is chiral if it is different from its mirror image. Another way of saying this is that a molecule is chiral if it does not admit an improper rotation axis, an improper rotation being a rotation followed by an inversion.

A tetragonal molecule is one where there is a centre atom with four atoms around it at the points of a tetrahedron, such as methane, CH_4, shown in Figure 7.4. An example of a tetragonal molecule which is achiral is CCl_2BrI, this is shown here:



(7.2.1)

On the other hand CFClBrI is chiral, this is shown here:



(7.2.2)

Notice that the green and blue are swapped after the inversion and rotation.

Part II

Representation Theory

Eight

Basics of Representation Theory

(R) Material in this section applies both to finite groups and compact groups. We won't worry too much about what it means for a group to be compact, we just note that the Lie groups $\mathrm{U}(1) \cong \mathrm{SO}(2)$ and $\mathrm{SO}(3) \cong \mathrm{SU}(2)/\mathbb{Z}_2$ are compact. More generally $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{Spin}(n)$, $\mathrm{U}(n)$, and $\mathrm{SU}(n)$ are compact.

8.1 Representation Definition

There are two essentially equivalent definitions of a representation:

Definition 8.1.1 — Representation A **representation** of a group, G, is a group action of G on a linear space, V. That is $\varphi\colon G \times V \to V$ is a representation if it is a group action.
A **representation** of a group, G, is a homomorphism with the automorphism group of a linear space. That is $\rho\colon G \to \mathrm{GL}(V)$ is a representation if it is a homomorphism.

The equivalence of these two definitions is simple, if $\varphi(g, \boldsymbol{v}) = g \,.\, \boldsymbol{v} = \rho(g)\boldsymbol{v}$ for all $g \in G$ and $\boldsymbol{v} \in V$ then φ and ρ are the same representation in the two slightly different definitions.

When the homomorphism is clear it is common to refer to V as the representation, rather than ρ.

Representation theory gives us a way to do concrete calculations in a group. We have implicitly been using representation theory already, for example we have used $\mathbb{Z}_n = \{\mathrm{e}^{2i\pi m/n} \mid m = 0, \dots, n-1\}$ as *the* cyclic group of order n. Strictly this is actually a representation of a more general cyclic group, defined by $\langle g \mid g^n = e\rangle$. The linear space in question is $\mathbb{C}$, and $\mathrm{GL}(\mathbb{C}) = \mathbb{C} \setminus \{0\}$. The group action is rotation by $2\pi m/n$.

Representation theory is particularly useful because we have developed a lot of tools for dealing with computations in linear spaces since they are common in other areas of maths and physics. Representation theory allows us to use these to work with groups.

▪ **Application 8.1.2** When we solve the quantum harmonic oscillator we can do so in different linear spaces, such position space, $|x\rangle$, momentum space, $|p\rangle$, or number-of-particles space, $|n\rangle$. Each one of these corresponds to studying the same underlying physics on a different linear space, which

> we can think of as being different representations.
> If a system has a certain symmetry described by a group, G, then this is expressed in the mathematics by G acting on the Hilbert space of states, which is to say as a representation of G on the Hilbert space of states.

If the relevant linear space is finite dimensional, say $\dim V = n$ then we can associate $\mathrm{GL}(V)$ with $\mathrm{GL}(n, \mathbb{F})$, the group of $n \times n$ matrices with entries in $\mathbb{F}$, and so we can associate $\rho(g)$ with a matrix.

8.2 Pedestrian Approach

R In this section we will get an idea of how representation can be used without being too worried about precise definitions.

Consider S_3. By Lemma 6.1.18 $\{(1\,2), (1\,3)\}$ generates S_n, so we can define a representation by how it maps these two elements to the linear space. Inspired by S_3 as a permutation group we look for a representation that conserves this permuting ability. In particular the obvious choice of 3 things to act on in a linear space are the 3 basis vectors of a 3-dimensional space, such as $\mathbb{R}^3$. We use the standard basis, $\boldsymbol{e_1} = (1, 0, 0)^\top$, $\boldsymbol{e_2} = (0, 1, 0)^\top$, and $\boldsymbol{e_3} = (0, 0, 1)^\top$. We can easily construct matrices that permute these, for example we want $\rho((1\,2))\boldsymbol{e_1} = \boldsymbol{e}_2$, $\rho((1\,2))\boldsymbol{e_2} = \boldsymbol{e}_2$ and $\rho((1\,2))\boldsymbol{e_3} = \boldsymbol{e}_3$. This fully determines the matrix $\rho((1\,2))$:

$$\rho((1\,2)) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{8.2.1}$$

Similarly, we have

$$\rho((1\,3)) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{8.2.2}$$

This is called the **permutation representation** of S_3, and can easily be generalised to S_n.

At this point there are a few questions we should consider. First, are there any other representations? The answer to this is yes, and we'll see some later. The second is can we find a simpler representation, for some sense of simpler. The answer is again yes. In particular notice that $\boldsymbol{v} = (1, 1, 1)^\top$ is a common eigenvector for both of these matrices, and hence there is an invariant subspace, $\operatorname{span}\{\boldsymbol{v}\}$, which is unchanged by this representation. It can be shown that this allows us to perform a basis change and write these matrices in block diagonal form. We can then define a new representation, ρ', in this block diagonal basis such that

$$\rho'((1\,2)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a_{11} & a_{12} \\ 0 & a_{21} & a_{22} \end{pmatrix}, \qquad \text{and} \qquad \rho'((1\,3)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & b_{11} & b_{12} \\ 0 & b_{21} & b_{22} \end{pmatrix}. \tag{8.2.3}$$

This is simpler because it is fully determined by the 2×2 block matrices on the diagonal. We say that the permutation representation is reducible.

8.3 Basic Definitions

> **Definition 8.3.1 — Faithful** If $\rho\colon G \to \mathrm{GL}(V)$ is a representation of G then we say ρ is **faithful** if it is injective, that is $\rho(g) = \rho(g')$ implies $g = g'$. If this is not the case then we say ρ is **unfaithful**.

The permutation representation of S_3 is faithful.

> **Definition 8.3.2 — Trivial Representation** Let $\rho\colon G \to \mathrm{GL}(V)$ be defined by $\rho(g) = \rho(e) = \mathbb{1}_V$, where $\mathbb{1}_V$ is the identity in $\mathrm{GL}(V)$ and V is an *arbitrary vector space*. We say that ρ is *A* **trivial representation**.
> *The* **trivial representation** is $\rho\colon G \to \mathrm{GL}(V)$ defined by $\rho(g) = \rho(e) = \mathbb{1}_V$, where $\mathbb{1}_V$ is the identity in $\mathrm{GL}(V)$ and V is a *one-dimensional vector space*.

A trivial representation is maximally unfaithful since all elements map to the same operator.

> **Definition 8.3.3 — Unitary Representation** If $\rho\colon G \to \mathrm{U}(V)$ is a homomorphism then we say that ρ is a **unitary representation**. That is a unitary representation is one in which $\rho(g)$ is a unitary operator for all $g \in G$.

We have been working with a unitary representation of $\mathbb{Z}_n$ as $\{\mathrm{e}^{2i\pi m/n} \mid m = 0, \dots, n-1\}$.

> **Definition 8.3.4 — Equivalence of Representations** If ρ and ρ' are representations of G on V we say that ρ and ρ' are **equivalent** if they are represented by a **similarity transform**, that is $\rho(g) = S\rho'(g)S^{-1}$ for some $S \in \mathrm{GL}(V)$. Notice that S must be the same for all $g \in G$.

As the name suggests the equivalence of representations, $\sim$, is an equivalence relation. Clearly $\rho \sim \rho$ since $\rho(g) = \mathbb{1}_V\rho(g)\mathbb{1}_V^{-1}$. Also, if $\rho \sim \rho'$ then $\rho(g) = S\rho'(g)S^{-1}$ for some $S \in \mathrm{GL}(V)$, and so $\rho'(g) = S^{-1}\rho(g)S$, identifying $S = (S^{-1})^{-1}$ and knowing that if $S \in \mathrm{GL}(V)$ we also have $S^{-1} \in \mathrm{GL}(V)$ we see that $\rho' \sim \rho$. Finally, if $\rho \sim \rho'$ and $\rho' \sim \rho''$ there exists $S, T \in \mathrm{GL}(V)$ such that $\rho(g) = S\rho'(g)S^{-1}$ and $\rho'(g) = T\rho''(g)T^{-1}$, and hence $\rho(g) = ST\rho''(g)T^{-1}S^{-1} = (ST)\rho(g)(ST)^{-1}$ and since $S, T \in \mathrm{GL}(V)$ it follows that $ST \in \mathrm{GL}(V)$, and hence $\rho \sim \rho''$.

> **Definition 8.3.5 — Invariant Subspace** Let $\rho\colon G \to \mathrm{GL}(V)$ be a representation of the group G on the linear space V. Let W be a subspace of V. Then we call W an **invariant subspace** if $\rho(g)\boldsymbol{w} \in W$ for all $g \in G$ and $\boldsymbol{w} \in W$.
> Using the group-action definition of a representation W is an invariant subspace if $\mathrm{Orb}(\boldsymbol{w}) \subseteq W$ for all $\boldsymbol{w} \in W$.

A trivial representation, $\rho(g) = \mathbb{1}_V$, has V as an invariant subspace. Our earlier example of S_3 in the permutation representation has $\operatorname{span}\{(1,1,1)^\top\}$ as an invariant subspace. All representations leave the trivial subspace, $\{\mathbf{0}\}$, invariant.

> **Definition 8.3.6 — Irreducible** We call a representation **irreducible** if it has no invariant subspaces, apart from the trivial zero-dimensional subspace, $\{\mathbf{0}\}$. If a representation can be written in block diagonal form then it is **reducible**

It is common to shorten "irreducible representation" to **irrep**.

For finite groups and continuous compact groups the reducible representations can be written as a direct sum of irreducible representations. For this reason we often care only about irreducible representations.

The permutation representation of S_3 is reducible.

Lemma 8.3.7 If a representation has invariant subspaces then we can write it in block diagonal form.

Definition 8.3.8 — Real and Complex Representations A representation, $\rho\colon G \to \mathrm{GL}(V)$, is **real** if either V is a real vector space or the representation is equivalent to a representation which can be thought of as acting on a real vector space by reducing the field of scalars to $\mathbb{R}$. Alternatively ρ is equivalent to ρ^*.

A **complex representation** is a representation, $\rho\colon G \to \mathrm{GL}(V)$, where ρ is not equivalent to ρ^*.

We will refine this notion later to include pseudo-real representations.

The permutation representation of S_3 is real, and hence so is ρ', even though it is possible that a_{ij} and/or b_{ij} are not real, since ρ and ρ' are equivalent.

▪ **Application 8.3.9** Unitary representations are important in quantum physics. They are the natural language to describe symmetries on the Hilbert space of states since they preserve the inner product, and hence the probability of being in a given state.

Further there is a certain view from which particles *are* irreducible representations of the Poincaré group, $\mathbb{R}^{1,3} \ltimes \mathrm{O}(1,3)$. We then associate complex representations with charged particles and real representations with neutral particles.

8.4 Some Theorems

Theorem 8.4.1 — Maschke's Theorem. Any representation of a finite group is equivalent to a unitary representation.

Proof. Let G be a finite group, V a vector space, and $\rho\colon G \to \mathrm{GL}(V)$ a representation. Let $\langle -, - \rangle$ be an inner product on G. The statement of the theorem is equivalent to stating that we can define a new inner product on V such that this inner product is invariant under the action of this representation. This works since the two inner products will be related by a change of basis, and hence this new inner product can be viewed as the old inner product after a similarity transform.

The inner product that we define is $\langle -, - \rangle_G$, it is defined for $x, y \in V$ as

$$\langle x, y \rangle_G := \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)x, \rho(g)y \rangle. \tag{8.4.2}$$

We can think of this inner product being defined by acting on the old inner product with the representation and then averaging over G.
We need to show that $\langle -, -\rangle_G$ is an inner product and that it is invariant with respect to the representation. By definition $\rho(g)x, \rho(g)y \in V$ and so $\langle \rho(g)x, \rho(g)y\rangle$ is positive definite. This carries through the sum and so $\langle x, y\rangle_G$ is positive definite. Linearity and conjugate symmetry of $\langle -, -\rangle_G$ similarly follows from these same properties for $\langle -, -\rangle$ without any complications since G is finite.
Now consider what happens when we first act on V with $\rho(g')$ for some $g' \in G$. Using the fact that ρ is a homomorphism we then have

$$\langle \rho(g')x, \rho(g')y\rangle_G = \frac{1}{|G|}\sum_{g\in G}\langle \rho(g)\rho(g')x, \rho(g)\rho(g')y\rangle \tag{8.4.3}$$
$$= \sum_{g\in G}\langle \rho(gg')x, \rho(gg')y\rangle \tag{8.4.4}$$
$$= \sum_{g''\in G}\langle \rho(g'')x, \rho(g'')y\rangle \tag{8.4.5}$$
$$= \langle x, y\rangle_G. \tag{8.4.6}$$

In the penultimate step we have used the fact that as g takes on all values in G gg' necessarily also takes on all values in G. This is due to the fact that in a Cayley table each column must contain every element of G exactly once. Hence, summing over g with factors of gg' is the same as summing over $g'' = gg'$, only the order of the terms changes and since the inner product gives an element of the base field this sum is commutative.
Finally, we remark that

$$\langle x, y\rangle_G = \langle \rho(g')x, \rho(g')y\rangle_G = \langle \rho(g')^\dagger\rho(g')x, y\rangle_G \tag{8.4.7}$$

using the property of inner products that $\langle x, Ay\rangle = \langle A^\dagger x, y\rangle$ for any inner product, $\langle -, -\rangle$, and operator A. Hence, we can identify that $\rho(g')^\dagger\rho(g') = \mathbb{1}$, and so ρ is a unitary representation. □

It turns out that being irreducible is an incredibly strong requirement, so much so that it doesn't really leave much wiggle room, as the next theorem shows. Schur's lemma states that there is no room for non-trivial homomorphisms between irreducible representations.

Theorem 8.4.8 — Schur's Lemma. Let $\rho\colon G \to \mathrm{GL}(V)$ and $\rho\colon G \to \mathrm{GL}(V')$ be irreducible representations of the group G on some finite dimensional vector spaces V and V'. Let $T\colon V \to V'$ be a linear map satisfying $\rho'(g) \circ T = T \circ \rho(g)$ for all $g \in G$ where $\circ$ is composition of functions. Then

1. either T is an isomorphism or T is trivial.

2. If $V = V'$ then $T = \lambda\mathbb{1}$ for $\lambda \in \mathbb{C}$ and $\mathbb{1}$ being the identity map on V.

Proof. The first step is to notice that $\ker T$ and $\operatorname{Im} T$ are invariant subspaces. Recall that $\ker T := \{v \in V \mid Tv = 0\}$. To show that $\ker T$ is an invariant subspace we need to show that $\rho(g)v \in \ker T$ for all $v \in \ker T$. To do this we notice that for $v \in \ker T$ we have

$$\begin{aligned} T\rho(g)v = (T \circ \rho(g))v = (\rho'(g) \circ T)v \\ = \rho'(g)Tv = \rho'(g)(Tv) = \rho'(g)0 = 0 \end{aligned} \tag{8.4.9}$$

where the final equality follows since linear maps map 0 to 0. We have therefore shown that $T\rho(g)v = 0$ for all $v \in \ker T$ and hence $\rho(g)v \in \ker T$ so $\ker T$ is an invariant subspace of V under ρ.
Now recall that $\operatorname{Im} T = \{v' \in V' \mid v' = T(v) \text{ for some } v \in V\}$. This is an invariant subspace if $\rho'(g)v \in \operatorname{Im} T$ for all $v' \in \operatorname{Im} T$. To show this we notice that for $v' \in \operatorname{Im} T$ we have $v \in V$ such that $v' = Tv$ and so

$$\rho'(g)v' = \rho'(g)Tv = (\rho'(g) \circ T)v = (T \circ \rho(g))v = T\rho(g)v \tag{8.4.10}$$

and so $\rho'(g)v'$ is of the form $T\rho(g)v$ and $\rho(g)v \in V$ meaning $\rho'(g)v' \in \operatorname{Im} T$. Hence, $\operatorname{Im} T$ is an invariant subspace of V' under ρ'.
By definition ρ and ρ' are irreducible representations and therefore have no nontrivial invariant subspaces. This means that the invariant subspace $\ker T$ must be $\{0\}$ or V, and similarly $\operatorname{Im} T$ must be $\{0\}$ or V'. We now treat this by cases.

- Suppose $\ker T = \{0\}$. Then $\operatorname{Im} T \neq \{0\}$ since all $v \in V$ with $v \neq 0$ map to something other than 0, meaning that $\operatorname{Im} T$ must contain nonzero elements. Hence, $\operatorname{Im} T = V'$. It follows that T is an isomorphism since a trivial kernel implies that $T\colon V \longrightarrow V'$ is injective since $V' = \operatorname{Im} T$ this map is also surjective.

- Suppose $\ker T = V$. Then $\operatorname{Im} T = \{0\}$, since by definition all $v \in V$ map to 0. Hence, T is the trivial zero function, $T(v) = 0$ for all $v \in V$.

This finishes the proof of the first statement.
For the second statement suppose T is nontrivial, that is $T \neq 0$. Then T has at least one nonzero eigenvalue, $\lambda \in \mathbb{C}$, since if all eigenvalues are zero then T is trivial. Now define a second linear map $U := T - \lambda\mathbb{1}$. Then by construction at least one eigenvalue of U is 0 and so U is not an isomorphism, since all vectors parallel to the eigenvector with eigenvalue 0 are mapped to 0. More formally this means that $\dim(\ker U) \geq 1$, in fact the dimension is the multiplicity of the eigenvalue 0.
Since $U = T - \lambda\mathbb{1}$ it is clear that $\rho'(g) \circ U = U \circ \rho(g)$ and so by the first part of this theorem $U = 0$, since U is not an isomorphism. Hence, $U = 0 = T - \lambda\mathbb{1}$ which we can rearrange to get $T = \lambda\mathbb{1}$. □

An equivalent statement to the first part of Schur's lemma is that the following diagram commuting for all $g \in G$ only if T is trivial or an isomorphism:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\rho(g)"'] & V' \arrow[d, "\rho'(g)"] \\ V \arrow[r, "T"] & V' \end{tikzcd} \tag{8.4.11}$$

An alternative statement of the second part of Schur's lemma, which is often state as the full version of Schur's lemma in the physics literature, is the following corollary.

Corollary 8.4.12 Let ρ be an irreducible representation of a group G on some finite dimensional vector space, V, and T be some linear map on this same vector space such that $[T, \rho(g)] = 0$ for all $g \in G$. Then $T = \lambda \mathbb{1}$.

R Here $[A, B] := AB - BA$ is the usual **commutator**.

The final theorem for this section relates writing representations as direct sums of irreducible representations. See Definition A.2.55 for the definition of direct sums.

Theorem 8.4.13 — Decomposability Theorem. Let $\rho\colon G \longrightarrow \mathrm{GL}(V)$ be a reducible representation for some compact group G. Then we can write

$$\rho(g) = m_1\rho_1(g) \oplus m_2\rho_2(g) \oplus \cdots \oplus m_k\rho_k(g) = \bigoplus_{i=1}^{k} m_i\rho_i(g) \tag{8.4.14}$$

where $\rho_i\colon G \longrightarrow \mathrm{GL}(V)$ are irreducible representations and $m_i \in \mathbb{Z}_{>0}$. By $m_i\rho_i(g)$ we mean

$$m_i\rho_i(g) := \underbrace{\rho_i(g) \oplus \cdots \oplus \rho_i(g)}_{m_i \text{ times}} = \bigoplus_{j=1}^{m_i} \rho_i(g). \tag{8.4.15}$$

Proof. □

Nine

Character Theory of Finite Groups

9.1 Basics of Character Theory

Definition 9.1.1 — Character Let G be a finite group and ρ a representation of G. Then we define the **character**, $\chi(g)$, of some $g \in G$ as

$$\chi(g) \coloneqq \mathrm{tr}[\rho(g)]. \tag{9.1.2}$$

Notation 9.1.3 When discussing multiple representations of G we will denote the representation being considered by a subscript, so $\chi_\rho(g) = \mathrm{tr}[\rho(g)]$ and $\chi_{\rho'}(g) = \mathrm{tr}[\rho'(g)]$.

Lemma 9.1.4 The characters of two equivalent representations are equal.

Proof. Recall that the representations ρ and ρ' of G on V are equivalent if there exists $S \in \mathrm{GL}(V)$ such that $\rho'(g) = S\rho(g)S^{-1}$ for all $g \in G$. Therefore,

$$\begin{aligned}
\chi_{\rho'}(g) &\coloneqq \mathrm{tr}[\rho'(g)] && (9.1.5)\\
&= \mathrm{tr}[S\rho(g)S^{-1}] && (9.1.6)\\
&= \mathrm{tr}[S^{-1}S\rho(g)] && (9.1.7)\\
&= \mathrm{tr}[\rho(g)] && (9.1.8)\\
&\eqqcolon \chi_\rho(g). && (9.1.9)
\end{aligned}$$

Here we have used the cyclic property of the trace, $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$. □

Recall that an equivalence relation is a relation which is symmetric, reflexive, and transitive. The equivalence class of $a \in A$ under some equivalence relation, $\sim$, is the set $[a] \coloneqq \{b \in A \mid a \sim b\}$. The set of all equivalence classes is denoted $A/{\sim}$.

Definition 9.1.10 A **class function** is a function, $f\colon A \to B$, which takes on the same value for all $b \in [a]$, where $[a]$ is an equivalence class of A under some equivalence relation.
We can therefore define a similar function $\tilde{f}\colon A/{\sim} \to B$ defined by $\tilde{f}([a]) = f(a)$, which is well defined for class functions f. We typically don't distin-

guish between f and $\tilde{f}$.

Recall that $x \sim y$ if $x = gyg^{-1}$ is an equivalence relation, called conjugacy, and the equivalence classes of this equivalence relation are called conjugacy classes.

Lemma 9.1.11 The character is a class function on the conjugacy classes.

Proof. Let $x, y \in G$ be in the same conjugacy class. Then $x = gyg^{-1}$ for some $g \in G$. Let ρ be a representation of G. Then

$$\begin{aligned}
\chi_\rho(x) &:= \mathrm{tr}[\rho(x)] && (9.1.12)\\
&= \mathrm{tr}[\rho(gyg^{-1})] && (9.1.13)\\
&= \mathrm{tr}[\rho(g)\rho(y)\rho(g^{-1})] && (9.1.14)\\
&= \mathrm{tr}[\rho(g)\rho(y)\rho(g)^{-1}] && (9.1.15)\\
&= \mathrm{tr}[\rho(g)^{-1}\rho(g)\rho(y)] && (9.1.16)\\
&= \mathrm{tr}[\rho(y)] && (9.1.17)\\
&=: \chi_\rho(y). && (9.1.18)
\end{aligned}$$

Here we have used the fact that ρ is a homomorphism so $\rho(ab) = \rho(a)\rho(b)$ and $\rho(a^{-1}) = \rho(a)^{-1}$. We have also used the cyclic property of the trace, $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$. Hence, $\chi_\rho(x) = \chi_\rho(y)$ for all $x, y \in [x] \in G/\sim$ where $\sim$ is conjugacy. □

Lemma 9.1.19 Let ρ and ρ' be representations of G. Then $\chi_{\rho\oplus\rho'}(g) = \chi_\rho(g) + \chi_{\rho'}(g)$ for all $g \in G$ and $\chi_{\rho\otimes\rho'}(g) = \chi_\rho(g)\chi_{\rho'}(g)$.

Proof. In this proof we use braket notation. Let $\{|i\rangle\}$ be an orthonormal basis. Then in braket notation with the Einstein summation convention

$$\mathrm{tr}(A) := \langle i|A|i\rangle. \tag{9.1.20}$$

Therefore,

$$\begin{aligned}
\chi_{\rho\oplus\rho'}(g) &:= \mathrm{tr}[(\rho\oplus\rho')(g)] && (9.1.21)\\
&= \mathrm{tr}[\rho(g)\oplus\rho'(g)] && (9.1.22)\\
&= \langle i|(\rho(g)\oplus\rho'(g))|i\rangle && (9.1.23)\\
&= \langle i|\rho(g)|i\rangle + \langle i|\rho'(g)|i\rangle && (9.1.24)\\
&= \mathrm{tr}[\rho(g)] + \mathrm{tr}[\rho'(g)] && (9.1.25)\\
&=: \chi_\rho(g) + \chi_{\rho'}(g). && (9.1.26)
\end{aligned}$$

Here we have used the fact that

$$(A\oplus B)(|v\rangle \oplus |w\rangle) = A|v\rangle \oplus B|w\rangle \tag{9.1.27}$$

and so

$$\begin{aligned}
(\langle v'| \oplus \langle w'|)(A\oplus B)(|v\rangle \oplus |w\rangle) &= (\langle v'| \oplus \langle w'|)(A|v\rangle \oplus B|w\rangle) && (9.1.28)\\
&= \langle v'|A|v\rangle + \langle w'|B|w\rangle. && (9.1.29)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\chi_{\rho\otimes\rho'}(g) &\coloneqq \operatorname{tr}[(\rho\otimes\rho')(g)] && (9.1.30)\\
&= \operatorname{tr}[\rho(g)\otimes\rho'(g)] && (9.1.31)\\
&= \langle i|(\rho(g)\otimes\rho'(g))|i\rangle && (9.1.32)\\
&= \langle i|\rho(g)|i\rangle\langle i|\rho'(g)|i\rangle && (9.1.33)\\
&= \operatorname{tr}[\rho(g)]\operatorname{tr}[\rho'(g)] && (9.1.34)\\
&\eqqcolon \chi_\rho(g)\chi_{\rho'}(g). && (9.1.35)
\end{aligned}$$

Here we have used the fact that

$$(A\otimes B)(|v\rangle\otimes|w\rangle) = A|v\rangle\otimes B|w\rangle. \tag{9.1.36}$$

□

Lemma 9.1.37 Let G be a finite group and ρ a representation of G. Then $\chi_\rho(g^{-1}) = \chi_\rho(g)^*$ for all $g \in G$.

Proof. By Theorem 8.4.1 ρ is equivalent to a unitary representation, ρ'. By Lemma 9.1.4 $\chi_\rho(g) = \chi_{\rho'}(g)$ for all $g \in G$. Since ρ' is a homomorphism $\rho'(g^{-1}) = \rho'(g)^{-1}$ by Lemma 2.1.20. ρ' is unitary so $\rho'(g)^{-1} = \rho'(g)^\dagger$. Transposing a matrix leaves the diagonal invariant so $\operatorname{tr}(A) = \operatorname{tr}(A^\top)$, it follows that $\operatorname{tr}(A^\dagger) = \operatorname{tr}(A^*) = \operatorname{tr}(A)^*$. Hence,

$$\begin{aligned}
\chi_\rho(g^{-1}) = \chi_{\rho'}(g^{-1}) = \operatorname{tr}[\rho'(g^{-1})] = \operatorname{tr}[\rho'(g)^{-1}]&\\
= \operatorname{tr}[\rho'(g)^\dagger] = \operatorname{tr}[\rho'(g)]^* \eqqcolon \chi_{\rho'}(g)^* = \chi_\rho(g)^*.&
\end{aligned} \tag{9.1.38}$$

□

Lemma 9.1.39 Let G be a finite group and ρ a one-dimensional representation of G. Then $\chi_\rho(g) = \rho(g)$, where we make the natural identification of $\rho(g)$ as a 1×1 matrix with the matrix element $\rho(g)_{11}$.

Proof. This is trivially true since the trace of (z) is $\operatorname{tr}[(z)] = z$, so $\rho(g) = (z)$ for some $z \in \mathbb{C}$ and $\chi_\rho(g) = \operatorname{tr}[\rho(g)] = \operatorname{tr}[(z)] = z$. □

9.2 Space of Characters

We can view the characters of a given representation as forming vectors. Given a finite group, G, and a representation ρ we define a vector

$$\chi_\rho = (\chi_\rho(e), \chi_\rho(g_1), \dots, \chi_\rho(g_{N_c})) \in \mathbb{C}^{N_c} \tag{9.2.1}$$

where g_i is in the kth conjugacy class of G and N_c is the total number of conjugacy classes. It is simply a matter of convention to assign the first conjugacy class to be the one containing the identity.

Definition 9.2.2 — Inner product on the space of characters We can define an inner product on the space of characters as follows:

$$\langle \chi_{\rho_a}, \chi_{\rho_b} \rangle := \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g) = \frac{1}{|G|} \sum_{k=1}^{N_c} c_k \chi_{ak}^* \chi_{bk} \tag{9.2.3}$$

where c_k is the number of elements in the kth conjugacy class and $\chi_{ak} := \chi_{\rho_a}(g_k)$ with g_k being a representative member of the kth conjugacy class.

The following theorem gives us a useful way to test if a representation is irreducible.

Theorem 9.2.4 — First Orthogonality Theorem. The irreducible representations form an orthonormal set in the space of characters.

Proof. R This proof is beyond the scope of this course.

Suppose V_a and V_b are irreducible representations[a] of some finite group, G. Denote these representations by ρ_{V_a} and ρ_{V_b} respectively. The space of all homomorphisms $V_a \longrightarrow V_b$ is denoted $\mathrm{Hom}(V_a, V_b)$, and is equal to $V_a^* \otimes V_b$ where V_a^* is the dual vector space of V_a, we can think of this informally as going from kets to bras, or more formally as the space of linear functions, $V_a \longrightarrow \mathbb{F}$, where $\mathbb{F}$ is the base field of V_a.
Let V_0 denote the vector space of trivial representations, that is the vector space with the associated representation $\rho_{V_0}(g) = \mathbb{1}_{V_0}$ for all $g \in G$.
Schur's lemma (Theorem 8.4.8) states that the only nontrivial homomorphisms compatible with the group structure are isomorphisms. Hence, there are only nontrivial homomorphisms if $V_a \cong V_b$, projecting onto the V_0 we then have $\dim(\mathrm{Hom}(V_a, V_b)_0) = 1$. On the other hand if $V_a \not\cong V_b$ then $\dim(\mathrm{Hom}(V_a, V_b)) = 0$. We can sum this up as $\dim(\mathrm{Hom}(V_a, V_b)_0) = \delta_{ab}$.
The character of the representation $V_a^* \otimes V_b$ is given by

$$\chi_{\rho_{V_a^* \otimes V_b}} = \chi_{\rho_{V_a}} \chi_{\rho_{V_b}}, \tag{9.2.5}$$

which follows from Lemma 9.1.19 applied to each component of the vectors χ_ρ with component wise multiplication, that is $(\chi_\rho \chi_{\rho'})(g) = \chi_\rho(g)\chi_{\rho'}(g)$ for all $g \in G$. Further

$$\chi_{\rho_{V_a^*}} = \chi_{\rho_{V_a}}^* \tag{9.2.6}$$

since

$$\chi_{\rho_{V_a^*}}(g) = \mathrm{tr}[\rho_{V_a^*}(g)] = \mathrm{tr}[\rho_{V_a}(g)^\dagger] = \mathrm{tr}[\rho_{V_a}(g)]^* = \chi_{\rho_{V_a}}(g)^*, \tag{9.2.7}$$

which follows from the fact that the adjoint (Hermitian conjugate) of a vector is an element of the dual space and vice versa. We therefore have

$$\chi_{\rho_{V_a^* \otimes V_b}} = \chi_{\rho_{V_a}}^* \chi_{\rho_{V_b}}. \tag{9.2.8}$$

Now define the operator φ according to

$$\varphi = \frac{1}{|G|}\sum_{g\in G}\rho_V(g). \tag{9.2.9}$$

This is a projection operator, meaning $\varphi^2 = \varphi$, in particular it projects onto V_0. This means $\varphi v_0 = v_0$ for all $v_0 \in V_0$ and $\varphi v_0^\perp = 0$ where $v_0^\perp$ is an element of the orthogonal space to V_0, which we denote $V_0^\perp$ and is the space such that $V_0^\perp \oplus V_0 = V$, note that all vectors in $V_0^\perp$ are by construction orthogonal to all vectors in V_0.
To see that φ has these properties notice that applying φ to a vector gives an invariant vector and the only invariant vectors correspond to vectors in the trivial representation subspace. Further, applying φ a second time just rearranges the order of the vectors in the sum, which has no effect since we are averaging over the whole group. The dimension of the space projected onto by a projector is simply the trace of said projector, since we can write a projector in a diagonal form with 1 on the diagonal for basis vectors spanning the subspace and 0 for the other diagonal components. Using the linearity of the trace we then have

$$\dim(V_0) = \operatorname{tr}\varphi = \frac{1}{|G|}\operatorname{tr}[\rho_V(g)] = \frac{1}{|G|}\sum_{g\in G}\chi_{\rho_V}(g). \tag{9.2.10}$$

Combining all of the above we have

$$\begin{aligned}
\delta_{ab} &= \dim(\mathrm{Hom}(V_a, V_b)) &(9.2.11)\\
&= \dim((V_a^* \otimes V_b)_0) &(9.2.12)\\
&= \frac{1}{|G|}\sum_{g\in G}\chi_{\rho_{V_a^*\otimes V_b}}(g) &(9.2.13)\\
&= \frac{1}{|G|}\sum_{g\in G}\chi_{\rho_{V_a}}(g)^*\chi_{\rho_{V_b}}(g) &(9.2.14)\\
&=: \langle \chi_{\rho_{V_a}}, \chi_{\rho_{V_b}}\rangle. &(9.2.15)
\end{aligned}$$

This completes the proof. □

[a]Of course, we're being a bit sloppy with the language here, but this is standard, what we really mean is there exist irreducible representations, $G \to \mathrm{GL}(V_a)$ and $G \to \mathrm{GL}(V_b)$.

The important thing about the orthogonality theorem is all of the corollaries we can prove from it. In the following let ρ be a representation of some finite group G. We can write ρ as

$$\rho = \bigoplus_{i=1}^{k} m_i\rho_i = \bigoplus_{i=1}^{k}\rho_i^{\oplus m_i} = m_1\rho_i \oplus \cdots \oplus m_k\rho_k \tag{9.2.16}$$

where ρ_i are irreducible representations and $m_i \in \mathbb{Z}_{>0}$.

Corollary 9.2.17 A representation is fully characterised by its character since the distinct ρ_i correspond to linearly independent, in fact orthonormal, vectors in the character space. We have

$$\chi_\rho(g) = \sum_{i=1}^{k} m_i \chi_{\rho_i}(g). \tag{9.2.18}$$

Corollary 9.2.19 The multiplicity of some particular irreducible representation, ρ_a, in the decomposition of ρ is

$$m_a = \langle \chi_{\rho_a}, \chi_\rho \rangle \tag{9.2.20}$$

Proof. By the linearity of the character

$$\chi_\rho = \sum_{i=1}^{k} m_i \chi \rho_i \tag{9.2.21}$$

Then by the linearity of the inner product we have

$$\langle \chi_{\rho_a}, \chi_\rho \rangle = \left\langle \chi_{\rho_a}, \sum_{i=1}^{k} m_i \chi_{\rho_i} \right\rangle \tag{9.2.22}$$

$$= \sum_{i=1}^{k} m_i \langle \chi_{\rho_a}, \chi_{\rho_i} \rangle \tag{9.2.23}$$

$$= \sum_{i=1}^{k} m_i \delta_{ai} \tag{9.2.24}$$

$$= m_a \tag{9.2.25}$$

which completes the proof. □

Compare this to the standard way of finding the components of a vector:

$$v^i = \boldsymbol{e_i} \cdot \boldsymbol{v}. \tag{9.2.26}$$

Corollary 9.2.27 We can define a norm on the space of characters by

$$\|\chi_\rho\|^2 \coloneqq \langle \chi_\rho, \chi_\rho \rangle = \sum_i m_i^2 \tag{9.2.28}$$

and ρ is irreducible if and only if $\|\chi_\rho\| = 1$.

Proof. Suppose ρ is irreducible. Then we can think of ρ as being decomposed as $\rho = 1\rho$, that is $k = 1$ and $m_1 = 1$. Hence, we trivially have $\|\chi_\rho\| = 1$. Suppose instead that $\|\chi_\rho\| = 1$. Then it follows that $m_1^2 + \cdots + m_k^2 = 1$. Since $m_i \in \mathbb{Z}_{>0}$ it follows that $m_i = 0$ for all but one value of i and for that value of i $m_i = 1$, which means $\rho = 1\rho_i = \rho_i$, so ρ is irreducible. □

Lemma 9.2.29 The character of the identity element corresponds to the dimension of the representation. That is

$$\chi_{\rho_i}(e) = \dim V_i. \tag{9.2.30}$$

Proof. For any representation, $\rho\colon G \to V$ since ρ is a homomorphism we have $\rho(e) = \mathbb{1}_V$ by Lemma 2.1.19 and for any finite dimensional vector space, V, we have $\operatorname{tr} \mathbb{1}_V = \dim V$, since $\mathbb{1}_V$ is just a matrix with ones on the diagonal. □

Lemma 9.2.31 Let ρ_0 be the trivial representation of some finite group, G. Then $\chi_{\rho_0} = (1, \dots, 1) \in \mathbb{C}^{N_c}$.

Proof. This follows since the characters of the trivial representation are all one since $\chi_{\rho_0}(g) = \operatorname{tr}[\rho_0(g)] = \operatorname{tr} \mathbb{1}_{V_0}$ and the vector space of *the* trivial representation is one dimensional so $\operatorname{tr} \mathbb{1}_{V_0} = 1$. □

9.2.1 Dimensionality Theorem

Definition 9.2.32 — Regular Representation Let V be a vector space of dimension $|G|$. Let $\{\boldsymbol{e}_g\}$ be a set of $|G|$ linearly independent vectors in V which we label with group elements, $g \in G$. Such a set is guaranteed to exist since V is $|G|$-dimensional. We can think of V as the space spanned by these vectors. Define a left action of G on $\{\boldsymbol{e}_g\}$ in the obvious way by

$$g \,.\, \boldsymbol{e}_{g'} \coloneqq \boldsymbol{e}_{gg'}. \tag{9.2.33}$$

The **regular representation**, $\rho_{\mathrm{R}}\colon G \to \mathrm{GL}(V)$ is defined as the representation associated with this action, that is

$$\rho_{\mathrm{R}}(g)\boldsymbol{e}_{g'} = \boldsymbol{e}_{gg'}. \tag{9.2.34}$$

The regular representation is, in general, reducible, however it does have the following useful property:

$$\chi_{\rho_{\mathrm{R}}}(g) = \begin{cases} 0 & \text{if } g \neq e, \\ |G| & \text{if } g = e. \end{cases} \tag{9.2.35}$$

This follows since for $g \neq e$ the diagonal of $\rho_{\mathrm{R}}(g)$ must be zero, since if they weren't, it would mean that $\boldsymbol{e}_{gg} = \boldsymbol{e}_g$, which means that $g^2 = g$, which can only happen for $g = e$. Additionally $\rho_{\mathrm{R}}(e) = \mathbb{1}_V$ and so $\chi_{\mathrm{R}}(e) = \operatorname{tr} \mathbb{1}_V = \dim V = |G|$.

The regular representation can also be viewed as the induced representation of the trivial representation. Induced representations will be defined in .

The main use of the regular representation for us is to prove the following theorem which aids in classifying the irreducible representations.

Theorem 9.2.36 — Dimensionality Theorem. Let G be a group and $\{V_i\}$ be the irreducible representations of G, then

$$|G| = \sum_{i=1}^{N_c} \dim(V_i)^2. \tag{9.2.37}$$

Proof. Let ρ_R be the regular representation. This can be written as

$$\rho_R = \bigoplus_{i=1}^{k} \rho_i^{\oplus m_i}. \tag{9.2.38}$$

Notice that we then have

$$m_i = \langle \chi_{\rho_i}, \chi_{\rho_R} \rangle \tag{9.2.39}$$

$$=: \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_i}(g)^* \chi_{\rho_R}(g) \tag{9.2.40}$$

$$= \chi_{\rho_i}(e)^* \tag{9.2.41}$$

$$= \dim V_i. \tag{9.2.42}$$

Here we have used the fact that $\chi_{\rho_R}(g) = 0$ for $g \neq e$ and so all terms in the sum vanish apart from the $g = e$ term. The regular representation contributes a factor of $|G|$ to this term, which cancels with the existing normalisation factor to leave just $\chi_{\rho_i}(e)^*$, since $g = e$ in this term. We then apply Lemma 9.2.29 to get $\chi_{\rho_i}(e) = \dim V_i$, and since this is real the complex conjugate does nothing.
Using $m_i = \dim V_i$ we then have

$$\chi_{\rho_R}(g) = \sum_{i=1}^{k} m_i \chi_{\rho_i}(g) = \sum_{i=1}^{k} \dim(V_i) \chi_{\rho_i}(g). \tag{9.2.43}$$

Considering the specific case of $g = e$ this then becomes

$$|G| = \chi_{\rho_R}(e) = \sum_{i=1}^{k} \dim(V_i) \chi_{\rho_i}(e) = \sum_{i=1}^{k} \dim(V_i)^2 \tag{9.2.44}$$

where we have used Lemma 9.2.29 again in the last equality. □

The dimensionality theorem quickly allows us to limit the possible irreducible representations. For example, S_3 ($|S_3| = 3! = 6$) could have either 6 one-dimensional irreducible representations ($6 \cdot 1^2 = 6$) or 2 one-dimensional irreducible representations and 1 two-dimensional irreducible representation ($2 \cdot 1^2 + 1 \cdot 2^2 = 6$). It turns out that the latter is correct.

9.3 Character Tables

The common way to give the information on the characters of the irreducible representations is as a **character table**. Since the character is a class function, that is, it is the same for all group elements sharing a conjugacy class, we only write out the conjugacy classes for the character table, rather than the full group. We then have an optional line for the order of the classes. The rows below then list the characters of the irreducible representations on the conjugacy classes. Typically, we list the conjugacy class containing the identity first and the trivial representation (one dimensional representation given by the trivial action) first which means that the first row contains all ones.

As an example we now give the character table of S_3, but there are some details here we have yet to discuss.

S_3 classes	$[()]$	$[(1\,2)]$	$[(1\,2\,3)]$
Order	1	3	2
Trivial	1	1	1
Alternating	1	-1	1
Standard	2	0	-1

(9.3.1)

The labels down the side label the irreducible representations, we haven't yet defined the alternating or standard representations so don't worry too much about them.

In this section we will write statements like "the character table is square". What we mean by the character table here is the numbers that we fill in for each conjugacy class and representation ignoring the labels that we give along the edges.

Theorem 9.3.2. The character table is square.

Proof. R This proof is beyond the scope of this course.

The total number of irreducible representations must be at most the number of conjugacy classes since the irreducible representations form a basis in the space of class functions, which is a space of dimension N_c.
Suppose there is a class function, $f\colon G \to \mathbb{C}$, such that f is orthogonal to all of the characters of the irreducible representations, that is

$$\langle f, \chi_{\rho_i} \rangle = 0 \tag{9.3.3}$$

for all irreducible representations, ρ_i. If we can show that this necessarily means $f = (0, \dots, 0) \in \mathbb{C}^{N_c}$ then this shows that the characters form a complete set of vectors on the class function space, and hence the number of characters is equal to the dimension of the space.
Consider the map

$$\varphi\colon V_i \to V_i \qquad \text{where} \qquad \varphi := \sum_{g \in G} f(g)^* \rho_i(\gamma). \tag{9.3.4}$$

Here V_i is the vector space associated with the irreducible representation ρ_i.

It can easily be seen that

$$\varphi\rho_i(g') = \sum_{g\in G} f(g)^*\rho_i(g)\rho_i(g') \tag{9.3.5}$$

$$= \sum_{g\in G} f(g)^*\rho_i(gg') \tag{9.3.6}$$

$$= \sum_{g''\in G} f(g'')^*\rho_i(g'') \tag{9.3.7}$$

$$= \sum_{g'''\in G} \rho_i(g')f(g''')^*\rho_i(g''') \tag{9.3.8}$$

$$= \rho_i(g') \sum_{g'''\in G} f(g''')^* \tag{9.3.9}$$

where we use the fact that we are averaging over the group so averaging gg' over g is the same as averaging over $g'' = gg'$. We then have by Schur's lemma (Theorem 8.4.8) that $\varphi = \lambda\mathbb{1}$. Hence,

$$\operatorname{tr}\varphi = \lambda \dim V_i \tag{9.3.10}$$

and

$$\operatorname{tr}\varphi = \sum_{g\in G} f(g)^*\chi_{\rho_i}(g) = |G|\langle f, \chi_{\rho_i}\rangle = 0. \tag{9.3.11}$$

The last equality being by our assumption that f is orthogonal to all of the characters. We therefore must have that either $\lambda = 0$, in which case $f(g) = 0$ for all $g \in G$, or

$$\sum_{g\in G} f^*(g)\rho_i(g) = 0 \tag{9.3.12}$$

holds for all irreducible representations, and hence for all representations since they can be written as a sum of irreducible representations.
In particular this must hold for the regular representation, ρ_{R}. However, since $\rho_{\mathrm{R}}(g)$ are all linearly independent by definition this means $f(g) = 0$ for all $g \in G$. Either way the result is that $f = (0, \dots, 0)$ and so the characters form a complete basis and hence the number of irreducible representations is equal to the number of conjugacy classes. □

The following theorem is useful, but the proof requires somewhat complicated linear algebra, so we omit it.

Theorem 9.3.13. This dimension of any irreducible representation divides the order of the group. That is $|G|/\dim V_i \in \mathbb{Z}_{>0}$ where $\rho_i\colon G \to \mathrm{GL}(V_i)$ is an irreducible representation.
Further $\dim V_i$ divides $G/Z(G)$, where $Z(G)$ is the centre of the group, which is the normal subgroup of all commuting elements.

Corollary 9.3.14 All irreducible representations of an Abelian group are one dimensional.

Proof. Let G be an Abelian group. Since the group is Abelian $gg'g^{-1} = g'gg^{-1} = g'$ for all $g, g' \in G$, and so every conjugacy class contains exactly one element. Hence, $N_c = |G|$. By the dimensionality theorem (Theorem 9.2.36)

$$|G| = \sum_{i=1}^{N_c} \dim(V_i)^2 = \sum_{i=1}^{|G|} \dim(V_i)^2. \tag{9.3.15}$$

Since $\dim V_i$ divides $|G|$ by Theorem 9.3.13 it follows that $\dim V_i \neq 0$ and so $\dim V_i = 1$ is the only way to have this equation hold. □

The past few theorems combined shows that we know quite a lot about the dimensions of the irreducible representations before we even start to work out what the representations are. This mostly stems from Schur's lemma. We can sum it up in a set of diophantine equations (equations with integer solutions) which must be satisfied.

- $|G| = 1 + \dim(V_2)^2 + \cdots + \dim(V_k)^2$,
- $k = N_c$, and
- $|G|/\dim V_i \in \mathbb{Z}_{>0}$.

The 1 in the first equation corresponds to the dimension of the trivial representation, which is always present. These equations are often enough to allow us to work out the dimensions of the irreducible representations without having to explicitly compute them.

Definition 9.3.16 — Inner product on the space of classes We can define an inner product on the class space as

$$\langle \chi_\rho([g]), \chi_\rho([g']) \rangle_c = \frac{c_{[g]}}{|G|} \sum_{i=1}^{N_c} \chi_{\rho_i}([g])^* \chi_{\rho_i}([g']). \tag{9.3.17}$$

Here $c_{[g]}$ is the size of the conjugacy class containing g, ρ_i are the irreducible representations appearing in the decomposition of ρ, and $\chi_\rho([g])$ is defined to be $\chi_\rho(g)$, which is well defined since χ is a class function.

The asymmetry in the above definition, that there is a scale factor of $c_{[g]}$ but note $c_{[g']}$ is in anticipation of the following theorem.

Theorem 9.3.18 — Second Orthogonality Theorem. The conjugacy classes are orthonormal with respect to the above inner product. That is

$$\langle \chi_\rho([g]), \chi_\rho([g']) \rangle_c = \delta_{[g][g']}. \tag{9.3.19}$$

Proof. The proof relies on the fact that for a finite unitary $n \times n$ matrix, U, we have

$$UU^\dagger = U^\dagger U = \mathbb{1}. \tag{9.3.20}$$

The first equation can be taken as the definition of unitarity and the second follows from taking the Hermitian conjugate.
Define the matrix U to have components

$$U_{a[g]} = \sqrt{\frac{c_{[g]}}{|G|}} \chi_{\rho_a}([g]). \tag{9.3.21}$$

This will give a square matrix since by Theorem 9.3.2 there are N_c irreducible representations, ρ_a, and N_c conjugacy classes, $[g]$.
Now consider the product $UU^\dagger$:

$$(UU^\dagger)_{ab} = \sum_{[g]\in G/\sim} U_{a[g]}(U^\dagger)_{[g]b} \tag{9.3.22}$$

$$= \sum_{g\in G/\sim} U_{a[g]}U^*_{b[g]} \tag{9.3.23}$$

$$= \sum_{g\in G/\sim} \sqrt{\frac{c_{[g]}}{|G|}} \chi_{\rho_a}([g]) \sqrt{\frac{c_{[g]}}{|G|}} \chi_{\rho_b}([g])^* \tag{9.3.24}$$

$$= \frac{1}{|G|} \sum_{k=1}^{N_\mathrm{c}} c_k \chi_{\rho_a}(g_k)\chi_{\rho_b}(g_k)^* \tag{9.3.25}$$

$$= \frac{1}{|G|} \sum_{k=1}^{N_\mathrm{c}} c_k \chi_{ak}\chi^*_{bk} \tag{9.3.26}$$

$$= \delta_{ab}. \tag{9.3.27}$$

Note that $G/\sim = \{[g] \mid g \in G\}$ is the set of conjugacy classes. In the penultimate step we have rewritten in the notation of Definition 9.2.2. The final step is then the result of the first orthogonality theorem (Theorem 9.2.4). This shows that U is a unitary matrix.
We now consider the product $U^\dagger U$, which we know is $\mathbb{1}$ since U is unitary. We have

$$(U^\dagger U)_{[g][g']} = \sum_{a=1}^{N_\mathrm{c}} (U^\dagger)_{[g]a} U_{a[g']} \tag{9.3.28}$$

$$= \sum_{a=1}^{N_\mathrm{c}} U^*_{a[g]} U_{a[g']} \tag{9.3.29}$$

$$= \sum_{a=1}^{N_\mathrm{c}} \sqrt{\frac{c_{[g]}}{|G|}} \chi_{\rho_a}([g])^* \sqrt{\frac{c_{[g']}}{|G|}} \chi_{\rho_a}([g']) \tag{9.3.30}$$

$$= \sqrt{\frac{c_{[g']}}{c_{[g]}}} \frac{c_{[g]}}{|G|} \sum_{a=1}^{N_\mathrm{c}} \chi_{\rho_a}([g])^* \chi_{\rho_a}([g']) \tag{9.3.31}$$

$$= \sqrt{\frac{c_{[g']}}{c_{[g]}}} \langle \chi_\rho([g]), \chi_\rho([g']) \rangle_c \tag{9.3.32}$$

$$= \delta_{[g][g']}. \tag{9.3.33}$$

The final equality holds since U is unitary so $(U^\dagger U)_{[g][g']} = \delta_{[g][g']}$. When $[g] = [g']$ we have

$$\sqrt{\frac{c_{[g]}}{c_{[g']}}} = 1 \implies \langle \chi_\rho([g]), \chi_\rho([g]) \rangle_c = 1, \tag{9.3.34}$$

and when $[g] \neq [g']$ we have

$$\langle \chi_\rho([g]), \chi_\rho([g']) \rangle_c = 0 \tag{9.3.35}$$

and so we have

$$\langle \chi_\rho([g]), \chi_\rho([g']) \rangle_c = \delta_{[g][g']} \tag{9.3.36}$$

completing the proof. □

9.4 Constructing Character Tables

Often we have enough information to complete the character table using the diophantine equations and orthogonality theorems. We will demonstrate the process with S_3. First we need a few results about the conjugacy classes of S_n. The lemmas just ensure that "cycle type" is well defined and that the first corollary, which is the statement we care about, holds.

Lemma 9.4.1 Every permutation in S_n has a cycle decomposition which is unique up to the ordering of the cycles and cyclic permutations of the elements within each cycle.

Proof. Let $\sigma \in S_n$ act on $X = (1, \dots, n)$ by permutation. Define $G = \langle \sigma \rangle$ to be the subgroup of S_n generated by σ. Then G acts on X by restricting the group action of S_n. By the orbit-stabiliser theorem (Theorem 3.2.15) this results in a partition of X into unique sets of orbits. For any orbit $\mathrm{Orb}_G(x)$ we then have a bijection associating $\sigma . x$ and $\sigma \,\mathrm{Stab}_G(x)$.
Since G is cyclic (as it is generated by a single element) it follows that $G/\mathrm{Stab}_G(x)$ is cyclic. Its order is the smallest positive integer, d, such that $\sigma^d \in \mathrm{Stab}_G(x)$. We know that $d = |\mathrm{Stab}(x)| = [G : \mathrm{Stab}_G(x)]$ and so with the aforementioned bijection we have that the unique cosets of $\mathrm{Stab}_G(x)$ in G are

$$\mathrm{Stab}_G(x), \sigma\,\mathrm{Stab}_G(x), \dots, \sigma^{d-1}\,\mathrm{Stab}_G(x). \tag{9.4.2}$$

The elements of $\mathrm{Orb}_G(x)$ are then

$$x, \sigma . x, \dots, \sigma^{d-1} . x. \tag{9.4.3}$$

Therefore on any orbit of size d σ is a d-cycle. This shows the existence of a cycle decomposition.
Uniqueness is then fairly simple. Each cycle determined by σ on an element of order d is determined uniquely by construction from x. Choosing a different element in the same orbit, say $\sigma^j x$ instead gives

$$\sigma^j \,.\, x, \sigma^{j+1} \,.\, x, \dots, \sigma^{d-1} \,.\, x, x, \sigma \,.\, x, \dots, \sigma^{j-1} \,.\, x. \tag{9.4.4}$$

This is the same cycle permuted left by j. □

Definition 9.4.5 — Cycle Type Given $\sigma \in S_n$ we can write σ as the product of k disjoint cycles of lengths $n_1, \dots, n_k$, where, since disjoint cycles commute, we can take $n_i \le n_{i+1}$. We then define $n_1, \dots, n_k$ to be the **cycle type** of σ. This is well defined by the uniqueness part of the previous lemma.

Lemma 9.4.6 Two permutations are conjugate if and only if they have the same cycle type.

Proof. Let $\sigma, \tau \in S_n$. Suppose σ and τ are conjugate. Then there exists $\rho \in S_n$ such that $\tau = \rho\sigma\rho^{-1}$. We can write σ as a product of disjoint cycles. To show that σ and τ have the same cycle type it suffices to show that if j follows i in the cycle decomposition of σ then $\rho(j)$ follows $\rho(i)$ in the cycle decomposition of τ. Suppose that $\rho(i) = j$. Then

$$\tau(\rho(i)) = \rho\sigma\rho^{-1}(\rho(i)) = \rho\sigma(i) = \rho(j) \tag{9.4.7}$$

and so $\rho(j)$ does indeed follow $\rho(i)$. This proves that conjugate elements of S_n have the same cycle type.
Suppose instead that σ and τ have the same cycle type. We can then write them as products of disjoint cycles in increasing cycle length, including 1-cycles. So, for example,

$$\sigma = (a_1)(a_2\, a_3\, a_4)(a_5\, \dots\, a_n), \qquad \text{and} \qquad \tau = (b_1)(b_2\, b_3\, b_4)(b_5\, \dots\, b_n). \tag{9.4.8}$$

Define ρ to be the permutation taking a_i to b_i. Since the cycle types match we have

$$\rho\sigma\rho^{-1}(b_i) = \rho\sigma(a_i) = \rho(a_j) = b_j. \tag{9.4.9}$$

With a_j and b_j being the elements after a_i and b_i in their respective cycles. We therefore have $\tau = \rho\sigma\rho^{-1}$ and so we have proven σ and τ are conjugate. □

Definition 9.4.10 — Integer Partition Given some positive integer, $n \in \mathbb{Z}_{>0}$ a **integer partition** is a sum $m_1 + \dots + m_k = n$, with $m_i \in \mathbb{Z}_{>0}$ where without loss of generality we can assume $m_i \le m_{i+1}$.

Corollary 9.4.11 The number of conjugacy classes in S_n is the number of partitions of n.

Proof. Each distinct cycle type in S_n represents a distinct partition of n, and each cycle type represents a conjugacy class since elements of S_n are conjugate if and only if they have the same conjugacy class. □

Corollary 9.4.12 Let $m_1, m_2, \dots, m_r$ be the distinct integers appearing cycle type of some permutation $\sigma \in S_n$ and let there be k_i cycles of order m_i. Then

$$|[\sigma]| = n! \prod_{i=1}^{r} \frac{1}{k_i! m_i^{k_i}}. \tag{9.4.13}$$

Consider a given arrangement of $1, \dots, n$. There are $n!$ such arrangements. For each arrangement we there are k_i cycles of order m_i and each can be written in m_i different ways by starting with a different element in the cycle. The m_i-cycles can appear in any of $k_i!$ possible orders since they are disjoint and so commute. This means that we have $n!$ overall arrangements but of these $k_i! m_i^{k_i}$ are equivalent and result only in m_i-cycles changing. The main result follows by accounting for cycles of all orders.

9.4.1 Character Table of S_3

The number 3 has 3 distinct partitions. The first is $3 = 1 + 1 + 1$, this corresponds to the cycle type $(a)(b)(c) = ()$, and so to the conjugacy class $[()]$. The second is $3 = 1 + 2$, this corresponds to the cycle type $(a)(b\,c) = (b\,c)$, and so to the conjugacy class $[(1\,2)]$. The third is $3 = 3$, this corresponds to the cycle type $(a\,b\,c)$, and so to the conjugacy class $[(1\,2\,3)]$.

As always we have the trivial representation, which is one dimensional. In this section, and much of the rest of the text, we will follow the convention of labelling representations by their dimension. The trivial representation is one-dimensional, so we call it $\mathbf{1}$. The trivial representation is $\mathbf{1} \colon S_3 \to \mathrm{GL}(\mathbb{C})$, defined by $\mathbf{1}(\sigma) = 1$ for all $\sigma \in S_3$. This has character $\chi_1(\sigma) = \mathrm{tr}[\mathbf{1}(\sigma)] = \mathrm{tr}[1] = 1$. The first row of the character table is thus

S_3	$[()]$	$[(1\,2)]$	$[(1\,2\,3)]$
$\lvert[g]\rvert$	1	3	2
$\mathbf{1}$	1	1	1

(9.4.14)

For the permutation group there is a second one-dimensional representation, $\mathbf{1}'$ that we can define.

Definition 9.4.15 — Alternating Representation The **alternating representation**, ρ_{alt} or $\mathbf{1}'$, is defined to be the representation $\rho_{\mathrm{alt}} \colon S_n \to \mathrm{GL}(\mathbb{C})$ such that $\rho_{\mathrm{alt}}(\sigma) = \mathrm{sgn}(\sigma)$. Here $\mathrm{sgn} \colon S_n \to \{\pm 1\}$ is the sign of the permutation.

The character of $\mathbf{1}'$ is simply $\chi_{\mathbf{1}'}(\sigma) = \mathrm{tr}[\mathbf{1}'(\sigma)] = \mathrm{tr}[\mathrm{sgn}(\sigma)] = \mathrm{sgn}(\sigma)$. Allowing us to add a second row to the character table:

$$\begin{array}{c|ccc} S_3 & [()] & [(1\,2)] & [(1\,2\,3)] \\ |[g]| & 1 & 3 & 2 \\ \hline \mathbf{1} & 1 & 1 & 1 \\ \mathbf{1}' & 1 & -1 & 1 \end{array} \tag{9.4.16}$$

Here we have used the fact that $() = (1\,2)(1\,2)$ is an even permutation, 2-cycles are odd permutations, and 3-cycles can be written as a product of two 2-cycles, so are even permutations.

Since S_3 has 3 conjugacy classes it also has 3 irreducible representations. The simplest way to work out the final row of the character table is to ensure that this third representation is orthogonal to the trivial and alternating representations. To do so we need to know the dimension of this representation. For this we turn to the diophantine equation

$$|G| = 1 + \dim(V_2)^2 + \cdots + \dim(V_k)^2. \tag{9.4.17}$$

For S_3 with the existing information this gives us

$$|S_3| = 6 = 1 + 1^2 + n^2 \implies n = 2 \tag{9.4.18}$$

so our final representation is $\mathbf{2}$, which we call the **standard representation** of S_3. More generally for S_n the standard representation is such that $\rho_{\mathrm{triv}} \oplus \rho_{\mathrm{stand}} = \rho_{\mathrm{perm}}$.

We now use the fact that the permutation representation, which is 3-dimensional, must decompose as a sum of irreducible representations. There are multiple ways that this could occur, it could be three copies of the trivial and alternating representations, or one copy of the trivial or alternating representation and one copy of $\mathbf{2}$. It turns out to be the latter.

We then use the simple formula

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2} \tag{9.4.19}$$

and $\chi_{\mathbf{1}} = (1, 1, 1)$ and $\chi_{\rho_{\mathrm{stand}}} = (3, 1, 0)$, which can be seen easily by either writing out all matrices in the standard representation or by noting that in the permutation representation the trace of a matrix is the number of ones on the diagonal, which is the number of elements left invariant by the permutation and the identity leaves all 3 elements invariant while 2-cycles leave 1 element invariant and 3-cycles change all elements. Hence, we have

$$\chi_{\rho_{\mathrm{stand}}} = \chi_{\mathbf{1}\oplus\mathbf{2}} = \chi_{\mathbf{1}} + \chi_{\mathbf{2}} \implies \chi_{\mathbf{2}} = (3, 1, 0) - (1, 1, 1) = (2, 0, -1). \tag{9.4.20}$$

This allows us to fill out the final line of our character table:

$$\begin{array}{c|ccc} S_3 & [()] & [(1\,2)] & [(1\,2\,3)] \\ |[g]| & 1 & 3 & 2 \\ \hline \mathbf{1} & 1 & 1 & 1 \\ \mathbf{1}' & 1 & -1 & 1 \\ \mathbf{2} & 2 & 0 & -1 \end{array} \tag{9.4.21}$$

9.4.2 Character Table of S_4

Four has the following partitions with the corresponding conjugacy classes

$$4 = 1 + 1 + 1 + 1 \leftrightarrow [()], \tag{9.4.22}$$
$$4 = 1 + 1 + 2 \leftrightarrow [(1\,2)], \tag{9.4.23}$$
$$4 = 2 + 2 \leftrightarrow [(1\,2)(3\,4)], \tag{9.4.24}$$
$$4 = 1 + 3 \leftrightarrow [(1\,2\,3)], \tag{9.4.25}$$
$$4 = 4 \leftrightarrow [(1\,2\,3\,4)] \tag{9.4.26}$$

These are of size 1, 6, 3, 8, and 6 respectively.

The dimensions of the irreducible representations are the solutions to

$$|S_4| = 24 = 1 + 1^2 + d_3^2 + d_4^2 + d_5^2 \tag{9.4.27}$$

where the first two terms correspond to the dimensions of the trivial and alternating representations. We stop at 5 since there are 5 conjugacy classes, and hence 5 irreducible representations. The solutions to this are fairly easy to find, we have $d_3 = 2$, $d_4 = 3$, and $d_5 = 3$.

Filling in the first two rows of the table we have

S_4	$[()]$	$[(1\,2)]$	$[(1\,2)(3\,4)]$	$[(1\,2\,3)]$	$[(1\,2\,3\,4)]$
$\lvert[g]\rvert$	1	6	3	8	6
1	1	1	1	1	1
1′	1	−1	1	1	−1

(9.4.28)

The standard representation for S_4 is 3-dimensional, since $\rho_{\text{stand}} \oplus \rho_{\text{triv}} = \rho_{\text{perm}}$ and the permutation representation is 4-dimensional, with the character given by the number of elements left invariant. We have that

$$\chi_{1\oplus 3} = \chi_1 + \chi_3 \implies \chi_3 = (4, 2, 0, 1, 0) - (1, 1, 1, 1, 1) = (3, 1, -1, 0, -1). \tag{9.4.29}$$

There is another representation which corresponds to multiplying the standard representation by the alternating representation, giving

$$\chi_{3'}(\sigma) = \chi_{1'\otimes 3}(\sigma) = \chi_{1'}(\sigma)\chi_3(\sigma) = \operatorname{sgn}(\sigma)\chi_3(\sigma) \implies \chi_{3'} = (3, -1, -1, 0, 1). \tag{9.4.30}$$

Putting these in the table we have

S_4	$[()]$	$[(1\,2)]$	$[(1\,2)(3\,4)]$	$[(1\,2\,3)]$	$[(1\,2\,3\,4)]$
$\lvert[g]\rvert$	1	6	3	8	6
1	1	1	1	1	1
1′	1	−1	1	1	−1
2	·	·	·	·	·
3	3	1	−1	0	−1
3′	3	−1	−1	0	1

(9.4.31)

We can find the character of the final irreducible representation, **2**, by requiring that it is orthogonal to the other irreducible representations.

Let $\chi_2 = (a, b, c, d, e)$. We know that $\chi_2(()) = 2$, since the character of an irreducible representation on the identity gives the dimension of the representation. The second orthogonality theorem gives

$$\langle \chi_2([()]), \chi_2([(1\,2)]) \rangle = \frac{1}{24}\sum_{i=1}^{5} \chi_{\rho_i}(())^* \chi_{\rho_i}((1\,2)) \tag{9.4.32}$$
$$= \frac{1}{24}(1\cdot 1 + 1\cdot(-1) + 2\cdot b + 3\cdot 1 + 3\cdot(-1)) \tag{9.4.33}$$
$$= \frac{b}{12} \tag{9.4.34}$$
$$= 0 \tag{9.4.35}$$

so we have $b = 0$.

The first orthogonality theorem gives

$$\langle \chi_1, \chi_2 \rangle = \frac{1}{24}\sum_{k=1}^{5} c_k \chi_1(g_k)^* \chi_2(g_k) \tag{9.4.36}$$
$$= \frac{1}{24}(1\cdot 1\cdot 2 + 6\cdot 1\cdot 0 + 3\cdot 1\cdot c + 8\cdot 1\cdot d + 6\cdot 1\cdot e) \tag{9.4.37}$$
$$= \frac{1}{24}(2 + 3c + 8d + 6e) \tag{9.4.38}$$
$$= 0. \tag{9.4.39}$$

This has as a solution $c = 2$, $d = -1$, and $e = 0$. Thus the complete character table is

S_4	$[()]$	$[(1\,2)]$	$[(1\,2)(3\,4)]$	$[(1\,2\,3)]$	$[(1\,2\,3\,4)]$
$\lvert[g]\rvert$	1	6	3	8	6
1	1	1	1	1	1
1′	1	−1	1	1	−1
2	2	0	2	−1	0
3	3	1	−1	0	−1
3′	3	−1	−1	0	1

(9.4.40)

9.5 Complex, Real, and Pseudo-Real Representations

Definition 9.5.1 — Complex, Real, and Pseudo-Real Representations Let $\rho\colon G \to \mathrm{GL}(V)$ be a representation of the group G on the vector space V. If ρ is inequivalent to its complex conjugate, $\overline{\rho} = \rho^*$, defined by $\overline{\rho}(g) = \overline{\rho(g)}$, then we say that ρ is a **complex representation**.
If there exists a basis such that the components of $\rho(g)$ are real for all $g \in G$ then we say that ρ is a **real representation**.
Surprisingly, there is a third type representation. It turns out that it is possible for ρ to be equivalent to $\overline{\rho}$ and yet there be no basis such that the elements of $\rho(g)$ are always real. In this case we call ρ a **pseudo-real representation** or **quaternionic representation**.

An alternative definition of these three types is as follows. A **complex representation** is a group homomorphism, $\rho\colon G \to \mathrm{GL}(V, \mathbb{C})$.

Both real and pseudo-real representations are isomorphic to their conjugates. This forces the existence of an equivariant anti-linear map, $j\colon V \to V$, which is simply the isomorphism composed with conjugation. **Equivariant** means that if the group H acts on V then $h\,.\,j(v) = j(h\,.\,v)$ for all $h \in H$ and $v \in V$. It should be noted that by Schur's lemma (Theorem 8.4.8) the only equivariant maps $V \to V$ square to a multiple of the identity, $\lambda \mathbb{1}_V$, and if $\lambda > 0$ we say j is a real form, and if $\lambda < 0$ we say j is a quaternionic form.

A **real representation** is a group homomorphism, $\rho \colon G \to \mathrm{GL}(V, \mathbb{R})$ with an antilinear equivariant map, $j \colon V \to V$, such that $j^2 = +1$. **Antilinear** means that $j(\lambda v) = \lambda^* j(v)$ for all $v \in V$ and $\lambda \in \mathbb{C}$, as well as $j(v + w) = j(v) + j(w)$ for $v, w \in V$. A **pseudo-real representation** is a group homomorphism, $\rho \colon G \to \mathrm{GL}(V, \mathbb{R})$ with an antilinear equivariant map, $j \colon V \to V$, such that $j^2 = -1$.

This structure on V for a pseudo-real representation makes V a quaternionic vector space, which is to say an $\mathbb{H}$-module. Recall that an R-module is what we get if we replace a field in the vector space definition with a ring, R, in this case the division ring of quaternions. We can instead view a **pseudo-real representation** as a group homomorphism, $\rho \colon G \to \mathrm{GL}(V, \mathbb{H})$.

Example 9.5.2 The most obvious example of a pseudo-real or quaternionic representation is the quaternion group,

$$Q := \langle -e, I, J, K \mid (-e)^2 = e, I^2 = J^2 = K^2 = IJK = -e \rangle. \tag{9.5.3}$$

It just so happens that Q has five irreducible representations, three of which are one-dimensional and one of which is two-dimensional.
The two-dimensional representation is pseudo-real. This representation is defined by $\rho_2(e) = \mathbb{1}$, $\rho_2(I) = i\sigma_1$, $\rho_2(J) = i\sigma_2$, and $\rho_2(K) = i\sigma_3$, where σ_i are the **Pauli matrices** defined by

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \text{and} \qquad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.5.4}$$

Recall that $\sigma_i^2 = \mathbb{1}$ and so $(i\sigma_i)^2 = -\mathbb{1}$.
The fact that this is a pseudo-real representation means that there is no basis in which all three Pauli matrices have real components.
There is a four-dimensional real reducible representation of Q given by

$$\rho(e) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \rho(I) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{9.5.5}$$

$$\rho(J) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \qquad \rho(K) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{9.5.6}$$

We can define two quantities, A and B, according to

$$A(\rho) := \langle \chi_{\overline{\rho}}, \chi_\rho \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g)^2, \tag{9.5.7}$$

and

$$B(\rho) := \frac{1}{|G|}\sum_{g\in G}\chi_\rho(g^2). \tag{9.5.8}$$

It turns out that for irreducible representations the values of these depend only on the type of the representation according to

Type	$A(\rho)$	$B(\rho)$
Complex	0	0
Real	1	1
Pseudo-Real	1	-1

(9.5.9)

The first column follows from the fact that for (pseudo-)real representations $\rho \sim \overline{\rho}$ whereas for complex representations $\rho \nsim \overline{\rho}$ and so the first orthogonality theorem (Theorem 9.2.4) gives this result. If you get anything other than 0 or 1 for the first result then the representation isn't irreducible. For the second column we can identify B as being related to the aforementioned equivariant anti-linear map, j.

▪ **Example 9.5.10** For the pseudo-real two-dimensional representation of the quaternion group we have

$$A(\rho_2) = \frac{1}{|G|}\sum_{g\in G}\chi_2(g)^2\frac{1}{8}(4\cdot 2) = 1 \tag{9.5.11}$$

which follows since $\chi_{\rho_2}(g) = 0$ for $g = I, J, K$ and $\chi_{\rho_2}(g) = \pm 2$ for $g = \pm e$. Similarly

$$B(\rho_2) = \frac{1}{|G|}\sum_{g\in G}\chi_2(g) = \frac{1}{8}(2\cdot 2 + 6(-2)) = -1 \tag{9.5.12}$$

which follows from the fact that $e^2 = (-e)^2 = e$ and $I^2 = J^2 = K^2 = -e$ which have characters $\chi_2(e) = 2$ and $\chi_2(-e) = -2$.

The reason that pseudo-real representations are a possibility is due to the fact that we can decompose $V\otimes V$ into a symmetric and antisymmetric part:

$$V\otimes V = \mathrm{Sym}(V\otimes V)\oplus \mathrm{Asym}(V\otimes V). \tag{9.5.13}$$

If V is the vector space for a non-complex representation and the trivial representation falls into the $\mathrm{Sym}(V\otimes V)$ part of the decomposition then the representation is real. On the other hand if the trivial representation falls in the $\mathrm{Asym}(V\otimes V)$ part then it is pseudo-real. The antisymmetric nature results in getting -1 instead of 1 for $B(\rho)$.

9.6 Branching Rules

Let G be a finite group and H a subgroup of G. If $\rho\colon G \to \mathrm{GL}(V)$ is a representation then $\rho|_H$ is certainly a representation, where by $\rho|_H$ we mean ρ restricted to H, so $\rho|_H\colon H \to \mathrm{GL}(V)$ is defined by $\rho|_H(h) = \rho(h)$ for all $h \in H$. Crucially it is possible that ρ is irreducible but $\rho|_H$ is reducible. The decomposability theorem then means we can decompose $\rho|_H$ into a sum of irreducible representations of H as follows:

$$\rho|_H(h) = m_1\rho_1(h) \oplus \cdots \oplus m_k\rho_k(h) \tag{9.6.1}$$

where ρ_i are irreducible representations of H and $m_i \in \mathbb{N}$ and k is the number of conjugacy classes of H.

In physics we often call such decompositions **branching rules**. It is possible to construct branching rules systematically from the character tables for G and H. Often it is possible to do so with even less information.

■ **Example 9.6.2** Consider the case of $G = S_3$ and $H = \mathbb{Z}_3$, which we view as a subgroup of S_3 by identifying $\mathbb{Z}_3 = \langle (1\,2\,3) \mid (1\,2\,3)^3 = ()\rangle$. Note that the transpositions do not appear in $\mathbb{Z}_3$. The character table for S_3 is

S_3	$[()]$	$[(1\,2)]$	$[(1\,2\,3)]$
$\lvert[g]\rvert$	1	3	2
$\mathbf{1}$	1	1	1
$\mathbf{1}'$	1	-1	1
$\mathbf{2}$	2	0	-1

(9.6.3)

Since $\mathbb{Z}_3$ is Abelian all of its irreducible representations are one-dimensional, and each element is in a conjugacy class of its own, so there are three conjugacy classes. All representations map the identity to 1. Clearly there is the trivial representation, $\rho_{\mathbf{1}}(h) = 1$ for all $h \in \mathbb{Z}_3$. Another representation is given by $\rho_{\mathbf{1}'}(1) = \mathrm{e}^{2\pi i/3}$ and $\rho_{\mathbf{1}'}(2) = \mathrm{e}^{2\pi i 2/3} = \mathrm{e}^{-2\pi i/3}$, where we use $\mathbb{Z}_3$ as the group of integers under addition modulo 3, and notice that this representation gives the other familiar definition of $\mathbb{Z}_3$ as roots of unity. The third and final representation of $\mathbb{Z}_3$ is defined by $\rho_{\mathbf{1}''}(1) = \mathrm{e}^{-2\pi i/3}$ and $\rho_{\mathbf{1}''}(2) = \mathrm{e}^{2\pi i/3}$. This allows us to complete the character table for $\mathbb{Z}_3$:

$\mathbb{Z}_3$	$[()]$	$[(1\,2\,3)]$	$[(1\,3\,2)]$
$\lvert[h]\rvert$	1	1	1
$\mathbf{1}$	1	1	1
$\mathbf{1}'$	1	ω	ω^*
$\mathbf{1}''$	1	ω^*	ω

(9.6.4)

where $\omega \coloneqq \mathrm{e}^{2\pi i/3}$.
Given some irreducible representation of S_3, $\rho_{S_3,i}$, we are looking for a way to decompose the restriction of this representation to $\mathbb{Z}_3$:

$$\rho_{S_3,i}|_{\mathbb{Z}_3} = \bigotimes_{j=1}^{3} m_{ij}\rho_{\mathbb{Z}_3,j} \tag{9.6.5}$$

where $\rho_{\mathbb{Z}_3,j}$ are the irreducible representations of $\mathbb{Z}_3$ and $m_{ij} \in \mathbb{N}$. The coefficients are given by

$$m_{ij} = \langle \chi_{\rho_{S_3,i}}, \chi_{\rho_{\mathbb{Z}_3,j}} \rangle_{\mathbb{Z}_3} \coloneqq \frac{1}{|\mathbb{Z}_3|} \sum_{h\in\mathbb{Z}_3} \chi_{\rho_{S_3,i}}(h)^* \chi_{\rho_{\mathbb{Z}_3,j}}(h). \tag{9.6.6}$$

Given the character tables above it is possible to calculate this inner product for all pairs of irreducible representations and the only nonvanishing cases are

$$\langle \chi_{\mathbf{1}_{S_3}}, \chi_{\mathbf{1}_{\mathbb{Z}_3}} \rangle = \langle \chi_{\mathbf{1}'_{S_3}}, \chi_{\mathbf{1}_{\mathbb{Z}_3}} \rangle = \langle \chi_{\mathbf{2}_{S_3}}, \chi_{\mathbf{1}'_{\mathbb{Z}_3}} \rangle = \langle \chi_{\mathbf{2}_{S_3}}, \chi_{\mathbf{1}''_{\mathbb{Z}_3}} \rangle = 1. \tag{9.6.7}$$

This gives the branching rules

$$\rho_{\mathbf{1}_{S_3}}|_{\mathbb{Z}_3} = \rho_{\mathbf{1}_{\mathbb{Z}_3}} \qquad \mathbf{1}_{S_3} \to \mathbf{1}_{\mathbb{Z}_3}, \tag{9.6.8}$$

$$\rho_{\mathbf{1}'_{S_3}}|_{\mathbb{Z}_3} = \rho_{\mathbf{1}_{\mathbb{Z}_3}} \qquad \mathbf{1}'_{S_3} \to \mathbf{1}_{\mathbb{Z}_3}, \tag{9.6.9}$$

$$\rho_{\mathbf{2}_{S_3}}|_{\mathbb{Z}_3} = \rho_{\mathbf{1}'_{\mathbb{Z}_3}} \oplus \rho_{\mathbf{1}''_{\mathbb{Z}_3}} \qquad \mathbf{2}_{S_3} \to \mathbf{1}'_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3}. \tag{9.6.10}$$

Here we have used two alternative notations for the same thing.

The method used here is guaranteed to work, but is often more work than the following method which relies on a few key observations.

■ **Example 9.6.11** As a general rule the trivial representation is always mapped to the trivial representation, since $\rho(g) = 1$ for all $g \in G$ implies $\rho|_H(h) = 1$ for all $h \in H$. Hence $\mathbf{1}_{S_3} \to \mathbf{1}_{\mathbb{Z}_3}$.
The alternating representation, $\mathbf{1}'_{S_3}$, differs from the trivial representation by a negative sign on odd permutations. However, $\mathbb{Z}_3$ has no odd permutations since it is generated by $(1\,2\,3) = (1\,2)(2\,3)$, which is even. Hence there is no difference between the trivial and alternating representation when restricted to $\mathbb{Z}_3$ and so $\mathbf{1}'_{S_3} \to \mathbf{1}_{\mathbb{Z}_3}$.
The standard representation, $\mathbf{2}_{S_3}$, has to branch into two one dimensional representations of $\mathbb{Z}_3$ since the dimension of $\mathbf{2}_{S_3}$ is 2, and the dimension of a direct sum is the sum of the dimensions of the things being summed. Since $\mathbf{2}_{S_3}$ is a real representation it can split either into $2 \cdot \mathbf{1}_{\mathbb{Z}_3}$ or $\mathbf{1}'_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3}$. It is not possible for it to split into $\mathbf{1}_{\mathbb{Z}_3} \oplus \mathbf{1}'_{\mathbb{Z}_3}$ since this is not real, which can be seen from

$$\overline{\mathbf{1}_{\mathbb{Z}_3} \oplus \mathbf{1}'_{\mathbb{Z}_3}} = \overline{\mathbf{1}_{\mathbb{Z}_3}} \oplus \overline{\mathbf{1}'_{\mathbb{Z}_3}} = \mathbf{1}_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3} \tag{9.6.12}$$

where we have used the fact that $\overline{\mathbf{1}_{\mathbb{Z}_3}} = \mathbf{1}_{\mathbb{Z}_3}$ since the trivial representation is real and $\overline{\mathbf{1}'_{\mathbb{Z}_3}} = \mathbf{1}''_{\mathbb{Z}_3}$. The same logic forbids splitting into $\mathbf{1}_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3}$.
We can rule out the split into $2 \cdot \mathbf{1}_{\mathbb{Z}_3}$ since this is not faithful, and neither are the two other irreducible representations, and it can be shown that every group with a cyclic centre has at least one faithful representation. Therefore, we must have $\mathbf{2}_{S_3} \to \mathbf{1}'_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3}$.

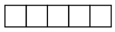
Branching rules are useful in physics when the system of interest has a given symmetry, G, and then something happens that disrupts the symmetry so that only a subgroup, H, of the original symmetry applies. An example might by rotational symmetry, $\mathrm{SO}(3)$, being broken by turning on an external field which gives a preferred direction, breaking to rotational symmetry about this axis, $\mathrm{SO}(2)$. In this case $\mathbf{3}_{\mathrm{SO}(3)} \to \mathbf{1}_{\mathrm{SO}(2)} \oplus \mathbf{2}_{\mathrm{SO}(2)}$.

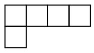
9.7 Constructing Representations

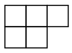
9.7.1 Irreducible Representations of S_n

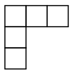
In this section we state a few facts about the irreducible representations of S_n without proof. We will work with S_5 as an example where appropriate. First recall that the number of conjugacy classes of S_n is given by the number of partitions of n, and is also equal to the number of irreducible representations. The conjugacy classes are given by cycle type.

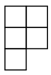
There is a graphical way to view partitions, using **Young tableau**. A partition of n given by $n = n_1 + n_2 + \cdots + n_k$ with $n_i \geq n_{i+1}$ is represented by a row of n_1 squares, with a row of n_2 squares below, with a row of n_3 squares below and so on for a total of n squares. For example,

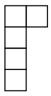
$5 = 5 =$  , (9.7.1)

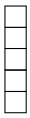
$= 4 + 1 =$  , (9.7.2)

$= 3 + 2 =$  , (9.7.3)

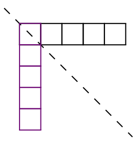
$= 3 + 1 + 1 =$  , (9.7.4)

$= 2 + 2 + 1 =$  , (9.7.5)

$= 2 + 1 + 1 + 1 =$  , (9.7.6)

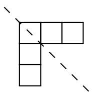
$= 1 + 1 + 1 + 1 + 1 =$  . (9.7.7)

Notice the symmetry here. The last Young tableau is given by reflecting the first in a diagonal line from top left to top right:



(9.7.8)

The same is then true of the second and penultimate tableau and so on. In the case where there is an odd number of tableaus the odd one out will end up being symmetric to itself under this reflection:



(9.7.9)

For each square in a Young tableau we can define a quantity called the **hook number**. This is the number of squares to the right of or below the square, plus the square itself. Most of the time these boxes form a right angled hook shape. This is shown in Figure 9.1 in detail for $5 = 3 + 2$. The hook lengths for all Young tableau's

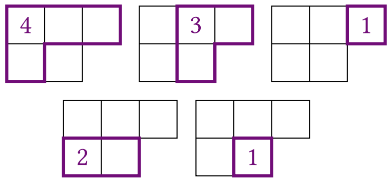


Figure 9.1: Calculating the hook lengths for the $5 = 3 + 2$ partition. The hook length is the number of squares to the right or below a given square, plus the square itself, so in this drawing the hook length of a square is the number of squares in the purple hook when that square is at the top left of the hook.

of 5 are given below:

$$\begin{ytableau} 5 & 4 & 3 & 2 & 1 \end{ytableau}, \quad \begin{ytableau} 5 & 3 & 2 & 1 \\ 1 \end{ytableau}, \quad \begin{ytableau} 4 & 3 & 1 \\ 2 & 1 \end{ytableau}, \quad \begin{ytableau} 5 & 2 & 1 \\ 2 \\ 1 \end{ytableau}, \tag{9.7.10}$$

$$\begin{ytableau} 4 & 2 \\ 3 & 1 \\ 1 \end{ytableau}, \quad \begin{ytableau} 5 & 1 \\ 3 & 2 & 1 \end{ytableau}, \quad \begin{ytableau} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{ytableau}. \tag{9.7.11}$$

The reason that we care about the hook length is because there is a rather remarkable theorem which states that the dimension of a irreducible representation of S_n is given by

$$\dim V_i = \frac{n!}{\prod \text{hook lengths}} \tag{9.7.12}$$

where the product is of all hook lengths appearing in Young tableau associated with the representation. In the case of S_5 this means

$$\dim V_1 = \frac{5!}{5!} = 1, \quad \begin{ytableau} 5 & 4 & 3 & 2 & 1 \end{ytableau}, \tag{9.7.13}$$

$$\dim V_2 = \frac{5!}{5 \cdot 3!} = 4, \quad \begin{ytableau} 5 & 3 & 2 & 1 \\ 1 \end{ytableau}, \tag{9.7.14}$$

$$\dim V_3 = \frac{5!}{4!} = 5, \quad \begin{ytableau} 4 & 3 & 1 \\ 2 & 1 \end{ytableau}, \tag{9.7.15}$$

$$\dim V_4 = \frac{5!}{5 \cdot 2 \cdot 2} = 6 \quad \begin{ytableau} 5 & 2 & 1 \\ 2 \\ 1 \end{ytableau}, \tag{9.7.16}$$

$$\dim V_5 = \frac{5!}{4!} = 5 \quad \begin{ytableau} 4 & 2 \\ 3 & 1 \\ 1 \end{ytableau}, \tag{9.7.17}$$

$$\dim V_6 = \frac{5!}{5 \cdot 3!} = 4 \quad \begin{ytableau} 5 & 1 \\ 3 & 2 & 1 \end{ytableau}, \tag{9.7.18}$$

$$\dim V_7 = \frac{5!}{5!} = 1 \quad \begin{ytableau} 5 \\ 4 \\ 3 \\ 2 \\ 1 \end{ytableau}. \tag{9.7.19}$$

While we won't prove that this works we can check that it doesn't violate the dimensionality theorem (Theorem 9.2.36):

$$1^2 + 4^2 + 5^2 + 6^2 + 5^2 + 4^2 + 1^1 = 120 = 5! = |S_5|. \tag{9.7.20}$$

The mirror symmetry between certain pairs of Young tableaus reflects a symmetry in the character tables. Namely that if ρ and ρ_{mirror} are irreducible representations corresponding to Young tableaus which are mirror images then

$$\chi_\rho(g) = \operatorname{sgn}(g)\chi_{\rho_{\mathrm{mirror}}}(g) \tag{9.7.21}$$

where $\operatorname{sgn}(g)$ is the sign of the permutation $g \in S_n$. Noticing that the $5 = 3 + 1 + 1$ partition, corresponding to the **6** representation, is its own mirror image this must mean that for all odd $g \in S_5$ we have $\chi_{\mathbf{6}}(g) = \operatorname{sgn}(g)\chi_{\mathbf{6}_{\mathrm{mirror}}}(g) = -\chi_{\mathbf{6}}(g)$, so we have $\chi_{\mathbf{6}}(g) = 0$ for all odd $g \in S_5$.

9.7.2 Direct Product Representations

So far we have focused on building the irreducible representations. Given two representations, $\rho_a \colon G \to \mathrm{GL}(V_1)$ and $\rho_b \colon G \to \mathrm{GL}(V_b)$, it is possible to build a new representation, $\rho_a \otimes \rho_b$, called the **tensor product** of the representations which acts on the tensor product space $V_a \otimes V_b$. In general this won't be irreducible but can be decomposed. The fact that $\rho_a \otimes \rho_b$ is a representation follows simply because the direct product of group homomorphisms is another group homomorphism. We can decompose $\rho_a \otimes \rho_b$ as

$$\rho_a \otimes \rho_b = m^1_{ab}\rho_1 \oplus \cdots \oplus m^k_{ab}\rho_{\rho_k} = \bigoplus_{c=1}^{k} m^c_{ab}\rho'_c, \tag{9.7.22}$$

with $m^c_{ab} \in \mathbb{Z}_{>0}$. Direct products of representations are commonly called **Kronecker products**, or when viewed as a direct sum like this decomposition **Clebsch–Gordan series**. Since $\dim(V_a \otimes V_b) = \dim(V_a)\dim(V_b)$ and $\dim(V_i \otimes V_j) = \dim(V_i) + \dim(V_j)$ it follows that

$$\dim(V_a)\dim(V_b) = \sum_{c=1}^{k} m^c_{ab} \dim V_c, \tag{9.7.23}$$

which is another diophantine equation that can be used to predict the dimensions of irreducible representations.

9.7.2.1 Motivation for Decompositions of Direct Product Representations

There are two motivating arguments for why direct products decompose into smaller representations. First, is an argument for finite groups. The direct product of the largest irreducible representation of a finite group must decompose since otherwise it is not the largest irreducible representation (unless one of the representations is one-dimensional, in which case the direct product isn't really interesting). This would also cause a conflict with the dimensionality theorem, with the dimension of the direct product being larger than the order of the group.

There is also a physical argument that the direct product representation should decompose based on the addition of angular momenta. This might make more sense after the next chapter on continuous groups. If a state has angular momentum $l = 1$ then it transforms as a vector under $\mathrm{SO}(3)$. Consider two three dimensional vectors, $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, which transforms under an $l = 1$ irreducible representation of the rotation group, $\mathrm{SO}(3)$. The direct product of these can be written as the vector with components

$$v_i w_j = A_{ij} + B_{ij} \tag{9.7.24}$$

where

$$A_{ij} = v_i w_j - \frac{1}{3}\delta_{ij}\boldsymbol{v}\cdot\boldsymbol{w} \tag{9.7.25}$$

is traceless and

$$B_{ij} = \frac{1}{3}\delta_{ij}\boldsymbol{v}\cdot\boldsymbol{w}. \tag{9.7.26}$$

The trace part cannot transform under rotations since it is defined by a scalar product. Also the trace is invariant under basis changes, which we can view as a superset of rotations. This means B_{ij} is a scalar (despite having two indices). This should be clear since B_{ij} is simply a scalar $(\boldsymbol{v}\cdot\boldsymbol{w})$ multiple of the identity matrix (δ_{ij}). Scalars transform under an $l = 0$ irreducible representation of SO(3).

The trace free part has an $l = 1$ antisymmetric and $l = 2$ symmetric irreducible representation. Motivated by the fact we can get a vector from two vectors via the cross product, $\boldsymbol{x} = \boldsymbol{v}\times\boldsymbol{w}$, which is antisymmetric we suppose that this corresponds to the antisymmetric $l = 1$ part. Recalling that addition of angular momentum means two states with angular momentum l_1 and l_2 can form states with angular momentum $|l_1 - l_2|$, $|l_1 - l_2| + 1$, and so on up to states with angular momentum $l_1 + l_2$ we see that for two $l = 1$ states we can form states with angular momentum $l = 0, 1, 2$. The reason for this should become clear in the next part when we consider the Lie group SO(3) and the corresponding Lie algebra, $\mathfrak{so}(3)$.

9.7.2.2 Multiplicities

We can calculate the multiplicities appearing in the decomposition of the direct product of representations using an inner product:

$$m_{ab}^{c} = \langle \chi_{\rho_a\otimes\rho_b}, \chi_{\rho_c}\rangle \tag{9.7.27}$$

$$= \frac{1}{|G|}\sum_{g\in G}\chi_{\rho_a\otimes\rho_b}(g)^*\chi_{\rho_c}(g) \tag{9.7.28}$$

$$= \frac{1}{|G|}\sum_{g\in G}\chi_{\rho_a}(g)^*\chi_{\rho_b}(g)^*\chi_{\rho_c}(g). \tag{9.7.29}$$

This leads to the following corollary.

Corollary 9.7.30 The Kronecker product of two irreducible representations, ρ_a and ρ_b, contains the trivial representation with multiplicity 1 if and only if $\rho_a = \overline{\rho_b}$.

R If ρ_a and ρ_b are not irreducible then the multiplicity can be greater than 1.

Proof. Let ρ_a and ρ_b be irreducible representations of some finite group G.

Then

$$m^{\mathbf{1}}_{ab} = \frac{1}{|G|}\sum_{g\in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g)^* \chi_{\mathbf{1}}(g) \tag{9.7.31}$$

$$= \frac{1}{|G|}\sum_{g\in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g)^* \tag{9.7.32}$$

where we have used the fact that $\chi_{\mathbf{1}}(g) = 1$ for all $g \in G$ since $\mathbf{1}$ maps all $g \in G$ to 1. Now using the fact that $\chi_\rho(g)^* = \chi_{\overline{\rho}}(g)$ we have

$$m^{\mathbf{q}}_{ab} = \frac{1}{|G|}\sum_{g\in G} \chi_{\rho_a}(g)\chi_{\overline{b}}(g) = \langle \chi_{\rho_a}, \chi_{\overline{\rho_b}}\rangle = \delta_{a\overline{b}}. \tag{9.7.33}$$

This is 1 if and only if $\rho_a = \overline{\rho_b}$, and zero otherwise. □

If instead we allowed on or both of the representations above to be reducible we could simply use the linearity of the inner product to derive a similar results.

We can use the Kronecker product to understand the mirror representations of S_n from Section 9.7.1. The mirror representation is related to the original representation by a Kronecker product with the alternating representation. That is

$$\rho_{\mathrm{mirror}} = \rho_{\mathbf{1}'} \otimes \rho. \tag{9.7.34}$$

9.7.3 Induced Representations

Induced representations can be thought of as the opposite of branching rules. Given two groups, G and H, with H a subgroup of G we can, given a representation, $\rho_a : H \to \mathrm{GL}(V_a)$, of H define a representation of G.

There are two essential steps. First let g_{n_i} be representatives of distinct cosets for $i = 1, \dots, [G : H]$. Define W as

$$W \cong \bigoplus_{i=1}^{[G:H]} g_{n_i} V_a \tag{9.7.35}$$

where each $g_{n_i} V_a$ is an isomorphic copy distinguished by the g_{n_i}s from each other. This means that $\dim W = [G : H] \dim V_a$. We can then define some vector $g_{n_i}\boldsymbol{v} \in g_{n_i} V_a$ for $\boldsymbol{v} \in V_a$. For different values of i these will correspond to, by construction, linearly independent vectors in W. The second step is to recognise that for arbitrary $g \in G$ we can write

$$g g_{n_i} = g_{m_i} h_i \tag{9.7.36}$$

for some $h_i \in H$. This is possible since $g g_{n_i} \in g_{m_i} H$ for some m_i since the left cosets, $g_{m_i} H$, partition G. The induced representation is then defined to act on $\boldsymbol{w} := \bigoplus_{i=1}^{[G:H]} g_{n_i} \boldsymbol{v}_{n'_i} \in W$ by

$$\rho_{\mathrm{induced}}(g)\boldsymbol{w} g \bigoplus_{i=1}^{[G:H]} g_{n_i} \boldsymbol{v}_{n'_i} = \bigoplus_{i=1}^{[G:H]} g_{m_i} \rho_a(h_i) \boldsymbol{v}_{n'_i}. \tag{9.7.37}$$

The dimension of the induced representation is the dimension of W. In general the induced representation is not irreducible.

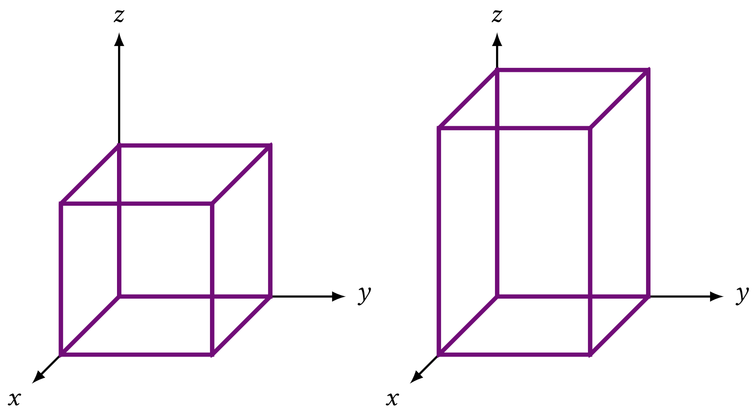


Figure 9.2: A cube distorted along the z-direction breaks symmetry $S_4 \to D_4$.

The regular representation can be seen as the representation induced by the case where H is the trivial subgroup and ρ_a is the trivial representation. The vector space is $V_0 = \mathbb{R}$ and we identify $g\boldsymbol{v} \in gV_0$ as $\boldsymbol{e}_g$ comparing to the notation of the original definition of the regular representation in Definition 9.2.32.

9.8 Applications

9.8.1 Distortion of Lattices

Consider a cubic lattice. A single cell of this lattice is a cube and so has the symmetry of a cube, which has symmetry group S_4. We will consider two distortions of this cube breaking the symmetry from S_4 to some subgroup. For simplicity we work with crystallographic coordinates[1], where the cube has corners at $(0, 0, 0)$, $(0, 0, 1)$, $(0, 1, 0)$ and so on.

[1] see the notes from the Introduction to Condensed Matter Physics.

First, consider a distortion in the z-direction. In this case the symmetry breaks to D_4, since we lose the ability to swap the side faces with the top and so the only symmetries are rotating around the z axis and rotating the whole cuboid to swap the top and bottom faces. This is depicted in Figure 9.2

Second, consider a distortion in the $(1, 1, 1)$-direction. This is depicted in Figure 9.3. This also breaks the symmetry and the resulting symmetry group is D_3. This is because we can view the distorted cube along the $(1, 1, 1)$-direction and end on we see a hexagon with three lines from its centre to three corners. This is depicted in Figure 9.4. This then has the same symmetry as an equilateral triangle, D_3.

9.8.2 Fermi–Bose Statistics

In quantum mechanics weird things happen when we have identical particles. In classical mechanics we can (in theory) track the trajectory of each particle and therefore distinguish even identical particles, based on some initial arbitrary labelling. In quantum mechanics we can't do this as there are no trajectories to follow.

Consider an ensemble of N non-interacting identical particles. The state of this system can be written as a tensor product of the states of the individual particles. Let $\mathcal{H}$ be the state space of a single particle and denote by $|\varphi_i\rangle$ a single particle state. Then the state space of the multi-particle system is $\mathcal{H}^{\otimes n}$. A given state then has the

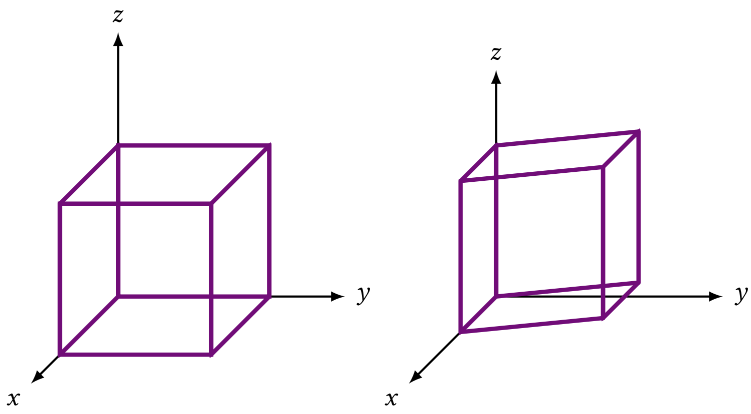


Figure 9.3: A cube distorted along the $(1,1,1)$-direction breaks symmetry $S_4 \longrightarrow D_3$.

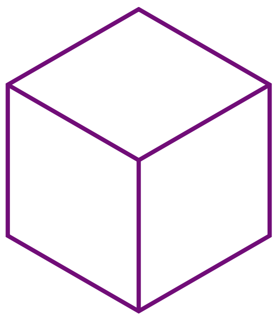


Figure 9.4: A cube distorted along the $(1,1,1)$-direction still has D_3 symmetry as it can be viewed corner on as depicted here.

form

$$|\Phi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle \otimes \cdots \otimes |\varphi_N\rangle = |\varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_N\rangle \tag{9.8.1}$$

where the last equality is just a notational convenience. It should be noted that by the assumption the particles are non-interacting we know they are not entangled and so it is possible to factorise the state into a product of single particle states. If there were interactions and entanglement this would not be the case and we would have some linear combination of factorised states instead.

The generic state vector can be in any irreducible representation of the symmetric permutation group, S_N, reflecting the fact that we can swap particles. The question is which representation will it be in? For large N, say $N = N_\mathrm{A} = 6 \times 10^{23}$ the number of irreducible representations is very large.

The two important cases of particles are bosons, particles with integer spin, $s = 0, 1, 2, \ldots$, and fermions, particles with half integer spin, $s = 1/2, 3/2, 5/3, \ldots$. For bosons the wave function is in the trivial representation, since exchanging two particles does nothing. For fermions the wave function is in the alternating representation, since exchanging two particles changes the sign.

As a concrete example consider the case of $N = 2$. We then have

$$|\Phi_{\text{bosons}}\rangle = \frac{\sqrt{2}}{2}\big(|\varphi_1 \otimes \varphi_2\rangle + |\varphi_2 \otimes \varphi_1\rangle\big), \tag{9.8.2}$$

$$|\Phi_{\text{fermions}}\rangle = \frac{\sqrt{2}}{2}\big(|\varphi_1 \otimes \varphi_2\rangle - |\varphi_2 \otimes \varphi_1\rangle\big). \tag{9.8.3}$$

One of the most important implications of this is that no two fermions can be in the same state, since in this case the wave function would vanish. This leads to Pauli's exclusion principle. The fact that this doesn't apply to bosons means that they can all occupy the same energy level, meaning at low temperatures they all fall into the ground state, or other low energy states. This leads to weird behaviour such as superfluidity and superconductivity. For fermions this isn't possible so we get a zero point pressure which prevents all of the fermions from being in the ground state. In a neutron star this zero point pressure balances the gravitational attraction and prevents the star from collapsing into a black hole.

The general connection between spin and statistics, namely that bosons have symmetric wave functions and fermions antisymmetric wave functions, is called the **spin statistics theorem**. Its proof is on quantum field theory and is beyond this course but it relies on a complexification of the Lorentz group.

Perhaps the most surprising thing about all of this is that none of the other representations of S_N appear. The statistics of particles acting under these in between representations, known as **parastatistics**, can be shown to be equivalent to either the trivial or alternating representation with some extra global symmetry, and therefore they don't bring anything new.

Part III

Continuous Groups

Ten

Lie Groups

10.1 Background

Finite groups were first utilised by mathematicians, particularly Galois, who wanted to classify the solutions to polynomials. Groups were used to show that, in general, polynomials of degree 5 or greater don't admit an exact solution in terms of elementary functions. That is, there is no "quintic formula" equivalent to the quadratic formula. There are linear, cubic, and quartic versions of the quadratic formula, but the linear formula is trivial and the cubic and quartic formulae contain too many terms to be useful for human computations. The main focus was on the symmetric groups, since permuting variables in equations played a big part in classifying the solutions, and it was soon noticed that the basic symmetry operations of a group applied to a far more general class of algebraic objects.

Sophus Lie introduced Lie groups, to be defined shortly, for similar reasons, he was attempting to classify solutions to differential equations. These have continuous solutions and so unsurprisingly Lie groups are continuous.

All simple finite groups have been classified. Likewise all simple and compact Lie groups have been classified, we will see the classification later.

10.2 Basics

> **Definition 10.2.1 — Continuous Group** A **continuous group** is a group, G, with an uncountable number of elements which can be continuously parametrised by some set of parameters, $\{\alpha\}$. We call the number of parameters the **dimension** of G, $\dim G$.

Notice that it doesn't make sense to discus the order of a continuous group since it is infinite.

> **Definition 10.2.2 — Lie Group** A **Lie group** is a continuous group which admits an analytic structure in the parameters $\{a\}$.
> Alternatively, a **Lie group** is a manifold equipped with a binary operation satisfying the group axioms.

For our purposes a manifold is a space which can be parametrised locally by a fixed number Euclidean coordinates. The number of coordinates is the dimension of the manifold. A fuller definition of a manifold is given in Chapter C.

All of the groups in the following definition are Lie groups. Technically, we are only defining a representation of these groups here but this is fine.

Definition 10.2.3 — Specific Lie Groups

- The **general linear group**, $\mathrm{GL}(n, \mathbb{F})$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{F}$, which is usually $\mathbb{R}$ or $\mathbb{C}$. It has the representation

$$\mathrm{GL}(n, \mathbb{F}) := \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M \neq 0\}. \tag{10.2.4}$$

  The requirement of non-zero determinant is so that the matrices are invertible.

- The **special linear group**, $\mathrm{SL}(n, \mathbb{F})$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{F}$ which preserve volumes. It has the representation

$$\mathrm{SL}(n, \mathbb{F}) := \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M \neq 1\} \subset \mathrm{GL}(n, \mathbb{F}). \tag{10.2.5}$$

- The **orthogonal group**, $\mathrm{O}(n)$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{R}$ which preserves lengths. It has the representation

$$\mathrm{O}(n) := \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = \mathbb{1}_n\} \subset \mathrm{GL}(n, \mathbb{R}). \tag{10.2.6}$$

- The **special orthogonal group**, $\mathrm{SO}(n)$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{R}$ which preserves lengths and orientations. It has the representation

$$\mathrm{SO}(n) := \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = \mathbb{1}_n \text{ and } \det O = 1\}. \tag{10.2.7}$$

  It is a subgroup of both $\mathrm{O}(n)$ and $\mathrm{SL}(n, \mathbb{R})$.

- The **unitary group**, $\mathrm{U}(n)$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{C}$ which preserves inner products. It has the representation

$$\mathrm{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = \mathbb{1}_n\} \subset \mathrm{GL}(n, \mathbb{C}). \tag{10.2.8}$$

- The **special unitary group**, $\mathrm{SU}(n)$, is defined as the group of transformations of an n-dimensional vector space over $\mathbb{C}$ which preserves inner products and orientations. It has the representation

$$\mathrm{SU}(n) := \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = \mathbb{1}_n \text{ and } \det U = 1\}. \tag{10.2.9}$$

  It is a subgroup of both $\mathrm{U}(n)$ and $\mathrm{SL}(n, \mathbb{C})$.

- The **symplectic group**, $\mathrm{USp}(2n) = \mathrm{Sp}(n)$, is defined as the group of transformations of an $2n$-dimensional vector space over $\mathbb{F}$ which preserves symplectic bilinear forms[a]. It has the representation

$$\mathrm{USp}(2n) = \mathrm{Sp}(n) := \{S \in \mathcal{M}_{2n}(\mathbb{F}) \mid S^\dagger J S = J\} \tag{10.2.10}$$

where

$$J := \begin{pmatrix} 0 & -\mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}. \tag{10.2.11}$$

Note that $J^2 = -\mathbb{1}_{2n}$, so we can think of J as a generalisation of i which has $i^2 = -1$.

[a] A **symplectic bilinear form** is a map $\omega\colon V \times V \to \mathbb{F}$ which is linear in both arguments, alternating, so $\omega(x, y) = -\omega(y, x)$, and non-degenerate, so $\omega(u, v) = 0$ for all $v \in V$ only if $u = 0$. We can view $\mathbb{R}^{2n}$ as a symplectic space by endowing it with the map J defined above such that $\omega(x, y) = x^\dagger J y$.

■ **Example 10.2.12** The following are group actions of some Lie group on $\mathbb{R}^3$.

- $\mathbb{R}^3$ with vector addition as the group operation is an 3-dimensional continuous group. The parameters can be seen as the components of the vectors.
- $\mathrm{SO}(3)$ with matrix multiplication as the group operation is a 3-dimensional continuous group. The parameters can be seen as the three Euler angles defining a rotation.

It should be noted that the dimension of $\mathrm{SO}(n)$ is not, in general, n. For example, $\mathrm{SO}(2)$ has dimension 1, since there is only one parameter defining rotations in the plane, the angle of the rotation. $\mathrm{SO}(4)$ has dimension 6, since there are initially $4^2 = 16$ degrees of freedom for the individual components of the matrix. It then turns out that orthogonality fixes one component in each row and column, since these must form unit vectors, and the requirement that the determinant is 1 fixes another three degrees of freedom.

10.3 Properties of Lie Groups

There are various properties which manifolds, and hence Lie groups, may or may not posses. In this section we cover some of these properties.

10.3.1 Basic Properties

Definition 10.3.1 — Finite or Infinite A manifold is finite dimensional if it is parametrised by a finite number of coordinates. If this is not the case then the manifold is infinite dimensional. A Lie group is **finite** or **infinite dimensional** if it is a finite or infinite dimensional manifold respectively.

■ **Example 10.3.2** The Lie groups $\mathrm{U}(1)$, $\mathrm{SO}(2)$, $\mathrm{SU}(2)$, and $\mathrm{SO}(3)$ are all finite dimensional.
The Lie group $\mathrm{U}(\mathcal{H})$ for an infinite dimensional Hilbert space is an infinite dimensional Lie group under the topology induced by the operator norm,

$$\|A\|_{\mathrm{op}} := \inf\{c \geq 0 \mid \|Av\| \leq c\|v\| \text{ for all } v \in \mathcal{H}\} \tag{10.3.3}$$

where $\|-\|$ is the norm on $\mathcal{H}$. This norm can be thought of as the smallest number, c, such that A doesn't scale the "length" of a vector by more than a factor of c. For example, this Lie group might represent all unitary transformations on the state of a particle with a continuous parameter.

Definition 10.3.4 — Real or Complex A manifold is real if it is parametrised by real numbers with smooth maps, and complex if it is parametrised by complex numbers with holomorphic maps. A Lie group is **real** or **complex** if it is a real or complex manifold, respectively.

Definition 10.3.5 The Lie groups $\mathrm{U}(1)$, $\mathrm{SO}(2)$, $\mathrm{SU}(2)$, and $\mathrm{SO}(3)$ are real Lie groups, although we will later see that $\mathrm{SU}(2)$ is pseudo-real.

10.3.2 Compactness

The following definition assumes the existence of a metric on the manifold.

Definition 10.3.6 — Compact A manifold, M, is compact if it is closed and bounded. **Bounded** means that $d(x, y) < r$ for all x, y in the manifold and $r \in \mathbb{R}$ being some finite number with d being a metric. **Closed** means that when viewing M as a submanifold of some larger manifold M contains its boundary. A Lie group is **compact** if it is compact as a manifold and **non-compact** otherwise.

As an example consider $\mathbb{R}$ with the standard metric $d(x, y) = |x - y|$. Then $[0, 1] := \{x \in \mathbb{R} \mid 0 \le x \le 1\}$ is closed and bounded, and hence compact. The intervals

$$[0, 1) := \{x \in \mathbb{R} \mid 0 \le x < 1\}, \tag{10.3.7}$$
$$(0, 1] := \{x \in \mathbb{R} \mid 0 < x \le 1\}, \quad \text{and} \tag{10.3.8}$$
$$(0, 1) := \{x \in \mathbb{R} \mid 0 < x < 1\} \tag{10.3.9}$$

are not closed, but are bounded. The interval $[0, \infty) := \{x \in \mathbb{R} \mid x > 0\}$ is not bounded since the distance between points can become arbitrarily large.

The circle, S^1, is compact but $\mathbb{R}$ isn't.

■ **Example 10.3.10** The Lie groups $\mathrm{U}(1)$, $\mathrm{SO}(2)$, $\mathrm{SU}(2)$, and $\mathrm{SO}(3)$ are compact Lie groups.

For a compact group all of the theorems of Chapter 8, such as Maschke's theorem (Theorem 8.4.1), Schur's lemma (Theorem 8.4.8), and the decomposability theorem (Theorem 8.4.13), can be modified to hold for compact continuous groups. To do so we replace the sum over group elements, $\sum_{g\in G}$, with the **Haar measure**, $\int \mathrm{d}\{\alpha\}$, we won't go into detail here on this though. For this reason we will restrict ourselves to compact groups, although here we give a few examples of non-compact groups.

■ **Example 10.3.11 — Non-Compact Lie Group** Consider the translation group on $\mathbb{R}$. That is $\mathbb{R}$ acting on $\mathbb{R}$ by $a\,.\,x = x + a$. This has a representation, $\rho := \mathbb{R} \to \mathrm{GL}(V)$, given by the following:

$$\rho(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}. \tag{10.3.12}$$

However, this doesn't decompose since the subspace defined by the span $\boldsymbol{e_x} = (1, 0)$ is an invariant subspace but its orthogonal complement, defined by the span of $\boldsymbol{e_y} = (0, 1)$, is not invariant, since $\rho(a)\boldsymbol{e_y} = (a, 1) \notin \operatorname{span}\{\boldsymbol{e_y}\}$. Hence we cannot write $\rho = \rho_1 \oplus \rho_2$ with $\rho_1 \colon \mathbb{R} \to \mathrm{GL}(W)$ and $\rho_2 \colon \mathbb{R} \to \mathrm{GL}(W^\perp)$ with $W \oplus W^\perp = V$, despite the fact that ρ is not irreducible. This is a failure of the decomposability theorem (Theorem 8.4.13) which holds only for compact groups.

■ **Example 10.3.13** The Lorentz group, $\mathrm{O}(1, 3)$, is not compact. This is because it is parametrised in part by the relative velocity of the frames and this is restricted to the non-compact interval $[0, c)$.
Maschke's theorem (Theorem 8.4.1) doesn't hold for the Lorentz group as it has no finite dimensional unitary representations. This is why we need to define an invariant of the form $\bar{\psi}\psi$ for spinors ψ with $\bar{\psi} := \psi^\dagger\gamma_0$ and $\gamma_0 := \mathbb{1}_2 \oplus (-\mathbb{1}_2) = \sigma_3 \otimes \mathbb{1}_2$. If ψ instead transformed under a finite dimensional unitary representation, which is guaranteed to exist for a compact group by Maschke's theorem (Theorem 8.4.1) then this would not be necessary as $\psi^\dagger\psi$ would be invariant without the need for γ_0.

10.3.3 Connectedness

Definition 10.3.14 A manifold is connected if there exists a continuous path between any two points in the manifold. A manifold is simply connected if any loop can be contracted continuously to a point. A manifold is disconnected if it is not connected. A Lie group is **disconnected**, **connected**, or **simply connectedsimply connected Lie group** if it is disconnected, connected, or simply connected as a manifold respectively.

Intuitively a space is simply connected if there are no holes.

■ **Example 10.3.15** The Lie groups $\mathrm{U}(1)$, $\mathrm{SO}(2)$, and $\mathrm{SO}(3)$ are connected. The Lie group $\mathrm{SU}(2)$ is simply connected. The Lie group $\mathrm{O}(2)$ is disconnected since there is no continuous path from $O \in \mathrm{O}(2)$ with $\det O = -1$ to $O' \in \mathrm{O}$ with $\det O' = +1$ since det is a continuous function and must jump from -1 to $+1$ along this path somewhere since all orthogonal matrices have $|\det O| = 1$.

10.3.4 Simplicity

Recall that a finite group is simple if it has no non-trivial proper normal subgroups, and that a normal subgroup is one which is invariant under conjugation. That is,

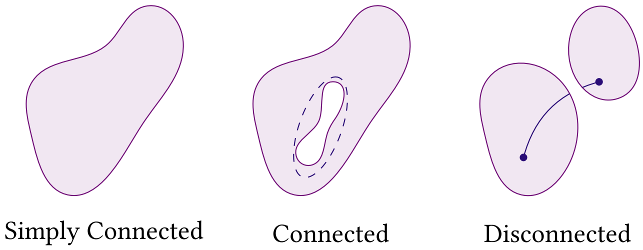


Figure 10.1: Examples of simply connected, connected, and disconnected spaces. The loop (- - -) in the connected example cannot be continuously contracted to a point without leaving the space. Similarly the two points (•) in the disconnected example cannot be connected by a continuous path. The path connecting them (——) has a jump.

$N \triangleleft G$ if $gng^{-1} \in N$ for all $n \in n$ and $g \in G$. We define simple Lie groups in a similar way but we explicitly disallow a few cases.

Definition 10.3.16 — Simple Lie Group A Lie group is **simple** if it is connected, non-Abelian, and has no non-trivial connected normal subgroups.
A **semi-simple Lie group** is a Lie group which can be written as a direct product of simple Lie groups.
A **composite Lie group** is a Lie group which can be written as a semi-direct product of simple Lie groups.

The reason we exclude more cases than for the definition for any group is motivated by the fact that it is not possible to define a quotient group from a simple group, apart from the trivial group and the simple group itself. We want this to hold replacing all groups with Lie groups. Importantly this means that we want a simple Lie group to be one where it is not possible to define a non-trivial quotient *Lie* group. If we expand upon what this requires we get the definition above.

We will only concern ourselves with simple Lie groups in this course.

■ **Example 10.3.17** The Lie groups $\mathrm{SL}(n)$, $\mathrm{SU}(n)$, and $\mathrm{USp}(2n) = \mathrm{Sp}(n)$ are simple.

10.3.4.1 Classification of Simple Compact Lie Groups

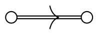
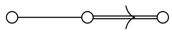
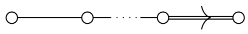
All simple compact Lie groups have been classified. It turns out there are four families of simple compact Lie groups, A_n, B_n, C_n, and D_n, and five exceptional groups, E_6, E_7, E_8, F_4, and G_2 which don't fall into any of these categories. We won't define the exceptional groups here, we just note that they exist. The families consist of familiar groups, $A_n = \mathrm{SU}(n+1)$, $B_n = \mathrm{SO}(2n+1)$, $C_n = \mathrm{USp}(2n)$, and $D_n = \mathrm{SO}(2n)$. These are sorted by their rank, which is given by the subscript. Rank is a concept related to the Lie algebras of these groups, which we shall meet later.

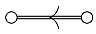
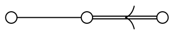
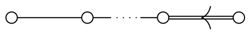
The Lie groups can be depicted by **Dynkin diagrams**, which are graphs with as many nodes as the rank[1] of the group. For example, A_1 corresponds to ○ , A_2 corresponds to ○——○ , and A_3 corresponds to ○——○——○ . In general A_n corresponds to ○——○– ⋯ –○——○ with n nodes.

[1]to be defined in Section 11.2

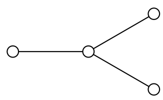
Group	Dimension	Rank
A_n	$(n+1)^2 - 1$	n
B_n	$n(2n+1)$	n
C_n	$n(2n+1)$	n
D_n	$n(2n-1)$	n
E_6	78	6
E_7	133	7
E_8	248	8
F_4	52	4
G_2	14	2

Table 10.1: Classification of simple compact Lie groups

The Dynkin diagram for B_n is similar but the final connection is doubled and directed, so B_2 corresponds to , and B_3 corresponds to . In general B_n corresponds to  with n nodes.

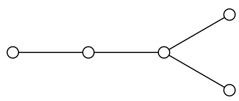
The Dynkin diagrams for C_n are almost identical to those of B_n but with the direction reversed, so C_2 corresponds to , and C_3 corresponds to . In general C_n corresponds to  with n nodes.

The Dynkin diagrams for D_n branch at the end into two. So D_4 corresponds to



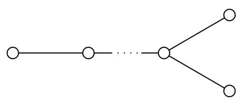
(10.3.18)

and D_5 corresponds to



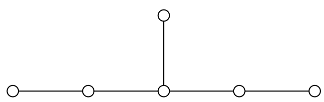
(10.3.19)

In general, D_n corresponds to



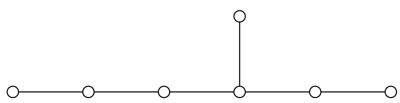
(10.3.20)

The exceptional groups also have Dynkin diagrams. For E_n ($n = 6, 7, 8$) the Dynkin diagram consists of a chain of $n - 1$ nodes with an extra node branching off three from the end, so E_6 corresponds to



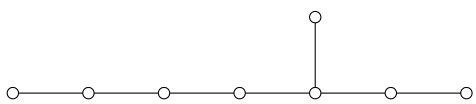
(10.3.21)

E_7 corresponds to

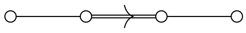


(10.3.22)

and E_8 to



(10.3.23)

The Dynkin diagram for F_4 is  .

The Dynkin diagram for G_2 is  .

Notice that for some low-rank cases the Dynkin diagrams are degenerate, for example A_1 and B_1 both correspond to $\circ$ which reflects the fact that $A_1 \cong B_1$, which is to say $SU(2) \cong SO(3)$, a fact that will be important later. It also explains why the E_n exceptional groups start at E_6, since the Dynkin diagram for E_5 is the same as for D_5, and the Dynkin diagram for E_4 is the same as for A_4.

Eleven

Lie Algebras

A lot of the uses of Lie groups make use of the underlying analytic structure to expand the group elements in the parameters, $\{\alpha\}$, and then keep only the first order terms, linearising the group. Mathematically this is what we do if we move to the tangent space of the manifold and we call the resulting linear space the **Lie algebra** associated with the Lie group. We will start with a simple example which we will then generalise.

11.1 The Abelian Group $\mathrm{U}(1) \cong \mathrm{SO}(2)$

We can define the unitary group $\mathrm{U}(1)$ as

$$\mathrm{U}(1) := \{U \in \mathcal{M}_1(\mathbb{C}) \mid U^\dagger U = \mathbb{1}\}. \tag{11.1.1}$$

This is obviously the same as the circle group

$$\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\} \tag{11.1.2}$$

under the obvious isomorphism associating $(z) \in \mathrm{U}(1)$ with $z \in \mathbb{T}$. We therefore don't distinguish between $\mathrm{U}(1)$ and $\mathbb{T}$ and will say things like $z \in \mathrm{U}(1)$, when it may be more proper to say $(z) \in \mathrm{U}(1)$. As the name suggests $\mathbb{T}$ is a Lie group with the circle, S^1, as its underlying manifold. This is a one-dimensional real manifold parametrised by $\alpha \in [0, 2\pi)$.

We can define the two dimensional rotation group, $\mathrm{SO}(2)$, as

$$\mathrm{SO}(2) := \{R \in \mathcal{M}_2(\mathbb{R}) \mid R^\top R = \mathbb{1} \text{ and } \det R = 1\}. \tag{11.1.3}$$

This acts on the plane, $\mathbb{R}^2$, through rotations, which are given by normal matrix multiplication, $R \,.\, \boldsymbol{x} = R\boldsymbol{x}$. This is a one dimensional real manifold parametrised by the rotation angle, $\alpha \in [0, 2\pi)$.

The unitary group $\mathrm{U}(1)$ has an infinite family of representations labelled by $n \in \mathbb{Z}$ given by

$$\rho_n^{\mathrm{U}(1)}(\alpha) = \mathrm{e}^{in\alpha}. \tag{11.1.4}$$

Here we are implicitly associating complex numbers with 1×1 complex matrices. The two-dimensional rotation group, $\mathrm{SO}(2)$, has an infinite family of representations labelled by $n \in \mathbb{Z}$ given by

$$\rho_n^{\mathrm{SO}(2)}(\alpha) = \begin{pmatrix} \cos(n\alpha) & \sin(n\alpha) \\ -\sin(n\alpha) & \cos(n\alpha) \end{pmatrix}. \tag{11.1.5}$$

It should be clear that $\mathrm{U}(1)$ and $\mathrm{SO}(2)$ are isomorphic. One isomorphism between them being $e^{i\alpha} \mapsto \rho_1^{\mathrm{SO}(2)}(\alpha)$. Since this is the case we will move between them as needed choosing which ever is most appropriate for the task at hand.

The Kronecker product of representation is particularly simple for this case with

$$\rho_n \otimes \rho_m = \rho_{n+m}. \tag{11.1.6}$$

In order to linearise $\mathrm{SO}(2)$ we make use of the fact that elements of $\mathrm{U}(1)$ can be expressed as $e^{i\alpha}$ and we suggest that $O \in \mathrm{SO}(2)$ can be expressed as

$$O = \exp[i\alpha T] \tag{11.1.7}$$

for $T \in \mathcal{M}_2(\mathbb{R})$. As usual the exponential of a matrix is to be understood either through its power series,

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}, \tag{11.1.8}$$

or a limit,

$$\exp(A) = \lim_{n\to\infty} \left(1 + \frac{A}{n}\right)^n. \tag{11.1.9}$$

Lemma 11.1.10 Let $A \in \mathcal{M}_m(\mathbb{C})$ be diagonalisable. Then $\det(\exp(A)) = \exp(\operatorname{tr}(A))$.

Proof. Let $A \in \mathcal{M}_m(\mathbb{C})$ be diagonalisable. We work in the basis in which A is diagonal since both det and tr are basis independent. In this basis the diagonal of A consists of its eigenvalues, λ_i. We then have

$$\det A = \prod_{i=1}^{m} \lambda_i, \tag{11.1.11}$$

and

$$\operatorname{tr} A = \sum_{i=1}^{m} \lambda_i. \tag{11.1.12}$$

Further in this basis $\exp(A)$ is diagonal and its diagonal components are $\exp(\lambda_i)$. This follows since the nth power of a diagonal matrix just raises the elements on the diagonal to the nth power, that is

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{pmatrix}. \tag{11.1.13}$$

It follows that

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \tag{11.1.14}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}^n \tag{11.1.15}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{pmatrix} \tag{11.1.16}$$

$$= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{\lambda_1^n}{n!} & & \\ & \ddots & \\ & & \sum_{n=0}^{\infty} \frac{\lambda_m^n}{n!} \end{pmatrix} \tag{11.1.17}$$

$$= \begin{pmatrix} \exp(\lambda_1) & & \\ & \ddots & \\ & & \exp(\lambda_m) \end{pmatrix} \tag{11.1.18}$$

Hence

$$\exp(\mathrm{tr}(A)) = \exp\left(\sum_i \lambda_i\right) \tag{11.1.19}$$

$$= \prod_i \exp(\lambda_i) \tag{11.1.20}$$

$$= \det(\exp(\lambda_i)). \tag{11.1.21}$$

□

The definition of $O \in \mathrm{SO}(2)$ is that $O^\top O = \mathbb{1}$, which can be expressed as

$$\mathbb{1} = O^\top O \tag{11.1.22}$$

$$= \exp(i\alpha T)^\top \exp(i\alpha T) \tag{11.1.23}$$

$$= \exp(i\alpha T^\top) \exp(i\alpha T) \tag{11.1.24}$$

$$= (\mathbb{1} + i\alpha T^\top + \mathcal{O}(\alpha^2))(\mathbb{1} + i\alpha T + \mathcal{O}(\alpha^2)) \tag{11.1.25}$$

$$= \mathbb{1} + i\alpha(T^\top + T) + \mathcal{O}(\alpha^2). \tag{11.1.26}$$

Therefore if we take α to be small we have that $T^\top = -T$, which is to say that T must be antisymmetric. This in fact generalises to $\mathrm{SO}(n)$, we can write any element as $\exp(i\alpha T)$ where $T \in \mathcal{M}_n(\mathbb{R})$ is antisymmetric.

One particular solution is

$$T = i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{11.1.27}$$

which leads to the previous representation since $T^2 = \mathbb{1}$ and so

$$\exp(i\alpha T) = \mathbb{1}\cos(\alpha) + iT\sin(\alpha) = \rho_1^{\mathrm{SO}(2)}(\alpha). \tag{11.1.28}$$

This follows by expanding the exponential and collecting even and odd terms.

We define the Lie algebra of SO(2) to be all (real) scalar multiples of T, $\mathfrak{so}(2) := \{\lambda T \mid \lambda \in \mathbb{R}\}$. This is fairly boring since there is only one dimension, but it is useful to demonstrate the tangent space notion of the Lie algebra.

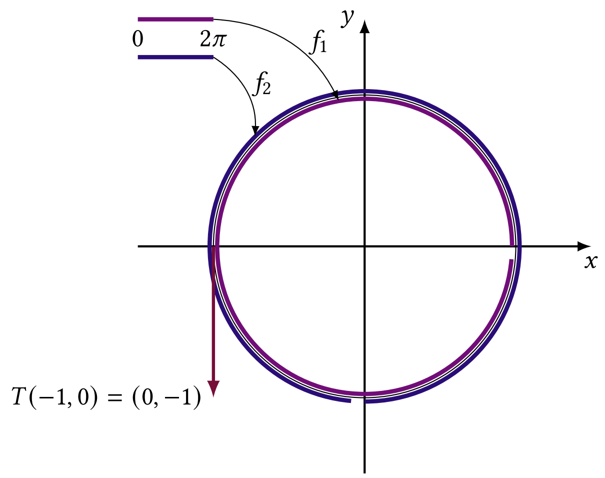


Figure 11.1: The manifold S^1 covered by two charts, f_1 and f_2. $T(x, y)$ gives a tangent vector as demonstrated by the case of $(x, y) = (-1, 0)$. The slight gaps in the two charts represent the discontinuity in α going from 0 to 2π. Either chart is valid away from these discontinuities and at a discontinuity simply use the continuous chart.

As previously mentioned the underlying manifold for $\mathrm{U}(1)$, and hence $\mathrm{SO}(2)$, is the circle, $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, strictly this is an embedding of S^1 in two-dimensions, on its own S^1 is a one-dimensional manifold. We can cover S^1 with two charts, two are needed to avoid a discontinuity at the join, we simply use whichever hasn't got the join at that point. For example, one chart could be $f_1(\vartheta) = (\cos\vartheta, \sin\vartheta)$ and another $f_2(\vartheta) = (\cos(\vartheta - \pi/1), \sin(\vartheta - \pi/2))$, which is just f_1 rotated around by $\pi/2$. Both of these maps have the codomain $[0, 2\pi)$.

Now consider the point (x, y) on S^1. We have $T(x, y) = (-y, x)$. For example, if $T(-1, 0) = (0, -1)$. This vector will be tangent to S^1 as embedded in $\mathbb{R}^2$.

The manifold $\mathrm{SO}(2)$ is connected, since S^1 is clearly connected. On the other hand the manifold $\mathrm{O}(2)$ is *not* connected. It is formed from two disconnected pieces, one, which has $\det O = 1$, is essentially a copy of $\mathrm{SO}(2)$, and one where $\det O = -1$. These two pieces are disconnected since det is a continuous function of the parameters yet makes a sudden jump from -1 to $+1$, which can only happen if there is a corresponding sudden jump in the parameters, meaning that $\mathrm{O}(2)$ is not connected.

Consider the parity operator

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.1.29}$$

This is an element of $\mathrm{O}(2)$, but not $\mathrm{SO}(2)$. Since $P^2 = \mathbb{1}$ we can view P as generating $\mathbb{Z}_2 = \langle P \mid P^2 = \mathbb{1}\rangle$ and we can then identify

$$\mathrm{SO}(2) \cong \mathrm{O}(2)/\mathbb{Z}_2. \tag{11.1.30}$$

In fact, this generalises to

$$\mathrm{SO}(n) \cong \mathrm{O}(n)/\mathbb{Z}_2. \tag{11.1.31}$$

Further we have

$$\mathrm{SU}(n) \cong \mathrm{U}(n)/\mathrm{U}(1) \tag{11.1.32}$$

since we can think of $\mathrm{U}(1)$ as the complex version of $\mathbb{Z}_2$.

For $\mathrm{SO}(2)$ the Haar measure is

$$\int_0^{2\pi} \frac{\mathrm{d}\alpha}{2\pi} \tag{11.1.33}$$

and so we can define an inner product on the character space as

$$\begin{aligned} \langle \chi_{\rho_n}, \chi_{\rho_m} \rangle_{\mathrm{U}(1)} &= \frac{1}{2\pi} \int_0^{2\pi} \chi_{\rho_n}(\alpha)^* \chi_{\rho_m}(\alpha) \,\mathrm{d}\alpha && (11.1.34) \\ &= \frac{1}{2\pi} \int_0^{2\pi} (\mathrm{e}^{in\alpha})^* \mathrm{e}^{im\alpha} \,\mathrm{d}\alpha && (11.1.35) \\ &= \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{i(n-m)\alpha} \,\mathrm{d}\alpha && (11.1.36) \\ &= \delta_{nm}. && (11.1.37) \end{aligned}$$

We see that the character are orthonormal with respect to this inner product. The characters turn out to be less important for Lie groups than finite groups and we won't consider them much more.

11.2 Lie Algebra Generalities

Theorem 11.2.1 — Exponential Map. Every one-parameter subgroup of $\mathrm{GL}(n, \mathbb{C})$ is given by a matrix exponential, $\rho_T(t) = \exp(tT)$ where $T \in \mathcal{M}_n(\mathbb{C})$ and T is not the zero matrix.

R We call T the **generator** of the one-parameter subgroup.

Proof. Consider the single parameter homomorphism $\rho_T : \mathbb{R} \longrightarrow \mathrm{GL}(n, \mathbb{C})$ defined by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t} \rho_T(t)\right|_{t=0} = T \tag{11.2.2}$$

with the boundary condition $\rho_T(0) = \mathbb{1}$. Since ρ_T is a group homomorphism we require

$$\rho_T(s+t) = \rho_T(s)\rho_T(t). \tag{11.2.3}$$

Differentiating this with respect to s and evaluating at $s = 0$ gives

$$\left.\frac{\mathrm{d}}{\mathrm{d}s} \rho_T(s+t)\right|_{s=0} = \left.\frac{\mathrm{d}}{\mathrm{d}s} \rho_T(s)\right|_{s=0} \rho_T(t) = T\rho_T(t) = \left.\frac{\mathrm{d}}{\mathrm{d}s} \rho_T(t+s)\right|_{s=0} = \frac{\mathrm{d}}{\mathrm{d}t} \rho_T(t) \tag{11.2.4}$$

where in the last step we set $s + t = t$. This gives us a differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \rho_T(t) = T\rho_T(t) \tag{11.2.5}$$

which has the unique solution

$$\rho_T(t) = \exp(tT). \tag{11.2.6}$$

This proves the theorem. □

This theorem generalises to multi-parameter subgroups of $\mathrm{GL}(n, \mathbb{C})$, where we can write all elements in the form

$$\exp(t^a T_a) \tag{11.2.7}$$

with the Einstein summation convention implying $t^a T_a = t^1 T_1 + t^2 T_2 + \cdots + t^k T_k$.

The possible generators depend on the group. We want to consider the product $\rho_1 \rho_2 \rho_1^{-1} \rho_2^{-1}$, this will be $\mathbb{1}$ for an Abelian group and some ρ_3 in the group by the closure of the group for a non-Abelian group. In order to avoid the **Baker–Campbell–Hausdorff formula**,

$$\mathrm{e}^A \mathrm{e}^B = \exp\left(A + B + \frac{1}{2}[A, B] + \frac{1}{2}\frac{1}{3!}\big([A, [A, B]] + [[A, B], B]\big) + \cdots\right), \tag{11.2.8}$$

we expand ρ_i as

$$\rho_1 = \mathbb{1} + i\alpha^a T_a + \frac{1}{2}(i\alpha^a T_a)^2 + O(\alpha^3), \tag{11.2.9}$$

$$\rho_2 = \mathbb{1} + i\beta^a T_a + \frac{1}{2}(i\beta^a T_a)^2 + O(\beta^3), \tag{11.2.10}$$

$$\rho_3 = \mathbb{1} + i\gamma^a T_a + \frac{1}{2}(i\gamma^a T_a)^2 + O(\gamma^3). \tag{11.2.11}$$

We call the matrices T_a the Lie algebra generators. Substituting this into $\rho_1 \rho_2 \rho_1^{-1} \rho_2^{-1}$ we get

$$[\alpha^a T_a, \beta^b T_b] = -i\gamma^c T_c + O(\alpha^3, \beta^3, \gamma^3). \tag{11.2.12}$$

All lower order terms cancel. We can therefore find ρ_3 by finding $\gamma^c = -f_{ab}{}^c \alpha^a \beta^b$ where we define the **structure constants**, $f_{ab}{}^c$, through the Lie algebra

$$[T_a, T_b] = i f_{ab}{}^c T_c. \tag{11.2.13}$$

This is the bare minimum amount of structure that the generators have to have in order to be compatible with the group structure. Notice that $f_{ab}{}^c = -f_{ba}{}^c$, since $[A, B] = -[B, A]$. Importantly the structure constants depend only on the commutator, and not on the anti-commutator. This is required since the commutator is independent of the representation but the anti-commutator is not.

The number of generators is equal to the dimension of the Lie group. Not all generators commute. We define the **rank** of the Lie group to be the size of the largest subset of generators such that all generators in the subset commute. Notice that the rank is at most equal to the dimension of the Lie group, with equality for an Abelian group where all generators, and hence all group elements, commute. Note that both $\mathrm{SO}(3)$ and $\mathrm{SU}(2)$ are rank one, which makes them about as simple as non-Abelian Lie groups can be.

We now posit a theorem without proof, since the proof requires more details about manifolds which are beyond the scope of this course.

Theorem 11.2.14. To any Lie algebra there corresponds a unique Lie group which is simply connected. This group is called the **universal covering group**.

In particular we find the universal covering group by exponentiating the Lie algebra. The important thing here is the simply connected part. This means that if we linearise a Lie group to get its Lie algebra the exponential map won't necessarily give back the same group, but it will give back a subgroup which is simply connected. As a subgroup this must contain the identity, and so we call this the component of the Lie group connected to the identity.

It turns out that compactness puts some restrictions on what the generators can be.

Theorem 11.2.15. For a compact group the associated Lie algebra is generated by Hermitian generators.

Proof. Given some $\rho \in G$ such that ρ is connected to the identity we can write $\rho = \exp(\iota\alpha^a T_a)$. Compact groups admit finite dimensional unitary representations, and so we have

$$\begin{aligned} \mathbb{1} &= \exp(i\alpha^a T_a)^\dagger \exp(i\alpha^a T_a) && (11.2.16)\\ &= \exp(-i\alpha^a T_a^\dagger)\exp(i\alpha^a T_a) && (11.2.17)\\ &= (\mathbb{1} - i\alpha^a T_a^\dagger + O(\alpha^2))(\mathbb{1} + i\alpha^a T_a + O(\alpha^2)) && (11.2.18)\\ &= \mathbb{1} + i\alpha^a(T_a - T_a^\dagger) + O(\alpha^2). && (11.2.19) \end{aligned}$$

Hence we must have $T_a - T_a^\dagger = 0$, and since this must hold for all ρ, and hence for all α^a we have $T_a = T_a^\dagger$ meaning T_a is Hermitian individually for each a. □

Corollary 11.2.20 The Hilbert–Schmidt inner product on $\mathcal{M}_n(\mathbb{C})$ is defined by

$$\langle T_a, T_b\rangle := \operatorname{tr}(T_a^\dagger T_b), \tag{11.2.21}$$

and if T_a are the generators of a Lie algebra associated with a compact group, that is T_a are Hermitian, we have

$$\langle T_a, T_b\rangle := \operatorname{tr}(T_a^\dagger T_b) = \operatorname{tr}(T_a T_b). \tag{11.2.22}$$

This leads to the following theorem,

Theorem 11.2.23. For a compact, semi-simple Lie algebra there is an orthogonal basis such that

$$\langle T_a, T_b\rangle := \operatorname{tr}(T_a T_b) = 2k_R\delta_{ab}. \tag{11.2.24}$$

We call k_R the **Dynkin index** and it depends on the representation, which is what the subscript R is there to remind us of. In this basis the structure constants, ${f_{ab}}^c$, are completely antisymmetric.

We can make the above theorem slightly more precise, but this theorem is beyond the scope of the course:

Theorem 11.2.25 — Killing Metric. We can define a unique symmetric tensor called the **Killing metric**:

$$k_{ab} := \frac{1}{k_R} \operatorname{tr}(T_a T_b) \tag{11.2.26}$$

where k_R is some constant depending on the choice of representation. This tensor is invariant under the action of the Lie group, that is

$$0 = \delta_{T_c} \operatorname{tr}(T_a T_b) = \operatorname{tr}([T^c, T_a] T_b) + \operatorname{tr}(T_a [T^c, T_b]) = i k_R ({f^c}_{ab} + {f^c}_{ba}). \tag{11.2.27}$$

This implies that the structure constants are completely antisymmetric.
For semi-simple Lie groups the Killing metric is non-singular. For semi-simple, compact Lie groups there is a basis where

$$k_{ab} = 2\delta_{ab}. \tag{11.2.28}$$

We can use the metric, k_{ab}, to raise and lower indices[1]. Since $k_{ab} \propto \delta_{ab}$ in this particular basis there is no numerical difference between ${f_{ab}}^c$ and f_{abc} and so we won't differentiate between them.

[1] see the notes for general relativity for lots on raising and lower indices

Theorem 11.2.29 — Jacobi Identity. The structure constants satisfy the **Jacobi identity**:

$$f_{aed} f_{bce} + f_{bed} f_{cae} + f_{ced} f_{abc} = 0. \tag{11.2.30}$$

R Note that this is simply $f_{aed} f_{bce}$ with a sum over cyclic permutations of a, b, and c.

Proof. For any three matrices, $A, B, C \in \mathcal{M}_n(\mathbb{C})$ we have

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]]. \tag{11.2.31}$$

This can be shown by expanding these commutators. We can then interpret f_{abc} as the components of the matrix $(T_a^{\text{adj}})_{bc} = -i f_{abc}$ and the result follows. ◻

Definition 11.2.32 — Adjoint Representation The **adjoint representation** is given by defining the generator T_c^{adj} to have components $(T_c^{\text{adj}})_{bc} = -i f_{abc}$.

Corollary 11.2.33 The structure constants are real.

Proof. If T_a generates a representation of the Lie algebra then so does $-T_a^*$, this is just the negative of the complex conjugate representation. Taking the complex conjugate of the definition of the adjoint representation, dropping the adj label, we have

$$[T_a, T_b]^* = (if_{abc}T_c)^* \implies [T_a^*, T_b^*] = -if_{abc}^* T_c^* \tag{11.2.34}$$

noting that $[-A, -B] = [A, B]$ since the negatives cancel when we expand the commutator we get

$$[-T_a^*, -T_b^*] = if_{abc}^*(-T_a^*). \tag{11.2.35}$$

Since f_{abc} are independent of the representation we must have $f_{abc} = f_{abc}^*$, which means that $f_{abc} \in \mathbb{R}$. □

The adjoint representation acts on the Lie algebra itself via the commutator. In particular if T_a^{adj} are the generators in the adjoint representation and T_d are the generators in some other representation then $T_a^{\text{adj}} \circ T_d = [T_a, T_d] = if_{ade}T_e$ and $if_{ade} = (T_a^{\text{adj}})_{de}$.

So far we have viewed Lie algebras as the result of linearising a Lie group. We can extract the important algebraic details into a more abstract object. This is the way a mathematician would approach the subject. They would first make the following definition and then derive what we have taken as the defining properties as a result of this definition.

Definition 11.2.36 — Lie Algebra A **Lie algebra**, $\mathfrak{g}$, is a vector space over $\mathbb{F}$ with a non-associative, alternating bilinear product satisfying the Jacobi identity. This product is called the **Lie bracket**. Its properties are

- **Bilinearity**: For all $x, y, z \in \mathfrak{g}$ and $a, b \in \mathbb{F}$

$$[ax + by, z] = a[x, z] + b[y, z], \quad \text{and} \tag{11.2.37}$$
$$[z, ax + by] = a[z, x] + b[z, y]. \tag{11.2.38}$$

- **Alternativity**: For all $x \in \mathfrak{g}$

$$[x, x] = 0. \tag{11.2.39}$$

- The **Jacobi identity**: For all $x, y, z \in \mathfrak{g}$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \tag{11.2.40}$$

Notice that the Jacobi identity just says that sums over symmetric permutations of x, y, and z in $[x, [y, z]]$ must vanish. Another important fact is that alternativity combined with bilinearity implies anticommutativity, so $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$.

With this definition we can define the universal enveloping algebra:

Definition 11.2.41 — Universal Enveloping Algebra Given a Lie algebra, $\mathfrak{g}$, we can embed it in an associative algebra, A, in such a way that the abstract Lie bracket of $\mathfrak{g}$, corresponds to the commutator in A, that is $[x, y] = xy - yx$ in A. The **universal enveloping algebra** is defined to be the associative algebra generated by t_i which are subject only to the conditions

$$t_i t_j - t_j t_i = i f_{ijk} t_k. \tag{11.2.42}$$

We denote elements of the universal enveloping algebra here by lower case to distinguish from elements of the Lie algebra.

Since we mostly consider matrix Lie groups which have Lie algebras where the Lie bracket can be interpreted as the commutator we often won't distinguish between the Lie algebra and its universal enveloping algebra.

The reason for defining the universal enveloping algebra is to make the following definition:

Definition 11.2.43 — Casimir Element A **Casimir element** is an element of the universal enveloping algebra of a Lie algebra which commutes with all generators of the Lie algebra. That is it is in the centre of the universal enveloping algebra.

Every Lie algebra associated with a semi-simple Lie group has at least one Casimir element, called the **quadratic Casimir**, which is defined by

$$C = \delta^{ab} T_a T_b. \tag{11.2.44}$$

For compact Lie groups Schur's lemma (Theorem 8.4.8) tells us that C is proportional to the identity, so $C = C_2(R)\mathbb{1}$, where the 2 stands for quadratic and the R tells us that the value of the number $C_2(R)$ is representation dependent.

It can be shown that the number of Casimir operators is equal to the number of invariant tensors, which is equal to the rank of the Lie algebra.

11.3 The Non–Abelian Groups $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$

11.3.1 The Lie Algebras of $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$

Recall that

$$\mathrm{SO}(3) := \{O \in \mathcal{M}_3(\mathbb{R}) \mid O^{\top} O = \mathbb{1}\}, \tag{11.3.1}$$

and

$$\mathrm{SU}(2) := \{U \in \mathcal{M}_2(\mathbb{C}) \mid U^{\dagger} U = \mathbb{1}\}. \tag{11.3.2}$$

The Lie algebra of $\mathrm{SO}(3)$, denoted $\mathfrak{so}(3)$, is generated by any three linearly independent antisymmetric matrices. Ideally we would work in a basis where the structure constants are completely antisymmetric. Such a basis is guaranteed to exist since $\mathrm{SO}(3)$ is compact and simple. One such basis is

$$T_1 = -i\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad T_2 = -i\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad T_3 = -i\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{11.3.3}$$

Notice that the first 2×2 submatrix of T_3 corresponds to a rotation by $\pi/2$ in SO(2), which corresponds to a three-dimensional rotation about the z-axis. The other two matrices are just permutations of this. It is easy to compute the Lie algebra for these three matrices:

$$[T_a, T_b] = i\varepsilon_{abc}T_c \tag{11.3.4}$$

where ε_{abc} is the Levi–Civita symbol. Hence the structure constants are $f_{abc} = \varepsilon_{abc}$. This shouldn't be too surprising since in three dimensions any antisymmetric tensor of rank 3 must be proportional to ε_{abc}. We can easily calculate the Dynkin index from

$$\mathrm{tr}(T_aT_b) = 2k_R\delta_{ab}, \tag{11.3.5}$$

since putting in $a = b = 1$ we get $\mathrm{tr}(T_1T_1) = 2$ and so $k_R = 1$ for this representation.

We know that the Lie algebra of SU(2), denoted $\mathfrak{su}(2)$, is generated by traceless Hermitian matrices. One choice is the Pauli matrices,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \text{and} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.3.6}$$

The generators of $\mathfrak{su}(2)$ are then given by

$$T_a \coloneqq \frac{1}{2}\sigma_a. \tag{11.3.7}$$

We can easily show that $\mathrm{tr}(T_aT_b) = \delta_{ab} = 2k_R\delta_{ab}$ and so $k_R = 1/4$ for this representation. The Lie algebra can be shown to satisfy

$$[T_a, T_b] = i\varepsilon_{abc}T_c, \tag{11.3.8}$$

which is the same as for $\mathfrak{so}(3)$.

What this means is that SO(3) and SU(2) are **locally isomorphic**, meaning they have the same Lie algebra, but globally different, meaning they have different Dynkin indices.

Given that $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ there must be a single simply connected Lie group which this Lie algebra exponentiates to. We can show that this is SU(2). Consider the mapping $\varphi\colon \mathbb{R}^4 \to \mathrm{SU}(2)$ given by

$$\varphi(x_0, x_1, x_2, x_3) = x_0\mathbb{1}_2 + x_1 i\sigma_1 + x_2 i\sigma_2 + x_3 i\sigma_3 = \begin{pmatrix} x_0 + ix_3 & ix_1 + x_2 \\ ix_1 - x_2 & x_0 - ix_3 \end{pmatrix} \coloneqq U. \tag{11.3.9}$$

The defining conditions for SU(2) lead to

$$U^\dagger U = \mathbb{1}_2 \text{ and } \det U = 1 \iff x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1. \tag{11.3.10}$$

What this means is that SU(2) is topologically (homeomorphic to) a three-sphere, S^3. Since three-spheres are simply connected this means SU(2) is simply connected and so $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ exponentiates to SU(2).

It is worth briefly considering why SO(3) is not simply connected. As a manifold SO(3) is a three-dimensional ball of radius π, that is $\{(x, y, z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq \pi^2\}$ with opposite points associated. That is if two points are opposite on the 2-sphere $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = \pi^2\}$, we consider them to be the same point. The reason for this is that we can associate a line through the centre of this ball to be the axis of rotation and the distance we move along the line from the origin is the magnitude of the rotation, with the origin representing the angle 0 and one end of the line π, and the other end $-\pi$, however a rotation by π is the same as a rotation by $-\pi$. This manifold is not simply connected. Consider a loop which goes along this line through the origin and then goes around the outside of the ball before joining up the other side of the loop. Moving along the line corresponds to rotating more or less and moving along the surface corresponds to changing the axis of rotation, essentially flipping the axis of rotation as we move from one side to the other. This is not contractible to a point and so SO(3) is not simply connected.

Since SO(3) and SU(2) have the same Lie algebra and SU(2) is the simply connected universal covering group we expect there to be an n-to-1 map $\mathrm{SU}(2) \to \mathrm{SO}(3)$. Indeed there is, for $n = 2$, and we call SU(2) a **double cover** of SO(3). One such map is given by the **Weyl homomorphism** which defines the function h to be $h(\boldsymbol{x}) = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$. This then satisfies

$$Uh(\boldsymbol{x})U^\dagger = h(O\boldsymbol{x}) \tag{11.3.11}$$

where U is any unitary 2×2 matrix and this equation defines O as some orthogonal 3×3 matrix. We can use this to define a map $U \mapsto O$, which is a map $\mathrm{SU}(2) \to \mathrm{SO}(3)$. The kernel of this map is $\{\pm\mathbb{1}\} \cong \mathbb{Z}_2$, and so

$$\mathrm{SO}(3) \cong \mathrm{SU}(2)/\mathbb{Z}_2 \tag{11.3.12}$$

by the first isomorphism theorem (Theorem 4.2.20).

Appendices

A

Mathematical Preliminaries

A.1 Basic Mathematics

A.1.1 Notation

Notation A.1.1 — Number Sets The set of natural numbers is

$$\mathbb{N} := \{0, 1, 2, \dots\}. \tag{A.1.2}$$

Note that the inclusion of zero in $\mathbb{N}$ is subject to debate. The set of integers is denoted

$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}. \tag{A.1.3}$$

The set of positive integers is denoted

$$\mathbb{Z}_{>0} := \{1, 2, \dots\}. \tag{A.1.4}$$

The set of rational numbers is denoted

$$\mathbb{Q} := \{p/q \mid p, q \in \mathbb{Z} \text{ and } q \neq 0\}. \tag{A.1.5}$$

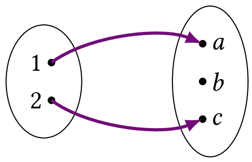
The set of real numbers is denoted $\mathbb{R}$, and the set of complex numbers $\mathbb{C}$. The set of *all* quaternions (as opposed to the quaternion group of order 8) is denoted $\mathbb{H}$.

Notation A.1.6 — Sphere The unit sphere in $n+1$ dimensions is

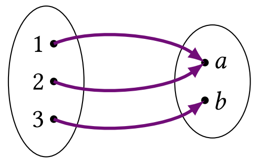
$$S^n := \{x \in \mathbb{R}^{n+1} \mid x_1^2 + \cdots x_{n+1}^2 = 1\}. \tag{A.1.7}$$

Note that S^n is an n-dimensional manifold, which we view as embedded in $(n+1)$-dimensional Euclidean space, $\mathbb{R}^{n+1}$.
What we normally call the circle is S^1 and what we normally call the sphere is S^2.

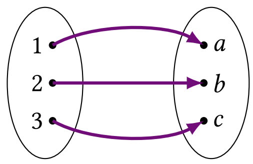
Notation A.1.8 — Sets of Matrices We denote the set of $m \times n$ matrices with entries in $\mathbb{F}$ (which is usually a field and usually $\mathbb{R}$ or $\mathbb{C}$) by $\mathcal{M}_{m\times n}(\mathbb{F})$.
We denote the set of square $n \times n$ matrices with entries in $\mathbb{F}$ by $\mathcal{M}_n(\mathbb{F})$.



(a) An injective function, $f\colon \{1,2\} \to \{a,b,c\}$. Note that $f(x) \neq b$ for any $x \in \{1,2\}$ and so the function fails to be surjective.



(b) A surjective function, $g\colon \{1,2,3\} \to \{a,b\}$. Note that $g(1) = g(2)$ but $1 \neq 2$ and so the function fails to be injective.



(c) A bijective function, $h\colon \{1,2,3\} \to \{a,b,c\}$.

Figure A.1: Injective, surjective, and bijective functions.

We denote the set of invertible $n\times n$ square matrices over $\mathbb{F}$, called the general linear group, by

$$\mathrm{GL}(n,\mathbb{F}) = \{A \in \mathcal{M}_n(\mathbb{F}) \mid \det A \neq 0\}. \tag{A.1.9}$$

If $\mathbb{F}$ is evident from context we may simply write $\mathrm{GL}(n)$. If V is an n-dimensional vector space over $\mathbb{F}$ then we may also write this set as $\mathrm{GL}(V)$.

Notation A.1.10 — Einstein Summation Convention When two identical indices appear in the same term then they are summed over, for example,

$$x_i y_i = \sum_i x_i y_i. \tag{A.1.11}$$

A.1.2 Definitions

Definition A.1.12 — Function Types Let $\varphi\colon A \to B$. Then φ is

- **injective** if for all $a, a' \in A$ $\varphi(a) = \varphi(a')$ implies $a = a'$,
- **surjective** if for all $b \in B$ there exists $a \in A$ such that $\varphi(a) = b$, and
- **bijective** if φ is both injective and surjective.

A function is invertible if and only if it is bijective.

Definition A.1.13 — Kernel Given a map $\varphi\colon A \to B$ the **kernel** is defined as the set of elements of A which map to the trivial element of B, which is the zero vector, $\mathbf{0}$, if B is a vector space, or the identity if B is a group:

$$\ker\varphi := \{a \in A \mid \varphi(a) \text{ is the trivial element of } B\} \subseteq A. \tag{A.1.14}$$

Definition A.1.15 — Image Given a map $\varphi\colon A \to B$ the **image** is set of $b \in B$ for which there exists some $a \in A$ such that $\varphi(a) = b$:

$$\operatorname{Im}\varphi = \varphi(A) := \{b \in B \mid \exists\, a \in A \text{ such that } \varphi(a) = b\} \subseteq B. \tag{A.1.16}$$

Definition A.1.17 — Empty Set The **empty set**, $\emptyset$, is the set containing no elements.

Definition A.1.18 — Kronecker Delta The **Kronecker delta**, δ_{ij}, is defined as

$$\delta_{ij} := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \tag{A.1.19}$$

Note that δ_{ij} are the elements of the identity matrix.

Definition A.1.20 — Levi-Civita Symbol The **Levi-Civita Symbol** in n-indices is the completely asymmetric (pseudo)tensor which is defined so that $\varepsilon_{123\dots n} := 1$. Antisymmetry then means that $\varepsilon_{1\cdots i\cdots j\cdots n} = -\varepsilon_{1\cdots j\cdots i\cdots n}$, for example $\varepsilon_{213\cdots n} = -1$. Antisymmetry also means that Levi–Civita symbol vanishes if it has repeated indices.
Most commonly $n = 3$ and

$$\varepsilon_{ijk} := \begin{cases} 1, & \text{if } (i,j,k) = (1,2,3), (2,3,1), (3,1,2), \\ -1, & \text{if } (i,j,k) = (1,3,2), (2,1,3), (3,2,1), \\ 0, & \text{if any index is repeated.} \end{cases} \tag{A.1.21}$$

Definition A.1.22 — Equivalence Relations Given two sets, A and B, a **relation**, R, is a subset of $A \times B := \{(a,b) \mid a \in A \text{ and } b \in B\} \supseteq R$. We say that $a \in A$ is related to $b \in B$, which we denote with infix notation, $a\,R\,b$, if $(a,b) \in R$.
If $A = B$ in the above definition then we call R a **binary!relation** on A.
A relation, $\sim$, on a set A is a binary relation on A such that the following axioms hold for all $a, b, c \in A$

- $\sim$ is **reflexive**, so $a \sim a$.
- $\sim$ is **symmetric**, so if $a \sim b$ then $b \sim a$.
- $\sim$ is **transitive**, so if $a \sim b$ and $b \sim c$ then $a \sim c$.

An **equivalence class** of an element $a \in A$ under some equivalence relation, $\sim$, is the set

$$[a] \coloneqq \{x \mid a \sim x\}. \tag{A.1.23}$$

We call elements of $[a]$ representatives of the equivalence class. We denote the set of all equivalence classes by $A/\sim$.

■ **Example A.1.24 — Equivalence Relations** $=$ is the prototypical equivalence relation.
Congruence modulo $m \in \mathbb{Z}_{>0}$ is an equivalence relation on $\mathbb{R}$.
$\sim$ defined by $z \sim w$ if $|z| = |w|$ is an equivalence relation on $\mathbb{C}$.
$\sim$ defined by $\boldsymbol{v} \sim \boldsymbol{u}$ if $\boldsymbol{u}$ and $\boldsymbol{v}$ are parallel is an equivalence relation on $\mathbb{R}^n$.

■ **Example A.1.25 — Isomorphism** Isomorphisms, as defined in the text, are equivalence relations:

- Let A be a group, then the identity function, $\mathrm{id}_A \colon A \to A$ defined by $\mathrm{id}_A(a) = a$ for all $a \in A$ is an isomorphism since $\mathrm{id}_A(aa') = aa' = \mathrm{id}_A(a)\mathrm{id}_A(a')$ and clearly id_A is invertible, and is its own inverse.
- Let A and B be isomorphic groups. Then there exists some bijection $\varphi \colon A \to B$ such that $\varphi(aa') = \varphi(a)\varphi(a')$. Since φ is a bijection $\varphi^{-1} \colon B \to A$ exists and is also a bijection. Applying the inverse to both sides of the defining relation we have $\varphi^{-1}(\varphi(aa')) = \varphi^{-1}(\varphi(a)\varphi(a'))$. Since φ is surjective any element of B can be written in the form $b = \varphi(a)$ for some $a \in A$ and so it follows that $\varphi^{-1}(\varphi(aa')) = \varphi^{-1}(b)\varphi(b')$ where $b, b' \in B$ are arbitrary, and we choose $a, a' \in A$ to be such that $b = \varphi(a)$ and $b' = \varphi(a')$. From the defining relation for φ we know that $\varphi(aa') = \varphi(a)\varphi(a') = bb'$. It follows that $\varphi^{-1}(\varphi(aa')) = \varphi^{-1}(bb') = \varphi^{-1}(b)\varphi^{-1}(b')$, which means that $B \cong A$.
- Let A, B, and C be groups such that $A \cong B$ and $B \cong C$. Then there exists isomorphisms $\varphi \colon A \to B$ and $\psi \colon B \to C$. We claim that $\psi \circ \varphi \colon A \to C$ is an isomorphism. Clearly $\psi \circ \varphi$ is bijective, since $\varphi^{-1} \circ \psi^{-1}$ is its inverse, as can be seen by considering $(\varphi^{-1} \circ \psi^{-1})((\psi \circ \varphi)(a)) = \varphi^{-1}(\psi^{-1}(\psi(\varphi(a)))) = \varphi^{-1}(\varphi(a)) = a$ for all $a \in A$.

  It remains to show that $\psi \circ \varphi$ is a homomorphism. To do so consider $(\psi \circ \varphi)(aa') = \psi(\varphi(aa')) = \psi(\varphi(a)\varphi(a'))$, which follows since φ is an isomorphism. Now write $\varphi(a) = b$ and $\varphi(a') = b'$, where $b, b' \in B$. We then have $(\psi \circ \varphi)(aa') = \psi(bb') = \psi(b)\psi(b')$, which follows since ψ is an isomorphism. We then have $(\psi \circ \varphi)(aa') = \psi(b)\psi(b') = \psi(\varphi(b))\psi\varphi(b') = (\psi \circ \varphi)(b)(\psi \circ \varphi)(b')$, and so $\psi \circ \varphi$ is a bijective homomorphism and hence an isomorphism, meaning $A \cong C$.

A.2 Linear Algebra

A.2.1 Vectors

Definition A.2.1 — Vector Space A vector space, V, over a field, $\mathbb{F}$, is a set of vectors, V, with two operations, $\cdot: \mathbb{F} \times V \longrightarrow V$, known as scalar multiplication, and $+: V \times V \longrightarrow V$, known as vector addition, which are defined such that the following hold for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and $k, k' \in \mathbb{F}$:

1. **Associativity**: $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$,
2. There exists a **zero vector**, $\mathbf{0} \in V$, such that $\boldsymbol{u} + \mathbf{0} = \boldsymbol{u}$.
3. There exists $-\boldsymbol{u} \in V$ such that $\boldsymbol{u} + (-\boldsymbol{u}) = \mathbf{0}$. We write this as $\boldsymbol{u} - \boldsymbol{u}$ for short.
4. **Commutativity**: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$.
5. Distributivity of scalar multiplication over vector addition $k(\boldsymbol{u} + \boldsymbol{v}) = k\boldsymbol{u} + k\boldsymbol{v}$.
6. Distributivity of scalar multiplication over field addition $(k + k')\boldsymbol{u} = k\boldsymbol{u} + k'\boldsymbol{u}$.
7. Compatibility of field and scalar multiplication $(kk')\boldsymbol{u} = k(k'\boldsymbol{u})$.
8. $1\boldsymbol{u} = \boldsymbol{u}$ where 1 is the multiplicative identity of $\mathbb{F}$.

Note that the first three axioms make $(V, +)$ a group and the fourth makes it Abelian.

Definition A.2.2 — Hilbert Space A **Hilbert space**, $\mathcal{H}$, is a vector space over either $\mathbb{R}$ or $\mathbb{C}$, equipped with an inner product that induces a complete metric. We shall assume a complex Hilbert space, for a real Hilbert space simply ignore any complex conjugates and replace $\mathbb{C}$ with $\mathbb{R}$.
An **inner product** is a function $\langle -, - \rangle : \mathcal{H} \times \mathcal{H} \longrightarrow \mathbb{C}$ such that for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{H}$ and $k, k' \in \mathbb{C}$

1. $\langle -, - \rangle$ is linear in its second argument, that is
$$\langle \boldsymbol{u}, k\boldsymbol{v} + k'\boldsymbol{w} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle + k'\langle \boldsymbol{u}, \boldsymbol{w} \rangle. \tag{A.2.3}$$
2. $\langle -, - \rangle$ is conjugate symmetric:
$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle^*. \tag{A.2.4}$$
3. $\langle -, - \rangle$ is positive definite, that is $\langle \boldsymbol{u}, \boldsymbol{u} \rangle \geq 0$ with equality only if $\boldsymbol{u} = \mathbf{0}$.

! Mathematicians often define an inner product to be linear in its first argument, so
$$\langle k\boldsymbol{u} + k'\boldsymbol{v}, \boldsymbol{w} \rangle = k\langle \boldsymbol{u}, \boldsymbol{w} \rangle + k'\langle \boldsymbol{v}, \boldsymbol{w} \rangle. \tag{A.2.5}$$

The first two axioms are sometimes combined to give an extra axiom that $\langle -, - \rangle$ is conjugate linear in its first argument (or second if we follow the other convention). That is

$$\langle k\boldsymbol{u} + k'\boldsymbol{v}, \boldsymbol{w} \rangle = k^* \langle \boldsymbol{u}, \boldsymbol{w} \rangle + k'^* \langle \boldsymbol{v}, \boldsymbol{w} \rangle. \tag{A.2.6}$$

We can then define a **norm** on $\mathcal{H}$ by $\|\boldsymbol{u}\| \coloneqq \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$.
The final condition for $\mathcal{H}$ to be a Hilbert space is completeness. Namely, that if the series $\sum_{n=0}^{\infty} \boldsymbol{u_n}$ converges absolutely, so that $\sum_{n=0}^{\infty} \|\boldsymbol{u_n}\|$ converges to a finite value then the original series, $\sum_{n=0}^{\infty} \boldsymbol{u_n}$, converges to some vector in $\mathcal{H}$.

■ **Example A.2.7** The set of n-tuples of complex numbers, $\mathbb{C}^n$, is a Hilbert space over $\mathbb{C}$ with the inner product

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle (u_1, \dots, u_n), (v_1, \dots, v_n) \rangle \coloneqq \sum_{i=1}^{n} u_i^* v_i = u_i^* v_i \tag{A.2.8}$$

where in the last term we use the Einstein summation convention.

■ **Example A.2.9 — Functions** The space of square integrable functions on $X \subseteq \mathbb{R}^n$ forms a Hilbert space, denoted $L^2(X)$. A function, $f \colon X \to \mathbb{C}$, is square integrable if[a]

$$\int_X |f(x)|^2 \, \mathrm{d}x \tag{A.2.10}$$

exists and is finite. For example, the function defined by $f(x) = e^{-x^2}$ is an element of $L^2(\mathbb{R})$.
Given $f, g \in L^2(X)$ we define the inner product in this space to be

$$\langle f, g \rangle \coloneqq \int_X f^*(x) g(x) \, \mathrm{d}x. \tag{A.2.11}$$

R Another subtly that arises here is that we actually need to consider elements of $L^2(X)$ to be equivalence classes of functions which are equal almost everywhere (meaning that the measure of the set of points where they are not equal is zero). Otherwise, we may have some functions such that $\langle f, g \rangle = 0$ but $f \neq g$ since f and g disagree on some set of points with a vanishing measure. We say that we are considering the functions mod the equivalence relation of being equal almost everywhere.

This is an important example since we can often identify “square integrable functions” with “possible wave functions”, since square-integrability is a requirement for us to be able to normalise a wave function, which we do so by the procedure

$$\psi \to \frac{\psi}{\|\psi\|} = \frac{\psi}{\sqrt{\langle \psi, \psi \rangle}} = \left(\int |\psi(x)|^2 \, \mathrm{d}x \right)^{-1/2} \psi. \tag{A.2.12}$$

This only makes sense if $\int |\psi(x)|^2 \, \mathrm{d}x$ is finite (and nonzero).

[a]For this space to be complete (a requirement for Hilbert spaces) this must be a Lebesgue integral but in physics functions are usually nice enough that we can use the standard Riemann integral, which agrees with the Lebesgue integral when both exists.

Notation A.2.13 — Bra-Ket Notation In physics, particularly in quantum mechanics, we often use **bra-ket notation**, developed by Dirac. We identify vectors, $\boldsymbol{u}$, with **kets**, $|u\rangle$, and dual vectors, $\boldsymbol{v}$, with **bras**, $\langle v|$. The inner product $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ is then written $\langle v|u\rangle$. This is the notation we will use for most of this course.

Definition A.2.14 — Linear Operator Given two vector spaces, V and W, over some field, $\mathbb{F}$, a function, $f\colon V \to W$, is said to be a linear operator if for $\boldsymbol{u}, \boldsymbol{v} \in V$ and $k \in \mathbb{F}$ we have

$$f(\boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}), \qquad \text{and} \qquad f(k\boldsymbol{u}) = kf(\boldsymbol{u}). \tag{A.2.15}$$

Instead of the function notation $f(\boldsymbol{u})$ we typically use a multiplicative notation, $A\boldsymbol{u}$, which is due to the fact that if V is finite dimensional then we can choose a basis and represent a linear map by a matrix. Using bra-ket notation also we have

$$A(|u\rangle + |v\rangle) = A|u\rangle + A|v\rangle, \qquad \text{and} \qquad A(k|u\rangle) = kA|u\rangle. \tag{A.2.16}$$

An operator is **antilinear** if

$$A(|u\rangle + |v\rangle) = A|u\rangle + A|v\rangle, \qquad \text{and} \qquad A(k|u\rangle) = k^*A|u\rangle. \tag{A.2.17}$$

An example of such an operator is the time reversal operator, T, which takes $t \to -t$.

For simplicity from now on we will consider the complex vector space $V = \mathbb{C}^N$. This is N-dimensional ($\dim V = N$), which we take to be finite, although many of these ideas work, possibly with slight modification, for infinite dimensional vector spaces. Most of the time we will also consider linear operators from V to V, since this is far more common in practice that linear operators from V to some different vector space, W.

Definition A.2.18 — Basis Given a vector space, V, we say that $\{|e_i\rangle\}$ is a **linearly independent** set if the only solution to

$$\lambda_i|e_i\rangle = |0\rangle, \tag{A.2.19}$$

where $|0\rangle$ is the zero vector, is $\lambda_i = 0$ for all i.
We say that $\{|e_i\rangle\}$ is a **basis** for V if $\{|e_i\rangle\}$ is a linearly independent set and spans V. That is given some $|u\rangle \in V$ we can write

$$|u\rangle = u_i|e_i\rangle \tag{A.2.20}$$

for some $u_i \in \mathbb{C}$.

The number of vectors in a basis is the **dimension** of the vector space, denoted $\dim V$.
We say that two vectors, $|u\rangle, |v\rangle \in V$, are **orthogonal** (with respect to some inner product) if $\langle u|v\rangle = 0$.
We say that a vector, $|u\rangle \in V$, is normalised if $\|u\| = \sqrt{\langle u|u\rangle} = 1$.
We say that $\{|e_i\rangle\}$ is an orthonormal basis for V if it is a basis for V, $|e_i\rangle$ is normalised for all i and all of the basis vectors are mutually orthogonal. These last two conditions are summarised by requiring that

$$\langle e_i|e_j\rangle = \delta_{ij}. \tag{A.2.21}$$

Definition A.2.22 — Completeness Relation Given a vector space, V, and an orthonormal basis, $\{|e_i\rangle\}$, we can write the identity operator, $\mathbb{1}$, as

$$\mathbb{1} = \sum_{i=1}^{N} |e_i\rangle\langle e_i|. \tag{A.2.23}$$

Recall that the identity operator is defined such that

$$\mathbb{1}|u\rangle = |u\rangle \tag{A.2.24}$$

for all $|u\rangle \in V$.

Definition A.2.25 — Partition of the Identity A **projection operator**, P_i, is an operator satisfying

$$P_i P_j = \delta_{ij} P_i, \qquad \text{and} \qquad P_i^\dagger = P_i. \tag{A.2.26}$$

A **partition of the identity** is a collection of projection operators, $\{P_j\}$, such that

$$\mathbb{1} = \sum_{j=1}^{N_P} P_j \tag{A.2.27}$$

where $N_P = |\{P_j\}|$ is the number of projection operators in the partition. We can write each partition operator as

$$P_j = \sum_{i=1}^{N_j} |e_i\rangle\langle e_i| \tag{A.2.28}$$

where N_j are such that

$$\dim V = N = \sum_{j=1}^{N_P} N_j. \tag{A.2.29}$$

Definition A.2.30 — Matrix Element Given a linear operator, $A\colon V \to V$, and an orthonormal basis, $\{|e_i\rangle\}$, we define the **matrix elements** to be

$$A_{ij} := \langle e_i|Ae_j\rangle = \langle e_i|A|e_j\rangle \tag{A.2.31}$$

where A is understood to act on the right and $|Ae_j\rangle := A|e_j\rangle$.
If we know the matrix elements of A we can reconstruct A using

$$A = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}|e_i\rangle\langle e_j|. \tag{A.2.32}$$

Definition A.2.33 — Eigenvalues and Eigenvectors Given a linear operator A, we call $|v_i\rangle$ an **eigenvector** and $\lambda_i \in \mathbb{C}$ an **eigenvalue** if

$$A|v_i\rangle = \lambda_i|v_i\rangle. \tag{A.2.34}$$

There are N solutions to this, which follows from the **characteristic polynomial**, $\det(A - \lambda\mathbb{1}) = 0$, having N solutions, which in turn follows from the fundamental theorem of algebra.

A.2.2 Matrices

Definition A.2.35 — Transpose and Hermitian Conjugate Given a matrix, A, with matrix elements A_{ij}, the **transpose** matrix, $A^\top$, has matrix elements $A^\top_{ij} = A_{ji}$.
A matrix is **symmetric** if $A^\top = A$, or **antisymmetric** if $A^\top = -A$.
Given a matrix, A, with matrix elements A_{ij}, the **Hermitian conjugate**, $A^\dagger$, has matrix elements $A^\dagger_{ij} = A^*_{ji}$. Here * denotes the **complex conjugate**, so $(x+iy)^* = x - iy$ and $(re^{i\vartheta})^* = re^{-\vartheta}$ for $x, y, r, \vartheta \in \mathbb{R}$. That is the Hermitian conjugate is the complex conjugate of the transpose.
A matrix is **Hermitian** if $A^\dagger = A$, or **anti-Hermitian** if $A^\dagger = -A$.

Lemma A.2.36 The eigenvalues of a Hermitian matrix are real.

Proof. Let A be a Hermitian matrix and $\boldsymbol{v}$ an eigenvector with nonzero eigenvalue λ. Note that this means $\boldsymbol{v}$ is nonzero. If 0 is an eigenvalue of A then this is real, so we need not consider this case further. By definition $A\boldsymbol{v} = \lambda\boldsymbol{v}$. Taking the Hermitian conjugate of both sides we get $\boldsymbol{v}^\dagger A^\dagger = \lambda^*\boldsymbol{v}^\dagger$, where we have used $(XY)^\dagger = Y^\dagger X^\dagger$. Multiplying both sides on the right by $\boldsymbol{v}$ we get $\boldsymbol{v}^\dagger A\boldsymbol{v} = \lambda^*\boldsymbol{v}^\dagger\boldsymbol{v}$. Identifying $A\boldsymbol{v} = \lambda\boldsymbol{v}$ on the left-hand side this becomes $\boldsymbol{v}^\dagger\lambda\boldsymbol{v} = \lambda\boldsymbol{v}^\dagger\boldsymbol{v} = \lambda^*\boldsymbol{v}^\dagger\boldsymbol{v}$. It follows that we must have $\lambda = \lambda^*$, which means we must have $\lambda \in \mathbb{R}$. □

We can choose the eigenvalues of a Hermitian matrix to be orthonormal, and hence they form a basis for the vector space. In this basis the matrix will be diagonal and the values on the diagonal are simply the eigenvalues.

Given a Hermitian matrix, A, with eigenvalues λ_i and corresponding eigenvectors $|v_i\rangle$ we can write this matrix as

$$A = \sum_{i=1}^{N} \lambda_i |v_i\rangle\langle v_i|. \tag{A.2.37}$$

This is diagonalised by the transformation $V^\dagger A V$ where

$$V = \sum_{i=1}^{N} |v_i\rangle\langle e_i| \tag{A.2.38}$$

where $|e_i\rangle$ are the basis vectors in the original basis. It is easy to see that this transform gives the desired result:

$$V^\dagger A V = \underbrace{|e_i\rangle\langle v_i|}_{=V^\dagger} \underbrace{(\lambda_j |v_j\rangle\langle v_j|)}_{=A} \underbrace{|v_k\rangle\langle e_k|}_{=V} \tag{A.2.39}$$

$$= \lambda_j |e_i\rangle\langle v_i|v_j\rangle\langle v_j|v_k\rangle\langle e_k| \tag{A.2.40}$$

$$= \lambda_j \delta_{ij}\delta_{jk} |e_i\rangle\langle e_k| \tag{A.2.41}$$

$$= \lambda_i |e_i\rangle\langle e_i| \tag{A.2.42}$$

This last term is just a diagonal matrix with the eigenvalues, λ_i, as the diagonal elements, which is exactly what we wanted.

For non-Hermitian matrices it is possible that the eigenvalues aren't linearly independent. In this case the best we can do is Jordan normal form where the eigenvalues are on the diagonal and all other entries are either zero or one for elements in the subspace of degenerate eigenvalues.

Definition A.2.43 — Inverse The **inverse** of a matrix, A, is the matrix A^{-1} such that $A^{-1}A = AA^{-1} = \mathbb{1}$. Such a matrix exists only if the determinant is non-zero.

An equivalent requirement for A^{-1} to exist is for A to have no zero eigenvalues. For a Hermitian matrix the inverse in the eigenbasis is simply

$$A^{-1} = \mathrm{diag}(1/\lambda_1, \dots, 1/\lambda_N). \tag{A.2.44}$$

Definition A.2.45 — Orthogonal and Unitary A matrix, O, is **orthogonal** if $O^\top O = \mathbb{1}$, that is $O^\top = O^{-1}$.
A matrix, U, is **unitary** if $U^\dagger U = \mathbb{1}$, that is $U^\dagger = U^{-1}$.

The following holds:

$$\langle u|Av\rangle = \langle u|A|v\rangle = \langle A^\dagger u|v\rangle. \tag{A.2.46}$$

For a unitary matrix, U, this implies

$$\langle Uu|Uv\rangle = \langle u|U^\dagger U|v\rangle = \langle u|\mathbb{1}|v\rangle = \langle u|v\rangle. \tag{A.2.47}$$

We say that unitary transforms preserve the inner product, or that the inner product is invariant under unitary transforms.

Definition A.2.48 — Trace The **trace** of a matrix, A is

$$\operatorname{tr} A := \sum_i \langle e_i | A_i | i \rangle = A_{ii} \tag{A.2.49}$$

where in the last term we are using the Einstein summation convention to sum over i.

The trace of a matrix is simply the sum of its eigenvalues, this doesn't just hold for Hermitian matrices.

The trace is cyclic, meaning $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$, etc.

The trace is linear, meaning $\operatorname{tr}(kA) = k \operatorname{tr}(A)$ for scalar k.

$\langle A, B \rangle := \operatorname{tr}(A^\dagger B)$ is an inner product on the vector space of matrices. This is called the **Gram–Schmidt inner product**.

Definition A.2.50 — Determinant The **determinant** of a matrix, A, is

$$\det A = |A| =:= \varepsilon_{i_1 \dots i_N} A_{1 i_1} \cdots A_{N i_N} \tag{A.2.51}$$

with summation over indices implied.

The determinant of a matrix is the product of its eigenvalues.

The determinant of a product is the product of the determinants:

$$\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA). \tag{A.2.52}$$

Definition A.2.53 — Diagonal A matrix, A, is **diagonal** if $A_{ij} = 0$ for $i \neq j$. A matrix, A, is **block diagonal** if it can be written in the form

$$A = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & A_n \end{pmatrix} \tag{A.2.54}$$

where A_i are square matrices and the 0s represent matrices where all elements are zero.

A.2.3 Combining Vector Spaces

Definition A.2.55 — Direct Sum Given vector spaces V and W we call $V \oplus W$ the **direct sum**. It is defined by associating with each pair of vectors, $|v_i\rangle \in V$ and $|w_a\rangle \in W$, a vector $|v_i\rangle \oplus |w_i\rangle = |v_i \oplus w_a\rangle \in V \oplus W$ and extending the inner product to

$$\langle v_i \oplus w_a | v_j \oplus w_b \rangle_{V \oplus W} := \langle v_i | v_j \rangle_V + \langle w_a | w_b \rangle_W \tag{A.2.56}$$

where the subscripts denote which vector space the inner product is in. Note that the notation $|v \oplus w\rangle$ is non-standard.

The dimension of $V \oplus W$ is

$$\dim(V \oplus W) = \dim V + \dim W. \tag{A.2.57}$$

Given $A \in \mathrm{GL}(V)$ and $B \in \mathrm{GL}(W)$ the direct sum, $A \oplus B$, acts on $|v\rangle \oplus |w\rangle \in V \oplus W$ as

$$(A \oplus B)(|v\rangle \oplus |w\rangle) := (A|v\rangle) \oplus (B|w\rangle). \tag{A.2.58}$$

This shows we can think of $V \oplus W$ as a $(\dim V + \dim W)$-dimensional vector space with operators represented by $(v+w) \times (v+w)$ block diagonal matrices:

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \tag{A.2.59}$$

We then think of $|v\rangle \oplus |w\rangle \in V \oplus W$ as $(v_1, \dots, v_{\dim V}, w_1, \dots, w_{\dim W})$.

An important question is can a given vector space be written as a direct sum of vector spaces, this occurs when considering the irreducibility of representations.

Definition A.2.60 — Direct Product Given vector spaces V and W we call $V \otimes W$ the **direct product**. It is defined by associating with each pair of vectors, $|v_i\rangle \in V$ and $|w_a\rangle \in W$, a vector $|v_i\rangle \otimes |w_a\rangle = |v_i \otimes w_a\rangle \in V \otimes W$ and extending the inner product to

$$\langle v_i \otimes w_a | v_j \otimes w_b \rangle_{V \otimes W} := \langle v_i | v_j \rangle_V \langle w_a | w_b \rangle \tag{A.2.61}$$

where the subscripts denote which vector space the inner product is in. Note that the notation $|v \otimes w\rangle$ is non-standard.
The dimension of $V \otimes W$ is

$$\dim(V \otimes W) = \dim(V)\dim(W). \tag{A.2.62}$$

Given $A \in \mathrm{GL}(V)$ and $B \in \mathrm{GL}(W)$ the direct product, $A \otimes B$, acts on $|v\rangle \otimes |w\rangle \in V \otimes W$ as

$$(A \otimes B)(|v\rangle \otimes |w\rangle) = (A|v\rangle) \otimes (B|w\rangle). \tag{A.2.63}$$

■ **Application A.2.64** In quantum mechanics we can combine states from different Hilbert spaces representing different properties with direct products. For example, given an electron wave function with a spatial component and a spin component the direct product of these gives the state of the electron.

The direct product plays a role in representation theory in terms of what we will call Kronecker products. These can be used to obtain all representations from the fundamental representations.

B

Groups

B.1 Finite Groups

Definition B.1.1 The **trivial group** is the group containing only the identity, $\{e\}$. It is the only group of order 1.

- Order 1.
- Rank 1.
- Cyclic.
- Abelian.

The trivial group is isomorphic to $\mathbb{Z}_1$, S_1, and $\mathrm{SO}(1)$.

Definition B.1.2 $\mathbb{Z}_2$ is the cyclic group of order 2, see Definition B.1.9. It is the only group of order 2.

- Order 2.
- Rank 1.
- Cyclic.
- Abelian.

$\mathbb{Z}_2$ is isomorphic to S_2.

Definition B.1.3 $\mathbb{Z}_3$ is the cyclic group of order 3, see Definition B.1.9. It is the only group of order 3.

- Order 3.
- Rank 1.
- Cyclic.
- Abelian.

Definition B.1.4 $\mathbb{Z}_4$ is the cyclic group of order 4, see Definition B.1.9. It is one of two groups of order 4.

- Order 4.
- Rank 1.
- Cyclic.
- Abelian.

Definition B.1.5 $\mathbb{Z}_2 \times \mathbb{Z}_2$ is the **Klein *Vierergruppe***. It is one of two groups of order 4. It is a direct product of two copies of $\mathbb{Z}_2$.

- Order 4.
- Abelian.
- Rank 2.

Definition B.1.6 S_3 is the permutation group on 3 elements, see Definition B.1.11.

- Order 6.
- Non-Abelian.
- Rank 2.

Definition B.1.7 The **quaternion group**, Q, has the group presentation

$$Q = \langle -e, i, j, k \mid (-e)^2 = e, i^2 = j^2 = k^2 = ijk = e \rangle. \tag{B.1.8}$$

- Order 8.
- Non-Abelian.
- Rank 2.

The Pauli matrices provide a two-dimensional complex representation by the correspondence $(-e, i, j, k) \longrightarrow (-\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$.

B.1.1 Other Finite Groups

Definition B.1.9 The **cyclic group** of order n, denoted $\mathbb{Z}_n$, is given by the presentation

$$\mathbb{Z}_n = \langle a \mid a^n = e \rangle. \tag{B.1.10}$$

Identifying $a = e^{2i\pi/n}$ and the operation as multiplication we get a group formed from the nth roots of unity. Identifying $a = 1$ and the operation as addition modulo n we get a group formed from $\{0, \dots, n-1\}$.

- Order n.
- Cyclic.
- Rank 1.
- Abelian.

All finite cyclic groups are isomorphic to $\mathbb{Z}_n$ for some n.

Definition B.1.11 The **permutation group** on n objects is the group of all permutations (bijections) of $\{1, \dots, n\}$, with function composition as the group operation.

- Order $n!$.
- Non-Abelian ($n > 2$).
- Rank 2.

S_1 and S_0 are isomorphic to the trivial group.
S_2 is isomorphic to $\mathbb{Z}_2$.

B.2 Discrete Groups

Definition B.2.1 The integers, $\mathbb{Z}$, under addition.

- Rank 1.
- Cyclic.
- Abelian.

Definition B.2.2 The rational numbers, $\mathbb{Q}$, under addition.

- Abelian.

Definition B.2.3 The nonzero rational numbers, $\mathbb{Q}^*$, under multiplication.

- Abelian.

B.3 Continuous Groups

B.3.1 Scalars

Definition B.3.1 The real numbers, $\mathbb{R}$, under addition.

- Abelian.

$(\mathbb{R}, +)$ is isomorphic to $(\mathbb{R}_{>0}, \cdot)$.

Definition B.3.2 The nonzero real numbers, $\mathbb{R}^*$, under multiplication.

- Abelian.

Definition B.3.3 The complex numbers, $\mathbb{C}$, under addition.

- Abelian.

Definition B.3.4 The nonzero complex numbers, $\mathbb{C}^*$, under multiplication.

- Abelian.

B.3.2 Matrices

Definition B.3.5 The **general linear group**

$$\mathrm{GL}(n, \mathbb{F}) = \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M \neq 0\}\,. \tag{B.3.6}$$

If V is a vector space of dimension n over $\mathbb{F}$ then this group is also denoted $\mathrm{GL}(V)$. If $\mathbb{F}$ is obvious from context then this group is denoted $\mathrm{GL}(n)$.

- Non-Abelian ($n > 1$).

Definition B.3.7 The **special linear group**

$$\mathrm{SL}(n, \mathbb{F}) = \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M = 1\}. \tag{B.3.8}$$

If V is a vector space of dimension n over $\mathbb{F}$ then this group is also denoted $\mathrm{SL}(V)$. If $\mathbb{F}$ is obvious from context then this group is denoted $\mathrm{SL}(n)$.

- Non-Abelian ($n > 1$).

$\mathrm{SL}(n, \mathbb{F})$ is a subgroup of $\mathrm{GL}(n, \mathbb{F})$.

Definition B.3.9 The **orthogonal group**

$$\mathrm{O}(n) = \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = OO^\top = \mathbb{1}\}. \tag{B.3.10}$$

- Non-Abelian ($n > 1$).

$\mathrm{O}(n)$ is a subgroup of $\mathrm{GL}(n, \mathbb{R})$.
$\mathrm{O}(n)$ is the group of distance preserving transformations of Euclidean space which leave the origin invariant.
$\mathrm{O}(n)$ is the group of rotations and inversions of $\mathbb{R}^n$.

Definition B.3.11 The **special orthogonal group**

$$\mathrm{SO}(n) = \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = OO^\top = \mathbb{1} \text{ and } \det O = 1\}. \tag{B.3.12}$$

- Non-Abelian ($n > 1$).

$\mathrm{SO}(n)$ is a subgroup of $\mathrm{O}(n)$ and $\mathrm{SL}(n, \mathbb{R})$.
$\mathrm{SO}(n)$ is the group of rotations of $\mathbb{R}^n$.
$\mathrm{SO}(2)$ is isomorphic to $\mathrm{U}(1)$ and the circle group, $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ under multiplication.

Definition B.3.13 The **unitary group**

$$\mathrm{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1}\}. \tag{B.3.14}$$

- Non-Abelian ($n > 1$).

$\mathrm{U}(n)$ is a subgroup of $\mathrm{GL}(n, \mathbb{C})$.
$\mathrm{U}(n)$ is the group which preserves the standard inner product on $\mathbb{C}^n$.
$\mathrm{U}(1)$ is isomorphic to $\mathrm{SO}(2)$ and the circle group, $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ under multiplication.

Definition B.3.15 The **special unitary group**

$$\mathrm{SU}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1} \text{ and } \det U = \mathbb{1}\}. \tag{B.3.16}$$

- Non-Abelian ($n > 1$).

$\mathrm{SU}(n)$ is a subgroup of $\mathrm{U}(n)$ and $\mathrm{SL}(n, \mathbb{C})$.

Definition B.3.17 The **isometries of Euclidean space**

$$\mathrm{ISO}(n) = \mathrm{O}(n) \ltimes \mathbb{R}^n \tag{B.3.18}$$

where $(R, \boldsymbol{a})(R', \boldsymbol{a}') \coloneqq (RR', \boldsymbol{a} + R\boldsymbol{a}')$.

- Non-Abelian

$\mathrm{ISO}(n)$ is the group of distance preserving transformations of Euclidean space.
$\mathrm{ISO}(n)$ is the group of rotations, reflections, and translations of $\mathbb{R}^n$.
$\mathrm{ISO}(n)$ has both $\mathrm{O}(n)$ and $\mathbb{R}^n$ as normal subgroups.

Definition B.3.19 The **Lorentz group**, $\mathrm{O}(1, 3)$, is the group of all Lorentz transformations of Minkowski space.

- Non-Abelian.

$\mathrm{O}(1, 3)$ is the group that preserves the quadratic form $(t, x, y, z) \mapsto t^2 - x^2 - y^2 - z^2$.
$\mathrm{SO}^+(1, 3)$ is the subgroup of $\mathrm{O}(1, 3)$ which preserves the orientation of space (S for special, that is unit determinant) and direction of time (that's what the + represents).

Definition B.3.20 The **Poincaré group** is the group of all isometries of Minkowski space, sometimes denoted $\mathrm{ISO}(1, 3)$. That is, it is the group of all Lorentz transformations and translations.

- Non-Abelian.

The Poincaré group can be identified as the semidirect product $\mathrm{ISO}(1, 3) =$

$\mathbb{R}^{1,3} \rtimes \mathrm{O}(1,3)$ where $\mathbb{R}^{1,3}$ is the group of spacetime translations of Minkowski space and $\mathrm{O}(1,3)$ is the Lorentz group.

C

Manifolds

R The notes in this section are repeated from the notes for general relativity so for more details look at those notes.

The theory of Lie groups starts by defining them as manifolds with a group structure, or groups with a manifold structure. In order to make this rigorous we need to define a manifold. We will define a real manifold for simplicity and the definition of a complex manifold is nearly identical replacing $\mathbb{R}$ with $\mathbb{C}$ and requiring maps be holomorphic instead of smooth.

I go into far more detail here than necessary. I also make statements like "**Top** is the category of topological spaces with continuous functions as morphisms and homeomorphisms as isomorphisms" without explanation, these sorts of statement aren't important if you don't understand them.

C.1 Manifolds

Manifolds are the stage for the maths of GR, and in fact most physics. We start with a rough definition of a manifold which lacks a few details which we will fill out later. For our purposes this rough definition is sufficient and the full definition is for completeness.

> **Definition C.1.1 — Manifold** A **manifold** is a space where we can locally introduce Cartesian coordinates such that different choices of local coordinates will give compatible descriptions and by defining sets of local coordinates covering the whole space we have complete information about the space.

A manifold is parametrised continuously and differentiably as an n-tuple of real numbers, $(x^1, \dots, x^n) \in \mathbb{R}^n$, which are the coordinates of the point. We can connect two points with a curve parametrised by some $\lambda \in \mathbb{R}$ such that derivatives of the coordinates along this path exist, that is $\mathrm{d}x^i/\mathrm{d}\lambda$ exists for all $i = 1, \dots, n$.

By this definition it should be clear that Euclidean space, where non-relativistic mechanics takes place, is a manifold. However, not all manifolds are as nice as Euclidean space, for example Euclidean space has a metric (way to measure distance), which not all manifolds do. Locally all manifolds behave like a subset of Euclidean space due to their differentiable nature.

C.1.1 Topology

Definition C.1.2 — Topological Space A **topological space**, $(X, \mathcal{T})$, is a set, X, and a collection of subsets of X, $\mathcal{T} \subset \mathcal{P}(X)$, called the **topology**, such that

- $X \in \mathcal{T}$,
- $\emptyset \in \mathcal{T}$,
- $\mathcal{T}$ is closed under unions, that is if $\{U_i\}_{i\in I} \subseteq \mathcal{T}$ then
$$\bigcup_{i\in I} U_i \in \mathcal{T}, \tag{C.1.3}$$
- $\mathcal{T}$ is closed under finite intersections, that is if $\{U_i, \dots, U_N\} \subseteq \mathcal{T}$ then
$$\bigcap_{i=1}^{N} U_i \in \mathcal{T}. \tag{C.1.4}$$

We call the elements of $\mathcal{T}$ **open sets**.

It is common to refer to X as a topological space and then specify $\mathcal{T}$ separately, or just leave $\mathcal{T}$ implicit, much in the same way that we might define a group G and operation $\cdot$, when more formally it might be more correct to call $(G, \cdot)$ and G the underlying set.

Topological spaces can be quite abstract, for example the following topology makes $(\{a, b, c, d\}, \mathcal{T})$ a topological space:

$$\mathcal{T} = \{\emptyset, \{a, b, c, d\}, \{a, b\}, \{c, d\}\}. \tag{C.1.5}$$

We only deal with more familiar topological spaces. We can define different topologies using a basis.

Definition C.1.6 — Topological Basis A **basis** for a topology on some set X is a collection of sets, $\mathcal{B} \subseteq \mathcal{P}(X)$ such that

- the elements of $\mathcal{B}$ cover X, that is for all $x \in X$ there exists some $B \in \mathcal{B}$ such that $x \in B$.
- given $B_1, B_2 \in \mathcal{B}$ for all $x \in B_1 \cap B_2$ there exists some $B \in \mathcal{B}$ such that $x \in B$ and $B_3 \subseteq B_1 \cap B_2$.

The topology generated by a basis, $\mathcal{B}$, is then defined to be the set of all subsets $U \subseteq X$ such that

$$U = \bigcup_{B_i \in B \subseteq \mathcal{B}} B_i. \tag{C.1.7}$$

That is all elements of the topology are unions of basis sets.

Using this definition we can define the **standard topology** on $\mathbb{R}^n$ as the topology generated by the basis of open balls in $\mathbb{R}^n$, where by an open ball we mean

$$B = \{\boldsymbol{r} \in \mathbb{R}^n \mid |\boldsymbol{r} - \boldsymbol{a}|^n \text{ for some } \boldsymbol{a} \in \mathbb{R}^n\} \tag{C.1.8}$$

with $|\boldsymbol{r}| = \sum_i x_i^2$ being the usual Euclidean metric. This is not the only topology on $\mathbb{R}^n$ but it is the only one we care about.

Let X and Y be topological spaces with topologies $\mathcal{T}_X$ and $\mathcal{T}_Y$ respectively. Then we can define a function, $f\colon X \to Y$. The **preimage** of some $S \subseteq Y$ under this function is the set

$$f^*(S) = f^{-1}(S) \coloneqq \{a \in X \mid f(a) \in S\} \subseteq X. \tag{C.1.9}$$

The notation f^{-1} suggests that this is related to the inverse of the function f, and indeed if the inverse exists then it is, as the preimage is the image of the inverse restricted to S.

The function $f\colon X \to Y$ is **continuous** if all preimages of open sets of Y are open sets of X, that is if $f^*(U_Y) \in \mathcal{T}_X$ for all $U_Y \in \mathcal{T}_Y$. With the standard topology for $\mathbb{R}^n$ this definition coincides with the normal ε-δ definition of continuity from analysis.

An invertible (and hence bijective) continuous function between two topological spaces is called a **homeomorphism**. If there exist homeomorphisms between two topological spaces we say those spaces are **homeomorphic**. **Top** is the category of topological spaces with continuous functions as morphisms and homeomorphisms as isomorphisms. This means that continuous functions and homeomorphisms are to topological spaces as homomorphisms and isomorphisms are to groups, since **Grp** is the category of groups with group homomorphisms as morphisms and group isomorphisms as isomorphisms.

C.1.2 Charts

In order to be able to define differentiation in topological spaces we need them to look locally like $\mathbb{R}^n$. By "look like" we mean we want there to be a homeomorphism between neighbourhoods of a topological space and $\mathbb{R}^n$. Homeomorphisms preserve topological properties so this allows us to translate things into Euclidean terms, perform calculations, and then convert the results back to the topological space. This isn't possible for all topological spaces but is a defining property of manifolds.

To better define the notion of "looks like" we define charts.

Definition C.1.10 — Chart Given a topological space, $(X, \mathcal{T})$, a **chart**, C, is an ordered pair $C = (U, \varphi)$, where $U \in \mathcal{T}$ and φ is a homeomorphism

$$\varphi\colon U \to \operatorname{im}\varphi \subseteq \mathbb{R}^n. \tag{C.1.11}$$

We call $n \in \mathbb{Z}_{>0}$ the **dimension** of U, it is independent of φ.

Simply put a chart associates with each neighbourhood in a topological space an open subset of $\mathbb{R}^n$ and gives us a way to move between U and this open subset while preserving key topological properties.

C.1.3 Topological Manifolds

For our first definition of a manifold we will need a few preliminary definitions.

Definition C.1.12 — Locally Euclidean A topological space, X, is **locally Euclidean** if there exists $n \in \mathbb{Z}_{>0}$ such that for every point, $p \in X$ there is a chart, $C = (U, \varphi)$, such that $p \in U$.

That is X must be covered by the open sets $\{U_i\}_{i\in I}$ and for each open set in this covering there must be an associated chart, $C_i = (U_i, \varphi_i)$.

> **Definition C.1.13 — Hausdorff** Let X be a topological space. Then $x, y \in X$ are separated by neighbourhoods if there exists some open sets $U \ni x$ and $V \ni y$ such that $U \cap V = \emptyset$.
> X is a **Hausdorff space** if all distinct points in X are pairwise neighbourhood separable.

Roughly put the condition of being Hausdorff means for any two points we can find sufficiently small neighbourhoods around the two points such that the two points aren't both in the neighbourhoods. Take for example, $\mathbb{R}^2$, the plane with the standard topology. Since the standard topology here is the topology generated by the open discs we can draw two discs around any two points such that the discs don't overlap and so $\mathbb{R}^2$ is Hausdorff.

> **Definition C.1.14 — Topological Manifold** A **topological manifold** is a Hausdorff space, M, which is locally Euclidean.

It is common to include extra conditions, such as the topology being **second-countable**, which means the topology can be generated by a countable basis.

A topological manifold is the simplest manifold, and all manifolds are topological manifolds. However, it is most common to work with smooth manifolds, which we will develop in the next few sections.

C.1.4 Local Coordinates

Charts allow us to rigorously define the notion local coordinates. Intuitively local coordinates are a way to parametrise a neighbourhood with tuples in $\mathbb{R}^n$, formally local coordinates are defined as follows.

> **Definition C.1.15 — Local Coordinates** Let M be a topological manifold and U an open subset of M such that (U, φ) is a chart. The coordinates of some point $p \in U$ are the Cartesian coordinates, $\boldsymbol{x}_p$, given by $\varphi(p) \in \mathbb{R}^n$.

In general charts are not unique and it is important that the physics we do doesn't depend on the choice of coordinate system, since coordinates are a man made construction and don't actually reflect nature, they just let us apply maths. This means we need charts to be compatible in the following sense.

> **Definition C.1.16 — Compatible Charts** Let M be a topological manifold with charts $C_1 = (U_1, \varphi_1)$ and $C_2 = (U_2, \varphi_2)$. If $U_1 \cap U_2 \neq \emptyset$ we can define the **transition maps** to be
>
> $$\varphi_1 \circ \varphi_2^{-1} \colon \varphi_2(U_1 \cap U_2) \to \varphi_1(U_1 \cap U_2), \tag{C.1.17}$$
> $$\varphi_2 \circ \varphi_1^{-1} \colon \varphi_1(U_1 \cap U_2) \to \varphi_2(U_1 \cap U_2). \tag{C.1.18}$$
>
> We say that C_1 and C_2 are **compatible** if either the transition maps are homeomorphisms or $U_1 \cap U_2 = \emptyset$.

We say that C_1 and C_2 are **smoothly compatible** if they are compatible and the transitions functions, if defined, are smooth.

We can think of the two charts as two separate sets of local coordinates and the transition maps as coordinate transformations, that is

$$(\varphi_1 \circ \varphi_2^{-1})(\boldsymbol{x}_p^{(2)}) = \varphi_1(p) = \boldsymbol{x}_p^{(1)}, \tag{C.1.19}$$

$$(\varphi_2 \circ \varphi_1^{-1})(\boldsymbol{x}_p^{(1)}) = \varphi_1(p) = \boldsymbol{x}_p^{(2)}. \tag{C.1.20}$$

We can use local coordinates to express functions on some manifold in terms of local coordinates.

Definition C.1.21 — Local Function Given a continuous function, $f : M \to \mathbb{R}$ for some topological manifold, M, and a chart, $C = (U, \varphi)$, we can associate the restriction of f to U with the function $f_U = f \circ \varphi^{-1}$ so

$$f_U : \varphi(U) \subseteq \mathbb{R}^n \to \mathbb{R} \tag{C.1.22}$$

is defined by

$$f(p) = f_U(\boldsymbol{x}_p) \tag{C.1.23}$$

for $p \in U$.

This is useful since we can deal with functions $\mathbb{R}^n \to \mathbb{R}$ using the normal tools of vector calculus.

For compatible charts we can represent the function in two different ways in terms of local coordinates. namely by the functions f_{U_i} with $i = 1, 2$. On the intersection $U_1 \cap U_2$ we define $f = f_1 \circ \varphi_1 = f_2 \circ \varphi_2$ so rearranging this final equality $f_2 = f_1 \circ \varphi_1 \circ \varphi_2^{-1}$ and $f_1 = f_2 \circ \varphi_2 \circ \varphi_1^{-1}$. We can think of these as the same function with a change of variables, so $f_1(\boldsymbol{x}_p^{(1)}) = f_2(\boldsymbol{x}_p^{(2)})$.

C.1.5 Smooth Manifold

Definition C.1.24 — Smooth Atlas A **smooth atlas**, $\mathcal{A}(M)$, for a topological manifold, M, is a family of charts, $\{C_a = (U_a, \varphi_a)\}$, which cover M, so $p \in U_a$ for some a for all $p \in M$, such that the charts are all mutually smoothly compatible.
Two smooth atlases, $\mathcal{A}_1(M)$ and $\mathcal{A}_2(M)$, are said to be **compatible** if all of the charts of $\mathcal{A}_1$ are compatible with all of the charts of $\mathcal{A}_2$.

We can think of charts as mapping the manifold and then the name atlas is naturally a collection of maps. Compatibility defines an equivalence relation on the set of all atlases for a given manifold.

Definition C.1.25 — Smooth Manifold A **smooth structure** on a topological manifold, M, is an equivalence class of smooth atlases:

$$\mathcal{S}(M) = [\mathcal{A}(M)], \tag{C.1.26}$$

that is $\mathcal{S}(M)$ is the set of all atlases compatible with $\mathcal{A}(M)$.
A **smooth manifold** is a topological manifold, M, equipped with a smooth structure, $\mathcal{S}$.

Definition C.1.27 — Smooth Functions and Maps A function, $f: M \to \mathbb{R}$, on a smooth manifold, M, is said to be **smooth** if all of its local coordinate representatives,

$$f_a = f \circ \varphi_a^{-1}: \varphi(U_a) \to \mathbb{R}, \tag{C.1.28}$$

are smooth.
Let M and V be smooth manifolds of dimensions m and n with charts (U_a, φ_a) and (V_b, ψ_b) respectively. Then the map $\mu: M \to N$ is **smooth** if all of its local coordinate representatives,

$$\mu_{ab} = \psi_b \circ \mu \circ \varphi_a^{-1}: \varphi_a(U_a) \subseteq \mathbb{R}^m \to \psi_b(V_b) \subseteq \mathbb{R}^n \tag{C.1.29}$$

are smooth. Smooth invertible maps between manifolds are also called **diffeomorphisms**.

It is possible to differentiate smooth functions and maps by first differentiating the local coordinates and then mapping the result back to the manifolds.

SmoothMan is the category of smooth manifolds with smooth maps as homomorphisms and diffeomorphisms as isomorphisms.

Smooth manifolds are the most common type of manifolds, and are the type we will be using, but there are other types. For example, we can define C^k-differentiable manifolds by replacing the requirement that the transition functions be smooth (C^∞) with requirements for transition functions to be C^k-differentiable. We can define real analytic manifolds by requiring that transition functions are real analytic functions. We can define complex smooth manifolds by replacing $\mathbb{R}$ with $\mathbb{C}$ and requiring that transition functions be holomorphic. The list of possible manifold types goes on.

C.1.6 Examples of Manifolds

Now that we've had quite an abstract introduction to manifolds it helps to discuss a few examples. The first and most obvious example is Euclidean space itself, which has an atlas consisting or a single chart, $(\mathbb{R}^n, \mathrm{id})$, where id is the identity function.

Our next example is the circle, S^1. This is a one-dimensional manifold. It takes at least two charts to form an atlas. One chart can be given by defining $\vartheta = 0$ to be the top of the circle, and to wrap around to 2π. The definition of a chart requires that the image of an open set is open in $\mathbb{R}$, in this case this means we can't include 0 or 2π in the neighbour hood, so our neighbour hood is $(0, 2\pi)$. We can introduce another chart that defines $\vartheta' = 0$ to be the right most point of the circle and similarly wraps around to 2π, this again maps to the neighbourhood $(0, 2\pi)$ and covers the point excluded in the first chart.

The circle is fairly obviously a manifold since it is usually viewed embedded in Euclidean space. This isn't necessary for all manifolds however. A more abstract manifold is SO(3), the group of rotations in three dimensions. This is continuously parametrised by three Euler angles, and so is a three-dimensional manifold. The

Lorentz group, $\mathrm{O}(1,3)$, is a three dimensional manifold parametrised by the three components of the velocity of the frame.

For a system of N particles their phase space, consisting of their positions and momenta, is a $6N$-dimensional manifold.

Given an equation with two variables, x and y, we can define the set of all (x, y) to be a manifold and any particular solution to the equation is a curve in this manifold.

Vector spaces are manifolds, in particular real finite-dimensional vector spaces. To see this we simply construct a basis and then parametrise the vectors by the coefficients of the basis vectors.

C.2 Tangent Spaces

One of the main changes when moving from discussing vectors in Euclidean space to vectors in some manifold is that it becomes important where the vector is on the manifold. In Euclidean geometry we usually think of vectors as being at the origin, but we can also move them around freely, say to add them by combining them tip to tail. This doesn't work on a general manifold as we will see later.

Let M be a smooth manifold. For each point, $p \in M$, we define a **tangent space** at p, T_pM, and we can treat the vectors in this tangent space as Euclidean vectors. A tangent vector is a vector in the tangent space. The formal definition of a tangent space is a bit abstract:

> **Definition C.2.1 — Tangent Space** Let M be a smooth manifold, and $f\colon M \to \mathbb{R}$ a smooth function on M. The set of all smooth functions on M, denoted $C^\infty(M)$, is a real associative algebra under point wise addition and products, that is for $f, g \in C^\infty(M)$ we define $(f+g)(p) = f(p) + g(p)$ and $fg(p) = f(p)g(p)$.
> A **derivation** at $p \in M$ is a linear map, $D\colon \mathrm{End}(M) \to \mathbb{R}$ satisfying the Leibniz identity,
>
> $$D(fg) = D(f)g(p) + f(p)D(g) \tag{C.2.2}$$
>
> for all $f, g \in C^\infty(M)$. Note the similarity to the product rule.
> Define addition and scalar multiplication of derivations at p by
>
> $$(D_1 + D_2)(f) = D_1(f) + D_2(f), \quad \text{and} \quad (\lambda D)(f) = \lambda D(f) \tag{C.2.3}$$
>
> where D_1, D_2, and D are derivations at p, and $\lambda \in \mathbb{R}$. This makes the space of derivations a real vector space and this is the space we define as the tangent space, T_pM.
> We can then define the **tangent bundle**, TM, as the disjoint union of all tangent spaces,
>
> $$TM := \bigsqcup_{p \in M} T_pM = \{(p, x) \mid p \in M, x \in T_pM\}. \tag{C.2.4}$$
>
> That is we combine all vectors in any tangent space, T_pM, into a single set and tag each one with the point p, to keep track of p, since this is important.

Given this definition we can define a vector field as an assignment of a tangent vector to each point in the manifold. More formally we define a vector field, $V\colon M \to$

TM such that $\pi \circ V$ is the identity mapping on M where $\pi\colon TM \to M$ is defined to be the projection $\pi(p, x) = p$.

D

Algebra

In this appendix we define the notion of *an* algebra and build our way to a definition of a Lie algebra independent from any definition of a Lie group.

Definition D.0.1 — Algebra An **algebra** is a vector space, A, over a field, $\mathbb{F}$, with a bilinear binary operation $\cdot A \times A \longrightarrow A$. To be bilinear means that this operation satisfies the following:

- **Right distributivity**: For all $x, y, z \in A$ we have $(x + y)z = xz + yz$.
- **Left distributivity**: For all $x, y, z \in A$ we have $z(x + y) = zx + zy$.
- **Compatibility with scalar multiplication**: For all $x, y \in A$ and $a, b \in \mathbb{F}$ we ahve $(ax)(by) = (ab)(xy)$.

If $x(yz) = (xy)z$ for all $x, y, z \in A$ then we say that A is an **associative algebra**.

When it is important to express the field we call A a $\mathbb{F}$-algebra.

Example D.0.2

- The complex numbers, $\mathbb{C}$, can be viewed as an associative $\mathbb{R}$-algebra with the vector space $\mathbb{R}^2$ with the product of complex numbers as the bilinear operation.
- Three dimensional Euclidean space, $\mathbb{R}^3$, can be viewed as a non-associative $\mathbb{R}$-algebra with the cross product as the bilinear operation.
- The quaternions, $\mathbb{H}$, can be viewed as an associative $\mathbb{R}$-algebra with the vector space $\mathbb{R}^4$ with quaternion multiplication as the bilinear operation.

Definition D.0.3 — $\mathbb{F}$-Algebra Morphisms Given two $\mathbb{F}$-algebras, A and B, a $\mathbb{F}$-**algebra homomorphism** is a $\mathbb{F}$-linear map, $\varphi\colon A \longrightarrow B$ such that

$$f(ax + by) = af(x) + bf(y), \qquad \text{and} \qquad f(\boldsymbol{x}\boldsymbol{y}) = f(\boldsymbol{x})f(\boldsymbol{y}) \quad \text{(D.0.4)}$$

for all $a, b \in \mathbb{F}$ and $\boldsymbol{x}, \boldsymbol{y} \in A$.
A $\mathbb{F}$-**algebra isomorphism** is a bijective $\mathbb{F}$-algebra homomorphism.

Definition D.0.5 — Lie Algebra A **Lie algebra** is a vector space, $\mathfrak{g}$, over some field, $\mathbb{F}$, equipped with a non-associative binary operation endowing $\mathfrak{g}$ with the structure of an algebra. This operation, $[-,-]\colon \mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$, called the **Lie bracket**, must also satisfy the following

- **Bilinearity**: For all $x, y, z \in \mathfrak{g}$ and $a, b \in \mathbb{F}$

$$[ax+by, z] = a[x,z] + b[y,z], \quad \text{and} \tag{D.0.6}$$
$$[z, ax+by] = a[z,x] + b[z,y]. \tag{D.0.7}$$

- **Alternativity**: For all $x \in \mathfrak{g}$

$$[x,x] = 0. \tag{D.0.8}$$

- The **Jacobi identity**: For all $x, y, z \in \mathfrak{g}$

$$[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0. \tag{D.0.9}$$

It is common, particularly in physics where we typically take $\mathbb{F}$ to be $\mathbb{R}$ or $\mathbb{C}$, to require anticommutativity, $[x,y] = -[y,x]$, instead of alternativity. For fields with characteristics other than 2 this then implies alternativity, however for fields with characteristic 2 this isn't the case so this is a slightly weaker condition, but the distinction isn't important for us.

Notice that the left hand side of the Jacobi identity, while it looks complex, is really a sum over cyclic permutations of $[x,[y,z]]$.

Consider the vector space of square $n\times n$ matrices. A Lie bracket defined on this vector space is the commutator, $[A,B] \coloneqq AB - BA$, this is the prototypical example of a Lie bracket and the reason that the Lie bracket is written in the way it is.

Consider the vector space $\mathbb{R}^3$. A Lie bracket defined on this vector space is the cross product, $[\boldsymbol{v},\boldsymbol{u}] \coloneqq \boldsymbol{v}\times\boldsymbol{u}$.

Let A be a commutative $\mathbb{R}$-algebra, that is $\boldsymbol{x}\boldsymbol{y} = \boldsymbol{y}\boldsymbol{x}$ for all $\boldsymbol{x},\boldsymbol{y} \in A$. We define a derivation, $D\colon A\to A$ to be a $\mathbb{R}$-linear map ($D(a\boldsymbol{x}) = aD(\boldsymbol{x})$ for all $a\in\mathbb{R}$ and $\boldsymbol{x}\in A$) such that D satisfies the Leibniz law:

$$D(\boldsymbol{x}\boldsymbol{y}) = \boldsymbol{x}D(\boldsymbol{y}) + D(\boldsymbol{x})\boldsymbol{y}. \tag{D.0.10}$$

The space of derivations, $\mathrm{Der}(A)$, is a Lie algebra with the Lie bracket given by $[D_1, D_2](\boldsymbol{x}) = D_1(D_2(\boldsymbol{x})) - D_2(D_1(\boldsymbol{x}))$.

Definition D.0.11 — Lie Algebra Morphisms Given two Lie algebras, $\mathfrak{g}$ and $\mathfrak{h}$, over the same base field, $\mathbb{F}$, a **Lie algebra homomorphism** is an $\mathbb{F}$-algebra homomorphism, $\varphi\colon\mathfrak{g}\to\mathfrak{h}$, which preserves the Lie bracket, so

$$\varphi([x,y]) = [\varphi(x), \varphi(y)]. \tag{D.0.12}$$

If this is bijective then it is an **isomorphism of Lie algebras**.

D.1 Connection to Lie Groups

Recall that we can think of a vector field on a smooth manifold as derivations. Any group, G, which acts smoothly on a manifold acts on the vector fields. Further, the vector space of vector fields fixed by the group action is closed under the Lie bracket of derivations and so forms a Lie algebra. Consider the case where G is a Lie group, so is itself a manifold, and acts smoothly on itself by left translation, $g\,.\,h = L_g(h) = gh$. Then the space of left invariant vector fields, that is vector fields satisfying $L_{g^*}X_h = X_{gh}$ for all $h \in G$ with L_{g^*} being the differential of L_g, is a Lie algebra under te Lie bracket of vector fields. We can extend any tangent vector at the identity, that is an element of T_eG, to a left invariant vector field by left translating the tangent vector to other points on the manifold. In particular the left invariant extension of some $v \in T_eG$ is defined by $v^\wedge{}_g := L_{g^*}v$. This identifies T_eG with the space of left invariant vector fields, which makes T_eG a Lie algebra.

We usually denote the Lie algebra T_eG by the lowercase fraktur letter $\mathfrak{g}$. The Lie bracket on $\mathfrak{g}$ is defined to be $[v, w] := [v^\wedge, w^\wedge]_e$, where $[v^\wedge, w^\wedge]_e$ is the Lie bracket of left invariant vector fields at the identity.

D.2 Universal Enveloping Algebra

Definition D.2.1 — Tensor Algebra Let V be a vector space over the field $\mathbb{F}$. For non-negative integers k we define the kth tensor power of V to be the tensor product of V with itself k times:

$$T^kV = V^{\otimes k} := \underbrace{V \otimes V \otimes \cdots \otimes V}_{k \text{ times}}. \tag{D.2.2}$$

That is T^kV consists of all tensors on V of order k. Which is to say we can think of T^kV as the space of all linear combinations of elements of the form $a_1 \otimes \cdots \otimes a_k$ with $\otimes$ being linear and distributing over vector addition. We define $T^0V = \mathbb{F}$ by convention, considering $\mathbb{F}$ as a one-dimensional vector space over itself.
The **tensor algebra**, $T(V)$, is then defined as the direct sum

$$T(V) = \bigoplus_{k=1}^{\infty} T^kV = \mathbb{F} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots. \tag{D.2.3}$$

Multiplication in $T(V)$ is given by the obvious isomorphism $T^kV \otimes T^lV \to T^{k+l}V$, given by

$$(a_1 \otimes \cdots \otimes a_k)(b_1 \otimes \cdots \otimes b_l) = a_1 \otimes \cdots \otimes a_k \otimes b_1 \otimes \cdots \otimes b_l \tag{D.2.4}$$

Definition D.2.5 — Universal Enveloping Algebra Given a Lie algebra $\mathfrak{g}$ we can construct the tensor algebra, $T(\mathfrak{g})$. We can then extend the Lie bracket to this tensor algebra one step at a time. We first define the bracket on $\mathfrak{g} \otimes \mathfrak{g}$ to be

$$a \otimes b - b \otimes a = [a, b]. \tag{D.2.6}$$

We then define the bracket on $\mathfrak{g} \otimes \cdots \otimes \mathfrak{g}$ recursively by defining

$$[a \otimes b, c] = a \otimes [b, c] + [a, c] \otimes b \tag{D.2.7}$$

and

$$[a, b \otimes c] = [a, b] \otimes c + b \otimes [a, c]. \tag{D.2.8}$$

The resulting definition is still a Lie bracket, but now defined on an associative algebra, $T(\mathfrak{g})$.
We define the **universal enveloping algebra**, $U(\mathfrak{g})$, to be the quotient space

$$U(\mathfrak{g}) \coloneqq T(\mathfrak{g})/\sim \tag{D.2.9}$$

where $a \sim b$ if

$$a \otimes b - b \otimes a = [a, b]. \tag{D.2.10}$$

Index