

Willoughby Seago

**Theoretical Physics**

# **Symmetries of Quantum Mechanics**

February 17, 2022

**COURSE NOTES**

# Symmetries of Quantum Mechanics

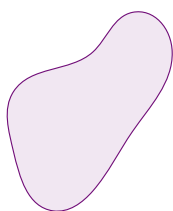
Willoughby Seago

February 17, 2022

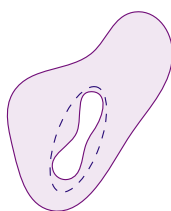
## Abstract

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.

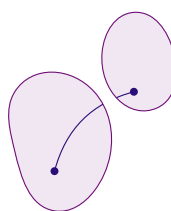
These notes were last updated at 17:43 on May 11, 2022. For notes on other topics see <https://github.com/WilloughbySeago/Uni-Notes>.



Simply Connected



Connected



Disconnected

---

# Chapters

---

	Page
<b>Chapters</b>	<b>ii</b>
<b>Contents</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>I Group Theory</b>	<b>1</b>
1 Introduction	2
2 Morphisms and Cosets	16
3 Group Action	24
4 Normal Subgroups	30
5 Products of Groups	37
6 Permutation Groups	42
7 Applications	49
<b>II Representation Theory</b>	<b>52</b>
8 Basics of Representation Theory	53
9 Character Theory of Finite Groups	60
<b>III Continuous Groups</b>	<b>92</b>
10 Lie Groups	93
11 Lie Algebras	101

<i>CHAPTERS</i>	iii
<b>IV Applications to Quantum Mechanics</b>	<b>122</b>
12 Quantum Mechanics Synopsis	123
13 Physics of Angular Momentum	128
14 Further Applications	135
15 Selection Rules	143
<b>Appendices</b>	<b>149</b>
A Mathematical Preliminaries	150
B Groups	162
C Manifolds	168
D Algebra	176
<b>Index</b>	<b>180</b>

---

# Contents

---

	Page
<b>Chapters</b>	<b>ii</b>
<b>Contents</b>	<b>iv</b>
<b>List of Figures</b>	<b>viii</b>
<b>I Group Theory</b>	<b>1</b>
<b>1 Introduction</b>	<b>2</b>
1.1 Binary Operations . . . . .	2
1.1.1 Associativity . . . . .	2
1.1.2 Identity . . . . .	4
1.1.3 Inverse . . . . .	5
1.2 Groups . . . . .	5
1.2.1 Examples of Groups . . . . .	6
1.2.2 Basic Definitions . . . . .	7
1.2.3 Cayley Tables . . . . .	10
1.3 Almost Groups . . . . .	13
<b>2 Morphisms and Cosets</b>	<b>16</b>
2.1 Morphisms . . . . .	16
2.2 Group Presentations . . . . .	20
2.3 Cosets . . . . .	21
<b>3 Group Action</b>	<b>24</b>
3.1 Group Action . . . . .	24
3.2 Orbits and Stabilisers . . . . .	26
<b>4 Normal Subgroups</b>	<b>30</b>
4.1 Normal Subgroups . . . . .	30
4.2 Coset Groups . . . . .	32
<b>5 Products of Groups</b>	<b>37</b>
5.1 Direct Products . . . . .	37

5.2	Semidirect Product . . . . .	38
<b>6</b>	<b>Permutation Groups</b>	<b>42</b>
6.1	Symmetric Group . . . . .	42
6.2	Alternating Group . . . . .	46
<b>7</b>	<b>Applications</b>	<b>49</b>
7.1	Platonic Solids . . . . .	49
7.2	Molecules . . . . .	50
<b>II</b>	<b>Representation Theory</b>	<b>52</b>
<b>8</b>	<b>Basics of Representation Theory</b>	<b>53</b>
8.1	Representation Definition . . . . .	53
8.2	Pedestrian Approach . . . . .	54
8.3	Basic Definitions . . . . .	55
8.4	Some Theorems . . . . .	57
<b>9</b>	<b>Character Theory of Finite Groups</b>	<b>60</b>
9.1	Basics of Character Theory . . . . .	60
9.2	Space of Characters . . . . .	62
9.2.1	Dimensionality Theorem . . . . .	66
9.3	Character Tables . . . . .	67
9.4	Constructing Character Tables . . . . .	72
9.4.1	Character Table of $S_3$ . . . . .	74
9.4.2	Character Table of $S_4$ . . . . .	76
9.4.3	Alternating Groups . . . . .	78
9.4.4	Cyclic Groups . . . . .	78
9.5	Complex, Real, and Pseudo-Real Representations . . . . .	79
9.6	Branching Rules . . . . .	81
9.7	Constructing Representations . . . . .	83
9.7.1	Irreducible Representations of $S_n$ . . . . .	83
9.7.2	Direct Product Representations . . . . .	85
9.7.3	Induced Representations . . . . .	88
9.8	Applications . . . . .	88
9.8.1	Distortion of Lattices . . . . .	88
9.8.2	Fermi–Bose Statistics . . . . .	89
<b>III</b>	<b>Continuous Groups</b>	<b>92</b>
<b>10</b>	<b>Lie Groups</b>	<b>93</b>
10.1	Background . . . . .	93
10.2	Basics . . . . .	93
10.3	Properties of Lie Groups . . . . .	95
10.3.1	Basic Properties . . . . .	95
10.3.2	Compactness . . . . .	96
10.3.3	Connectedness . . . . .	97
10.3.4	Simplicity . . . . .	97

<b>11 Lie Algebras</b>	<b>101</b>
11.1 The Abelian Group $U(1) \cong SO(2)$	101
11.2 Lie Algebra Generalities	105
11.3 The Non-Abelian Groups $SU(2)$ and $SO(3)$	111
11.3.1 The Lie Algebras of $SU(2)$ and $SO(3)$	111
11.3.2 Irreducible Representations of $\mathfrak{su}(2)$	113
11.3.3 Clebsch–Gordan Series of $SU(2)$	117
 <b>IV Applications to Quantum Mechanics</b>	 <b>122</b>
<b>12 Quantum Mechanics Synopsis</b>	<b>123</b>
12.1 Rules of Quantum Mechanics	123
12.2 From hamiltonian to the Schrödinger Equation	124
12.3 Symmetries and Conservation Laws	125
<b>13 Physics of Angular Momentum</b>	<b>128</b>
13.1 Basics of Angular Momentum in Quantum Mechanics	128
13.2 Wigner–Eckart Theorem	130
13.2.1 Using Clebsch–Gordan Coefficients	132
13.2.2 Geometric Interpretation of the $j = 1$ Representation	133
<b>14 Further Applications</b>	<b>135</b>
14.1 Polynomial Invariants	135
14.2 Landé $g$ -Factor	136
14.3 Spherical Harmonics	138
14.3.1 Application to $\mathbb{R}^3 \cong \mathbb{R}^+ \otimes S^2$	141
14.3.2 Why $S^2$ ?	142
<b>15 Selection Rules</b>	<b>143</b>
15.1 Parity Selection	143
15.2 Superselection	144
15.3 Electric Dipole	145
15.4 Pauli’s Hydrogen Atom	146
 <b>Appendices</b>	 <b>149</b>
<b>A Mathematical Preliminaries</b>	<b>150</b>
A.1 Basic Mathematics	150
A.1.1 Notation	150
A.1.2 Definitions	151
A.2 Linear Algebra	154
A.2.1 Vectors	154
A.2.2 Matrices	158
A.2.3 Combining Vector Spaces	160
<b>B Groups</b>	<b>162</b>
B.1 Finite Groups	162
B.1.1 Other Finite Groups	163
B.2 Discrete Groups	164
B.3 Continuous Groups	164

B.3.1	Scalars . . . . .	164
B.3.2	Matrices . . . . .	165
<b>C</b>	<b>Manifolds</b>	<b>168</b>
C.1	Manifolds . . . . .	168
C.1.1	Topology . . . . .	169
C.1.2	Charts . . . . .	170
C.1.3	Topological Manifolds . . . . .	170
C.1.4	Local Coordinates . . . . .	171
C.1.5	Smooth Manifold . . . . .	172
C.1.6	Examples of Manifolds . . . . .	173
C.2	Tangent Spaces . . . . .	174
<b>D</b>	<b>Algebra</b>	<b>176</b>
D.1	Connection to Lie Groups . . . . .	178
D.2	Universal Enveloping Algebra . . . . .	178
	<b>Index</b>	<b>180</b>



---

## List of Figures

---

	Page
5.1 Ethane molecule. . . . .	41
6.1 The $k$ -cycle $(a_1 \dots a_k)$ . . . . .	43
6.2 Permutation by transposition. . . . .	45
7.1 The octahedron is the dual of the cube. . . . .	50
7.2 Two possible symmetries of the tetrahedron . . . . .	50
7.3 Water, with symmetry group $\mathbb{Z}_2$ , has a permanent dipole. . . . .	50
7.4 Methane is tetragonal. . . . .	51
9.1 Hook length. . . . .	85
9.2 A cube distorted along the $z$ -direction. . . . .	89
9.3 A cube distorted along the $(1, 1, 1)$ -direction . . . . .	89
9.4 A cube distorted along the $(1, 1, 1)$ -direction viewed point on. . . . .	90
10.1 Connectedness . . . . .	98
11.1 The manifold $S^1$ . . . . .	105
13.1 Infinitesimal rotation. . . . .	134
A.1 Injective, surjective, and bijective functions. . . . .	151

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,  $f: X \times X \rightarrow X$ .



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  $\cdot: X \times X \rightarrow X$  maps  $(x, x') \mapsto x \cdot x'$  whereas for a normal function, say,  $f: X \times X \rightarrow X$ , 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  $\cdot: X \times X \rightarrow X$  is **associative** if

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z. \quad (1.1.4)$$

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

$$(x_1 x_2)(x_3 x_4) = x_1(x_2(x_3 x_4)) = x_1 x_2 x_3 x_4. \quad (1.1.5)$$

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  $\text{Hom}(A, B)$  the set of functions from  $A$  to  $B$ . We define **function composition** to be the binary operation  $\circ: \text{Hom}(B, C) \times \text{Hom}(A, B) \rightarrow \text{Hom}(A, C)$  for sets  $A$ ,  $B$ , and  $C$ , such that for  $f \in \text{Hom}(B, C)$  and  $g \in \text{Hom}(A, B)$  we have

$$(f \circ g)(a) = f(g(a)) \quad (1.1.7)$$

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:

$$\begin{array}{ccc} A & \xrightarrow{g} & B \\ & \searrow f \circ g & \downarrow f \\ & & C. \end{array} \quad (1.1.8)$$

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

$$f \circ (g \circ h) = (f \circ g) \circ h \quad (1.1.9)$$

and so  $\circ$  is associative.

The commutative diagram expressing this fact is

$$\begin{array}{ccccc} & & g \circ f & & h \circ g \\ & \nearrow & & \searrow & \\ A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D. \\ & \searrow & & \nearrow & \\ & & h \circ g \circ f & & \end{array} \quad (1.1.10)$$

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,  $\times: \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$  (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  $\mathbf{a} = (1, 1, 0)^T$ ,

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

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

whereas

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

### 1.1.2 Identity

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

$$x \cdot e = e \cdot x = x \quad (1.1.15)$$

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  $\text{id}$ , or  $\iota$ ,
- 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  $\delta_{ij}$  as elements.
- The identity function is  $\text{id}_X : X \times X \rightarrow X$  defined by  $\text{id}_X(x) = x$  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  $X = \mathbb{Z}_{>0}$ <sup>1</sup> and take our operation to be addition then there is no identity since  $0 \notin \mathbb{Z}_{>0}$ .

$$^1\mathbb{Z}_{>0} := \{1, \dots\} = \mathbb{N} \setminus \{0\}$$

### 1.1.3 Inverse

**Definition 1.1.18 — Inverse** Given a set,  $X$ , and a binary operation on  $X$ ,  $\cdot : X \times X \rightarrow X$ , 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

$$x \cdot x^{-1} = x^{-1} \cdot x = e. \quad (1.1.19)$$

**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  $X = \mathbb{N}$ <sup>2</sup> 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  $-3 \notin \mathbb{N}$ .

$$^2\mathbb{N} := \{0, 1, \dots\} = \mathbb{Z}_{>0} \cup \{0\}.$$

## 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  $g_1, g_2, g_3 \in G$

$$g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3. \quad (1.2.2)$$

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

$$g \cdot g^{-1} = g^{-1} \cdot g = e \quad (1.2.3)$$

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  $X \times X \rightarrow X$ , 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.  $\square$

**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.  $\square$

**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  $(gh)(h^{-1}g^{-1}) = g(hh^{-1})g^{-1} = geg^{-1} = gg^{-1} = e$ .  $\square$

**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  $e^{-1} = e$ .  $\square$

## 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

$$\sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\}. \quad (1.2.12)$$

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  $m$ th 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  $O(3) := \{O \in \mathcal{M}_3(\mathbb{R}) \mid O^T O = \mathbb{1}\}$ . 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  $O \in O(3)$  we have  $\det O = \pm 1 \neq 0$ . This is called the **rotation group**.



Strictly this is the fundamental representation of  $O(3)$ .

## 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  $\mathbb{Z}$ . That is, if  $G$  is a finite set, countably infinite set, or uncountable set.



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  $O(3)$  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  $e^{2i\pi/3} = e^{2i\pi/3} \in \mathbb{Z}_9$  is 3 since  $(e^{2i\pi/3})^1 = e^{2i\pi/3}$ ,  $(e^{2i\pi/3})^2 = e^{4i\pi/3}$ , and  $(e^{2i\pi/3})^3 = e^{6i\pi/3} = e^{2i\pi} = 1$ , 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  $O(3)$  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  $\mathbb{Z}_3 = \{1, e^{2i\pi/3}, e^{4i\pi/3}\} \subset \mathbb{Z}_9$ .

**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  $g_1 = gg_2g^{-1}$ .

**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  $a = bgb^{-1}$  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  $a = bgb^{-1}$  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.  $\square$

**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

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

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

$$(zz')g = z(z'g) = z(gz') = (zg)z' = (gz)z' = g(zz') \quad (1.2.30)$$

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  $z^{-1}gzz^{-1} = z^{-1}g = z^{-1}zg z^{-1} = gz^{-1}$ , 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  $G$ .  $\square$

■ **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  $O(3)$  is  $Z(O(3)) = \{1, -1\}$ .

**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  $g g^{-1} = e \in H$ , 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  $G$ .  $\square$

### 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  $i$ th row and  $j$ th column as the product of the  $i$ th and  $j$ th element of the group in the chosen order. That is

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

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

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

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:

$$\begin{array}{c|ccc} \mathbb{Z}_3 & 1 & e^{2i\pi/3} & e^{2i\pi 2/3} \\ \hline 1 & 1 & e^{2i\pi/3} & e^{2i\pi 2/3} \\ e^{2i\pi/3} & e^{2i\pi/3} & e^{2i\pi 2/3} & 1 \\ e^{2i\pi 2/3} & e^{2i\pi 2/3} & 1 & e^{2i\pi/3} \end{array} \quad \begin{array}{c|ccc} \mathbb{Z}_3 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \quad (1.2.38)$$

We will see in a bit that what we really are saying here is that the multiplicative group  $\{1, e^{2i\pi/3}, e^{2i\pi 2/3}\}$  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 think 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  $(12)$  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

$$(12)(a, b, c) = (b, a, c). \quad (1.2.40)$$

Applying this 2-cycle a second time we get

$$(12)^2(a, b, c) = (12)(12)(a, b, c) = (12)(b, a, c) = (a, b, c), \quad (1.2.41)$$

and so we see that  $(12)^2 = ()$ .

Now consider the 3-cycle  $(123)$  acting on  $(a, b, c)$ . We see that

$$(123)(a, b, c) = (c, a, b), \quad (1.2.42)$$

$$(123)^2 = (123)(c, a, b) = (b, c, a), \quad (1.2.43)$$

$$(123)^3 = (123)(b, c, a) = (a, b, c), \quad (1.2.44)$$

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:

	$S_3$	$()$	$(1\ 2)$	$(2\ 3)$	$\dots$
$\mathbb{Z}_2$	$()$	$()$	$(1\ 2)$	$(2\ 3)$	$\dots$
	$(1\ 2)$	$(1\ 2)$	$()$	$(1\ 2\ 3)$	$\dots$
	$(2\ 3)$	$(1\ 3)$	$(3\ 2\ 1)$	$()$	$\dots$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

(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.  $\square$

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  $|G|$ . Then for each  $g \in G$  we can define  $\sigma_g: G \rightarrow G$  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  $\sigma_g$  is bijective and hence is a permutation on the set  $G$ .

Now define  $\varphi: G \rightarrow S_{|G|}$  by  $\varphi(g) = \sigma_g$ . Then  $\varphi$  is a homomorphism since

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

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  $\varphi$  is injective.

The function  $\tilde{\varphi}: G \rightarrow \text{Im}(\varphi)$  given by  $\tilde{\varphi}(g) = \varphi(g)$  is a surjective. It remains only to show that  $\text{Im } \varphi$  is a subgroup of  $S_{|G|}$ . This will be proven in [Lemma 2.1.21](#), and so we have proven the theorem.  $\square$

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,  $\cdot: X \times X \rightarrow X$ .

#### ■ 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

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

- $(\mathbb{R}, f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}; f(x, y) = x^3y + 2x)$  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  $e^2 = e$ .
- $(\{\text{True}, \text{False}\}, \text{XOR}, \text{False})$  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.
- $(\text{Hom}_{\text{Set}}(X, X), \circ, \text{id}_X)$  is a monoid where  $\text{Hom}_{\text{Set}}(X, X)$  is the set of all functions  $X \rightarrow X$  for some set  $X$ , more generally  $\text{Hom}_{\mathcal{C}}(A, B)$  is the class of all morphisms from  $A$  to  $B$  in some category  $\mathcal{C}$ . Also,  $\circ$  is composition of functions and  $\text{id}_X$  is the identity function, defined by  $\text{id}_X(x) = x$  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.

- $(\text{End}_{\mathbf{Grp}}(G), \circ, \text{id}_G)$  is a monoid where  $G$  is a group,  $\text{End}_{\mathbf{Grp}}(G)$  is the set of all **endomorphisms** of  $G$ , which is to say homomorphisms (to be defined in the next section) of  $G$  with itself,  $\text{End}_{\mathbf{Grp}}(G) := \text{Hom}_{\mathbf{Grp}}(G, G)$ .
- More generally  $(\text{End}_{\mathbf{C}}(X), \circ, \text{id}_X)$  is a monoid where  $\mathbf{C}$  is a category,  $X$  is an object of  $\mathbf{C}$ , and  $\text{End}_{\mathbf{C}}(X)$  is the set of morphisms from  $X$  to  $X$  in  $\mathbf{C}$ , that is  $\text{End}_{\mathbf{C}}(X) = \text{Hom}_{\mathbf{C}}(X, X)$ . The previous example is subsumed by this one taking  $\mathbf{C} = \mathbf{Grp}$ , the category of groups with group homomorphisms as morphisms.
- $(\text{Hom}_{\mathbf{C}}(\bullet, \bullet), \circ, \text{id}_{\bullet})$  is a monoid where  $\mathbf{C}$  is a category with a single object,  $\bullet \in \text{Obj}(\mathbf{C})$ . The elements of this monoid are the morphisms from this object to itself with morphism composition as a binary operation and the identity morphism,  $\text{id}_{\bullet}$ , 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  ${}^{-1}: X \rightarrow X$  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  $\varphi: G \rightarrow H$  which preserves the group product. That is for all  $g, g' \in G$  we have

$$\varphi(gg') = \varphi(g)\varphi(g'). \quad (2.1.2)$$

**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  $\varphi: G \rightarrow \{e\}$  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,  $\text{id}_G: G \rightarrow G$ , defined by  $\text{id}_G(g) = g$  for all  $g \in G$  is an isomorphism. That is  $\text{id}_G(gg') = gg' = \text{id}_G(g)\text{id}_G(g')$ , and  $\text{id}_G$  is a self inverse, so  $\text{id}_G$  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,  $(12)$ . Then  $() \mapsto 1$  and  $(12) \mapsto -1$  is an isomorphism. To see this note that  $(12)(12) = ()$ , that is swapping and swapping back has no net effect, and so

$$\varphi(1 \cdot 1) = \varphi(1) = () = ()() = \varphi(1)\varphi(1), \quad (2.1.5)$$

$$\varphi((-1) \cdot (-1)) = \varphi(1) = () = (12)(12) = \varphi(-1)\varphi(-1), \quad (2.1.6)$$

$$\varphi(1 \cdot (-1)) = \varphi(-1) = (12) = ()(12) = \varphi(1)\varphi(-1). \quad (2.1.7)$$

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,  $\varphi: \mathbb{Z} \rightarrow 2\mathbb{Z}$ , 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

$$\varphi(n + m) = 2(n + m) = 2n + 2m = \varphi(n) + \varphi(m), \quad (2.1.9)$$

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  $\varphi$  is injective.

Finally, we check that this is surjective. Consider some  $n \in 2\mathbb{Z}$ , since this is even<sup>a</sup>  $2|n$  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  $\varphi$  is surjective.

Note that we could also have identified that  $\varphi^{-1}(n) = n/2$  is the inverse of  $\varphi$ . Either way  $\varphi$  is a bijective homomorphism and hence an isomorphism.

<sup>a</sup>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  $x \mapsto e^x$ , which is a homomorphism since

$$e^{x+y} = e^x e^y \quad (2.1.11)$$

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  $2 \times 2$  real matrices

$$\left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}. \quad (2.1.13)$$

An isomorphism between these two groups is given by

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

We first check that this is a homomorphism:

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

and

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

so this is indeed a homomorphism.

We can see that this is bijective by noticing that the inverse is simply

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

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  $1 \times 1$  unitary matrices by  $U(1)$  and the group of two-dimensional rotations by  $SO(2)$  what we see here is that  $U(1) \cong \mathbb{T} \cong SO(2)$ , where the isomorphism between  $\mathbb{T}$  and  $U(1)$  is the obvious one mapping  $z \in \mathbb{T}$  to  $(z) \in U(1)$  and the isomorphism between  $\mathbb{T}$  and  $SO(2)$  is the restriction of  $\varphi$  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**,  $\mathbb{Z}_2 \times \mathbb{Z}_2$ <sup>1</sup>, which has the unique property for a group of this order that all non-trivial (i.e. not

<sup>1</sup>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

$$\begin{array}{c|cccc} \mathbb{Z}_4 & 1 & -1 & i & -i \\ \hline 1 & 1 & -1 & i & -i \\ -1 & -1 & 1 & -i & i \\ i & i & -i & -1 & 1 \\ -i & -i & i & 1 & -1 \end{array} \quad \begin{array}{c|cccc} \mathbb{Z}_2 \times \mathbb{Z}_2 & e & a & b & c \\ \hline e & e & a & b & c \\ a & a & e & c & b \\ b & b & c & e & a \\ c & c & b & a & e \end{array} \quad (2.1.18)$$

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  $\varphi: G \rightarrow H$  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  $\varphi(e_Gg) = \varphi(e_G)\varphi(g)$ , and  $\varphi(e_Gg) = \varphi(g) = e_H\varphi(g)$ . Right multiplying these two results by  $\varphi(g)^{-1}$  we have  $\varphi(e_G) = e_H$ .  $\square$

**Lemma 2.1.20 — Homomorphisms Map Inverses to Inverses** Let  $G$  and  $H$  be groups and  $\varphi: G \rightarrow H$  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.  $\square$

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

*Proof.* Consider  $h, h' \in \text{Im } \varphi$ . 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  $hh'^{-1} \in \text{Im } \varphi$  and hence  $\text{Im } \varphi$  is a subgroup of  $H$  by the subgroup criterion of [Theorem 1.2.32](#).  $\square$

**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  $g^i = g^j$  if  $i \equiv j \pmod n$ .

The map  $\varphi: G \rightarrow \mathbb{Z}_n$  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  $\varphi$  occurs modulo  $n$ .

The inverse of this map is simply  $\varphi^{-1}(i) = g^i$  and so this is a bijection. Meaning that  $\varphi$  is an isomorphism.  $\square$

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

*Proof.* Let  $G$  and  $H$  be groups and  $\varphi: G \rightarrow H$  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  $\varphi$  is injective.

Suppose instead that  $\varphi$  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  $\varphi$  means  $e_G = g$ , and so  $\ker \varphi = \{e_G\}$ .  $\square$

## 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

$$G = \langle S \mid C \rangle \quad (2.2.2)$$

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

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

We can identify  $a = e^{2i\pi/n}$  as one possible generator with the group operation of multiplication, but it need not be the only one. For example if  $n = 3$  then  $a = e^{2i\pi/3}$  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  $S_3$ :

$$S_3 = \langle a, b, c \mid a^2 = b^2 = c^3 = abc = e \rangle. \quad (2.2.6)$$

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  $S_3$ :

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

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

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

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  $\sigma_i$  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

$$gH := \{gh \mid h \in H\} \quad (2.3.2)$$

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

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  $g_1H \rightarrow g_2H$  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  $|gH| = |H|$ .  $\square$

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

$$[G : H] := \frac{|G|}{|H|}. \quad (2.3.7)$$

**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  $|G|/|H| = n$ .  $\square$

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,  $\varphi: G \times X \rightarrow X$ , where we use the notation  $\varphi(g, x) = g \cdot x$ . This map must be compatible with the group structure, by which we mean

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



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



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,  $\varphi: X \times G \rightarrow X$ , where we use the notation  $\varphi(x, g) = x \cdot g$ . This map must be compatible with the group structure, by which we mean

- $x \cdot e = x$  ( $\varphi(x, e) = x$ ) for all  $x \in X$ , and
- $x \cdot (gg') = (x \cdot g) \cdot g'$  ( $\varphi(x, gg') = \varphi(\varphi(x, g), g')$ ) 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  $\varphi: G \rightarrow \text{Aut}(X)$ , where  $\text{Aut}(X)$  is the group of automorphisms on  $X$ , which is to say exactly the set of bijections  $X \rightarrow X$ , 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

$$g \cdot x = x \quad (3.1.3)$$

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

■ **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  $m \cdot r = (-1)^m r$  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,  $O(1, 3)$ , 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**

$$U(n) := \{U \in M_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1}\}. \quad (3.1.8)$$

In particular, we often deal with

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

which is the group of complex numbers with unit modulus<sup>a</sup>. This particular group is the gauge group of (quantum) electrodynamics, its action on the gauge potential is

$$e^{i\varphi(x)} \cdot A_\mu = A_\mu + \partial_\mu \varphi(x). \quad (3.1.10)$$

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



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.

<sup>a</sup>Strictly the group of complex numbers with unit modulus is the circle group,  $\mathbb{T}$ , and  $U(1)$  is the group of  $1 \times 1$  unitary matrices over the complex numbers, however, under the obvious correspondence that  $(z) \in U(1)$  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

$$g(x) := \{g \cdot x \mid g \in G\} \subseteq X. \quad (3.2.2)$$

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

■ **Example 3.2.4** Consider the group of rotations in the plane about some point  $a$ . The orbit of some other point  $b$  is the circle of radius  $|a - b|$  around the point  $a$ .

Consider the group of integers,  $\mathbb{Z}$ , acting on  $\mathbb{R}$  by  $n \cdot r = r + n$  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  $r \cdot v = rv$  for  $r \in \mathbb{R}$  and  $v \in V$ . Then the orbit of  $v$  is all vectors parallel to  $v$ .

**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

$$G_x := \{g \in G \mid g \cdot x = x\}. \quad (3.2.6)$$

**Notation 3.2.7** Sometimes the stabiliser is denoted  $\text{Stab}_G(x)$ , or  $\text{Stab}(x)$  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  $\text{Stab}_{S_4}((a, b, c, c)) = \{(), (3\ 4)\}$ .

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

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

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

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

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

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

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  $g, h \in \text{Stab}(x)$ , that is  $g \cdot x = h \cdot x = x$ . Then  $(gh^{-1}) \cdot x = g \cdot (h^{-1} \cdot x)$ . Since  $h^{-1}h = e$  we have that  $(h^{-1}h) \cdot x = e \cdot x = x$ , we also have that  $(h^{-1}h) \cdot x = h^{-1} \cdot (h \cdot x) = h^{-1} \cdot x$ . Hence  $h^{-1} \cdot x = x$ . Using this we have  $(gh^{-1}) \cdot x = g \cdot x = x$ , and so  $gh^{-1} \in \text{Stab}(x)$ . Thus, by the subgroup criterion ([Theorem 1.2.32](#))  $\text{Stab}(x)$  is a subgroup of  $G$ .  $\square$

**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  $\phi g(x) \rightarrow G/\text{Stab}(x)$ , where  $G/\text{Stab}(x)$  is the set of all cosets of  $\text{Stab}(x)$ , defined by  $g \cdot x \mapsto g \text{Stab}(x)$  defines a bijection.

*Proof.* Since  $\text{Stab}(x)$  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  $g_1 \text{Stab}(x) =$

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

$$g_1 \cdot x = (g_2 g) \cdot x = g_2 \cdot g \cdot x = g_2 \cdot x \quad (3.2.16)$$

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

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

$$|G| = |\text{Stab}(x)| |\text{Orb}(x)|. \quad (3.2.18)$$

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

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

If  $\text{Stab}(x) = \{e\}$  then  $|\text{Orb}(x)| = |G|$ , since all elements of  $G \setminus \{e\}$  must change  $x$ . Then  $|G| = |\text{Stab}(x)| |\text{Orb}(x)|$  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 choose  $x_p$  such that  $x_p = (x_1 \cdots x_{p-1})^{-1}$ . We then have

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

Hence, we make a choice from a set of size  $|G|^{p-1}$   $p-1$  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  $x = (x_1, \dots, x_p) \in X$  define

$$m \cdot x = (x_{1+m}, \dots, x_{p+m}) = (x_{m+1}, \dots, x_p, x_1, \dots, x_m) \quad (3.2.21)$$

where addition in the indices is done modulo  $p$  so the indices remain in  $\{1, \dots, p\}$ .

Clearly  $|\mathbb{Z}_p| = p$ . Then by [Corollary 3.2.17](#) it follows that

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

Since  $p$  is prime either  $|\text{Stab}_{\mathbb{Z}_p}(x)| = p$  and  $|\text{Orb}_{\mathbb{Z}_p}(x)| = 1$  or  $|\text{Stab}_{\mathbb{Z}_p}(x)| = 1$  and  $|\text{Orb}_{\mathbb{Z}_p}(x)| = p$ .

If  $|\text{Orb}_{\mathbb{Z}_p}(x)| = 1$  then  $x \in X$  must be of the form  $(g, \dots, g)$  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  $x = (e, \dots, e)$ . This cannot be the only example since then  $|X|$  is not divisible by  $p$ , since we can write  $|X|$  as  $|X| = np + m$ , 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  $(g, \dots, g) \in X$ . By the definition of  $X$  this means that  $g \cdots g = g^p = e$ , and so  $g$  is some element of order  $p$ .  $\square$

■ **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  $\mathbb{Z}_p$ .  $\square$

# 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,  $G \times N \rightarrow N$ , 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,  $\{e_i\}$ , and that  $\{e'_i\}$  is some other basis related to the first by  $e'_i = T_{ij}e_j$ . 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

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

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

$$eee^{-1} = ee(-e) = -e^3 = -e, \quad (4.1.6)$$

$$(-e)e(-e)^{-1} = (-e)ee = -e^3 = -e, \quad (4.1.7)$$

$$e(-e)e^{-1} = e(-e)(-e) = e^3 = e, \quad (4.1.8)$$

$$(-e)(-e)(-e)^{-1} = (-e)(-e)e = e^3 = e, \quad (4.1.9)$$

$$iei^{-1} = ie(-i) = -i^2 = e, \quad (4.1.10)$$

$$i(-e)i^{-1} = i(-e)(-i) = i^2 = -e, \quad (4.1.11)$$

$$(-i)e(-i)^{-1} = (-i)ei = -i^2 = e, \quad (4.1.12)$$

$$(-i)(-e)(-i)^{-1} = (-i)(-e)i = i^2 = -e. \quad (4.1.13)$$

$$(4.1.14)$$

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

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**:

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

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**:

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

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



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

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  $G$ .  $\square$

## 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

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

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

$$(gN)(g'N) := (gg')N. \quad (4.2.3)$$

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

$$g'\tilde{g}' = (gn)(\tilde{g}\tilde{n}) \quad (4.2.5)$$

$$= g(n\tilde{g})\tilde{n} \quad (4.2.6)$$

$$= g(\tilde{g}n')\tilde{n} \quad (4.2.7)$$

$$= (g\tilde{g})(n'\tilde{n}) \quad (4.2.8)$$

$$= (g\tilde{g})n'' \quad (4.2.9)$$

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

$$[(g_1N)(g_2N)](g_3N) = [(g_1g_2)N](g_3N) \quad (4.2.10)$$

$$= ((g_1g_2)g_3)N \quad (4.2.11)$$

$$= (g_1(g_2g_3))N \quad (4.2.12)$$

$$= (g_1N)[(g_2g_3)N] \quad (4.2.13)$$

$$= (g_1N)[(g_2N)(g_3N)]. \quad (4.2.14)$$

Hence,  $G/N$  is a group.  $\square$

■ **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  $\mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{Z}_n$ , namely  $m + n\mathbb{Z} \mapsto e^{2i\pi m/n}$ . 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](#).  $\square$

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  $\varphi: G \rightarrow H$  a homomorphism. Then

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

*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

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

so  $k_1 k_2^{-1} \in \ker \varphi$  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

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

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  $\psi: G/\ker \varphi \rightarrow \text{Im } \varphi$ . To do so we consider the obvious choice that  $\psi(gK) = \varphi(g)$ , where  $K = \ker \varphi$ . Since  $\psi$  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  $\psi$  is well-defined. Next, we verify that  $\psi$  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

$$\psi((g_1K)(g_2K)) = \psi((g_1g_2)K) \quad (4.2.23)$$

$$= \varphi(g_1g_2) \quad (4.2.24)$$

$$= \varphi(g_1)\varphi(g_2) \quad (4.2.25)$$

$$= \psi(g_1K)\psi(g_2K), \quad (4.2.26)$$

so  $\psi$  is a homomorphism.

Finally, we show that  $\psi$  is an isomorphism, that is that it is bijective. The kernel of  $\psi$  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  $\psi$  is the trivial group  $\{K\} \subset G/K$ . This proves that  $\psi$  is injective by [Lemma 2.1.23](#). Finally, let  $h \in \text{Im } \varphi$ . Then there exists  $g \in G$  such that  $\varphi(g) = h$ . We then have  $\psi(gK) = \varphi(g) = h$ , and so  $\psi$  is surjective.

So  $\psi$  is a bijective homomorphism and hence an isomorphism.  $\square$

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

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & \text{Im } \varphi \subseteq H \\ \pi \downarrow & \nearrow \psi & \\ G/\ker \varphi & & \end{array} \quad (4.2.27)$$

Here  $\pi: G \rightarrow G/\ker \varphi$  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  $\varphi: G \rightarrow \tilde{G}$  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

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

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

$$(h, j)(h', j') = (hh', jj') \quad (5.1.3)$$

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

$$(h_1, j_1)[(h_2, j_2)(h_3, j_3)] = (h_1, j_1)(h_2h_3, j_2j_3) \quad (5.1.5)$$

$$= (h_1(h_2h_3), j_1(j_2j_3)) \quad (5.1.6)$$

$$= ((h_1h_2)h_3, (j_1j_2)j_3) \quad (5.1.7)$$

$$= (h_1h_2, j_1j_2)(h_3, j_3) \quad (5.1.8)$$

$$= [(h_1, j_1)(h_2, j_2)](h_3, j_3) \quad (5.1.9)$$

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

$$(h, j)(e_H, e_J) = (he_H, je_J) = (h, j) \quad (5.1.10)$$

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

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

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.  $\square$

■ **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

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

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

$$(x, y) + (x', y') = (x + x', y + y'). \quad (5.1.15)$$

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:

$$\mathbb{Z}_n \times \mathbb{Z}_m \cong \mathbb{Z}_{mn}. \quad (5.1.16)$$

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  $H$ :

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

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  $\varphi: H \times J \rightarrow J$  a group action of  $H$  on  $J$ . Then we can define a new group,  $H \ltimes J$ , called the **semidirect product** of  $H$  and  $J$  as the set

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

and the group product

$$(h, j)(h', j') = (hh', j\varphi(h, j')) = (hh', j(h \cdot j')) \quad (5.2.3)$$

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



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  $\text{ISO}(n)$  or  $\text{E}(n)$ , 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  $\text{O}(n)$ , 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  $x \in \mathbb{R}^n$ , 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  $\text{ISO}(n)$  as a semidirect product,  $\text{ISO}(n) = \text{O}(n) \ltimes \mathbb{R}^n$ . We just need to work out the correct group action. First we need to state how  $\text{ISO}(n)$  acts on  $\mathbb{R}^n$  (as Euclidean space). This is simple, given a rotation and/or reflection  $R \in \text{O}(n)$  and translation  $a \in \mathbb{R}^n$  this acts on  $x \in \mathbb{R}^n$  as  $(R, a) \cdot x = Rx + a$ , that is we rotate  $x$  with  $R$  and then translate by  $a$ , notice that the order is important, which we will see means that  $\text{ISO}(n)$  is *not* a direct product of  $\text{O}(n)$  and  $\mathbb{R}^n$ .

Consider what happens when we act on some  $x \in \mathbb{R}^n$  (as a point in Euclidean space) by two isometries,  $(R, a)$  and  $(R', a')$ , where  $R, R' \in \text{O}(n)$  and  $a, a' \in \mathbb{R}^n$  (as the group of translations). We then have

$$(R', a')(R, a) \cdot x = (R', a') \cdot (R, a) \cdot x \quad (5.2.5)$$

$$= (R', a) \cdot (Rx + a) \quad (5.2.6)$$

$$= R'(Rx + a) + a' \quad (5.2.7)$$

$$= R'Rx + R'a + a' \quad (5.2.8)$$

$$= (R'R, R'a + a') \cdot x. \quad (5.2.9)$$

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

$$R' \cdot a = R'a, \quad \text{or} \quad \varphi(R', a) = Ra, \quad (5.2.10)$$



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,  $\text{Aff}(V)$ , which is the semidirect product  $\text{GL}(V) \ltimes V$  where  $\text{GL}(V)$  acts on the vector space  $V$  with the expected action,  $M \cdot v = Mv$  for  $M \in \text{GL}(V)$  and  $v \in V$ . This contains  $\text{ISO}(V)$  as a subgroup.

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

$$D_n := \langle r, s \mid r^n = s^2 = e, s^{-1}rs = r^{-1} \rangle. \quad (5.2.12)$$

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  $\mathbb{Z}_n = \{e^{2i\pi m/n}\}$ . A few examples of products in  $D_n$  are then

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

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

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

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

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  $120^\circ$  and  $s$  to be a reflection in the vertical. Graphically the action of  $r$  and  $s$  on the triangle is

$$r \cdot \begin{array}{c} 2 \\ \triangle \\ 1 \quad 3 \end{array} = \begin{array}{c} 1 \\ \triangle \\ 3 \quad 2 \end{array}, \quad \text{and} \quad s \cdot \begin{array}{c} 2 \\ \triangle \\ 1 \quad 3 \end{array} = \begin{array}{c} 2 \\ \triangle \\ 3 \quad 1 \end{array} \quad (5.2.16)$$

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  $3 \cdot 120^\circ = 360^\circ$ , which is the same as no rotation at all. Second, notice that repeating

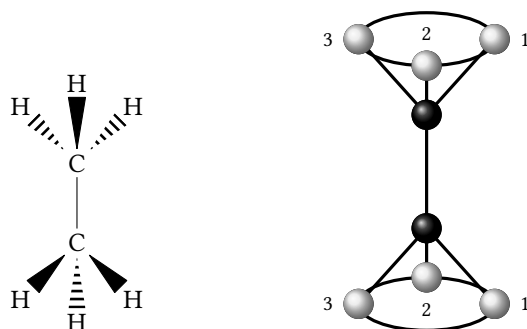


Figure 5.1: Ethane molecule.

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

$$\begin{aligned}
 srs^{-1} \cdot \begin{array}{c} 2 \\ \triangle \\ 1 \quad 3 \end{array} &= sr \cdot \begin{array}{c} 2 \\ \triangle \\ 3 \quad 1 \end{array} \\
 &= s \cdot \begin{array}{c} 3 \\ \triangle \\ 1 \quad 2 \end{array} \\
 &= \begin{array}{c} 3 \\ \triangle \\ 2 \quad 1 \end{array}
 \end{aligned} \tag{5.2.17}$$

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

$$r^{-1} \cdot \begin{array}{c} 2 \\ \triangle \\ 1 \quad 3 \end{array} = \begin{array}{c} 3 \\ \triangle \\ 2 \quad 1 \end{array} \tag{5.2.18}$$

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  $180^\circ$  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**,  $\sigma$ , on  $n$  objects is a bijection  $\sigma: X \rightarrow X$  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.

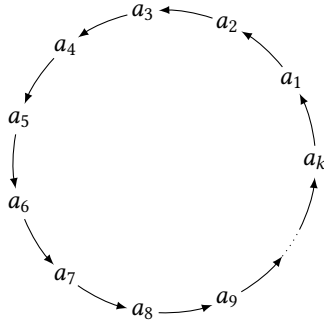
*Proof.* Let  $\sigma, \rho \in S_n$ . Then both of these are bijections on some set  $X$  with  $|X| = n$ . Their composition is defined by  $(\sigma \circ \rho)(x) = \sigma(\rho(x))$  for all  $x \in X$ . Straight away we see that this is indeed a permutation, since  $\sigma \circ \rho: X \rightarrow X$  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,  $\text{id}_X: X \rightarrow X$  defined by  $\text{id}_X(x) = x$  for all  $x \in X$  is a permutation and  $\text{id}_X \circ \sigma = \sigma$  for all  $\sigma \in S_n$ . As previously discussed  $\sigma$  has the inverse  $\sigma^{-1}$ , which is such that  $\sigma \circ \sigma^{-1} = \text{id}_X$ . Hence,  $S_n$  is a group.  $\square$

$S_n$  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:

$$\sigma \cdot (x_1, \dots, x_n) = (\sigma(x_1), \dots, \sigma(x_n)). \quad (6.1.4)$$

**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  $(a, b, c, d)$ :

$$(1\ 4\ 2)(1\ 2\ 3) \cdot (a\ b\ c\ d) = (1\ 4\ 2) \cdot (c, a, b, d) \quad (6.1.7)$$

$$= (a, d, b, c) \quad (6.1.8)$$

$$= (2\ 3\ 4) \cdot (a, b, c, d), \quad (6.1.9)$$

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  $\sigma$  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  $\rho = (nk)\sigma$ . We have that  $\rho(n) = (nk)\sigma(n) = (nk)(k) = n$ , here we are treating  $(nk)$  as a function,  $(nk): \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ , defined by  $(nk)(n) = k$ ,  $(nk)(k) = n$  and  $(nk)(a) = a$  for  $a \neq n, k$ . So

we can think of  $\rho$  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  $\tau_i$  only contain the numbers  $1, \dots, n-1$ , and each appears in at most one of these cycles.

Clearly  $(nk)(nk) = ()$ , and so it follows that  $\sigma = (nk)(nk)\sigma = (nk)\tau_1 \cdots \tau_r$ . If  $k$  doesn't appear in any of the cycles  $\tau_i$  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  $\tau_1$ .

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

$$(nk)\tau_1 = (nk)(k a_1 \dots a_m) = (nk a_1 \dots a_m). \quad (6.1.12)$$

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

$$\sigma = (nk a_1 \dots a_m)\tau_2 \cdots \tau_r \quad (6.1.13)$$

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  $n \in \mathbb{N}$ .  $\square$

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

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

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

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

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  $a_{i+1}$ .

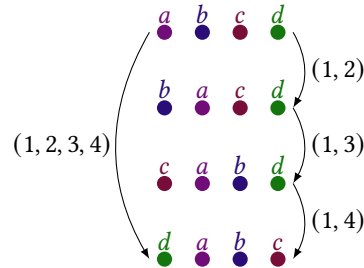


Figure 6.2: The permutation  $(1\ 2\ 3\ 4)$  acts on  $(a, b, c, d)$ , which can be done in steps where each step is a transposition.

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.  $\square$

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

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

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

$$(a_1\ a_2) = (1\ a_1)(1\ a_2)(1\ a_1). \quad (6.1.19)$$

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

$$(1\ a_1)(1\ a_2)(1\ a_1)(1 \dots a_1 \dots a_2 \dots) \quad (6.1.20)$$

$$= (1\ a_1)(1\ a_2)(a_1 \dots 1 \dots a_2 \dots) \quad (6.1.21)$$

$$= (1\ a_1)(a_1 \dots a_2 \dots 1 \dots) \quad (6.1.22)$$

$$= (1 \dots a_2 \dots a_1 \dots) \quad (6.1.23)$$

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  $S_n$ .  $\square$

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

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

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  $S_n$ .  $\square$

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

$$(k\ k+1) = g(1\ 2)g^{-1} \quad (6.1.27)$$

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  $S_n$ .  $\square$

**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.  $\square$

## 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

$$P(x_1, \dots, x_n) = P(\mathbf{x}) \quad (6.2.2)$$

$$:= (x_1 - x_2)(x_1 - x_3) \cdots (x_1 - x_n) \quad (6.2.3)$$

$$\cdot (x_2 - x_3) \cdots (x_2 - x_n) \cdots (x_{n-1} - x_n) \quad (6.2.4)$$

$$= \prod_{\substack{i, j \in \{1, \dots, n\} \\ i < j}} (x_i - x_j). \quad (6.2.5)$$

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  $(x_1, \dots, x_n)$ . We can define a group action,  $\varphi: S_n \times X \rightarrow X$ , in the usual way as  $S_n$  acting on  $\mathbf{x} \in X$  by permutation.

Now define  $\text{sgn}: S_n \rightarrow \{\pm 1\}$  by  $\text{sgn}(\sigma) = P(\sigma \cdot \mathbf{x})/|P(\mathbf{x})|$ . Then if  $\text{sgn}(\sigma) = 1$  we say  $\sigma$  is an **even permutation** and if  $\text{sgn}(\sigma) = -1$  we say that  $\sigma$  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  $\text{sgn}: S_n \rightarrow \mathbb{Z}_2$  as defined in Definition 6.2.1 is a group homomorphism. First notice that  $\text{sgn}(\sigma\rho) = P(\sigma\rho \cdot \mathbf{x})/|P(\mathbf{x})|$ . Now consider  $\text{sgn}(\sigma) \text{sgn}(\rho) = P(\sigma \cdot \mathbf{x})P(\rho \cdot \mathbf{x})/|P(\mathbf{x})|^2$ . Hence,  $\text{sgn}(\sigma) \text{sgn}(\rho) = 1$  if  $\sigma$  and  $\rho$  have the same parity and  $\text{sgn}(\sigma) \text{sgn}(\rho) = -1$  if  $\sigma$  and  $\rho$  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,  $\text{sgn}(\sigma\rho) = P(\sigma\rho \cdot \mathbf{x})/|P(\mathbf{x})|$  is 1 if  $\sigma$  and  $\rho$  are the same parity and  $-1$  if they are opposite parities. Hence,  $\text{sgn}$  is a homomorphism. The kernel of  $\text{sgn}$  is  $A_n$ , since even permutations map to 1 by definition. Hence, by Theorem 4.2.20  $A_n$  is a normal subgroup of  $S_n$ .  $\square$

**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  $A_n$ .  $\square$

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	$A_5$

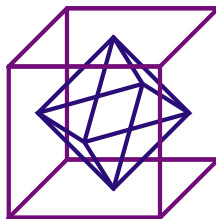


Figure 7.1: The octahedron is the dual of the cube.

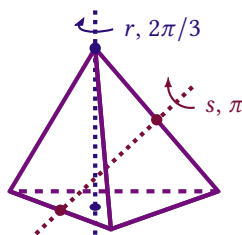
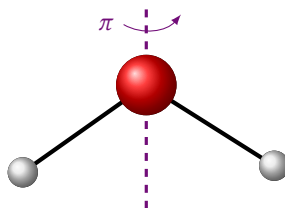


Figure 7.2: Two possible symmetries of the tetrahedron

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,  $\text{C}_2\text{H}_4$ , shown in Figure 5.1. This has symmetry group  $D_3 \cong \mathbb{Z}_2 \times \mathbb{Z}_3$ . 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  $\text{C}_2\text{H}_4$  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.

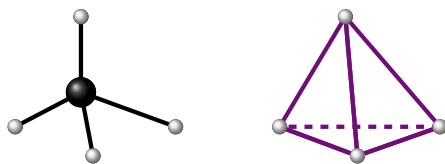
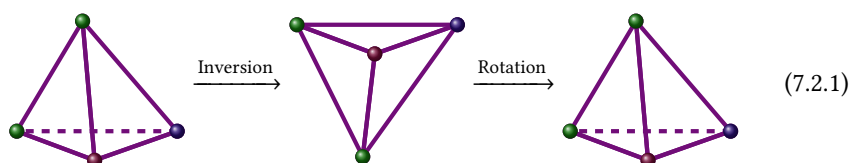


Figure 7.4: Methane is a tetragonal molecule, meaning that it has a central atom, here carbon, and four atoms arranged around it making the points of a tetrahedron.

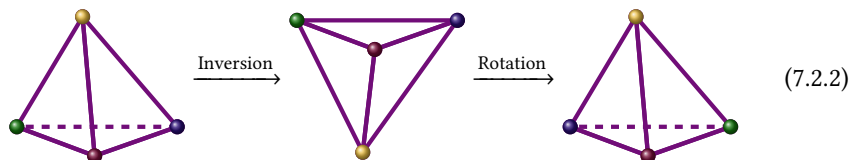
It can be shown that only molecules with a cyclic symmetry group,  $\mathbb{Z}_n$ , can have permanent dipoles.

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,  $\text{CH}_4$ , shown in Figure 7.4. An example of a tetragonal molecule which is achiral is  $\text{CCl}_2\text{BrI}$ , this is shown here:



On the other hand  $\text{CFClBrI}$  is chiral, this is shown here:



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  $U(1) \cong SO(2)$  and  $SO(3) \cong SU(2)/\mathbb{Z}_2$  are compact. More generally  $O(n)$ ,  $SO(n)$ ,  $Spin(n)$ ,  $U(n)$ , and  $SU(n)$  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  $\varphi: G \times V \rightarrow V$  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  $\rho: G \rightarrow GL(V)$  is a representation if it is a homomorphism.

**R** If we are going to be rigorous about it we should really define a (linear) representation as an ordered pair,  $(\rho, V)$ , where  $\rho: G \rightarrow GL(V)$  is a homomorphism and  $V$  is a linear space. This is similar to how the technically correct definition of a group is an ordered pair  $(G, \cdot)$ , but mostly we just talk about the group  $G$ . People refer to both  $\rho$  and  $V$  as the representation, depending on which they care about at this point in time. For example, we will say things like “let  $\rho$  be a representation of  $G$ ”, if we want to talk about properties of representations which are independent of the vector space, and we will say “the dimension of the representation”, when really we mean the dimension of  $V$ .

The equivalence of these two definitions is simple, if  $\varphi(g, v) = g \cdot v = \rho(g)v$  for all  $g \in G$  and  $v \in V$  then  $\varphi$  and  $\rho$  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  $\rho$ .

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  $\mathbb{Z}_n = \{e^{2i\pi m/n} \mid m = 0, \dots, n-1\}$  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  $GL(\mathbb{C}) = \mathbb{C} \setminus \{0\}$ . 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  $\text{GL}(V)$  with  $\text{GL}(n, \mathbb{F})$ , 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,  $\mathbf{e}_1 = (1, 0, 0)^\top$ ,  $\mathbf{e}_2 = (0, 1, 0)^\top$ , and  $\mathbf{e}_3 = (0, 0, 1)^\top$ . We can easily construct matrices that permute these, for example we want  $\rho((1\ 2))\mathbf{e}_1 = \mathbf{e}_2$ ,  $\rho((1\ 2))\mathbf{e}_2 = \mathbf{e}_1$  and  $\rho((1\ 2))\mathbf{e}_3 = \mathbf{e}_3$ . This fully determines the matrix  $\rho((1\ 2))$ :

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

Similarly, we have

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

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  $\mathbf{v} = (1, 1, 1)^\top$  is a common eigenvector for both of these matrices, and hence there is an invariant subspace,  $\text{span}\{\mathbf{v}\}$ , 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,  $\rho'$ , in this block diagonal basis such that

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

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

### 8.3 Basic Definitions

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

The permutation representation of  $S_3$  is faithful.

**Definition 8.3.2 — Trivial Representation** Let  $\rho: G \rightarrow \text{GL}(V)$  be defined by  $\rho(g) = \rho(e) = \mathbb{1}_V$ , where  $\mathbb{1}_V$  is the identity in  $\text{GL}(V)$  and  $V$  is an *arbitrary vector space*. We say that  $\rho$  is a **trivial representation**.

The **trivial representation** is  $\rho: G \rightarrow \text{GL}(V)$  defined by  $\rho(g) = \rho(e) = \mathbb{1}_V$ , where  $\mathbb{1}_V$  is the identity in  $\text{GL}(V)$  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  $\rho: G \rightarrow \text{U}(V)$  is a homomorphism then we say that  $\rho$  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  $\{e^{2i\pi m/n} \mid m = 0, \dots, n-1\}$ .

**Definition 8.3.4 — Equivalence of Representations** If  $\rho$  and  $\rho'$  are representations of  $G$  on  $V$  we say that  $\rho$  and  $\rho'$  are **equivalent** if they are represented by a **similarity transform**, that is  $\rho(g) = S\rho'(g)S^{-1}$  for some  $S \in \text{GL}(V)$ . 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  $S \in \text{GL}(V)$ , and so  $\rho'(g) = S^{-1}\rho(g)S$ , identifying  $S = (S^{-1})^{-1}$  and knowing that if  $S \in \text{GL}(V)$  we also have  $S^{-1} \in \text{GL}(V)$  we see that  $\rho' \sim \rho$ . Finally, if  $\rho \sim \rho'$  and  $\rho' \sim \rho''$  there exists  $S, T \in \text{GL}(V)$  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  $S, T \in \text{GL}(V)$  it follows that  $ST \in \text{GL}(V)$ , and hence  $\rho \sim \rho''$ .

**Definition 8.3.5 — Invariant Subspace** Let  $\rho: G \rightarrow \text{GL}(V)$  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  $\rho(g)w \in W$  for all  $g \in G$  and  $w \in W$ . Using the group-action definition of a representation  $W$  is an invariant subspace if  $\text{Orb}(w) \subseteq W$  for all  $w \in W$ .

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  $\text{span}\{(1, 1, 1)^\top\}$  as an invariant subspace. All representations leave the trivial subspace,  $\{0\}$ , invariant.

**Definition 8.3.6 — Irreducible** We call a representation **irreducible** if it has no invariant subspaces, apart from the trivial zero-dimensional subspace,  $\{0\}$ . 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,  $\rho: G \rightarrow \text{GL}(V)$ , 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  $\rho$  is equivalent to  $\rho^*$ .

A **complex representation** is a representation,  $\rho: G \rightarrow \text{GL}(V)$ , where  $\rho$  is not equivalent to  $\rho^*$ .

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

The permutation representation of  $S_3$  is real, and hence so is  $\rho'$ , even though it is possible that  $a_{ij}$  and/or  $b_{ij}$  are not real, since  $\rho$  and  $\rho'$  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,  $\mathbb{R}^{1,3} \ltimes \text{O}(1,3)$ . 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  $\rho: G \rightarrow \text{GL}(V)$  a representation. Let  $\langle -, - \rangle$  be an inner product on  $V$ . 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

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

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  $\rho$  is a homomorphism we then have

$$\langle \rho(g')x, \rho(g')y \rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)\rho(g')x, \rho(g)\rho(g')y \rangle \quad (8.4.3)$$

$$= \sum_{g \in G} \langle \rho(gg')x, \rho(gg')y \rangle \quad (8.4.4)$$

$$= \sum_{g'' \in G} \langle \rho(g'')x, \rho(g'')y \rangle \quad (8.4.5)$$

$$= \langle x, y \rangle_G. \quad (8.4.6)$$

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

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

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  $\rho$  is a unitary representation.  $\square$

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  $\rho: G \rightarrow \text{GL}(V)$  and  $\rho': G \rightarrow \text{GL}(V')$  be irreducible representations of the group  $G$  on some finite dimensional vector spaces  $V$  and  $V'$ . Let  $T: V \rightarrow V'$  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  $\text{Im } T$  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

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

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  $\rho$ .

Now recall that  $\text{Im } T = \{v' \in V' \mid v' = T(v) \text{ for some } v \in V\}$ . This is an invariant subspace if  $\rho'(g)v' \in \text{Im } T$  for all  $v' \in \text{Im } T$ . To show this we notice that for  $v' \in \text{Im } T$  we have  $v' = Tv$  for some  $v \in V$  and so

$$\rho'(g)v' = \rho'(g)Tv = (\rho'(g) \circ T)v = (T \circ \rho(g))v = T\rho(g)v \quad (8.4.10)$$

and so  $\rho'(g)v'$  is of the form  $T\rho(g)v$  and  $\rho(g)v \in V$  meaning  $\rho'(g)v' \in \text{Im } T$ . Hence,  $\text{Im } T$  is an invariant subspace of  $V'$  under  $\rho'$ .

By definition  $\rho$  and  $\rho'$  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  $\text{Im } T$  must be  $\{0\}$  or  $V'$ . We now treat this by cases.

- Suppose  $\ker T = \{0\}$ . Then  $\text{Im } T \neq \{0\}$  since all  $v \in V$  with  $v \neq 0$  map to something other than 0, meaning that  $\text{Im } T$  must contain nonzero elements. Hence,  $\text{Im } T = V'$ . It follows that  $T$  is an isomorphism since a trivial kernel implies that  $T: V \rightarrow V'$  is injective since  $V' = \text{Im } T$  this map is also surjective.
- Suppose  $\ker T = V$ . Then  $\text{Im } T = \{0\}$ , 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  $T = \lambda \mathbb{1}$ .  $\square$

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:

$$\begin{array}{ccc} V & \xrightarrow{T} & V' \\ \rho(g) \downarrow & & \downarrow \rho'(g) \\ V & \xrightarrow{T} & V' \end{array} \quad (8.4.11)$$

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  $\rho$  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  $\rho: G \rightarrow \text{GL}(V)$  be a reducible representation for some compact group  $G$ . Then we can write

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

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

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

*Proof.*

$\square$

# Nine

## Character Theory of Finite Groups

### 9.1 Basics of Character Theory

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

$$\chi(g) := \text{tr}[\rho(g)]. \quad (9.1.2)$$

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

**Lemma 9.1.4** The characters of two equivalent representations are equal.

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

$$\chi_{\rho'}(g) := \text{tr}[\rho'(g)] \quad (9.1.5)$$

$$= \text{tr}[S\rho(g)S^{-1}] \quad (9.1.6)$$

$$= \text{tr}[S^{-1}S\rho(g)] \quad (9.1.7)$$

$$= \text{tr}[\rho(g)] \quad (9.1.8)$$

$$=: \chi_\rho(g). \quad (9.1.9)$$

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

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  $[a] := \{b \in A \mid a \sim b\}$ . The set of all equivalence classes is denoted  $A/\sim$ .

**Definition 9.1.10** A **class function** is a function,  $f: A \rightarrow B$ , 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  $\tilde{f}: A/\sim \rightarrow B$  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  $\rho$  be a representation of  $G$ . Then

$$\chi_\rho(x) := \text{tr}[\rho(x)] \quad (9.1.12)$$

$$= \text{tr}[\rho(gyg^{-1})] \quad (9.1.13)$$

$$= \text{tr}[\rho(g)\rho(y)\rho(g^{-1})] \quad (9.1.14)$$

$$= \text{tr}[\rho(g)\rho(y)\rho(g)^{-1}] \quad (9.1.15)$$

$$= \text{tr}[\rho(g)^{-1}\rho(g)\rho(y)] \quad (9.1.16)$$

$$= \text{tr}[\rho(y)] \quad (9.1.17)$$

$$=: \chi_\rho(y). \quad (9.1.18)$$

Here we have used the fact that  $\rho$  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,  $\text{tr}(ABC) = \text{tr}(CAB)$ . Hence,  $\chi_\rho(x) = \chi_\rho(y)$  for all  $x, y \in [x] \in G/\sim$  where  $\sim$  is conjugacy.  $\square$

**Lemma 9.1.19** Let  $\rho$  and  $\rho'$  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 bracket notation. Let  $\{|i\rangle\}$  be an orthonormal basis. Then in bracket notation with the Einstein summation convention

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

Therefore,

$$\chi_{\rho \oplus \rho'}(g) := \text{tr}[(\rho \oplus \rho')(g)] \quad (9.1.21)$$

$$= \text{tr}[\rho(g) \oplus \rho'(g)] \quad (9.1.22)$$

$$= \langle i|(\rho(g) \oplus \rho'(g))|i\rangle \quad (9.1.23)$$

$$= \langle i|\rho(g)|i\rangle + \langle i|\rho'(g)|i\rangle \quad (9.1.24)$$

$$= \text{tr}[\rho(g)] + \text{tr}[\rho'(g)] \quad (9.1.25)$$

$$=: \chi_\rho(g) + \chi_{\rho'}(g). \quad (9.1.26)$$

Here we have used the fact that

$$(A \oplus B)(|v\rangle \oplus |w\rangle) = A|v\rangle \oplus B|w\rangle \quad (9.1.27)$$

and so

$$(\langle v'| \oplus \langle w'|)(A \oplus B)(|v\rangle \oplus |w\rangle) = (\langle v'| \oplus \langle w'|)(A|v\rangle \oplus B|w\rangle) \quad (9.1.28)$$

$$= \langle v'|A|v\rangle + \langle w'|B|w\rangle. \quad (9.1.29)$$

Similarly,

$$\chi_{\rho \otimes \rho'}(g) := \text{tr}[(\rho \otimes \rho')(g)] \quad (9.1.30)$$

$$= \text{tr}[\rho(g) \otimes \rho'(g)] \quad (9.1.31)$$

$$= \langle i | (\rho(g) \otimes \rho'(g)) | i \rangle \quad (9.1.32)$$

$$= \langle i | \rho(g) | i \rangle \langle i | \rho'(g) | i \rangle \quad (9.1.33)$$

$$= \text{tr}[\rho(g)] \text{tr}[\rho'(g)] \quad (9.1.34)$$

$$=: \chi_\rho(g) \chi_{\rho'}(g). \quad (9.1.35)$$

Here we have used the fact that

$$(A \otimes B)(|v\rangle \otimes |w\rangle) = A|v\rangle \otimes B|w\rangle. \quad (9.1.36)$$

□

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

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

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

□

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

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

## 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  $\rho$  we define a vector

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

where  $g_i$  is in the  $k$ th 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:

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

where  $c_k$  is the number of elements in the  $k$ th conjugacy class and  $\chi_{ak} := \chi_{\rho_a}(g_k)$  with  $g_k$  being a representative member of the  $k$ th 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<sup>a</sup> of some finite group,  $G$ . Denote these representations by  $\rho_{V_a}$  and  $\rho_{V_b}$  respectively. The space of all homomorphisms  $V_a \rightarrow V_b$  is denoted  $\text{Hom}(V_a, V_b)$ , 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,  $V_a \rightarrow \mathbb{F}$ , 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  $\dim(\text{Hom}(V_a, V_b)_0) = 1$ . On the other hand if  $V_a \not\cong V_b$  then  $\dim(\text{Hom}(V_a, V_b)) = 0$ . We can sum this up as  $\dim(\text{Hom}(V_a, V_b)_0) = \delta_{ab}$ .

The character of the representation  $V_a^* \otimes V_b$  is given by

$$\chi_{\rho_{V_a^* \otimes V_b}} = \chi_{\rho_{V_a}}^* \chi_{\rho_{V_b}}, \quad (9.2.5)$$

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

$$\chi_{\rho_{V_a^*}} = \chi_{\rho_{V_a}}^* \quad (9.2.6)$$

since

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

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

$$\chi_{\rho_{V_a^* \otimes V_b}} = \chi_{\rho_{V_a}}^* \chi_{\rho_{V_b}}. \quad (9.2.8)$$



Now define the operator  $\varphi$  according to

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

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  $\varphi$  has these properties notice that applying  $\varphi$  to a vector gives an invariant vector and the only invariant vectors correspond to vectors in the trivial representation subspace. Further, applying  $\varphi$  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

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

Combining all of the above we have

$$\delta_{ab} = \dim(\text{Hom}(V_a, V_b)) \quad (9.2.11)$$

$$= \dim((V_a^* \otimes V_b)_0) \quad (9.2.12)$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_{V_a^* \otimes V_b}}(g) \quad (9.2.13)$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_{V_a}}(g)^* \chi_{\rho_{V_b}}(g) \quad (9.2.14)$$

$$=: \langle \chi_{\rho_{V_a}}, \chi_{\rho_{V_b}} \rangle. \quad (9.2.15)$$

This completes the proof.  $\square$

<sup>a</sup>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,  $G \rightarrow \text{GL}(V_a)$  and  $G \rightarrow \text{GL}(V_b)$ .

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

$$\rho = \bigoplus_{i=1}^k m_i \rho_i = \bigoplus_{i=1}^k \rho_i^{\oplus m_i} = m_1 \rho_1 \oplus \cdots \oplus m_k \rho_k \quad (9.2.16)$$

where  $\rho_i$  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  $\rho_i$  correspond to linearly independent, in fact orthonormal, vectors in the character space. We have

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

**Corollary 9.2.19** The multiplicity of some particular irreducible representation,  $\rho_a$ , in the decomposition of  $\rho$  is

$$m_a = \langle \chi_{\rho_a}, \chi_\rho \rangle \quad (9.2.20)$$

*Proof.* By the linearity of the character

$$\chi_\rho = \sum_{i=1}^k m_i \chi_{\rho_i} \quad (9.2.21)$$

Then by the linearity of the inner product we have

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

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

$$= \sum_{i=1}^k m_i \delta_{ai} \quad (9.2.24)$$

$$= m_a \quad (9.2.25)$$

which completes the proof.  $\square$

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

$$v^i = \mathbf{e}_i \cdot \mathbf{v}. \quad (9.2.26)$$

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

$$\|\chi_\rho\|^2 := \langle \chi_\rho, \chi_\rho \rangle = \sum_i m_i^2 \quad (9.2.28)$$

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

*Proof.* Suppose  $\rho$  is irreducible. Then we can think of  $\rho$  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  $m_1^2 + \dots + m_k^2 = 1$ . Since  $m_i \in \mathbb{Z}_{\geq 0}$  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  $\rho$  is irreducible.  $\square$

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

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

*Proof.* For any representation,  $\rho: G \rightarrow V$  since  $\rho$  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  $\text{tr } \mathbb{1}_V = \dim V$ , since  $\mathbb{1}_V$  is just a matrix with ones on the diagonal.  $\square$

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

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

### 9.2.1 Dimensionality Theorem

**Definition 9.2.32 — Regular Representation** Let  $V$  be a vector space of dimension  $|G|$ . Let  $\{e_g\}$  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  $\{e_g\}$  in the obvious way by

$$g \cdot e_{g'} := e_{gg'}. \quad (9.2.33)$$

The **regular representation**,  $\rho_R: G \rightarrow \text{GL}(V)$  is defined as the representation associated with this action, that is

$$\rho_R(g)e_{g'} = e_{gg'}. \quad (9.2.34)$$

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

$$\chi_{\rho_R}(g) = \begin{cases} 0 & \text{if } g \neq e, \\ |G| & \text{if } g = e. \end{cases} \quad (9.2.35)$$

This follows since for  $g \neq e$  the diagonal of  $\rho_R(g)$  must be zero, since if they weren't, it would mean that  $e_{gg} = e_g$ , which means that  $g^2 = g$ , which can only happen for  $g = e$ . Additionally  $\rho_R(e) = \mathbb{1}_V$  and so  $\chi_R(e) = \text{tr } \mathbb{1}_V = \dim V = |G|$ .

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

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

*Proof.* Let  $\rho_R$  be the regular representation. This can be written as

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

Notice that we then have

$$m_i = \langle \chi_{\rho_i}, \chi_{\rho_R} \rangle \quad (9.2.39)$$

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

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

$$= \dim V_i. \quad (9.2.42)$$

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

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

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

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

where we have used [Lemma 9.2.29](#) again in the last equality.  $\square$

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.* Ⓡ 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,  $f: G \rightarrow \mathbb{C}$ , such that  $f$  is orthogonal to all of the characters of the irreducible representations, that is

$$\langle f, \chi_{\rho_i} \rangle = 0 \quad (9.3.3)$$

for all irreducible representations,  $\rho_i$ . 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

$$\varphi: V_i \rightarrow V_i \quad \text{where} \quad \varphi := \sum_{g \in G} f(g)^* \rho_i(g). \quad (9.3.4)$$

Here  $V_i$  is the vector space associated with the irreducible representation  $\rho_i$ .

It can easily be seen that

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

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

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

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

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

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,

$$\text{tr } \varphi = \lambda \dim V_i \quad (9.3.10)$$

and

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

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

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

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,  $\rho_R$ . However, since  $\rho_R(g)$  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.  $\square$

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  $\rho_i: G \rightarrow \text{GL}(V_i)$  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)

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

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.  $\square$

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

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

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

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

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

$$UU^\dagger = U^\dagger U = \mathbb{1}. \quad (9.3.20)$$

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

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

This will give a square matrix since by [Theorem 9.3.2](#) there are  $N_c$  irreducible representations,  $\rho_a$ , and  $N_c$  conjugacy classes,  $[g]$ .

Now consider the product  $UU^\dagger$ :

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

$$= \sum_{g \in G/\sim} U_{a[g]} U_{b[g]}^* \quad (9.3.23)$$

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

$$= \frac{1}{|G|} \sum_{k=1}^{N_c} c_k \chi_{\rho_a}(g_k) \chi_{\rho_b}(g_k)^* \quad (9.3.25)$$

$$= \frac{1}{|G|} \sum_{k=1}^{N_c} c_k \chi_{ak} \chi_{bk}^* \quad (9.3.26)$$

$$= \delta_{ab}. \quad (9.3.27)$$

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

$$(U^\dagger U)_{[g][g']} = \sum_{a=1}^{N_c} (U^\dagger)_{[g]a} U_{a[g']} \quad (9.3.28)$$

$$= \sum_{a=1}^{N_c} U_{a[g]}^* U_{a[g']} \quad (9.3.29)$$

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

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



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

$$= \delta_{[g][g']}. \quad (9.3.33)$$

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

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

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

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

and so we have

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

completing the proof.  $\square$

## 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  $\sigma$ . 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  $\text{Orb}_G(x)$  we then have a bijection associating  $\sigma \cdot x$  and  $\sigma \text{Stab}_G(x)$ .

Since  $G$  is cyclic (as it is generated by a single element) it follows that  $G/\text{Stab}_G(x)$  is cyclic. Its order is the smallest positive integer,  $d$ , such that  $\sigma^d \in \text{Stab}_G(x)$ . We know that  $d = |\text{Stab}(x)| = [G : \text{Stab}_G(x)]$  and so with the aforementioned bijection we have that the unique cosets of  $\text{Stab}_G(x)$  in  $G$  are

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

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

$$x, \sigma \cdot x, \dots, \sigma^{d-1} \cdot x. \quad (9.4.3)$$

Therefore on any orbit of size  $d$   $\sigma$  is a  $d$ -cycle. This shows the existence of a cycle decomposition.

Uniqueness is then fairly simple. Each cycle determined by  $\sigma$  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

$$\sigma^j \cdot x, \sigma^{j+1} \cdot x, \dots, \sigma^{d-1} \cdot x, x, \sigma \cdot x, \dots, \sigma^{j-1} \cdot x. \quad (9.4.4)$$

This is the same cycle permuted left by  $j$ . □

**Definition 9.4.5 — Cycle Type** Given  $\sigma \in S_n$  we can write  $\sigma$  as the product of  $k$  disjoint cycles of lengths  $n_1, \dots, n_k$ , where, since disjoint cycles commute, we can take  $n_i \leq n_{i+1}$ . We then define  $n_1, \dots, n_k$  to be the **cycle type** of  $\sigma$ . 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  $\sigma$  and  $\tau$  are conjugate. Then there exists  $\rho \in S_n$  such that  $\tau = \rho\sigma\rho^{-1}$ . We can write  $\sigma$  as a product of disjoint cycles. To show that  $\sigma$  and  $\tau$  have the same cycle type it suffices to show that if  $j$  follows  $i$  in the cycle decomposition of  $\sigma$  then  $\rho(j)$  follows  $\rho(i)$  in the cycle decomposition of  $\tau$ . Suppose that  $\rho(i) = j$ . Then

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

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  $\sigma$  and  $\tau$  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,

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

Define  $\rho$  to be the permutation taking  $a_i$  to  $b_i$ . Since the cycle types match we have

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

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  $\sigma$  and  $\tau$  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  $m_i \leq m_{i+1}$ .

**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.  $\square$

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

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

*Proof.* Consider a given arrangement of  $\{1, \dots, n\}$ . There are  $n!$  such arrangements. For each arrangement 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.  $\square$

### 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 **1**. The trivial representation is  $\mathbf{1}: S_3 \rightarrow \text{GL}(\mathbb{C})$ , defined by  $\mathbf{1}(\sigma) = 1$  for all  $\sigma \in S_3$ . This has character  $\chi_1(\sigma) = \text{tr}[\mathbf{1}(\sigma)] = \text{tr}[1] = 1$ . The first row of the character table is thus

$S_3$	$[(())]$	$[(1\ 2)]$	$[(1\ 2\ 3)]$
$ [g] $	1	3	2
<b>1</b>	1	1	1

(9.4.14)

For the permutation group there is a second one-dimensional representation, **1'** that we can define.

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

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

$$\begin{array}{c|ccc}
 S_3 & [()] & [(12)] & [(123)] \\
 |[g]| & 1 & 3 & 2 \\
 \hline
 \mathbf{1} & 1 & 1 & 1 \\
 \mathbf{1}' & 1 & -1 & 1
 \end{array} \tag{9.4.16}$$

Here we have used the fact that  $() = (12)(12)$  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 2, which we call the **standard representation** of  $S_3$ .

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 2. It turns out to be the latter. More generally for  $S_n$  the standard representation is such that  $\rho_{\text{triv}} \oplus \rho_{\text{stand}} = \rho_{\text{perm}}$ .

We then use the simple formula

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

where  $\chi_1 = (1, 1, 1)$  and  $\chi_{\rho_{\text{perm}}} = (3, 1, 0)$ , 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

$$\chi_{\rho_{\text{stand}}} = \chi_{1 \oplus 2} = \chi_1 + \chi_2 \implies \chi_2 = (3, 1, 0) - (1, 1, 1) = (2, 0, -1). \tag{9.4.20}$$

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

$$\begin{array}{c|ccc}
 S_3 & [()] & [(12)] & [(123)] \\
 |[g]| & 1 & 3 & 2 \\
 \hline
 \mathbf{1} & 1 & 1 & 1 \\
 \mathbf{1}' & 1 & -1 & 1 \\
 \mathbf{2} & 2 & 0 & -1
 \end{array} \tag{9.4.21}$$

### 9.4.2 Character Table of $S_4$

Four has the following partitions with the corresponding conjugacy classes

$$4 = 1 + 1 + 1 + 1 \leftrightarrow [()], \quad (9.4.22)$$

$$4 = 1 + 1 + 2 \leftrightarrow [(12)], \quad (9.4.23)$$

$$4 = 2 + 2 \leftrightarrow [(12)(34)], \quad (9.4.24)$$

$$4 = 1 + 3 \leftrightarrow [(123)], \quad (9.4.25)$$

$$4 = 4 \leftrightarrow [(1234)] \quad (9.4.26)$$

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

The dimensions of the irreducible representations are the solutions to

$$|S_4| = 24 = 1^2 + 1^2 + d_3^2 + d_4^2 + d_5^2 \quad (9.4.27)$$

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  $d_3 = 2$ , and  $d_4 = d_5 = 3$ .

Filling in the first two rows of the table we have

$S_4$	$[()]$	$[(12)]$	$[(12)(34)]$	$[(123)]$	$[(1234)]$	(9.4.28)
$ [g] $	1	6	3	8	6	
$\mathbf{1}$	1	1	1	1	1	
$\mathbf{1}'$	1	-1	1	1	-1	

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

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

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

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

In general given a representation,  $\rho$ , of the symmetric group  $S_n$  there is another representation defined by  $\text{sgn}(\sigma)\rho(\sigma)$ . Note that this is only distinct from  $\rho$  if there exists a conjugacy class  $[\sigma]$  such that  $\rho(\sigma)$  is not zero and  $\text{sgn}(\sigma) = -1$ . Putting these in the table we have

$S_4$	$[()]$	$[(12)]$	$[(12)(34)]$	$[(123)]$	$[(1234)]$	(9.4.31)
$ [g] $	1	6	3	8	6	
$\mathbf{1}$	1	1	1	1	1	
$\mathbf{1}'$	1	-1	1	1	-1	
$\mathbf{2}$	.	.	.	.	.	
$\mathbf{3}$	3	1	-1	0	-1	
$\mathbf{3}'$	3	-1	-1	0	1	

We can find the character of the final irreducible representation,  $\mathbf{2}$ , 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

$$\langle \chi_2([()]), \chi_2([(1\ 2)]) \rangle = \frac{1}{24} \sum_{i=1}^5 \chi_{\rho_i}()^* \chi_{\rho_i}((1\ 2)) \quad (9.4.32)$$

$$= \frac{1}{24} (1 \cdot 1 + 1 \cdot (-1) + 2 \cdot b + 3 \cdot 1 + 3 \cdot (-1)) \quad (9.4.33)$$

$$= \frac{b}{12} \quad (9.4.34)$$

$$= 0 \quad (9.4.35)$$

so we have  $b = 0$ . Next

$$\langle \chi_2([()]), \chi_2([(1\ 2)(3\ 4)]) \rangle = \frac{1}{24} \sum_{i=1}^5 \chi_{\rho_i}()^* \chi_{\rho_i}((1\ 2)(3\ 4)) \quad (9.4.36)$$

$$= \frac{1}{24} (1 \cdot 1 + 1 \cdot 1 + 2 \cdot c + 3 \cdot (-1) + 3 \cdot (-1)) \quad (9.4.37)$$

$$= \frac{1}{24} (2c - 4) \quad (9.4.38)$$

$$= 0 \quad (9.4.39)$$

so we have  $c = 2$ . Continuing on we have

$$\langle \chi_2([()]), \chi_2([(1\ 2\ 3)]) \rangle = \frac{1}{24} \sum_{i=1}^5 \chi_{\rho_i}()^* \chi_{\rho_i}((1\ 2\ 3)) \quad (9.4.40)$$

$$= \frac{1}{24} (1 \cdot 1 + 1 \cdot 1 + 2 \cdot d + 3 \cdot 0 + 3 \cdot 0) \quad (9.4.41)$$

$$= \frac{1}{24} (2 + 2d) \quad (9.4.42)$$

$$= 0 \quad (9.4.43)$$

giving  $d = -1$ . Finally

$$\langle \chi_2([()]), \chi_2([(1\ 2\ 3\ 4)]) \rangle = \frac{1}{24} \sum_{i=1}^5 \chi_{\rho_i}()^* \chi_{\rho_i}((1\ 2\ 3\ 4)) \quad (9.4.44)$$

$$= \frac{1}{24} (1 \cdot 1 + 1 \cdot (-1) + 2 \cdot e + 3 \cdot (-1) + 3 \cdot 1) \quad (9.4.45)$$

$$= \frac{1}{24} (2e) \quad (9.4.46)$$

$$= 0$$

so  $e = 0$ .

Thus, the complete character table is

$S_4$ $ [g] $	$[(\ )]$	$[(1\ 2)]$	$[(1\ 2)(3\ 4)]$	$[(1\ 2\ 3)]$	$[(1\ 2\ 3\ 4)]$
<b>1</b>	1	1	1	1	1
<b>1'</b>	1	-1	1	1	-1
<b>2</b>	2	0	2	-1	0
<b>3</b>	3	1	-1	0	-1
<b>3'</b>	3	-1	-1	0	1

(9.4.47)

### 9.4.3 Alternating Groups

Suppose we want to know the character table of the alternating group,  $A_4$ . We first have to work out the conjugacy classes of  $A_4$ . It can be shown that the conjugacy classes of  $A_n$  are the same as  $S_n$ , unless the cycle type consists of cycles of distinct odd lengths, including one element cycles, which are just the identity. In this case the conjugacy class splits into two equal size classes. For  $A_4$  the only conjugacy class with this property is  $[(1\ 2\ 3)]_{S_4}$ , which splits into  $[(1\ 2\ 3)]_{A_4}$  and  $[(1\ 3\ 2)]_{A_4}$ . We also have to get rid of any conjugacy classes corresponding to an odd permutation, so  $[(1\ 2)]_{S_4}$  and  $[(1\ 2\ 3\ 4)]_{S_4}$ .

We can work out the dimensions of the irreducible representations by solving

$$|A_4| = \frac{4!}{2} = 12 = 1^2 + d_2^2 + d_3^2 + d_4^2, \quad (9.4.48)$$

and requiring that  $d_i$  divides  $|A_4| = 12$  we see that  $d_2 = d_3 = 1$  and  $d_4 = 3$ . Note that there is no alternating representation since all elements of  $A_4$  are even so the alternating representation is identical to the trivial representation.

In order to find the two unknown one dimensional representations we consider the fact that  $(1\ 2\ 3)^3 = (\ )$ , and so we have  $\rho((1\ 2\ 3)^3) = \rho((\ )) = 1$ , but we also know that  $\rho$  is a homomorphism, so  $\rho((1\ 2\ 3)^3) = \rho((1\ 2\ 3))^3$ , from which we conclude that  $\rho((1\ 2\ 3))$  is a primitive root of unity. This means that  $\rho((1\ 2\ 3)) \in \{1, \omega, \omega^2\}$ , where  $\omega = \exp[2\pi i/3]$ . Clearly  $\rho((1\ 2\ 3)) = 1$  gives the trivial representation. Suppose that  $\rho((1\ 2\ 3)) = \omega$ . Then we can use the fact that  $(1\ 2\ 3)^2 = (1\ 3\ 2)$  and so  $\rho((1\ 3\ 2)) = \rho((1\ 2\ 3)^2) = \rho((1\ 2\ 3))^2 = \omega^2$ . Similarly if we started with  $\rho((1\ 2\ 3)) = \omega^2$  we would find that  $\rho((1\ 3\ 2)) = \rho((1\ 2\ 3))^2 = \omega^4 = \omega$ . Since  $((1\ 2)(3\ 4))^2 = (\ )$  we have that  $\rho((1\ 2)(3\ 4))$  is a second root of unity, so  $\pm 1$ , and it turns out that 1 is the correct value.

Finally we can find the character of the three dimensional representation by an orthogonality argument. This gives the final character table as

$A_4$ $ [g] $	$[(\ )]$	$[(1\ 2\ 3)]$	$[(1\ 3\ 2)]$	$[(1\ 2)(3\ 4)]$
<b>1</b>	1	1	1	1
<b>1'</b>	1	$\omega$	$\omega^2$	1
<b>1''</b>	1	$\omega^2$	$\omega$	1
<b>3</b>	3	0	0	-1

(9.4.49)

### 9.4.4 Cyclic Groups

Since  $\mathbb{Z}_n$  is Abelian all of its representations are one-dimensional. This follows since being Abelian means that all conjugacy classes contain a single element, and therefore there are  $n$  conjugacy classes and  $n$  irreducible representations. The only way

for the dimensionality theorem to hold is if each irreducible representation has dimension 1. Think of  $\mathbb{Z}_n$  as the group of  $\{0, \dots, n-1\}$  with addition modulo  $n$ . Apart from the trivial representation the most obvious representation is one that we have already been using, namely  $\rho_1(n) = e^{2\pi i/n}$ . It can easily be seen that  $\rho_2(n) = e^{2\pi i 2/n}$  is also a representation, and so on giving the representations  $\rho_k(n) = e^{2\pi i k/n}$ , for  $k = 1, \dots, n-1$ , if we take  $k = 0$  we even get the trivial representation in this form.

As a particular example consider  $\mathbb{Z}_3$ . The representations are then

$$\rho_0(0) = 1, \quad \rho_0(1) = 1, \quad \rho_0(2) = 1, \quad (9.4.50)$$

$$\rho_1(0) = 1, \quad \rho_1(1) = \omega, \quad \rho_1(2) = \omega^2, \quad (9.4.51)$$

$$\rho_2(0) = 1, \quad \rho_2(1) = \omega^2, \quad \rho_2(2) = \omega. \quad (9.4.52)$$

From here we can easily fill out the character table:

$A_4$	[0]	[1]	[2]	(9.4.53)
$ [g] $	1	1	1	
1	1	1	1	
1'	1	$\omega$	$\omega^2$	
1''	1	$\omega^2$	$\omega$	

## 9.5 Complex, Real, and Pseudo-Real Representations

**Definition 9.5.1 — Complex, Real, and Pseudo-Real Representations** Let  $\rho: G \rightarrow \text{GL}(V)$  be a representation of the group  $G$  on the vector space  $V$ . If  $\rho$  is inequivalent to its complex conjugate,  $\bar{\rho} = \rho^*$ , defined by  $\bar{\rho}(g) = \overline{\rho(g)}$ , then we say that  $\rho$  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  $\rho$  is a **real representation**.

Surprisingly, there is a third type representation. It turns out that it is possible for  $\rho$  to be equivalent to  $\bar{\rho}$  and yet there be no basis such that the elements of  $\rho(g)$  are always real. In this case we call  $\rho$  a **pseudo-real representation** or **quaternionic representation**.

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

Both real and pseudo-real representations are isomorphic to their conjugates. This forces the existence of an equivariant anti-linear map,  $j: V \rightarrow V$ , which is simply the isomorphism composed with conjugation. **Equivariant** means that if the group  $H$  acts on  $V$  then  $h \cdot j(v) = j(h \cdot v)$  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  $V \rightarrow V$  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,  $\rho: G \rightarrow \text{GL}(V, \mathbb{R})$  with an antilinear equivariant map,  $j: V \rightarrow V$ , 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,  $\rho: G \rightarrow \text{GL}(V, \mathbb{R})$  with an antilinear equivariant map,  $j: V \rightarrow V$ , 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,  $\rho: G \rightarrow \text{GL}(V, \mathbb{H})$ .

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

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

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  $\sigma_i$  are the **Pauli matrices** defined by

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

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

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

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

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

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

and

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

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  $\rho \sim \bar{\rho}$  whereas for complex representations  $\rho \not\sim \bar{\rho}$  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

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

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

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

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:

$$V \otimes V = \text{Sym}(V \otimes V) \oplus \text{Asym}(V \otimes V). \quad (9.5.13)$$

If  $V$  is the vector space for a non-complex representation and the trivial representation falls into the  $\text{Sym}(V \otimes V)$  part of the decomposition then the representation is real. On the other hand if the trivial representation falls in the  $\text{Asym}(V \otimes V)$  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  $\rho: G \rightarrow \text{GL}(V)$  is a representation then  $\rho|_H$  is certainly a representation, where by  $\rho|_H$  we mean  $\rho$  restricted to  $H$ , so  $\rho|_H: H \rightarrow \text{GL}(V)$  is defined by  $\rho|_H(h) = \rho(h)$  for all  $h \in H$ . Crucially it is possible that  $\rho$  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:

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

where  $\rho_i$  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

$S_3$ $ [g] $	$[(1)]$	$[(1\ 2)]$	$[(1\ 2\ 3)]$
$\mathbf{1}$	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,  $\rho_1(h) = 1$  for all  $h \in \mathbb{Z}_3$ . Another representation is given by  $\rho_{1'}(1) = e^{2\pi i/3}$  and  $\rho_{1'}(2) = e^{2\pi i 2/3} = e^{-2\pi i/3}$ , 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  $\rho_{1''}(1) = e^{-2\pi i/3}$  and  $\rho_{1''}(2) = e^{2\pi i/3}$ . This allows us to complete the character table for  $\mathbb{Z}_3$ :

$\mathbb{Z}_3$ $ [h] $	$[(1)]$	$[(1\ 2\ 3)]$	$[(1\ 3\ 2)]$
$\mathbf{1}$	1	1	1
$\mathbf{1}'$	1	$\omega$	$\omega^*$
$\mathbf{1}''$	1	$\omega^*$	$\omega$

(9.6.4)

where  $\omega := e^{2\pi i/3}$ .

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  $\mathbb{Z}_3$ :

$$\rho_{S_3,i}|_{\mathbb{Z}_3} = \bigoplus_{j=1}^3 m_{ij} \rho_{\mathbb{Z}_3,j} \quad (9.6.5)$$

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

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

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

$$\langle \chi_{1_{S_3}}, \chi_{1_{\mathbb{Z}_3}} \rangle = \langle \chi_{1'_{S_3}}, \chi_{1'_{\mathbb{Z}_3}} \rangle = \langle \chi_{2_{S_3}}, \chi_{1'_{\mathbb{Z}_3}} \rangle = \langle \chi_{2_{S_3}}, \chi_{1''_{\mathbb{Z}_3}} \rangle = 1. \quad (9.6.7)$$

This gives the branching rules

$$\rho_{1_{S_3}}|_{\mathbb{Z}_3} = \rho_{1_{\mathbb{Z}_3}} \quad \mathbf{1}_{S_3} \rightarrow \mathbf{1}_{\mathbb{Z}_3}, \quad (9.6.8)$$

$$\rho_{1'_{S_3}}|_{\mathbb{Z}_3} = \rho_{1'_{\mathbb{Z}_3}} \quad \mathbf{1}'_{S_3} \rightarrow \mathbf{1}'_{\mathbb{Z}_3}, \quad (9.6.9)$$

$$\rho_{2_{S_3}}|_{\mathbb{Z}_3} = \rho_{1'_{\mathbb{Z}_3}} \oplus \rho_{1''_{\mathbb{Z}_3}} \quad \mathbf{2}_{S_3} \rightarrow \mathbf{1}'_{\mathbb{Z}_3} \oplus \mathbf{1}''_{\mathbb{Z}_3}. \quad (9.6.10)$$

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  $1_{S_3} \rightarrow 1_{\mathbb{Z}_3}$ .

The alternating representation,  $1'_{S_3}$ , 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  $1'_{S_3} \rightarrow 1_{\mathbb{Z}_3}$ .

The standard representation,  $2_{S_3}$ , has to branch into two one dimensional representations of  $\mathbb{Z}_3$  since the dimension of  $2_{S_3}$  is 2, and the dimension of a direct sum is the sum of the dimensions of the things being summed. Since  $2_{S_3}$  is a real representation it can split either into  $2 \cdot 1_{\mathbb{Z}_3}$  or  $1'_{\mathbb{Z}_3} \oplus 1''_{\mathbb{Z}_3}$ . It is not possible for it to split into  $1_{\mathbb{Z}_3} \oplus 1'_{\mathbb{Z}_3}$  since this is not real, which can be seen from

$$\overline{1_{\mathbb{Z}_3} \oplus 1'_{\mathbb{Z}_3}} = \overline{1_{\mathbb{Z}_3}} \oplus \overline{1'_{\mathbb{Z}_3}} = 1_{\mathbb{Z}_3} \oplus 1''_{\mathbb{Z}_3} \quad (9.6.12)$$

where we have used the fact that  $\overline{1_{\mathbb{Z}_3}} = 1_{\mathbb{Z}_3}$  since the trivial representation is real and  $\overline{1'_{\mathbb{Z}_3}} = 1''_{\mathbb{Z}_3}$ . The same logic forbids splitting into  $1_{\mathbb{Z}_3} \oplus 1''_{\mathbb{Z}_3}$ .

We can rule out the split into  $2 \cdot 1_{\mathbb{Z}_3}$  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  $2_{S_3} \rightarrow 1'_{\mathbb{Z}_3} \oplus 1''_{\mathbb{Z}_3}$ .

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 be rotational symmetry,  $SO(3)$ , being broken by turning on an external field which gives a preferred direction, breaking to rotational symmetry about this axis,  $SO(2)$ . In this case  $3_{SO(3)} \rightarrow 1_{SO(2)} \oplus 2_{SO(2)}$ .

## 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,

$$5 = 5 \quad = \begin{array}{|c|c|c|c|c|} \hline \square & \square & \square & \square & \square \\ \hline \end{array}, \quad (9.7.1)$$

$$= 4 + 1 \quad = \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & & & \\ \hline \end{array}, \quad (9.7.2)$$

$$= 3 + 2 \quad = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \\ \hline \end{array}, \quad (9.7.3)$$

$$= 3 + 1 + 1 \quad = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \square & & \\ \hline \end{array}, \quad (9.7.4)$$

$$= 2 + 2 + 1 \quad = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \\ \hline \end{array}, \quad (9.7.5)$$

$$= 2 + 1 + 1 + 1 \quad = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad (9.7.6)$$

$$= 1 + 1 + 1 + 1 + 1 = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}. \quad (9.7.7)$$

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 tableaux 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 tableaux of 5 are given below:

$$\begin{array}{|c|c|c|c|c|} \hline 5 & 4 & 3 & 2 & 1 \\ \hline \end{array}, \quad \begin{array}{|c|c|c|c|} \hline 5 & 3 & 2 & 1 \\ \hline 1 & & & \end{array}, \quad \begin{array}{|c|c|c|} \hline 4 & 3 & 1 \\ \hline 2 & 1 & \end{array}, \quad \begin{array}{|c|c|c|} \hline 5 & 2 & 1 \\ \hline 2 & & \\ \hline 1 & & \end{array}, \quad (9.7.10)$$

$$\begin{array}{|c|c|} \hline 4 & 2 \\ \hline 3 & 1 \\ \hline 1 & \end{array}, \quad \begin{array}{|c|c|c|} \hline 5 & 1 & \\ \hline 3 & 2 & 1 \\ \hline \end{array}, \quad \begin{array}{|c|} \hline 5 \\ \hline 4 \\ \hline 3 \\ \hline 2 \\ \hline 1 \\ \hline \end{array}. \quad (9.7.11)$$

The reason that we care about the hook length is because there is a rather remarkable theorem which states that the dimension of an irreducible representation of  $S_n$  is given by

$$\dim V_i = \frac{n!}{\prod \text{hook lengths}} \quad (9.7.12)$$

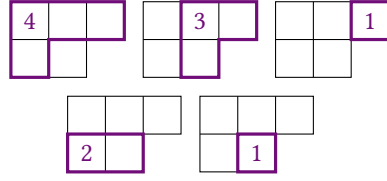


Figure 9.1: Calculating the hook lengths for the  $5 = 3 + 2$  partition. The hook length is the number of squares to the right or below a given square, plus the square itself, so in this drawing the hook length of a square is the number of squares in the purple hook when that square is at the top left of the hook.

where the product is of all hook lengths appearing in Young tableau associated with the representation. In the case of  $S_5$  this means

$$\dim V_1 = \frac{5!}{5!} = 1, \quad \begin{array}{|c|c|c|c|c|} \hline 5 & 4 & 3 & 2 & 1 \\ \hline \end{array}, \quad (9.7.13)$$

$$\dim V_2 = \frac{5!}{5 \cdot 3!} = 4, \quad \begin{array}{|c|c|c|c|c|} \hline 5 & 3 & 2 & 1 \\ \hline 1 & & & & \end{array}, \quad (9.7.14)$$

$$\dim V_3 = \frac{5!}{4!} = 5, \quad \begin{array}{|c|c|c|} \hline 4 & 3 & 1 \\ \hline 2 & 1 & \end{array}, \quad (9.7.15)$$

$$\dim V_4 = \frac{5!}{5 \cdot 2 \cdot 2} = 6, \quad \begin{array}{|c|c|c|} \hline 5 & 2 & 1 \\ \hline 2 & & \\ \hline 1 & & \end{array}, \quad (9.7.16)$$

$$\dim V_5 = \frac{5!}{4!} = 5, \quad \begin{array}{|c|c|} \hline 4 & 2 \\ \hline 3 & 1 \\ \hline 1 & \end{array}, \quad (9.7.17)$$

$$\dim V_6 = \frac{5!}{5 \cdot 3!} = 4, \quad \begin{array}{|c|c|c|} \hline 5 & 1 & \\ \hline 3 & 2 & 1 \\ \hline \end{array}, \quad (9.7.18)$$

$$\dim V_7 = \frac{5!}{5!} = 1, \quad \begin{array}{|c|} \hline 5 \\ \hline 4 \\ \hline 3 \\ \hline 2 \\ \hline 1 \\ \hline \end{array}. \quad (9.7.19)$$

While we won't prove that this works we can check that it doesn't violate the dimensionality theorem ([Theorem 9.2.36](#)):

$$1^2 + 4^2 + 5^2 + 6^2 + 5^2 + 4^2 + 1^1 = 120 = 5! = |S_5|. \quad (9.7.20)$$

The mirror symmetry between certain pairs of Young tableaux reflects a symmetry in the character tables. Namely that if  $\rho$  and  $\rho_{\text{mirror}}$  are irreducible representations corresponding to Young tableaux which are mirror images then

$$\chi_\rho(g) = \text{sgn}(g) \chi_{\rho_{\text{mirror}}}(g) \quad (9.7.21)$$

where  $\text{sgn}(g)$  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  $\chi_6(g) = \text{sgn}(g) \chi_{6_{\text{mirror}}}(g) = -\chi_6(g)$ , so we have  $\chi_6(g) = 0$  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,  $\rho_a: G \rightarrow \text{GL}(V_1)$  and  $\rho_b: G \rightarrow \text{GL}(V_b)$ , 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

$$\rho_a \otimes \rho_b = m_{ab}^1 \rho_1 \oplus \cdots \oplus m_{ab}^k \rho_k = \bigoplus_{c=1}^k m_{ab}^c \rho'_c, \quad (9.7.22)$$

with  $m_{ab}^c \in \mathbb{Z}_{>0}$ . 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

$$\dim(V_a) \dim(V_b) = \sum_{c=1}^k m_{ab}^c \dim V_c, \quad (9.7.23)$$

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  $\text{SO}(3)$ . Consider two three dimensional vectors,  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ , which transforms under an  $l = 1$  irreducible representation of the rotation group,  $\text{SO}(3)$ . The direct product of these can be written as the vector with components

$$v_i w_j = A_{ij} + B_{ij} \quad (9.7.24)$$

where

$$A_{ij} = v_i w_j - \frac{1}{3} \delta_{ij} \mathbf{v} \cdot \mathbf{w} \quad (9.7.25)$$

is traceless and

$$B_{ij} = \frac{1}{3} \delta_{ij} \mathbf{v} \cdot \mathbf{w}. \quad (9.7.26)$$

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  $(\mathbf{v} \cdot \mathbf{w})$  multiple of the identity matrix  $(\delta_{ij})$ . Scalars transform under an  $l = 0$  irreducible representation of  $\text{SO}(3)$ .

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,  $\mathbf{x} = \mathbf{v} \times \mathbf{w}$ , 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  $\mathrm{SO}(3)$  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:

$$m_{ab}^c = \langle \chi_{\rho_a \otimes \rho_b}, \chi_{\rho_c} \rangle \quad (9.7.27)$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a \otimes \rho_b}(g)^* \chi_{\rho_c}(g) \quad (9.7.28)$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g)^* \chi_{\rho_c}(g). \quad (9.7.29)$$

This leads to the following corollary.

**Corollary 9.7.30** The Kronecker product of two irreducible representations,  $\rho_a$  and  $\rho_b$ , contains the trivial representation with multiplicity 1 if and only if  $\rho_a = \overline{\rho_b}$ .

**(R)** If  $\rho_a$  and  $\rho_b$  are not irreducible then the multiplicity can be greater than 1.

*Proof.* Let  $\rho_a$  and  $\rho_b$  be irreducible representations of some finite group  $G$ . Then

$$m_{ab}^1 = \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g)^* \chi_1(g) \quad (9.7.31)$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a}(g)^* \chi_{\rho_b}(g)^* \quad (9.7.32)$$

where we have used the fact that  $\chi_1(g) = 1$  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

$$m_{ab}^1 = \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_a}(g) \chi_{\overline{\rho_b}}(g) = \langle \chi_{\rho_a}, \chi_{\overline{\rho_b}} \rangle = \delta_{a\overline{b}}. \quad (9.7.33)$$

This is 1 if and only if  $\rho_a = \overline{\rho_b}$ , and zero otherwise.  $\square$

If instead we allowed one or both of the representations above to be reducible we could simply use the linearity of the inner product to derive a similar result.

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

$$\rho_{\text{mirror}} = \rho_{1'} \otimes \rho. \quad (9.7.34)$$

### 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,  $\rho_a: H \rightarrow \text{GL}(V_a)$ , 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

$$W \cong \bigoplus_{i=1}^{[G:H]} g_{n_i} V_a \quad (9.7.35)$$

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  $g_{n_i} \mathbf{v} \in g_{n_i} V_a$  for  $\mathbf{v} \in V_a$ . 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

$$gg_{n_i} = g_{m_i} h_i \quad (9.7.36)$$

for some  $h_i \in H$ . This is possible since  $gg_{n_i} \in g_{m_i} H$  for some  $m_i$  since the left cosets,  $g_{m_i} H$ , partition  $G$ . The induced representation is then defined to act on  $\mathbf{w} := \bigoplus_{i=1}^{[G:H]} g_{n_i} \mathbf{v}_{n'_i} \in W$  by

$$\rho_{\text{induced}}(g) \mathbf{w} g \bigoplus_{i=1}^{[G:H]} g_{n_i} \mathbf{v}_{n'_i} = \bigoplus_{i=1}^{[G:H]} g_{m_i} \rho_a(h_i) \mathbf{v}_{n'_i}. \quad (9.7.37)$$

The dimension of the induced representation is the dimension of  $W$ . In general the induced representation is not irreducible.

The regular representation can be seen as the representation induced by the case where  $H$  is the trivial subgroup and  $\rho_a$  is the trivial representation. The vector space is  $V_0 = \mathbb{R}$  and we identify  $g\mathbf{v} \in gV_0$  as  $\mathbf{e}_g$  comparing to the notation of the original definition of the regular representation in [Definition 9.2.32](#).

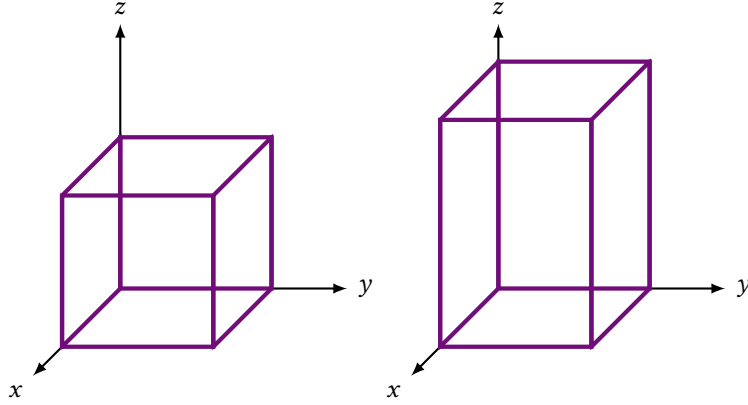
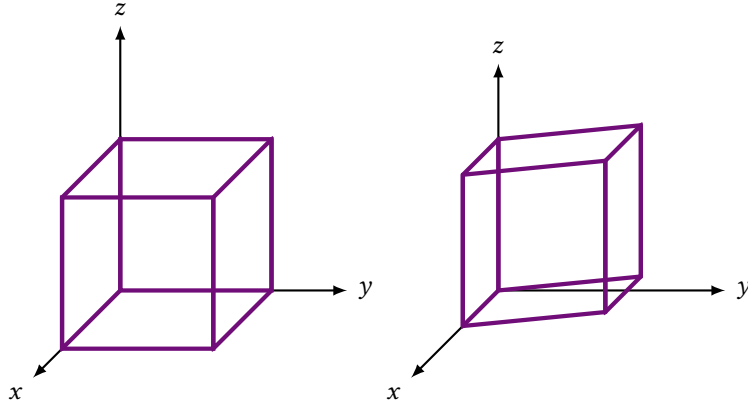
## 9.8 Applications

### 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<sup>1</sup>, where the cube has corners at  $(0, 0, 0)$ ,  $(0, 0, 1)$ ,  $(0, 1, 0)$  and so on.

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](#)

<sup>1</sup>see the notes from the Introduction to Condensed Matter Physics.

Figure 9.2: A cube distorted along the  $z$ -direction breaks symmetry  $S_4 \rightarrow D_4$ .Figure 9.3: A cube distorted along the  $(1, 1, 1)$ -direction breaks symmetry  $S_4 \rightarrow D_3$ .

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 form

$$|\Phi\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle \otimes \cdots \otimes |\varphi_N\rangle = |\varphi_1 \otimes \varphi_2 \otimes \cdots \otimes \varphi_N\rangle \quad (9.8.1)$$

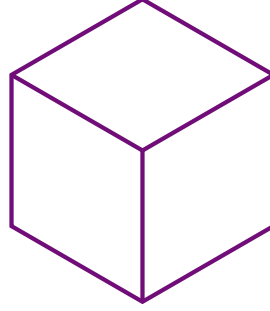


Figure 9.4: A cube distorted along the  $(1, 1, 1)$ -direction still has  $D_3$  symmetry as it can be viewed corner on as depicted here.

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  $N = N_A = 6 \times 10^{23}$  the number of irreducible representations is very large.

The two important cases of particles are bosons, particles with integer spin,  $s = 0, 1, 2, \dots$ , and fermions, particles with half integer spin,  $s = 1/2, 3/2, 5/2, \dots$ . 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

$$|\Phi_{\text{bosons}}\rangle = \frac{\sqrt{2}}{2}(|\varphi_1 \otimes \varphi_2\rangle + |\varphi_2 \otimes \varphi_1\rangle), \quad (9.8.2)$$

$$|\Phi_{\text{fermions}}\rangle = \frac{\sqrt{2}}{2}(|\varphi_1 \otimes \varphi_2\rangle - |\varphi_2 \otimes \varphi_1\rangle). \quad (9.8.3)$$

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 discuss 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  $\{\alpha\}$ . 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**,  $\text{GL}(n, \mathbb{F})$ , 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

$$\text{GL}(n, \mathbb{F}) := \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M \neq 0\}. \quad (10.2.4)$$

The requirement of non-zero determinant is so that the matrices are invertible.

- The **special linear group**,  $\text{SL}(n, \mathbb{F})$ , is defined as the group of transformations of an  $n$ -dimensional vector space over  $\mathbb{F}$  which preserve volumes. It has the representation

$$\text{SL}(n, \mathbb{F}) := \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M = 1\} \subset \text{GL}(n, \mathbb{F}). \quad (10.2.5)$$

- The **orthogonal group**,  $\text{O}(n)$ , is defined as the group of transformations of an  $n$ -dimensional vector space over  $\mathbb{R}$  which preserves lengths. It has the representation

$$\text{O}(n) := \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^T O = \mathbb{1}_n\} \subset \text{GL}(n, \mathbb{R}). \quad (10.2.6)$$

- The **special orthogonal group**,  $\text{SO}(n)$ , 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

$$\text{SO}(n) := \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^T O = \mathbb{1}_n \text{ and } \det O = 1\}. \quad (10.2.7)$$

It is a subgroup of both  $\text{O}(n)$  and  $\text{SL}(n, \mathbb{R})$ .

- The **unitary group**,  $\text{U}(n)$ , is defined as the group of transformations of an  $n$ -dimensional vector space over  $\mathbb{C}$  which preserves inner products. It has the representation

$$\text{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = \mathbb{1}_n\} \subset \text{GL}(n, \mathbb{C}). \quad (10.2.8)$$

- The **special unitary group**,  $\text{SU}(n)$ , 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

$$\text{SU}(n) := \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = \mathbb{1}_n \text{ and } \det U = 1\}. \quad (10.2.9)$$

It is a subgroup of both  $\text{U}(n)$  and  $\text{SL}(n, \mathbb{C})$ .

- The **symplectic group**,  $\text{USp}(2n) = \text{Sp}(n)$ , is defined as the group of transformations of an  $2n$ -dimensional vector space over  $\mathbb{F}$  which preserves symplectic bilinear forms<sup>a</sup>. It has the representation

$$\text{USp}(2n) = \text{Sp}(n) := \{S \in \mathcal{M}_{2n}(\mathbb{F}) \mid S^\dagger J S = J\} \quad (10.2.10)$$

where

$$J := \begin{pmatrix} 0 & -\mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix}. \quad (10.2.11)$$

Note that  $J^2 = -\mathbb{1}_{2n}$ , so we can think of  $J$  as a generalisation of  $i$  which has  $i^2 = -1$ .

<sup>a</sup>A **symplectic bilinear form** is a map  $\omega: V \times V \rightarrow \mathbb{F}$  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.
- $\text{SO}(3)$  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  $\text{SO}(n)$  is not, in general,  $n$ . For example,  $\text{SO}(2)$  has dimension 1, since there is only one parameter defining rotations in the plane, the angle of the rotation.  $\text{SO}(4)$  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 possess. 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  $\text{U}(1)$ ,  $\text{SO}(2)$ ,  $\text{SU}(2)$ , and  $\text{SO}(3)$  are all finite dimensional.

The Lie group  $\text{U}(\mathcal{H})$  for an infinite dimensional Hilbert space is an infinite dimensional Lie group under the topology induced by the operator norm,

$$\|A\|_{\text{op}} := \inf\{c \geq 0 \mid \|Av\| \leq c\|v\| \text{ for all } v \in \mathcal{H}\} \quad (10.3.3)$$



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  $U(1)$ ,  $SO(2)$ ,  $SU(2)$ , and  $SO(3)$  are real Lie groups, although we will later see that  $SU(2)$  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  $[0, 1] := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$  is closed and bounded, and hence compact. The intervals

$$[0, 1) := \{x \in \mathbb{R} \mid 0 \leq x < 1\}, \quad (10.3.7)$$

$$(0, 1] := \{x \in \mathbb{R} \mid 0 < x \leq 1\}, \quad \text{and} \quad (10.3.8)$$

$$(0, 1) := \{x \in \mathbb{R} \mid 0 < x < 1\} \quad (10.3.9)$$

are not closed, but are bounded. The interval  $[0, \infty) := \{x \in \mathbb{R} \mid x \geq 0\}$  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  $U(1)$ ,  $SO(2)$ ,  $SU(2)$ , and  $SO(3)$  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**,  $\int d\{\alpha\}$ , 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  $a \cdot x = x + a$ . This has a representation,  $\rho := \mathbb{R} \rightarrow \text{GL}(V)$ , given by the following:

$$\rho(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}. \quad (10.3.12)$$

However, this doesn't decompose since the subspace defined by the span  $\mathbf{e}_x = (1, 0)$  is an invariant subspace but its orthogonal complement, defined by the span of  $\mathbf{e}_y = (0, 1)$ , is not invariant, since  $\rho(a)\mathbf{e}_y = (a, 1) \notin \text{span}\{\mathbf{e}_y\}$ . Hence we cannot write  $\rho = \rho_1 \oplus \rho_2$  with  $\rho_1: \mathbb{R} \rightarrow \text{GL}(W)$  and  $\rho_2: \mathbb{R} \rightarrow \text{GL}(W^\perp)$  with  $W \oplus W^\perp = V$ , despite the fact that  $\rho$  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,  $O(1, 3)$ , 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  $\psi$  with  $\bar{\psi} := \psi^\dagger \gamma_0$  and  $\gamma_0 := \mathbb{1}_2 \oplus (-\mathbb{1}_2) = \sigma_3 \otimes \mathbb{1}_2$ . If  $\psi$  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  $\gamma_0$ .

### 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 connected** 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  $U(1)$ ,  $SO(2)$ , and  $SO(3)$  are connected. The Lie group  $SU(2)$  is simply connected. The Lie group  $O(2)$  is disconnected since there is no continuous path from  $O \in O(2)$  with  $\det O = -1$  to  $O' \in O$  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,

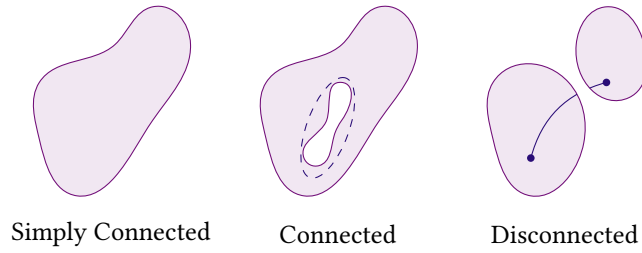


Figure 10.1: Examples of simply connected, connected, and disconnected spaces. The loop (---) in the connected example cannot be continuously contracted to a point without leaving the space. Similarly the two points (•) in the disconnected example cannot be connected by a continuous path. The path connecting 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  $SL(n)$ ,  $SU(n)$ , and  $USp(2n) = Sp(n)$  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,  $A_n = SU(n+1)$ ,  $B_n = SO(2n+1)$ ,  $C_n = USp(2n)$ , and  $D_n = SO(2n)$ . 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 groups can be depicted by **Dynkin diagrams**, which are graphs with as many nodes as the rank<sup>1</sup> of the group. For example,  $A_1$  corresponds to  $\circ$ ,  $A_2$  corresponds to  $\circ \text{---} \circ$ , and  $A_3$  corresponds to  $\circ \text{---} \circ \text{---} \circ$ . In general  $A_n$  corresponds to  $\circ \text{---} \circ \text{---} \cdots \text{---} \circ$  with  $n$  nodes.

<sup>1</sup>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  $\circ \rightleftarrows \circ$ , and  $B_3$  corresponds to  $\circ \text{---} \circ \rightleftarrows \circ$ . In general  $B_n$  corresponds to  $\circ \text{---} \circ \cdots \circ \rightleftarrows \circ$  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  $\circ \leftleftarrows \circ$ , and  $C_3$  corresponds to  $\circ \text{---} \circ \leftleftarrows \circ$ . In general  $C_n$  corresponds to  $\circ \text{---} \circ \cdots \circ \leftleftarrows \circ$  with  $n$  nodes.

The Dynkin diagrams for  $D_n$  branch at the end into two. So  $D_4$  corresponds to

$$\begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \text{---} \circ \end{array} \quad (10.3.18)$$

and  $D_5$  corresponds to

$$\begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \text{---} \circ \text{---} \circ \end{array} \quad (10.3.19)$$

In general,  $D_n$  corresponds to

$$\begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \circ \text{---} \circ \cdots \circ \end{array} \quad (10.3.20)$$

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

$$\begin{array}{c} \circ \\ | \\ \circ \text{---} \circ \text{---} \circ \text{---} \circ \end{array} \quad (10.3.21)$$

$E_7$  corresponds to

$$\begin{array}{c} \circ \\ | \\ \circ \text{---} \circ \text{---} \circ \text{---} \circ \text{---} \circ \end{array} \quad (10.3.22)$$

and  $E_8$  to

$$\begin{array}{c} \circ \\ | \\ \circ \text{---} \circ \text{---} \circ \text{---} \circ \text{---} \circ \text{---} \circ \end{array} \quad (10.3.23)$$

The Dynkin diagram for  $F_4$  is  $\circ \text{---} \circ \text{---} \text{---} \text{---} \circ \text{---} \circ$ .

The Dynkin diagram for  $G_2$  is  $\circ \text{---} \text{---} \text{---} \text{---} \circ$ .

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 $U(1) \cong SO(2)$

We can define the unitary group  $U(1)$  as

$$U(1) := \{U \in \mathcal{M}_1(\mathbb{C}) \mid U^\dagger U = \mathbb{1}\}. \quad (11.1.1)$$

This is obviously the same as the circle group

$$\mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\} \quad (11.1.2)$$

under the obvious isomorphism associating  $(z) \in U(1)$  with  $z \in \mathbb{T}$ . We therefore don't distinguish between  $U(1)$  and  $\mathbb{T}$  and will say things like  $z \in U(1)$ , when it may be more proper to say  $(z) \in U(1)$ . 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,  $SO(2)$ , as

$$SO(2) := \{R \in \mathcal{M}_2(\mathbb{R}) \mid R^\top R = \mathbb{1} \text{ and } \det R = 1\}. \quad (11.1.3)$$

This acts on the plane,  $\mathbb{R}^2$ , through rotations, which are given by normal matrix multiplication,  $R \cdot \mathbf{x} = R\mathbf{x}$ . This is a one dimensional real manifold parametrised by the rotation angle,  $\alpha \in [0, 2\pi)$ .

The unitary group  $U(1)$  has an infinite family of representations labelled by  $n \in \mathbb{Z}$  given by

$$\rho_n^{U(1)}(\alpha) = e^{in\alpha}. \quad (11.1.4)$$

Here we are implicitly associating complex numbers with  $1 \times 1$  complex matrices. The two-dimensional rotation group,  $SO(2)$ , has an infinite family of representations labelled by  $n \in \mathbb{Z}$  given by

$$\rho_n^{SO(2)}(\alpha) = \begin{pmatrix} \cos(n\alpha) & \sin(n\alpha) \\ -\sin(n\alpha) & \cos(n\alpha) \end{pmatrix}. \quad (11.1.5)$$

It should be clear that  $U(1)$  and  $SO(2)$  are isomorphic. One isomorphism between them being  $e^{i\alpha} \mapsto \rho_1^{SO(2)}(\alpha)$ . 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

$$\rho_n \otimes \rho_m = \rho_{n+m}. \quad (11.1.6)$$

In order to linearise  $SO(2)$  we make use of the fact that elements of  $U(1)$  can be expressed as  $e^{i\alpha}$  and we suggest that  $O \in SO(2)$  can be expressed as<sup>1</sup>

$$O = \exp[i\alpha T] \quad (11.1.7)$$

<sup>1</sup>the factor of  $i$  appearing in the exponent is a matter of convention, typically physicists include it and mathematicians absorb it in the definition of  $T$ , which results in slightly different requirements on  $T$ , in particular  $T$  will be anti-Hermitian (or skew-Hermitian), so  $T^\dagger = -T$

for some  $T \in \mathcal{M}_2(\mathbb{R})$ . As usual the exponential of a matrix is to be understood either through its power series,

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}, \quad (11.1.8)$$

or a limit,

$$\exp(A) = \lim_{n \rightarrow \infty} \left(1 + \frac{A}{n}\right)^n. \quad (11.1.9)$$

**Lemma 11.1.10** Let  $A \in \mathcal{M}_m(\mathbb{C})$  be diagonalisable. Then  $\det(\exp(A)) = \exp(\text{tr}(A))$ .



This theorem also holds for non-diagonalisable matrices. One way to show this involves the Jordan normal form. Another way is to note that diagonalisable matrices are dense in  $\mathcal{M}_n(\mathbb{C})$  and so every  $A \in \mathcal{M}_n(\mathbb{C})$  can be written as an infinite series of diagonalisable matrices. We can then apply this proof to the terms of the series and the fact that  $\text{tr}$  and  $\det$  are continuous guarantees the result holds as if we were to apply the theorem to the original matrix.

*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  $\text{tr}$  are basis independent. In this basis the diagonal of  $A$  consists of its eigenvalues,  $\lambda_i$ . We then have

$$\det A = \prod_{i=1}^m \lambda_i, \quad (11.1.11)$$

and

$$\text{tr} A = \sum_{i=1}^m \lambda_i. \quad (11.1.12)$$

Further in this basis  $\exp(A)$  is diagonal and its diagonal components are  $\exp(\lambda_i)$ . This follows since the  $n$ th power of a diagonal matrix just raises

the elements on the diagonal to the  $n$ th power, that is

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}^n = \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{pmatrix}. \quad (11.1.13)$$

It follows that

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} \quad (11.1.14)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}^n \quad (11.1.15)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_m^n \end{pmatrix} \quad (11.1.16)$$

$$= \begin{pmatrix} \sum_{n=0}^{\infty} \frac{\lambda_1^n}{n!} & & \\ & \ddots & \\ & & \sum_{n=0}^{\infty} \frac{\lambda_m^n}{n!} \end{pmatrix} \quad (11.1.17)$$

$$= \begin{pmatrix} \exp(\lambda_1) & & \\ & \ddots & \\ & & \exp(\lambda_m) \end{pmatrix} \quad (11.1.18)$$

Hence

$$\exp(\operatorname{tr}(A)) = \exp\left(\sum_i \lambda_i\right) \quad (11.1.19)$$

$$= \prod_i \exp(\lambda_i) \quad (11.1.20)$$

$$= \det(\exp(A)). \quad (11.1.21)$$

□

An alternative statement of this theorem is that the following diagram commutes. That is, the result is the same no matter what path you follow.

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\exp} & G \\ \operatorname{tr} \downarrow & & \downarrow \det \\ \mathbb{C} & \xrightarrow{\exp} & \mathbb{C}. \end{array} \quad (11.1.22)$$

Here  $G$  is a Lie group, and  $\mathfrak{g}$  is its Lie algebra (to be defined). Note that there are two subtly different functions called  $\exp$  here, one is a function  $\mathbb{C} \rightarrow \mathbb{C}$ , whereas the other is a function  $\mathfrak{g} \rightarrow G$ , or more generally from the tangent space,  $T_e M$ , to the manifold,  $M$ . Both can be defined through equivalent power series and agree when  $G$  is one-dimensional.



The definition of  $O \in \text{SO}(2)$  is that  $O^\top O = \mathbb{1}$ , which can be expressed as

$$\mathbb{1} = O^\top O \quad (11.1.23)$$

$$= \exp(i\alpha T)^\top \exp(i\alpha T) \quad (11.1.24)$$

$$= \exp(i\alpha T^\top) \exp(i\alpha T) \quad (11.1.25)$$

$$= (\mathbb{1} + i\alpha T^\top + O(\alpha^2))(\mathbb{1} + i\alpha T + O(\alpha^2)) \quad (11.1.26)$$

$$= \mathbb{1} + i\alpha(T^\top + T) + O(\alpha^2). \quad (11.1.27)$$

Therefore if we take  $\alpha$  to be small we have that  $T^\top = -T$ , which is to say that  $T$  must be antisymmetric. This in fact generalises to  $\text{SO}(n)$ , we can write any element as  $\exp(i\alpha T)$  where  $T \in \mathcal{M}_n(\mathbb{R})$  is antisymmetric.

One particular solution is

$$T = i \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad (11.1.28)$$

which leads to the previous representation since  $T^2 = -\mathbb{1}$  and so

$$\exp(i\alpha T) = \mathbb{1} \cos(\alpha) + iT \sin(\alpha) = \rho_1^{\text{SO}(2)}(\alpha). \quad (11.1.29)$$

This follows by expanding the exponential and collecting even and odd terms.

We define the Lie algebra of  $\text{SO}(2)$  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.

As previously mentioned the underlying manifold for  $\text{U}(1)$ , and hence  $\text{SO}(2)$ , 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  $f_2(\vartheta) = (\cos(\vartheta - \pi/2), \sin(\vartheta - \pi/2))$ , 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  $\text{SO}(2)$  is connected, since  $S^1$  is clearly connected. On the other hand the manifold  $\text{O}(2)$  is *not* connected. It is formed from two disconnected pieces, one, which has  $\det O = 1$ , is essentially a copy of  $\text{SO}(2)$ , 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  $\text{O}(2)$  is not connected.

Consider the parity operator

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (11.1.30)$$

This is an element of  $\text{O}(2)$ , but not  $\text{SO}(2)$ . 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

$$\text{SO}(2) \cong \text{O}(2)/\mathbb{Z}_2. \quad (11.1.31)$$

In fact, this generalises to

$$\text{SO}(n) \cong \text{O}(n)/\mathbb{Z}_2. \quad (11.1.32)$$

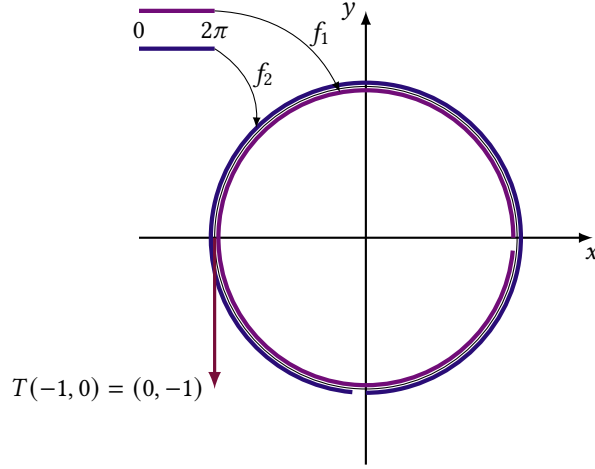


Figure 11.1: The manifold  $S^1$  covered by two charts,  $f_1$  and  $f_2$ .  $T(x, y)$  gives a tangent vector as demonstrated by the case of  $(x, y) = (-1, 0)$ . The slight gaps in the two charts represent the discontinuity in  $\alpha$  going from 0 to  $2\pi$ . Either chart is valid away from these discontinuities and at a discontinuity simply use the continuous chart.

Further we have

$$\mathrm{SU}(n) \cong \mathrm{U}(n)/\mathrm{U}(1) \quad (11.1.33)$$

since we can think of  $\mathrm{U}(1)$  as the complex version of  $\mathbb{Z}_2$ .

For  $\mathrm{SO}(2)$  the Haar measure is

$$\int_0^{2\pi} \frac{d\alpha}{2\pi} \quad (11.1.34)$$

and so we can define an inner product on the character space as

$$\langle \chi_{\rho_n}, \chi_{\rho_m} \rangle_{\mathrm{U}(1)} = \frac{1}{2\pi} \int_0^{2\pi} \chi_{\rho_n}(\alpha)^* \chi_{\rho_m}(\alpha) d\alpha \quad (11.1.35)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (e^{in\alpha})^* e^{im\alpha} d\alpha \quad (11.1.36)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)\alpha} d\alpha \quad (11.1.37)$$

$$= \delta_{nm}. \quad (11.1.38)$$

We see that the characters 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  $\text{GL}(n, \mathbb{C})$  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  $\rho_T: \mathbb{R} \rightarrow \text{GL}(n, \mathbb{C})$  defined by

$$\left. \frac{d}{dt} \rho_T(t) \right|_{t=0} = T \quad (11.2.2)$$

with the boundary condition  $\rho_T(0) = \mathbb{1}$ . Since  $\rho_T$  is a group homomorphism we require

$$\rho_T(s+t) = \rho_T(s)\rho_T(t). \quad (11.2.3)$$

Differentiating this with respect to  $s$  and evaluating at  $s = 0$  gives

$$\left. \frac{d}{ds} \rho_T(s+t) \right|_{s=0} = \left. \frac{d}{ds} \rho_T(s) \right|_{s=0} \rho_T(t) = T \rho_T(t) = \left. \frac{d}{ds} \rho_T(t+s) \right|_{s=0} = \left. \frac{d}{dt} \rho_T(t) \right|_{t=0} \rho_T(t) \quad (11.2.4)$$

where in the last step we set  $s+t = t$ . This gives us a differential equation

$$\frac{d}{dt} \rho_T(t) = T \rho_T(t) \quad (11.2.5)$$

which has the unique solution

$$\rho_T(t) = \exp(tT). \quad (11.2.6)$$

This proves the theorem.  $\square$

This theorem generalises to multi-parameter subgroups of  $\text{GL}(n, \mathbb{C})$ , where we can write all elements in the form

$$\exp(t^a T_a) \quad (11.2.7)$$

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  $\rho_3$  in the group by the closure of the group for a non-Abelian group. In order to avoid the **Baker–Campbell–Hausdorff formula**,

$$e^A e^B = \exp \left( A + B + \frac{1}{2} [A, B] + \frac{1}{2} \frac{1}{3!} ([A, [A, B]] + [[A, B], B]) + \cdots \right), \quad (11.2.8)$$

we expand  $\rho_i$  as

$$\rho_1 = \mathbb{1} + i\alpha^a T_a + \frac{1}{2}(i\alpha^a T_a)^2 + O(\alpha^3), \quad (11.2.9)$$

$$\rho_2 = \mathbb{1} + i\beta^a T_a + \frac{1}{2}(i\beta^a T_a)^2 + O(\beta^3), \quad (11.2.10)$$

$$\rho_3 = \mathbb{1} + i\gamma^a T_a + \frac{1}{2}(i\gamma^a T_a)^2 + O(\gamma^3). \quad (11.2.11)$$

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

$$[\alpha^a T_a, \beta^b T_b] = -i\gamma^c T_c + O(\alpha^3, \beta^3, \gamma^3). \quad (11.2.12)$$

All lower order terms cancel. We can therefore find  $\rho_3$  by finding  $\gamma^c = -f_{ab}^{\phantom{c}c} \alpha^a \beta^b$  where we define the **structure constants**,  $f_{ab}^{\phantom{c}c}$ , through the Lie algebra

$$[T_a, T_b] = if_{ab}^{\phantom{c}c} T_c. \quad (11.2.13)$$

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  $f_{ab}^{\phantom{c}c} = -f_{ba}^{\phantom{c}c}$ , 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  $SO(3)$  and  $SU(2)$  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  $\rho$  is connected to the identity we can write  $\rho = \exp(i\alpha^a T_a)$ . Compact groups admit finite dimensional unitary representations, and so we have

$$\mathbb{1} = \exp(i\alpha^a T_a)^\dagger \exp(i\alpha^a T_a) \quad (11.2.16)$$

$$= \exp(-i\alpha^a T_a^\dagger) \exp(i\alpha^a T_a) \quad (11.2.17)$$

$$= (\mathbb{1} - i\alpha^a T_a^\dagger + \mathcal{O}(\alpha^2))(\mathbb{1} + i\alpha^a T_a + \mathcal{O}(\alpha^2)) \quad (11.2.18)$$

$$= \mathbb{1} + i\alpha^a (T_a - T_a^\dagger) + \mathcal{O}(\alpha^2). \quad (11.2.19)$$

Hence we must have  $T_a - T_a^\dagger = 0$ , and since this must hold for all  $\rho$ , and hence for all  $\alpha^a$  we have  $T_a = T_a^\dagger$  meaning  $T_a$  is Hermitian individually for each  $a$ .  $\square$

**Corollary 11.2.20** The Hilbert–Schmidt inner product on  $\mathcal{M}_n(\mathbb{C})$  is defined by

$$\langle T_a, T_b \rangle := \text{tr}(T_a^\dagger T_b), \quad (11.2.21)$$

and if  $T_a$  are the generators of a Lie algebra associated with a compact group, that is  $T_a$  are Hermitian, we have

$$\langle T_a, T_b \rangle := \text{tr}(T_a^\dagger T_b) = \text{tr}(T_a T_b). \quad (11.2.22)$$

This leads to the following theorem,

**Theorem 11.2.23.** For a compact, semi-simple Lie algebra there is an orthogonal basis such that

$$\langle T_a, T_b \rangle := \text{tr}(T_a T_b) = 2k_R \delta_{ab}. \quad (11.2.24)$$

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**:

$$k_{ab} := \frac{1}{k_R} \text{tr}(T_a T_b) \quad (11.2.26)$$

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

$$0 = \delta_{T_c} \text{tr}(T_a T_b) = \text{tr}([T^c, T_a] T_b) + \text{tr}(T_a [T^c, T_b]) = ik_R (f_{ab}^c + f_{ba}^c). \quad (11.2.27)$$

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

$$k_{ab} = 2\delta_{ab}. \quad (11.2.28)$$

We can use the metric,  $k_{ab}$ , to raise and lower indices<sup>2</sup>. 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.

<sup>2</sup>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**:

$$f_{aed}f_{bce} + f_{bed}f_{cae} + f_{ced}f_{abc} = 0. \quad (11.2.30)$$

**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

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]]. \quad (11.2.31)$$

This can be shown by expanding these commutators. We can then interpret  $f_{abc}$  as the components of the matrix  $(T_a^{\text{adj}})_{bc} = -if_{abc}$  and the result follows.  $\square$

**Definition 11.2.32 — Adjoint Representation** The **adjoint representation** is given by defining the generator  $T_a^{\text{adj}}$  to have components  $(T_a^{\text{adj}})_{bc} = -if_{abc}$ .

**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

$$[T_a, T_b]^* = (if_{abc}T_c)^* \implies [T_a^*, T_b^*] = -if_{abc}^*T_c^* \quad (11.2.34)$$

noting that  $[-A, -B] = [A, B]$  since the negatives cancel when we expand the commutator we get

$$[-T_a^*, -T_b^*] = if_{abc}^*(-T_a^*). \quad (11.2.35)$$

Since  $f_{abc}$  are independent of the representation we must have  $f_{abc} = f_{abc}^*$ , which means that  $f_{abc} \in \mathbb{R}$ .  $\square$

The adjoint representation acts on the Lie algebra itself via the commutator. In particular if  $T_a^{\text{adj}}$  are the generators in the adjoint representation and  $T_d$  are the

generators in some other representation then  $T_a^{\text{adj}} \circ T_d = [T_a, T_d] = \text{if}_{ade} T_e$  and  $\text{if}_{ade} = (T_a^{\text{adj}})_{de}$ .

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  $a, b \in \mathbb{F}$

$$[ax + by, z] = a[x, z] + b[y, z], \quad \text{and} \quad (11.2.37)$$

$$[z, ax + by] = a[z, x] + b[z, y]. \quad (11.2.38)$$

- **Alternativity:** For all  $x \in \mathfrak{g}$

$$[x, x] = 0. \quad (11.2.39)$$

- The **Jacobi identity:** For all  $x, y, z \in \mathfrak{g}$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \quad (11.2.40)$$

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

$$t_i t_j - t_j t_i = \text{if}_{ijk} t_k. \quad (11.2.42)$$

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

$$C = \delta^{ab} T_a T_b. \quad (11.2.44)$$

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 $SU(2)$ and $SO(3)$

### 11.3.1 The Lie Algebras of $SU(2)$ and $SO(3)$

Recall that

$$SO(3) := \{O \in \mathcal{M}_3(\mathbb{R}) \mid O^\top O = \mathbb{1}\}, \quad (11.3.1)$$

and

$$SU(2) := \{U \in \mathcal{M}_2(\mathbb{C}) \mid U^\dagger U = \mathbb{1}\}. \quad (11.3.2)$$

The Lie algebra of  $SO(3)$ , 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  $SO(3)$  is compact and simple. One such basis is

$$T_1 = -i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad T_2 = -i \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad T_3 = -i \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (11.3.3)$$

Notice that the first  $2 \times 2$  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:

$$[T_a, T_b] = i\epsilon_{abc} T_c \quad (11.3.4)$$

where  $\epsilon_{abc}$  is the Levi-Civita symbol. Hence the structure constants are  $f_{abc} = \epsilon_{abc}$ . This shouldn't be too surprising since in three dimensions any antisymmetric tensor of rank 3 must be proportional to  $\epsilon_{abc}$ . We can easily calculate the Dynkin index from

$$\text{tr}(T_a T_b) = 2k_R \delta_{ab}, \quad (11.3.5)$$

since putting in  $a = b = 1$  we get  $\text{tr}(T_1 T_1) = 2$  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,

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (11.3.6)$$

The generators of  $\mathfrak{su}(2)$  are then given by

$$T_a := \frac{1}{2} \sigma_a. \quad (11.3.7)$$

We can easily show that  $\text{tr}(T_a T_b) = \delta_{ab} = 2k_R \delta_{ab}$  and so  $k_R = 1/4$  for this representation. The Lie algebra can be shown to satisfy

$$[T_a, T_b] = i\epsilon_{abc} T_c, \quad (11.3.8)$$

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  $\varphi: \mathbb{R}^4 \rightarrow SU(2)$  given by

$$\varphi(x_0, x_1, x_2, x_3) = x_0 \mathbb{1}_2 + x_1 i \sigma_1 + x_2 i \sigma_2 + x_3 i \sigma_3 = \begin{pmatrix} x_0 + ix_3 & ix_1 + x_2 \\ ix_1 - x_2 & x_0 - ix_3 \end{pmatrix} := U. \quad (11.3.9)$$

The defining conditions for  $SU(2)$  lead to

$$U^\dagger U = \mathbb{1}_2 \text{ and } \det U = 1 \iff x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1. \quad (11.3.10)$$

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  $\pi$ , 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  $\pi$ , and the other end  $-\pi$ , however a rotation by  $\pi$  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.

The manifold given by taking the  $n$ -sphere,  $S^n$ , and associating opposite (antipodal) points is called the **real projective space** of dimension  $n$ , and is denoted  $\mathbb{RP}^n$ . So, the manifold of  $SO(3)$  is  $\mathbb{RP}^3$ .

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  $SU(2) \rightarrow SO(3)$ .

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  $h(\mathbf{x}) = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$ . This then satisfies

$$Uh(\mathbf{x})U^\dagger = h(O\mathbf{x}) \quad (11.3.11)$$

where  $U$  is any unitary  $2 \times 2$  matrix and this equation defines  $O$  as some orthogonal  $3 \times 3$  matrix. We can use this to define a map  $U \mapsto O$ , which is a map  $SU(2) \rightarrow SO(3)$ . The kernel of this map is  $\{\pm 1\} \cong \mathbb{Z}_2$ , and so

$$SO(3) \cong SU(2)/\mathbb{Z}_2 \quad (11.3.12)$$

by the first isomorphism theorem ([Theorem 4.2.20](#)).

### 11.3.2 Irreducible Representations of $\mathfrak{su}(2)$

In this section we will construct the irreducible representations of  $\mathfrak{su}(2)$ , from which we can obtain the irreducible representations of  $\mathfrak{so}(3)$ . We will use a method known as the **highest weight representation**, which is a very general method which can be applied to any simple Lie algebra. For this process we will use a notation inspired by angular momentum in quantum mechanics<sup>3</sup>, which it should quickly become clear is intimately linked to irreducible representations of  $\mathfrak{su}(2)$ . In particular, we will denote the generators of  $\mathfrak{su}(2)$  by  $J_a$ , for  $a = 1, 2, 3$ , and we will denote vectors with bra-ket notation.

<sup>3</sup>To brush up on angular momentum in quantum mechanics see the notes from the principles of quantum mechanics course.

The process uses Casimir operators, and there is one for each rank. Since  $\mathfrak{su}(2)$  is rank 1 there is a single Casimir operator. All semi-simple Lie algebras have at least one Casimir operator given by

$$C = k^{ab}T_aT_b \quad (11.3.13)$$

where  $k^{ab}$  is the Killing metric. In the case of  $\mathfrak{su}(2)$  we have the Casimir operator

$$J^2 := J_1^2 + J_2^2 + J_3^2 \quad (11.3.14)$$

since  $SU(2)$  is simple and compact. By definition this Casimir operator commutes with all of the generators:

$$[J^2, J_a] = 0. \quad (11.3.15)$$

By Schur's lemma ([Theorem 8.4.8](#)) the Casimir operator is of the form

$$J^2 = C_2(R) \mathbb{1}_{\dim V_R} \quad (11.3.16)$$

where  $C_2(R)$  is some number depending on the representation and  $V_R$  is the vector space upon which the representation acts.

We choose  $\{J^2, J_3\}$  as our set of maximally commuting operators, the choice of  $J_3$  instead of  $J_1$  or  $J_2$ , or some linear combination of  $J_a$ , is arbitrary but conventional. Let  $|m\rangle$  be an eigenvector of  $J_3$  with eigenvalue  $m$ , that is

$$J_3|m\rangle = m|m\rangle. \quad (11.3.17)$$

Since  $J_a$  are Hermitian we know that  $m \in \mathbb{R}$ . We also know that  $\{|m\rangle\}$  forms a basis for  $V_R$ .

Now we define

$$J_{\pm} := J_1 \pm iJ_2. \quad (11.3.18)$$

Since  $J_i$  are Hermitian we immediately have that  $J_{\pm}^{\dagger} = J_{\mp}$ . Using the properties of the Lie bracket and the Lie algebra relations of  $\mathfrak{su}(2)$ , namely  $[J_a, J_b] = i\epsilon_{abc}J_c$ , we have

$$[J_3, J_{\pm}] = [J_3, J_1 \pm iJ_2] \quad (11.3.19)$$

$$= [J_3, J_1] \pm i[J_3, J_2] \quad (11.3.20)$$

$$= i\epsilon_{312}J_2 \mp \epsilon_{321}J_1 \quad (11.3.21)$$

$$= iJ_2 \pm J_1 \quad (11.3.22)$$

$$= \pm(J_1 \pm iJ_2) \quad (11.3.23)$$

$$= \pm J_{\pm}. \quad (11.3.24)$$

We also have

$$J_+J_- = (J_1 + iJ_2)(J_1 - iJ_2) \quad (11.3.25)$$

$$= J_1^2 + J_2^2 + iJ_2J_1 - iJ_1J_2 \quad (11.3.26)$$

$$= J_1^2 + J_2^2 + i[J_2, J_1] \quad (11.3.27)$$

$$= J_1^2 + J_2^2 - \epsilon_{213}J_3 \quad (11.3.28)$$

$$= J_1^2 + J_2^2 + J_3 \quad (11.3.29)$$

which we can rearrange to get  $J_1^2 + J_2^2 = J_+J_- - J_3$  which gives

$$J^2 = J_1^2 + J_2^2 + J_3^2 = J_+J_- - J_3 + J_3^2. \quad (11.3.30)$$

Similarly by considering  $J_-J_+$  we can show that

$$J^2 = J_-J_+ + J_3 + J_3^2. \quad (11.3.31)$$

Now consider what happens when we act on some basis vector with  $J_3J_{\pm}$ , we can use  $[J_3, J_{\pm}]$  to rewrite this as  $J_3J_{\pm} = [J_3, J_{\pm}] + J_{\pm}J_3 = \pm J_{\pm} \pm J_{\pm}J_3$ . Hence

$$J_3J_{\pm}|m\rangle = (\pm J_{\pm} \pm J_{\pm}J_3)|m\rangle \quad (11.3.32)$$

$$= \pm J_{\pm}|m\rangle \pm J_{\pm}J_3|m\rangle \quad (11.3.33)$$

$$= \pm J_{\pm}|m\rangle \pm mJ_{\pm}|m\rangle \quad (11.3.34)$$

$$= \pm(1 \pm m)(J_{\pm}|m\rangle) \quad (11.3.35)$$

$$= (m \pm 1)(J_{\pm}|m\rangle). \quad (11.3.36)$$

What this shows is that  $J_{\pm}|m\rangle$  is an eigenvector of  $J_3$  with eigenvalue  $m \pm 1$ . Since  $J_3$  is Hermitian its eigenvectors are orthogonal and so we know that  $|m\rangle$  is perpendicular to  $J_{\pm}|m\rangle$ . Further, so long as  $J_{\pm}|m\rangle \neq 0$  we can identify  $J_{\pm}|m\rangle = |m \pm 1\rangle$ . That is, we can use the operators  $J_{\pm}$  to produce states with eigenvalues either one higher or one lower. For this reason  $J_{\pm}$  are called **ladder operators** collectively, with  $J_+$  being a **raising operator** and  $J_-$  a **lowering operator**.

We now have the tools required to produce all finite dimensional representations. To do this we need only find the action of the generators,  $J_a$ , on the basis,  $\{|m\rangle\}$ . The assumption of a finite dimensional representation means that there are a finite number of basis vectors,  $|m\rangle$ . This means that there exists some maximal eigenvalue,

$m^*$ , at the “top of the ladder”, such that we can’t go higher. This means that using the raising operator to generate a higher eigenvalue state must fail on  $|m^*\rangle$ , which can only happen if

$$J_+|m^*\rangle = 0. \quad (11.3.37)$$

This state,  $|m^*\rangle$ , is called the highest weight state.

Now, consider what happens when we act with  $J^2$  on this state. Using [Equation \(11.3.31\)](#) we have  $J^2 = J_-J_+ + J_3 + J_3^2$  and so

$$J^2|m^*\rangle = (J_-J_+ + J_3 + J_3^2)|m^*\rangle \quad (11.3.38)$$

$$= J_-J_+|m^*\rangle + J_3|m^*\rangle + J_3^2|m^*\rangle \quad (11.3.39)$$

$$= 0 + m^*|m^*\rangle + m^{*2}|m^*\rangle \quad (11.3.40)$$

$$= m^*(m^* + 1)|m^*\rangle. \quad (11.3.41)$$

We can then use this relation for all states,  $|m\rangle$ , not just  $|m^*\rangle$ , since  $J^2$  is a Casimir operator, so commutes with  $J_- = J_1 - iJ_2$ , which means if we consider the state  $|m-n\rangle$  for  $n \in \mathbb{Z}_{>0}$  we have

$$J^2|m^*-n\rangle = J^2J_-^n|m^*\rangle = J_-^nJ^2|m^*\rangle = m^*(m^*+1)J_-^n|m^*\rangle = m^*(m^*-1)|m^*-n\rangle. \quad (11.3.42)$$

Since we assume a finite dimensional representation we must also have some  $n^* \in \mathbb{Z}_{>0}$  such that  $|m^* - n^*\rangle$  is the lowest eigenvalue state. In this case we must have that

$$J_-|m^* - n^*\rangle = 0. \quad (11.3.43)$$

Using [Equation \(11.3.30\)](#) we have  $J^2 = J_+J_- - J_3 + J_3^2$  and so

$$J^2|m^* - n^*\rangle = (J_+J_- - J_3 + J_3^2)|m^* - n^*\rangle \quad (11.3.44)$$

$$= [0 - (m^* - n^*) + (m^* - n^*)^2]|m^* - n^*\rangle \quad (11.3.45)$$

$$= (m^* - n^*)[(m^* - n^*) - 1]|m^* - n^*\rangle. \quad (11.3.46)$$

However, since  $|m^* - n^*\rangle$  is also just another state we have

$$J^2|m^* - n^*\rangle = m^*(m^* + 1)|m^* - n^*\rangle. \quad (11.3.47)$$

Combining these we have

$$(m^* - n^*)[(m^* - n^*) - 1] = m^*(m^* + 1) \implies 2m^* = n^* \in \mathbb{Z}_{>0}. \quad (11.3.48)$$

This implies that  $m^*$  can take on half-integer values. We can use the value of  $m^*$  to label the representation.

At this point we move slightly closer to quantum mechanics notation by defining  $j = m^*$  and labelling the representation explicitly for each state:  $|m\rangle = |j, m\rangle$ . The  $j$ th representation has dimension  $2j + 1$ . Since  $j = 0, 1/2, 1, 3/2, 2, \dots$  this means that the dimension is  $1, 2, 3, 4, \dots$ . The fundamental irreducible representation of  $\mathfrak{su}(2)$  is the  $j = 1/2$  representation. The eigenstates can have values of  $m$  from  $j$  to  $j - n^* = j - 2j = -j$ , so  $m = -j, -j + 1, \dots, j - 1, j$ .

Now that we have the finite dimensional irreducible representations of  $\mathfrak{su}(2)$  we can find the finite dimensional irreducible representations of  $\mathfrak{so}(3)$ . We find that the

representations are just those we get if we restrict  $j$  to be an integer,  $j = 1, 2, 3, \dots$ , which means they have dimension  $1, 3, 5, \dots$ . The fundamental representation of  $\mathfrak{so}(3)$  is the  $j = 1$  representation.

For a given representation we can compute the matrix elements of the operators  $J_3$  and  $J_\pm$ . We add a label,  $j$ , for the representation since the matrix form of an operator depends on the representation. Explicitly we have

$$(J_3^{(j)})_{mm'} = \langle j, m | J_3 | j, m' \rangle = m \langle j, m | j, m \rangle = m \delta_{mm'}, \quad (11.3.49)$$

where we have chosen to have  $|j, m\rangle$  be normalised. For  $J_\pm$  we have

$$(J_\pm^{(j)})_{mm'} = \langle j, m | J_\pm | j, m' \rangle = N_\pm(j, m) \langle j, m | j, m' \pm 1 \rangle = N_\pm(j, m) \delta_{m, m' \pm 1} \quad (11.3.50)$$

where  $N_\pm(j, m)$  is given by

$$N_\pm(j, m) = \sqrt{(j \mp m)(j \pm m + 1)} = \sqrt{j(j+1) - m(m \pm 1)}. \quad (11.3.51)$$

This normalisation factor is calculated by considering

$$\langle j, m | J_\pm^\dagger J_\pm | j, m \rangle = \langle j, m | J_\mp J_\pm | j, m \rangle \quad (11.3.52)$$

$$= \langle j, m | (J^2 - J_3^2 \mp J_3^2) | j, m \rangle \quad (11.3.53)$$

$$= j(j+1) - m^2 \mp m \quad (11.3.54)$$

$$= j(j+1) - m(m \pm 1) \quad (11.3.55)$$

which can be calculated in a different way by considering

$$\langle j, m | J_\pm^\dagger J_\pm | j, m \rangle = N_\pm(j, m) \langle j, m | J_\pm^\dagger | j, m \pm 1 \rangle \quad (11.3.56)$$

$$= N_\pm(j, m) \langle j, m \pm 1 | J_\pm | j, m \rangle^* \quad (11.3.57)$$

$$= N_\pm(j, m) N_\pm^*(j, m) \langle j, m \pm 1 | j, m \pm 1 \rangle \quad (11.3.58)$$

$$= |N_\pm(j, m)|^2. \quad (11.3.59)$$

Choosing  $N_\pm(j, m)$  to be real and positive comparing these calculations we have

$$N_\pm(j, m) = \sqrt{j(j+1) - m(m \pm 1)}. \quad (11.3.60)$$

On the level of the Lie group,  $SU(2)$ , we have that a representation of the group is given by

$$\rho_j(\alpha)_{mm'} = \langle j, m | \exp[i\alpha^a T_a] | j, m' \rangle, \quad (11.3.61)$$

which makes  $\rho_j(\alpha)$  a  $(2j+1) \times (2j+1)$  matrix over  $\mathbb{C}$ . Note that  $j$  denotes the representation,  $\alpha$  parametrises a group element, and  $mm'$  gives an entry into the  $(2j+1) \times (2j+1)$  matrix. It is this representation that is called the highest weight representation for a given value of  $j$ .

For  $SU(3)$  and  $SU(2)$  it is common to define the **Wigner D-matrix**,  $D^j(\alpha, \beta, \gamma)$  with components

$$\begin{aligned} D^j(\alpha, \beta, \gamma)_{mm'} &:= \langle j, m | \exp[-i\alpha J_3] \exp[-i\beta J_2] \exp[-i\gamma J_3] | j, m' \rangle \\ &=: \exp[-im\alpha + m'\gamma] d^j(\beta)_{mm'}, \end{aligned} \quad (11.3.62)$$

where  $d^j(\beta)$  is called the **little Wigner  $d$ -matrix**. In this case  $\alpha$ ,  $\beta$ , and  $\gamma$  parametrise the group elements instead of the more general  $\alpha$ . Notice that  $J_1$  doesn't need to appear in the expression for  $D^j(\alpha, \beta, \gamma)$  since it is implicitly generated when we combine the exponentials into one using the Baker–Campbell–Hausdorff formula through commutators of  $J_2$  and  $J_3$ . The Wigner  $D$ -matrices are widely used in physics literature since they are useful for calculating angular decay distribution amplitudes.

Finally, for a given representation the quadratic Casimir is given by  $C_2(j)\mathbb{1}$  with

$$C_2(j) = j(j+1), \quad (11.3.63)$$

and the Dynkin index is given by

$$k_j = \frac{1}{6}j(j+1)(2j+1). \quad (11.3.64)$$

### 11.3.3 Clebsch–Gordan Series of $SU(2)$

Recall that if two particles have angular momenta  $j_1$  and  $j_2$  then the combined system of both is given by the tensor product of the particles state spaces and the total angular momentum of this combined system,  $j$ , can take on values from  $|j_1 - j_2|$  to  $j_1 + j_2$  in integer steps. The reason for this is that combining the state spaces in this way corresponds to taking tensor products of the spin representations. The product in the group,  $SU(2)$ , becomes a sum on the level of the Lie algebra,  $\mathfrak{su}(2)$ , due to the exponential map. This is made rigorous with the next theorem.

**Theorem 11.3.65 — Clebsch–Gordan Series for  $\mathfrak{su}(2)$ .** The Clebsch–Gordan series for  $\mathfrak{su}(2)$ , also called the Kronecker product or tensor product of representations, is given by

$$\rho_{j_1} \otimes \rho_{j_2} = \rho_{|j_1 - j_2|} \oplus \rho_{|j_1 - j_2| + 1} \oplus \cdots \oplus \rho_{j_1 + j_2} = \bigoplus_{j=|j_1 - j_2|}^{j_1 + j_2} \rho_j \quad (11.3.66)$$

where  $\rho_j$  are irreducible representations of  $\mathfrak{su}(2)$  with Casimir elements  $j(j+1)\mathbb{1}_{2j+1}$ .

**R** Notice that the multiplicity (coefficient in the direct sum) of all irreducible representations involved is 1. If this is always the case for the tensor product of two irreducible representations then we call the group **simply reducible**.

*Proof.* Consider the Wigner  $D$ -matrix. It is sufficient for this proof to consider rotations through a single angle since we only want to compute the character of the Wigner  $D$ -matrix and any other rotation by the same amount is equivalent by a similarity transform, and hence has the same character. We therefore consider

$$D^j(\vartheta, 0, 0)_{mm'} = e^{-im\vartheta} \delta_{mm'}. \quad (11.3.67)$$

The fact that this is diagonal comes from the fact that  $|j, m'\rangle$  is an eigenstate of  $J_3$ , which is the only operator appearing in

$$D^j(\vartheta, 0, 0)_{mm'} = \langle j, m | \exp[-i\vartheta J_3] | j, m' \rangle \propto \langle j, m | j, m' \rangle = \delta_{mm'}. \quad (11.3.68)$$

The character can then be computed as the trace of  $D^j$ , recalling that the rows and columns of  $D^j$  are indexed by  $m$ , which runs from  $-j$  to  $j$ . Hence,

$$\chi_j(\vartheta) := \sum_{m=-j}^j D^j(\vartheta, 0, 0)_{mm} \quad (11.3.69)$$

$$= \sum_{m=-j}^j e^{-im\vartheta} \quad (11.3.70)$$

$$= \sum_{m=0}^{2j} e^{-im\vartheta} e^{-ij\vartheta} = e^{-ij\vartheta} \sum_{m=0}^{2j} (e^{i\vartheta})^m. \quad (11.3.71)$$

We can now identify this as a geometric series and use the identity

$$\sum_{n=0}^N r^n = \frac{1 - r^{N+1}}{1 - r}, \quad (11.3.72)$$

which gives

$$\chi_j(\vartheta) = e^{-ij\vartheta} \frac{1 - e^{i(2j+1)\vartheta}}{1 - e^{i\vartheta}} \quad (11.3.73)$$

$$= \frac{e^{-ij\vartheta} - e^{i(j+1)\vartheta}}{1 - e^{i\vartheta}} \quad (11.3.74)$$

$$= \frac{e^{-ij\vartheta} - e^{i(j+1)\vartheta}}{1 - e^{i\vartheta}} \frac{e^{-i\vartheta/2}}{e^{-i\vartheta/2}} \quad (11.3.75)$$

$$= \frac{e^{-i(j+1/2)\vartheta} - e^{i(j+1/2)\vartheta}}{e^{-i\vartheta/2} - e^{i\vartheta/2}} \quad (11.3.76)$$

$$= \frac{\sin([j + 1/2]\vartheta)}{\sin(\vartheta/2)}. \quad (11.3.77)$$

Two important cases are

$$\chi_{1/2}(\vartheta) = \frac{\sin(\vartheta)}{\sin(\vartheta/2)} = 2 \cos \frac{\vartheta}{2}, \quad (11.3.78)$$

and

$$\chi_1(\vartheta) = \frac{\sin(3\vartheta/2)}{\sin(\vartheta/2)} = 1 + 2 \cos \vartheta. \quad (11.3.79)$$

Now consider the product  $\chi_{j_1}(\vartheta)\chi_{j_2}(\vartheta) = \chi_{\rho_{j_1} \otimes \rho_{j_2}}(\vartheta)$ , applying trig identities it is possible to show that

$$\chi_{j_1}(\vartheta)\chi_{j_2}(\vartheta) = \chi_{j_1+j_2}(\vartheta) + \chi_{j_1-1/2}(\vartheta)\chi_{j_2-1/2}(\vartheta). \quad (11.3.80)$$

Assuming without loss of generality that  $j_1 \geq j_2$  we can iterate this equation  $2j_2 - 1$  times, using this product formula to first to compute  $\chi_{j_1-1/2}(\vartheta)\chi_{j_2-1/2}(\vartheta)$ , and then to compute the products  $\chi_{j_1-1}(\vartheta)\chi_{j_2-1}(\vartheta)$  which appear after this and so on we will find

$$\chi_{j_1}(\vartheta)\chi_{j_2}(\vartheta) = \chi_{j_1+j_2} + \chi_{j_1+j_2-1} + \cdots + \chi_{j_1-j_2}. \quad (11.3.81)$$

This proves the theorem holds since we can identify the left hand side as the character of  $\rho_{j_1} \otimes \rho_{j_2}$  and the right hand side as the character of  $\rho_{|j_1|-j_2} \oplus \cdots \oplus \rho_{j_1+j_2}$ .  $\square$

While the above proof ensures that the theorem is correct we can also perform a non-trivial test of the theorem just by thinking about dimensions. Recall that the dimension of a tensor product is the product of the dimensions, whereas the dimension of a direct sum is the sum of the dimensions. Since the dimension of  $\rho_j$  is  $2j+1$  in order for the theorem to hold we must have

$$(2j_1+1)(2j_2+1) = [2|j_1-j_2|+1] + [2(|j_1-j_2|+1)+1] + \cdots + [2(j_1+j_2)+1]. \quad (11.3.82)$$

In order to check this we assume without loss of generality that  $j_1 \geq j_2$ , and so the right hand side can be expressed as

$$[2(j_1-j_2)+1] + [2(j_1-j_2+1)+1] + \cdots + [2(j_1-j_2+2j_2)+1]. \quad (11.3.83)$$

This can then be written as

$$2(j_1-j_2)+1 + 2(j_1-j_2)+3 + \cdots + 2(j_1-j_2)+4j_2+1. \quad (11.3.84)$$

There are  $2j_2+1$  copies of  $2(j_1-j_2)$  in this sum and then a sum over odd integers from 1 to  $4j_2+1$ . Hence we can write the sum as

$$2(j_1-j_2)(2j_2+1) + \sum_{n=0}^{2j_2} (2n+1). \quad (11.3.85)$$

Using the identities

$$\sum_{n=0}^N 1 = N+1, \quad \text{and} \quad \sum_{n=0}^N n = \frac{N}{2}(N+1) \quad (11.3.86)$$

we can write the sum as

$$2(j_1-j_2)(2j_2+1) + 2j_2(2j_2+1) + 2j_2+1 \quad (11.3.87)$$

$$= 4j_1j_2 + 2j_1 - 4j_2^2 - 2j_2 + 4j_2^2 + 2j_2 + 2j_2 + 1 \quad (11.3.88)$$

$$= 4j_1j_2 + 2j_1 + 2j_2 + 1 \quad (11.3.89)$$

$$= (2j_1+1)(2j_2+1). \quad (11.3.90)$$

Which shows that

$$\dim(\rho_{j_1} \otimes \rho_{j_2}) = \dim \left[ \bigoplus_{J=|j_1-j_2|}^{j_1+j_2} \rho_J \right]. \quad (11.3.91)$$

Consider the result

$$\chi_j(\vartheta) = \frac{\sin([j+1/2]\vartheta)}{\sin(\vartheta/2)}. \quad (11.3.92)$$



In the limit of  $\vartheta \rightarrow 0$  we first notice that  $D^j(0, 0, 0) = \mathbb{1}_{2j+1}$  and we have

$$\lim_{\vartheta \rightarrow 0} \chi_j(\vartheta) = \lim_{\vartheta \rightarrow 0} \frac{\sin([j+1]\vartheta)}{\sin(\vartheta/2)} \quad (11.3.93)$$

$$= \lim_{\vartheta \rightarrow 0} \frac{(j+1/2) \cos([j+1/2]\vartheta)}{\cos(\vartheta/2)/2} \quad (11.3.94)$$

$$= 2j+1. \quad (11.3.95)$$

So we see that indeed the character of the identity is the dimension of the representation, which is what we expect.

**Definition 11.3.96 — Fundamental Representation** A **fundamental representation** is a representation from which all other finite dimensional irreducible representations can be constructed through tensor products.

**Corollary 11.3.97 — Fundamental Representations of  $\mathfrak{su}(2)$  and  $\mathfrak{so}(3)$**

The fundamental representation of  $\mathfrak{su}(2)$  is the  $j = 1/2$  representation. The fundamental representation of  $\mathfrak{so}(3)$  is the  $j = 1$  representation.

*Proof.* For  $\mathfrak{su}(2)$  we have  $\rho_{1/2} \otimes \rho_{1/2} = \rho_0 \oplus \rho_1$ , so we can construct  $\rho_1$ , since  $\rho_0$  is just the trivial representation. We then have  $\rho_{1/2} \otimes \rho_1 = \rho_{1/2} \otimes \rho_{1/2} \otimes \rho_{1/2} = \rho_{1/2} \oplus \rho_{3/2}$ , and so we can construct  $\rho_{3/2}$ . Carrying on like this we can construct  $\rho_{n/2}$  as  $\bigotimes_{i=1}^{n/2} \rho_{1/2}$ .

The same can be done for  $\mathfrak{so}(3)$ , but limiting to integer values of  $j$ .  $\square$

■ **Example 11.3.98** The previous theorem and its corollary are important for spin. Given two particles with angular momenta  $j_1$  and  $j_2$  each is individually described by the  $\rho_{j_i}$  representation, and the combined system is described by the  $\rho_{j_1} \otimes \rho_{j_2}$  representation. The Clebsch–Gordan series then limits the possible angular momentum values the system can take to be  $|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2$ . This also allows us to combine the orbital angular momentum and spin of a particle into the total angular momentum.

The corollary demonstrates that any system with angular momentum  $j$  can be modelled, at least in terms of angular momentum, as a system of  $j$  particles with spin  $1/2$  and no orbital angular momentum.

The next theorem completes our study of  $\mathfrak{su}(2)$ .

**Theorem 11.3.99.** The representation of  $\mathfrak{su}(2)$  as Pauli matrices is pseudo-real.

*Proof.* Consider the similarity transform<sup>a</sup>

$$-T_i^* = ST_i S^{-1}, \quad \text{where} \quad S = \sigma_2 = 2T_2. \quad (11.3.100)$$

We can show this by explicit computations. First note that  $\sigma_2^2 = \mathbb{1}$  so  $\sigma_2^{-1} = \sigma_2$ .

We then have

$$\frac{1}{2}\sigma_2\sigma_1\sigma_2^{-1} = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (11.3.101)$$

$$= \frac{1}{2}\begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (11.3.102)$$

$$= \frac{1}{2}\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \quad (11.3.103)$$

$$= -T_1. \quad (11.3.104)$$

For  $i = 2$  we simply have  $\sigma_2^3 = \sigma_2 = -\sigma_2^*$ . For  $i = 3$  we have

$$\frac{1}{2}\sigma_2\sigma_3\sigma_2^{-1} = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (11.3.105)$$

$$= \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (11.3.106)$$

$$= \frac{1}{2}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (11.3.107)$$

$$= -T_3. \quad (11.3.108)$$

This shows that this representation is equivalent to its complex conjugate, which means that this representation is not complex. However, there is no basis in which all three Pauli matrices are real and so this is a pseudo-real representation.  $\square$

---

<sup>a</sup>the minus sign is due to the physicist's convention that the exponential map is  $\exp(i\alpha T)$ .

## Part IV

# Applications to Quantum Mechanics

# Twelve

## Quantum Mechanics Synopsis

In this chapter we give a short synopsis of quantum mechanics. For more details see the notes from principles of quantum mechanics.

### 12.1 Rules of Quantum Mechanics

A state is described by a vector in some complex Hilbert space,  $\mathcal{H}$ , (Definition A.2.2). We denote states with bra-ket notation, in particular a state is denoted by a ket, such as  $|\Psi\rangle$ . This corresponds to some physical state which follows a dynamical equation. For the non-relativistic case the dynamical equation is the (time independent)

**Schrödinger equation:**

$$\hat{H}|\Psi\rangle = E|\Psi\rangle. \quad (12.1.1)$$

Here  $\hat{H}$  is an operator, called the **Hamiltonian**, and  $E$  is the energy eigenvalue associated with the eigenstate  $|\Psi\rangle$ .

The Hamiltonian, and indeed all observables, are Hermitian operators on  $\mathcal{H}$ , so  $\hat{H}^\dagger = \hat{H}$ . This is required so that the observed value,  $E$ , is real, which all eigenvalues of Hermitian operators are since

$$E\langle\Psi|\Psi\rangle = \langle\Psi|\hat{H}|\Psi\rangle = \langle\Psi|\hat{H}^\dagger|\Psi\rangle = \langle\Psi|\hat{H}|\Psi\rangle^* = E^*\langle\Psi|\Psi\rangle = E^*\langle\Psi|\Psi\rangle, \quad (12.1.2)$$

which can only hold if  $E^* = E$ , so  $E \in \mathbb{R}$ . Further, for a Hermitian operator eigenstates with different eigenvalues are orthogonal since

$$E_\Psi\langle\Phi|\Psi\rangle = \langle\Phi|\hat{H}|\Psi\rangle = \langle\Phi|\hat{H}^\dagger|\Psi\rangle = \langle\Psi|\hat{H}|\Phi\rangle = E_\Phi\langle\Psi|\Phi\rangle. \quad (12.1.3)$$

Hence,

$$(E_\Psi - E_\Phi)\langle\Psi|\Phi\rangle = 0, \quad (12.1.4)$$

so either  $E_\Psi = E_\Phi$  or  $\langle\Psi|\Phi\rangle = 0$ .

In quantum mechanics we can only determine the probability of certain processes occurring. One important case is the probability that a system transitions from state  $|\Psi\rangle$  to state  $|\Phi\rangle$ , which is given by

$$P(\Psi \rightarrow \Phi) = |A_{\Psi \rightarrow \Phi}|^2, \quad \text{where} \quad A_{\Psi \rightarrow \Phi} = \langle\Phi|\Psi\rangle. \quad (12.1.5)$$

In order for the probabilities to be properly normalised we must have  $\langle\Psi|\Psi\rangle = \langle\Phi|\Phi\rangle = 1$ .

## 12.2 From hamiltonian to the Schrödinger Equation

We can often find the Hamiltonian, up to ambiguity in operator ordering, from the classical case by replacing observables with the associated operators. The distinction being that operators don't commute. The most famous example of this being the **Heisenberg commutation relation** for the position operator,  $\hat{x}$ , and momentum operator,  $\hat{p}$ :

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}, \quad [\hat{x}_i, \hat{x}_j] = 0, \quad \text{and} \quad [\hat{p}_i, \hat{p}_j] = 0. \quad (12.2.1)$$

From these relations we can derive the Heisenberg uncertainty relation,  $\Delta x \Delta p \geq \hbar/2$ .

For a non-relativistic particle of mass  $m$  we can write the Hamiltonian as a kinetic part,  $T$ , and a potential part,  $V$ , which we assume to be a function of only the position (and not the momentum):

$$\hat{H} = \hat{T} + \hat{V} = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{x}}). \quad (12.2.2)$$

We wish to use this to perform calculations. To do so we need to choose a representation, which corresponds to choosing a maximal set of commuting operators. The most common choice is to work in the position representation, where we have as a basis states  $|\mathbf{x}\rangle$ , for which

$$\hat{x}_i|\mathbf{x}\rangle = x_i|\mathbf{x}\rangle. \quad (12.2.3)$$

Completeness of the Hilbert space ensures that we can write the identity as an integral (or sum in the discrete case) over projection operators:

$$\mathbb{1}_x = \int d^3x |\mathbf{x}\rangle\langle\mathbf{x}|. \quad (12.2.4)$$

Choosing the basis to be orthonormal we have

$$\langle\mathbf{x}|\mathbf{x}'\rangle = \delta(\mathbf{x} - \mathbf{x}'), \quad (12.2.5)$$

where  $\delta$  is the Dirac delta distribution.

Using the completeness relation,  $\mathbb{1}_x = \int d^3x |\mathbf{x}\rangle\langle\mathbf{x}|$ , we have

$$|\Psi\rangle = \int d^3x |\mathbf{x}\rangle\langle\mathbf{x}|\Psi\rangle = \int d^3x \Psi(\mathbf{x})|\mathbf{x}\rangle \quad (12.2.6)$$

where we define  $\Psi(\mathbf{x}) := \langle\mathbf{x}|\Psi\rangle$ .

In the position representation the momentum operator is given by

$$\hat{p}_j|\mathbf{x}\rangle \rightarrow -i\hbar\partial_j, \quad \text{where} \quad \partial_j := \frac{\partial}{\partial x_j}. \quad (12.2.7)$$

By this we mean

$$\langle\mathbf{x}|\hat{p}_j|\Psi\rangle = -i\hbar \frac{\partial \Psi(\mathbf{x})}{\partial x_j}. \quad (12.2.8)$$

Starting with the Schrödinger equation we act on the left with a bra from the position eigenbasis,  $\langle\mathbf{x}|$ , giving  $\langle\mathbf{x}|\hat{H}|\Psi\rangle = E\langle\mathbf{x}|\Psi\rangle = E\Psi(\mathbf{x})$ . Applying the completeness relation to the left hand side of the Schrödinger equation we get

$$\int d^3x \langle\mathbf{x}|\hat{H}|\mathbf{x}'\rangle\langle\mathbf{x}'|\Psi\rangle. \quad (12.2.9)$$

Now inserting the Hamiltonian we get two terms:

$$\int d^3x \sum_{j=1}^3 \frac{1}{2m} \langle \mathbf{x} | \hat{p}_j | \mathbf{x}' \rangle \langle \mathbf{x}' | \Psi \rangle + \int d^3x \langle \mathbf{x} | V(\hat{\mathbf{x}}) | \mathbf{x}' \rangle \langle \mathbf{x}' | \Psi \rangle. \quad (12.2.10)$$

We now use the fact that  $\langle \mathbf{x} | V(\hat{\mathbf{x}}) | \mathbf{x}' \rangle = V(\mathbf{x}') \langle \mathbf{x} | \mathbf{x}' \rangle = V(\mathbf{x}') \delta(\mathbf{x} - \mathbf{x}')$  and  $\hat{p}_j \rightarrow -i\hbar \partial_j$ ,

$$\int d^3x \sum_{j=1}^3 \frac{-\hbar^2}{2m} \delta(\mathbf{x} - \mathbf{x}') \frac{\partial^2 \psi}{\partial x_j^2} + \int d^3x V(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}') \langle \mathbf{x}' | \Psi \rangle. \quad (12.2.11)$$

Now using the sifting property of the Dirac delta distribution we have

$$\left[ -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right] \psi(\mathbf{x}) = E \psi(\mathbf{x}). \quad (12.2.12)$$

A common application is calculating transition matrices,  $A_{\Psi \rightarrow \Phi}$ . These arise when a system is described by the Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{H}_{\text{trans}}$ . If  $|\Psi\rangle$  and  $|\Phi\rangle$  are eigenstates of  $\hat{H}_0$  then we define

$$A_{\Psi \rightarrow \Phi} := \langle \Phi | \hat{H}_{\text{trans}} | \Psi \rangle + \text{higher order terms}. \quad (12.2.13)$$

This is called **Fermi's golden rule**. It gives reasonable results for sufficiently small perturbations.

## 12.3 Symmetries and Conservation Laws

**Theorem 12.3.1 — Noether's Theorem.** Each symmetry has an associated conservation law and each conservation law can be seen as the consequence of some symmetry.

Suppose we have a system governed by the Hamiltonian  $\hat{H}$ . If the system is invariant under a unitary symmetry,  $U \in U(\mathcal{H})$ , that is  $|\Psi\rangle$  and  $U|\Psi\rangle$  describe the same state, then

$$\langle \Phi | \hat{H} | \Psi \rangle = \langle \Phi | U^\dagger \hat{H} U | \Psi \rangle. \quad (12.3.2)$$

This can only be the case if  $U^\dagger \hat{H} U = \hat{H}$ , and so we see that  $[U, \hat{H}] = 0$ . That is the unitary symmetry transformation commutes with the Hamiltonian. We can write the unitary transformation in terms of a generator,  $T$ , as  $U = \exp(i\alpha T)$ . We then see that the generator must commute with the Hamiltonian,  $[T, \hat{H}] = 0$ .

Heisenberg's equation of motion for a time independent operator,  $T$ , ( $\partial_t T = 0$ ) is

$$\frac{d}{dt} \langle T \rangle = i \langle [\hat{H}, T] \rangle, \quad (12.3.3)$$

where  $\langle A \rangle := \langle \Psi | A | \Psi \rangle$  denotes the expectation value of an operator. This shows us that  $\langle T \rangle$  is a conserved quantity if it commutes with the hamiltonian.

We therefore conclude that the symmetry generated by  $T$  is linked with a conservation law for  $\langle T \rangle$ .

There are many categories of symmetries worth considering. We list some important examples now.

### ■ Example 12.3.4 — Poincaré Group



In this example we work in units where  $c = 1$ .

The **Lorentz group**,  $O(1, 3)$ , is the group preserving the metric<sup>a</sup> in special relativity, that is it preserves  $x^\mu x_\mu = t^2 - x^2 - y^2 - z^2$ , by this we mean if  $\Lambda \in O(1, 3)$  then the transformation  $x^\mu \rightarrow x'^\mu = \Lambda_{\mu\nu} x^\nu$  doesn't change  $x^\mu x_\mu$ . The **Poincaré group** is the Lorentz group and spacetime translations, it is given by the semi-direct product  $O(1, 3) \ltimes \mathbb{R}^{1,3}$ .

In the rest frame of some observer the Poincaré group reduces to the rotation group,  $SO(3)$  (or  $SU(2)$ ), spatial translations, and temporal translations. These all have associated conservation laws.

- Symmetry under rotations,  $SO(3)$ , corresponds to conservation of angular momentum.
- Symmetry under spatial translations,  $\mathbb{R}^3$ , corresponds to conservation of momentum.
- Symmetry under temporal translations,  $\mathbb{R}$ , corresponds to conservation of energy.

Note that the last two can be combined in to symmetry under spacetime translations,  $\mathbb{R}^{1,3}$ , corresponding to conservation of four-momentum,  $p^\mu = (E, \mathbf{p})$ .

When we consider different inertial frames then we get additional symmetries given by boosting between frames. All boosts can be given as a combination of boosts along the three coordinate axes. Correspondingly  $\dim(O(3)) = 3$  generalises to  $\dim(O(1, 3)) = 6$ . The Lie algebra associated with  $O(1, 3)$  is rank 2, since the generators of rotations commute with the generators of boosts.

<sup>a</sup>We use the metric sign convention  $(+---)$ .

■ **Example 12.3.5 — CPT Symmetry** There are three fundamental symmetries in particle physics that one may wish to consider. These are

- **Parity symmetry**,  $P(t, x, y, z) = (t, -x, -y, -z)$ , that is  $P\mathbf{r} = -\mathbf{r}$ . The parity is conserved by any interactions not involving the weak force.
- **Time symmetry**,  $T(t, x, y, z) = (-t, x, y, z)$ , that is  $Tt = -t$ . This is violated slightly in weak interactions, which may well be responsible for the matter-antimatter asymmetry of the universe.
- **Charge symmetry**,  $Cp = \bar{p}$ , where  $p$  is any particle and  $\bar{p}$  is its antiparticle. This is again conserved up to weak interactions.

We can combine these into one single symmetry, CPT, where we exchange all particles for their antiparticles, reflect everything through the origin and reverse time. As far as we know CPT is always conserved.

■ **Example 12.3.6 — Gauge Symmetries** Local symmetries of Lie groups lead to charge conservation. Here we are being generic with the word charge, allowing for electric charge, but also other properties like colour and isospin. At high energies the standard model exhibits a symmetry described by  $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$ . The  $c$  stands for colour charge, which is conserved under  $SU(3)$  symmetry (3 because there are three colours). The  $L$  stands for left-handed, and the  $Y$  for hypercharge. At lower energies the Higgs mechanism reduces the symmetry to  $SU(3)_c \otimes U(1)_{em}$ . Here  $U(1)_{em}$  describes conservation of electric charge.



# Thirteen

## Physics of Angular Momentum

So far we have studied  $\mathfrak{su}(2)$ . We have seen how to construct representations and the eigenvalues of the  $J_3$  generator and  $J^2$  Casimir element. We've seen that the Casimir number,  $C_2(j) = j(j+1)$ , or the value of  $j$ , can be used to label the eigenstates, of the mutual basis, in order to lift degeneracy we need a second label which we take to be  $m$ , the projection of the spin on the  $z$ -axis. The value of  $j$  is representation dependent, but frame independent, since it is associated with a Casimir element which commutes with all generators. The value of  $m$  is frame dependent since  $J_3$  is not a Casimir operator.

### 13.1 Basics of Angular Momentum in Quantum Mechanics

Recall that the **orbital angular momentum operator** is defined as

$$\hat{\mathbf{L}} := \hat{\mathbf{x}} \times \hat{\mathbf{p}}, \quad \text{or} \quad \hat{L}_a = (\hat{\mathbf{x}} \times \hat{\mathbf{p}})_a = \varepsilon_{abc} \hat{x}_b \hat{p}_c. \quad (13.1.1)$$

As well as the orbital angular momentum we have the **intrinsic angular momentum**, or **spin**, with the associated operator  $\hat{\mathbf{S}}$ . Combined, these give the **total angular momentum operator**,  $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ . If there is no spin or no orbital angular momentum it is common to simply talk about the total angular momentum,  $\hat{\mathbf{J}}$ , in place of the non-zero angular momentum operator.

Consider the commutation relation for two components of the orbital angular momentum operator. For this derivation we will need the canonical commutation relations,

$$[\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij}, \quad [\hat{x}_i, \hat{x}_j] = 0, \quad \text{and} \quad [\hat{p}_i, \hat{p}_j] = 0, \quad (13.1.2)$$

as well as the commutation identities

$$[AB, C] = A[B, C] + [A, C]B, \quad \text{and} \quad [A, BC] = B[A, C] + [A, B]C, \quad (13.1.3)$$

which can be proven by expanding the commutators. Finally we need the identity

$$\varepsilon_{ijk} \varepsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}. \quad (13.1.4)$$

Combing these we have

$$[\hat{L}_a, \hat{L}_b] = [\varepsilon_{aij} \hat{x}_i \hat{p}_k, \varepsilon_{bkl} \hat{x}_k \hat{p}_l] \quad (13.1.5)$$

$$= \varepsilon_{aij} \varepsilon_{bkl} [\hat{x}_i \hat{p}_j, \hat{x}_k \hat{p}_l] \quad (13.1.6)$$

$$= \varepsilon_{aij} \varepsilon_{bkl} (\hat{x}_i [\hat{p}_j, \hat{x}_k \hat{p}_l] + [\hat{x}_i, \hat{x}_k \hat{p}_l] \hat{p}_j) \quad (13.1.7)$$

$$= \varepsilon_{aij} \varepsilon_{bkl} (\hat{x}_i \hat{x}_j [\hat{p}_j, \hat{p}_l] + \hat{x}_i [\hat{p}_j, \hat{x}_k] \hat{p}_l + \hat{x}_k [\hat{x}_i, \hat{p}_l] \hat{p}_k + [\hat{x}_i, \hat{x}_k] \hat{p}_l \hat{p}_j) \quad (13.1.8)$$

$$= \varepsilon_{aij} \varepsilon_{bkl} (\hat{x}_i [\hat{p}_j, \hat{x}_k] \hat{p}_l + \hat{x}_k [\hat{x}_i, \hat{p}_l] \hat{p}_j) \quad (13.1.9)$$

$$= \varepsilon_{aij} \varepsilon_{bkl} (-i\hbar \delta_{jk} \hat{x}_i \hat{p}_l + i\hbar \delta_{il} \hat{x}_k \hat{p}_j) \quad (13.1.10)$$

$$= -i\hbar \varepsilon_{aij} \varepsilon_{jlb} \hat{x}_i \hat{p}_l + i\hbar \varepsilon_{jai} \varepsilon_{ibk} \hat{x}_k \hat{p}_j \quad (13.1.11)$$

$$= -i\hbar (\delta_{ai} \delta_{ib} - \delta_{ab} \delta_{il}) \hat{x}_i \hat{p}_l + i\hbar (\delta_{jb} \delta_{ak} - \delta_{jk} \delta_{ab}) \hat{x}_k \hat{p}_j \quad (13.1.12)$$

$$= -i\hbar (\hat{x}_b \hat{p}_a - \delta_{ab} \hat{x}_i \hat{p}_i) + i\hbar (\hat{x}_a \hat{p}_b - \delta_{ab} \hat{x}_j \hat{p}_j) \quad (13.1.13)$$

$$= i\hbar (\hat{x}_a \hat{p}_b - \hat{x}_b \hat{p}_a). \quad (13.1.14)$$

Now consider  $\varepsilon_{abc} \hat{L}_c$ :

$$\varepsilon_{abc} \hat{L}_c = \varepsilon_{abc} \varepsilon_{cij} \hat{x}_i \hat{p}_j \quad (13.1.15)$$

$$= (\delta_{ai} \delta_{bj} - \delta_{aj} \delta_{bi}) \hat{x}_i \hat{p}_j \quad (13.1.16)$$

$$= \hat{x}_a \hat{p}_b - \hat{x}_b \hat{p}_a. \quad (13.1.17)$$

So we see that

$$[\hat{L}_a, \hat{L}_b] = i\hbar \varepsilon_{abc} \hat{L}_c, \quad (13.1.18)$$

which means that the angular momentum operators satisfy the  $\mathfrak{su}(2)$  commutation relations, up to an unimportant factor of  $\hbar$ .

We can now ascribe a physical meaning to the value of  $j$ . Consider the Wigner  $D$ -matrix  $D^j(\vartheta, 0, 0)_{mm'} = e^{-im\vartheta} \delta_{mm'}$ . We see that

$$D^j(2\pi, 0, 0)_{mm'} = e^{-2\pi im} \delta_{mm'} = (e^{-i\pi})^{2m} = (-1)^{2m} = (-1)^{2j} \quad (13.1.19)$$

where we've used the fact that  $(-1)^{2m} = 1$  if  $m$  is an integer and  $(-1)^{2m} = -1$  if  $m$  is a half-integer, and  $m$  is an integer or half-integer precisely when  $j$  is. From this we see that if a spin 1/2 particle is rotated  $360^\circ$  its state picks up a minus sign, and another  $360^\circ$  rotation is needed in order to return to the original state. This means that to obtain an invariant scalar we need an even number of fermions so that these minus signs cancel.

For a system which is spherically symmetric the *total* angular momentum must be conserved. For a system of a single particle this means  $\mathbf{J} = \mathbf{L} + \mathbf{S}$  is conserved. For a system of two particles with angular momenta  $\mathbf{J}^{(i)}$  this means  $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$  is conserved. For a system of multiple particles their individual angular are compatible, in the sense that  $[\hat{J}_i^{(a)}, \hat{J}_j^{(b)}] = 0$  for all  $a \neq b$ . A simple example of a system with multiple sources of angular momentum is a particle with orbital angular momentum and spin 1/2, so the spin is represented by  $|s = 1/2\rangle$ . Say this particle is in an  $l = 1$  state, so the position dependent part of the state is  $|\Psi_{l=1}\rangle$ , then we can write its state as  $|\Psi_{l=1}\rangle \otimes |s = 1/2\rangle = |j = 1/2\rangle \oplus |j = 3/2\rangle$ , where we have used the Clebsch-Gordan series for  $\rho_1 \otimes \rho_{1/2} = \rho_{1/2} \oplus \rho_{3/2}$ . Notice that, despite one of the sources having angular momentum 1, there is no way the total angular momentum can be 1, it can only be 1/2 or 3/2.

## 13.2 Wigner–Eckart Theorem

**Definition 13.2.1 — Tensor Operator** Let  $G$  be a Lie group,  $\rho$  a representation of  $G$ , and  $\rho_c$  an irreducible representation of  $G$ . We say that  $T_{m_c}^c$  is a **tensor operator** in the irreducible representation,  $c$ , if it transforms as

$$\rho(g)T_{m_c}^c\rho(g)^\dagger = \rho_c(g)_{m_c m'_c}T_{m'_c}^c, \quad (13.2.2)$$

with summation over  $m'_c$  implied.

An example of a tensor operator is the momentum operator,  $\hat{p}_i$ , which transforms as  $\hat{p}_i \rightarrow R_{ij}\hat{p}_j$ , where  $R_{ij}$  is in the fundamental representation of  $\text{SO}(3)$  (that is the  $j = 1$  representation). We call objects that transform under the  $j = 1$  representation of  $\text{SO}(3)$  vectors, and hence this is sometimes referred to as the **vector representation**. Similarly the trivial representation is the scalar representation, since it leaves everything unchanged, which is how scalars transform.

The direct product of irreducible representations will, in general, decompose into a direct sum of irreducible representations. This means that, given two irreducible representations,  $a$  and  $b$ , we can decompose the tensor product as a sum over irreducible representations,  $c$ . We will use the notation of  $\mathfrak{su}(2)$  here, but this idea generalises to many other Lie algebras in a natural way. Suppose we have two states,  $|a, m_a\rangle$ , and  $|b, m_b\rangle$ . Consider their direct product,  $|a, m_a\rangle \otimes |b, m_b\rangle = |a, m_a \otimes b, m_b\rangle$ . We can insert completeness giving

$$|a, m_a\rangle \otimes |b, m_b\rangle = \sum_{c, m_c, i} |c, m_c\rangle \langle c, m_c| (|a, m_a\rangle \otimes |b, m_b\rangle), \quad (13.2.3)$$

where  $i = 1, \dots, m_{ab}^c$  is the multiplicity of the representation  $c$  in the decomposition of  $a \otimes b$ . If  $m_{ab}^c = 0, 1$  then we call the group **simply reducible**. In this case we may write

$$|a, m_a\rangle \otimes |b, m_b\rangle = \sum_{c, m_c} |c, m_c\rangle \langle c, m_c| (|a, m_a\rangle \otimes |b, m_b\rangle), \quad (13.2.4)$$

where representations with multiplicity 0 are implicitly left out of the sum. We will only consider simply reducible groups. In more terse notation we can write this as

$$|a, m_a \otimes b, m_b\rangle = \sum_{c, m_c} |c, m_c\rangle \langle c, m_c| a, m_a \otimes b, m_b\rangle. \quad (13.2.5)$$

We can identify the final bra-ket as simply a constant for fixed values of  $a, b, c, m_a, m_b$ , and  $m_c$ . This is useful since we can compute it once and then not worry about it again, giving rise to the following definition.

**Definition 13.2.6 — Clebsch–Gordan Coefficient** Let  $a$  and  $b$  be irreducible representations of some Lie group and let  $c$  index the irreducible representations in the decomposition of  $a \otimes b$ . Then we define the **Clebsch–Gordan coefficients**,  $C_{m_c m_a m_b}^{cab}$ , are defined as the coefficients appearing in the decomposition of  $|a, m_a\rangle \otimes |b, m_b\rangle$ :

$$C_{m_c m_a m_b}^{cab} := \langle c, m_c| (|a, m_a\rangle \otimes |b, m_b\rangle) = \langle c, m_c| a, m_a \otimes b, m_b\rangle. \quad (13.2.7)$$



The Clebsch–Gordan coefficients are defined only up to a phase. Therefore there is a phase convention when we give their values. The common phase convention is the **Condon–Shortley phase convention**, which is such that the Clebsch–Gordan coefficients are real for  $\mathfrak{su}(2)$ , this is what we will use.

**Theorem 13.2.8 — Wigner–Eckart Theorem.** For simply reducible groups the matrix elements of a tensor operator are proportional to a Clebsch–Gordan coefficient. That is,

$$\langle c, m_c | T_{m_a}^a | b, m_b \rangle = C_{m_c m_a m_b}^{cab} \langle c || T^a || b \rangle, \quad (13.2.9)$$

where the constant of proportionality,  $\langle c || T^a || b \rangle$ , is called the **reduced matrix element**, and is independent of  $m_a$ ,  $m_b$ , and  $m_c$ .

*Proof.* The state  $T_{m_a}^a | b, m_b \rangle$  transforms like a direct product of two states. We can see this by considering its transformation under some unitary representation,  $\rho$ :

$$\rho(g) T_{m_a}^a | b, m_b \rangle = \rho(g) T_{m_a}^a \rho(g)^\dagger \rho(g) | b, m_b \rangle \quad (13.2.10)$$

$$= \rho_a(g)_{m_a m'_a} T_{m'_a}^a \rho_b(g)_{m_b} | b, m'_b \rangle. \quad (13.2.11)$$

Here we have used  $\rho(g)^\dagger \rho(g) = \mathbb{1}$  for a unitary representation, as well as the definition of the tensor operator's transformation law, and that the state  $| j, m \rangle$  transforms under the representation  $\rho_j$ .

By the definition of Clebsch–Gordan coefficients we know that

$$| j_1, m_1 \otimes j_2, m_2 \rangle = C_{M' m_1 m_2}^{J' j_1 j_2} | J', M' \rangle, \quad (13.2.12)$$

with summation over  $J'$  and  $M'$  implied. We therefore have that

$$T_{m_a}^a | b, m_b \rangle \propto C_{m'_c m_a m_b}^{c' ab} | c, m'_c \rangle. \quad (13.2.13)$$

Since this is the only way to have  $T_{m_a}^a | b, m_b \rangle$  transform like a product state. Now apply  $\langle c, m_c |$  to this relation and we get

$$\langle c, m_c | T_{m_a}^a | b, m_b \rangle \propto C_{m'_c m_a m_b}^{c' ab} \langle c, m_c | c, m'_c \rangle \quad (13.2.14)$$

$$= C_{m'_c m_a m_b}^{c' ab} \delta_{cc'} \delta_{m_c m'_c} \quad (13.2.15)$$

$$= C_{m_c m_a m_b}^{cab}. \quad (13.2.16)$$

The constant of proportionality in this equation is what we define as the reduced matrix element,  $\langle c || T^a || b \rangle$ .  $\square$

The Wigner–Eckart theorem is useful since it allows us to compute a single matrix element,  $\langle c, m_c | T_{m_a}^a | b, m_b \rangle$ , which we can use to compute  $\langle c || T^a || b \rangle$ , and then all other matrix elements are simply proportional to this, and the constants of proportionality are known. Things get even simpler if we only consider the ratio of matrix elements, since in this case we don't even need to compute the reduced matrix element. This is useful to calculate, for example, relative decay times of particles when the particles are part of a representation of some internal symmetry group.

This occurs for the  $SU(2)$  isospin symmetry where we exchange up-type quarks and down-type quarks in pions.

Often the Clebsch–Gordan coefficients vanish. This provides **selection rules**, which restrict the possible states. In particular we must have  $|j_1 + j_2| \leq J \leq |j_1 - j_2|$  and  $M = m_1 + m_2$  for the Clebsch–Gordan coefficient to be non-zero.

To summarise this section, the matrix element  $\langle c, m_c | T_{m_a}^a | b, m_b \rangle$  is non-zero only if the irreducible representation appears in the Clebsch–Gordan series decomposition of  $a \otimes b$  with non-zero multiplicity. The direction of the matrix element is determined completely by the Clebsch–Gordan coefficient. This is useful since Clebsch–Gordan coefficients have already been computed for most representations of interest and can be readily looked up.

### 13.2.1 Using Clebsch–Gordan Coefficients

Suppose we have an electron and a photon, that is a spin  $1/2$  particle and a spin  $1$  particle. The electron spin is then described by the representation  $\rho_{1/2}$ , and the photon's spin by  $\rho_1$ . The combined electron-photon system's is then described by the tensor product representation  $\rho_1 \otimes \rho_{1/2}$ . By the addition of angular momentum theorem (Theorem 11.3.65) we know that this representation can be decomposed as a sum of  $1 - 1/2 = 1/2$  and  $1 + 1/2 = 3/2$  representations:  $\rho_1 \otimes \rho_{1/2} = \rho_{3/2} \oplus \rho_{1/2}$ . Consider the state  $|J, M\rangle$  in this representation. We can decompose this as a sum of tensor products of states in the individual particle representations, with the Clebsch–Gordan coefficients as the weights in the sum. States in the  $j_1 = 1$  representation are  $|j_1, m_1\rangle = |1, m_1\rangle$  and states in the  $j_2 = 1/2$  representation are  $|j_2, m_2\rangle = |1/2, m_2\rangle$ . As always we have that  $m_1 = -j_1, -j_1 + 1, \dots, j_1 - 1, j_1$ , which means  $m_1 = -1, 0, 1$  and similarly we have  $m_2 = -1/2, 1/2$ . The Clebsch–Gordan coefficients,  $C_{M m_1 m_2}^{J j_1 j_2}$ , are non-zero only when  $m_1 + m_2 = M$ .

For example, consider the state  $|J, M\rangle = |3/2, 1/2\rangle$ . One way to have  $m_1 + m_2 = 1/2$  we is  $m_1 = 1$  and  $m_2 = -1/2$ . This corresponds to the states  $|1, 1\rangle$  and  $|1/2, -1/2\rangle$ . The other way to have  $m_1 + m_2 = 1/2$  is to have  $m_1 = 0$  and  $m_2 = 1/2$ , which corresponds to  $|1, 0\rangle$  and  $|1/2, 1/2\rangle$ . We then find

$$|3/2, 1/2\rangle = \sqrt{\frac{3}{2}} \sqrt{\frac{1}{2}} |1, 0\rangle \otimes |1/2, 1/2\rangle + \sqrt{\frac{3}{2}} \sqrt{\frac{1}{2}} |1, 1\rangle \otimes |1/2, -1/2\rangle. \quad (13.2.17)$$

Note that we use a bar to denote a negative index, so  $\bar{\frac{1}{2}} = -1/2$ .

We can then look up the values of the Clebsch–Gordan coefficients. To do so we need to understand the standard layout of a table of Clebsch–Gordan coefficients. First we need to find the correct table, they are labelled by  $j_1$  and  $j_2$ , usually as  $j_1 \times j_2$  or  $j_1 \otimes j_2$ . The tables are then of the form

		$J$	$J$	$\dots$
		$M$	$M$	$\dots$
$m_1$	$m_2$			
$m_1$	$m_2$			
$\vdots$	$\vdots$			

(13.2.18)

The entries into the table are then almost the Clebsch–Gordan coefficients. However, since most Clebsch–Gordan coefficients tend to be square roots there is an implicit square root. Choosing the Condon–Shortly phase coefficient the Clebsch–Gordan

coefficients are also real, and so the square root misses any negatives. This means that, for example, the entry  $-1/2$  in the table is understood to mean  $-\sqrt{1/2}$ .

We are interested in the case of  $j_1 = 1/2$  and  $j_2 = 1$ . The Clebsch–Gordan table for  $1 \otimes 1/2$  is as follows:

$$\begin{array}{c|cc} & 3/2 & 1/2 \\ \hline 1 & 3/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ \hline 1 & -1/2 & 1/3 \\ 0 & 1/2 & 2/3 \\ & & -1/3 \end{array} \quad (13.2.19)$$

$$\begin{array}{c|cc} & 3/2 & 1/2 \\ \hline 0 & -1/2 & -1/2 \\ -1/2 & 2/3 & 1/3 \\ -1 & 1/3 & -2/3 \end{array} \quad \begin{array}{c|c} & 3/2 \\ \hline -1 & -3/2 \\ & 1 \end{array} \quad (13.2.20)$$

From this we can read off the Clebsch–Gordan coefficients as

$$C_{\frac{1}{2}0\frac{1}{2}}^{\frac{3}{2}1\frac{1}{2}} = \sqrt{2/3} \quad \text{and} \quad C_{\frac{1}{2}1\frac{1}{2}}^{\frac{3}{2}1\frac{1}{2}} = \sqrt{1/3}. \quad (13.2.21)$$

Hence the state  $|3/2, 1/2\rangle$  is given by

$$|3/2, 1/2\rangle = \sqrt{\frac{2}{3}}|1, 0\rangle \otimes |1/2, 1/2\rangle + \sqrt{\frac{1}{3}}|1, 1\rangle \otimes |1/2, -1/2\rangle. \quad (13.2.22)$$

As another example the state  $|1/2, 1/2\rangle$  is

$$|1/2, 1/2\rangle = C_{\frac{1}{2}0\frac{1}{2}}^{\frac{1}{2}1\frac{1}{2}}|1, 0\rangle \otimes |1/2, 1/2\rangle + C_{\frac{1}{2}1\frac{1}{2}}^{\frac{1}{2}1\frac{1}{2}}|1, 1\rangle \otimes |1/2, -1/2\rangle \quad (13.2.23)$$

$$= -\sqrt{\frac{1}{3}}|1, 0\rangle \otimes |1/2, 1/2\rangle + \sqrt{\frac{2}{3}}|1, 1\rangle \otimes |1/2, -1/2\rangle. \quad (13.2.24)$$

Of course, now days the use of look up tables is quickly vanishing and many programs exist to compute/look up Clebsch–Gordan coefficients for you.

■ **Code 13.2.25 — Clebsch–Gordan** Looking up Clebsch–Gordan coefficients with *Mathematica*.

```
1 (* ClebschGordan[{j1, m1}, {j2, m2}, {J, M}]
2 gives C_{Mm1m2}^{Jj1j2} *)
3 In[1]:= ClebschGordan[{1, 0}, {1/2, 1/2}, {3/2, 1/2}]
4 Out[1]= Sqrt[2/3]
5 In[2]:= ClebschGordan[{1, 1}, {1/2, -1/2}, {3/2, 1/2}]
6 Out[2]= 1/Sqrt[3]
```

### 13.2.2 Geometric Interpretation of the $j = 1$ Representation

We now consider the geometric interpretation of the  $j = 1$  representation. Consider the definition of a tensor operator in the  $j = 1$  representation as something which transforms according to

$$\rho(g)T_k^1\rho(g)^\dagger = \rho_c(g)_{kl}T_l^1. \quad (13.2.26)$$

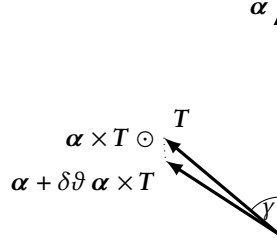


Figure 13.1: An infinitesimal rotation by  $\delta\vartheta$  of  $T$  about  $\alpha$  results in adding a term proportional to  $\alpha \times T$  to  $T$ .

In infinitesimal form this is

$$[J_a, T_k^1] = i\varepsilon_{akl}T_l^1. \quad (13.2.27)$$

For  $\mathfrak{su}(2)$  we can write  $\rho(g)$  as  $\exp[i\alpha^a J_a]$  and  $\rho(g)^\dagger$  as  $\exp[i\alpha^b J_b]$ , with implied sums over  $a$  and  $b$ . Expanding the transformation law to first order in  $\alpha$  we get

$$(\mathbb{1} + i\alpha^a J_a + O(\alpha^2))T_k^1(\mathbb{1} - i\alpha^b J_b + O(\alpha^2)) \quad (13.2.28)$$

$$= (T_k^1 + i\alpha^a J_a T_k^1 + O(\alpha^2))(\mathbb{1} - i\alpha^b J_b) \quad (13.2.29)$$

$$= T_k^1 + i\alpha^a J_a T_k^1 - \alpha^b J_b T_k^1 + O(\alpha^2) \quad (13.2.30)$$

$$= T_k^1 + i\alpha^a [J_a, T_k^1] + O(\alpha^2) \quad (13.2.31)$$

$$= T_k^1 + i\alpha^a (i\varepsilon_{akl}T_l^1) + O(\alpha^2) \quad (13.2.32)$$

$$= T_k^1 - \alpha^a \varepsilon_{akl}T_l^1 + O(\alpha^2) \quad (13.2.33)$$

$$= T_k^1 + \alpha^a \varepsilon_{alk}T_l^1 + O(\alpha^2). \quad (13.2.34)$$

Now, in a slight abuse of notation, we can define a “vector”,  $T = (T_1^1, T_2^1, T_3^1)$ , and a “vector”  $\alpha = (\alpha^1, \alpha^2, \alpha^3)$ . We can then identify

$$\varepsilon_{alk}\alpha^a T_l^1 = (\alpha \times T)_k = |\alpha||T| \sin(\gamma), \quad (13.2.35)$$

where  $\gamma$  is the “angle” between  $T$  and  $\alpha$ . We can then identify the first order term in the tensor transformation law as a small translation in the direction “perpendicular” to  $T$  and  $\alpha$ , which is then added onto the original  $T_k^1$ . This is simply the infinitesimal version of a rotation of  $T$  about  $\alpha$ . This is shown in [Figure 13.1](#).

# Fourteen

## Further Applications

### 14.1 Polynomial Invariants

**R** This section is nonexaminable.

In physics we are very interested in invariant quantities, such as energy, momentum, mass, etc. These are all invariant under some particular symmetry and this fact can greatly simplify physics problems. One question that we might ask is given a number of irreducible representations how many invariant quantities can we construct? The simplest example of an invariant is the inner product. This is an invariant of a vector representation and its complex conjugate.

Suppose we have an irreducible representation, which we denote by its dimension,  $X$ . Consider the representation  $X \otimes \bar{X}$ . We know by [Corollary 9.7.30](#) that the trivial representation appears exactly once in the Clebsch–Gordan series, that is:

$$X \otimes \bar{X} = 1 \cdot 1 \oplus \cdots . \quad (14.1.1)$$

The inner product corresponds to the trivial representation,  $1$ , in this decomposition.

For a given group,  $G$ , we can define an invariant,  $\mathcal{I}$ , as an  $n$ -multilinear operator from the space of irreducible representations to  $\mathbb{C}$ , such that for a given group element,  $g \in G$ ,  $\mathcal{I}$  is unchanged if we act on each representation with itself. That is

$$\mathcal{I}_n = \mathcal{I}(X_1, \dots, X_n) = \mathcal{I}(\rho_1(g)X_1, \dots, \rho_n(g)X_n) \quad (14.1.2)$$

for all  $g \in G$ .

The inner product then corresponds to  $\mathcal{I}_2$ , it takes two values in and returns an invariant. More explicitly as long as  $\dim X_a = \dim X_b$  we define the inner product as the invariant

$$\mathcal{I}_2(X_a, X_b) := \sum_{i=1}^{\dim X_a} (X_a)_i^* (X_b)_i. \quad (14.1.3)$$

If  $\dim X_a \neq \dim X_b$  then no degree two invariant can be formed. Note that  $(X_a)_i$  here are the components of some element in the  $X_a$  representation. The basis we choose isn't important since by definition of being an invariant this quantity is the same no matter what basis we choose.

One way to find higher order invariants is to take Kronecker products and count how many times  $1$  appears. Each occurrence leads to an invariant. The specific representation giving rise to  $1$  corresponds to the degree of the associated invariant.



As an example we will consider the group  $G = S_4$ , and limit ourselves to irreducible representations of dimension 3 or fewer. The following Kronecker products can be computed:

$$3_1 \otimes 3_1 = \text{Sym}(1 \oplus 2 \oplus 3_1) \oplus \text{Asym}(3_2), \quad (14.1.4)$$

$$3_1 \otimes 2 = 3_1 \oplus 3_2, \quad (14.1.5)$$

$$3_1 \otimes 3_2 = 1_1 \oplus 2 \oplus 3_1 \oplus 3_2. \quad (14.1.6)$$

We restrict ourselves to considering invariants involving  $3_1$ . We can take  $(x, y, z)$  as a generic element of this vector space associated with this representation.

The antisymmetric part vanishes since it corresponds to  $\varepsilon_{ijk}x_ix_j = 0$ . The invariants of degree at most three are

$$\mathcal{I}_2 = x^2 + y^2 + z^2, \quad \text{and} \quad \mathcal{I}_3 = xyz. \quad (14.1.7)$$

We have seen already that  $\mathcal{I}_2$  corresponds to the inner product. That this also appears for  $S_4$  is simply due to the fact that  $S_4$  is a subgroup of  $O(3)$ . The invariant of degree three comes from considering  $3_1 \otimes 3_1 \otimes 3_1 = 1 \cdot 1 \oplus \dots$ .

Invariants are basis independent, as they must be to be useful. This follows since a similarity transformation on the representation will always affect the invariants in a definite way and we can construct them such that there is no change.

A more systematic way of obtaining the degrees and interdependencies of invariants in a given representation uses the **Molien function**, which is beyond the scope of this course.

For a given degree we can construct an invariant by making a generic ansatz and averaging over the group by the **Reynolds operator**. For example, consider the representation containing the vector  $(x, y, z)$ , again restricting ourselves to three dimensions. Now consider any polynomial function,  $f$ , of  $x$ ,  $y$ , and  $z$ , that is  $f \in \mathbb{C}[x, y, z]$ . The following is then an invariant:

$$\mathcal{I}(x, y, z) = \frac{1}{|G|} \sum_{g \in G} f(\rho(g) \circ x, \rho(g) \circ y, \rho(g) \circ z), \quad (14.1.8)$$

where we understand  $\rho(g) \circ x$  as  $\rho(g)\mathbf{e}_x$ , where  $\mathbf{e}_x$  is a unit vector in the  $x$ -direction.

This operation corresponds to projection operators and will always give rise to an invariant. However, most of the time it will lead to 0, which is trivially invariant, unless the degree of  $f$  matches one of the degrees predicted by either the Kronecker product method or the Molien function. As an example consider  $f(x, y, z) = x^2$ . From this we derive the invariant  $\mathcal{I}(x, y, z) \propto x^2 + y^2 + z^2 = \mathcal{I}_2$ , with some unimportant constant of proportionality.

One can think of the Clebsch–Gordan coefficient appearing in

$$\langle c, m_c | T_{m_a}^a | b, m_b \rangle = C_{m_c m_a m_b}^{cab} \langle c || T^a || b \rangle \quad (14.1.9)$$

as an invariant of degree three of type  $\mathcal{I}(\bar{c}, \mathbf{a}, \mathbf{b})$ .

## 14.2 Landé $g$ -Factor



In this section we choose units such that  $\hbar = 1$ .

Consider a Hydrogen atom with a Coulomb potential. This system has the Hamiltonian  $H = H_0$ , where  $H_0 = H_{\text{Coulomb}} + H_{\text{spin-orbit}}$ , where  $H_{\text{Coulomb}}$  is the kinetic energy term and the Coulomb potential term, and  $H_{\text{spin-orbit}}$  is an extra term corresponding to spin-orbit coupling, the details of which are not important. This system exhibits spherical symmetry. We can break this symmetry by applying an external magnetic field. In this case we get a modified Hamiltonian  $H = H_0 + H'$ , where  $H' = -\boldsymbol{\mu} \cdot \mathbf{B}$  where  $\mathbf{B}$  is the external magnetic field and  $\boldsymbol{\mu}$  is the dipole moment given by

$$\boldsymbol{\mu} = \frac{e}{2\pi m_e c} (\mathbf{J} + \mathbf{S}) = \frac{e}{2\pi m_e c} (\mathbf{L} + 2\mathbf{S}). \quad (14.2.1)$$

Defining  $x = -eB/(2m_e c)$  we can write the perturbing Hamiltonian as  $H' = x(L_3 + 2S_3) = x(J_3 + S_3)$ , choosing the 3-axis to be along the direction of the magnetic field.

We assume that the energy eigenstates,  $|n, l, j, m\rangle$ , of  $H_0$  are known and treat  $H'$  as a perturbation. The first order perturbation correction to the energy gives an extra term of the form

$$E' = \langle n, l, j, m | H' | n, l, j, m \rangle = x(m + \langle S_3 \rangle). \quad (14.2.2)$$

Here we have used  $J_3 |n, l, j, m\rangle = m |n, l, j, m\rangle$ .

Using the Wigner-Eckart theorem, and a condensed notation,  $|m\rangle = |n, l, j, m\rangle$ , we have

$$\langle m | S_i | m' \rangle = C_{imm'}^{1jj'} \langle \|S_1\| \rangle, \quad \text{and} \quad \langle m | J_i | m' \rangle = C_{imm'}^{1jj'} \langle \|J_1\| \rangle. \quad (14.2.3)$$

Note that  $\langle \|S_1\| \rangle$  and  $\langle \|J_1\| \rangle$  are the reduced matrix elements in this condensed notation suppressing  $j$ . What these relations tell us is that  $\langle m | S_i | m' \rangle = a \langle m | J_i | m' \rangle$  with

$$a := \frac{\langle \|S_1\| \rangle}{\langle \|J_1\| \rangle}. \quad (14.2.4)$$

Importantly  $a$  is independent of the values of  $m$  and  $m'$ .

We have

$$\langle m | J^2 | m \rangle = \sum_{m', i} \langle m | J_i | m' \rangle \langle m' | J_i | m \rangle, \quad (14.2.5)$$

where we have used  $J^2 = \sum_i J_i J_i$  and inserted the completeness relation,

$$\mathbb{1} = \sum_{m'} |m'\rangle \langle m'|. \quad (14.2.6)$$

We similarly have

$$\langle m | \mathbf{J} \cdot \mathbf{S} | m \rangle = \sum_{m', i} \langle m | J_i | m' \rangle \langle m' | S_i | m \rangle \quad (14.2.7)$$

$$= \sum_{m', i} \langle m | J_i | m' \rangle a \langle m' | J_i | m \rangle \quad (14.2.8)$$

$$= a \langle m | J^2 | m \rangle \quad (14.2.9)$$

$$= a j(j+1). \quad (14.2.10)$$

Alternatively, using  $L = J - S$  we have

$$L^2 = (J - S)^2 = J^2 + S^2 - 2J \cdot S. \quad (14.2.11)$$

Note that here we have used the fact that  $J$  and  $S$  are both from two different copies of the  $SU(2)$  Lie algebra, and hence commute with each other, we can view both as elements of the universal enveloping algebra. Hence we find that

$$\langle m | J \cdot S | m \rangle = -\frac{1}{2} [l(l+1) - j(j+1) - s(s+1)]. \quad (14.2.12)$$

We can compare these two results and solve for  $a$ . Doing so we find that

$$a = \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}, \quad (14.2.13)$$

which is, as expected, independent of  $m$  and  $m'$ . We find that

$$E'_{jls m} = -\frac{emB}{2m_e c} g \quad (14.2.14)$$

where

$$g = 1 + a \quad (14.2.15)$$

is the **Landé  $g$ -factor** to first order in perturbation theory. This energy shift,  $E'$ , can be added to the energy of the unperturbed Hamiltonian to lift the degeneracy.

### 14.3 Spherical Harmonics

The trace is linear, meaning  $\text{tr}(A+B) = \text{tr}(A) + \text{tr}(B)$ . For finite matrices we also have the cyclic property, which means in particular that  $\text{tr}(AB) = \text{tr}(BA)$ . This implies that for finite matrices  $\text{tr}([A, B]) = 0$ . This cyclic property doesn't hold for operators on infinite dimensional spaces. The canonical commutation relation,  $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$ , has trace  $3i\hbar \neq 0$  in three spatial dimensions. This then implies that the position space representation is infinite dimensional. This shouldn't come as a surprise since the position basis is continuously labelled by  $x$ .

The question then is how come the representations associated with the angular momentum are finite dimensional? These representations are on vector spaces spanned by  $|l, m\rangle$  with  $l < \infty$  and  $m = -l, \dots, l$ , making up a finite basis. The answer is that we can decompose the Hilbert space on the sphere,  $S^2$ , into an infinite sum of finite dimensional Hilbert spaces. We will now investigate a basis for each of these finite dimensional Hilbert spaces.

Consider  $S^2$  parametrised by the usual spherical coordinates  $(\vartheta, \varphi)$ . Note that we keep  $r = 1$  fixed and so this is a two-dimensional manifold embedded in 3-dimensional space. A state  $|\vartheta, \varphi\rangle$  obeys the completeness and orthogonality relations

$$1_{S^2} = \int d\Omega |\vartheta, \varphi\rangle \langle \vartheta, \varphi|, \quad (14.3.1)$$

$$\langle \vartheta, \varphi | \vartheta', \varphi' \rangle = \delta(\varphi - \varphi') \delta(\cos \vartheta - \cos \vartheta'). \quad (14.3.2)$$

Note that  $d\Omega = \sin \vartheta d\varphi d\vartheta$  is a solid angle element, so

$$\int_{S^2} d\Omega = \int_0^{2\pi} d\varphi \int_0^\pi d\vartheta \sin \vartheta = \int_0^{2\pi} d\varphi \int_{-1}^1 d(\cos \vartheta). \quad (14.3.3)$$

We can decompose the Hilbert space  $\mathcal{H}_{S^2} = \text{span}\{|\vartheta, \varphi\rangle\}$ , as a direct sum of the Hilbert spaces  $\mathcal{H}_l = \text{span}\{|l, m\rangle \mid m = -l, \dots, l\}$ . Clearly

$$\dim \mathcal{H}_l = |\text{span}\{|l, m\rangle \mid m = -l, \dots, l\}| = | \{-l, \dots, l\} | = 2l + 1. \quad (14.3.4)$$

We then have

$$\mathcal{H}_{S^2} = \bigoplus_{l=0}^{\infty} \mathcal{H}_l. \quad (14.3.5)$$

We can rewrite the completeness and orthogonality relations for  $\mathcal{H}_{S^2}$  in terms of its decomposition, giving

$$\mathbb{1}_{S^2} = \sum_{l=0}^{\infty} \sum_{m=-l}^l |l, m\rangle \langle l, m|, \quad (14.3.6)$$

$$\langle l, m | l', m' \rangle = \delta_{ll'} \delta_{mm'}. \quad (14.3.7)$$

**Definition 14.3.8 — Spherical Harmonic** We define the **spherical harmonics**,  $Y_{lm}: S^2 \rightarrow \mathbb{C}$ , as the projection of the state  $|l, m\rangle$  onto  $|\vartheta, \varphi\rangle$ :

$$Y_{lm}(\vartheta, \varphi) := \langle \vartheta, \varphi | l, m \rangle. \quad (14.3.9)$$

**Lemma 14.3.10** The following orthogonality relations hold for the spherical harmonics:

$$\int Y_{l'm'}^*(\vartheta, \varphi) Y_{lm}(\vartheta, \varphi) d\Omega = \delta_{ll'} \delta_{mm'}, \quad (14.3.11)$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}(\vartheta, \varphi) Y_{lm}^*(\vartheta', \varphi') = \delta(\varphi - \varphi') \delta(\cos \vartheta - \cos \vartheta'). \quad (14.3.12)$$

*Proof.* We start with the orthogonality relation for  $|l, m\rangle$ :

$$\langle l, m | l', m' \rangle = \delta_{ll'} \delta_{mm'}. \quad (14.3.13)$$

We then insert the completeness relation for  $|\vartheta, \varphi\rangle$ :

$$\mathbb{1}_{S^2} = \int d\Omega |\vartheta, \varphi\rangle \langle \vartheta, \varphi|. \quad (14.3.14)$$

This gives

$$\int d\Omega \langle l, m | \vartheta, \varphi \rangle \langle \vartheta, \varphi | l', m' \rangle = \delta_{ll'} \delta_{mm'}. \quad (14.3.15)$$

Now we use  $\langle l, m | \vartheta, \varphi \rangle = \langle \vartheta, \varphi | l, m \rangle^*$  and the definition of the spherical harmonics as  $Y_{lm}(\vartheta, \varphi) := \langle \vartheta, \varphi | l, m \rangle$  to get

$$\int d\Omega Y_{lm}^*(\vartheta, \varphi) Y_{l'm'}(\vartheta, \varphi) = \delta_{ll'} \delta_{mm'}. \quad (14.3.16)$$

Exchanging which symbols are primed we get the desired result. Next we start with the orthogonality relation for  $|\vartheta, \varphi\rangle$ :

$$\langle \vartheta, \varphi | \vartheta', \varphi' \rangle = \delta(\varphi - \varphi') \delta(\cos \vartheta - \cos \vartheta'). \quad (14.3.17)$$

We then insert the completeness relation for  $|l, m\rangle$ :

$$\mathbb{1}_{S_2} = \sum_{l=0}^{\infty} \sum_{m=-l}^l |l, m\rangle \langle l, m|. \quad (14.3.18)$$

This gives

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l \langle \vartheta, \varphi | l, m \rangle \langle l, m | \vartheta', \varphi' \rangle = \delta(\varphi - \varphi') \delta(\cos \vartheta - \cos \vartheta'). \quad (14.3.19)$$

Again using  $\langle l, m | \vartheta', \varphi' \rangle = \langle \vartheta', \varphi' | l, m \rangle^*$  and the definition of the spherical harmonics as  $Y_{lm}(\vartheta, \varphi) := \langle \vartheta, \varphi | l, m \rangle$  we get

$$\sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}(\vartheta, \varphi) Y_{lm}^*(\vartheta', \varphi') = \delta(\varphi - \varphi') \delta(\cos \vartheta - \cos \vartheta'). \quad (14.3.20)$$

□

For  $l = 0, 1$ , using the Condon–Shortley phase convention, the spherical harmonics are given by

$$Y_{00}(\vartheta, \varphi) = \frac{1}{2} \sqrt{\frac{1}{\pi}}, \quad (14.3.21)$$

$$Y_{10}(\vartheta, \varphi) = \frac{1}{2} \sqrt{\frac{3}{\pi}} \cos \vartheta = \frac{1}{2} \sqrt{\frac{3}{\pi}} \frac{z}{r}, \quad (14.3.22)$$

$$Y_{1,-1}(\vartheta, \varphi) = \frac{1}{2} \sqrt{\frac{3}{2\pi}} e^{-i\varphi} \sin \vartheta = \frac{1}{2} \sqrt{\frac{3}{2\pi}} \frac{x - iy}{r}, \quad (14.3.23)$$

$$Y_{11}(\vartheta, \varphi) = -\frac{1}{2} \sqrt{\frac{3}{2\pi}} e^{i\varphi} \sin \vartheta = -\frac{1}{2} \sqrt{\frac{3}{2\pi}} \frac{x + iy}{r}. \quad (14.3.24)$$

In general the  $\varphi$  dependence of  $Y_{lm}$  only appears through a  $e^{im\varphi}$  term and the  $\vartheta$  dependence appears through trigonometric terms  $\sin \vartheta$  and  $\cos \vartheta$ , which occur with maximum power  $l$ . This follows from the fact that we can write the spherical harmonics as

$$Y_{lm}(\vartheta, \varphi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_{lm}(\cos \vartheta) e^{im\varphi}. \quad (14.3.25)$$

Here  $P_{lm}$  are the **associated Legendre polynomials**, which are polynomials of order  $l$ . A useful property of the spherical harmonics is that

$$Y_{l,-m} = (-1)^m Y_{lm}^*. \quad (14.3.26)$$

The quadratic Casimir operator,  $L^2$ , on  $S^2$  is given by

$$\int d\Omega \langle \vartheta', \varphi' | \left( -\frac{L^2}{\hbar^2} \right) | \vartheta, \varphi \rangle = \frac{1}{\sin^2 \vartheta} \frac{\partial^2}{\partial \varphi^2} + \frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left( \sin \vartheta \frac{\partial}{\partial \vartheta} \right). \quad (14.3.27)$$

This corresponds to the Laplacian,  $\nabla_{S^2}^2$ , on  $S^2$ . We therefore have that

$$\nabla_{S^2}^2 Y_{lm} = -\frac{L^2}{\hbar^2} Y_{lm} = -l(l+1)Y_{lm}. \quad (14.3.28)$$

Which follows since we know that the action of the quadratic Casimir,  $L^2$ , for  $SU(2)$  is to multiply by a factor of  $l(l+1)$ . It makes sense that this should be the Casimir. First, the Laplacian is invariant under rotations, and it is of second order, as a quadratic Casimir must be. Solutions to Laplace's equation are called harmonic functions, which explains the name "spherical harmonics", they are harmonic functions on the sphere.

### 14.3.1 Application to $\mathbb{R}^3 \cong \mathbb{R}^+ \otimes S^2$

We can decompose three-dimensional Euclidean space,  $\mathbb{R}^3$ , into a product of the non-negative real numbers,  $\mathbb{R}^+ := \{r \in \mathbb{R} \mid r \geq 0\}$ , and the sphere,  $S^2$ :

$$\mathbb{R}^3 \cong \mathbb{R}^+ \otimes S^2. \quad (14.3.29)$$

This is most obvious when working in spherical polar coordinates, where we associate the radial coordinate  $r$  with  $\mathbb{R}^+$  and the angular components,  $(\vartheta, \varphi)$ , with  $S^2$ . Since the spherical harmonics are a complete set of functions on  $S^2$  we can express any<sup>1</sup> function,  $f: \mathbb{R}^3 \rightarrow \mathbb{C}$  as

<sup>1</sup>Well, any function of interest in physics.

$$f(x, y, z) = \sum_{n=1}^{\infty} \sum_{l=0}^{n-1} \sum_{m=-l}^l u_{nlm}(r) Y_{lm}(\vartheta, \varphi) \quad (14.3.30)$$

where  $u_{nlm}: \mathbb{R}^+ \rightarrow \mathbb{C}$  and  $Y_{lm}: S^2 \rightarrow \mathbb{C}$  are the spherical harmonics.

The Laplacian in spherical coordinates is

$$\nabla_{\mathbb{R}^3}^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \nabla_{S^2}^2. \quad (14.3.31)$$

This is used, for example, to express the Schrödinger equation in spherical coordinates.

**Definition 14.3.32 — Solid Harmonics** The **solid harmonics** are the functions  $\mathcal{Y}_{lm}: \mathbb{R}^3 \rightarrow \mathbb{C}$  defined by

$$\mathcal{Y}_{lm}(r, \vartheta, \varphi) := r^l Y_{lm}(\vartheta, \varphi). \quad (14.3.33)$$

The solid harmonics are solutions to Laplace's equation in  $\mathbb{R}^3$ , that is

$$\nabla_{\mathbb{R}^3}^2 \mathcal{Y}_{lm} = 0. \quad (14.3.34)$$

These are particularly useful when we have a spherically symmetric potential, which we can write as  $V(r)$ , that is as a function of only the distance from the origin, with no angular dependence. In this case the solution doesn't depend on the magnetic quantum number,  $m$ , since  $m$  singles out a direction in space, defining the  $L_3$  operator as the angular momentum along a particular axis. We therefore reduce  $u_{nlm}$  to  $u_{nl}$  and the Schrödinger equation in three dimensions reduces to a Schrödinger equation in one dimension for the radial component:

$$\left( -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + V_l^{\text{eff}}(r) \right) u_{nl}(r) = E_{nl} u_{nl}. \quad (14.3.35)$$

Here  $V_l^{\text{eff}}$  is the effective potential

$$V_l^{\text{eff}}(r) := V(r) + \frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}. \quad (14.3.36)$$

Note that we are now restricting  $u_{nl}$  to be eigenfunctions of the Schrödinger equation, rather than an arbitrary complete set of functions on  $\mathbb{R}^+$  as they originally were.

### 14.3.2 Why $S^2$ ?

Why is it that  $S^2$  appears as the manifold of interest here. We are considering  $\text{SO}(3)$  symmetry, and the manifold of  $\text{SO}(3)$  is  $\mathbb{RP}^3$ , that is  $S^3$  with antipodal points associated. So, why is this not the manifold that we are considering?

Consider an arbitrary vector in  $\mathbb{R}^3$ . We may as well define a coordinate system such that this arbitrary vector is  $v = (1, 0, 0)$ . We can act on this vector using  $\text{SO}(2)$  to act on the  $y$  and  $z$  components, that is we act on it with  $\mathbb{1}_1 \otimes R$  with  $R \in \text{SO}(2)$ . So

$$\mathbb{1}_1 \otimes R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \vartheta & \sin \vartheta \\ 0 & -\sin \vartheta & \cos \vartheta \end{pmatrix}. \quad (14.3.37)$$

The interpretation of this is of course as a rotation about  $v = (1, 0, 0)$ .

The fact that our arbitrary vector is invariant under this operation means we have some redundancy, and to get rid of it we mod out by  $\text{SO}(2)$ . Doing so the relevant manifold is

$$\text{SO}(3)/\text{SO}(2) \cong S^2. \quad (14.3.38)$$

Note that this is *not* a group manifold since  $\text{SO}(2)$  is *not* normal in  $\text{SO}(3)$ .

# Fifteen

## Selection Rules

### 15.1 Parity Selection

Recall that we defined **parity symmetry** as

$$P \circ (t, x, y, z) = (t, -x, -y, -z), \quad P \circ \mathbf{r} = -\mathbf{r}. \quad (15.1.1)$$

Until 1956 it was believed that parity was strictly conserved in all interactions. In 1956 theorists posited that parity conservation may be violated in weak interactions. Experiments and further development of the theory found that parity is in fact maximally violated in weak interactions. The weak force ignores so called “right handed particles” and acts only on “left handed particles”. For this reason many attach an L subscript to the symmetry group of the weak force,  $SU(2)_L$ . In practice the weak force is, well, weak, and can be neglected in many strong and electromagnetic interactions, so the parity is still usually a good quantum number.

By allowing parity violation we effectively extend the symmetry group from  $SO(3)$  to  $O(3)$ , since we can write elements of  $O(3)$  as a reflection times an element of  $SO(3)$ , with the reflection corresponding to swapping the parity. We implement the parity transformation as a unitary transformation,  $U(P)$ . Further, reversing parity and then reversing back cannot change the physics, and hence we must have that  $U(P)^2 = e^{i\eta} \mathbb{1}$ , so that upon doubly reversing the parity the only effect is a phase change. Often we are free to choose this phase factor, so in which case we usually choose  $\eta = 0$ , corresponding to  $U(P) = \pm \mathbb{1}$ .

Given a state  $|l, m\rangle$  the parity transformation is

$$U(P)|l, m\rangle = (-1)^l |l, m\rangle \quad (15.1.2)$$

for  $l \in \mathbb{N}$ . This can be inferred from the parity transformation properties of the spherical harmonics:

$$P \circ Y_{lm}(\vartheta, \varphi) = Y_{lm}(\pi - \vartheta, \varphi + \pi) = (-1)^l Y_{lm}(\vartheta, \varphi). \quad (15.1.3)$$

This can be seen by either thinking about the coordinates  $(r, \vartheta, \varphi)$  being inverted through the origin, or algebraically by changing  $x \rightarrow -x$ ,  $y \rightarrow -y$ , and  $z \rightarrow -z$  in the equations  $\vartheta = \arctan(\sqrt{x^2 + y^2}/z)$  and  $\varphi = \arctan(y/x)$ .

Suppose we have a transition governed by the operator  $X$ . The parity of the operator is  $\eta_X$ , which is given by  $U(P)XU(P)^\dagger = (-1)^{\eta_X} X$ , or equivalently  $X = (-1)^{\eta_X} U(P)^\dagger X U(P)$ . We then have the selection rule that

$$\langle l, m | X | l', m' \rangle \neq 0 \quad (15.1.4)$$



only if  $l + l' + \eta_X$  is even, since we have

$$\langle l, m | X | l', m' \rangle = \langle l, m | (-1)^{\eta_X} U(P)^\dagger X U(P) | l', m' \rangle \quad (15.1.5)$$

$$= (-1)^{l+l'+\eta_X} \langle l, m | X | l', m' \rangle. \quad (15.1.6)$$

Hence we have  $(-1)^{l+l'+\eta_X} = 1$ , which implies  $l + l' + \eta_X$  is even. Note that if  $X$  is a tensor operator we also have the selection rules that  $|j_1 - j_2| \leq J \leq j_1 + j_2$ , and  $M = m_1 + m_2$ .

Common examples of operators with definite parity are the position,  $\mathbf{x}$ , momentum,  $\mathbf{p}$ , angular momentum,  $\mathbf{L}$ , electric field,  $\mathbf{E}$ , and magnetic field,  $\mathbf{B}$ . Of these  $\mathbf{x}$ ,  $\mathbf{p}$ ,  $\mathbf{E}$  are proper vectors, so have parity  $\eta_x = \eta_p = \eta_E = 1$ , whereas  $\mathbf{L}$  and  $\mathbf{B}$  are **pseudo-vectors**, meaning they change by a sign under a parity transformation, and hence  $\eta_L = \eta_B = 0$ .

It is easy enough to show the pseudo-vector nature of  $\mathbf{L}$ , since we know that  $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ , so

$$P \circ \mathbf{L} = P \circ (\mathbf{x} \times \mathbf{p}) = (P \circ \mathbf{x}) \times (P \circ \mathbf{p}) = (-\mathbf{x}) \times (-\mathbf{p}) = \mathbf{x} \times \mathbf{p} = \mathbf{L}. \quad (15.1.7)$$

In fact the cross product of two vectors is always a pseudo-vector. Similarly the cross product of two pseudo-vectors is a pseudo-vector, and the cross product of a pseudo-vector and vector is a vector. This is the logic that means  $\mathbf{B}$  is a pseudo vector since we have that  $\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$ , and since  $\mathbf{F}$  is a proper vector and so is  $\mathbf{v}$  we must have that  $\mathbf{B}$  is a pseudo-vector.

Each particle has an intrinsic parity. Bosons have the same intrinsic parity as their antiparticles, whereas fermions have the opposite parity to their antiparticles. This is a result of causality in QFT. The angular momentum of a particle is a quantum number,  $J$ . It is also possible for the intrinsic parity to have this pseudo property. If we are allowing for this then we label the particles with the quantum number  $J^\pm$ , with  $-$  for when  $\eta$  is odd and  $+$  when  $\eta$  is even. For example, the  $\rho$ -meson ( $u\bar{d}$ ) is a vector meson, meaning  $S = J = 1$ , and the  $A_1$ -meson ( $\bar{u}d$ ) is also a vector meson. However, the  $\rho$ -meson has negative parity, so is  $1^-$ , whereas the  $A_1$ -meson has positive parity, so is  $1^+$ . We say that the  $A_1$ -meson is a pseudo-vector particle.

When we are only interested in computing transitions between like particles, for example, an electron going from one energy level to another, the parity cancels from both sides of the equation. For this reason we often ignore parity when we are doing quantum mechanics. When we are doing particle physics, such as QFT, however, we often consider cases where we generate new particles, potentially with different parities. A general example would be the decay  $A \rightarrow B + C$ . In this case the parity of the final state is the sum of the intrinsic parities and the angular momentum,  $\eta_{\text{tot}} = \eta_B + \eta_C + \eta_L$ . If  $A$  has definite parity and the interaction conserves parity (i.e. it is not a weak interaction) then we must have that  $(-1)^{\eta_A} = (-1)^{\eta_{\text{tot}}}$ .

## 15.2 Superselection

Superselection rules generalise selection rules. Selection rules refer to a particular Hamiltonian which is such that only certain transitions between states occur with non-zero probability. Superselection rules state that certain matrix elements vanish for *all* observables. That is given the states  $|\psi\rangle$  and  $|\phi\rangle$  we say these states are separated by a **superselection rule** if

$$\langle \psi | O | \phi \rangle = 0 \quad (15.2.1)$$

for all observables  $O$ .

This means that a relative phase between  $|\psi\rangle$  and  $|\varphi\rangle$  is unobservable, and hence there is no coherent state of the form  $a|\psi\rangle + b|\varphi\rangle$ .

Notice that we can always define an Hermitian operator  $X := |\psi\rangle\langle\varphi| + |\varphi\rangle\langle\psi|$ , and clearly

$$\langle\psi|X|\varphi\rangle = \langle\psi|\psi\rangle\langle\varphi|\varphi\rangle + \langle\psi|\varphi\rangle\langle\varphi|\psi\rangle = 1 \neq 0. \quad (15.2.2)$$

What this means is that the presence of superselection rules really means that not all Hermitian operators are valid observables. That is the space of observables is a strict subset of the space of Hermitian operators in the presence of superselection rules.

An example of a superselection is the fermionic/bosonic nature of a particle. Suppose that  $|1/2\rangle$  is a fermionic state, and  $|1\rangle$  is a bosonic state. We could, at least mathematically in the Hilbert space, define a vector  $|\psi\rangle = a|1/2\rangle + b|1\rangle$ , however, this doesn't correspond to a physically meaningful state. This is because upon a  $2\pi$  rotation, due to some operator,  $U(2\pi)$ , a fermionic state will pick up a negative sign, but a bosonic state won't, so

$$U(2\pi)|\psi\rangle = -a|1/2\rangle + b|1\rangle. \quad (15.2.3)$$

Clearly this doesn't make any physical sense, since the relative phase between  $a$  and  $b$  changes by  $\pi$  upon rotation.

Another example is conserved charge. It doesn't make sense to write, for example,  $|\psi\rangle = |e^-\rangle + |e^+\rangle$ , since these transform under complex conjugates representations, and so pick up opposite phases when transforming. The only time that this isn't a problem is if the representation is real, but this corresponds to the particle being neutral.

## 15.3 Electric Dipole

Consider an extended charge distribution,  $\rho$ . To first order in the multipole expansion we can write the Hamiltonian in an external field,  $E$ , as

$$H = H_0 + H' \quad (15.3.1)$$

where  $H_0$  is a Coulomb term and  $H' = -\mathbf{d} \cdot \mathbf{E}$ , where  $\mathbf{d}$  is the electric dipole moment, given by

$$\mathbf{d} := \int \rho(\mathbf{x}) \mathbf{x} d^3x. \quad (15.3.2)$$

Importantly  $\mathbf{d}$  transforms the same as  $\mathbf{x}$ , due to the factor of  $\mathbf{x}$  in the definition. We know that  $\mathbf{x}$  is a  $j = 1$  tensor operator (that is, a vector) and hence the Wigner–Eckart theorem applies, as do the associated selection rules.

For simplicity we consider the case of an external field aligned along the  $z$ -axis, so  $\mathbf{E} = E_z \mathbf{e}_z$ . It follows from this that  $H' \propto z$  since

$$H' = - \int \rho(\mathbf{x}) \mathbf{x} d^3x \cdot \mathbf{e}_E = -E_z \int \rho(\mathbf{x}) \mathbf{x} \cdot \mathbf{e}_z d^3x = -E_z \int \rho(\mathbf{x}) z d^3x. \quad (15.3.3)$$

So  $H'$  transforms like  $z$ . Looking at the spherical harmonics in Equation (14.3.22) we see that

$$Y_{10} \propto \frac{z}{r}. \quad (15.3.4)$$

Considering the definition of the solid harmonics as  $\mathcal{Y}_{lm} = r^l Y_{lm}$  we see that

$$\mathcal{Y}_{10} \propto z. \quad (15.3.5)$$

So  $H' \propto \mathcal{Y}_{10}$ .

Applying the Wigner–Eckart theorem we then have

$$\langle l, m | H' | l', m' \rangle \propto \langle l, m | \mathcal{Y}_{10} | l', m' \rangle = C_{m0m'}^{l1l'} \langle l || \mathcal{Y}_1 || l' \rangle. \quad (15.3.6)$$

We know that the Clebsch–Gordan coefficient,  $C_{Mm_1m_2}^{Jj_1j_2}$ , is non-zero only if  $M = m_1 + m_2$ , so we must have  $m = m'$ , and if  $|j_1 - j_2| \leq J \leq j_1 + j_2$ , which means that we must have  $\Delta l := |l - l'| \leq 1$ . Further  $C_{m0m'}^{010} = 0$  for all allowed values of  $m$  and  $m'$ , which is just  $m = m' = 0$ . This means we cannot have  $l = l' = 0$ .

Further considering parity selection rules since  $\eta_x = 1$ , and  $H'$  transforms like  $x$ , so has the same parity which means we must have  $l + l' + \eta_x = l + l' + 1$  even. Hence we cannot have  $l - l' = 0$ , since if this is the case then  $l + l' + 1 = 2l + 1$ , which is odd for all integer  $l$ . So  $\Delta l \neq 0$ . Hence we have the selection rules

$$m = m', \quad \text{and} \quad \Delta l = |l - l'| = 1. \quad (15.3.7)$$

## 15.4 Pauli's Hydrogen Atom



In this section we use units such that  $c = \hbar = 1$ .

Shortly after Schrödinger solved his equation for the hydrogen atom Pauli did so using group theory and a hidden  $SU(2) \otimes SU(2)$  symmetry. The energy levels of the hydrogen atom are given by

$$E_n = -\frac{me^4}{2} \frac{1}{n^2}. \quad (15.4.1)$$

Importantly there is no  $l$  dependence. This means that for each value of the principal quantum number,  $n$ , we have  $l = 0, \dots, n-1$ , which means the total degeneracy is

$$\sum_{l=0}^{n-1} (2l+1) = 2 \sum_{l=0}^{n-1} l + \sum_{l=0}^{n-1} 1 = 2 \frac{n-1}{2} n + n = n^2. \quad (15.4.2)$$

Note that the  $2l+1$  term is due to degeneracy in  $m$ , which runs from  $-l$  to  $l$ , so takes on  $2l+1$  values for fixed  $l$ .

The degeneracy in  $m$  is expected, since we have a spherically symmetric potential, and  $m$  picks out a direction. The fact that we have no  $l$  dependence suggests that there is some further symmetry which we are missing. Pauli used the fact that there is a conserved vector,  $\mathbf{A}$ , with components

$$A_i = \varepsilon_{ijk} p_j L_k - me^2 \frac{x_i}{r}, \quad (15.4.3)$$

which is conserved in a spherically symmetric system in classical mechanics. This is known as the **Laplace–Runge–Lenz vector**. In quantum mechanics there is a related quantity given by symmetrisation:

$$A_i = \varepsilon_{ijk} p_j L_k - me^2 \frac{x_i}{r} - ip_i \quad (15.4.4)$$

Note that this is an operator in quantum mechanics. It can be shown that the following hold:

$$\mathbf{L} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{L} = 0, \quad (15.4.5)$$

$$[L_i, A_j] = i\epsilon_{ijk}A_k, \quad (15.4.6)$$

$$[H, A_i] = 0, \quad \text{where} \quad H = \sum_i \frac{p_i^2}{2m} - \frac{e^2}{r}. \quad (15.4.7)$$

$$(15.4.8)$$

The second of these means that  $A_i$  is a  $j = 1$  vector operator. The third means that  $\mathbf{A}$  is conserved.

By Noether's theorem we can identify from this conservation law that there must be a hidden symmetry. To find this symmetry we make some new definitions and extract a Lie algebra. Start by defining

$$\tilde{A}_i := \frac{A_i}{\sqrt{-2mH}}. \quad (15.4.9)$$

Note that for a bound state the total energy is negative, so we take the minus sign in the root to keep  $\tilde{A}_i$  real. It can then be shown that

$$[\tilde{A}_i, \tilde{A}_j] = i\epsilon_{ijk}\tilde{A}_k. \quad (15.4.10)$$

This shows that  $\tilde{A}_i$  are the generators of a  $\mathfrak{su}(2)$  Lie algebra. We then define  $X_i^\pm := (L_i \pm \tilde{A}_i)/2$ . From this the commutation relation

$$[X_i^\pm, X_j^\pm] = i\epsilon_{ijk}X_k^\pm \quad (15.4.11)$$

follows, since both  $L_i$  and  $A_i$  individually satisfy the requirements of a  $\mathfrak{su}(2)$  Lie algebra, and since  $\mathbf{L} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{L} = 0$  cross terms vanish.

The result is that we identify an overall  $SU(2) \otimes SU(2)$  symmetry, corresponding to the two copies of the  $\mathfrak{su}(2)$  Lie algebra, spanned by  $X_i^+$  and  $X_i^-$  respectively. We can think of one copy as resulting from rotational symmetry and the second from the symmetry associated with conservation of  $\mathbf{A}$ .

One question we may want to ask is what are the Casimir operators of these two Lie algebras are. The fact that  $\mathbf{L} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{L} = 0$  means that the two Casimir operators are actually the same, in particular the quadratic Casimir is

$$C_2^+ = (X^+)^2 = (X^-)^2 = C_2^- = C_2. \quad (15.4.12)$$

We can parametrise this Casimir by  $C_2 = x(x+1)$  with  $x = 0, 1/2, 1, 3/2, \dots$

From this it is possible to show that

$$A_i A_i = 2mH(L_i L_i + 1) + m^2 e^4. \quad (15.4.13)$$

Dividing through by  $-2mH$  and solving for  $H$  we get

$$H = -\frac{me^4/2}{L_i L_i + \tilde{A}_i \tilde{A}_i + 1} = -\frac{me^4/2}{4C_2 + 1} = -\frac{me^4/2}{(2x+1)^2} \quad (15.4.14)$$

for  $x = 0, 1/2, 1, 3/2, \dots$ . The second equality follows using  $L_i = X_i^+ + X_i^-$ . Counting the degeneracy of the states corresponds to counting the states in the direct product

representation  $\rho_{j=x} \otimes \rho_{j=x}$ , which has  $(2x+1)^2$  states. Comparing this with our initial work in this section we identify the principal quantum number  $n = 2x + 1$ .

Pauli's approach here was to write the Hamiltonian in terms of two Casimir operators (which just happen to be equal). That this is possible is not an accident. In the absence of spin we need three quantum numbers to describe a state, say  $\{x_1, x_2, x_3\}$ ,  $\{p_1, p_2, p_3\}$ , or as is most common when talking about the hydrogen atom,  $\{n, l, m\}$ . Spherical symmetry guarantees that  $m$  doesn't appear in the final result, and hence we are reduced to two quantum numbers. This means that a rank two symmetry, with two Casimirs, is sufficient to obtain solutions based solely on symmetry. In order to include the effects of spin, as well as other relativistic corrections, we still need to solve Schrödinger's equation.

## Appendices

# A

## Mathematical Preliminaries

### A.1 Basic Mathematics

#### A.1.1 Notation

**Notation A.1.1 — Number Sets** The set of natural numbers is

$$\mathbb{N} := \{0, 1, 2, \dots\}. \quad (\text{A.1.2})$$

Note that the inclusion of zero in  $\mathbb{N}$  is subject to debate. The set of integers is denoted

$$\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}. \quad (\text{A.1.3})$$

The set of positive integers is denoted

$$\mathbb{Z}_{>0} := \{1, 2, \dots\}. \quad (\text{A.1.4})$$

The set of rational numbers is denoted

$$\mathbb{Q} := \{p/q \mid p, q \in \mathbb{Z} \text{ and } q \neq 0\}. \quad (\text{A.1.5})$$

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  $\mathbb{H}$ .

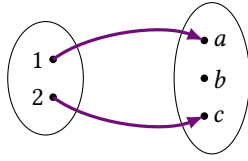
**Notation A.1.6 — Sphere** The unit sphere in  $n + 1$  dimensions is

$$S^n := \{\mathbf{x} \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}. \quad (\text{A.1.7})$$

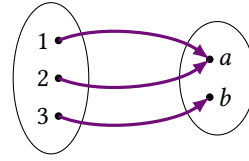
Note that  $S^n$  is an  $n$ -dimensional manifold, which we view as embedded in  $(n + 1)$ -dimensional Euclidean space,  $\mathbb{R}^{n+1}$ .

What we normally call the circle is  $S^1$  and what we normally call the sphere is  $S^2$ .

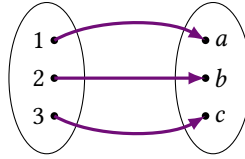
**Notation A.1.8 — Sets of Matrices** We denote the set of  $m \times n$  matrices with entries in  $\mathbb{F}$  (which is usually a field and usually  $\mathbb{R}$  or  $\mathbb{C}$ ) by  $\mathcal{M}_{m \times n}(\mathbb{F})$ . We denote the set of square  $n \times n$  matrices with entries in  $\mathbb{F}$  by  $\mathcal{M}_n(\mathbb{F})$ .



(a) An injective function,  $f: \{1, 2\} \rightarrow \{a, b, c\}$ . Note that  $f(x) \neq b$  for any  $x \in \{1, 2\}$  and so the function fails to be surjective.



(b) A surjective function,  $g: \{1, 2, 3\} \rightarrow \{a, b\}$ . Note that  $g(1) = g(2)$  but  $1 \neq 2$  and so the function fails to be injective.



(c) A bijective function,  $h: \{1, 2, 3\} \rightarrow \{a, b, c\}$ .

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

$$\text{GL}(n, \mathbb{F}) = \{A \in \mathcal{M}_n(\mathbb{F}) \mid \det A \neq 0\}. \quad (\text{A.1.9})$$

If  $\mathbb{F}$  is evident from context we may simply write  $\text{GL}(n)$ . If  $V$  is an  $n$ -dimensional vector space over  $\mathbb{F}$  then we may also write this set as  $\text{GL}(V)$ .

**Notation A.1.10 — Einstein Summation Convention** When two identical indices appear in the same term then they are summed over, for example,

$$x_i y_i = \sum_i x_i y_i. \quad (\text{A.1.11})$$

## A.1.2 Definitions

**Definition A.1.12 — Function Types** Let  $\varphi: A \rightarrow B$ . Then  $\varphi$  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  $\varphi$  is both injective and surjective.

A function is invertible if and only if it is bijective.

**Definition A.1.13 — Kernel** Given a map  $\varphi: A \rightarrow B$  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:

$$\ker \varphi := \{a \in A \mid \varphi(a) \text{ is the trivial element of } B\} \subseteq A. \quad (\text{A.1.14})$$

**Definition A.1.15 — Image** Given a map  $\varphi: A \rightarrow B$  the **image** is set of  $b \in B$  for which there exists some  $a \in A$  such that  $\varphi(a) = b$ :

$$\text{Im } \varphi = \varphi(A) := \{b \in B \mid \exists a \in A \text{ such that } \varphi(a) = b\} \subseteq B. \quad (\text{A.1.16})$$

**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**,  $\delta_{ij}$ , is defined as

$$\delta_{ij} := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases} \quad (\text{A.1.19})$$

Note that  $\delta_{ij}$  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  $\varepsilon_{1\dots i\dots j\dots n} = -\varepsilon_{1\dots j\dots i\dots n}$ , for example  $\varepsilon_{213\dots n} = -1$ . Antisymmetry also means that Levi-Civita symbol vanishes if it has repeated indices. Most commonly  $n = 3$  and

$$\varepsilon_{ijk} := \begin{cases} 1, & \text{if } (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2), \\ -1, & \text{if } (i, j, k) = (1, 3, 2), (2, 1, 3), (3, 2, 1), \\ 0, & \text{if any index is repeated.} \end{cases} \quad (\text{A.1.21})$$

**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

$$[a] := \{x \mid a \sim x\}. \quad (\text{A.1.23})$$

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**  $\sim$  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  $v \sim u$  if  $u$  and  $v$  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,  $\text{id}_A: A \rightarrow A$  defined by  $\text{id}_A(a) = a$  for all  $a \in A$  is an isomorphism since  $\text{id}_A(aa') = aa' = \text{id}_A(a)\text{id}_A(a')$  and clearly  $\text{id}_A$  is invertible, and is its own inverse.
- Let  $A$  and  $B$  be isomorphic groups. Then there exists some bijection  $\varphi: A \rightarrow B$  such that  $\varphi(aa') = \varphi(a)\varphi(a')$ . Since  $\varphi$  is a bijection  $\varphi^{-1}: B \rightarrow A$  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  $\varphi$  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  $\varphi^{-1}(\varphi(aa')) = \varphi^{-1}(b)\varphi^{-1}(\varphi(a'))$  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  $\varphi$  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  $\varphi: A \rightarrow B$  and  $\psi: B \rightarrow C$ . We claim that  $\psi \circ \varphi: A \rightarrow C$  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  $\varphi$  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  $\psi$  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,  $\cdot: \mathbb{F} \times V \rightarrow V$ , known as scalar multiplication, and  $+: V \times V \rightarrow V$ , known as vector addition, which are defined such that the following hold for all  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  and  $k, k' \in \mathbb{F}$ :

1. **Associativity:**  $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ ,
2. There exists a **zero vector**,  $\mathbf{0} \in V$ , such that  $\mathbf{u} + \mathbf{0} = \mathbf{u}$ .
3. There exists  $-\mathbf{u} \in V$  such that  $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ . We write this as  $\mathbf{u} - \mathbf{u}$  for short.
4. **Commutativity:**  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ .
5. Distributivity of scalar multiplication over vector addition  $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$ .
6. Distributivity of scalar multiplication over field addition  $(k + k')\mathbf{u} = k\mathbf{u} + k'\mathbf{u}$ .
7. Compatibility of field and scalar multiplication  $(kk')\mathbf{u} = k(k'\mathbf{u})$ .
8.  $1\mathbf{u} = \mathbf{u}$  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  $\langle -, - \rangle: \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$  such that for all  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{H}$  and  $k, k' \in \mathbb{C}$

1.  $\langle -, - \rangle$  is linear in its second argument, that is

$$\langle \mathbf{u}, k\mathbf{v} + k'\mathbf{w} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle + k'\langle \mathbf{u}, \mathbf{w} \rangle. \quad (\text{A.2.3})$$

2.  $\langle -, - \rangle$  is conjugate symmetric:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle^*. \quad (\text{A.2.4})$$

3.  $\langle -, - \rangle$  is positive definite, that is  $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$  with equality only if  $\mathbf{u} = \mathbf{0}$ .



Mathematicians often define an inner product to be linear in its first argument, so

$$\langle k\mathbf{u} + k'\mathbf{v}, \mathbf{w} \rangle = k\langle \mathbf{u}, \mathbf{w} \rangle + k'\langle \mathbf{v}, \mathbf{w} \rangle. \quad (\text{A.2.5})$$

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

$$\langle ku + k'v, w \rangle = k^* \langle u, w \rangle + k'^* \langle v, w \rangle. \quad (\text{A.2.6})$$

We can then define a **norm** on  $\mathcal{H}$  by  $\|u\| := \sqrt{\langle u, u \rangle}$ .

The final condition for  $\mathcal{H}$  to be a Hilbert space is completeness. Namely, that if the series  $\sum_{n=0}^{\infty} u_n$  converges absolutely, so that  $\sum_{n=0}^{\infty} \|u_n\|$  converges to a finite value then the original series,  $\sum_{n=0}^{\infty} u_n$ , 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

$$\langle u, v \rangle = \langle (u_1, \dots, u_n), (v_1, \dots, v_n) \rangle := \sum_{i=1}^n u_i^* v_i = u_i^* v_i \quad (\text{A.2.8})$$

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,  $f: X \rightarrow \mathbb{C}$ , is square integrable if<sup>u</sup>

$$\int_X |f(x)|^2 dx \quad (\text{A.2.10})$$

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

$$\langle f, g \rangle := \int_X f^*(x) g(x) dx. \quad (\text{A.2.11})$$

**R** Another subtlety 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

$$\psi \rightarrow \frac{\psi}{\|\psi\|} = \frac{\psi}{\sqrt{\langle \psi, \psi \rangle}} = \left( \int |\psi(x)|^2 dx \right)^{-1/2} \psi. \quad (\text{A.2.12})$$

This only makes sense if  $\int |\psi(x)|^2 dx$  is finite (and nonzero).

<sup>a</sup>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,  $\mathbf{u}$ , with **kets**,  $|u\rangle$ , and dual vectors,  $\mathbf{v}$ , with **bras**,  $\langle v|$ . The inner product  $\langle \mathbf{u}, \mathbf{v} \rangle$  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,  $f: V \rightarrow W$ , is said to be a linear operator if for  $\mathbf{u}, \mathbf{v} \in V$  and  $k \in \mathbb{F}$  we have

$$f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v}), \quad \text{and} \quad f(k\mathbf{u}) = kf(\mathbf{u}). \quad (\text{A.2.15})$$

Instead of the function notation  $f(\mathbf{u})$  we typically use a multiplicative notation,  $A\mathbf{u}$ , 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

$$A(|u\rangle + |v\rangle) = A|u\rangle + A|v\rangle, \quad \text{and} \quad A(k|u\rangle) = kA|u\rangle. \quad (\text{A.2.16})$$

An operator is **antilinear** if

$$A(|u\rangle + |v\rangle) = A|u\rangle + A|v\rangle, \quad \text{and} \quad A(k|u\rangle) = k^*A|u\rangle. \quad (\text{A.2.17})$$

An example of such an operator is the time reversal operator,  $T$ , which takes  $t \rightarrow -t$ .

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 than 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

$$\lambda_i |e_i\rangle = |0\rangle, \quad (\text{A.2.19})$$

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

$$|u\rangle = u_i |e_i\rangle \quad (\text{A.2.20})$$

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

$$\langle e_i|e_j\rangle = \delta_{ij}. \quad (\text{A.2.21})$$

**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

$$\mathbb{1} = \sum_{i=1}^N |e_i\rangle\langle e_i|. \quad (\text{A.2.23})$$

Recall that the identity operator is defined such that

$$\mathbb{1}|u\rangle = |u\rangle \quad (\text{A.2.24})$$

for all  $|u\rangle \in V$ .

**Definition A.2.25 — Partition of the Identity** A **projection operator**,  $P_i$ , is an operator satisfying

$$P_i P_j = \delta_{ij} P_i, \quad \text{and} \quad P_i^\dagger = P_i. \quad (\text{A.2.26})$$

A **partition of the identity** is a collection of projection operators,  $\{P_j\}$ , such that

$$\mathbb{1} = \sum_{j=1}^{N_P} P_j \quad (\text{A.2.27})$$

where  $N_P = |\{P_j\}|$  is the number of projection operators in the partition. We can write each partition operator as

$$P_j = \sum_{i=1}^{N_j} |e_i\rangle\langle e_i| \quad (\text{A.2.28})$$

where  $N_j$  are such that

$$\dim V = N = \sum_{j=1}^{N_P} N_j. \quad (\text{A.2.29})$$

**Definition A.2.30 — Matrix Element** Given a linear operator,  $A: V \rightarrow V$ , and an orthonormal basis,  $\{|e_i\rangle\}$ , we define the **matrix elements** to be

$$A_{ij} := \langle e_i | A e_j \rangle = \langle e_i | A | e_j \rangle \quad (\text{A.2.31})$$

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

$$A = \sum_{i=1}^N \sum_{j=1}^N A_{ij} |e_i\rangle \langle e_j|. \quad (\text{A.2.32})$$

**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

$$A|v_i\rangle = \lambda_i |v_i\rangle. \quad (\text{A.2.34})$$

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,  $A^T$ , has matrix elements  $A_{ij}^T = A_{ji}$ .

A matrix is **symmetric** if  $A^T = A$ , or **antisymmetric** if  $A^T = -A$ .

Given a matrix,  $A$ , with matrix elements  $A_{ij}$ , the **Hermitian conjugate**,  $A^\dagger$ , has matrix elements  $A_{ij}^\dagger = A_{ji}^*$ . Here  $*$  denotes the **complex conjugate**, so  $(x + iy)^* = x - iy$  and  $(re^{i\vartheta})^* = re^{-i\vartheta}$  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  $\mathbf{v}$  an eigenvector with nonzero eigenvalue  $\lambda$ . Note that this means  $\mathbf{v}$  is nonzero. If 0 is an eigenvalue of  $A$  then this is real, so we need not consider this case further. By definition  $A\mathbf{v} = \lambda\mathbf{v}$ . Taking the Hermitian conjugate of both sides we get  $\mathbf{v}^\dagger A^\dagger = \lambda^* \mathbf{v}^\dagger$ , where we have used  $(XY)^\dagger = Y^\dagger X^\dagger$ . Multiplying both sides on the right by  $\mathbf{v}$  we get  $\mathbf{v}^\dagger A\mathbf{v} = \lambda^* \mathbf{v}^\dagger \mathbf{v}$ . Identifying  $A\mathbf{v} = \lambda\mathbf{v}$  on the left-hand side this becomes  $\mathbf{v}^\dagger \lambda\mathbf{v} = \lambda \mathbf{v}^\dagger \mathbf{v} = \lambda^* \mathbf{v}^\dagger \mathbf{v}$ . It follows that we must have  $\lambda = \lambda^*$ , which means we must have  $\lambda \in \mathbb{R}$ .  $\square$

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  $\lambda_i$  and corresponding eigenvectors  $|v_i\rangle$  we can write this matrix as

$$A = \sum_{i=1}^N \lambda_i |v_i\rangle \langle v_i|. \quad (\text{A.2.37})$$

This is diagonalised by the transformation  $V^\dagger A V$  where

$$V = \sum_{i=1}^N |v_i\rangle \langle e_i| \quad (\text{A.2.38})$$

where  $|e_i\rangle$  are the basis vectors in the original basis. It is easy to see that this transform gives the desired result:

$$V^\dagger A V = \underbrace{|e_i\rangle \langle v_i|}_{=V^\dagger} \underbrace{(\lambda_j |v_j\rangle \langle v_j|)}_{=A} \underbrace{|v_k\rangle \langle e_k|}_{=V} \quad (\text{A.2.39})$$

$$= \lambda_j |e_i\rangle \langle v_i| v_j\rangle \langle v_j| v_k\rangle \langle e_k| \quad (\text{A.2.40})$$

$$= \lambda_j \delta_{ij} \delta_{jk} |e_i\rangle \langle e_k| \quad (\text{A.2.41})$$

$$= \lambda_i |e_i\rangle \langle e_i| \quad (\text{A.2.42})$$

This last term is just a diagonal matrix with the eigenvalues,  $\lambda_i$ , 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

$$A^{-1} = \text{diag}(1/\lambda_1, \dots, 1/\lambda_N). \quad (\text{A.2.44})$$

**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:

$$\langle u|Av\rangle = \langle u|A|v\rangle = \langle A^\dagger u|v\rangle. \quad (\text{A.2.46})$$

For a unitary matrix,  $U$ , this implies

$$\langle Uu|Uv\rangle = \langle u|U^\dagger U|v\rangle = \langle u|\mathbb{1}|v\rangle = \langle u|v\rangle. \quad (\text{A.2.47})$$

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

$$\text{tr } A := \sum_i \langle e_i | A | e_i \rangle = A_{ii} \quad (\text{A.2.49})$$

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  $\text{tr}(AB) = \text{tr}(BA)$ ,  $\text{tr}(ABC) = \text{tr}(BCA) = \text{tr}(CAB)$ , etc.

The trace is linear, meaning  $\text{tr}(kA) = k \text{tr}(A)$  for scalar  $k$ .

$\langle A, B \rangle := \text{tr}(A^\dagger B)$  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

$$\det A = |A| := \varepsilon_{i_1 \dots i_N} A_{1i_1} \cdots A_{Ni_N} \quad (\text{A.2.51})$$

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:

$$\det(AB) = \det(A) \det(B) = \det(B) \det(A) = \det(BA). \quad (\text{A.2.52})$$

**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

$$A = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & A_n \end{pmatrix} \quad (\text{A.2.54})$$

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  $|v_i\rangle \oplus |w_a\rangle = |v_i \oplus w_a\rangle \in V \oplus W$  and extending the inner product to

$$\langle v_i \oplus w_a | v_j \oplus w_b \rangle_{V \oplus W} := \langle v_i | v_j \rangle_V + \langle w_a | w_b \rangle_W \quad (\text{A.2.56})$$

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

$$\dim(V \oplus W) = \dim V + \dim W. \quad (\text{A.2.57})$$

Given  $A \in \text{GL}(V)$  and  $B \in \text{GL}(W)$  the direct sum,  $A \oplus B$ , acts on  $|v\rangle \oplus |w\rangle \in V \oplus W$  as

$$(A \oplus B)(|v\rangle \oplus |w\rangle) := (A|v\rangle) \oplus (B|w\rangle). \quad (\text{A.2.58})$$

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:

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}. \quad (\text{A.2.59})$$

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

$$\langle v_i \otimes w_a | v_j \otimes w_b \rangle_{V \otimes W} := \langle v_i | v_j \rangle_V \langle w_a | w_b \rangle \quad (\text{A.2.61})$$

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

$$\dim(V \otimes W) = \dim(V) \dim(W). \quad (\text{A.2.62})$$

Given  $A \in \text{GL}(V)$  and  $B \in \text{GL}(W)$  the direct product,  $A \otimes B$ , acts on  $|v\rangle \otimes |w\rangle \in V \otimes W$  as

$$(A \otimes B)(|v\rangle \otimes |w\rangle) = (A|v\rangle) \otimes (B|w\rangle). \quad (\text{A.2.63})$$

■ **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  $\text{SO}(1)$ .

**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.
- Rank 2.
- Abelian.

**Definition B.1.6**  $S_3$  is the permutation group on 3 elements, see [Definition B.1.11](#).

- Order 6.
- Rank 2.
- Non-Abelian.

**Definition B.1.7** The **quaternion group**,  $Q$ , has the group presentation

$$Q = \langle -e, i, j, k \mid (-e)^2 = e, i^2 = j^2 = k^2 = ijk = e \rangle. \quad (\text{B.1.8})$$

- Order 8.
- Rank 2.
- Non-Abelian.

The Pauli matrices provide a two-dimensional complex representation by the correspondence  $(-e, i, j, k) \rightarrow (-\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$ .

### 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

$$\mathbb{Z}_n = \langle a \mid a^n = e \rangle. \quad (\text{B.1.10})$$

Identifying  $a = e^{2i\pi/n}$  and the operation as multiplication we get a group formed from the  $n$ th roots of unity. Identifying  $a = 1$  and the operation as addition modulo  $n$  we get a group formed from  $\{0, \dots, n-1\}$ .

- Order  $n$ .
- Rank 1.
- Cyclic.
- 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  $n!$ .
- Rank 2.
- Non-Abelian ( $n > 2$ ).

$S_1$  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.
- Abelian.
- Cyclic.

**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

$$\mathrm{GL}(n, \mathbb{F}) = \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M \neq 0\}. \quad (\text{B.3.6})$$

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

$$\mathrm{SL}(n, \mathbb{F}) = \{M \in \mathcal{M}_n(\mathbb{F}) \mid \det M = 1\}. \quad (\text{B.3.8})$$

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

$$\mathrm{O}(n) = \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = OO^\top = \mathbb{1}\}. \quad (\text{B.3.10})$$

- 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

$$\mathrm{SO}(n) = \{O \in \mathcal{M}_n(\mathbb{R}) \mid O^\top O = OO^\top = \mathbb{1} \text{ and } \det O = 1\}. \quad (\text{B.3.12})$$

- 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

$$\mathrm{U}(n) = \{U \in \mathcal{M}_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1}\}. \quad (\text{B.3.14})$$

- Non-Abelian ( $n > 1$ ).

$U(n)$  is a subgroup of  $GL(n, \mathbb{C})$ .

$U(n)$  is the group which preserves the standard inner product on  $\mathbb{C}^n$ .

$U(1)$  is isomorphic to  $SO(2)$  and the circle group,  $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$  under multiplication.

**Definition B.3.15** The **special unitary group**

$$SU(n) = \{U \in M_n(\mathbb{C}) \mid U^\dagger U = UU^\dagger = \mathbb{1} \text{ and } \det U = 1\}. \quad (\text{B.3.16})$$

- Non-Abelian ( $n > 1$ ).

$SU(n)$  is a subgroup of  $U(n)$  and  $SL(n, \mathbb{C})$ .

**Definition B.3.17** The **isometries of Euclidean space**

$$ISO(n) = O(n) \ltimes \mathbb{R}^n \quad (\text{B.3.18})$$

where  $(R, \mathbf{a})(R', \mathbf{a}') := (RR', \mathbf{a} + R\mathbf{a}')$ .

- Non-Abelian

$ISO(n)$  is the group of distance preserving transformations of Euclidean space.

$ISO(n)$  is the group of rotations, reflections, and translations of  $\mathbb{R}^n$ .

$ISO(n)$  has both  $O(n)$  and  $\mathbb{R}^n$  as normal subgroups.

**Definition B.3.19** The **Lorentz group**,  $O(1, 3)$ , is the group of all Lorentz transformations of Minkowski space.

- Non-Abelian.

$O(1, 3)$  is the group that preserves the quadratic form  $(t, x, y, z) \mapsto t^2 - x^2 - y^2 - z^2$ .

$SO^+(1, 3)$  is the subgroup of  $O(1, 3)$  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  $ISO(1, 3)$ . That is, it is the group of all Lorentz transformations and translations.

- Non-Abelian.

The Poincaré group can be identified as the semidirect product  $ISO(1, 3) =$

$\mathbb{R}^{1,3} \rtimes O(1, 3)$  where  $\mathbb{R}^{1,3}$  is the group of spacetime translations of Minkowski space and  $O(1, 3)$  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

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  $dx^i/d\lambda$  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

$$\bigcup_{i \in I} U_i \in \mathcal{T}, \quad (\text{C.1.3})$$

- $\mathcal{T}$  is closed under finite intersections, that is if  $\{U_i, \dots, U_N\} \subseteq \mathcal{T}$  then

$$\bigcap_{i=1}^N U_i \in \mathcal{T}. \quad (\text{C.1.4})$$

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:

$$\mathcal{T} = \{\emptyset, \{a, b, c, d\}, \{a, b\}, \{c, d\}\}. \quad (\text{C.1.5})$$

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  $B_3 \in \mathcal{B}$  such that  $x \in B_3$  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

$$U = \bigcup_{B_i \in \mathcal{B} \subseteq U} B_i. \quad (\text{C.1.7})$$

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

$$B = \{\mathbf{r} \in \mathbb{R}^n \mid |\mathbf{r} - \mathbf{a}| \leq 1 \text{ for some } \mathbf{a} \in \mathbb{R}^n\} \quad (\text{C.1.8})$$

with  $|r| = \sum_i x_i^2$  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,  $f: X \rightarrow Y$ . The **preimage** of some  $S \subseteq Y$  under this function is the set

$$f^*(S) = f^{-1}(S) := \{a \in X \mid f(a) \in S\} \subseteq X. \quad (\text{C.1.9})$$

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  $f: X \rightarrow Y$  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  $\varepsilon$ - $\delta$  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  $\varphi$  is a homeomorphism

$$\varphi: U \rightarrow \text{im } \varphi \subseteq \mathbb{R}^n. \quad (\text{C.1.11})$$

We call  $n \in \mathbb{Z}_{>0}$  the **dimension** of  $U$ , it is independent of  $\varphi$ .

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  $(U, \varphi)$  is a chart. The coordinates of some point  $p \in U$  are the Cartesian coordinates,  $x_p$ , 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

$$\varphi_1 \circ \varphi_2^{-1}: \varphi_2(U_1 \cap U_2) \rightarrow \varphi_1(U_1 \cap U_2), \quad (\text{C.1.17})$$

$$\varphi_2 \circ \varphi_1^{-1}: \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2). \quad (\text{C.1.18})$$

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

$$(\varphi_1 \circ \varphi_2^{-1})(\mathbf{x}_p^{(2)}) = \varphi_1(p) = \mathbf{x}_p^{(1)}, \quad (\text{C.1.19})$$

$$(\varphi_2 \circ \varphi_1^{-1})(\mathbf{x}_p^{(1)}) = \varphi_2(p) = \mathbf{x}_p^{(2)}. \quad (\text{C.1.20})$$

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,  $f: M \rightarrow \mathbb{R}$  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

$$f_U: \varphi(U) \subseteq \mathbb{R}^n \rightarrow \mathbb{R} \quad (\text{C.1.22})$$

is defined by

$$f(p) = f_U(\mathbf{x}_p) \quad (\text{C.1.23})$$

for  $p \in U$ .

This is useful since we can deal with functions  $\mathbb{R}^n \rightarrow \mathbb{R}$  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  $f_1(\mathbf{x}_p^{(1)}) = f_2(\mathbf{x}_p^{(2)})$ .

### 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:

$$\mathcal{S}(M) = [\mathcal{A}(M)], \quad (\text{C.1.26})$$

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,  $f: M \rightarrow \mathbb{R}$ , on a smooth manifold,  $M$ , is said to be **smooth** if all of its local coordinate representatives,

$$f_a = f \circ \varphi_a^{-1}: \varphi_a(U_a) \rightarrow \mathbb{R}, \quad (\text{C.1.28})$$

are smooth.

Let  $M$  and  $V$  be smooth manifolds of dimensions  $m$  and  $n$  with charts  $(U_a, \varphi_a)$  and  $(V_b, \psi_b)$  respectively. Then the map  $\mu: M \rightarrow N$  is **smooth** if all of its local coordinate representatives,

$$\mu_{ab} = \psi_b \circ \mu \circ \varphi_a^{-1}: \varphi_a(U_a) \subseteq \mathbb{R}^m \rightarrow \psi_b(V_b) \subseteq \mathbb{R}^n \quad (\text{C.1.29})$$

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 ( $C^\infty$ ) 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 of a single chart,  $(\mathbb{R}^n, \text{id})$ , where  $\text{id}$  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  $2\pi$ . 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  $2\pi$  in the neighbourhood, so our neighbourhood 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  $2\pi$ , 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  $\text{SO}(3)$ , 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,  $O(1,3)$ , 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  $p$ ,  $T_p M$ , 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. The set of all smooth functions  $M \rightarrow \mathbb{R}$ , 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,  $D: C^\infty(M) \rightarrow \mathbb{R}$  satisfying the Leibniz identity,

$$D(fg) = D(f)g(p) + f(p)D(g) \quad (\text{C.2.2})$$

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

$$(D_1 + D_2)(f) = D_1(f) + D_2(f), \quad \text{and} \quad (\lambda D)(f) = \lambda D(f) \quad (\text{C.2.3})$$

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,  $T_p M$ .

We can then define the **tangent bundle**,  $TM$ , as the disjoint union of all tangent spaces,

$$TM := \bigsqcup_{p \in M} T_p M = \{(p, x) \mid p \in M, x \in T_p M\}. \quad (\text{C.2.4})$$

That is we combine all vectors in any tangent space,  $T_p M$ , 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,  $V: M \rightarrow$

$TM$  such that  $\pi \circ V$  is the identity mapping on  $M$  where  $\pi: TM \rightarrow M$  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  $\cdot : A \times A \rightarrow A$ . 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 have  $(ax)(by) = (ab)(xy)$ .

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,  $\varphi: A \rightarrow B$  such that

$$f(ax + by) = af(x) + bf(y), \quad \text{and} \quad f(xy) = f(x)f(y) \quad (\text{D.0.4})$$

for all  $a, b \in \mathbb{F}$  and  $x, y \in A$ .

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,  $[-, -] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ , called the **Lie bracket**, must also satisfy the following

- **Bilinearity:** For all  $x, y, z \in \mathfrak{g}$  and  $a, b \in \mathbb{F}$

$$[ax + by, z] = a[x, z] + b[y, z], \quad \text{and} \quad (\text{D.0.6})$$

$$[z, ax + by] = a[z, x] + b[z, y]. \quad (\text{D.0.7})$$

- **Alternativity:** For all  $x \in \mathfrak{g}$

$$[x, x] = 0. \quad (\text{D.0.8})$$

- The **Jacobi identity:** For all  $x, y, z \in \mathfrak{g}$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \quad (\text{D.0.9})$$

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,  $[A, B] := AB - BA$ , 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,  $[\mathbf{v}, \mathbf{u}] := \mathbf{v} \times \mathbf{u}$ .

Let  $A$  be a commutative  $\mathbb{R}$ -algebra, that is  $\mathbf{x}\mathbf{y} = \mathbf{y}\mathbf{x}$  for all  $\mathbf{x}, \mathbf{y} \in A$ . We define a derivation,  $D : A \rightarrow A$  to be a  $\mathbb{R}$ -linear map ( $D(a\mathbf{x}) = aD(\mathbf{x})$  for all  $a \in \mathbb{R}$  and  $\mathbf{x} \in A$ ) such that  $D$  satisfies the Leibniz law:

$$D(\mathbf{x}\mathbf{y}) = \mathbf{x}D(\mathbf{y}) + D(\mathbf{x})\mathbf{y}. \quad (\text{D.0.10})$$

The space of derivations,  $\text{Der}(A)$ , is a Lie algebra with the Lie bracket given by  $[D_1, D_2](\mathbf{x}) = D_1(D_2(\mathbf{x})) - D_2(D_1(\mathbf{x}))$ .

**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,  $\varphi : \mathfrak{g} \rightarrow \mathfrak{h}$ , which preserves the Lie bracket, so

$$\varphi([x, y]) = [\varphi(x), \varphi(y)]. \quad (\text{D.0.12})$$

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,  $g \cdot h = L_g(h) = gh$ . 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 the 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  $k$ th tensor power of  $V$  to be the tensor product of  $V$  with itself  $k$  times:

$$T^k V = V^{\otimes k} := \underbrace{V \otimes V \otimes \cdots \otimes V}_{k \text{ times}}. \quad (\text{D.2.2})$$

That is  $T^k V$  consists of all tensors on  $V$  of order  $k$ . Which is to say we can think of  $T^k V$  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  $T^0 V = \mathbb{F}$  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

$$T(V) = \bigoplus_{k=1}^{\infty} T^k V = \mathbb{F} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots. \quad (\text{D.2.3})$$

Multiplication in  $T(V)$  is given by the obvious isomorphism  $T^k V \otimes T^l V \rightarrow T^{k+l} V$ , given by

$$(a_1 \otimes \cdots \otimes a_k)(b_1 \otimes \cdots \otimes b_l) = a_1 \otimes \cdots \otimes a_k \otimes b_1 \otimes \cdots \otimes b_l \quad (\text{D.2.4})$$

**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

$$a \otimes b - b \otimes a = [a, b]. \quad (\text{D.2.6})$$

We then define the bracket on  $\mathfrak{g} \otimes \cdots \otimes \mathfrak{g}$  recursively by defining

$$[a \otimes b, c] = a \otimes [b, c] + [a, c] \otimes b \quad (\text{D.2.7})$$

and

$$[a, b \otimes c] = [a, b] \otimes c + b \otimes [a, c]. \quad (\text{D.2.8})$$

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

$$U(\mathfrak{g}) := T(\mathfrak{g}) / \sim \quad (\text{D.2.9})$$

where  $a \sim b$  if

$$a \otimes b - b \otimes a = [a, b]. \quad (\text{D.2.10})$$

Note that we trivially have  $a \sim b$  for  $a, b \in \mathfrak{g}$  by definition of  $[-, -]$ . However, if  $a, b \in \mathfrak{g} \otimes \mathfrak{g}$ , we don't necessarily have that  $a \sim b$ .

---

# Index

---

## Symbols

$-^*$ , complex conjugate, 156  
 $-^*$ , nonzero elements, 7  
 $-\dagger$ , Hermitian conjugate, 156  
 $-\top$ , transpose, 156  
 $-^{-1}$ , inverse, 157  
 $|\cdot|$ , determinant, 158  
 $\langle - | - \rangle$ , inner product, 154  
 $\langle - |$ , bra, 154  
 $\otimes$ , direct product, 159  
 $\oplus$ , direct sum, 158  
 $\mathbb{I}$ , identity matrix, 155  
 $\langle -, - \rangle$ , inner product, 152  
 $| - \rangle$ , ket, 154  
 $\| - \|$ , norm, 153

## A

Abelian, 8  
 accociative, 152  
 adjoint representation, 107  
 algebra, 174  
 alternating group, 47  
 alternating representation, 74  
 $A_n$ , alternating group, 47  
 anti-Hermitian, 156  
 antilinear, 78, 154  
 antisymmetric, 156  
 associated Legendre polynomial, 138  
 associative, 2  
 associative algebra, 174

## B

Baker–Campbell–Hausdorff formula, 104  
 basis, 154  
 bijective, 149  
 binary  
   operation, 2  
   relation, 150

block diagonal, 158  
 bounded, 94  
 bra, 154  
 bra-ket notation, 154  
 branching rules, 80

## C

Casimir element, 109  
 Cayley table, 10  
 Cayley’s Theorem, 12  
 centraliser, 27  
 centre, 9  
 character, 60  
   table, 67  
 characteristic polynomial, 156  
 charge symmetry, 124  
 chart, 168  
 class function, 60  
 Clebsch–Gordan  
   coefficient, 128  
   series, 84  
 closed, 2, 94  
 commutative, 152  
 commutator, 59  
 compact Lie group, 94  
 compatible  
   atlases, 170  
   charts, 169  
 complex  
   conjugate, 156  
   Lie group, 94  
 complex representation, 56, 77  
 composite Lie group, 96  
 Condon–Shortley phase convention, 129  
 conjugate, 8  
 connected Lie group, 95  
 continuous, 168  
 continuous group, 7, 91

coset, 21  
 cycle, 11, 42  
 cycle type, 73  
 cyclic, 9  
   group, 7, 161

**D**

decomposability theorem, 59  
 decomposition, *see* partition  
 derivation, 172  
 det, determinant, 158  
 determinant, 158  
 diagonal, 158  
 diffeomorphism, 171  
 dihedral group, 40  
 dim, dimension, 155  
 dimension, 155, 168  
   of a group, 91  
 direct  
   product, 37, 159  
   sum, 158  
 disconnected Lie group, 95  
 discrete group, 7  
 double cover, 111  
 Dynkin diagram, 96  
 Dynkin index, 106

**E**

eigenvalue, 156  
 eigenvector, 156  
 empty set, 150  
 endomorphism, 15  
 equivalence class, 151  
 equivalence of representations, 55  
 equivariant, 78  
 even permutation, 47

**F**

factor group, *see* quotient group  
 faithful representation, 55  
 Fermi's golden rule, 123  
 finite dimensional Lie group, 93  
 finite group, 7  
 first isomorphism theorem, 34  
 function composition, 3  
 fundamental representation, 118

**G**

general linear group, 31, 92, 163  
 generate, 9

generator, 9, 104  
 $GL(n, \mathbb{F})$ , general linear group, 163  
 Gram–Schmidt inner product, 158  
 group, 5, 15  
   action, 24  
   axioms, 5  
   presentation, 20

**H**

$\mathbb{H}$ , quaternions, 21  
 Haar measure, 94  
 Hamiltonian, 121  
 Hausdorff space, 169  
 Heisenberg commutation relation, 122  
 Hermitian, 156  
   conjugate, 156  
 highest weight representation, 111  
 Hilbert space, 152  
 homeomorphic, 168  
 homeomorphism, 168  
 homomorphism  
   of algebras, 174  
   of groups, 16  
   of Lie algebras, 175  
 hook number, 82

**I**

identity, 4  
 image, 150  
 index, 22  
 infinite dimensional Lie group, 93  
 injective, 149  
 inner product, 152  
 integer partition, 73  
 intrinsic angular momentum, *see* spin  
 invariant subgroup, *see* normal subgroup  
 invariant subspace, 55  
 inverse, 5, 157  
 irreducible representation, 55  
 irrep, *see* irreducible representation  
 $ISO(1, 3)$ , Poincaré group, 164  
 isometries of Euclidean space, 39, 164  
 isomorphic, 16  
 isomorphism  
   of algebras, 174  
   of groups, 16  
   of Lie algebras, 175

**J**

Jacobi identity, 107, 108

## K

kernel, 149  
ket, 154  
Killing metric, 106  
Klein *Vierergruppe*, 18, 38, 161  
Kronecker delta, 150

## L

ladder operator, 112  
Landé  $g$ -factor, 136  
Laplace–Runge–Lenz vector, 144  
Levi-Civita Symbol, 150  
Lie  
    algebra, 99, 108, 175  
    bracket, 108, 175  
    group, 91  
linearly independent, 154  
little Wigner  $d$ -matrix, 115  
locally Euclidean, 168  
locally isomorphic, 110  
Lorentz group, 124, 164  
lowering operator, 112

## M

magma, 13  
manifold, 166  
    smooth, 171  
    topological, 169  
Maschke’s theorem, 56  
matrix element, 156  
Molien function, 134  
monoid, 14  
multiplication table, *see* Cayley table

## N

natural projection, 35  
non-compact Lie group, 94  
norm, 153  
normal subgroup, 30  
 $n\mathbb{Z}$ , integer multiples of  $n$ , 17

## O

$O(1, 3)$ , Lorentz group, 164  
 $O(3)$ , rotation group, 7  
 $O(n)$ , orthogonal group, 163  
odd permutation, 47  
open set, 167  
orbit, 26

orbital angular momentum operator,  
    126

order

    element, 8  
    group, 8

orthogonal

    group, 92, 163  
    matrix, 157  
    vectors, 155

## P

parastatistics, 89  
parity symmetry, 124, 141  
partition, 21  
partition of the identity, 155  
Pauli matrices, 78  
permutation, 7, 42  
    group, 7, 161  
    representation, 54  
Platonic solid, 49  
Poincaré group, 124, 164  
preimage, 168  
projection operator, 155  
pseudo-real representation, 77, 78  
pseudo-vector, 142

## Q

$Q$ , quaternion group, 161  
quadratic Casimir, 109  
quaternion group, 21, 161  
quaternionic representation, *see* pseudo-  
    real representation  
quotient group, 32

## R

raising operator, 112  
rank, 9, 105  
real  
    Lie group, 94  
    projective space, 110  
    representation, 56, 77, 78  
reduced matrix element, 129  
reducible representation, 55  
reflexive, 150  
regular representation, 66  
relation, 150  
representation, 53  
Reynolds operator, 134  
rotation group, 7  
 $\mathbb{RP}^n$ , real projective space, 110

## S

Schrödinger equation, 121  
 Schur's lemma, 57  
 second-countable, 169  
 selection rule, 130  
 semi-simple Lie group, 96  
 semidirect product, 39  
 semigroup, 14  
 similarity transform, 55  
 simple group, 34, 96  
 simply connected Lie group, 95  
 simply reducible, 115, 128  
 $SL(n, \mathbb{F})$ , special linear group, 163  
 smooth  
   atlas, 170  
   function, 171  
   manifold, 171  
   map, 171  
   structure, 170  
 smoothly compatible, 170  
 $S_n$ , permutation group, 7  
 $SO(n)$ , special orthogonal group, 163  
 solid harmonic, 139  
 $Sp(n)$ , symplectic group of rank  $n$ , 92  
 special linear group, 31, 92, 163  
 special orthogonal group, 92, 163  
 special unitary group, 92, 164  
 spherical harmonic, 137  
 spin, 126  
 spin statistics theorem, 89  
 stabiliser, 26  
 standard representation, 75  
 standard topology, 167  
 structure constants, 105  
 $SU(n)$ , special unitary group, 164  
 subgroup, 8  
   proper, 8  
 superselection rule, 142  
 surjective, 149  
 symmetric, 156  
   relation, 150  
 symmetric group, 42  
 symplectic  
   bilinear form, 93  
   group, 92

## T

tangent  
   bundle, 172  
   space, 172

## tensor

  algebra, 176  
   operator, 128  
   product, 84  
 time symmetry, 124  
 topological  
   basis, 167  
   manifold, 169  
   space, 167  
 topology, 167  
 total angular momentum operator, 126  
 tr, trace, 158  
 trace, 158  
 transition maps, 169  
 transitive, 150  
 transpose, 156  
 transposition, 43  
 trivial  
   group, 8, 160  
   representation, 55

## U

$U(n)$ , unitary group, 25, 163  
 unfaithful representation, 55  
 unitary, 157  
   group, 25, 92, 163  
   representation, 55  
 universal  
   covering group, 105  
   enveloping algebra, 108, 177  
 $USp(2n)$ , symplectic group of rank  $n$ , 92

## V

Vandermonde polynomial, 47  
 vector representation, 128

## W

Weyl homomorphism, 111  
 Wigner  $D$ -matrix, 114

## Y

Young tableau, 82

## Z

$Z(G)$ , centre, 9  
 $\mathbb{Z}_2 \times \mathbb{Z}_2$ , Klein *Vierergruppe*, 18  
 $\mathbb{Z}_n$ , cyclic group, 7