Willoughby Seago

Theoretical Physics

Symmetries of Particles and Fields

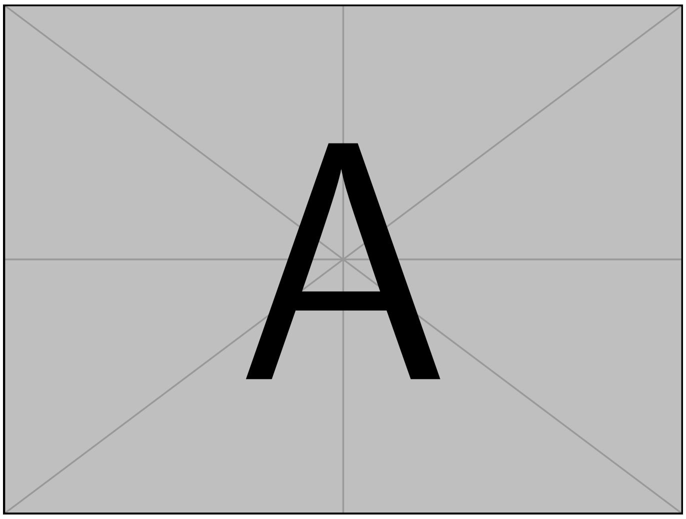
Course Notes

Symmetries of Particles and Fields

Willoughby Seago

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

These notes were last updated at 22:40 on October 3, 2022. For notes on other topics see https://github.com/WilloughbySeago/Uni-Notes.



Chapters

Page

Contents

Page

List of Figures

One

Introduction

Symmetries are incredibly important in physics. Both approximate and exact symmetries highly restrict the forms that our equations can take, to the point that often the correct equation can be written down from symmetries alone. The mathematical language expressing symmetries is group theory. In particular we are interested in continuous symmetries, which are modelled by Lie groups, and we can get a better understanding of Lie groups by studying the associated Lie algebras.

The following is a selection of symmetries we'll discuss in this course, to demonstrate how broad a concept symmetry is:

- Spacetime translations and rotations, these give rise to conservation of energy, momentum, and angular momentum.
- Internal symmetries, these give rise to conservation of electric charge, and particle number to give a couple of examples.
- Charge conjugation, parity, and time reversal symmetries.
- Homogeneity and isotropy of the universe on a large scale, these allow us to write down a metric tensor which then describes the curvature of spacetime across the universe.
- Approximate internal symmetries, such as nuclear isospin and flavour SU(3).
- General coordinate invariance (or invariance under diffeomorphisms) leads to an interpretation of general relativity as a gauge theory.

The course is structured as follows:

1. We introduce Lie groups and Lie algebras.
2. We study spacetime symmetries and look at the actions that we can form obeying these symmetries.
3. We study compact groups and their representations.
4. We look at applications in particle physics, including Noether's theorem, isospin, the quark model, spontaneous symmetry breaking, chiral symmetry, gauge theories, QCD, electroweak theory, and the Higgs mechanism.
5. We loo at applications in cosmology.

Part I

Lie Groups

Two

Groups and Lie Groups

What's the difference between theoretical physics and maths? The distinction is that thoretical physicists rarely prove anything, we cheat by looking at nature.

Neil Turok

2.1 Groups

R For more details see the *Symmetries of Quantum Mechanics* course.

A group is the mathematical language used to describe symmetries. Roughly speaking a symmetry is something that we can do to something else, such that the second thing is unchanged. We should expect that chaining together symmetries should behave nicely. In particular, it should be possible to do nothing, and it should be possible to undo anything we do. Abstracting this gives us the definition of a group.

Definition 2.1.1 — Group A **group**, $(G, \cdot)$, is a set, G, along with an associative binary operation, $\cdot \colon G \times G \to G$, such that there is an identity element and every element has an inverse.
Associativity means that for all $g, h, k \in G$ we have

$$g \cdot (h \cdot k) = (g \cdot h) \cdot k. \tag{2.1.2}$$

The **identity** element is the distinguished element $1 \in G$ such that

$$1 \cdot g = g \cdot 1 = g \tag{2.1.3}$$

for all $g \in G$. For each $g \in G$ we have some $g^{-1} \in G$, called the **inverse** of g, such that

$$g \cdot g^{-1} = g^{-1} \cdot g = 1. \tag{2.1.4}$$

Definition 2.1.5 — Abelian A group, G, is **Abelian** if $gh = hg$ for all $g, h \in G$.

Notation 2.1.6 We usually refer to a group, G, rather than a group $(G, \cdot)$, with the operation either implicit or stated separately.
We usually write the operation as juxtaposition, so gh instead of $g \cdot h$. We may use multiple symbols for different group operations, such as $+$, $*$, or $\circ$. We then write the inverse and identities with the appropriate notation, so $-g$ for g inverse if we use $+$, and 0 for the identity.

Example 2.1.7 — Groups

- The single element set, $\{e\}$, is a group if we define $e \cdot e = e$. This is called the **trivial group**.
- The symmetries of a cube form a group, with the operation being concatenation of symmetries.
- The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ all form groups under addition.
- The sets $\mathbb{Q}^\times$, $\mathbb{R}^\times$, and $\mathbb{C}^\times$ all form groups under multiplication where $S^\times = S \setminus \{0\}$ denotes the set S with 0 removed.
- Invertible $n \times n$ matrices over some ring, R, form a group under matrix multiplication.
- The set $\mathbb{Z}_n = C_n = \{0, 1, \dots, n-1\}$ forms a group under addition modulo n. This is called the **cyclic group** of order n.
- The set of all permutations on n elements forms a group, denoted S_n. This is called the **symmetric group** on n objects.
- The set of invertible functions, $A \to A$, forms a group under function composition.
- The set of translations of space forms a group under repeated application of translations.
- The set of rotations of space forms a group under repeated application of rotations.

Definition 2.1.8 — Subgroup Let G be a group, and let H be a subset of G. We say that H is a **subgroup** of G, and write $H \subseteq G$ if H is a group under the operation of G.
If $H \neq G$ we say that H is a **proper subgroup** of G, and we write $H \subset G$.
If $H \neq \{e\}$ then we say that H is a of G.

Definition 2.1.9 — Order The order of a finite group, that is a group with a finite number of elements, is the number of elements. We write $|G|$ for the order of the group G.

Example 2.1.10

- The order of the trivial group is $|\{e\}| = 1$.
- The order of C_n is $|C_n| = n$.
- The order of S_n is $|S_n| = n!$.

We usually think of symmetries, and therefore groups, as acting on something. The group definition doesn't tell us how the group elements act on an object (mathematical or physical), this will come later. We typically don't distinguish much between a group acting on something, and a stand alone group with nothing to act on, since the later is not that interesting.

Definition 2.1.11 — Group Action Given a group, G, and a set, S, we define a **left group action** as a function, $\varphi : G \times S \to S$, such that

- $\varphi(1, s) = s$ for all $s \in S$; and
- $\varphi(g, \varphi(h, s)) = \varphi(gh)$.

Writing $\varphi(g, s) = g \,.\, s$ these requirements become $1 \,.\, s = s$ and $g \,.\, (h \,.\, s) = (gh) \,.\, s$.

Example 2.1.12 — Group Actions The group $\mathbb{R}^3$ of three-dimensional real vectors under vector addition acts on itself by translation:

$$\boldsymbol{a} \,.\, \boldsymbol{x} = \boldsymbol{x} + \boldsymbol{a}. \tag{2.1.13}$$

The group S_n acts on $(1, \dots, n)$ by permutation, say $\sigma \in S_n$ swaps the first and second element, then

$$\sigma \,.\, (1, 2, 3, \dots, n) = (2, 1, 3, \dots, n). \tag{2.1.14}$$

The group $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ acts on $\mathbb{R}^2$ by rotations by angles 0, $\pi/2$, π, and $3\pi/2$, corresponding to 0, 1, 2, and 3 respectively.
The group $C_4 = \{1, i, -1, -i\}$, acts on $\mathbb{C}$ by complex multiplication:

$$z \,.\, w = zw \tag{2.1.15}$$

where $z \in C_4$ and $w \in \mathbb{C}$. However, notice that we can interpret $\mathbb{C}$ as a plane and rotation by multiples of i as rotations by multiples of $\pi/2$, so really these two groups are the same.

Often two groups may look different, as in the case of $\mathbb{Z}_4$ and C_4 above, but really they are the same on the level of the group structure. This is expressed through the notion of isomorphisms.

Definition 2.1.16 — Morphisms Let G and H be groups. A **group homomorphism** is a function $\varphi : G \to H$ such that $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$ for all $g_1, g_2 \in G$. If this function is invertible it is called a **group isomorphism**. If there exists an isomorphism $G \to H$ then we say G and H are **isomorphic**, and write $G \cong H$.

So, when we say that $\mathbb{Z}_4$ and C_4 are the same, we really mean that the mapping defined by $0 \mapsto 1$, $1 \mapsto i$, $2 \mapsto -1$, and $3 \mapsto -i$ is a group isomorphism, and $\mathbb{Z}_4 \cong C_4$. In fact, this is a slightly cheaty example as $\mathbb{Z}_n \cong C_n$ where $C_n = \{e^{2\pi i m/n} \mid m = 0, \dots, n-1\}$. Alternatively we can define $C_n = \mathbb{Z}_n$, and then say that the complex exponentials before are a representation of the group.

When two groups are isomorphic, as far as group theory is concerned, they are equivalent and interchangeable.

Example 2.1.17

- The group S_2 is isomorphic to the group C_2.
- The group of symmetries of the triangle is isomorphic to S_3.
- All groups are isomorphic to some subgroup of S_n for some n.
- All groups of order 1 are isomorphic, since they have a single element and that element must be the identity.

Definition 2.1.18 — Direct Product Let G and H be groups. Then there exists a group $G \times H$ formed from pairs (g, h) with $g \in G$ and $h \in H$ with the group operation just elementwise application of G and H's operations:

$$(g, h)(g', h') = (gg', hh'). \tag{2.1.19}$$

Lemma 2.1.20 The direct product as defined above is a group.

Proof. Let G and H be groups and $G \times H$ their direct product. Then for all $g, g', g'' \in G$ and $h, h', h'' \in H$ we have

$$\begin{aligned}
(g, h)[(g', h')(g'', h'')] &= (g, h)(g'g'', h'h'') && (2.1.21)\\
&= (g(g'g''), h(h'h'')) && (2.1.22)\\
&= ((gg')g'', (hh')h'') && (2.1.23)\\
&= (gg', hh')(g'', h''') && (2.1.24)\\
&= [(g, h)(g', h')](g'', h''). && (2.1.25)
\end{aligned}$$

So, the operation is associative. Let $1_G \in G$ and $1_H \in H$ be the identities in their respective groups, then

$$(1_G, 1_H)(g, h) = (1_G g, 1_H h) = (g, h) = (g 1_G, h 1_H) = (g, h)(1_G, 1_H), \tag{2.1.26}$$

so $1_{G\times H} = (1_G, 1_H)$ is the identity in $G\times H$. Finally, take $g \in G$ and $h \in H$. Then we have

$$(g,h)(g^{-1},h^{-1}) = (gg^{-1},hh^{-1}) = (1_G,1_H) = 1_{G\times H}, \tag{2.1.27}$$

so $G \times H$ has inverses. $\square$

Example 2.1.28

- Considering $\mathbb{R}$ as a group under addition we have $\mathbb{R}\times\mathbb{R} = \{(x,y) \mid x,y \in \mathbb{R}\}$ with addition defined as $(x,y)+(x',y') = (x+x', y+y')$, so we identify this with the plane, $\mathbb{R}^2$, under vector addition.
- The product $\mathbb{Z}_3\times\mathbb{Z}_2$ is the set $\{(n,m) \mid n = 0,1,2 \text{ and } m = 0,1\}$. This group is generated[a] by $(1,1)$, and $(1,1)$ has order 6, so this group is isomorphic to $\mathbb{Z}_6$. In fact, if p and q are relatively prime them $\mathbb{Z}_p \times \mathbb{Z}_q \cong \mathbb{Z}_{pq}$.
- The first noncyclic group is the Klein four-group, $V = \mathbb{Z}_2 \times \mathbb{Z}_2$.
- If G is a group and $\{1\}$ is the trivial group then $G\times\{1\} \cong \{1\}\times G \cong G$, so $\{1\}$ acts as an identity for the direct product.

[a] A group, G, is **generated** by some $g \in G$ if all $h \in G$ are of the form g^n for some $n \in \mathbb{N}$, where $g^n := gg\cdots g$ with n factors and $g^0 := 1$. We call the smallest such n the **order** of h.

As well as direct products, which just pair up elements in a non-interacting way, we can act on one of the groups with the other, and then pair up the elements. This gives a new type of product, and so allows us to construct new groups.

Definition 2.1.29 — Semidirect Product Let N and H be groups, and let consider a group action of N on H. We can define a group, $N \rtimes H$, consisting of pairs (n,h) with $n \in N$ and $h \in H$, and with the group operation

$$(n,h)(n',h') = (n(h\,.\,n'), hh'). \tag{2.1.30}$$

This is called the **semidirect product**.

That is, we act on the second element of N with the first element of H, and then proceed with the usual direct product.

Example 2.1.31 — Dihedral Group The dihedral group, D_n, of order $2n$ is the semidirect product of C_n and C_2 where C_2 acts on C_n by the nonidentity element of C_2 sending elements of C_n to their inverses. It can be shown that this is the symmetry group of the regular n-gon, where we interpret the action of C_n on the n-gon as rotations by $2\pi/n$ and the action of C_2 as reflections.

Groups can act on sets, and groups *are* sets, so it makes sense to have groups act on groups, and in particular, its pretty obvious we can have a group act on itself. One example is the left action of G on G, given by $g\,.\,h = gh$ for all $g,h \in G$. Here we interpret this as g acting on h. A slightly less trivial example is **conjugation**, where $g \,.\, h = ghg^{-1}$ for all $g, h \in G$. We can view conjugacy as a symmetry of the group, but what does it act on? Well, we can just act on the group, but this isn't that helpful. Instead, we define the **conjugacy class** of some $g \in G$ to be the set of all $h \in G$ conjugate to G, that is, all h such that there exists $k \in G$ with $khk^{-1} = g$. This partitions G into sets, since this defines an equivalence relation.

Now that we have a particular symmetry of G we should ask if there are any invariants, and indeed there is.

Definition 2.1.32 — Normal Subgroup Let G be a group and N a subgroup of G. Then we call N a **normal subgroup**, or **invariant subgroup** if N is invariant under conjugation. That is if $n \in N$ then $gng^{-1} \in N$ for all $g \in G$.

Example 2.1.33 — Normal Subgroups

- The trivial group, consisting of just the identity, is a normal subgroup of any group.
- The **centre of a group**, defined as the set of all elements which commute with every element of the group is a normal subgroup. If the only normal subgroups are the trivial group and the centre then we call the group a **simple**.
- The group $\{(), (1\,2\,3), (3\,2\,1)\}$ is a normal subgroup of S_3.

We've seen products of groups, we now ask if there is a sensible notion of a quotient of groups, and indeed there is! The definition only makes sense if one of the groups is a normal subgroup of the other. But first, we need another definition.

Definition 2.1.34 — Coset Let G be a group and H a subgroup of G. For each $g \in G$ we define the **left coset** to be $gH \coloneqq \{gh \mid h \in H\}$. That is, it's the set of all elements of the form gh for $h \in H$. Similarly the **right coset** is $Hg \coloneqq \{hg \mid h \in H\}$.

! Cosets are not, in general, groups, since only the coset H contains the identity.

Example 2.1.35 — Cosets

- Let $\mathbb{Z}$ be the group of integers under addition. Then $2\mathbb{Z}$ is the subgroup of even integers under addition. We can define cosets, $0 + 2\mathbb{Z} = 2\mathbb{Z}$ and $1 + 2\mathbb{Z}$ corresponding to the even integers and odd integers.
- Consider $\mathbb{R}^3$, then we can view $\mathbb{R}$ as a subgroup of this given by the x-axis. Then the cosets $\boldsymbol{a} + \mathbb{R}$ for $\boldsymbol{a} \in \mathbb{R}$ consist of lines parallel to the x-axis.

Definition 2.1.36 — Quotient Group Let G be a group, and N a normal subgroup of G. Define G/N to be the set of all left cosets in G, that $G/N := \{aN, a \in G\}$. We can define a group operation on this set: $(gN)(g'N) := (gg')N$. We then call G/N the **quotient group** of G by N, also known as the **factor group**.

! This definition only defines a group if N is a normal subgroup of G. If this isn't the case then the operation will not be well defined.

Definition 2.1.37 — Quotient Groups

- Consider the group of integers, $\mathbb{Z}$, under addition, and the normal subgroup $2\mathbb{Z}$, of even integers. The quotient group $\mathbb{Z}/2\mathbb{Z}$ consists of the cosets $2\mathbb{Z}$ and $1 + 2\mathbb{Z}$, where the second is all odd integers. Considering $2\mathbb{Z}$ to be $0 + 2\mathbb{Z}$ we see that, for example,

  $$(1 + 2\mathbb{Z})(0 + 2\mathbb{Z}) = (1 + 0) + 2\mathbb{Z} = 1 + 2\mathbb{Z} \tag{2.1.38}$$

  and

  $$(1+2\mathbb{Z})(1+2\mathbb{Z}) = (1+1)+2\mathbb{Z} = 2+2\mathbb{Z} = 0+2\mathbb{Z} = 2\mathbb{Z}. \tag{2.1.39}$$

  We can conclude that $\mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}_2$, and indeed some people use $\mathbb{Z}/n\mathbb{Z}$ to denote $\mathbb{Z}_n$.

- Consider the real numbers, $\mathbb{R}$, viewed as an additive group, with the subgroup $\mathbb{Z}$. Each coset of $\mathbb{Z}$ is then of the form $a + \mathbb{Z}$ where $a \in \mathbb{R}$. When the noninteger parts of a and b are equal the cosets $a + \mathbb{Z}$ and $b + \mathbb{Z}$ are equal, and so we can subtract the integer parts of a and b and just consider the noninteger parts, identifying the groups with the noninteger part of their representative we notice that $a + \mathbb{Z}$ is essentially covering the interval $[0, 1)$ over and over again as a is varied. In fact, $\mathbb{R}/\mathbb{Z} \cong S^1$, identifying $a + \mathbb{Z}$ with a rotation by $2\pi\tilde{a}$, where $\tilde{a}$ is the noninteger part of a.

- The group G/G is the trivial group, and the group $G/\{e\}$ is (isomorphic to) G.

2.2 Lie Groups

Often the groups of interest in physics have an uncountable number of elements, in this case they are continuos groups. We can define this more rigorously by parametrising the elements in a continuous way.

Definition 2.2.1 — Continuous Groups A **continuous group**, G, is a group with an infinite number of elements parametrised by a (possibly infinite) set of parameters, $\{\alpha\} = \{\alpha_1, \alpha_2, \dots\}$, such that each element of G can be written as a function of these parameters, $g(\alpha) = g(\alpha_1, \alpha_2, \dots)$.

The **dimension** of G is the number of independent parameters.

Example 2.2.2 Spatial translations of $\mathbb{R}^3$ form a three dimensional continuous group. The translation $\boldsymbol{x} \to \boldsymbol{x} + \boldsymbol{a}$ is parametrised by $\{a_1, a_2, a_3\}$. Spatial rotations of $\mathbb{R}^3$ form a three dimensional continuous group. Consider the rotation $x^i \to R^i{}_j x^j$. For a rotation we know that $R^i{}_j R^k{}_l \delta_{ik} = \delta_{jl}$, or $R^\mathsf{T} R = 1$. A rotation in three dimensions can be parametrised by 3 parameters, two used to pick out a unit vector in $\mathbb{R}^3$, the third component being fixed by normalisation, and one parameter to give the angle of rotation about said unit vector.

As is usually the case in physics we assume that most things are analytic, they can be expanded in Taylor series. This assumption defines Lie groups, one of the main subjects of our study in this course.

Definition 2.2.3 — Lie Group A **Lie group**[a], G, is a continuous group for which the group multiplication has an analytic structure. That is, if $g(\alpha) = g(\beta)g(\gamma)$ then $\alpha_i = \varphi_i(\beta, \gamma)$ where φ_i are analytic functions of β and γ.

[a] pronounced *lee*

Formally we say that the parameter space for a Lie group is a smooth manifold and multiplication is given by a smooth function on this manifold.

Continuous groups can be either compact or noncompact, depending on the structure of the parameter space. For our purposes a **compact space** is one which is closed and bounded, that is it contains its boundary and the parameters are restricted in their size. For example, $\mathbb{R}$ is noncomapct, since it isn't bounded, the intervals $[0, 1)$, $(0, 1]$, and $(0, 1)$ are noncompact as they don't contain their boundaries, and $[0, 1]$ is compact. The sphere, S^n, is compact.

For example, the group of translations of $\mathbb{R}^3$ is noncompact, since we can have infinite translations. The group of rotations of $\mathbb{R}^3$ on the other hand is compact, the parameters defining the axis are constrained to $[0, 1]$ and the parameters defining the angle are constrained to $[0, 2\pi)$, which looks like it isn't closed but actually we identify 0 and 2π as the same rotation, so really this is drawing from the circle, S^1, which is closed. The parameter space is then $[0, 1]^2 \times S^1$, which is compact, since the product of compact spaces is compact.

2.3 Metric Spaces

R This section is concerned with the notion of a metric on a real vector space. There is a more general mathematical notion of a metric on a topological space which we shall not discuss here, but when restricted to $\mathbb{R}^n$ these notions coincide.

Consider some finite dimensional real vector space, V, and fix some basis. Take a vector with components x^μ where $\mu = 1, \dots, N$ where $N = \dim V$. Alternatively, particularly when doing relativity, we may take $\mu = 0, \dots, N-1$, in which case we interpret x^0 as the time. A **metric**, $g_{\mu\nu}$, is a real, symmetric, $N \times N$ matrix. We define the length of a vector according to

$$\|x\|^2 = x^2 = g_{\mu\nu} x^\mu x^\nu, \tag{2.3.1}$$

where summation over μ and ν is implied by the Einstein summation convention, which we follow in this course: repeated indices appearing once raised and once lowered are summed over unless otherwise specified.

It is possible to choose a basis such that:

- $g_{\mu\nu}$ is diagonal, that is $g_{\mu\nu} = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$ where $\lambda_i \in \mathbb{R}$ are the eigenvalues of $g_{\mu\nu}$; and
- the eigenvalues are either 0 or ± 1. Reordering the basis as required allows us to put the metric in the canonical form

$$g_{\mu\nu} = \operatorname{diag}(1, \dots, 1, 0, \dots, 0, -1, \dots, -1). \tag{2.3.2}$$

We call this the **canonical basis**.

The first point, that $g_{\mu\nu}$ is diagonalisable, is not that surprising. The second point, that the eigenvalues are restricted to $\{0, \pm 1\}$, is worth explaining. Suppose that $g_{\mu\nu} = \operatorname{diag}(\lambda, \dots)$ where $\lambda \neq 0$. We can rescale $x^1 \mapsto sx^1$. Then, $g_{\mu\nu} \mapsto \operatorname{diag}(\lambda/s^2, \dots)$, in order for x^2 to remain invariant, since using the diagonal nature of $g_{\mu\nu}$ we have

$$\begin{aligned} x^2 = g_{\mu\nu}x^\mu x^\nu &= \lambda x^1 x^1 + \sum_{i=2}^{N} g_{ii} x^i x^i \\ &\mapsto \frac{\lambda}{s^2} sx^1 sx^1 + \sum_{i=2}^{N} g_{ii} x^i x^i = \lambda x^1 x^1 + \sum_{i=2}^{N} g_{ii} x^i x^i = x^2. \end{aligned} \tag{2.3.3}$$

So, by choosing $s = \sqrt{|\lambda|}$ we can scale $g_{\mu\nu}$ such that

$$g_{\mu\nu} = \operatorname{diag}(\operatorname{sgn}(\lambda), \dots) \tag{2.3.4}$$

where

$$\operatorname{sgn}(\lambda) = \begin{cases} 1 & \lambda > 0, \\ -1 & \lambda < 0. \end{cases} \tag{2.3.5}$$

Doing this same process for all nonzero eigenvalues of the original $g_{\mu\nu}$ we can scale all diagonal elements to be 0 or ± 1. Reordering the basis then gives us the canonical form.

Example 2.3.6 — Metrics Euclidean space, $\mathbb{R}^n$, has the metric $g_{ij} = \delta_{ij}$. In particular, when $n = 3$ we have

$$x^2 = (x^1)^2 + (x^2)^2 + (x^3)^2, \tag{2.3.7}$$

which is just the usual length-squared of a vector, $\boldsymbol{x} = (x^1, x^2, x^3)$. Minkowski space, $\mathbb{R}^{1,3}$, has the metric $g_{\mu\nu} = \operatorname{diag}(1, -1, -1, -1)$, so

$$x^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2, \tag{2.3.8}$$

which is just the usual scalar product of two four-vectors.

! The choice of $g_{\mu\nu} = \operatorname{diag}(-1, 1, 1, 1)$ is also common, leading to lots of annoying sign discrepancies. In fact, given any metric, g, the metric $-g$ is equivalent.

Transformations preserving the metric in its canonical form are the symmetries that we are most interested in. This makes sense when you think about what we use the metric for. In Euclidean space the metric defines lengths, something that shouldn't change under symmetries. In special relativity transformations preserving the metric don't preserve lengths, at least in the traditional sense, they instead preserve the speed of light. In quantum mechanics the metric used to define the inner product between state vectors being preserved means that probability is preserved by metric preserving transformations.

Definition 2.3.9 — Types of Metric If the metric has only positive eigenvalues then we say that it is a **positive definite** metric. If the metric has positive and negative eigenvalues then we say that it is **indefinite**.

An indefinite metric allows for x^2 to be positive, negative or zero, contrary to our normal intuition about length squared.

The **metric signature** is either a pair or triple of natural numbers giving the number of 1s, -1s, and 0s in the metric. For example, the the Euclidean metric on $\mathbb{R}^n$ is $(n, 0, 0)$ or $(n, 0)$, where no third number is taken to mean no zero eigenvalues. So the Minkowski metric is $(1, 3)$, or $(1, 3, 0)$.

If we have a nonsingular metric, that is $\det g \neq 0$, then we can use the metric, and its inverse, to raise and lower indices. Define the inverse metric $g^{\mu\nu}$ to be such that

$$g^{\mu\rho} g_{\rho\nu} = \delta^{\mu}{}_{\nu}. \tag{2.3.10}$$

Then we define

$$x_\mu = g_{\mu\nu} x^\nu, \quad T_{\mu\nu} = g_{\mu\rho} g_{\nu\sigma} T^{\rho\sigma}, \tag{2.3.11}$$

and so on for more indices.

After lengths the next most important quantity we can define is volumes. Recall that a parallelepiped with sides $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3$ has volume

$$|\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})| = |a^i b^j c^k \varepsilon^{ijk}|. \tag{2.3.12}$$

This suggests that we can use the Levi-Civita symbol to generalise volumes. Define the **Levi-Civita symbol** on n indices to be

$$\varepsilon_{\mu_1\mu_2\dots\mu_n} := \begin{cases} +1 & \text{if } \mu_1\mu_2 \dots \mu_n \text{ is an even permutation of } 1 \dots n, \\ -1 & \text{if } \mu_1\mu_2 \dots \mu_n \text{ is an odd permutation of } 1 \dots n, \\ 0 & \text{otherwise.} \end{cases} \tag{2.3.13}$$

For a nonsingular metric we can then define the Levi-Civita symbol with raised indices:

$$\varepsilon^{\mu_1\mu_2\cdots\mu_n} :=^{\mu_1\nu_1} g^{\mu_2\nu_2} \cdots g^{\mu_n\nu_n} \varepsilon_{\nu_1\nu_2\dots\nu_n}. \tag{2.3.14}$$

We can use this to define the determinant of a matrix, A:

$$\det A := \varepsilon_{\mu_1\mu_2\dots\mu_n} A^{\mu_1}{}_1 A^{\mu_2}{}_2 \cdots A^{\mu_n}{}_n \tag{2.3.15}$$

In a general metric the infinitesimal volume element, which is invariant under metric preserving transformations, is

$$\mathrm{d}V := \sqrt{|\det g|}\varepsilon_{\mu_1\mu_2\dots\mu_n}\,\mathrm{d}x^{\mu_1}\,\mathrm{d}x^{\mu_2}\cdots\mathrm{d}x^{\mu_n}. \tag{2.3.16}$$

If g is the Minkowski metric then its determinant is negative in any coordinate system and so this is often written $\sqrt{-\det g}$ instead. It's also common to use g as short for $\det g$, and anywhere else the metric appears it has indices.

Three

Matrix Groups

3.1 General Theory

In this chapter we will study various metric preserving groups of linear transformations on some vector space, V. If V is finite dimensional then we can choose a basis and identify these linear transformations with matrices, which gives a group.

A **linear transformation** is a function, $T\colon U \to V$, where U and V are vector spaces over the same base field, $\Bbbk$, such that

$$T(\alpha u_1 + \beta u_2) = \alpha T(u_1) + \beta T(u_2) \tag{3.1.1}$$

for all $\alpha, \beta \in \Bbbk$ and $u_1, u_2 \in U$. Note that we often drop the brackets for linear transformations, and just write $Tu = T(u)$, which reflects the notion of linear transformations as matrices in the finite dimensional case.

The group operation in question is composition of linear transformation, which is matrix multiplication in the finite dimensional case. This inherits associativity from the underlying operation of function composition. Let $f\colon A \to B$, $g\colon B \to C$, and $h\colon C \to D$ be functions. Then,

$$\begin{aligned}[(h \circ g) \circ f](a) &= (h \circ g)(f(a)) = h(g(f(a))) \\ &= h((g \circ f)(a)) = [h \circ (g \circ f)](a).\end{aligned} \tag{3.1.2}$$

for all $a \in A$, so $(h \circ g) \circ f = h \circ (g \circ f)$.

The identity linear transformation, $1\colon V \to V$, is given by $1(v) = v$ for all $v \in V$. Note that $T \circ 1_U = T$ and $1_V \circ T = T$ for all $T\colon U \to V$ where 1_V and 1_U are the identities on their respective vector spaces.

The inverse linear transformation of $T\colon U \to V$, if it exists, is the linear transformation $T^{-1}\colon V \to U$ such that $T \circ T^{-1} = 1_V$ and $T^{-1} \circ T = 1_U$. We get around the "if it exists" problem by just defining our groups to be formed from invertible transformations. This works since if $T\colon U \to V$ and $S\colon V \to W$ are invertible linear transformations then the inverse of $S \circ T$ is $T^{-1} \circ S^{-1}$:

$$(S \circ T) \circ (T^{-1} \circ S^{-1}) = S \circ (T \circ T^{-1}) \circ S^{-1} = S \circ 1_V \circ S = S \circ S^{-1} = 1_W, \tag{3.1.3}$$

and similarly $(T^{-1} \circ S^{-1}) \circ (S \circ T) = 1_U$, so $S \circ T\colon U \to W$ has an inverse.

In our examples we will consider linear transformations from a space to itself, known as **endomorphisms**. The set of all such functions is

$$\operatorname{End}(V) \coloneqq \{T\colon V \to V \mid T \text{ is linear}\}. \tag{3.1.4}$$

This same notion applies to other types of objects, such as groups, with linear transformations replaced with the appropriate type, so group homomorphisms for groups. With function composition as multiplication $\mathrm{End}(V)$ forms a monoid, which is a group the requirement for inverses. Restricting ourselves to invertible linear transformations we consider all invertible linear transformations from a space to itself, known as **automorphisms**. The set of all such functions is

$$\mathrm{Aut}(V) \coloneqq \{T\colon\ V \to V \mid T \text{is linear and invertible}\} \subseteq \mathrm{End}(V). \tag{3.1.5}$$

With function composition as multiplication $\mathrm{Aut}(V)$ forms a group.

In the finite dimensional case we are interested in transformations which can be expressed as invertible matrices, $D^{\mu}{}_{\nu}(\alpha)$, where $\alpha = \alpha_1, \dots, \alpha_{\dim G}$ parametrises the group. These act on vectors as

$$x^{\mu} \to x'^{\mu} = D^{\mu}{}_{\nu}(\alpha)x^{\nu} \tag{3.1.6}$$

where $\mu, \nu = 1, \dots, N$. Note that $N \neq \dim G$ in general.

3.2 Matrix Groups

In this section we will define specific matrix groups of interest in physics. In all of these suppose we have a vector space, V, of potentially infinite dimension over either the real or complex numbers. In all cases the elements of the groups are linear transformations and the group operation is composition of transformations. The identity is the identity transformation. Our first group is the broadest possible such group.

Definition 3.2.1 — General Linear Group The **general linear group**, $\mathrm{GL}(V)$, is the group of all invertible transformations of V, that is

$$\mathrm{GL}(V) \coloneqq \{T\colon\ V \to V \mid T \text{ is invertible}\} = \mathrm{Aut}(V). \tag{3.2.2}$$

If V is a vector space over the field $\Bbbk$, and is of finite dimension N, then by choosing a basis we can identify each linear transformation with a matrix, and we can identify $\mathrm{GL}(V)$ with the set of invertible $N \times N$ matrices over $\Bbbk$, denoted $\mathrm{GL}(N, \Bbbk)$. In this case invertible is equivalent to nonzero determinant, and so

$$\mathrm{GL}(N, \Bbbk) \coloneqq \{T \in \mathcal{M}_N(\Bbbk) \mid \det T \neq 0\} \tag{3.2.3}$$

where $\mathcal{M}_N(\Bbbk)$ is the set of $N \times N$ matrices over $\Bbbk$.

There are a variety of notations for the general linear group, and the subsequent subgroups, such as $\mathrm{GL}_n(\Bbbk)$ or $\mathrm{GL}(n)$, leaving the precise field to context.

All other groups to be defined are subgroups of $\mathrm{GL}(V)$, and so to show they are groups we need only show that they're closed under the induced operation and contain all inverses. This will require various linear algebra facts which we state without proof. We also won't worry too much about intricacies with definitions in the infinite dimensional case, we'll just work with matrices.

Definition 3.2.4 — Special Linear Group The **special linear group**, $\mathrm{SL}(V)$, for finite dimensional V, is the group of all invertible linear transformations with unit determinant:

$$\mathrm{SL}(V) := \{T\colon V \to V \mid T \text{ is invertible and } \det T = 1\} \subseteq \mathrm{GL}(V). \quad (3.2.5)$$

Choosing a basis we can identify each linear transformation with an $N \times N$ matrix over $\Bbbk$ and we get

$$\mathrm{SL}(V) := \{T \in \mathcal{M}_N(\Bbbk) \mid \det T = 1\} \subseteq \mathrm{GL}(N, \Bbbk). \quad (3.2.6)$$

We know that $\det(S \circ T) = \det(S)\det(T)$, so if $\det T = \det S = 1$ then $\det(S \circ T) = 1$, so $\mathrm{SL}(V)$ is closed under composition. Since $\det(T^{-1}) = 1/\det(T)$ if $\det T = 1$ then $\det(T^{-1}) = 1$, so $\mathrm{SL}(V)$ contains all inverses. Hence, $\mathrm{SL}(V)$ is a group.

Definition 3.2.7 — Orthogonal Group The **orthogonal group**, $\mathrm{O}(N)$, is the group preserving the Euclidean metric. It consists of all orthogonal $N \times N$ matrices over $\mathbb{R}$:

$$\mathrm{O}(N) := \{O \in \mathcal{M}_N(\mathbb{R}) \mid O^\top O = OO^\top = 1\} \subseteq \mathrm{GL}(N, \mathbb{R}). \quad (3.2.8)$$

Note that $O^\top O = 1$ can be written as $O^\top 1 O = 1$, where the 1 on the left is understood as the matrix form of the Euclidean metric, so the Euclidean metric is invariant under the action of the Orthogonal group

$$\delta_{\mu\nu} \xrightarrow{D \in \mathrm{O}(N)} \delta_{\alpha\beta} D^\alpha{}_\mu D^\beta{}_\nu = \delta_{\mu\nu}. \quad (3.2.9)$$

This allows us to interpret orthogonal transformations as rotations, since they preserve distances and angles.

Suppose $O_1, O_2 \in \mathrm{O}(N)$, then $O^\top O = 1$. Then $(O_1 O_2)^\top (O_1 O_2) = O_2^\top O_1^\top O_1 O_2 = O_2^\top 1 O_2 = 1$, so $\mathrm{O}(N)$ is closed under composition. We claim that if $O \in \mathrm{O}(N)$ then $O^{-1} \in \mathrm{O}(N)$, first note that $O^{-1} = O^\top$, then we have $(O^{-1})^\top O^{-1} = (O^\top)^\top O^\top = OO^\top = 1$, so $O^{-1} \in \mathrm{O}(N)$. Hence $\mathrm{O}(N)$ contains all inverses.

Definition 3.2.10 — Special Orthogonal Group The **special orthogonal group**, $\mathrm{SO}(N)$, is the subgroup of the orthogonal group given by restricting to $O \in \mathrm{O}(N)$ with $\det O = 1$:

$$\mathrm{SO}(N) := \{O \in \mathrm{O}(N) \mid \det O = 1\}. \quad (3.2.11)$$

Note that $\mathrm{SO}(N) \subseteq \mathrm{O}(N)$ and $\mathrm{SO}(N) \subseteq \mathrm{SL}(N, \mathbb{R})$.

The special orthogonal group is a group for exactly the same reasons that $\mathrm{O}(N)$ and $\mathrm{SL}(N, \Bbbk)$ are.

Definition 3.2.12 — Unitary Group The **unitary group**, $\mathrm{U}(N)$, is the group preserving the standard inner product on a complex vector space:

$$(x, y) = \sum_i x_i^* y_i. \tag{3.2.13}$$

! We follow the physics convention that an inner product is conjugate linear in its first argument, mathematicians often define it to be conjugate linear in the second instead.

The unitary group consists of all $N \times N$ unitary matrices over $\mathbb{C}$:

$$\mathrm{U}(N) := \{U \in \mathcal{M}_N(\mathbb{C}) \mid U^\dagger U = UU^\dagger = 1\} \subseteq \mathrm{GL}(N, \mathbb{C}). \tag{3.2.14}$$

Note that $U^\dagger U = 1$ can be written as $U^\dagger 1 U = 1$, with the 1 on the left understood as the matrix form of the metric on this complex vector space, so the metric is invariant under the action of the unitary group

$$\delta_{\mu\nu} \xrightarrow{D \in \mathrm{U}(N)} \delta_{\alpha\beta} (D^\alpha{}_\mu)^* D^\beta{}_\nu = \delta_{\mu\nu}. \tag{3.2.15}$$

The same logic used to show the orthogonal group is a group works for the unitary group if we just replace transposes with Hermitian conjugates.

Definition 3.2.16 — Special Unitary Group The **special unitary group**, $\mathrm{SU}(N)$, is the subgroup of the unitary group given by restricting to $U \in \mathrm{U}(N)$ with $\det U = 1$:

$$\mathrm{SU}(N) := \{U \in \mathrm{U}(N) \mid \det U = 1\}. \tag{3.2.17}$$

Note that $\mathrm{SU}(N) \subseteq \mathrm{U}(N)$ and $\mathrm{SU}(N) \subseteq \mathrm{SL}(N, \mathbb{C})$.

The special unitary group is a group for exactly the same reasons that $\mathrm{U}(N)$ and $\mathrm{SL}(N, \Bbbk)$ are.

Example 3.2.18 — SU(2) Consider $\mathrm{SU}(2)$. Start with some 2×2 matrix over $\mathbb{C}$,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{3.2.19}$$

The requirement that $U^\dagger U = 1$ tells us that

$$U^{-1} = U^\dagger \implies \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} \tag{3.2.20}$$

so $a = d^*$ and $c^* = -b$. Now add in the requirement that $\det U = 1$ and we have

$$ad - bc = 1 \implies aa^* + bb^* = 1. \tag{3.2.21}$$

Now write $a = \alpha_1 + i\alpha_2$ and $b = \alpha_3 + i\alpha_4$ for $\alpha_i \in \mathbb{R}$. Then we have

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \alpha_4^2 = 1. \tag{3.2.22}$$

This defines the three sphere,

$$S^3 := \{\boldsymbol{x} \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + x_3^2 + x_4^2 = \|\boldsymbol{x}\|^2 = 1\}. \tag{3.2.23}$$

This is a three-dimensional real manifold which is the parameter space for SU(2). Note that S^3 is simply connected, that is any loop can be contracted to a point.

Definition 3.2.24 — Pseudo-Orthogonal Group The **pseudo-orthogonal group**, $\mathrm{O}(n, m)$, is the group preserving the metric with signature (n, m).

Example 3.2.25 — Lorentz Group The **proper Lorentz group**, SO(3, 1), is the group of Lorentz transformations, Λ, preserving the Minkowski metric:

$$\Lambda^{\top}\eta\Lambda = \eta \iff \Lambda_{\mu}{}^{\rho}\eta_{\rho\sigma}\Lambda^{\sigma}{}_{\nu} = \eta_{\mu\nu}, \tag{3.2.26}$$

and with $\det\Lambda = 1$.
Note that we can also consider the general **Lorentz group** $\mathrm{O}(3, 1)$, which allows transformations inverting spacetime with $\det\Lambda \neq 1$, and the **proper orthochronous Lorentz group**, $\mathrm{SO}^+(3, 1)$, which has $\Lambda^0{}_0 \geq 1$.

Definition 3.2.27 — Pseudo-Unitary Group The **pseudo-unitary group** is the group preserving an indefinite metric on a complex vector space.

Definition 3.2.28 — Symplectic Groups The **symplectic group** $\mathrm{Sp}(2N, \Bbbk)$, preserves an antisymmetric metric given by the block diagonal matrix

$$g = \begin{pmatrix} 0 & 1_N \\ -1_N & 0 \end{pmatrix} \tag{3.2.29}$$

on a $2N$-dimensional vector space over $\Bbbk$.
The **compact symplectic group**, $\mathrm{Sp}(2N)$, is $\mathrm{Sp}(2N) = \mathrm{Sp}(2N, \mathbb{C}) \cap \mathrm{U}(2N)$.
We can think of it as being the result of taking matrices in $\mathrm{Sp}(2N), \mathbb{C}$ and replacing complex numbers with 2×2 real matrices according to

$$x + iy \mapsto \begin{pmatrix} x & y \\ -y & x \end{pmatrix}. \tag{3.2.30}$$

The symplectic group arises in physics when we consider phase space. In classical mechanics in three spatial dimensions phase space is the six dimensional space spanned by the three components of position and three components of momentum, so a point in phase space is $(q_1, q_2, q_3, p_1, p_2, p_3)$. Symplectic transformations, $\mathrm{Sp}(6, \mathbb{R})$, are the set of transformations preserving Hamilton's equations:

$$\dot{q}_i = \frac{\mathrm{d}H}{\mathrm{d}p_i}, \qquad \dot{p}_i = -\frac{\mathrm{d}H}{\mathrm{d}q_i}. \tag{3.2.31}$$

With some restrictions most Lie groups fit into one of the previously mentioned categories. However, there are five Lie groups that don't, these are called the **exceptional groups**. There is no particularly simple definition of any of them, we just note here that they exist, and are called F_4, G_2, E_6, E_7, and E_8. These names come from the classification of semisimple Lie algebras (to be defined later), where we call the Lie algebras of $\mathrm{SL}(n+1, \mathbb{C})$, $\mathrm{SO}(2n+1)$, $\mathrm{Sp}(2n)$, and $\mathrm{SO}(2n)$ by the names A_n, B_n, C_n, and D_n respectively. Here n is the rank of the Lie algebra (to be defined later).

3.3 One Dimensional Groups

In this section we will discuss one-dimensional groups, these are groups parametrised by a single value.

Example 3.3.1

- Translations along $\mathbb{R}$ form a one-dimensional Lie group parametrised by the size of the translation, $a \in \mathbb{R}$, with the action $x \mapsto x + a$. This group is just $\mathbb{R}$. This example is noncompact.
- Rotations about some fixed axis form a one-dimensional Lie group parametrised by the size of the rotation, $\vartheta \in [0, 2\pi)$, with the action

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ -\sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{3.3.2}$$

  This group is $\mathrm{SO}(2)$. This example is compact.
- Multiplication by a phase factor forms a one-dimensional Lie group parametrised by the argument of the phase factor, $\varphi \in [0, 2\pi)$, with the action $z \mapsto \mathrm{e}^{i\varphi} z$. This group is $\mathrm{U}(1)$. This example is compact.

These last two examples are actually isomorphic, using $\mathrm{e}^{i\varphi} = \cos\varphi + i \sin\varphi$ we can map from $\mathrm{U}(1)$ to $\mathrm{SO}(2)$ with

$$\mathrm{e}^{i\varphi} \mapsto \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}. \tag{3.3.3}$$

So, $\mathrm{U}(1) \cong \mathrm{SO}(2)$.

One-dimensional Lie groups are, unsurprisingly, some of the simplest Lie groups. In fact, they're so simple that they're not that interesting, a theory we'll develop rigorously through a collection of theorems.

Theorem 3.3.4. All one-dimensional Lie groups can be parametrised so that

$$g(a)g(b) = g(a + b) \tag{3.3.5}$$

for all a and b.

Proof. Consider a one-dimensional Lie group, G, parametrised by some value in the interval I. Associativity tells us that

$$g(x)[g(y)g(z)] = [g(x)g(y)]g(z) \tag{3.3.6}$$

for all $x, y, z \in I$. We can write $g(y)g(z)$ and $g(x)g(y)$ as single group elements $g(\varphi(y,z))$ and $g(\varphi(x,y))$ for some analytic function $\varphi : I^2 \to I$, this is just the definition of a Lie group. Doing so we have

$$g(x)g(\varphi(y,z)) = g(\varphi(x,y))g(z). \tag{3.3.7}$$

Now we can use analyticity again to write $g(x)g(\varphi(y,z)) = g(x)g(a) = g(\varphi(x,a)) = g(\varphi(x,\varphi(y,z)))$ and $g(\varphi(x,y))g(z) = g(\varphi(\varphi(x,y),z))$:

$$g(\varphi(x,\varphi(y,z))) = g(\varphi(\varphi(x,y),z)). \tag{3.3.8}$$

Hence, we must have

$$\varphi(x,\varphi(y,z)) = \varphi(\varphi(x,y),z). \tag{3.3.9}$$

Now, take the derivative of this expression with respect to z, the right hand side gives

$$\frac{\partial}{\partial z}\varphi(\varphi(x,y),z), \tag{3.3.10}$$

and the chain rule applied to the left hand side gives

$$\frac{\partial}{\partial z}\varphi(x,\varphi(y,z)) = \frac{\partial\varphi(x,\varphi(y,z))}{\partial\varphi(y,z)}\frac{\partial\varphi(y,z)}{\partial z}. \tag{3.3.11}$$

We are free to shift our parametrisation interval around, so imagine we choose it to contain zero and choose a parametrisation such that $g(0) = 1$. Consider the case where $z = 0$. Then $g(y)g(z) = g(y)g(0) = g(y)1 = g(y)$, but we also have $g(y)g(z) = g(\varphi(y,z)) = g(\varphi(y,0))$, so we must have $\varphi(y,0) = y$. Thus, evaluating are derivatives at $z = 0$, we have

$$\frac{\partial\varphi(x,y)}{\partial y}\psi(y) = \psi(\varphi(x,y)) \tag{3.3.12}$$

where

$$\psi(y) := \left.\frac{\partial\varphi(y,z)}{\partial z}\right|_{z=0}. \tag{3.3.13}$$

This differential equation can be solved by writing it as

$$\frac{1}{\psi(\varphi(x,y))}\frac{\partial\varphi(x,y)}{\partial y} = \frac{1}{\psi(y)}. \tag{3.3.14}$$

We can then integrate with respect to y and we get

$$\rho(\varphi(x,y)) = \rho(y) + c(x) \tag{3.3.15}$$

where c is an arbitrary function of x, taking the role of our integration constant, but not constant as x is allowed to vary, and

$$\rho(x) := \int_0^x \frac{1}{\psi(t)}\,\mathrm{d}t \tag{3.3.16}$$

with $\rho(0) = 0$.
We can determine $c(x)$ by choosing $y = 0$, which gives $\rho(\varphi(x, 0)) = \rho(x) = \rho(0) + c(x) = 0 + c(x)$, so $c(x) = \rho(x)$. Hence,

$$\rho(\varphi(x, y)) = \rho(x) + \rho(y). \tag{3.3.17}$$

Now we just reparametrise our group from $\mathrm{g}(x)$ to $\bar{\mathrm{g}}(\rho(x))$, we then have

$$\bar{\mathrm{g}}(\rho(x))\bar{\mathrm{g}}(\rho(y)) = \mathrm{g}(x)\mathrm{g}(y) = \mathrm{g}(\varphi(x, y)) = \bar{\mathrm{g}}(\rho(\varphi(x, y))) = \bar{\mathrm{g}}(\rho(x)+\rho(y)). \tag{3.3.18}$$

Hence, our group operation becomes addition in parameter space and we are finished. □

Corollary 3.3.19 All one dimensional Lie algebras are Abelian.

Proof. Let G be a one-dimensional Lie group parametrised such that $\mathrm{g}(a)\mathrm{g}(b) = \mathrm{g}(a + b)$ for all a and b. Then

$$\mathrm{g}(a)\mathrm{g}(b) = \mathrm{g}(a + b) = \mathrm{g}(b + a) = \mathrm{g}(b)\mathrm{g}(a), \tag{3.3.20}$$

and so G is Abelian. □

The following theorem won't be proved here, but intuitively all it says is that every compact connected Abelian Lie group can be parametrised by phase factors.

Theorem 3.3.21. All compact connected Abelian Lie groups are isomorphic to

$$\bigotimes_{i=1}^{n} \mathrm{U}(1) = \underbrace{\mathrm{U}(1) \otimes \cdots \otimes \mathrm{U}(1)} \tag{3.3.22}$$

for some $n \in \mathbb{N}$.

Identifying each copy of $\mathrm{U}(1)$ with a circle we see that every compact connected Abelian Lie group is a product of circles, which is a Torus.
This means that in most of the course we'll be interested in non-Abelian Lie groups.

Four

Representations

4.1 What is a Representation

Intuitively a representation of a group, G, is a set of $N \times N$ matrices, D, parametrised by group elements, g, such that matrix multiplication is compatible with the group operation:

$$D(g)D(h) = D(gh) \iff D^{\mu}{}_{\nu}(g)D^{\nu}{}_{\rho}(h) = D^{\mu}{}_{\rho}(gh) \tag{4.1.1}$$

for all $g, h \in G$.

Example 4.1.2 — S_3 The symmetric group on 3 objects has a representation, called the **permutation representation**, given by the matrices

$$() \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1\,2) \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1\,3) \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$(2\,3) \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (1\,2\,3) \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad (3\,2\,1) \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

To motivate this suppose that the three objects are the basis vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{and} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{4.1.3}$$

Then this representation acts by permuting the vectors in the obvious way. For example, the matrix representing $(1\,2)$ sends e_1 to e_2, e_2 to e_1, and e_3 to itself.

Another representation of S_3 is given by the **trivial representation**, which represents all group elements with the 1×1 zero matrix, (0), sending all vectors to the zero vector.

A third representation of S_3 is given by sending each element either to (1) or (-1), depending on the sign of the permutation.

There are two ways to formalise the notion of a representation, they are as follows.

Definition 4.1.4 — Representation Let G be a group. Then a group **representation**, (D, V), is a vector space, V, and a homomorphism, $D: G \to \mathrm{GL}(V)$.

Definition 4.1.5 — Representation Let G be a group and V a vector space. A group **representation**, (D, V), is a vector space, V, and a group action, $G \times V \to V$ given by $g.v = D(g)v$ where $D(g)$ is some linear transformation.

Notation 4.1.6 We're generally pretty loose with the exact language as to what objects make up a representation. People will refer to (D, V), D, V, and the set of matrices $\{D(g) \mid g \in G\}$ as the representation interchangeably, with context telling us which we care about.

Definition 4.1.7 — Representation Dimension The dimension of a representation is the dimension of the vector space, V, which is also the number of rows in the matrix.

! The dimension of the representation is, in general, *not* the same as the dimension of a continuous group.

Example 4.1.8 — S_3 The permutation representation is of dimension 3, the trivial representation is of dimension 1, as is the sign representation.

4.2 Properties of Representations

Example 4.2.1 — Faithful Representation A representation is faithful if ρ is injective.

Example 4.2.2 — S_3 The permutation representation is faithful, the trivial and sign permutations are not.

Representations give us another way to think about the matrix groups of the last chapter. Instead of defining them as matrices we can define them by symmetries, so, for example, $\mathrm{O}(N)$ is the group preserving the Euclidean metric. Then the usual interpretation as $N \times N$ orthogonal matrices is just a representation. We call it the **defining representation**, since it can be used to define the representation. The defining representation must be an isomorphism in order for it to contain all group elements exactly once.

The **fundamental representation** is one from which all other all other representations can be built from tensor products.

Definition 4.2.3 — Equivalence of Representations Two representations are **equivalent** if they are related by a similarity transformation. That is, if G is a group with representations D and D' then there exists some S, independent of g, such that for all G

$$D'(g) = SD(g)S^{-1}. \tag{4.2.4}$$

Definition 4.2.5 — Reducible A representation, (D, V), is **reducible** if there exists some subspace $U \subset V$ with $U \neq V$ and $U \neq \{0\}$ such that

$$D(g)u \in U \text{ for all } u \in U \text{ and } g \in G. \tag{4.2.6}$$

If this is the case we call U an **invariant subspace**. A representation is **irreducible** if it is not reducible.

Notation 4.2.7 People often shorten “irreducible representation” to **irrep**.

Notice that if D is a representation of G then so is D^*, since

$$D(g)^* D(h)^* = (D(g)D(h))^* = D(gh)^* \tag{4.2.8}$$

for all $g, h \in G$.

Definition 4.2.9 Let D be a representation of some group G.
If $D(g)^* = D(g)$ for all $g \in G$ then we say that D is a **real representation**.
If D is not equivalent to D^* then we say that D is a **complex representation**.
If $D \neq D^*$ but D is equivalent to D^* then we say that D is a **pseudo-real representation**.

Suppose that we have some symmetry described by the group G. Then we classify different types of objects based on which representation they transform under. If

$$s \mapsto s, \tag{4.2.10}$$

that is there is no transformation, then s transforms under the trivial representation and we call s a **scalar**. If

$$v^\mu \mapsto D^\mu{}_\nu(g) v^\nu, \tag{4.2.11}$$

where D is the fundamental representation we say that v is a **vector**. If

$$T^{\mu_1\mu_2\cdots\mu_n} \mapsto D^{\mu_1}{}_{\nu_1}(g) D^{\mu_2}{}_{\nu_2}(g) \cdots D^{\mu_n}{}_{\nu_n}(g) T^{\nu_1\nu_2\cdots\nu_n} \tag{4.2.12}$$

then we say that T is a **tensor**, note that both scalars and vectors are special cases of tensors.

We can write this last line more succinctly as

$$T^\alpha = D^\alpha{}_\beta(g) T^\beta. \tag{4.2.13}$$

We understand this as T transforming under the tensor product

$$(D(g) \otimes D(g) \otimes \cdots \otimes D(g))^{\mu_1\mu_2\cdots\mu_n}{}_{\nu_1\nu_2\ldots\nu_n} = D^{\mu_1}{}_{\nu_1}(g) D^{\mu_2}{}_{\nu_2}(g). \tag{4.2.14}$$

4.3 Unitary Representations

Definition 4.3.1 — Unitary Representation Let G be a group and (D, V) a representation of G. We say that D is a **unitary representation** if it preserves the inner product on V, that is if

$$(u, v) = (D(g)u, D(g)v) \tag{4.3.2}$$

for all $g, h \in G$.

Theorem 4.3.3 — 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 (D, V) a representation. Let $(-, -)$ be the inner product on V. Define a new inner product,

$$\langle x|y\rangle \coloneqq \sum_{g\in G}(D(g)x, D(g)y). \tag{4.3.4}$$

Then we claim that (D, V) is a unitary representation with respect to this new inner product and since different inner products are related by similarity transformations (D, V) is equivalent to a unitary representation.
Take some fixed $h \in G$. The rearrangement theorem states that if the elements of the group are listed, $(g_1, \dots, g_n)$, then the action $(g_1, \dots, g_n) \mapsto (g_1 h, \dots, g_n h)$ is just a permutation. In particular, if $f : G \to \mathbb{C}$ then

$$\sum_{g\in G} f(g) = \sum_{g\in G} f(gh), \tag{4.3.5}$$

since multiplying each argument by h just permutes the terms.
Using this we can equivalently write the inner product, $\langle -|-\rangle$, as

$$\begin{align}
\langle x|y\rangle &= \sum_{g\in G}(D(gh)x, D(gh)y) \tag{4.3.6}\\
&= \sum_{g\in G}(D(g)D(h)x, D(g)D(h)y) \tag{4.3.7}\\
&= \sum_{g\in G}(D(g)u, D(g)v) \tag{4.3.8}\\
&= \langle u|v\rangle, \tag{4.3.9}
\end{align}$$

where $u = D(h)x$ and $v = D(h)y$. Then we have

$$\langle x|y\rangle = \langle u|v\rangle = \langle D(h)x|D(h)y\rangle = \langle x|D^\dagger(h)D(h)|y\rangle, \tag{4.3.10}$$

so for this to hold true for all $x, y \in V$ we must have $D^\dagger(h)D(h) = 1$. Hence, D is unitary with respect to this new inner product. □

A very similar theorem holds for compact Lie groups. The proof is almost identical but the sum is replaced with an integral. We won't prove it here as it requires some measure theory to make everything work. It comes down to the compactness requirement ensuring convergence of the integrals at every step. Without this requirement the integrals may not converge, in which case there is no reason for an equivalent unitary representation to exist.

Theorem 4.3.11. Every representation of a compact Lie group is equivalent to a unitary representation.

Why do we care about unitary representations? Mostly for quantum mechanics. If a group action on a Hilbert space of states is to preserve probability then it must be either unitary, so $(Ux, Uy) = (x, y)$, or **antiunitary**, that is $(Ux, Uy) = (x, y)^*$. An example of an antiunitary transformation is time reversal. Consider a system evolving from state $|\psi\rangle$ to $|\varphi\rangle$. The amplitude for this is $\langle\psi|\varphi\rangle$. The time reversed system evolves from state $|\varphi\rangle$ to state $|\psi\rangle$. The amplitude for this is $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$.

Part II

Lie Algebras

Five

Lie Algebras

There are, broadly, two approaches to Lie algebras. One can derive their properties and then abstract, or abstract and then show that there is a connection to Lie groups. The first is how a typical physics course goes about it, and fits the historical process of discovering Lie algebras, but I prefer to abstract and then look at applications where possible, so that's what these notes will do.

We'll define Lie algebras as abstract objects, and briefly discuss some properties which follow immediately. Then we'll discuss the link to Lie groups, and then go into more detail about the properties of Lie algebras.

5.1 What is an Algebra?

An algebra is a vector space equipped with a product compatible with the vector space structure. Formally, the definition is as follows.

Definition 5.1.1 — Algebra Let $\Bbbk$ be a field and A a vector space over $\Bbbk$. Equip A with a binary operation $A \times A \to A$, denoted by juxtaposition here. If this product has the following properties then A is an **algebra**:

- Left distributivity: $x(y+z) = xy + xz$ for all $x, y, z \in A$;
- Right distributivity: $(x+y)z = xz + yz$ for all $x, y, z \in A$;
- Compatibility with scalar multiplication: $(\alpha x)(\beta y) = (\alpha\beta)(xy)$ for all $\alpha, \beta \in \Bbbk$ and $x, y \in A$.

We call A an **associative algebra** if the product is associative.

Example 5.1.2 — Algebra

- The real numbers are a vector space over themselves, and also an algebra over themselves with the product being normal multiplication.
- The complex numbers are a vector space over themselves, and also an algebra over themselves with the product being normal multiplication.
- The complex numbers are a vector space over $\mathbb{R}$, and also an algebra

over $\mathbb{R}$ with the product being normal complex multiplication when we think of complex numbers as ordered pairs of real numbers.

- The vector space $\mathbb{R}^3$ when equipped with the cross product forms an algebra over $\mathbb{R}$.
- The vector space $\mathbb{R}^4$ when identified with the quaternions forms an algebra over $\mathbb{R}$ with the product being the quaternion product.
- Polynomials over $\mathbb{R}$, that is $\mathbb{R}[x]$, form a vector space over $\mathbb{R}$ and when equipped with polynomial multiplication this is an algebra.
- The set of $m \times n$ matrices over $\mathbb{R}$ forms an mn-dimensional vector space. This forms an algebra over $\mathbb{R}$ when equipped with the normal matrix product.
- The set of $m \times n$ matrices over $\mathbb{R}$ forms an mn-dimensional vector space. This forms an algebra over $\mathbb{R}$ when equipped with the **commutator**, $[A, B] \coloneqq AB - BA$.

5.2 What is a Lie Algebra

A Lie algebra is an algebra where the product is given by a Lie bracket, which is just a particular form of product with a couple of properties which we write as a bracket, so $[a, b]$ instead of ab.

Definition 5.2.1 — Lie Algebra A **Lie algebra**, $\mathfrak{g}$, is a vector space over some field, $\Bbbk$, equipped with a binary operation, $[-,-] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, satisfying the following:

- **Bilinearity**: for all $\alpha, \beta \in \Bbbk$ and $x, y, z \in \mathfrak{g}$ we have

$$[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z], \quad \text{and} \quad [z, \alpha x + \beta y] = \alpha[z, x] + \beta[z, y].$$

- **Alternativity**: for all $x \in \mathfrak{g}$ we have

$$[x, x] = 0, \tag{5.2.2}$$

  where 0 is the zero vector.

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

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \tag{5.2.3}$$

  Note that this is just $[x, [y, z]]$ plus cyclic permutations.

Notation 5.2.4 Lie algebras are, unsurprisingly, related to Lie groups. If we have a Lie group denoted with capital letters, such as G, GL, SO, U, and so on, then we denote the associated Lie algebra by the same letter but lowercase and in the Fraktur script, so $\mathfrak{g}$, $\mathfrak{gl}$, $\mathfrak{so}$, $\mathfrak{u}$, and so on. For future reference here is the alphabet, in order, in Fraktur:

$$\mathfrak{a}\,\mathfrak{b}\,\mathfrak{c}\,\mathfrak{d}\,\mathfrak{e}\,\mathfrak{f}\,\mathfrak{g}\,\mathfrak{h}\,\mathfrak{i}\,\mathfrak{j}\,\mathfrak{k}\,\mathfrak{l}\,\mathfrak{m}\,\mathfrak{n}\,\mathfrak{o}\,\mathfrak{p}\,\mathfrak{q}\,\mathfrak{r}\,\mathfrak{s}\,\mathfrak{t}\,\mathfrak{u}\,\mathfrak{v}\,\mathfrak{w}\,\mathfrak{x}\,\mathfrak{y}\,\mathfrak{z}. \tag{5.2.5}$$

It is also common to use lowercase, non-Fraktur letters for Lie algebras, such as g, gl, so, u, and so on.

Note that some sources define the Lie bracket to be anticommutative, instead of alternating. That is, $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$. This is necessarily the case for an alternating Lie bracket:

$$\begin{aligned} 0 = [x + y, x + y] &= [x, x + y] + [y, x + y] \\ &= [x, x] + [x, y] + [y, x] + [y, y] = [x, y] + [y, x], \end{aligned} \tag{5.2.6}$$

and so $[x, y] = -[y, x]$. However, the reverse implication, that an anticommutative product is alternating, only holds if the characteristic[1] of the field is not 2, so our definition is *slightly* more general. However, we'll almost entirely work over $\mathbb{R}$ and $\mathbb{C}$ which have characteristic 0, so it's not an important distinction. So, suppose that we work over a field with characteristic not equal to 2, then $[x, x] = -[x, x]$ by anticommutativity (in the second bracket we swapped the xs, you just can't tell). Hence, we have

[1]the **characteristic** of a ring, and hence field, is the number of times you have to add 1 to itself to get 0, or 0 if you never get 0 this way.

$$2[x, x] = [x, x] + [x, x] = [x, x] - [x, x] = 0, \tag{5.2.7}$$

and so long as we can divide by 2 we have $[x, x] = 0$. It is this dividing by 2 step that fails in a field with characteristic 2.

The Jacobi identity is a bit weird as a requirement. Define $D_x(y) = [x, y]$, then we can write the Jacobi identity as

$$D_x([y, z]) = [D_x(y), z] + [y, D_x(z)], \tag{5.2.8}$$

which looks a lot like the product rule. Going back to writing Lie brackets this is

$$[x, [y, z]] = [[x, y], z] + [y, [x, z]]. \tag{5.2.9}$$

Using the anticommutativity of the Lie bracket (in fields of characteristic other than 2, an implicit assumption from now on) and the bilinearity we can rewrite

$$[[x, y], z] = -[z, [x, y]] \tag{5.2.10}$$

and

$$[y, [x, z]] = [y, -[z, x]] = -[y, [z, x]]. \tag{5.2.11}$$

Hence,

$$[x, [y, z]] = -[z, [x, y]] - [y, [z, x]], \tag{5.2.12}$$

which is just the Jacobi identity with some terms moved to the other side.

Example 5.2.13 — Lie Algebra

- Consider the set of $m \times n$ matrices over $\mathbb{R}$. This forms a Lie algebra when equipped with the commutator, $[A, B] = AB - BA$.
- Consider $\mathbb{R}^3$. This forms a Lie algebra when equipped with the cross product, $[\boldsymbol{v}, \boldsymbol{u}] = \boldsymbol{v} \times \boldsymbol{u}$.

5.3 Subalgebras

As with groups we can often find Lie algebras hiding inside other Lie algebras, and certain such Lie algebras are special.

Definition 5.3.1 — Lie Subalgebra and Ideals Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h}$ a subspace of $\mathfrak{g}$ (as vector spaces). Then $\mathfrak{h}$ is a **Lie subalgebra**, or simply a **subalgebra**, of $\mathfrak{g}$ if $[x, y] \in \mathfrak{h}$ for all $x, y \in \mathfrak{h}$, where $[-, -]$ is the Lie bracket of $\mathfrak{g}$.
An **ideal** is a Lie subalgebra $\mathfrak{i}$ with the stronger condition that $[g, i] \in \mathfrak{i}$ for all $g \in \mathfrak{g}$.

The definition of an ideal here is analogous to that of a normal subgroup. A normal subgroup is invariant under conjugation with any group element and an ideal is invariant under Lie brackets with any algebra element. Note that if $[g, i] \in \mathfrak{i}$ then $[i, g] = -[g, i] \in \mathfrak{i}$ also as $\mathfrak{i}$ is a subspace of $\mathfrak{g}$. Similar to the group case a Lie algebra is **simple!Lie algebra** if it has no nontrivial ideals, that is the only ideals are itself and the zero dimensional vector space, and it is not Abelian. A Lie algebra is **Abelian** if the Lie bracket vanishes, that is $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. Note that this is equivalent to saying $[x, y] = [y, x]$, the more usual notion of being Abelian, but anticommutativity means $[y, x] = -[x, y]$, so we have $[x, y] = -[x, y]$, and so $[x, y] = 0$.

5.4 Morphisms Between Lie Algebras

As often happens in maths after defining an object we should define maps between these objects. These maps should preserve the structure of the object, and if such a map is invertible then the structure, at least at the Lie algebra level, of the two objects is the same, and so, when thinking about Lie algebras, we treat the two objects as if they were identical.

Definition 5.4.1 — Morphisms Let $\mathfrak{g}$ and $\mathfrak{h}$ be Lie algebras. A **Lie algebra homomorphism** is a function $\varphi : \mathfrak{g} \to \mathfrak{h}$ preserving the bracket, that is

$$\varphi([x, y]) = [\varphi(x), \varphi(y)] \tag{5.4.2}$$

for all $x, y \in \mathfrak{g}$. Here the Lie bracket on the left is that of $\mathfrak{g}$, and on the right the Lie bracket is that of $\mathfrak{h}$.
An **Lie algebra isomorphism** is an invertible Lie algebra homomorphism.

5.5 Generators and Structure Constants

Lie algebras are vector spaces, and as such have lots of nice things, like bases, that we can use to make calculations easier.

Definition 5.5.1 — Generators Let $\mathfrak{g}$ be a Lie algebra. A set of elements of $\mathfrak{g}$ are said to **generate** $\mathfrak{g}$ if the smallest subalgebra containing these elements is $\mathfrak{g}$ itself. That is, if all elements of $\mathfrak{g}$ can be be generated from linear combinations of the generators, and Lie brackets of the generators, and Lie brackets of Lie brackets of the generators and so on. The generators are the basis for $\mathfrak{g}$ as a vector space.

Notation 5.5.2 Generators are commonly denoted T_a, where a indexes the generating set.

Let $\mathfrak{g}$ be a Lie algebra with generators T_a. Since A Lie algebra is a vector space, and any element of it can be expressed as a linear combination of basis vectors, which for a Lie algebra are the generators. In particular, the Lie bracket of two elements of the Lie algebra is another element of the Lie algebra, and so can be expressed in this way:

$$[x, y] = f^a(x, y)T_a \tag{5.5.3}$$

where $f^a(x, y)$ is the coefficient of T_a in this expansion, and there is an implicit sum over a. Naturally we can consider the case when x and y are themselves generators, and this leads to the following definition.

Definition 5.5.4 — Structure Constants Let $\mathfrak{g}$ be a Lie algebra with generators T_a. Then the **structure constant** $c^c{}_{ab}$ is defined as the coefficient in one of the following:

$$[T_a, T_b] = c^c{}_{ab}T_c, \qquad \text{or} \qquad [T_a, T_b] = ic^c{}_{ab}T_c, \tag{5.5.5}$$

where there is an implicit sum over the index c.

Which of these two statements defines the structure constants depends on convention. For Lie algebras formed from real matrices we usually choose the first, and for Lie algebras formed from complex matrices the second. If we include the factor of i here then there are other places we need to include it later, but it can make the generators slightly nicer for physics, such as making them Hermitian when they would otherwise have been anti-Hermitian.

We can use the Jacobi identity to derive an equivalent identity for the structure constants. Start with the Jacobi identity applied to three generators:

$$0 = [T_a, [T_b, T_c]] + [T_b, [T_c, T_a]] + [T_c, [T_a, T_b]]. \tag{5.5.6}$$

Replacing the inner Lie bracket with structure constants then applying linearity we have

$$\begin{aligned} 0 &= [T_a, c^d{}_{bc}T_d] + [T_b, c^d{}_{ca}T_d] + [T_c, c^d{}_{ab}T_d] && (5.5.7)\\ &= c^d{}_{bc}[T_a, T_d] + c^d{}_{ca}[T_b, T_d] + c^d{}_{ab}[T_c, T_d]. && (5.5.8) \end{aligned}$$

Now replace the Lie brackets with structure constants again:

$$0 = c^d{}_{bc} c^e{}_{ad} T_e + c^d{}_{ca} c^e{}_{bd} T_e + c^d{}_{ab} c^e{}_{cd} T_e. \tag{5.5.9}$$

Requiring this to hold for initial choices of generators the coefficients here must vanish, and so

$$0 = c^d{}_{bc} c^e{}_{ad} + c^d{}_{ca} c^e{}_{bd} + c^d{}_{ab} c^e{}_{cd}. \tag{5.5.10}$$

Lemma 5.5.11 Let $\mathfrak{g}$ be the Lie algebra generated by $\{T_a\}$ with structure constants $c^c{}_{ab}$. If T_a are Hermitian then the structure constants are real.

Proof. If $\{T_a\}$ generates the Lie algebra then so does $\{-T_a^*\}$. Taking the complex conjugate of the defining relation for the structure constants we have

$$[T_a, T_b]^* = (ic^c{}_{ab} T_a)^* \implies [T_a^*, T_b^*] = -ic^c{}_{ab}{}^* T_c^*. \tag{5.5.12}$$

Linearity tells us that $[A, B] = [-A, -B]$, and so

$$[-T_a^*, T_b^*] = ic^c{}_{ab}{}^*(-T_c^*). \tag{5.5.13}$$

The structure constants are independent of the choice of generators, so we must have $c^c{}_{ab} = c^c{}_{ab}{}^*$, that is $c^c{}_{ab} \in \mathbb{R}$. □

5.6 Representations

Representations of Lie algebras aren't that different to representations of groups. The idea is to find some set of matrices, D, parametrised by the elements of the Lie algebra, say $x, y \in \mathfrak{g}$, such that the Lie bracket in $\mathfrak{g}$ corresponds to the commutator of these matrices, that is

$$D([x, y]) = D(x)D(y) - D(y)D(x) = [D(x), D(y)], \tag{5.6.1}$$

where the first $[-,-]$ is an abstract Lie bracket, which may or may not take the form of a commutator, and the second $[-,-]$ is the normal commutator. The formal definition of a representation of a Lie algebra is as follows.

Definition 5.6.2 — Representation Let $\mathfrak{g}$ be a Lie algebra. Then a **Lie algebra representation**, (D, V), is a vector space, V, and a homomorphism of Lie algebras

$$D\colon \mathfrak{g} \to \mathfrak{gl}(V) \tag{5.6.3}$$

where $\mathfrak{gl}(V) = \operatorname{End}(V)$ is made into a Lie algebra by equipping it with the commutator as a Lie bracket.

Representations essentially equate to a choice of generators. We have already met one of the most important representations of any Lie group, the structure constants themselves.

Definition 5.6.4 — Adjoint Representation The **adjoint representation** of a Lie algebra, $\mathfrak{g}$, is the representation on $\mathbb{C}^{\dim \mathfrak{g}}$ given by choosing the generators to satisfy

$$(T_a)^b{}_c = ic^b{}_{ac}. \tag{5.6.5}$$

To show that this is indeed a representation we need to show that T_a defined in this way satisfy $[T_a, T_b] = ic^c{}_{ab}T_c$. We can use the antisymmetry of the structure constants to write the Jacobi identity as

$$(T_a)^d{}_e(A_b)^e{}_c - (A_b)^d{}_e(A_a)^e{}_c = i(A_e)^d{}_c c^e{}_{ab}, \tag{5.6.6}$$

which in matrix notation gives

$$[T_a, T_b] = ic^e{}_{ab}T_e. \tag{5.6.7}$$

Six

Lie Algebras and Lie Groups

In this chapter we discuss the relationship between a Lie group and the associated Lie algebra. Starting with how we can get a Lie algebra from a Lie group by linearisation, then we'll see some examples, then we'll look at how to get back to the Lie group (or close to it) from the Lie algebra.

6.1 Linearising

Let G be a Lie group. An important idea is that every point in a Lie group is essentially equivalent, since if we study a neighbourhood of $g \in G$ then we can transform this into a neighbourhood of $h \in G$ by multiplying on the left by hg^{-1}. This suggests that we can learn about a Lie group by studying just a single neighbourhood, and if we have to select a single point in the group it makes sense to choose the identity.

If we want to study the behaviour of something near a fixed point it makes sense to expand about this point, which we're allowed to do for Lie groups by the analyticity part of the definition. Doing so up to first order we **linearise** the group. If the group elements are $g(\alpha)$, where $\alpha = \{\alpha^a\}$ is parametrising the group, then expanding about the identity, 1, we get

$$g(\alpha) = 1 + i\alpha^a T_a + \mathcal{O}(\alpha^2) \tag{6.1.1}$$

where

$$T_a = -i\left(\frac{\partial g(\alpha)}{\partial \alpha^a}\right)_{\alpha=0} \tag{6.1.2}$$

and a runs from 1 to $\dim G$. We call T_a the **generators**.

! The factor of i here in these definitions is convention. It allows us to work with Hermitian matrices rather than anti-Hermitian matrices. Not everyone follows this convention, and we won't include it if we're working with real matrices.

6.2 Lie Bracket

In a group we can measure how Abelian the group is using the **group commutator**, defined for $g, h \in G$ to be the product $f = ghg^{-1}h^{-1}$. Notice that this gives the identity if g and h commute. If G is a Lie group then we can expand each term to evaluate this:

$$g = 1 + i\alpha^a T_a + \frac{1}{2}(i\alpha^a T_a)^2 + \mathcal{O}(\alpha^3), \tag{6.2.1}$$

$$h = 1 + i\beta^a T_a + \frac{1}{2}(i\beta^a T_a)^2 + \mathcal{O}(\beta^3), \tag{6.2.2}$$

$$f = 1 + i\gamma^a T_a + \frac{1}{2}(i\gamma^a T_a)^2 + \mathcal{O}(\gamma^3). \tag{6.2.3}$$

The inverse of g can similarly be expanded as

$$g^{-1} = 1 - i\alpha^a T_a + \frac{1}{2}(i\alpha^a T_a)^2 + \mathcal{O}(\alpha^3), \tag{6.2.4}$$

this can be checked by expanding gg^{-1} and showing that we get 1. Write α for $\alpha^a T_a$ and β for $\beta^a T_a$, then we have

$$ghg^{-1}h^{-1} \approx \left(1 + i\alpha + \frac{1}{2}(i\alpha)^2\right)\left(1 + i\beta + \frac{1}{2}(i\beta)^2\right) \tag{6.2.5}$$

$$\left(1 - i\alpha + \frac{1}{2}(i\alpha)^2\right)\left(1 - i\beta + \frac{1}{2}(i\beta)^2\right) \tag{6.2.6}$$

$$= 1 + i\alpha + i\beta - i\alpha - i\beta \tag{6.2.7}$$

$$- \alpha\beta + \alpha\alpha + \alpha\beta + \beta\alpha + \beta\beta - \alpha\beta \tag{6.2.8}$$

$$+ \frac{1}{2}[(i\alpha)^2 + (i\beta)^2 + (i\alpha)^2 + (i\beta)^2] + \mathcal{O}(\alpha^3, \beta^3) \tag{6.2.9}$$

$$= -\alpha\beta + \beta\alpha + \mathcal{O}(\alpha^3, \beta^3). \tag{6.2.10}$$

That is,

$$f = -\alpha^a T_a \beta^b T_b + \beta^b T_b \alpha^a T_a + \mathcal{O}(\alpha^3, \beta^3) = -[\alpha^a T_a, \beta^b T_b] + \mathcal{O}(\alpha^3, \beta^3) \tag{6.2.11}$$

where $[-,-]$ is the usual commutator.

From here we define the Lie algebra of the Lie group G as the Lie algebra, $\mathfrak{g}$, generated by

$$T_a = -i\left(\frac{\partial g(\alpha)}{\partial \alpha^a}\right)_{\alpha=0} \tag{6.2.12}$$

with the Lie bracket given by the commutator, which is such that

$$[T_a, T_b] = ic^c{}_{ab} T_c \tag{6.2.13}$$

6.3 Exponential Map

Let G be a Lie group. Consider some one-dimensional Lie subgroup, $G_1 \subseteq G$. We can think of G_1 as a path on the manifold G parametrised by some parameter t, so

$$G_1 = \{g(\alpha(t))t \in \mathbb{R}\}, \tag{6.3.1}$$

we then write $g(t)$ for $g(\alpha(t))$.

Since this is a one-dimensional Lie group by Theorem 3.3.4 we can choose the parametrisation to be such that $g(t)g(s) = g(t+s)$ and $g(0) = 1$. Now differentiate this with respect to s and then set $s = 0$, we get

$$g(t)g'(0) = g'(t). \tag{6.3.2}$$

A solution to this is

$$g(t) = e^{g'(0)t}. \tag{6.3.3}$$

Now, define $g'(0) = i\gamma^a T_a$, since this is what we get if $g(s) = 1 + i\gamma^a T_a + \mathcal{O}(\gamma^2)$ and we find that we can write an arbitrary element of G_1 as

$$g(t) = e^{it\gamma^a T_a}. \tag{6.3.4}$$

This motivates the following theorem, which we won't prove.

Theorem 6.3.5. Let G be a compact Lie group with Lie algebra $\mathfrak{g}$. Let $g \in G$ be continuously connected to the identity. Then

$$g = e^{i\gamma^a T_a} \tag{6.3.6}$$

for some γ^a and T_a the generators of $\mathfrak{g}$.

In words, if G is a compact Lie group then every element connected to the identity can be obtained by exponentiating the Lie algebra.

A compact Lie group is a Riemannian manifold. The exponential map, $t \mapsto \exp[it\gamma^a T_a]$, then gives geodesics through the origin on this manifold. Every point connected to the identity is then on one of these geodesics. Intuitively this makes sense because by definition such a point is connected to the identity and so by simply minimising the path connecting it we get a geodesic.

6.3.1 Noncompact Lie Groups

The compact requirement is important, for noncompact Lie groups the theorem does not hold, which we demonstrate now by example. The group $\mathrm{SL}(2, \mathbb{R})$ is noncompact, it is formed of elements of the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{6.3.7}$$

with $ab - cd = 1$. The values of a, b, c, and d are unbounded, and so $\mathrm{SL}(2, \mathbb{R})$ is not compact.

Suppose that all $g \in \mathrm{SL}(2, \mathbb{R})$ can be expressed as $\exp[l]$ for some $l \in \mathfrak{sl}(2, \mathbb{R})$. Then we can define a one-dimensional path in $\mathrm{SL}(2, \mathbb{R})$ as the image of the map $t \mapsto \exp[tl]$ with $t \in [0, 1]$. In particular, $h = \exp[l/2]$ satisfies $h^2 = g$, so we can talk of "square roots" of group elements.

We now proceed to demonstrate that such an h does not exist for all choices of g. Let

$$g = \begin{pmatrix} -4 & 0 \\ 0 & -1/4 \end{pmatrix} \tag{6.3.8}$$

Then suppose that

$$h = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{6.3.9}$$

We then have

$$h^2 = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix}. \tag{6.3.10}$$

Since g has zero off diagonal elements we have $b(a + d) = c(a + d) = 0$, meaning either b and c are both zero or $a + d = 0$. Suppose that $b = c = 0$, then looking at the first entry in g we have $a^2 + bc = a^2 = -4$, but this can't be the case as a is real. Suppose then that $a + d = 0$, then $a = -d$ and so $a^2 = d^2$. The first entry in g gives us $a^2 + bc = -4$, but the last gives $d^2 + bc = a^2 + bc = -1/4$. It is not possible for $a^2 + bc$ to have two separate values so we conclude that there is no way to satisfy the conditions and there is no $h \in \mathrm{SL}(2, \mathbb{R})$ satisfying $h^2 = g$.

Note that $\mathrm{SL}(2, \mathbb{R})$ is connected, so it's the lack of compactness that is causing the problems here. We can show that g is connected to the identity by considering the piecewise path first defined by

$$\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \tag{6.3.11}$$

for $\vartheta \in [0, \pi]$, which gets us from the identity to

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{6.3.12}$$

and then the path given by

$$\begin{pmatrix} -\lambda & 0 \\ 0 & -\frac{1}{\lambda} \end{pmatrix}, \tag{6.3.13}$$

for $\lambda \in [1, 4]$, which ends us at

$$\begin{pmatrix} -4 & 0 \\ 0 & -\frac{1}{4} \end{pmatrix}. \tag{6.3.14}$$

That is, the path given by the image of

$$t \mapsto \begin{cases} \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} & \vartheta = t \in [0, \pi], \\ \begin{pmatrix} -\lambda & 0 \\ 0 & -\frac{1}{\lambda} \end{pmatrix} & \lambda = t + 1 - \pi, t \in [\pi, \pi + 3]. \end{cases} \tag{6.3.15}$$

This path is continuous (but not differentiable at $t = \pi$).

6.3.2 Baker–Campbell–Hausdorff

For two matrices, A and B, we have

$$\exp[A]\exp[B] = \exp\left[A + B + \frac{1}{2}[A, B] + \frac{1}{12}([A, [A, B]] + [[A, B], B]) + \cdots\right]$$

If we know that $A, B \in \mathfrak{g}$ for some Lie algebra $\mathfrak{g}$ then we can evaluate the commutators and hence can determine the product of $\exp[A]$ and $\exp[B]$ in G.

6.4 Matrix Lie Algebras

6.4.1 General Linear

Consider the general linear group, $\mathrm{GL}(n, \Bbbk)$. This consists of all invertible $n \times n$ matrices over $\Bbbk$. Now let A be any $n \times n$ matrix over $\Bbbk$. Then $\exp[A]$ is an invertible $n \times n$ matrix over $\Bbbk$, in particular $\exp[A]^{-1} = \exp[-A]$. This shows that the Lie algebra of $\mathrm{GL}(n, \Bbbk)$, denoted $\mathfrak{gl}(n, \Bbbk)$, is the set of all $n \times n$ matrices over $\Bbbk$. Note that this can be identified with the set of all linear transformations, $\mathfrak{gl}(n, \Bbbk) \cong \mathrm{End}(\Bbbk^n)$. Since we have n^2 entries into an $n \times n$ matrix, and no conditions, we conclude that $\dim \mathfrak{gl}(n, \mathbb{R}) = n^2$. Similarly, if $\Bbbk = \mathbb{C}$ then each entry has two real parameters and so we have $2n^2$ degrees of freedom so the dimension of $\mathfrak{gl}(n, \mathbb{C})$ as a real vector space is $2n^2$.

Now consider the special linear group, $\mathrm{SL}(n, \Bbbk)$. We then have the extra condition that $\det M = 1$ for all $M \in \mathrm{SL}(n, \Bbbk)$. To proceed we need the following lemma

Lemma 6.4.1 Let A be an $n \times n$ matrix and I the n-dimensional identity matrix. Then

$$\det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr} A \tag{6.4.2}$$

to first order in ε.

Proof. Let a_i be the eigenvalues of A. Then the characteristic polynomial of A is

$$\det(tI - A) = (t - a_1)(t - a_2) \cdots (t - a_n). \tag{6.4.3}$$

Setting $t = -1$ we get

$$\begin{aligned}
\det(-I - A) = (-1)^n \det(I + A) & \qquad (6.4.4)\\
= (-1 - a_1)(-1 - a_2) \cdots (-1 - a_n) & \qquad (6.4.5)\\
= (-1)^n (1 + a_1)(1 + a_2) \cdots (1 + a_n) & \qquad (6.4.6)\\
= (-1)^n [1 + a_1 + a_2 + \cdots + a_n + \mathcal{O}(a_i a_j)] & \qquad (6.4.7)\\
= (-1)^n [1 + \operatorname{tr} A + \mathcal{O}(a_i a_j)]. & \qquad (6.4.8)
\end{aligned}$$

Hence $\det(I + A) = 1 + \operatorname{tr} A + \mathcal{O}(a_i a_j)$ Rescaling so that $A \to \varepsilon A$ we get the result $\det(I + \varepsilon A) = 1 + \varepsilon \operatorname{tr} A + \mathcal{O}(\varepsilon^2)$. □

Using this we can see that if $M \in \mathrm{SL}(n, \Bbbk)$ we can expand $M = I + i\alpha^a T_a + \mathcal{O}(\alpha^2)$ and we get

$$1 = \det M \approx \det(I + i\alpha^a T_a) \approx 1 + \alpha^a \operatorname{tr} T_a, \tag{6.4.9}$$

and so the generators of $\mathfrak{sl}(n, \Bbbk)$ must be traceless, and hence all matrices in $\mathfrak{sl}(n, \Bbbk)$ are traceless. So, $\mathfrak{sl}(n, \Bbbk)$ consists of all traceless $n \times n$ matrices. We have n^2 entries, but being traceless fixes one, since we can force a matrix to be traceless by setting $A_{11} = -A_{22} - A_{33} - \cdots - A_{nn}$, and so $\dim \mathfrak{sl}(n, \mathbb{R}) = n^2 - 1$ and being traceless in both the real and imaginary components fixes two degrees of freedom in the complex case, so $\dim \mathfrak{sl}(n, \mathbb{C}) = 2(n^2 - 1)$ as a real vector space.

Now consider the orthogonal group, $\mathrm{O}(n)$. Writing some generic $O \in \mathrm{O}(n)$ as $I + \varepsilon A + \mathcal{O}(\varepsilon^2)$ and expanding the defining condition, $O^\top O = I$, we get

$$\begin{aligned} I &= (I + \varepsilon A + \mathcal{O}(\varepsilon^2))(I + \varepsilon A^\top + \mathcal{O}(\varepsilon^2)) \\ &= I + \varepsilon A + \varepsilon A^\top + \mathcal{O}(\varepsilon^2) = I + \varepsilon(A + A^\top) + \mathcal{O}(\varepsilon^2). \end{aligned} \tag{6.4.10}$$

So we must have that $A = -A^\top$, that is A is antisymmetric. So the Lie algebra, $\mathfrak{so}(n)$ consists of all $n \times n$ real antisymmetric matrices. Being antisymmetric sets all values below the diagonal, and the diagonal must be zero, so we get an upper triangle to fill in of base $n - 1$, the number of entries in this triangle is $n(n - 1)/2$.

Consider the special orthogonal group, $\mathrm{SO}(n)$. As with $\mathrm{O}(n)$ the matrices in the Lie algebra must be antisymmetric. But, this then enforces that these matrices are traceless, since an antisymmetric matrix has zeros down the diagonal, and hence $\mathfrak{o}(n) \cong \mathfrak{so}(n)$.

Now consider the unitary group, $\mathrm{U}(n)$. Suppose that $H \in \mathfrak{u}(n)$, that is $\exp[i\varepsilon H] \in \mathrm{U}(n)$, then $\exp[i\varepsilon H]\exp[i\varepsilon H]^\dagger = I$ by definition, and we have

$$I = \exp[i\varepsilon H]\exp[i\varepsilon H]^\dagger \approx (I + i\varepsilon H)(I - i\varepsilon H^\dagger) = I + i\varepsilon(H - H^\dagger) + \mathcal{O}(\varepsilon^2), \tag{6.4.11}$$

and so we must have $H = H^\dagger$, so H is Hermitian. Note that if we had not included the factor of i we would instead have found H to be anti-Hermitian, which isn't as nice and this is why we include the factor of i. So, $\mathfrak{u}(n)$ consists of all $n \times n$ Hermitian matrices. Being Hermitian fixes the values below the diagonal, removing $n(n - 1)/2$ entries, but $n(n - 1)$ degrees of freedom, since these are complex numbers with two real parameters. Being Hermitian also forces the diagonal to be real, removing another n parameters from the imaginary part of the diagonal. This leaves $n(n-1)$ real parameters defining the upper triangle and n defining the diagonal, for a total of $n(n - 1) + n = n^2$ real parameters, so $\dim \mathfrak{u}(n) = n^2$.

Finally, consider $\mathrm{SU}(n)$. This consists of all Hermitian, traceless, $n \times n$ matrices. Being traceless fixes one real parameter on the diagonal, since the diagonal is real anyway, and so $\dim \mathfrak{su}(n) = n^2 - 1$.

6.5 Universal Covering Group

Every Lie group has a unique Lie algebra which we find by linearising the group. Every Lie algebra can be exponentiated to form a Lie group. However, this Lie group is not necessarily the same one that we linearised to find the Lie algebra. The exponentiated Lie algebra forms a simply connected Lie group, even if the original Lie group wasn't simply connected. This means that multiple Lie groups can have the same Lie algebra, but only one of these groups is attained by exponentiating the Lie algebra, this Lie group is called the **universal covering group**.

If two Lie groups have the same Lie algebra it is because they are indistinguishable in the neighbourhood of the identity.

> **Example 6.5.1** Consider $\mathrm{U}(1)$. This has elements $\mathrm{e}^{i\vartheta}$ for some $\vartheta \in [0, 2\pi)$. Since we identify 0 and 2π this group is compact, and topologically equivalent to the circle, which is not simply connected. The Lie algebra is then just $[0, 2\pi)$.
> Consider instead the set of elements of the form e^{ϑ} for $\vartheta \in \mathbb{R}$. This gives

$\mathbb{R}_{>0}$, the set of positive real numbers. However, the Lie algebra is, up to an unimportant factor of i, the same as for $\mathrm{U}(1)$. The ray consisting of the positive real numbers is simply connected, and hence exponentiating this Lie algebra gives this group, rather than $\mathrm{U}(1)$.

Example 6.5.2 We've seen that both $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ have the same Lie algebra. This is because the requirement of unit determinant is not relevant in the neighbourhood of the identity, where it is satisfied in both groups.
The orthogonal group, $\mathrm{O}(n)$, is formed of two disconnected components, one with determinant $+1$ and one with determinant -1. The piece with determinant $+1$ is just the subgroup $\mathrm{SO}(n)$. The characterisation of the piece determinant -1 depends on the dimension. If n is odd then we can form pairs $(A, +1)$ and $(A, -1)$ for all $A \in \mathrm{SO}(n)$. Then identifying $(A, +1) = A \in \mathrm{O}(n)$ and $(A, -1) = -A \in \mathrm{O}(n)$ we get a one-to-one correspondence between these pairs and $\mathrm{O}(n)$. That means we have

$$\mathrm{O}(n) \cong \mathrm{SO}(n) \times \mathbb{Z}_2 \qquad n \text{ odd.} \tag{6.5.3}$$

If instead n is even then, for example, both I and $-I$ are in $\mathrm{SO}(n)$, and so we cannot do the same thing with pairs. The result is that the two components don't commute with each other.
Consider the case where $n = 2$. We can then split $\mathrm{O}(2)$ into two parts, the first formed of matrices of the form

$$g = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \in \mathrm{SO}(2), \tag{6.5.4}$$

and the second formed of matrices of the form

$$\bar{g} = \begin{pmatrix} -\cos\psi & \sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \in \overline{\mathrm{SO}(2)}. \tag{6.5.5}$$

That is, $\mathrm{SO}(2)$ consists of all matrices in $\mathrm{O}(2)$ with determinant 1 and $\overline{\mathrm{SO}(2)}$ consists of all matrices in $\mathrm{O}(2)$ with determinant -1. Notice that if $\bar{g}, \bar{h} \in \overline{\mathrm{SO}(2)}$ then we have $\det(\bar{g}\bar{h}) = \det(\bar{g})\det(\bar{h}) = (-1)(-1) = 1$, so $\bar{g}\bar{h} \in \mathrm{SO}(2)$.
It can be shown that $\mathrm{SO}(2)$ is a normal subgroup of $\mathrm{O}(2)$, that is $ghg^{-1} \in \mathrm{SO}(2)$ for all $g \in \mathrm{O}(2)$ and $h \in \mathrm{SO}(2)$. In particular, if $g \in \mathrm{SO}(2)$ and $\bar{g} \in \overline{\mathrm{SO}(2)}$ then $\bar{g}g\bar{g}^{-1} \in \mathrm{SO}(2)$.
If we have a normal subgroup then we can write the original group as a semidirect product of the normal subgroup and some other group, in this case

$$\mathrm{O}(2) \cong \mathrm{SO}(2) \rtimes \mathbb{Z}_2, \tag{6.5.6}$$

where the action of $\mathbb{Z}_2$ on $\mathrm{SO}(2)$ is

$$0 \,.\, g = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} g, \qquad \text{and} \qquad 1 \,.\, g = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} g. \tag{6.5.7}$$

Example 6.5.8 Perhaps the most important example, at least for physics, of two Lie groups with the same Lie algebras is SO(3) and SU(2). Start with SU(2), this has a Lie algebra formed from all Hermitian traceless 2×2 matrices. A basis for these matrices is given by the Pauli matrices, although to we include a factor of $1/2$ as a matter of convention, giving

$$\mathfrak{su}(2) = \operatorname{span}_{\mathbb{R}}\left\{\frac{1}{2}\sigma_a\right\}. \tag{6.5.9}$$

This Lie algebra then satisfies

$$[\sigma_a/2, \sigma_b/2] = i\varepsilon_{cab}\sigma_c/2, \tag{6.5.10}$$

that is, the structure constants are ${c^c}_{ab} = \varepsilon_{cab}$. Note that we are free to raise and lower indices here since the Killing form (defined later) is positive definite.

For SO(3) the Lie algebra is formed from the matrices T_a with components $(T_a)_{bc} = i\varepsilon_{bac}$. This is simply the adjoint representation of $\mathfrak{so}(2)$ as defined above, so clearly these two Lie algebras are the same.

Consider what happens when we exponentiate them. We can consider

$$\exp\left[i\omega\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}/2\right] \tag{6.5.11}$$

where our parameters, $\omega\hat{\boldsymbol{n}}$, are formed from a unit vector, $\hat{\boldsymbol{n}} \in S^2$, and some $\omega \in \mathbb{R}$. Note that a unit vector in n dimensions requires $n-1$ parameters to define. Using the identity $\sigma_i\sigma_j = \delta_{ij}I + i\varepsilon_{ijk}\sigma_k$ we can expand the exponential and collect terms to show that

$$\exp[i\omega\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}/2] = \cos\left(\frac{\omega}{2}\right)I + i\hat{\boldsymbol{n}} \cdot \boldsymbol{\sigma}\sin\left(\frac{\omega}{2}\right). \tag{6.5.12}$$

Thus we have to take $\omega \in [0, 4\pi)$ in order to cover every element of SU(2), at least those which can be reached in this way.

Instead we can consider SO(3) and exponentiate

$$\exp[i\omega\hat{\boldsymbol{n}}(i\varepsilon_{bac})] = n_b n_c + (\delta_{bc} - n_b n_c)\cos\omega + \varepsilon_{abc}n_a\sin\omega. \tag{6.5.13}$$

This requires a lot of algebra to show, but in the *Vectors, Tensors, and Continuum Mechanics* part of the *Methods of Theoretical Physics* course we show that this is the general form of a rotation. Note then that if $\omega \in [0, 4\pi)$ as for SU(2) we will hit every element of SO(3) twice. We say that SU(2) is a double cover of SO(3).

As a manifold SO(3) corresponds to the three-dimensional ball of radius π, where a vector in this ball picks out an axis, and its length picks out the angle of rotation. In particular, a vector, $\boldsymbol{x}$, of length π is equivalent as a rotation to the vector $-\boldsymbol{x}$, and so we identify opposite points on the ball. This means that a line going from one side of the ball to the other is technically a loop, but is not contractible to a point so SO(3) is not simply connected.

On the other hand, we showed in Example 3.2.18 that SU(2) as a manifold is the three sphere, S^3, which is simply connected. We conclude that both the Lie algebras of SU(2) and SO(3) both exponentiate to give SU(2).

Index