# A Scholar's Review of Lie Groups and Algebras

Barak Shoshany*

*Perimeter Institute for Theoretical Physics*
*31 Caroline St. N., Waterloo, ON, Canada, N2L 2Y5*

August 24, 2019

**Abstract**

These lecture notes, prepared for the Perimeter Scholars International master's program at Perimeter Institute, provide a detailed and self-contained introduction to Lie groups, Lie algebras, and their representations. The notes first review fundamental concepts in differential geometry and abstract algebra, define Lie groups and algebras, and discuss the relation between a Lie group and its associated Lie algebra in full mathematical detail, along with some subtleties. An exhaustive list of common matrix Lie groups, their associated Lie algebras, and their topological properties is provided, along with a detailed discussion of the Lorentz group. Representation theory is then introduced, along with some relevant concepts, followed by a thorough derivation of some of the irreducible representations of several popular Lie groups.

## Contents



*bshoshany@perimeterinstitute.ca

# 1 Introduction

These lecture notes are intended to provide students with some essential knowledge of Lie groups, algebras, and their representations, which are ubiquitous in theoretical physics, especially in the context of quantum mechanics, gauge theories, and quantum gravity.

The notes are structured as follows:

- Chapter 2 provides the mathematical concepts and definitions necessary to understand Lie groups and algebras at the abstract level, and in particular the precise relation between Lie groups and their corresponding Lie algebras.
- Chapter 3 presents an organized and comprehensive discussion of matrix Lie groups, which are the most common Lie groups encountered in theoretical physics, along with their associated algebras and topological properties.
- Chapter 4 discusses representations of Lie groups and algebras at the abstract level, along with some relevant concepts.
- Finally, Chapter 5 derives some of the representations of several popular matrix Lie groups: U (1), SO (2), SU (2), and SO (3).

Should the reader encounter any typos or mistakes, please report them to the author at the email address provided above.

# 2 Abstract Lie Groups and Algebras

## 2.1 Manifolds

### 2.1.1 Smooth Manifolds

The Earth is shaped roughly like a spheroid. Therefore it has curvature, and its topology is compact. However, a person standing on the surface of the planet and looking around will not see this curvature; their immediate surroundings will look flat. Moreover, they will not notice the compact topology; the Earth seems to keep going as far as the eye can see. They might then conclude that the entire Earth is, in fact, a flat Euclidean plane, and not a spheroid! However, this is only true locally; satellite photos, for example, reveal the global structure of the Earth.

An $n$-dimensional *manifold*, or $n$-manifold, generalizes this idea: it is a space which **locally** looks like $\mathbb{R}^n$, although **globally** it may be more complicated. More precisely, an $n$-manifold is a topological space equipped with an *atlas*, which is a collection of *charts*. These charts, also known as *coordinate* *charts* or *neighborhoods*, are open sets which cover the manifold – that is, their union is the whole manifold. This means that any point on the manifold is in at least one chart. Each chart is *homeomorphic* to $\mathbb{R}^n$, or in other words, there is a continuous invertible function from each chart to $\mathbb{R}^n$.

The charts are open sets, so to cover the whole manifold they must overlap. When two charts overlap, there is a *transition function* which allows us to move from one chart to the other in a well-defined way. More precisely, this transition function provides a *change of coordinates* from one coordinate chart to another.

Generally, it is impossible to cover a whole manifold with just one chart. As a simple example, consider the circle $S^1$. Points on the circle may be given by an angle $\theta$; however, the angles are in the interval $[0, 2\pi)$, which is not open and thus cannot be a chart. Furthermore, we cannot take our chart to be a larger interval, such as $(-\pi, 2\pi)$, since then some points will be described by two different values of the coordinate, and the map cannot be invertible. Instead, we must cover the circle with two or more overlapping charts, related by transition functions.

In a *differentiable manifold*, the transition functions are all continuously differentiable a finite number of times, that is, they belong to $C^k$ for some $k \in \mathbb{N}$. In a *smooth manifold*, the transition functions are smooth, that is, infinitely differentiable or $C^\infty$. We will assume all manifolds are smooth in these notes.

### 2.1.2 Tangent Spaces

A *tangent space* to a manifold at a point is simply the collection of vectors that are tangent to the manifold at that point. The tangent space to the manifold $M$ at a point $p \in M$ is denoted $T_pM$. For example, the tangent space $T_pS^2$ to the sphere $S^2$ at the point $p \in S^2$ is the plane orthogonal to the radial vector pointing from the origin to $p$. Of course, this definition relies on the fact that we can embed the sphere in a 3-dimensional ambient space. We should instead define a tangent space in an abstract way, without using an embedding.

A nice way to define the tangent space $T_pM$ is as follows. Let $f, g \in C^\infty(M)$ be smooth functions on the manifold, $f, g : M \to \mathbb{R}$. We define a *tangent vector* $\mathbf{x}$ at the point $p \in M$ as a map which takes a smooth function $f \in C^\infty(M)$ to a number $\mathbf{x}[f]$,

$$\mathbf{x} : C^\infty(M) \to \mathbb{R} \quad \Longrightarrow \quad f \mapsto \mathbf{x}[f], \tag{2.1}$$

and satisfies the following axioms:

1. $\mathbf{x}[f + g] = \mathbf{x}[f] + \mathbf{x}[g]$,
2. $\mathbf{x}[\alpha f] = \alpha\,\mathbf{x}[f]$ where $\alpha \in \mathbb{R}$,
3. $\mathbf{x}[fg] = \mathbf{x}[f]\,g + f\mathbf{x}[g]$.

A map satisfying these axioms is also called a *derivation*. The first two axioms ensure that it is a linear map, while the third is a generalization of the familiar Leibniz rule. Furthermore, given a number $\alpha \in \mathbb{R}$ and another tangent vector $\mathbf{y}$, we define

$$(\alpha\mathbf{x})[f] \equiv \alpha\,\mathbf{x}[f], \qquad (\mathbf{x} + \mathbf{y})[f] \equiv \mathbf{x}[f] + \mathbf{y}[f]. \tag{2.2}$$

Then it is easy to see that the tangent vectors indeed form a vector space. We call that space the tangent space to $M$ at $p$, or $T_pM$.

Now, a *curve* $\gamma$ is a function $\gamma : \mathbb{R} \to M$ mapping real numbers to points on the manifolds in a smooth way, that is, such that for any smooth function $f \in C^\infty(M)$, the composition $(f \circ \gamma)(t) \equiv f(\gamma(t))$ depends smoothly on the real parameter $t$. The curve has a tangent vector $\gamma'(t)$ at each point $\gamma(t) \in M$. We may imagine the curve as indicating the position of a particle on the manifold at time $t$, in which case $\gamma'(t)$ is the velocity vector, indicating the particle's instantaneous velocity at time $t$.

The meaning of the derivative $\gamma'(t)$ is intuitively clear, but to define it rigorously, we must use its action on a function $f \in C^{\infty}(M)$:

$$\gamma'(t) : C^{\infty}(M) \to \mathbb{R} \quad \Longrightarrow \quad \gamma'(t)[f] \equiv (f \circ \gamma)'(t). \tag{2.3}$$

One can easily check that $\gamma'(t)$ satisfies the three axioms above. Therefore, we may define the tangent space $T_pM$ as the collection of tangent vectors $\gamma'(0)$ to all the curves passing through $p$ at time $t = 0$, that is, curves satisfying $\gamma(0) = p$.

## 2.2 Lie Groups

### 2.2.1 Groups

Let us recall that a *group* $G$ is a set of elements, along with a product, satisfying the following axioms:

1. **Closure:** For any two elements $g, h \in G$, the product $gh$ is also in $G$.
2. **Associativity:** For any three elements $f, g, h \in G$, we have $(fg)h = f(gh)$.
3. **Identity element**[1]**:** There exists a (unique) element $I \in G$ such that, for any element $g \in G$, we have $Ig = gI = g$.
4. **Inverse element:** For any element $g \in G$ there exists a unique element $g^{-1}$ such that $gg^{-1} = g^{-1}g = I$.

If the product is commutative, $gh = hg$ for all $g, h \in G$, we say that the group is *Abelian*. However, the product is usually not commutative.

Note that the product can be any binary operation, not just multiplication; for example, the integers with the addition operation, $m + n$ for $m, n \in \mathbb{Z}$, form an (Abelian) group. However, in theoretical physics, and especially in the context of representation theory, we mostly use the notation $gh$ for the group product, since the group elements are represented as matrices, and the product is then matrix multiplication.

### 2.2.2 Lie Groups and Their Actions

We are now ready to define a *Lie group*: it is a group which is also a manifold. In other words, it is a set of elements which satisfy the group axioms, with the additional requirement that a neighborhood of each element is homeomorphic to $\mathbb{R}^n$. Furthermore, the group product and the inversion $g \mapsto g^{-1}$ are both required to be smooth maps. The *dimension* of the Lie group is $n$, the dimension of the manifold.

One of the most powerful concepts in physics is *symmetries*: transformations under which a particular object is invariant. Some symmetries are discrete; for example, a square is invariant under four rotations: 0, $\pi/2$, $\pi$, and $3\pi/2$ radians, and four flips: vertical, horizontal, and the two diagonals. This forms a *discrete symmetry group* with 8 elements, known as the *dihedral group* of order 8 and denoted $D_4$. However, symmetries can also be continuous; for example, a circle is invariant under rotations by any angle. This forms a *continuous symmetry group* with a continuum of elements, known as the *circle group*, or the *unitary group* of degree 1, and denoted $U(1)$.

Lie group are generally used in physics as continuous symmetry groups, and thus their elements represent symmetries of a certain object. The *group action* defines how elements of the group act on the object; for example, the action of an element $g \in U(1)$ on the circle rotates the circle by the angle corresponding to this element.

Given a group $G$ and a space $X$, we define the action of $G$ on $X$ as follows. The *left group action* is a function $G \times X \to X$, denoted for $g \in G$ and $x \in X$ as $g \triangleright x$, or simply $gx$, such that

[1] Usually, the identity element is labeled $e$. Here we instead use $I$, in order to avoid confusion with the exponential map. $I$ is the standard notation for the identity matrix, and in a matrix Lie group, the identity element is indeed the identity matrix.

1. $I \triangleright x = x$ for all $x \in X$, where $I \in G$ is the identity element, and
2. $(gh) \triangleright x = g \triangleright (h \triangleright x)$ for all $g, h \in G$ and $x \in X$, that is, the action of the product $gh$ is the same as first applying $h$ and then applying $g$ to the result.

Note that if $X = G$, that is, the group is acting on itself (via the usual group product), then these two axioms are automatically satisfied. The axioms mean that, for every $g \in G$, the left action is an invertible map from $X$ to itself, which is exactly what we would expect from a group of transformations.

We can similarly define the *right group action*, a function $X \times G \to X$, denoted for $g \in G$ and $x \in X$ as $x \triangleleft g$, or simply $xg$, such that $x \triangleleft I = x$ and $x \triangleleft (gh) = (x \triangleleft g) \triangleleft h$. Whether the action is left or right ultimately amounts to which elements acts first when the product $gh$ acts on $x$: $h$ for a left action or $g$ for a right action, as can be seen from the axioms $(gh) \triangleright x = g \triangleright (h \triangleright x)$ and $x \triangleleft (gh) = (x \triangleleft g) \triangleleft h$. Thus, if we are given a right action, we can always construct a left action from it simply by mapping $x \triangleleft g \mapsto g^{-1} \triangleright x$. Indeed, we have that $(gh)^{-1} \triangleright x = h^{-1} g^{-1} \triangleright x$, so $g$ would still end up acting first, as would be expected from a right action. Therefore, we can limit ourselves to discuss left actions exclusively without loss of generality.

## 2.3 Lie Algebras

### 2.3.1 Algebras and Lie Algebras

An *algebra* is a vector space $\mathfrak{g}$ over a field $\mathbb{F}$ (which we will take to be either $\mathbb{R}$ or $\mathbb{C}$) equipped with a binary operation $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the following axioms:

- **Right distributivity:** For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathfrak{g}$, $[\mathbf{x} + \mathbf{y}, \mathbf{z}] = [\mathbf{x}, \mathbf{z}] + [\mathbf{y}, \mathbf{z}]$.
- **Left distributivity:** For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathfrak{g}$, $[\mathbf{x}, \mathbf{y} + \mathbf{z}] = [\mathbf{x}, \mathbf{y}] + [\mathbf{x}, \mathbf{z}]$.
- **Compatibility with scalars:** For all $\mathbf{x}, \mathbf{y} \in \mathfrak{g}$ and $\alpha, \beta \in \mathbb{F}$, $[\alpha \mathbf{x}, \beta \mathbf{y}] = \alpha\beta [\mathbf{x}, \mathbf{y}]$.

Note that the operation $[\cdot, \cdot]$ is not necessarily commutative, or even associative!

In the case of a Lie algebra, the binary operator $[\cdot, \cdot]$ is called the *Lie bracket* or *commutator*, and it is required to satisfy two additional axioms:

- **Alternativity:** For all $\mathbf{x} \in \mathfrak{g}$, $[\mathbf{x}, \mathbf{x}] = 0$.
- **The Jacobi identity:** For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathfrak{g}$, $[\mathbf{x}, [\mathbf{y}, \mathbf{z}]] + [\mathbf{y}, [\mathbf{z}, \mathbf{x}]] + [\mathbf{z}, [\mathbf{x}, \mathbf{y}]] = 0$.

The alternativity axiom implies anti-commutativity of the bracket:

$$0 = [\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}] = [\mathbf{x}, \mathbf{y}] + [\mathbf{y}, \mathbf{x}] \quad \Longrightarrow \quad [\mathbf{x}, \mathbf{y}] = -[\mathbf{y}, \mathbf{x}]. \tag{2.4}$$

A well-known example of a Lie algebra is the vector space $\mathbb{R}^3$ with the *cross product* $\times$. It's easy to check that it indeed satisfies all of the axioms above. Another well-known example is the set $\mathrm{M}(n, \mathbb{F})$ of real or complex $n \times n$ matrices, with the binary operation being the matrix commutator $[A, B] \equiv AB - BA$.

### 2.3.2 Generators of Lie Algebras and Structure Constants

Since a Lie algebra is a vector space, it has a basis. In this context, we call the basis vectors *generators*. The number of generators defines the *dimension* of the Lie algebra. We will denote the generators $\boldsymbol{\tau}_i$, where the index $i$ goes from 1 to the number of dimensions. Now, if we take the commutator of two generators, the result is by definition another vector in the algebra, and therefore it must be a linear combination of the generators. Thus we define:

$$\left[\boldsymbol{\tau}_i, \boldsymbol{\tau}_j\right] \equiv f_{ij}{}^k \boldsymbol{\tau}_k, \tag{2.5}$$

where $f_{ij}{}^k$ are called the *structure constants*. (Note that the Einstein summation convention is implied: any index which appears twice, once as an upper index and once as a lower index, is summed over.) Since the commutator is anti-symmetric and satisfies the Jacobi identity, the structure constants satisfy

$$f_{ij}{}^k = -f_{ji}{}^k, \qquad f_{ij}{}^l f_{kl}{}^m + f_{jk}{}^l f_{il}{}^m + f_{ki}{}^l f_{jl}{}^m = 0. \tag{2.6}$$

Sometimes in physics we instead use the convention[2] $\left[\boldsymbol{\tau}_i, \boldsymbol{\tau}_j\right] \equiv \mathrm{i}\, f_{ij}{}^k \boldsymbol{\tau}_k$, which differs from the math convention by a factor of i. In these notes we will use the math convention, except in Subsection 5.2.5 where we will use the physics convention.

## 2.4 The Lie Algebra Associated to a Lie Group

Every Lie group has a Lie algebra associated to it. In this section, we will see that the Lie algebra is the tangent space to the Lie group at the identity element, and it inherits its Lie bracket from the action of tangent vectors. However, in order to define this relation rigorously, we will first have to introduce some fundamental concepts in differential geometry.

It is important to note that while every Lie group has a Lie algebra, the relation is not one-to-one; two groups can have the same algebra. There are some other subtleties involved, depending on the topology of the group. We will discuss some of them below.

In general, the notation for the Lie algebra will use the same letters designating the Lie group it is associated with, written in lowercase Fraktur font. For example, the Lie algebra of $\mathrm{SU}(2)$ is $\mathfrak{su}(2)$. For a general Lie group $G$, the associated Lie algebra is written $\mathfrak{g}$. We will denote group elements with Roman font, $g \in G$, and algebra elements in bold font, $\mathbf{x} \in \mathfrak{g}$. This is to remind the reader that the Lie algebra elements are actually vectors in a vector space.

### 2.4.1 The Differential and Left-Invariant Vector Fields

Given a map $\phi : M \to N$ between a manifold $M$ and another manifold $N$, the *differential* or *pushforward* of $\phi$ at a point $p \in M$, denoted $\mathrm{d}\phi_p$, is the linear map

$$\mathrm{d}\phi_p : T_p M \to T_{\phi(p)} N \quad \Longrightarrow \quad \mathrm{d}\phi_p \left(\gamma'(0)\right) \equiv \left(\phi \circ \gamma\right)'(0), \tag{2.7}$$

where $\gamma : \mathbb{R} \to M$ is a curve such that $\gamma(0) = p$. To understand what this means, recall that $\gamma'(0)$ is a tangent vector at $p$. We can map the curve $\gamma : \mathbb{R} \to M$ to a curve $\phi \circ \gamma : \mathbb{R} \to N$, and the differential "pushes us forward" from the tangent vector $\gamma'(0) \in T_p M$ at $p = \gamma(0) \in M$ to the tangent vector $(\phi \circ \gamma)'(0) \in T_{\phi(p)} N$ at $\phi(p) = (\phi \circ \gamma)(0) \in N$.

In Subsection 2.2.2 we defined the group action, and we noted that any group can also act on itself, with the group action being the usual group product. For each group element $g \in G$, we define a corresponding left action:

$$L_g : G \to G \quad \Longrightarrow \quad L_g(h) \equiv gh. \tag{2.8}$$

[2] To avoid confusion, in these lecture notes the imaginary unit i is written in roman typeface, while the index $i$ is written in italics.

Now, let $\mathbf{x} \in \Gamma(TG)$ be a vector field[3] on the Lie group $G$, that is, a function assigning a tangent vector $\mathbf{x}(g) \in T_g G$ to each point $g \in G$ in the manifold. Given some $h \in G$, we can employ two different actions to combine it with $g$:

1. First compose $g$ with $h$ to get $L_g(h) = gh$, and then calculate $\mathbf{x}(L_g(h)) = \mathbf{x}(gh)$;
2. First calculate $\mathbf{x}(h)$, and then push it forward from the point $h$ to the point $gh$ using the differential of the left action, $(\mathrm{d}L_g)(\mathbf{x}(h))$.

In general, these two actions will produce difference results. The vector field $\mathbf{x}$ is called *left-invariant* if, for all $g, h \in G$, these two actions coincide:

$$\mathbf{x}(gh) = (\mathrm{d}L_g)(\mathbf{x}(h)). \tag{2.9}$$

In other words, for a left-invariant vector field, if we calculate $\mathbf{x}$ at $gh$, or first calculate $\mathbf{x}$ at $h$ and then push the result forward to $gh$, we will get the same result.

The reason we care about left-invariant vector fields is that, if $\mathbf{x}$ is a left-invariant vector field, and we take $h = I$ (the identity element), then

$$\mathbf{x}(g) = (\mathrm{d}L_g)(\mathbf{x}(I)). \tag{2.10}$$

Hence, the value of $\mathbf{x}$ at any point $g \in G$ is completely determined by its value at the identity. Instead of looking at an entire vector field – a tangent vector at each group element – we can just look at the field's value at the identity, provided that it's left-invariant.

Conversely, if we have a tangent vector at the identity element, $\mathbf{v} \in T_I G$, we may define a vector field from it as follows:

$$\mathbf{x}(g) \equiv (\mathrm{d}L_g)(\mathbf{v}), \qquad \forall g \in G. \tag{2.11}$$

Obviously, $\mathbf{x}(I) = \mathbf{v}$. Furthermore, we have

$$\begin{aligned} \mathbf{x}(gh) &= (\mathrm{d}L_{gh})(\mathbf{v}) \\ &= \mathrm{d}(L_g \circ L_h)(\mathbf{v}) \\ (*) &= (\mathrm{d}L_g)(\mathrm{d}L_h(\mathbf{v})) \\ &= (\mathrm{d}L_g)(\mathbf{x}(h)), \end{aligned}$$

where in $(*)$ we used the chain rule. Hence, the vector field $\mathbf{x}$ thus defined is left-invariant. We conclude that there is an isomorphism – an invertible homomorphism – between the space of left-invariant vector fields on $G$ and the space of tangent vectors at the identity, $T_I G$.

### 2.4.2 The Lie Algebra of a Lie Group

Now we can properly define the Lie algebra $\mathfrak{g}$ of a Lie group $G$: it is **the set of all left-invariant vector fields** on $G$. One can check that this set is indeed a vector space. However, we are still missing one crucial ingredient – the Lie bracket. In Subsection 2.1.2 we defined tangent vectors as derivations: linear maps satisfying the Leibniz rule. Given two vector fields $\mathbf{x}, \mathbf{y}$ on $G$, we define their commutator by its action on functions $f \in C^\infty(G)$ as follows:

$$[\mathbf{x}, \mathbf{y}][f] \equiv \mathbf{x}[\mathbf{y}[f]] - \mathbf{y}[\mathbf{x}[f]]. \tag{2.12}$$

Here, by $\mathbf{x}[\mathbf{y}[f]]$ we mean: first act with the derivation $\mathbf{y}$ on the function $f$, which produces another function (a real number at each point on the manifold), and then act with $\mathbf{x}$ on the resulting function. One can check that $[\mathbf{x}, \mathbf{y}]$ satisfies the axioms for a derivation, and therefore it is a vector field as well.

[3] $TM$ is called the *tangent bundle* to the manifold $M$. It is simply the set of pairs $(p, \mathbf{x})$ such that $p \in M$ and $\mathbf{x} \in T_p M$, equipped with a *projection* $\pi : TM \to M$ such that $\pi(p, \mathbf{x}) = p$. A *section* of a bundle is the inverse of the projection, in the sense that it is a map $\sigma : M \to TM$ such that $\pi(\sigma(p)) = p$ for all $p \in M$. In other words, a section assigns, to each point on the manifold, a tangent vector in the tangent space at that point. This is exactly what a vector field does, and hence, a vector field is a section of the tangent bundle. The set of sections of the tangent bundle is denoted $\Gamma(TM)$.

### 2.4.3 Integral Curves and Flows

Let $\mathbf{x} \in \Gamma(TG)$ be a vector field on a manifold $G$, let $g \in G$ be a point on the manifold, and let $R \subseteq \mathbb{R}$ be an open interval containing 0. The *integral curve* of $\mathbf{x}$ passing through $g$ is a curve $\gamma : R \to G$ such that

$$\gamma'(t) = \mathbf{x}(\gamma(t)), \qquad \gamma(0) = g. \tag{2.13}$$

In other words, the curve passes through $g$ and the tangent vector to the curve at each point is given by the value of the vector field $\mathbf{x}$ at that point. If we imagine a vector field as a collection of arrows sprinkled on the manifold, then the velocity of the integral curve follows these arrows. (2.13) is an ordinary differential equation, and by the standard existence and uniqueness theorems for such equations, there exists a unique integral curve for any vector field $\mathbf{x}$ passing through any point $g$ for some interval $R \subseteq \mathbb{R}$ with $0 \in R$.

In particular, there exists a unique *maximal integral curve* $\gamma : R \to G$, where the interval $R$ is maximal. A vector field $\mathbf{x} \in \Gamma(TG)$ is called *complete* if every maximal integral curve $\gamma$ of that vector field is defined on all of $\mathbb{R}$, that is, $\gamma : \mathbb{R} \to G$. Given an integral curve $\gamma$ of a left-invariant vector field $\mathbf{x}$ passing through a group element $g$, the curve $\gamma_g \equiv L_g \circ \gamma$ is also an integral curve, due to $\mathbf{x}$ being left-invariant. Using this property, we may in fact extend any integral curve to $\mathbb{R}$, and thus $\mathbf{x}$ is complete (this is left as an exercise for the reader). Since the Lie algebra is the set of left-invariant vector fields, any vector in a Lie algebra is complete.

### 2.4.4 The Exponential Map

Using the fact that $\mathbf{x}$ is complete, we may define a smooth map $\phi_{\mathbf{x}} : \mathbb{R} \times G \to G$ such that, for all $g \in G$,

1. $\phi_{\mathbf{x}}(0, g) = g$,
2. $t \mapsto \phi_{\mathbf{x}}(t, g)$ is an integral curve of $\mathbf{x}$ passing through $g$.

We may collect all of these maps into a map $\phi : \mathbb{R} \times G \times \mathfrak{g} \to G$ defined by $\phi(t, g, \mathbf{x}) \equiv \phi_{\mathbf{x}}(t, g)$. This map satisfies:

1. $\phi(0, g, \mathbf{x}) = g$,
2. $t \mapsto \phi(t, g, \mathbf{x})$ is an integral curve of $\mathbf{x}$ passing through $g$,
3. $\phi(t, g, \mathbf{x}) = g\phi(t, I, \mathbf{x})$,
4. $\phi(t, g, s\mathbf{x}) = \phi(st, g, \mathbf{x})$,
5. $\phi(s, I, \mathbf{x})\,\phi(t, I, \mathbf{x}) = \phi(s + t, I, \mathbf{x})$.

Properties 1 and 2 follow automatically from the definition of $\phi_{\mathbf{x}}$. Property 3 follows from the fact that $\gamma(t) \equiv g\phi(t, I, \mathbf{x})$ is an integral curve of $\mathbf{x}$ with initial condition

$$\gamma(0) = g\phi(0, I, \mathbf{x}) = gI = g, \tag{2.14}$$

and since $\phi(t, g, \mathbf{x})$ is also an integral curve of $\mathbf{x}$ with the same initial condition, by the uniqueness property they must be equivalent. Properties 4 and 5 follow similarly.

By combining properties 3 and 4, we find:

$$\phi(t, g, \mathbf{x}) = g\phi(t, I, \mathbf{x}) = g\phi(1, I, t\mathbf{x}). \tag{2.15}$$

Therefore, the map $\phi$ is in fact completely determined by $\mathbf{x}$ alone: if we know $\phi(1, I, \mathbf{x})$, then for any $g$ and $t$ we can easily find $\phi(t, g, \mathbf{x})$. We call this map the *exponential map*:

$$\exp \mathbf{x} : \mathfrak{g} \to G, \qquad \exp \mathbf{x} \equiv \phi(1, I, \mathbf{x}). \tag{2.16}$$

Then $g \exp(t\mathbf{x})$ is the integral curve for $\mathbf{x}$ which passes through $g$ at $t = 0$. Furthermore, from property 5 we have

$$\exp(s\mathbf{x}) \exp(t\mathbf{x}) = \exp((s+t)\mathbf{x}). \tag{2.17}$$

We may also write $e^{\mathbf{x}} \equiv \exp \mathbf{x}$ for short. However, note that the usual identity $e^{\mathbf{x}+\mathbf{y}} = e^{\mathbf{x}} e^{\mathbf{y}}$ does **not** apply here, unless $\mathbf{x}$ and $\mathbf{y}$ are linearly dependent.

In summary, using the exponential map, we may obtain a group element $\exp \mathbf{x} \in G$ from an algebra element $\mathbf{x} \in \mathfrak{g}$.

# 3 The Matrix Lie Groups and Their Algebras

So far, we've dealt with abstract Lie groups and algebras. Now that we have a fairly good understanding of how they are defined, we will talk about the Lie groups that are the most common in physics: the matrix Lie groups.

Let $\mathrm{M}(n, \mathbb{F})$ be the set of $n \times n$ matrices with entries from the field $\mathbb{F}$, which we will assume is either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. This set is not a group with respect to matrix multiplication, since some matrices are not invertible. However, if we take only the invertible matrices, we obtain the Lie group $\mathrm{GL}(n, \mathbb{F})$, also known as the *general linear group*. Note that for $n \geq 2$ this group is non-Abelian, since matrix multiplication is non-commutative.

The group $\mathrm{GL}(n, \mathbb{F})$ and its subgroups are collectively known as the *matrix Lie groups*. Their associated Lie algebras are matrix Lie algebras, and the Lie bracket is the usual matrix commutator. In this chapter we will give a more down-to-earth definition of the exponential map for the case of a matrix Lie group, review some subtleties, recall the definition of a normal subgroup, and then list the most important matrix Lie groups, along with their dimensions, topological properties, and associated Lie algebras.

## 3.1 The Exponential Map for Matrix Lie Groups

If our Lie group $G$ is a matrix group with a matrix Lie algebra $\mathfrak{g}$, then the exponential map has a concrete definition:

$$e^{\mathbf{x}} \equiv \sum_{k=0}^{\infty} \frac{\mathbf{x}^k}{k!}, \tag{3.1}$$

where $\mathbf{x} \in \mathfrak{g}$, and $\mathbf{x}^k$ is simply the product of the matrix $\mathbf{x}$ with itself $k$ times. The identity element $I \in G$, which is the $n \times n$ identity matrix, is the exponential of the zero element $\mathbf{0} \in \mathfrak{g}$, which is the $n \times n$ zero matrix:

$$I = e^{\mathbf{0}}. \tag{3.2}$$

The exponential of a general algebra element $\mathbf{x} \in \mathfrak{g}$ will always result in some group element $g \in G$:

$$e^{\mathbf{x}} = g \in G. \tag{3.3}$$

If the matrices $\mathbf{x} \in \mathfrak{g}$ and $\mathbf{y} \in \mathfrak{g}$ commute, then

$$e^{\mathbf{x}+\mathbf{y}} = e^{\mathbf{x}} e^{\mathbf{y}}. \tag{3.4}$$

In particular, since $\mathbf{x}$ commutes with $-\mathbf{x}$, we have $\mathrm{e}^{\mathbf{x}}\,\mathrm{e}^{-\mathbf{x}} = I$. Therefore, the inverse $g^{-1} \in G$ will be the exponential of $-\mathbf{x} \in \mathfrak{g}$:

$$g^{-1} = \mathrm{e}^{-\mathbf{x}}. \tag{3.5}$$

Furthermore, for two scalars $s, t \in \mathbb{F}$, the matrices $s\mathbf{x}$ and $t\mathbf{x}$ commute, so

$$\mathrm{e}^{s\mathbf{x}}\,\mathrm{e}^{t\mathbf{x}} = \mathrm{e}^{(s+t)\mathbf{x}}. \tag{3.6}$$

We also have, for a scalar $t \in \mathbb{R}$ and a matrix $\mathbf{x} \in \mathfrak{g}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{t\mathbf{x}} = \mathbf{x}\,\mathrm{e}^{t\mathbf{x}} = \mathrm{e}^{t\mathbf{x}}\,\mathbf{x} \quad\Longrightarrow\quad \frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{t\mathbf{x}}\bigg|_{t=0} = \mathbf{x}. \tag{3.7}$$

Finally, we have the important relation[4]

$$\det \mathrm{e}^{\mathbf{x}} = \mathrm{e}^{\operatorname{tr}\mathbf{x}}. \tag{3.10}$$

Now, if $G$ is connected and compact[5], then the map $\mathbf{x} \mapsto \mathrm{e}^{\mathbf{x}}$ is *surjective (onto)*, meaning that each group element $g \in G$ has **at least** one element $\mathbf{x} \in \mathfrak{g}$ such that $g = \mathrm{e}^{\mathbf{x}}$. This is a sufficient condition, but not a necessary one; for example, even though $\mathrm{GL}(n, \mathbb{C})$ is non-compact, the map from the Lie algebra $\mathfrak{gl}(n, \mathbb{C})$ to the Lie group $\mathrm{GL}(n, \mathbb{C})$ is surjective.

Moreover, if the map $\mathbf{x} \mapsto \mathrm{e}^{\mathbf{x}}$ is *injective (one-to-one)*, meaning that each group element $g \in G$ has **at most** one element $\mathbf{x} \in \mathfrak{g}$ such that $g = \mathrm{e}^{\mathbf{x}}$, then $G$ is simply connected[6]. This is a necessary condition, but not a sufficient one; for example, even though $\mathrm{SU}(2)$ is simply connected, the matrix $\mathbf{x} \equiv \operatorname{diag}(\mathrm{i}\,\pi, -\mathrm{i}\,\pi) \in \mathfrak{su}(2)$ and its negative both map to the same matrix $\mathrm{e}^{\mathbf{x}} = \mathrm{e}^{-\mathbf{x}} \in \mathrm{SU}(2)$.

## 3.2 Normal Subgroups

### 3.2.1 Subgroups and Cosets

A *subgroup* is a subset of elements in the group which themselves form a group. For example, consider the general linear group $\mathrm{GL}(n, \mathbb{F})$, which consists of all $n \times n$ matrices over the field $\mathbb{F}$. If we restrict to the subset of matrices with determinant 1, we get the special linear group $\mathrm{SL}(n, \mathbb{F})$. It is easy to see that the properties of closure, associativity, identity element and inverse element are all satisfied for $\mathrm{SL}(n, \mathbb{F})$, so it is indeed a group, and hence a subgroup of $\mathrm{GL}(n, \mathbb{F})$.

Now, given a group $G$ and a subgroup $H$, the *left coset* $gH$ and *right coset* $Hg$ are the sets of elements formed by acting on every single element of $H$ with a particular element $g \in G$, either from the left or from the right:

$$gH \equiv \{gh \,|\, h \in H\}, \qquad Hg \equiv \{hg \,|\, h \in H\}, \qquad g \in G. \tag{3.11}$$

[4]To remember this, recall that the determinant is the product of the eigenvalues of the matrix, while the trace is the sum of the eigenvalues. In particular, for a diagonal matrix $\mathbf{x} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, we have

$$\operatorname{tr}\mathbf{x} = \sum_{i=1}^{n} \lambda_i. \tag{3.8}$$

One can easily calculate that $\mathrm{e}^{\mathbf{x}} = \operatorname{diag}\left(\mathrm{e}^{\lambda_1}, \ldots, \mathrm{e}^{\lambda_n}\right)$, and thus

$$\det \mathrm{e}^{\mathbf{x}} = \prod_{i=1}^{n} \mathrm{e}^{\lambda_i} = \mathrm{e}^{\sum_{i=1}^{n} \lambda_i} = \mathrm{e}^{\operatorname{tr}\mathbf{x}}. \tag{3.9}$$

The reader is invited to prove this relation for a general (not necessarily diagonal) matrix.

[5]A space is *connected* if it cannot be divided into two disjoint non-empty open sets. A space $X$ is *compact* if every one of its *open covers* – collections of open subsets of $X$ such that $X$ is the union of those subsets, just like a manifold is covered by an atlas of charts – has a finite subcover. Intuitively, this generalizes the notion of a *closed and bounded set* in Euclidean space to more abstract spaces.

[6]In a *simply connected* space, every path between two points can be continuously deformed into any other path between the same points, or equivalently, every loop can be continuously shrunk to a point.

Note that if $G$ is Abelian, then the left and right cosets are the same. Also, since $(gH)^{-1} = Hg^{-1}$, the number of left cosets is equal to the number of right cosets.

As a simple example, consider the group $(\mathbb{Z}, +)$ of the integers with the addition operation. The even integers $2\mathbb{Z}$ are clearly a subgroup of $(\mathbb{Z}, +)$. There are two possible cosets with respect to this subgroup: $2\mathbb{Z}$ itself (acting on every element of $2\mathbb{Z}$ with the identity element 0) and $2\mathbb{Z} + 1$, which is the subgroup of odd integers. However, note that, in general, cosets are not necessarily subgroups!

### 3.2.2 Definition of a Normal Subgroup

A *normal subgroup* $N$ of $G$ is a subgroup for which the left and right cosets are the same:

$$gN = Ng, \qquad \forall g \in G. \tag{3.12}$$

In other words, the normal subgroup $N$ "commutes" with every element of the group $G$. (Obviously, if $G$ is Abelian, every subgroup is automatically normal.) An equivalent definition is that a normal subgroup $N$ of $G$ is invariant under *conjugation* with any element of $G$:

$$ghg^{-1} \in N, \qquad \forall h \in N, \qquad \forall g \in G. \tag{3.13}$$

As an example, consider the group $\mathrm{GL}(n, \mathbb{F})$ and its subgroup $\mathrm{SL}(n, \mathbb{F})$. Given an element $g \in \mathrm{GL}(n, \mathbb{F})$ and an element $h \in \mathrm{SL}(n, \mathbb{F})$, we have:

$$\det\left(ghg^{-1}\right) = \det(g)\det(h)\det(g^{-1}) = 1, \tag{3.14}$$

since $\det(g^{-1}) = 1/\det(g)$ and $\det(h) = 1$. We see that $ghg^{-1}$ has determinant 1, and since it is also invertible, we conclude that it is in $\mathrm{SL}(n, \mathbb{F})$. Hence $\mathrm{SL}(n, \mathbb{F})$ is a normal subgroup of $\mathrm{GL}(n, \mathbb{F})$.

### 3.2.3 Quotient Groups

The *quotient group* $G/N$, where $N$ is a normal subgroup, is defined as the set of left cosets:

$$G/N \equiv \{gN \,|\, g \in G\}, \qquad \text{where } gN \equiv \{gh \,|\, h \in N\}. \tag{3.15}$$

Every element $gN$ in $G/N$ is actually the set of all of the possible ways elements from the normal subgroup $N$ can be applied to $g$ from the right. In order to make this set into a group, we must also define a group product. We take the product to be:

$$(g_1 N)(g_2 N) \equiv (g_1 g_2) N, \qquad g_1 N, g_2 N \in G/N. \tag{3.16}$$

Here, the fact that $N$ is normal becomes crucial. If $N$ was not normal, then we could have had a situation where $g_1 N = g_1' N$ and $g_2 N = g_2' N$ for $g_1, g_2, g_1', g_2' \in G$, and yet $(g_1 g_2) N \neq (g_1' g_2') N$, which means that the product does not have a well-defined outcome! However, if $N$ is normal, then by definition we know that $N$ "commutes" with any element of $G$. Thus

$$(g_1 g_2)N = g_1 g_2' N = g_1 N g_2' = g_1' N g_2' = N g_1' g_2' = (g_1' g_2')N, \tag{3.17}$$

and we have a well-defined product.

### 3.2.4 Semidirect Products and Simple Groups

Finally, let $Q$ be a group isomorphic to a quotient group, $Q \cong G/N$. Then we say that $G$ is the *semidirect product* of $Q$ and the normal subgroup $N$:

$$G = Q \ltimes N. \tag{3.18}$$

Any group $G$ possesses two trivial normal subgroups: the group $\{I\}$, consisting of only the identity element, and the group $G$ itself. If $G$ does not possess any non-trivial normal subgroups, we say that it is a *simple group*. If $G$ is not simple, then we can break it into two smaller groups by taking the quotient with a non-trivial normal subgroup. In the example of $\mathrm{GL}(n,\mathbb{F})$ and its (non-trivial) normal subgroup $\mathrm{SL}(n,\mathbb{F})$, we find

$$\mathrm{GL}(n,\mathbb{F}) = \mathbb{F}^{\times} \ltimes \mathrm{SL}(n,\mathbb{F}), \qquad \mathbb{F}^{\times} \cong \mathrm{GL}(n,\mathbb{F})/\mathrm{SL}(n,\mathbb{F}), \tag{3.19}$$

where $\mathbb{F}^{\times} \equiv \mathbb{F}\setminus\{0\}$ is the multiplicative group of the field $\mathbb{F}$.

## 3.3 The General and Special Linear Groups

The *general linear group* $\mathrm{GL}(n,\mathbb{F})$ consists of all invertible $n \times n$ matrices:

$$\mathrm{GL}(n,\mathbb{F}) \equiv \{A \in \mathrm{M}(n,\mathbb{F}) \,|\, \det A \neq 0\}. \tag{3.20}$$

- **Action:** $\mathrm{GL}(n,\mathbb{F})$ is the group of all invertible linear transformations from the vector space $\mathbb{F}^n$ to itself (also known as *automorphisms*).
- **Dimension:** The dimension of $\mathrm{GL}(n,\mathbb{R})$ is $n^2$, since real $n^2$ parameters are needed to specify an $n \times n$ matrix. The requirement that it is invertible does not reduce the number of parameters. To see that, note that removing the matrices with $\det A = 0$ is analogous to removing one point (the origin) from $\mathbb{R}$; you still need 1 parameter to specify where on the real line you are. Similarly, $\mathrm{GL}(n,\mathbb{C})$ has dimension $2n^2$, since its $n^2$ complex entries require $2n^2$ real parameters.
- **Compactness:** $\mathrm{GL}(n,\mathbb{F})$ is non-compact.
- **Connectedness:** $\mathrm{GL}(n,\mathbb{R})$ is disconnected. It consists of two connected components, corresponding to positive and negative determinants respectively. To see this, note that the image of $\det A$ is $\mathbb{R}^{\times} \equiv \mathbb{R}\setminus\{0\}$, which is disconnected, since it can be divided into the two disjoint open sets $(-\infty, 0)$ and $(0, +\infty)$. Since $\det A$ is a continuous function, $\mathrm{GL}(n,\mathbb{R})$ must also be disconnected[7]. In contrast, $\mathrm{GL}(n,\mathbb{C})$ is connected, since the image of $\det A$ is $\mathbb{C}^{\times} \equiv \mathbb{C}\setminus\{0\}$, which is connected (but not simply connected – why?)
- **Lie algebra:** The associated Lie algebra $\mathfrak{gl}(n,\mathbb{F})$ is simply the set $\mathrm{M}(n,\mathbb{F})$ of all $n \times n$ matrices over the field $\mathbb{F}$. Indeed, for any matrix $\mathbf{x} \in \mathrm{M}(n,\mathbb{F})$, the corresponding group element $\mathrm{e}^{\mathbf{x}}$ is automatically invertible, with its inverse being $\mathrm{e}^{-\mathbf{x}}$.

The *special linear group* $\mathrm{SL}(n,\mathbb{F})$ consists of all $n \times n$ matrices with determinant 1:

$$\mathrm{SL}(n,\mathbb{F}) \equiv \{A \in \mathrm{M}(n,\mathbb{F}) \,|\, \det A = 1\}. \tag{3.21}$$

- **Action:** $\mathrm{SL}(n,\mathbb{F})$ is the group of all automorphisms of $\mathbb{F}^n$ that preserve volume and orientation.
- **Dimension:** The dimension of $\mathrm{SL}(n,\mathbb{R})$ is $n^2 - 1$, and the dimension of $\mathrm{SL}(n,\mathbb{C})$ is $2\left(n^2 - 1\right)$.
- **Compactness:** $\mathrm{SL}(n,\mathbb{F})$ is still non-compact.
- **Connectedness:** $\mathrm{SL}(n,\mathbb{F})$ is connected, even for $\mathbb{F} = \mathbb{R}$.
- **Lie algebra:** The associated Lie algebra $\mathfrak{sl}(n,\mathbb{F})$ is the set of all $n \times n$ matrices over $\mathbb{F}$ with vanishing trace, since $\det \mathrm{e}^{\mathbf{x}} = \mathrm{e}^{\operatorname{tr}\mathbf{x}}$.

[7] A continuous function maps connected spaces to connected spaces.

Note that $\mathrm{SL}(n,\mathbb{F})$ is a *normal subgroup* of $\mathrm{GL}(n,\mathbb{F})$, and we have

$$\mathrm{GL}(n,\mathbb{F}) = \mathbb{F}^{\times} \ltimes \mathrm{SL}(n,\mathbb{F}), \tag{3.22}$$

where $\mathbb{F}^{\times} \equiv \mathbb{F}\setminus\{0\}$ is the multiplicative group of $\mathbb{F}$. This means that we can decompose any general linear transformation in $\mathrm{GL}(n,\mathbb{F})$ into a volume- and orientation-preserving transformation in $\mathrm{SL}(n,\mathbb{F})$, times an element of $\mathbb{F}^{\times}$ which may change the volume and/or orientation. Or, more simply, any matrix in $\mathrm{GL}(n,\mathbb{F})$ can be written as its determinant, which is a number in $\mathbb{F}^{\times}$, times a matrix in $\mathrm{SL}(n,\mathbb{F})$.

## 3.4 The Orthogonal and Special Orthogonal Groups

The *orthogonal group* $\mathrm{O}(n)$ consists of all real orthogonal $n \times n$ matrices:

$$\mathrm{O}(n) \equiv \left\{ A \in \mathrm{M}(n,\mathbb{R}) \,\middle|\, A^{\mathrm{T}}A = I \right\}, \tag{3.23}$$

where $A^{\mathrm{T}}$ is the *transpose* of $A$ such that $\left(A^{\mathrm{T}}\right)_{ij} \equiv A_{ji}$.

- **Action:** $\mathrm{O}(n)$ is the group of distance-preserving automorphisms of $\mathbb{R}^n$ (also known as *isometries*) which leave the origin invariant. Indeed, the inner product of two vectors $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ is given by $\mathbf{x}\cdot\mathbf{y} \equiv \mathbf{x}^{\mathrm{T}}\mathbf{y}$, and if $A$ is orthogonal, the inner product is invariant under the orthogonal transformation $\mathbf{x} \mapsto A\mathbf{x}$, $\mathbf{y} \mapsto A\mathbf{y}$. Hence the norm $\|\mathbf{x}\| \equiv \sqrt{\mathbf{x}\cdot\mathbf{x}}$, and in particular the Euclidean distance $\|\mathbf{x}-\mathbf{y}\|$, is invariant under the action of $\mathrm{O}(n)$.
- **Dimension:** The dimension of $\mathrm{O}(n)$ is $n(n-1)/2$, since that is the number of real parameters needed to specify an orthogonal matrix[8].
- **Compactness:** $\mathrm{O}(n)$ is our first example of a compact Lie group. Indeed, first recall that, by the *Heine-Borel theorem*, a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded. We can identify $\mathrm{M}(n,\mathbb{R})$ with $\mathbb{R}^{n^2}$ by taking the $n^2$ entries of the vector in $\mathbb{R}^{n^2}$ to be the matrix entries of a matrix in $\mathrm{M}(n,\mathbb{R})$; for example $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a,b,c,d)$. Then $\mathrm{O}(n)$ may be similarly identified with a subset of $\mathbb{R}^{n^2}$. To see that it is closed, consider the continuous function $f(A) \equiv A^{\mathrm{T}}A$. Then we have $\mathrm{O}(n) = f^{-1}(\{I\})$, that is, $\mathrm{O}(n)$ is the preimage of the closed set $\{I\}$; therefore, since $f$ is continuous, $\mathrm{O}(n)$ must itself be closed[9]. Furthermore, $\mathrm{O}(n)$ is bounded, since each of the rows (or columns) of an orthogonal matrix is a vector of norm 1, and thus no entry in the matrix may be larger than 1. Thus $\mathrm{O}(n)$ is compact.
- **Connectedness:** $\mathrm{O}(n)$ is disconnected, since an orthogonal matrix has $\det\left(A^{\mathrm{T}}A\right) = (\det A)^2 = 1$ and thus $\det A = \pm 1$; the two connected components correspond to the two possible values of the determinant.
- **Lie algebra:** The associated Lie algebra $\mathfrak{o}(n)$ is the set of real anti-symmetric $n \times n$ matrices. Indeed, if $\mathbf{x}$ is anti-symmetric then $\mathbf{x}^{\mathrm{T}} = -\mathbf{x}$ and thus, if $A = \mathrm{e}^{\mathbf{x}}$, we have $A^{\mathrm{T}} = \left(\mathrm{e}^{\mathbf{x}}\right)^{\mathrm{T}} = \mathrm{e}^{-\mathbf{x}} = A^{-1}$, so $A$ is orthogonal.

---

[8] To see this, consider that a general real $n \times n$ matrix $A$ has $n^2$ arbitrary parameters. The matrix $A^{\mathrm{T}}A$ satisfies $\left(A^{\mathrm{T}}A\right)^{\mathrm{T}} = A^{\mathrm{T}}A$, so it is symmetric and thus has $n(n+1)/2$ independent parameters. Hence the equation $A^{\mathrm{T}}A = I$ results in $n(n+1)/2$ equations for the original $n^2$ parameters, reducing the number of independent parameters to $n(n-1)/2$.

[9] A function $f$ from a topological space $X$ to a topological space $Y$ is continuous if and only if the preimage $f^{-1}(S) \equiv \{A \in X | f(A) \in S\}$ of every closed set $S$ in $Y$ is closed in $X$. (Alternatively, if and only if the preimage $f^{-1}(S)$ of every open set $S$ in $Y$ is open in $X$.)

The *special orthogonal group* $\mathrm{SO}(n)$ consists of all real orthogonal $n \times n$ matrices with determinant 1:

$$\mathrm{SO}(n) \equiv \left\{ A \in \mathrm{M}(n, \mathbb{R}) \,\middle|\, A^{\mathrm{T}} A = I \text{ and } \det A = 1 \right\}. \tag{3.24}$$

- **Action:** $\mathrm{SO}(n)$ is the group of isometries of $\mathbb{R}^n$ which leave the origin invariant and preserve orientation. These transformations are also known as *rotations*.
- **Dimension:** The dimension of $\mathrm{SO}(n)$, like that of $\mathrm{O}(n)$, is $n(n-1)/2$. The condition $\det A = 1$ doesn't reduce the number of parameters. To see why, note that $\mathrm{O}(n)$ is a space of dimension $n(n-1)/2$ consisting of two disjoint open sets, one with determinant $+1$ and another with determinant $-1$. Just like dividing the real numbers into positive and negative numbers does not reduce the dimension of the real line, so does selecting only the component with determinant $+1$ not reduce the dimension of the overall space.
- **Compactness:** Like $\mathrm{O}(n)$, $\mathrm{SO}(n)$ is compact.
- **Connectedness:** $\mathrm{SO}(n)$ is connected, and it is in fact one of the two connected components of $\mathrm{O}(n)$: the one with determinant 1.
- **Lie algebra:** The associated Lie algebra $\mathfrak{so}(n)$ is the set of real anti-symmetric $n \times n$ matrices, and it is identical to $\mathfrak{o}(n)$. Note that since an anti-symmetric matrix always has zero trace, its exponential always has determinant 1. Thus, even though both groups have the same Lie algebra, exponentiating that algebra only allows access to the component of $\mathrm{O}(n)$ connected to the identity matrix, namely $\mathrm{SO}(n)$. This is to be expected, since the Lie algebra is the tangent space at the identity, and the identity has determinant 1[10].

$\mathrm{SO}(n)$ is a normal subgroup of $\mathrm{O}(n)$, and we have

$$\mathrm{O}(n) = \mathbb{Z}_2 \ltimes \mathrm{SO}(n), \tag{3.25}$$

where $\mathbb{Z}_2$ is the group $\{+1, -1\}$ with the group product being multiplication. This results from the fact that any matrix in $\mathrm{O}(n)$ can be written as its determinant, which is a number in $\mathbb{Z}_2$, times a matrix in $\mathrm{SO}(n)$.

## 3.5 The Indefinite Orthogonal and Special Orthogonal Groups

### 3.5.1 General $p$ and $q$

Let $p$ and $q$ be positive integers. The *indefinite orthogonal group* $\mathrm{O}(p,q)$ consists of all real $(p+q) \times (p+q)$ matrices which leave a bilinear form (or metric) of signature $(p,q)$ invariant:

$$\mathrm{O}(p,q) \equiv \left\{ A \in \mathrm{M}(p+q, \mathbb{R}) \,\middle|\, A^{\mathrm{T}} \eta A = \eta \right\}, \qquad \eta \equiv \mathrm{diag}(\underbrace{-1, \ldots, -1}_{p}, \underbrace{+1, \ldots, +1}_{q}). \tag{3.26}$$

- **Action:** $\mathrm{O}(p,q)$ is the group of isometries of $\mathbb{R}^{p+q}$ which leave the origin invariant. $\mathbb{R}^{p+q}$ is the $(p+q)$-dimensional real vector space with the dot product $\mathbf{x} \cdot \mathbf{y} \equiv \mathbf{x}^{\mathrm{T}} \eta \mathbf{y}$, and it is invariant under the transformation $\mathbf{x} \mapsto A\mathbf{x}$, $\mathbf{y} \mapsto A\mathbf{y}$. Note that for $p = 0$ and $q = n$, the metric $\eta$ is the $n \times n$ identity matrix, and we recover the condition for $\mathrm{O}(n)$, $A^{\mathrm{T}} A = I$.
- **Dimension:** The dimension of $\mathrm{O}(p,q)$ is $n(n-1)/2$ where $n = p + q$.

[10] Recall that, as we mentioned above, if a group is connected and compact, like $\mathrm{SO}(n)$, then the map from the algebra to the group is surjective (onto), but this is not necessarily true for groups that are disconnected, like $\mathrm{O}(n)$, or non-compact.

- **Compactness:** $\mathrm{O}(p,q)$ is non-compact for $p,q>0$, since the elements of the matrices are not bounded.
- **Connectedness:** $\mathrm{O}(p,q)$ is disconnected and consists of four connected components, described below.
- **Lie algebra:** The associated Lie algebra $\mathfrak{o}(p,q)$ is the set of real $n\times n$ matrices $\mathbf{x}$ satisfying $\mathbf{x}^{\mathrm{T}}=-\eta\mathbf{x}\eta^{-1}$. Indeed, we then have

  $$(\mathrm{e}^{\mathbf{x}})^{\mathrm{T}}\,\eta\,\mathrm{e}^{\mathbf{x}}=\mathrm{e}^{-\eta\mathbf{x}\eta^{-1}}\,\eta\,\mathrm{e}^{\mathbf{x}}=\left(\eta\,\mathrm{e}^{-\mathbf{x}}\,\eta^{-1}\right)\eta\,\mathrm{e}^{\mathbf{x}}=\eta, \tag{3.27}$$

  as required. Note that for $p=0$ and $q=n$, this condition reduces to the anti-symmetry condition defining $\mathfrak{o}(n)$. As for $\mathfrak{o}(n)$, the Lie algebra only allows us access to the component connected to the identity.

The *indefinite special orthogonal group* $\mathrm{SO}(p,q)$ consists of all of the matrices in $\mathrm{O}(p,q)$ which have determinant 1:

$$\mathrm{SO}(p,q)\equiv\left\{A\in\mathrm{M}(p+q,\mathbb{R})\,\middle|\,A^{\mathrm{T}}\eta A=\eta\text{ and }\det A=1\right\}, \tag{3.28}$$

where $\eta$ is defined as before.

- **Action:** It is the group of isometries of $\mathbb{R}^{p+q}$ which leave the origin invariant and preserve orientation.
- **Dimension:** $\mathrm{SO}(p,q)$ still has dimension $n(n-1)/2$ where $n=p+q$.
- **Compactness:** Like $\mathrm{O}(p,q)$, the group $\mathrm{SO}(p,q)$ is non-compact for $p,q>0$.
- **Connectedness:** $\mathrm{SO}(p,q)$ is still disconnected for $p,q>0$, but with only two connected components, the ones where $\det A=1$.
- **Lie algebra:** The associated Lie algebra $\mathfrak{so}(p,q)$ is the set of real $n\times n$ matrices $\mathbf{x}$ satisfying $\mathbf{x}^{\mathrm{T}}=-\eta\mathbf{x}\eta^{-1}$. It is identical to $\mathfrak{o}(p,q)$, and only allows us access to the component of $\mathrm{SO}(p,q)$ connected to the identity.

### 3.5.2 $p=1$: The Lorentz Group

The special case $p=1$ is the most important one in physics, since it defines the *Lorentz group*, with $\eta$ the *Minkowski metric*. The group $\mathrm{O}(1,n-1)$ is the Lorentz group in $n$ dimensions. To describe the four connected components of this group, let us consider the simplest non-trivial case, that of $\mathrm{O}(1,1)$, the Lorentz group in 2 dimensions. The Minkowski metric is

$$\eta\equiv\begin{pmatrix}-1&0\\0&1\end{pmatrix}, \tag{3.29}$$

and we take a general matrix $A\in\mathrm{O}(1,1)$,

$$A\equiv\begin{pmatrix}a&b\\c&d\end{pmatrix}. \tag{3.30}$$

From the defining relation $A^{\mathrm{T}}\eta A=\eta$, we get

$$\begin{pmatrix}a&c\\b&d\end{pmatrix}\begin{pmatrix}-1&0\\0&1\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}c^2-a^2&cd-ab\\cd-ab&d^2-b^2\end{pmatrix}=\begin{pmatrix}-1&0\\0&1\end{pmatrix}. \tag{3.31}$$

Thus

$$ab = cd, \qquad c^2 - a^2 = -1, \qquad d^2 - b^2 = 1. \tag{3.32}$$

These are 3 equations for 4 unknowns, so we must have one free parameter. Indeed, this is consistent with the fact that the dimension of the group is 1. Let us choose the free parameter to be $a$. Then we immediately get $c^2 = a^2 - 1$ and thus we must have $a^2 \geq 1$. Furthermore, $b = cd/a$, and plugging that into $1 = d^2 - b^2$ we obtain

$$1 = d^2 - \frac{c^2 d^2}{a^2} = d^2 - \frac{\left(a^2 - 1\right) d^2}{a^2} = \frac{d^2}{a^2}, \tag{3.33}$$

so $d = \pm a$. Then from $b = cd/a$ we get $b = \pm c$, where the sign corresponds to the same sign as that chosen for $d = \pm a$. In conclusion, we have two sets of solutions, one with positive determinant:

$$b = c = \pm\sqrt{a^2 - 1}, \qquad d = +a, \qquad \det A = ad - bc = +1, \tag{3.34}$$

and one with negative determinant:

$$b = \pm\sqrt{a^2 - 1}, \qquad c = \mp\sqrt{a^2 - 1}, \qquad d = -a, \qquad \det A = ad - bc = -1. \tag{3.35}$$

In addition, the condition $a^2 \geq 1$ is satisfied by either $a \geq +1$ or $a \leq -1$. We can now classify the solutions into four types, depending on the sign of $a$ and the sign of the determinant:

- *Proper*: $\det A = +1$,
- *Improper*: $\det A = -1$,
- *Orthochronous*: $a \geq +1$,
- *Non-orthochronous*: $a \leq -1$.

The four connected components of $\mathrm{O}(1,1)$ are then:

- Proper, orthochronous transformations: $\det A = +1$ and $a \geq +1$. This is the component connected to the identity matrix $I = \mathrm{diag}(1,1)$, which is obtained for $a = +1$.
- Improper, orthochronous transformations: $\det A = -1$ and $a \geq +1$. This component includes the *parity inversion* matrix $P \equiv \mathrm{diag}(1,-1)$, which is obtained for $a = +1$.
- Improper, non-orthochronous transformations: $\det A = -1$ and $a \leq -1$. This component includes the *time reversal* matrix $T \equiv \mathrm{diag}(-1,1)$, which is obtained for $a = -1$.
- Proper, non-orthochronous transformations: $\det A = +1$ and $a \leq -1$. This component includes the matrix $PT = \mathrm{diag}(-1,-1) = -I$, which is obtained for $a = -1$.

Now, by definition $\mathrm{SO}(1,1)$ contains the two proper components, both orthochronous and non-orthochronous. The subgroup of proper orthochronous transformations is denoted $\mathrm{SO}^+(1,1)$. It can be shown that this is a normal subgroup of $\mathrm{O}(1,1)$.

Furthermore, note that we can move between the four connected components by applying $P$, $T$, or both. One can check that the elements $\{I, P, T, PT\}$ form a group; in fact, this group is isomorphic to the quotient group $\mathrm{O}(1,1)/\mathrm{SO}^+(1,1)$. Thus we have the semidirect product

$$\mathrm{O}(1,1) = \{I, P, T, PT\} \ltimes \mathrm{SO}^+(1,1). \tag{3.36}$$

This analysis readily generalizes to the 4-dimensional Lorentz group $\mathrm{O}(1,3)$.

### 3.6 The Unitary and Special Unitary Groups

The *unitary group* $\mathrm{U}(n)$ consists of all complex unitary $n \times n$ matrices:

$$\mathrm{U}(n) \equiv \left\{ A \in \mathrm{M}(n, \mathbb{C}) \,\middle|\, A^\dagger A = I \right\}, \tag{3.37}$$

where $A^\dagger$ is the *conjugate transpose* of $A$ such that $\left(A^\dagger\right)_{ij} \equiv \bar{A}_{ji}$.

- **Action:** $\mathrm{U}(n)$ is the group of isometries of $\mathbb{C}^n$, with respect to the Hermitian inner product given by $\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{x}^\dagger \mathbf{y}$ where $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, which leave the origin invariant. It is the analogue of $\mathrm{O}(n)$ for complex spaces.
- **Dimension:** The dimension of $\mathrm{U}(n)$ is $n^2$. To see this, note that a general complex matrix requires $2n^2$ real parameters. The matrix $A^\dagger A$ is Hermitian, since $\left(A^\dagger A\right)^\dagger = A^\dagger A$. The diagonal entries of a Hermitian matrix must be real, so they correspond to $n$ real parameters, and there are $\sum_{k=1}^{n-1} k = n(n-1)/2$ entries above the diagonal corresponding to $n(n-1)$ real parameters. Thus the Hermitian matrix $A^\dagger A$ has $n^2$ real parameters, and the equation $A^\dagger A = I$ results in $n^2$ equations for the original $2n^2$ parameters, reducing the number of independent parameters to $n^2$.
- **Compactness:** $\mathrm{U}(n)$ is compact; the proof is similar to the one we had above for $\mathrm{O}(n)$.
- **Connectedness:** Unlike $\mathrm{O}(n)$, the group $\mathrm{U}(n)$ is connected. We still have $\det\left(A^\dagger A\right) = |\det A|^2 = 1$, but the determinant is now complex and thus can take any value on the unit circle.
- **Lie algebra:** The associated Lie algebra $\mathfrak{u}(n)$ is the set of complex anti-Hermitian matrices, that is, matrices $\mathbf{x}$ such that $\mathbf{x}^\dagger = -\mathbf{x}$. Indeed, we then get $\left(\mathrm{e}^{\mathbf{x}}\right)^\dagger \mathrm{e}^{\mathbf{x}} = \mathrm{e}^{-\mathbf{x}} \mathrm{e}^{\mathbf{x}} = I$, as required. Furthermore, the trace of an anti-Hermitian matrix is imaginary, so the determinant $\det \mathrm{e}^{\mathbf{x}} = \mathrm{e}^{\operatorname{tr} \mathbf{x}}$ is on the unit circle, and we see that the map from the Lie algebra to the group is surjective, unlike the case of $\mathfrak{o}(n)$; this is due to $\mathrm{U}(n)$ being connected and compact.

The *special unitary group* $\mathrm{SU}(n)$ consists of all complex unitary $n \times n$ matrices with determinant 1:

$$\mathrm{SU}(n) \equiv \left\{ A \in \mathrm{M}(n, \mathbb{C}) \,\middle|\, A^\dagger A = I \text{ and } \det A = 1 \right\}. \tag{3.38}$$

- **Action:** $\mathrm{SU}(n)$ is the group of isometries of $\mathbb{C}^n$ which leave the origin invariant but do not rotate the components of the transformed vectors by an additional overall complex phase $\mathrm{e}^{\mathrm{i}\phi}$, which would be given by the determinant.
- **Dimension:** $\mathrm{SU}(n)$ has dimension $n^2 - 1$. Here, unlike the case of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, the condition $\det A = 1$ actually reduces the number of parameters by 1. This is because for $\mathrm{U}(n)$ the determinant can be anywhere on the unit circle, i.e. $\det A = \mathrm{e}^{\mathrm{i}\phi}$, so the phase $\phi$ is one of the $n^2$ real parameters determining an element of $\mathrm{U}(n)$. Hence, demanding $\det A = 1$ removes one parameter.
- **Compactness:** $\mathrm{SU}(n)$ is compact.
- **Connectedness:** $\mathrm{SU}(n)$ is connected.
- **Lie algebra:** The associated Lie algebra $\mathfrak{su}(n)$ is the set of complex anti-Hermitian matrices with vanishing trace. The vanishing trace guarantees that the determinant of the exponentiated group element is 1.

SU $(n)$ is a normal subgroup of U $(n)$, and we have

$$\mathrm{U}(n) = \mathrm{U}(1) \ltimes \mathrm{SU}(n), \tag{3.39}$$

where U $(1)$ is the unit circle on the complex plane, corresponding to all the possible determinants of matrices in U $(n)$.

One can also define the *indefinite unitary group* U $(p,q)$ and the *indefinite special unitary group* SU $(p,q)$ where $p+q=n$, similar to how we defined O $(p,q)$ and SO $(p,q)$. We will not go into the details here.

## 3.7 Summary

The following table summarizes the matrix Lie groups we have discussed in this chapter:

| Name | Notation | Dimension | Compact? | Connected? | Algebra Matrices |
|---|---|---|---|---|---|
| General Linear ($\mathbb{C}$) | $\mathrm{GL}(n,\mathbb{C})$ | $2n^2$ | No | Yes | Complex |
| General Linear ($\mathbb{R}$) | $\mathrm{GL}(n,\mathbb{R})$ | $n^2$ | No | 2 Comp. | Real |
| Special Linear ($\mathbb{C}$) | $\mathrm{SL}(n,\mathbb{C})$ | $2(n^2-1)$ | No | Yes | Complex Traceless |
| Special Linear ($\mathbb{R}$) | $\mathrm{SL}(n,\mathbb{R})$ | $n^2-1$ | No | Yes | Real Traceless |
| Orthogonal | $\mathrm{O}(n)$ | $n(n-1)/2$ | Yes | 2 Comp. | Real Anti-Symmetric |
| Special Orthogonal | $\mathrm{SO}(n)$ | $n(n-1)/2$ | Yes | Yes | Real Anti-Symmetric |
| Indefinite Orthogonal | $\mathrm{O}(p,q)$ | $n(n-1)/2$ | No | 4 Comp. | Real with $\mathbf{x}^{\mathrm{T}}=-\eta\mathbf{x}\eta^{-1}$ |
| Indefinite Special Orthogonal | $\mathrm{SO}(p,q)$ | $n(n-1)/2$ | No | 2 Comp. | Real with $\mathbf{x}^{\mathrm{T}}=-\eta\mathbf{x}\eta^{-1}$ |
| Unitary | $\mathrm{U}(n)$ | $n^2$ | Yes | Yes | Anti-Hermitian |
| Special Unitary | $\mathrm{SU}(n)$ | $n^2-1$ | Yes | Yes | Anti-Hermitian Traceless |

# 4 Representation Theory of Lie Groups and Algebras

## 4.1 Basic Definitions and Concepts

### 4.1.1 Representations

An $n$-dimensional *representation* of a Lie group $G$ is a map[11]

$$R : G \to \mathrm{GL}(n,\mathbb{F}), \tag{4.1}$$

such that the group product on $G$ maps to the group product on GL $(n,\mathbb{F})$:

$$R(gh) = R(g)\,R(h), \qquad \forall g,h \in G. \tag{4.2}$$

This type of map is called a *Lie group homomorphism*. From the definition, one can prove that the identity element of $G$ maps to the identity element of GL $(n,\mathbb{F})$:

$$R(I_G) = I_{\mathrm{GL}(n,\mathbb{F})}, \tag{4.3}$$

and that an inverse element in $G$ maps to the corresponding inverse element in GL $(n,\mathbb{F})$:

$$R(g^{-1}) = R(g)^{-1}. \tag{4.4}$$

[11] As usual, we will take $\mathbb{F}$ to be either $\mathbb{R}$, for a real representation, or $\mathbb{C}$, for a complex representation.

Similarly, an $n$-dimensional representation of a Lie algebra $\mathfrak{g}$ is a map

$$R : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{F}), \tag{4.5}$$

such that the Lie brackets on $\mathfrak{g}$ map to the Lie brackets on $\mathfrak{gl}(n, \mathbb{F})$:

$$R([\mathbf{x}, \mathbf{y}]) = [R(\mathbf{x}), R(\mathbf{y})]. \tag{4.6}$$

This type of map is called a *Lie algebra homomorphism*.

If the homomorphism is injective (one-to-one), that is, each matrix in $\mathrm{GL}(n, \mathbb{F})$ or $\mathfrak{gl}(n, \mathbb{F})$ is mapped to by at most one element of $G$ or $\mathfrak{g}$ respectively, the representation is called *faithful*.

It is important to note that the dimension of the representation is **not** the same as the dimension of the Lie group. The dimension of the representation is equal to the dimension $n$ of the vector space $\mathbb{F}^n$ it acts on, while the dimension of the Lie group is the number of real parameters needed to specify an element of the group.

A subspace $W$ of $\mathbb{F}^n$ is called *invariant* if the action of the group or algebra on a vector in $W$ stays in $W$. That is,

$$R(g)\,\mathbf{v} \in W, \qquad \forall g \in G, \qquad \forall \mathbf{v} \in W, \tag{4.7}$$

or

$$R(\mathbf{x})\,\mathbf{v} \in W, \qquad \forall \mathbf{x} \in \mathfrak{g}, \qquad \forall \mathbf{v} \in W. \tag{4.8}$$

The trivial invariant subspaces are $\{\mathbf{0}\}$, where $\mathbf{0} \in \mathbb{F}^n$ is the zero vector, and $\mathbb{F}^n$ itself. A representation with no non-trivial invariant subspaces is called *irreducible*. Sometimes you will see an irreducible representation referred to as an "*irrep*" for short. For compact groups, every representation is equivalent to a direct sum of irreducible representations.

Representations allow us to define a *linear action* of a Lie group or algebra on a vector space $\mathbb{F}^n$. The action of an element $g \in G$ or $\mathbf{x} \in \mathfrak{g}$ on a vector $\mathbf{v} \in \mathbb{F}^n$ is given by acting on it with the representing matrix:

$$g \triangleright \mathbf{v} = R(g)\,\mathbf{v}, \qquad \mathbf{x} \triangleright v = R(\mathbf{x})\,\mathbf{v}. \tag{4.9}$$

### 4.1.2 Algebra Representations from Group Representations

If a Lie group $G$ has Lie algebra $\mathfrak{g}$, and $R$ is a representation of $G$ acting on $\mathbb{F}^n$, then there exists a unique representation of $\mathfrak{g}$ on $\mathbb{F}^n$ such that

$$R(\mathrm{e}^{\mathbf{x}}) = \mathrm{e}^{R(\mathbf{x})}, \qquad \forall \mathbf{x} \in \mathfrak{g}. \tag{4.10}$$

In other words, if a group element $g$ is expressed as the exponential of an algebra element $\mathbf{x}$, then the representation of $g$ is expressed as the exponential of the representation of $\mathbf{x}$. In fact, the representation is given by

$$R(\mathbf{x}) \equiv \frac{\mathrm{d}}{\mathrm{d}t} R\left(\mathrm{e}^{t\mathbf{x}}\right)\bigg|_{t=0}. \tag{4.11}$$

Note that we are using the same letter $R$ for the representation of the Lie group and the Lie algebra. This makes sense given that, for example,

$$R\left(g\mathbf{x}g^{-1}\right) = R(g)\,R(\mathbf{x})\,R(g)^{-1}, \qquad \forall g \in G, \qquad \forall \mathbf{x} \in \mathfrak{g}. \tag{4.12}$$

In other words, $R$ preserves not only the product of group elements but also their action on algebra elements (this is called the *adjoint action*, and we will discuss it in Subsection 4.2.3). If $G$ is simply connected, every representation of $\mathfrak{g}$ may be derived from a representation of $G$.

### 4.1.3 Unitary Representations and Quantum Mechanics

A *unitary representation* is one which maps every element $g \in G$ to the unitary group $\mathrm{U}(n)$, instead of $\mathrm{GL}(n, \mathbb{C})$. More generally, we can consider a *Hilbert space* $\mathcal{H}$, a complete[12] complex vector space with an *inner product* $\langle \cdot, \cdot \rangle$. The group elements are then represented by unitary operators acting on the Hilbert space, which, by definition, preserve the inner product:

$$\langle R(g)\,\mathbf{v}, R(g)\,\mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle, \qquad \forall g \in G, \qquad \forall \mathbf{v}, \mathbf{w} \in \mathcal{H}. \tag{4.13}$$

In quantum mechanics, whenever a physical system can be acted on with a group $G$, the Hilbert space $\mathcal{H}$ will carry a unitary representation of $G$. The vectors in $\mathcal{H}$ are the quantum states, and the group action preserves the inner product of states, and thus probability amplitudes.

Now, given a unitary representation $R$, since $R(g)$ is in $\mathrm{U}(n)$ for every $g \in G$, we may write

$$R(g) = \mathrm{e}^{\mathbf{x}}, \tag{4.14}$$

where $\mathbf{x}$ is an anti-Hermitian matrix, $\mathbf{x}^\dagger = -\mathbf{x}$, in the Lie algebra $\mathfrak{u}(n)$. Thus the matrix $\mathrm{i}\,\mathbf{x}$ is *Hermitian*:

$$(\mathrm{i}\,\mathbf{x})^\dagger = \mathrm{i}\,\mathbf{x}. \tag{4.15}$$

Hermitian (or *self-adjoint*) operators in quantum mechanics correspond to *observables*. Thus, the unitary action of a Lie group on a Hilbert space provides us with observables for free! In fact, most observables of interest, such as energy, momentum and angular momentum, arise from the action of some Lie group.

For example, let us consider *translation in time*, given by the group $\mathbb{R}$, with the group operation being addition. A unitary representation of this group on a Hilbert space $\mathcal{H}$ is given by

$$U : \mathbb{R} \to \mathrm{U}(n), \qquad U(t) \equiv \mathrm{e}^{-\mathrm{i}\,Ht}, \tag{4.16}$$

where $H$ is called the *Hamiltonian*. Since $H$ is a Hermitian operator, $U(t)$ is a unitary operator. Furthermore, it is easy to see that the map $U$ is a group homomorphism:

$$U(t_1 + t_2) = \mathrm{e}^{-\mathrm{i}\,H(t_1+t_2)} = U(t_1)\,U(t_2). \tag{4.17}$$

Note that we have mapped the additive group $\mathbb{R}$ to the multiplicative group $\mathrm{U}(n)$, so addition in $\mathbb{R}$ is mapped to multiplication in $\mathrm{U}(n)$. The action of $U(\Delta t)$ on a state $\boldsymbol{\psi}(t) \in \mathcal{H}$ simply translates the state in time by a duration $\Delta t$:

$$U(\Delta t)\,\boldsymbol{\psi}(t) = \boldsymbol{\psi}(t + \Delta t). \tag{4.18}$$

Then, given $U(t) \equiv \mathrm{e}^{-\mathrm{i}\,Ht}$, we can calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\psi}(t) = \lim_{\Delta t \to 0} \frac{\boldsymbol{\psi}(t + \Delta t) - \boldsymbol{\psi}(t)}{\Delta t} = \left( \lim_{\Delta t \to 0} \frac{\mathrm{e}^{-\mathrm{i}\,H\Delta t} - 1}{\Delta t} \right) \boldsymbol{\psi}(t) = -\mathrm{i}\,H\boldsymbol{\psi}(t), \tag{4.19}$$

which is none other than the Schrödinger equation.

Notice that translation in time is not necessarily a symmetry of the theory. The study of group representations is not limited to symmetry groups, although symmetries are a very important application.

[12] A *Cauchy sequence* is a sequence of vectors $\mathbf{v}_k \in \mathcal{H}$ such that, for every positive real number $\varepsilon > 0$, almost all (i.e. all but a finite number) of the vectors in the sequence satisfy $\|\mathbf{v}_i - \mathbf{v}_j\| < \varepsilon$, where $\|\mathbf{v}\| \equiv \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$. A space $\mathcal{H}$ is called *complete* if every Cauchy sequence in $\mathcal{H}$ converges to a vector in $\mathcal{H}$.

The real numbers are complete, but the rational numbers are not complete, since one can always find a Cauchy sequence of rational numbers which converges to an irrational number; for example, the sequence $\{3, 3.1, 3.14, 3.141, \ldots\}$ converges to $\pi$.

### 4.1.4 Direct Sums and Tensor Products

Given two representations $R_1$ and $R_2$, of dimensions $n_1$ and $n_2$ respectively, the *direct sum* $R_1 \oplus R_2$ is a representation of dimension $n_1 + n_2$ given by the homomorphism

$$R_1 \oplus R_2 : G \to \mathrm{GL}(n_1 + n_2, \mathbb{F}), \qquad (R_1 \oplus R_2)(g) = \begin{pmatrix} R_1(g) & 0 \\ 0 & R_2(g) \end{pmatrix}, \tag{4.20}$$

where the matrix $(R_1 \oplus R_2)(g)$ is an $(n_1 + n_2) \times (n_1 + n_2)$ block-diagonal matrix with the blocks given by the $n_1 \times n_1$ matrix $R_1(g)$ and the $n_2 \times n_2$ matrix $R_2(g)$. Importantly, any unitary representation $R$ can be written as a direct sum of irreducible representations,

$$R = R_1 \oplus \ldots \oplus R_m, \qquad m \in \mathbb{N}. \tag{4.21}$$

Next, let $U$ and $V$ be finite-dimensional vector spaces. Then the *tensor product* $U \otimes V$ consists of linear combinations of elements $\mathbf{u} \otimes \mathbf{v}$ with $\mathbf{u} \in U$ and $\mathbf{v} \in V$. If $\mathbf{u}_i$ and $\mathbf{v}_i$ are basis vectors for $U$ and $V$ respectively, then $\mathbf{u}_i \otimes \mathbf{v}_j$ for all $i, j$ will be a basis for $U \otimes V$. Furthermore, let $G$ and $H$ be Lie groups. The *direct product* $G \times H$ is a new group, whose elements are all possible ordered pairs $(g, h)$ with $g \in G$ and $h \in H$, and the group product is $(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2)$.

Given a $p$-dimensional representation $R_1$ of the group $G$ on the space $U$ and a $q$-dimensional representation $R_2$ of the group $H$ on the space $V$, we may define a $pq$-dimensional representation $R_1 \otimes R_2$ of $G \times H$ by

$$(R_1 \otimes R_2)(g, h) \equiv R_1(g) \otimes R_2(h), \qquad g \in G, \qquad h \in H. \tag{4.22}$$

Let $\mathfrak{g}$ be the Lie algebra of $G$ and let $\mathfrak{h}$ be the Lie algebra of $H$. The Lie algebra of $G \times H$ is isomorphic to $\mathfrak{g} \otimes \mathfrak{h}$, and we have

$$(R_1 \otimes R_2)(\mathbf{x}, \mathbf{y}) = R_1(\mathbf{x}) \otimes I + I \otimes R_2(\mathbf{y}), \qquad \mathbf{x} \in \mathfrak{g}, \qquad \mathbf{y} \in \mathfrak{h}. \tag{4.23}$$

## 4.2 The Trivial, Fundamental and Adjoint Representations

### 4.2.1 The Trivial Representation

The *trivial representation* of a group $G$ is a map $R : G \to \mathrm{GL}(n, \mathbb{F})$ which sends all elements in $G$ to the identity matrix in $\mathrm{GL}(n, \mathbb{F})$. Similarly, the trivial representation of an algebra $\mathfrak{g}$ is a map $R : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{F})$ which sends all elements in $\mathfrak{g}$ to the zero matrix in $\mathfrak{gl}(n, \mathbb{F})$. We use the trivial representation when we want a particular quantity to be invariant under the action of the group.

### 4.2.2 The Fundamental Representation

It may not come as a shock to the reader that, since the matrix Lie groups are themselves subgroups of $\mathrm{GL}(n, \mathbb{F})$, they may **represent themselves**, with the "representation" being the identity map $R(g) \equiv g$. Similarly, the matrix Lie algebras are subsets of $\mathfrak{gl}(n, \mathbb{F})$ and may thus represent themselves. In physics, this representation is called the *fundamental representation*. In math, it is called the *defining* or *standard representation*.

For example, the fundamental representation of the group $\mathrm{SO}(n)$ on $\mathbb{R}^n$ is simply the one where the rotation matrices in $\mathrm{SO}(n)$ do exactly what they're supposed to: rotate vectors in $\mathbb{R}^n$.

### 4.2.3 The Adjoint Representation

Let $G$ be a Lie group with Lie algebra $\mathfrak{g}$. For each element $g \in G$, we define the *adjoint map* $\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$:

$$\mathrm{Ad}_g \, \mathbf{x} \equiv g \mathbf{x} g^{-1}, \qquad \forall \mathbf{x} \in \mathfrak{g}. \tag{4.24}$$

From this definition we easily obtain that, for all $\mathbf{x}, \mathbf{y} \in \mathfrak{g}$,

$$\mathrm{Ad}_g\left([\mathbf{x}, \mathbf{y}]\right) = \left[\mathrm{Ad}_g \mathbf{x}, \mathrm{Ad}_g \mathbf{y}\right]. \tag{4.25}$$

Thus, $\mathrm{Ad}_g$ preserves the Lie bracket.

Now, recall that a Lie algebra is nothing but a vector space equipped with a Lie bracket. If the dimension of this vector space is $n$, meaning that each algebra element may be represented by $n$ real parameters, then given a suitable choice of basis, we may represent linear transformations on $\mathfrak{g}$ using matrices in $\mathrm{GL}(n, \mathbb{R})$.

We are now ready to define the *adjoint representation* of $G$: it is the map $\mathrm{Ad} : G \to \mathrm{GL}(n, \mathbb{R})$ given by $g \mapsto \mathrm{Ad}_g$ and acting on the $n$-dimensional real vector space $\mathfrak{g}$. It satisfies

$$\mathrm{Ad}_{gh} \mathbf{x} = gh\mathbf{x}h^{-1}g^{-1} = \mathrm{Ad}_g \mathrm{Ad}_h \mathbf{x}, \qquad \forall g, h \in G, \qquad \forall \mathbf{x} \in \mathfrak{g}, \tag{4.26}$$

and thus it is a group homomorphism, as required for a representation.

Moreover, we can show that there is an associated map $\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(n, \mathbb{R})$ given by $\mathbf{x} \mapsto \mathrm{ad}_{\mathbf{x}}$, such that

$$\mathrm{e}^{\mathrm{ad}_{\mathbf{x}}} = \mathrm{Ad}_{(\mathrm{e}^{\mathbf{x}})}. \tag{4.27}$$

We may derive the action of this map by considering a group element $\mathrm{e}^{t\mathbf{x}}$ and using (4.11):

$$\mathrm{ad}_{\mathbf{x}} \mathbf{y} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{Ad}_{(\mathrm{e}^{t\mathbf{x}})} \mathbf{y}\right)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{t\mathbf{x}} \mathbf{y}\, \mathrm{e}^{-t\mathbf{x}}\right)\bigg|_{t=0} = [\mathbf{x}, \mathbf{y}]. \tag{4.28}$$

Thus, we see that the Lie algebra associated with a Lie group inherits its Lie bracket from the group product through the action of the adjoint map.

The map ad provides the adjoint representation of the Lie algebra $\mathfrak{g}$, since it satisfies

$$\begin{aligned}
\mathrm{ad}_{[\mathbf{x},\mathbf{y}]} \mathbf{z} &= [[\mathbf{x}, \mathbf{y}], \mathbf{z}] \\
(*) &= [\mathbf{x}, [\mathbf{y}, \mathbf{z}]] - [\mathbf{y}, [\mathbf{x}, \mathbf{z}]] \\
&= \left(\mathrm{ad}_{\mathbf{x}} [\mathbf{y}, \mathbf{z}] - \mathrm{ad}_{\mathbf{y}} [\mathbf{x}, \mathbf{z}]\right) \\
&= \mathrm{ad}_{\mathbf{x}} \mathrm{ad}_{\mathbf{y}} \mathbf{z} - \mathrm{ad}_{\mathbf{y}} \mathrm{ad}_{\mathbf{x}} \mathbf{z} \\
&= \left[\mathrm{ad}_{\mathbf{x}}, \mathrm{ad}_{\mathbf{y}}\right] \mathbf{z},
\end{aligned}$$

where in $(*)$ we used the Jacobi identity. Hence, it is a Lie algebra homomorphism, as required.

### 4.2.4 Constructing the Adjoint Representation from the Structure Constants

If the structure constants ${f_{ij}}^k$ are known, then the adjoint representation can be explicitly constructed from them by defining the matrix elements of the representatives of the generators as:

$$\left(\mathbf{T}_i\right)^k{}_j \equiv {f_{ij}}^k. \tag{4.29}$$

By this, we mean that the element at row number $k$ and column number $j$ of the matrix representing the generator $\mathbf{T}_i$ is given by the number ${f_{ij}}^k$.

Recall that the structure constants ${f_{ij}}^k$ are defined by

$$\left[\mathbf{T}_i, \mathbf{T}_j\right] \equiv {f_{ij}}^k \mathbf{T}_k, \tag{4.30}$$

where $\mathbf{T}_i$ are the generators (with $i$ going from 1 to the dimension of the group), and they satisfy anti-symmetry in the first two indices:

$${f_{ij}}^k = -{f_{ji}}^k, \tag{4.31}$$

and the Jacobi identity:

$$f_{ij}{}^{l} f_{kl}{}^{m} + f_{jk}{}^{l} f_{il}{}^{m} + f_{ki}{}^{l} f_{jl}{}^{m} = 0. \tag{4.32}$$

We would like to show that the representatives of the generators $\mathbf{T}_i$ defined in this way satisfy the appropriate commutation relations. Let us write the commutator explicitly (which we can do since the algebra is a matrix Lie algebra):

$$\left[\mathbf{T}_i, \mathbf{T}_j\right] = \mathbf{T}_i \mathbf{T}_j - \mathbf{T}_j \mathbf{T}_i. \tag{4.33}$$

First, note that $\mathbf{T}_i \mathbf{T}_j$ is also a matrix, with elements

$$\left(\mathbf{T}_i \mathbf{T}_j\right)^m{}_k = \left(\mathbf{T}_i\right)^m_l \left(\mathbf{T}_j\right)^l{}_k = f_{il}{}^{m} f_{jk}{}^{l} = f_{jk}{}^{l} f_{il}{}^{m}. \tag{4.34}$$

Similarly,

$$\left(\mathbf{T}_j \mathbf{T}_i\right)^m{}_k = f_{ik}{}^{l} f_{jl}{}^{m} = -f_{ki}{}^{l} f_{jl}{}^{m}, \tag{4.35}$$

and thus we get

$$\left[\mathbf{T}_i, \mathbf{T}_j\right]^m{}_k = \left(\mathbf{T}_i \mathbf{T}_j - \mathbf{T}_j \mathbf{T}_i\right)^m{}_k = f_{jk}{}^{l} f_{il}{}^{m} + f_{ki}{}^{l} f_{jl}{}^{m}. \tag{4.36}$$

Using the Jacobi identity, we now see that

$$\left[\mathbf{T}_i, \mathbf{T}_j\right]^m{}_k = -f_{ij}{}^{l} f_{kl}{}^{m} = f_{ij}{}^{l} f_{lk}{}^{m} = f_{ij}{}^{l} \left(\mathbf{T}_l\right)^m_k, \tag{4.37}$$

and therefore $\left[\mathbf{T}_i, \mathbf{T}_j\right] = f_{ij}{}^{l} \mathbf{T}_l$, as desired.

The generators are the basis vectors for the Lie algebra, which is also the vector space on which the adjoint representation acts. Let us define an inner product on this vector space using the representation matrices:

$$\mathbf{T}_i \cdot \mathbf{T}_j \equiv \operatorname{Tr}\left(\mathbf{T}_i \mathbf{T}_j\right) \equiv \left(\mathbf{T}_i\right)^k_l \left(\mathbf{T}_j\right)^l{}_k, \tag{4.38}$$

with summation on both $k$ and $l$ implied. Note that the trace Tr is with respect to the matrix indices $k, l$, and **not** the indices $i, j$, which simply enumerate the generators. Now, for a compact Lie algebra, we can always find an orthonormal basis such that

$$\mathbf{T}_i \cdot \mathbf{T}_j = \delta_{ij}. \tag{4.39}$$

In this basis, we have the *triple product*

$$\left[\mathbf{T}_i, \mathbf{T}_j\right] \cdot \mathbf{T}_k = f_{ij}{}^{l} \mathbf{T}_l \cdot \mathbf{T}_k = f_{ij}{}^{l} \delta_{lk} = f_{ijk}. \tag{4.40}$$

This is a generalization of the triple product $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$ in $\mathbb{R}^3$. On the other hand, using the cyclic property of the trace, we find:

$$\begin{aligned}
\left[\mathbf{T}_i, \mathbf{T}_j\right] \cdot \mathbf{T}_k &= \left(\mathbf{T}_i \mathbf{T}_j - \mathbf{T}_j \mathbf{T}_i\right) \cdot \mathbf{T}_k \\
&= \operatorname{Tr}\left(\mathbf{T}_i \mathbf{T}_j \mathbf{T}_k - \mathbf{T}_j \mathbf{T}_i \mathbf{T}_k\right) \\
&= \operatorname{Tr}\left(\mathbf{T}_j \mathbf{T}_k \mathbf{T}_i - \mathbf{T}_k \mathbf{T}_j \mathbf{T}_i\right) \\
&= \left[\mathbf{T}_j, \mathbf{T}_k\right] \cdot \mathbf{T}_i,
\end{aligned}$$

and therefore

$$f_{ijk} = f_{jki}. \tag{4.41}$$

Furthermore, since the structure constants are anti-symmetric in the first two indices, we also have $f_{ijk} = -f_{jik}$ and thus

$$-f_{jik} = f_{jki}, \tag{4.42}$$

so in this basis the structure constants are also anti-symmetric in the last two indices, and therefore *completely anti-symmetric*, like the Levi-Civita symbol $\epsilon_{ijk}$.

Finally, let us show that this construction indeed gives us the adjoint action, $\mathrm{ad}_{\mathbf{x}}\,\mathbf{y} = [\mathbf{x},\mathbf{y}]$. Let us write the three vectors $\mathbf{x}$, $\mathbf{y}$ and $\mathrm{ad}_{\mathbf{x}}\,\mathbf{y}$ explicitly using their components with respect to the generators:

$$\mathbf{x} \equiv x^i \mathbf{T}_i, \qquad \mathbf{y} \equiv y^i \mathbf{T}_i, \qquad \mathrm{ad}_{\mathbf{x}}\,\mathbf{y} \equiv \left(\mathrm{ad}_{\mathbf{x}}\,\mathbf{y}\right)^i \mathbf{T}_i. \tag{4.43}$$

The adjoint action of $\mathbf{x}$ on $\mathbf{y}$ is given by taking the elements of the **matrix** representing $\mathbf{x}$ and acting with it, as a linear transformation, on the components of the **vector** $\mathbf{y}$:

$$\left(\mathrm{ad}_{\mathbf{x}}\,\mathbf{y}\right)^i = \left(\mathbf{x}\right)^i{}_j\, y^j = x^k \left(\mathbf{T}_k\right)^i{}_j\, y^j = x^k f_{kj}{}^i y^j. \tag{4.44}$$

Hence

$$\mathrm{ad}_{\mathbf{x}}\,\mathbf{y} = \left(\mathrm{ad}_{\mathbf{x}}\,\mathbf{y}\right)^i \mathbf{T}_i = x^k \left(f_{kj}{}^i \mathbf{T}_i\right) y^j = x^k \left[\mathbf{T}_k, \mathbf{T}_j\right] y^j = \left[\mathbf{x},\mathbf{y}\right], \tag{4.45}$$

as required.

# 5 Representations of Some Matrix Lie Groups

After defining representations of Lie groups and algebras, we will now construct some representations of common matrix Lie groups encountered in physics.

## 5.1 Representations of $\mathrm{U}(1)$ and $\mathrm{SO}(2)$

The Lie group $\mathrm{U}(1)$ consists of the $1\times 1$ unitary matrices. Such a matrix $A \in \mathrm{U}(1)$ satisfies $A^\dagger A = I$. Of course, these matrices are isomorphic to the complex numbers $z$ with unit norm, that is, $z \in \mathbb{C}$ with $|z|^2 \equiv z^* z = 1$. On the complex plane, these correspond to the unit circle:

$$z \equiv a + \mathrm{i}\, b, \qquad |z|^2 = a^2 + b^2 = 1, \tag{5.1}$$

and may be parametrized by a real number $\theta \in \mathbb{R}$ such that

$$z \equiv \mathrm{e}^{\mathrm{i}\,\theta}. \tag{5.2}$$

The fundamental representation of $\mathrm{U}(1)$ is, of course, itself, and it acts on the complex numbers $\mathbb{C}$ by rotating them around the origin in the complex plane. The parameter $\theta$ corresponds to the angle of rotation. Obviously, $\theta$ and $\theta + 2\pi k$ for $k \in \mathbb{Z}$ correspond to the same rotation and thus the same group element, which follows automatically from the fact that $\mathrm{e}^{2\pi\,\mathrm{i}} = 1$.

### 5.1.1 Schur's Lemma

Let $V$ and $W$ be vector spaces over $\mathbb{C}$ and let $R_V$ and $R_W$ be irreducible representations of a Lie group $G$ on these vector spaces. Then *Schur's Lemma* states that

1. If $V$ and $W$ are not isomorphic, then there are no non-trivial maps $\phi$ from $V$ to $W$ such that, for every $g \in G$, $R_W(g) \circ \phi = \phi \circ R_V(g)$.
2. If $V = W$ and $R_V = R_W$, then the only non-trivial maps $\phi$ as described above are the identity and scalar multiples of it.

From this lemma we can show that, since $\mathrm{U}(1)$ is Abelian, all of its complex irreducible representations are 1-dimensional. Indeed, by Schur's lemma, if a matrix $A$ commutes with all of the matrices of a complex irreducible representation $R$ of a group $G$, that is, $AR(g) = R(g)A$ for all $g \in G$, then

$A$ must be a scalar multiple of the identity matrix: $A = \lambda I$ for some $\lambda \in \mathbb{C}$. Now, if $G$ is Abelian, then we have

$$R(g)\,R(h) = R(h)\,R(g)\,, \qquad \forall g, h \in G. \tag{5.3}$$

In other words, the matrix $R(h)$ commutes with $R(g)$ for all $g \in G$. Thus, if $R$ is irreducible, we must have $R(h) = \lambda(h)\, I$ with a particular $\lambda(h) \in \mathbb{C}$ for every $h \in G$. This then means that $R$ must be 1-dimensional, and given by $R(h) = \lambda(h) \in \mathbb{C}$.

### 5.1.2 Irreducible Representations of U (1)

Given our result above, all irreducible representations of $\mathrm{U}(1)$ will be maps

$$R : \mathrm{U}(1) \to \mathrm{GL}(1, \mathbb{C})\,, \tag{5.4}$$

where $\mathrm{GL}(1, \mathbb{C})$ is the group of invertible $1 \times 1$ matrices, which is isomorphic to $\mathbb{C}^\times \equiv \mathbb{C} \setminus \{0\}$. However, we can in fact show that the irreducible representations are all unitary, and given by

$$R(\mathrm{e}^{\mathrm{i}\,\theta}) = \mathrm{e}^{\mathrm{i}\,k\theta} \in \mathrm{U}(1)\,, \qquad k \in \mathbb{Z}. \tag{5.5}$$

Therefore, the only irreducible representations of $\mathrm{U}(1)$ are homomorphisms from the group to itself! Indeed, let $R : \mathrm{U}(1) \to \mathrm{GL}(1, \mathbb{C})$ be a representation of $\mathrm{U}(1)$, and let us write

$$\rho(\theta) \equiv R(\mathrm{e}^{\mathrm{i}\,\theta}). \tag{5.6}$$

Then $\rho$ is periodic and a homomorphism:

$$\rho(\theta + 2\pi) = \rho(\theta)\,, \qquad \rho(\theta_1 + \theta_2) = \rho(\theta_1)\,\rho(\theta_2)\,. \tag{5.7}$$

Hence

$$\begin{aligned} \rho'(\theta) &= \lim_{\Delta\theta \to 0} \frac{\rho(\theta + \Delta\theta) - \rho(\theta)}{\Delta\theta} \\ &= \rho(\theta) \lim_{\Delta\theta \to 0} \frac{\rho(\Delta\theta) - 1}{\Delta\theta} \\ &= \rho(\theta)\,\rho'(0)\,. \end{aligned}$$

This differential equation is easily solved, with the initial condition $\rho(0) = 1$, by

$$\rho(\theta) = \mathrm{e}^{\rho'(0)\theta}\,. \tag{5.8}$$

Furthermore, from periodicity we find that

$$1 = \rho(2\pi) = \mathrm{e}^{2\pi\rho'(0)} \quad \Longrightarrow \quad \rho'(0) = \mathrm{i}\,k, \qquad k \in \mathbb{Z}. \tag{5.9}$$

Hence

$$\rho(\theta) = \mathrm{e}^{\mathrm{i}\,k\theta}, \tag{5.10}$$

as we wanted to show. In conclusion, the irreducible representations of $\mathrm{U}(1)$ are all of the form $R(\mathrm{e}^{\mathrm{i}\,\theta}) = \mathrm{e}^{\mathrm{i}\,k\theta}$ for some integer $k$.

### 5.1.3 Reducible Representations of $\mathrm{U}(1)$

The unitary reducible representations may be written as a direct sum of irreducible ones. Let $R$ be a unitary representation on a Hilbert space $\mathcal{H}$ of dimension $n$. We can split $\mathcal{H}$ into

$$\mathcal{H} = \mathcal{H}_{q_1} \oplus \cdots \oplus \mathcal{H}_{q_n}, \qquad q_1, \ldots, q_n \in \mathbb{Z}, \tag{5.11}$$

where each $\mathcal{H}_{q_i}$ is 1-dimensional. Choosing a basis and defining a Hermitian *charge operator* $Q$:

$$Q \equiv \operatorname{diag}(q_1, \ldots, q_n), \tag{5.12}$$

we obtain the general $n$-dimensional representation of $\mathrm{U}(1)$:

$$R(e^{i\theta}) = e^{iQ\theta} = \operatorname{diag}(e^{iq_1\theta}, \ldots, e^{iq_n\theta}). \tag{5.13}$$

The operator $e^{iQ\theta}$ simply adds a phase $q_i$ to each component of a vector $\mathbf{v} \in \mathcal{H}$. The charge $Q$, since it is Hermitian, is an observable. Furthermore, it corresponds to a conserved quantity if it commutes with the Hamiltonian:

$$[H, Q] = 0. \tag{5.14}$$

The time evolution operator defined above, $U(t) \equiv e^{-iHt}$, then also commutes with $Q$:

$$[U(t), Q] = 0. \tag{5.15}$$

This means that the value of the observable $Q$ will be constant in time, since, if $Q\mathbf{v}(0) = q\mathbf{v}(0)$ where $q \in \mathbb{R}$ is an eigenvalue of $Q$, we have at a later time $t$:

$$Q\mathbf{v}(t) = QU(t)\mathbf{v}(0) = U(t)Q\mathbf{v}(0) = U(t)q\mathbf{v}(0) = q\mathbf{v}(t). \tag{5.16}$$

The operator $e^{iQ\theta}$, in this case, corresponds to a symmetry transformation, and the $q_i$ will be conserved quantities.

### 5.1.4 Isomorphism with $\mathrm{SO}(2)$

Finally, we show that $\mathrm{U}(1) \cong \mathrm{SO}(2)$, that is, an element of $\mathrm{U}(1)$ may also represent a rotation on $\mathbb{R}^2$. This comes from Euler's formula:

$$e^{i\theta} = \cos\theta + i\sin\theta. \tag{5.17}$$

If we act on a complex number $z \equiv a + ib$ with a rotation $e^{i\theta} \in \mathrm{U}(1)$, we get:

$$e^{i\theta} z = (\cos\theta + i\sin\theta)(a + ib) = (a\cos\theta - b\sin\theta) + i(a\sin\theta + b\cos\theta). \tag{5.18}$$

Let us now represent $z$ as a vector in $\mathbb{R}^2$:

$$z \equiv a + ib \cong \begin{pmatrix} a \\ b \end{pmatrix}. \tag{5.19}$$

Then it is easy to see that $e^{i\theta}$ may be represented as a $2 \times 2$ real matrix:

$$e^{i\theta} \cong \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{5.20}$$

such that

$$e^{i\theta} z \cong \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a\cos\theta - b\sin\theta \\ a\sin\theta + b\cos\theta \end{pmatrix}, \tag{5.21}$$

which is indeed the representation of (5.18) as a vector in $\mathbb{R}^2$. This is an irreducible real representation of $\mathrm{U}(1)$, with real dimension 2 (which corresponds to complex dimension 1), acting on $\mathbb{R}^2$ – and it is also an element of $\mathrm{SO}(2)$. In fact, all of the irreducible real representations of $\mathrm{U}(1)$ on $\mathbb{R}^2$ are of the form

$$R(e^{i\theta}) = \begin{pmatrix} \cos(k\theta) & -\sin(k\theta) \\ \sin(k\theta) & \cos(k\theta) \end{pmatrix}, \qquad k \in \mathbb{Z}. \tag{5.22}$$

Since $\mathrm{U}(1) \cong \mathrm{SO}(2)$, their representations coincide.

### 5.1.5 The Lie Algebras $\mathfrak{u}(1)$ and $\mathfrak{so}(2)$

The Lie algebra $\mathfrak{u}(1)$ consists of $1 \times 1$ anti-Hermitian matrices, which are isomorphic to the imaginary numbers $\mathrm{i}\,\mathbb{R}$. It is 1-dimensional, and a suitable generator is simply the number i. Indeed, each element $g \in \mathrm{U}(1)$ can be written as $\mathrm{e}^{\mathrm{i}\theta}$, with $\theta$ being a real parameter and i the generator. (This is, trivially, a "linear combination" of the generators.)

The Lie algebra $\mathfrak{so}(2)$ consists of $2 \times 2$ real anti-symmetric matrices. It is also 1-dimensional, and a suitable generator is

$$\boldsymbol{\tau} \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{5.23}$$

Indeed, one can show that

$$\mathrm{e}^{\theta\boldsymbol{\tau}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{5.24}$$

It is easy to see that these Lie algebras are isomorphic. The isomorphism as vector spaces is trivial, since any two finite-dimensional vector spaces of the same dimension are isomorphic. The Lie brackets are also trivially the same, since the only possible bracket is between the one generator and itself, and the structure constants (2.5) are just ${f_{11}}^1 = 0$ for both.

## 5.2 Representations of $\mathrm{SU}(2)$ and $\mathrm{SO}(3)$

### 5.2.1 The Generators of $\mathfrak{so}(3)$

The group $\mathrm{SO}(3)$ consists of $3 \times 3$ orthogonal matrices with determinant 1, which rotate vectors in $\mathbb{R}^3$. The Lie algebra $\mathfrak{so}(3)$, as we have seen above, is the vector space of $3 \times 3$ real anti-symmetric matrices.

Let us take the following basis of generators for $\mathfrak{so}(3)$:

$$\mathbf{L}_1 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathbf{L}_2 \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{L}_3 \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.25}$$

A general $3 \times 3$ anti-symmetric matrix $\mathbf{x} \in \mathfrak{so}(3)$ can be constructed as a linear combination of these generators:

$$\mathbf{x} = \begin{pmatrix} 0 & -x^3 & x^2 \\ x^3 & 0 & -x^1 \\ -x^2 & x^1 & 0 \end{pmatrix} = x^i \mathbf{L}_i, \tag{5.26}$$

where a sum over $i \in \{1, 2, 3\}$ is implied. Furthermore, the generators satisfy the commutation relations

$$[\mathbf{L}_i, \mathbf{L}_j] = {\epsilon_{ij}}^k \mathbf{L}_k, \tag{5.27}$$

where ${\epsilon_{ij}}^k$ is the totally anti-symmetric Levi-Civita symbol with ${\epsilon_{12}}^3 \equiv +1$. From the definition (2.5), we see that the structure constants of the Lie algebra $\mathfrak{so}(3)$ are

$${f_{ij}}^k = {\epsilon_{ij}}^k. \tag{5.28}$$

### 5.2.2 The Generators of $\mathfrak{su}(2)$

Consider now $\mathrm{SU}(2)$. It consists of the $2 \times 2$ unitary matrices with determinant 1, which rotate vectors in $\mathbb{C}^2$. The Lie algebra $\mathfrak{su}(2)$ is the vector space of $2 \times 2$ anti-Hermitian traceless matrices. To define a basis of generators for $\mathfrak{su}(2)$, let us first define the *Pauli matrices*:

$$\sigma_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 \equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.29}$$

These matrices are clearly traceless, but they are Hermitian and not anti-Hermitian. To obtain a basis for anti-Hermitian traceless matrices, we multiply the Pauli matrices by $-\mathrm{i}/2$:

$$\boldsymbol{\tau}_i \equiv -\frac{\mathrm{i}}{2}\sigma_i, \tag{5.30}$$

or explicitly:

$$\boldsymbol{\tau}_1 = \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}, \qquad \boldsymbol{\tau}_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{\tau}_3 = \frac{1}{2}\begin{pmatrix} -\mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix}. \tag{5.31}$$

A general $2 \times 2$ traceless anti-Hermitian matrix $\mathbf{x} \in \mathfrak{su}(2)$ can be constructed as a linear combinations of these generators:

$$\mathbf{x} = \frac{1}{2}\begin{pmatrix} -\mathrm{i}\, x^3 & -x^2 - \mathrm{i}\, x^1 \\ x^2 - \mathrm{i}\, x^1 & \mathrm{i}\, x^3 \end{pmatrix} = x^i \boldsymbol{\tau}_i, \tag{5.32}$$

where $x^i$ are real parameters. Due to our choice of normalization, these generators satisfy the commutation relations

$$[\boldsymbol{\tau}_i, \boldsymbol{\tau}_j] = \epsilon_{ij}{}^k \boldsymbol{\tau}_k, \tag{5.33}$$

and thus the structure constants of $\mathfrak{su}(2)$ are the same as those of $\mathfrak{so}(3)$!

Both $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are 3-dimensional vector spaces (since elements in both are determined by 3 real parameters) equipped with Lie brackets that have the same structure constants, and hence, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic Lie algebras.

### 5.2.3 $\mathbb{R}^3$ with the Cross Product

In fact, these Lie algebras are isomorphic to another familiar algebra: the vector space $\mathbb{R}^3$ with the usual cross product. Indeed, the generators of $\mathbb{R}^3$ can be given, for example, by

$$\mathbf{e}_1 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{e}_3 \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{5.34}$$

The cross product equips this vector space with a binary operation $[\mathbf{x}, \mathbf{y}] \equiv \mathbf{x} \times \mathbf{y}$, which is also anti-commutative and satisfies the Jacobi identity. Moreover, the structure constants are

$$[\mathbf{e}_i, \mathbf{e}_j] \equiv \mathbf{e}_i \times \mathbf{e}_j = \epsilon_{ij}{}^k \mathbf{e}_k, \tag{5.35}$$

same as for $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$. Thus the three Lie algebras are isomorphic.

### 5.2.4 A Map from SU (2) to SO (3)

We have found that the algebras $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are isomorphic. However, it turns out that the groups SU (2) and SO (3) are not isomorphic! Let us show that.

We can write the most general matrix $g \in \mathrm{SU}(2)$ as follows:

$$g(x, y) \equiv \begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix}, \qquad |x|^2 + |y|^2 = 1, \qquad x, y \in \mathbb{C}. \tag{5.36}$$

With some work, we can show that there is a representation $R$ of SU (2) such that

$$R(g(x, y)) = \begin{pmatrix} \mathrm{Re}\left(x^2 - y^2\right) & \mathrm{Im}\left(x^2 + y^2\right) & -2\,\mathrm{Re}(xy) \\ -\mathrm{Im}\left(x^2 - y^2\right) & \mathrm{Re}\left(x^2 + y^2\right) & 2\,\mathrm{Im}(xy) \\ 2\,\mathrm{Re}(x\bar{y}) & 2\,\mathrm{Im}(x\bar{y}) & |x|^2 - |y|^2 \end{pmatrix} \in \mathrm{SO}(3). \tag{5.37}$$

This map is surjective (onto), since for each rotation in SO (3) we can find at least one pair $x, y \in \mathbb{C}$ which maps to it. Let us assume that we have mapped some element $g(x, y) \in \mathrm{SU}(2)$ to the identity matrix in SO (3). Then in particular, by comparing (5.37) with the identity matrix, we see that we must have

$$xy = 0, \tag{5.38}$$

so either $x = 0$ or $y = 0$. However, we must also have $|x|^2 - |y|^2 = 1$, so we cannot take $x = 0$, and thus necessarily $y = 0$, so we are left with

$$|x|^2 = 1. \tag{5.39}$$

Furthermore, let us denote

$$x \equiv a + \mathrm{i}\, b \quad \Longrightarrow \quad x^2 = a^2 - b^2 + 2\,\mathrm{i}\, ab, \qquad |x|^2 = a^2 + b^2. \tag{5.40}$$

Then we must have

$$\mathrm{Re}\left(x^2 - y^2\right) = \mathrm{Re}\, x^2 = a^2 - b^2 = 1, \tag{5.41}$$

but on the other hand

$$|x|^2 - |y|^2 = |x|^2 = a^2 + b^2 = 1, \tag{5.42}$$

so we conclude that $b^2 = 0$ and $a^2 = 1$. In summary, for $g(x, y) \in \mathrm{SU}(2)$ to map to $I \in \mathrm{SO}(3)$, we must have

$$x = \pm 1, \qquad y = 0. \tag{5.43}$$

Therefore, there are **two** elements in SU (2) which map to the identity of SO (3):

$$g(\pm 1, 0) = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}. \tag{5.44}$$

These two elements form the *kernel* of $R$, that is, the subset of elements which map to the identity. They are isomorphic to the group $\mathbb{Z}_2 \equiv \{+1, -1\}$. Furthermore, they form a normal subgroup of SU (2). Thus we can take the quotient (see Subsection 3.2):

$$\mathrm{SO}(3) \cong \mathrm{SU}(2) / \mathbb{Z}_2, \qquad \mathrm{SU}(2) = \mathrm{SO}(3) \ltimes \mathbb{Z}_2. \tag{5.45}$$

In other words, SU (2) **contains two copies of** SO (3). We sometimes say that SU (2) is a *double cover* of SO (3).

### 5.2.5 Irreducible Representations of $\mathfrak{su}(2)$

Let us first change conventions and write the $\mathfrak{su}(2)$ generators (5.31) as they are commonly written in physics:

$$\mathbf{J}_i \equiv \mathrm{i}\, \boldsymbol{\tau}_i = \frac{1}{2}\sigma_i. \tag{5.46}$$

Since we already know how the Pauli matrices look like, we immediately have the matrices in the fundamental representation:

$$\mathbf{J}_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{J}_2 = \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \mathbf{J}_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.47}$$

However, let us forget these matrices and treat the algebra as an abstract entity. All we know about it are the Lie brackets:

$$\left[\mathbf{J}_i, \mathbf{J}_j\right] = \mathrm{i}\, \epsilon_{ij}{}^{k} \mathbf{J}_k, \tag{5.48}$$

which tell us that the structure constants are $\epsilon_{ij}{}^{k}$ (in the physics notation). We will now detail the steps for finding the irreducible representations of $\mathfrak{su}(2)$ using these structure constants. This prescription can be generalized to other Lie algebras.

**Step 1:** Find the maximal set of mutually commuting generators. These generators form the *Cartan subalgebra*, often denoted $\mathfrak{h}$, and the number of generators in $\mathfrak{h}$ is called the *rank* of the Lie algebra.

In the case of $\mathfrak{su}(2)$, the generators $\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3$ satisfy $\left[\mathbf{J}_i, \mathbf{J}_j\right] = \mathrm{i}\,\epsilon_{ij}{}^{k}\mathbf{J}_k$, so no two of them commute. Conventionally, $\mathbf{J}_3$ is chosen (since, in the fundamental representation, it is the only diagonal matrix). Thus the rank of $\mathfrak{su}(2)$ is 1.

**Step 2:** For the generators which are **not** in the Cartan subalgebra $\mathfrak{h}$, find linear combinations such that they are "eigengenerators" of the generators in $\mathfrak{h}$, in the sense that, for each $\boldsymbol{\tau} \in \mathfrak{h}$ and $\boldsymbol{\sigma} \in \mathfrak{g}$, there is a scalar $\lambda \in \mathbb{C}$ such that

$$[\boldsymbol{\tau}, \boldsymbol{\sigma}] = \lambda \boldsymbol{\sigma}. \tag{5.49}$$

For $\mathfrak{su}(2)$, we will choose the linear combinations

$$\mathbf{J}_\pm \equiv \frac{1}{\sqrt{2}}\left(\mathbf{J}_1 \pm \mathrm{i}\,\mathbf{J}_2\right) \quad \Longrightarrow \quad \mathbf{J}_1 = \frac{1}{\sqrt{2}}\left(\mathbf{J}_+ + \mathbf{J}_-\right), \qquad \mathbf{J}_2 = \frac{1}{\mathrm{i}\sqrt{2}}\left(\mathbf{J}_+ - \mathbf{J}_-\right), \tag{5.50}$$

such that $\mathbf{J}_+ = \mathbf{J}_-^\dagger$; note that they are no longer Hermitian. We then get that

$$[\mathbf{J}_3, \mathbf{J}_\pm] = \pm \mathbf{J}_\pm. \tag{5.51}$$

The eigenvalues $\pm 1$ are called the *weights* of the eigenvectors. Furthermore,

$$[\mathbf{J}_+, \mathbf{J}_-] = \mathbf{J}_3. \tag{5.52}$$

**Step 3:** Build the vector space upon which the representation matrices acts. To do that, consider the **representation matrices** of the generators, and the vectors they act on. In particular, consider eigenvectors of the Cartan generators.

For $\mathfrak{su}(2)$, it is instructive to use bra-ket notation, since as we will see, the representations of $\mathfrak{su}(2)$ are directly related to spin in quantum mechanics. Instead of eigenvectors of the Cartan generator, we will talk about *eigenstates*. Let $|m\rangle$ be the eigenstate of the **representation matrix** of the Cartan generator $\mathbf{J}_3$ (**not** $\mathbf{J}_3$ itself!) with eigenvalue $m$:

$$\mathbf{J}_3 \left|m\right\rangle = m \left|m\right\rangle. \tag{5.53}$$

Now, we have

$$\mathbf{J}_3 \left(\mathbf{J}_\pm \left|m\right\rangle\right) = \left(\mathbf{J}_\pm \mathbf{J}_3 + [\mathbf{J}_3, \mathbf{J}_\pm]\right)\left|m\right\rangle = (m \pm 1)\,\mathbf{J}_\pm \left|m\right\rangle, \tag{5.54}$$

and therefore $\mathbf{J}_\pm |m\rangle$ is an eigenstate with eigenvalue $m \pm 1$, which means it's proportional to $|m \pm 1\rangle$ up to normalization:

$$\mathbf{J}_\pm \left|m\right\rangle \propto \left|m \pm 1\right\rangle. \tag{5.55}$$

In other words, $\mathbf{J}_\pm$ act as raising and lowering operators, just as for a harmonic oscillator. The proportionality constant will be determined by demanding that the states are orthonormal:

$$\langle m | m' \rangle = \delta_{mm'}. \tag{5.56}$$

We define

$$\mathbf{J}_- \left|m\right\rangle \equiv N_m \left|m - 1\right\rangle. \tag{5.57}$$

Then

$$\langle m - 1 | \mathbf{J}_- | m \rangle = \langle m - 1 | N_m | m - 1 \rangle = N_m, \tag{5.58}$$

and, taking the Hermitian conjugate, using the fact that $\mathbf{J}_+ = \mathbf{J}_-^\dagger$ and assuming that $N_m$ is real,

$$\langle m|\mathbf{J}_+|m-1\rangle = N_m. \tag{5.59}$$

From this we learn that

$$\mathbf{J}_+|m-1\rangle = N_m\,|m\rangle \quad\Longrightarrow\quad \mathbf{J}_+|m\rangle = N_{m+1}\,|m+1\rangle\,. \tag{5.60}$$

Therefore, on the one hand

$$\langle m|\,[\mathbf{J}_+,\mathbf{J}_-]\,|m\rangle = \langle m|\mathbf{J}_3|m\rangle = m, \tag{5.61}$$

and on the other hand

$$\begin{aligned}
\langle m|\,[\mathbf{J}_+,\mathbf{J}_-]\,|m\rangle &= \langle m|\,(\mathbf{J}_+\mathbf{J}_- - \mathbf{J}_-\mathbf{J}_+)\,|m\rangle \\
&= \langle m|\mathbf{J}_+N_m\,|m-1\rangle - \langle m|\mathbf{J}_-N_{m+1}\,|m+1\rangle \\
&= N_m\langle m|N_m\,|m\rangle - N_{m+1}\langle m|N_{m+1}\,|m\rangle \\
&= N_m^2 - N_{m+1}^2,
\end{aligned}$$

so

$$m = N_m^2 - N_{m+1}^2. \tag{5.62}$$

To solve this recursion relation, we write it out explicitly step by step, starting from $j$ and going down to some $m < j$:

$$\begin{aligned}
j &= N_j^2 - N_{j+1}^2 \\
j-1 &= N_{j-1}^2 - N_j^2 \\
j-2 &= N_{j-2}^2 - N_{j-1}^2 \\
\vdots\ &= \quad\vdots \\
m+2 &= N_{m+2}^2 - N_{m+3}^2 \\
m+1 &= N_{m+1}^2 - N_{m+2}^2 \\
m &= N_m^2 - N_{m+1}^2.
\end{aligned}$$

Adding all of the lines together, we see that the terms on the right-hand side cancel each other out except for the highest one, $N_{j+1}^2$, and the lowest one, $N_m^2$, and we are left with:

$$\begin{aligned}
N_m^2 - N_{j+1}^2 &= j + (j-1) + (j-2) + \cdots + (m+2) + (m+1) + m \\
&= \sum_{k=m}^{j} k \\
&= \frac{1}{2}\,(j+m)\,(j-m+1)\,.
\end{aligned}$$

It remains only to choose an initial condition. Let us assume that there is a *highest-weight state*, that is, a state $|j\rangle$ which is annihilated by the raising operator:

$$\mathbf{J}_+|j\rangle = N_{j+1}\,|j+1\rangle = 0 \quad\Longrightarrow\quad N_{j+1} = 0. \tag{5.63}$$

Then we finally obtain the solution

$$N_m = \frac{1}{\sqrt{2}}\sqrt{(j+m)\,(j-m+1)}. \tag{5.64}$$

Of course, this vanishes for $m = j+1$ by our initial condition. However, we also notice that it vanishes for $m = -j$. Therefore, the existence of a highest weight $+j$ automatically implies the existence of a lowest weight $-j$ such that

$$\mathbf{J}_- \left|-j\right\rangle = 0. \tag{5.65}$$

Now, note that, whatever real number we chose for $j$, we must be able to go down in unit steps from $j$ to $-j$. This means that $2j$ must be an integer; we can't have $j = 1.1$, for example, since there is no way to get from 1.1 to $-1.1$ with unit steps! Therefore, the allowed values for the highest weights are the half integers:

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots. \tag{5.66}$$

### 5.2.6 Representations of SU(2) and Spin

We call $j$ the *spin* of the representation. A representation with spin $j$ has $2j+1$ states, denoted $\left|j, m\right\rangle$, where $j$ is the highest weight and $m$ is the eigenvalue of $\mathbf{J}_3$, which takes values from $-j$ to $+j$ in discrete unit steps. Here are some examples of representation spaces:

$$j = 0 \quad \Longrightarrow \quad \left|0, 0\right\rangle, \tag{5.67}$$

$$j = \frac{1}{2} \quad \Longrightarrow \quad \left|\frac{1}{2}, -\frac{1}{2}\right\rangle, \qquad \left|\frac{1}{2}, +\frac{1}{2}\right\rangle, \tag{5.68}$$

$$j = 1 \quad \Longrightarrow \quad \left|1, -1\right\rangle, \qquad \left|1, 0\right\rangle, \qquad \left|1, +1\right\rangle. \tag{5.69}$$

Given a spin $j$, there is a unique representation $R : \mathfrak{su}(2) \to \mathfrak{gl}(2j+1, \mathbb{C})$, assigning to each element in $\mathfrak{su}(2)$ a $(2j+1) \times (2j+1)$ complex matrix. The matrix elements of the representative of $\mathbf{J}_3$ in each representation will be given by

$$(\mathbf{J}_3)_{m'm} = \langle m'|\mathbf{J}_3|m\rangle = m\delta_{m'm}, \tag{5.70}$$

so it is always diagonal. The representatives of $\mathbf{J}_\pm$ will have the matrix elements

$$(\mathbf{J}_+)_{m'm} = \langle m'|\mathbf{J}_+|m\rangle = N_{m+1}\delta_{m',m+1} = \frac{1}{\sqrt{2}}\sqrt{(j+m+1)(j-m)}\,\delta_{m',m+1}, \tag{5.71}$$

$$(\mathbf{J}_-)_{m'm} = \langle m'|\mathbf{J}_-|m\rangle = N_m\delta_{m',m-1} = \frac{1}{\sqrt{2}}\sqrt{(j+m)(j-m+1)}\,\delta_{m',m-1}. \tag{5.72}$$

Now we can easily write down the basis of states, and the irreducible representations of $\mathfrak{su}(2)$ on that basis, for each value of $j$. These representations also correspond to representations of SU(2), because SU(2) is simply connected. For other groups, the algebra representations might not correspond to group representations. Let us discuss this subtlety next.

### 5.2.7 Exponentiating the Algebra Representation: A Subtlety

Let $G$ be a Lie group and $\mathfrak{g}$ its associated Lie algebra. Recall that, given a representation of the group, we can use (4.11) to obtain a representation of the corresponding Lie algebra. In addition, if $G$ is simply connected, we can exponentiate an algebra representation to get a group representation. However, this may not be true if the Lie group is not simply connected.

The group SU(2) is simply connected, and thus representations of $\mathfrak{su}(2)$ may be exponentiated to find representations of SU(2). Hence, the irreducible representations of SU(2) are characterized, as above, by a spin $j$, and act on a vector space with dimension $2j+1$.

However, the group SO(3), while connected, is not simply connected! To see this, note that a rotation in 3D is characterized by the angle of rotation, $\theta \in [-\pi, \pi]$, and a unit vector indicating the axis around which to rotate. Thus SO (3) can be viewed as a *solid ball* in $\mathbb{R}^3$, with radius $\pi$. A point inside the ball which is at the direction of the unit vector $\hat{\mathbf{r}}$ and has magnitude $\theta \in [0, \pi]$ will correspond to a rotation by the angle $\theta$ around the axis determined by $\hat{\mathbf{r}}$, with the origin of the ball corresponding to the identity (no rotation). Furthermore, a rotation by a negative angle $-\theta \in [-\pi, 0]$ is equivalent to a rotation with positive angle $\theta \in [0, \pi]$, but with an opposite direction vector $-\hat{\mathbf{r}}$. Thus, every possible rotation is encoded in the ball.

However, we now run into a problem: a rotation with $\theta = -\pi$ is the same as a rotation with $\theta = \pi$, and therefore we must identify two antipodal points, which we will call the north and south poles. With this identification, the ball is now homeomorphic to SO (3). Now we can see why SO (3) is not simply connected: a path from the north pole to the south pole will be a closed loop, since the poles are identified, and this loop cannot be shrunk to a point!

Because SO (3) is not simply connected, not all representations of its Lie algebra $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ correspond to representations of SO (3). In fact, it turns out that only representations with **integer** $j$ are representations of SO (3).

Now we will list the first three representations of SU (2).

### 5.2.8 The Case $j = 0$: The Trivial Representation

For $j = 0$ we have a 1-dimensional representation (a *singlet*) acting on one state,

$$|0, 0\rangle \equiv 1, \tag{5.73}$$

and the representation matrices will be $1 \times 1$ matrices, which all evaluate to zero:

$$\mathbf{J}_+ = \mathbf{J}_- = \mathbf{J}_3 = 0. \tag{5.74}$$

This is the trivial representation, since it maps every element of $\mathfrak{su}(2)$ to the zero matrix (i.e. the number 0) and, by exponentiation, every element of SU (2) to the identity matrix (i.e. the number 1). This representation is also called the *scalar* representation, since the states in this representation are invariant under the action of SU (2). It is also the trivial representation of SO (3).

### 5.2.9 The Case $j = 1/2$: The Fundamental Representation

For $j = 1/2$ we will have a 2-dimensional representation (a *doublet*). Let us define the basis states as:

$$\left|\frac{1}{2}, +\frac{1}{2}\right\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{5.75}$$

Then one can calculate the following representation matrices:

$$\mathbf{J}_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{J}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{J}_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5.76}$$

This representation is the fundamental representation of SU (2). It is also called the *spinor* representation. Note that it does **not** correspond to a representation of SO (3), since it has non-integer spin.

### 5.2.10 The Case $j = 1$: The Adjoint Representation

For $j = 1$ we will have a 3-dimensional representation (a *triplet*). Let us define the basis states as:

$$|1, +1\rangle \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |1, 0\rangle \equiv \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |1, -1\rangle \equiv \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{5.77}$$

Then a tedious but straightforward calculation yields the $3 \times 3$ matrices:

$$\mathbf{J}_+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{J}_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathbf{J}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{5.78}$$

This is the adjoint representation of $\mathrm{SU}(2)$. Indeed, it can be shown (using a similarity transformation $\mathbf{J}_i \mapsto \mathbf{S}^{-1}\mathbf{J}_i\mathbf{S}$ with an appropriate matrix $\mathbf{S}$) that this representation is isomorphic to the one defined using the structure constants in (4.29). It corresponds to the fundamental (**not** adjoint!) representation of $\mathrm{SO}(3)$, and therefore it describes rotation of vectors in $\mathbb{R}^3$, which is why it is also called the *vector* representation.

#### 5.2.11 The Casimir Operator

For our final task, let us define the *Casimir operator*:

$$\mathbf{J}^2 \equiv \mathbf{J}_1^2 + \mathbf{J}_2^2 + \mathbf{J}_3^2 = \mathbf{J}_3^2 + \mathbf{J}_+\mathbf{J}_- + \mathbf{J}_-\mathbf{J}_+. \tag{5.79}$$

It is straightforward to check that it commutes with all of the generators:

$$\left[\mathbf{J}^2, \mathbf{J}_i\right] = 0, \qquad \forall i \in \{1, 2, 3\}. \tag{5.80}$$

Now, by Schur's lemma (see Subsection 5.1.1), since $\mathbf{J}^2$ is a map from $\mathfrak{su}(2)$ to itself which commutes with every representation matrix, it must be a scalar multiple of the identity[13]. This means that every state in the representation space is an eigenstate of $\mathbf{J}^2$ with the same eigenvalue. This eigenvalue turns out to be $j(j+1)$. For example, for $j = 0$ we get 0, for $j = 1/2$ we get $3/4$, and for $j = 1$ we get 2. Of course, this is familiar from the definition of spin in quantum mechanics.

## 6 Further Reading and Acknowledgments

The interested reader is invited to consult the books [1, 2, 3] for further information. The author wishes to thank Jacob Barnett for reading the first draft and providing helpful suggestions.

This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ministry of Research, Innovation and Science.

## References

[1] H. Georgi, *Lie Algebras In Particle Physics: from Isospin To Unified Theories*. Frontiers in Physics. Avalon Publishing, 1999.

[2] B. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer International Publishing, 2015.

[3] P. Woit, *Quantum Theory, Groups and Representations: An Introduction*. Springer International Publishing, 2017.

[13] Incidentally, this also means that if we find a matrix which commutes with every representation matrix, but is **not** a multiple of the identity, then the representation is not irreducible.