MATTHEW B WINGATE

SYMMETRIES, PARTICLES, AND FIELDS

These notes are for use with the Cambridge University Part III lecture course *Symmetries, Particles, and Fields*, given in the Department of Applied Mathematics and Theoretical Physics in Michaelmas Term 2025. They draw on the notes of previous lecturers, in particular Ben Allanach and Nick Dorey. Many thanks to James Moore and Matthew Blacker for feedback on the notes and the course; these notes also benefited from their typed solutions to the Examples Sheet problems.

This version was compiled 5 December 2025 (12:30).

Contents

Introduction: Symmetries 5

Lie groups and Lie algebras 9

Representations 25

Angular momentum: SO(3) and SU(2) 35

Relativistic symmetries 45

Classification of Lie algebras 57

Classification of representations 83

Gauge theories 93

Appendix: Discrete groups 103

Appendix: Vector spaces 111

Appendix: Quantum mechanics review 119

Bibliography 125

Introduction: Symmetries

Symmetry is an important concept in physics. When we say that Nature possesses a certain symmetry, we mean that the laws of physics are unchanged after a corresponding transformation. By Nöther's theorem, this implies conservation of some quantity.

We encounter symmetries as soon as we start studying physics. In nonrelativistic classical dynamics, solutions to Newton's law

$$m\frac{d^2\vec{x}}{dt^2} = \vec{F}(\vec{x}) \tag{1.1}$$

simplify greatly if the force is spherically symmetric about some point, say the origin of a coordinate system. If

$$\vec{F}(\vec{x}) = F(r)\hat{r} \tag{1.2}$$

where $r = |\vec{x}|$ and $\hat{r}$ is the unit radial vector, then angular momentum $\vec{x} \times \vec{p}$ is conserved[1] and motion is confined to a plane containing the origin.

[1] Differentiate with respect to time and use (1.1).

Lagrangian mechanics allows us to investigate symmetries more broadly. Recall the Lagrangian $L(q_i, \dot{q}_i; t)$ is a function of a set of coordinates q_i and their time-derivatives $\dot{q}_i$, and possibly depends on time explicitly. The principle of least action says that classical trajectories minimize

$$S = \int_{t_1}^{t_2} dt\; L(q_i(t), \dot{q}_i(t); t)\,. \tag{1.3}$$

That is, under variations of the coordinates $q_i \mapsto q_i + \delta q_i$, $\delta S = 0$ implies that trajectories satisfy the Euler–Lagrange equations

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0\,. \tag{1.4}$$

Noether's theorem says that an invariance in L under some transformation in coordinates yields an associated conserved quantity. For example, let's consider a 3-dimensional system which has Lagrangian

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - U(x, y, z)\,. \tag{1.5}$$

1. Note that this Lagrangian does not explicitly depend on time. We say that it is invariant under translations in time: $t \mapsto t + \delta t$. We can show that this implies that the Hamiltonian, the total energy, is conserved.

The conjugate momentum to each coordinate is defined as

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i\,. \tag{1.6}$$

The Hamiltonian is the Legendre transform of the Lagrangian

$$H(x_i, p_i; t) = \sum_i \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L\,. \tag{1.7}$$

The time derivative of H is

$$\begin{aligned}\frac{dH}{dt} &= -\frac{\partial L}{\partial t} + \sum_i \left[\ddot{x}_i \frac{\partial L}{\partial \dot{x}_i} + \dot{x}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right) - \dot{x}_i \frac{\partial L}{\partial x_i} - \ddot{x}_i \frac{\partial L}{\partial \dot{x}_i}\right] \\ &= -\frac{\partial L}{\partial t}\,,\end{aligned} \tag{1.8}$$

where the terms in the square brackets above cancel by (1.4). Thus, if $\partial L/\partial t = 0$, then $H = \frac{1}{2m}|\vec{p}|^2 + U(x, y, z)$ is conserved.

2. If L is invariant under translations in one variable, then its conjugate momentum is conserved. Say L is independent of x, then the Euler–Lagrange equation (1.4) for x leads to

$$\frac{\partial L}{\partial x} = 0 \implies \frac{\partial L}{\partial \dot{x}} = p_x = \text{const}\,. \tag{1.9}$$

3. If L is invariant under rotations about the z-axis, then z-component of angular momentum is conserved. This is easiest to see in cylindrical coordinates, where $x = \rho\cos\theta$ and $y = \rho\sin\theta$. In this case

$$L = \frac{1}{2}m(\dot{\rho}^2 + \rho^2\dot{\theta}^2 + \dot{z}^2) + U(\rho, z)\,. \tag{1.10}$$

Using the Euler–Lagrange equation (1.4) for θ, we have

$$\frac{\partial L}{\partial \theta} = 0 \implies \frac{\partial L}{\partial \dot{\theta}} = p_\theta = m\rho^2\dot{\theta} = xp_y - yp_x = \text{const}\,. \tag{1.11}$$

Noether's theorem is very powerful when one goes from the dynamics of classical particles to classical field theory and quantum field theory.

Let's stick with particles here, though, but turn to quantum mechanics. We assume familiarity with quantum mechanics, but a very short review is given in § 11.1. When we talk about a symmetry of a quantum system with Hilbert space $\mathcal{H}$, we mean that there is some invertible operator $U : \mathcal{H} \to \mathcal{H}$ which preserves inner products up to an overall phase.

Definition 1.1. Let $|\Psi\rangle$ and $|\Phi\rangle$ be any normalized vectors in Hilbert space $\mathcal{H}$ and let $|U\Psi\rangle = U\,|\Psi\rangle$ and $|U\Phi\rangle = U\,|\Phi\rangle$. We say operator U is a **symmetry transformation operator** if and only if

2025-10-16 Removed errant prime.

$$|\langle U\Phi \,|\, U\Psi\rangle| = |\langle \Phi \,|\, \Psi\rangle|\,. \tag{1.12}$$

Proposition 1.2. (Wigner's theorem) Symmetry transformation operators are either (a) linear and unitary or (b) antilinear and antiunitary.

The proof of Wigner's theorem is beyond the scope of this course (see Weinberg[2]). In most cases of interest, the symmetry operators are linear and unitary. (Time-reversal is one of the few prominent examples of case (b).)

[2] S Weinberg. *The Quantum Theory of Fields I*. Cambridge University Press, 1995. ISBN 0-521-55001-7. URL https://www.cambridge.org/core/books/quantum-theory-of-fields/22986119910BF6A2EFE42684801A3BDF

Let us assume we are interested quantum systems which are invariant under translations in time; i.e. the corresponding Hamiltonian H is independent of time t. In that case, the time evolution of a state is determined by the time-evolution operator[3]

[3] By default, we assume natural units $\hbar = c = 1$ in this course.

$$|\Psi(t)\rangle = \exp(-iHt)\,|\Psi(0)\rangle \ . \tag{1.13}$$

If U is a linear, unitary symmetry operation, then it preserves inner products, so for any states $|\Psi(t)\rangle$ and $|\Phi\rangle$

$$\langle U\Phi|U\Psi(t)\rangle = \langle \Phi|\Psi(t)\rangle = \langle \Phi|e^{-iHt}|\Psi(0)\rangle \ . \tag{1.14}$$

Given the time-independence of the Hamiltonian, we should obtain the same inner product by applying the symmetry operator to $|\Psi(0)\rangle$ first, then evolving in time. Writing things out explicitly, denote $|\Psi'(0)\rangle = U\,|\Psi(0)\rangle$. Evolving this in time gives

2025-10-12: Some clarification

$$e^{-iHt}\,|\Psi'(0)\rangle = e^{-iHt}U\,|\Psi(0)\rangle \ . \tag{1.15}$$

If U is a symmetry transformation, then this state and the state obtained by applying U to (1.13) should be the same up to a phase, i.e. symmetry dictates that

$$e^{-iHt}U\,|\Psi(0)\rangle = Ue^{-iHt}\,|\Psi(0)\rangle \ , \tag{1.16}$$

where we have assumed a trivial overall phase.[4] Taking the inner product with $|U\Phi\rangle$ and insisting on invariance of observables under U implies

[4] One can include a factor $\omega(t)$ above, then use unitarity and Schödinger's equation that $\omega(t)$ is a constant phase. Then the argument below says $[U, H]$ must be proportional to the identity operator. Requiring the energy spectrum to remain invariant then tells us $H = U^\dagger HU$ and hence $[U, H] = 0$.

$$|\langle \Phi|U^\dagger e^{-iHt}U|\Psi(0)\rangle| = |\langle \Phi|e^{-iHt}|\Psi(0)\rangle| \tag{1.17}$$

Either (1.16) or (1.17) imply we must have

$$U^\dagger e^{-iHt}U = e^{-iHt} \, . \tag{1.18}$$

In other words, U and H must commute. If U is antilinear and antilinear, then one can appropriately modify the argument above to again find that U and H must commute.

For example, if the Hamiltonian commutes with the momentum operator $\vec{p}$, then H cannot depend on the position operator $\vec{x}$ and H is invariant under spatial translations: $\vec{x} \mapsto \vec{x} + \vec{a}$. We can form a corresponding unitary operator with the operator-exponential $\exp(i\vec{a} \cdot \vec{p})$. Similarly, if the Hamiltonian commutes with the angular momentum operator $\vec{J}$, this implies rotational invariance. We will return to this example in depth later in the course.

Even approximate symmetries are useful. The isospin and flavour symmetries of hadrons, bound states of quarks, are not exact due to the different masses and electromagnetic charges of the quarks. Nevertheless, the symmetry limit provides a useful approximation which is a good organizing principle.

Some symmetries are not properties of Nature, but reflect redundancies in our mathematical descriptions. At the end of the course we will see how what we learn about Lie groups and representations are used to account for gauge symmetry in quantum field theories.

Lie groups and Lie algebras

Lie groups and Lie algebras are named after Norwegian mathematician Sophus Lie (1842-1899). His surname is pronounced /*LEE*/.

In previous experience, you may have been introduced to the theory of discrete groups. Discrete groups can be finite or infinite. In either case discrete groups have a countable number of elements. We could label the group elements of a discrete group G by an integer index

$$G = \{ g_i \} \quad \text{with } i = 1, 2, \ldots, |G| \,. \tag{2.1}$$

where $|G|$ is the order of the group. An appendix reviewing some main ideas in discrete group theory is given at the end of these notes, starting in § 9.1. Extensive revision should not be necessary, but sometimes it can be useful to invoke examples from discrete group theory.

2.1 Lie groups

Recall the definition and axioms of group theory:

Definition 2.1. A **group** is a set G of objects or elements together with a binary operation, here denoted by $\cdot$, which we will usually call a **product**, where the four group axioms below hold.

(i) **Closure**: The product of any pair of group elements must also be an element of the group. We use the notation

$$\forall g_1, g_2 \in G \,, \quad g_2 \cdot g_1 \in G \tag{2.2}$$

which reads "for all elements g_1 and g_2 in G, their product $g_2 \cdot g_1$ is also an element of G."

(ii) **Associativity**: The group operation must be associative, that is we must have

$$g_3 \cdot (g_2 \cdot g_1) = (g_3 \cdot g_2) \cdot g_1 \tag{2.3}$$

$\forall g_1, g_2, g_3 \in G$.

(iii) **Identity**: One, and only one, element of G, denoted e, must satisfy the following for all $g \in G$

$$e \cdot g = g \cdot e = g \,. \tag{2.4}$$

You can prove by contradiction that the identity must be unique.

(iv) **Inverse**: Each element $g \in G$ must have a unique inverse $g^{-1} \in G$, such that

$$g \cdot g^{-1} = e = g^{-1} \cdot g \,. \tag{2.5}$$

You can also use proof by contradiction to show that for every g, g^{-1} is unique.

Note that we have not insisted that the group operation be commutative. If the following property holds for all $g_1, g_2 \in G$

$$g_2 \cdot g_1 = g_1 \cdot g_2 \tag{2.6}$$

then we say that the group is **commutative** or **Abelian**, otherwise we say the group is **noncommutative** or **nonabelian**.

Our interest in this course are groups which represent continuous symmetry transformations, Lie groups.

Lie groups, also called **continuous groups**, are those with an infinite number of elements, each characterized by one or more continuous parameters. That is, in addition to satisfying all the group axioms, a Lie group is also a manifold, one where group operations are described by smooth maps.

First we give working definitions which will suffice for this course.

Working definition 2.2. A **manifold** is a space which looks Euclidean, like $\mathbb{R}^n$, on small scales, i.e. in a small neighbourhood about any point p. The **dimension** of the the manifold is given by the dimension of the Euclidean space which is isomorphic to the manifold in a small neighbourhood about any point p.

Working definition 2.3. A **differentiable manifold** is a manifold which satisfies certain smoothness conditions, such that notions of differentiability on $\mathbb{R}^n$ extand to open subsets of the manifold.

Definition 2.4. A **Lie group** G is a differentialble manifold with the addition of a group operation $G \times G \to G$ which is a smooth map.

A couple simple but atypical examples are:

1. The set of n-dimensional vectors $\{\vec{x} \in \mathbb{R}^n\}$ form a Lie group under vector addition. $\vec{x}_3 = \vec{x}_1 + \vec{x}_2$ varies smoothly with $\vec{x}_1$ and $\vec{x}_2$. The dimension of the group is n.

2. The set of points on a circle $S^1 = \{\, \theta \mid 0 \le \theta < 2\pi \,\}$ form a Lie group under addition of angles, with the identification of θ with $\theta + 2n\pi$, integer n. In any small neighbourhood, S^1 is isomorphic to $\mathbb{R}^1$, and the dimension of the group is 1. Equivalently, we could define $S^1 = \{\, z \in \mathbb{C} \mid |z| = 1 \,\}$ with the group operation being complex multiplication. The two definitions are isomorphic, being related by the map $z = e^{i\theta}$.

Just like discrete groups, Lie groups can have subgroups. If the subgroup is continuous, and the underlying manifold is a submanifold of the original, then the subgroup is a **Lie subgroup**.

Matrix groups

Most of what we will discuss in this course concern **matrix Lie groups**, those Lie groups whose elements are square, invertible (obviously!) matrices, and whose operation is matrix multiplication. The n-dimensional **general linear group**, $GL(n, \mathbb{F})$ is the set of invertible $n \times n$ matrices over a field $\mathbb{F}$. This is a subset of the set of all $n \times n$ matrices over $\mathbb{F}$, $\mathrm{Mat}_n(\mathbb{F})$. We shall normally take $\mathbb{F}$ to be the real $\mathbb{R}$ or complex $\mathbb{C}$ numbers. The dimension of $GL(n, \mathbb{R})$ is n^2, since the entries of the square matrix are n^2 independent real numbers. The **real dimension** of $GL(n, \mathbb{C})$ is $2n^2$; this is what we will usually mean when we write about dimensionality. Sometimes one might read that the **complex dimension** of $GL(n, \mathbb{C})$ is n^2, in which case the author is counting the number complex parameters needed to specify a group element.

Let us introduce the most important subgroups of $GL(n, \mathbb{R})$ and $GL(n, \mathbb{C})$:

1. The **special linear group**:

$$SL(n, \mathbb{F}) := \{ M \in GL(n, \mathbb{F}) \mid \det M = 1 \} . \tag{2.7}$$

   The restriction that the matrices have unit determinant is a constraint which takes away one $\mathbb{F}$ degree-of-freedom. Hence $\dim SL(n, \mathbb{R}) = \dim GL(n, \mathbb{R}) - 1$. For complex matrices, the dimension decreases by 1 complex degree-of-freedom, or equivalently 2 real degrees-of-freedom.

2. The **orthogonal group**:

$$O(n) := \left\{ M \in GL(n, \mathbb{R}) \;\middle|\; M^{\mathsf{T}} M = I \right\} , \tag{2.8}$$

   where $^{\mathsf{T}}$ denotes a matrix transpose and I is the identity matrix. Counting the number of matrix elements and the number of constraints implied by $M^{\mathsf{T}} M = I$, one can argue that $\dim O(n) = n(n-1)/2$. One possible loophole is whether the constraint equations are all linearly independent. Later in the course we will have a more rigourous proof that this is the correct result for the dimension. Note that $M^{\mathsf{T}} M = I$ implies $\det M = \pm 1$.

   The **special orthogonal group** is the subgroup of $O(n)$ corresponding to matrices M satisfying $\det M = +1$:

$$SO(n) := \{ M \in O(n) \mid \det M = 1 \} . \tag{2.9}$$

3. The **pseudo-orthogonal group**: We begin by defining an $(n+m) \times (n+m)$ **metric matrix**

$$\eta := \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix} . \tag{2.10}$$

   The **pseudo-orthogonal** group $O(n, m)$ is defined to be the set of real matrices such that

$$O(n, m) := \left\{ M \in GL(n+m, \mathbb{R}) \;\middle|\; M^{\mathsf{T}} \eta M = \eta \right\} . \tag{2.11}$$

4. The **unitary group**:
$$U(n) := \left\{ M \in GL(n,\mathbb{C}) \;\middle|\; M^\dagger M = I \right\} \tag{2.12}$$
where † is the complex transpose.[5] The **special unitary group** is the subgroup of $U(n)$ of matrices of determinant equal to $+1$
$$SU(n) := \{ M \in U(n) \mid \det M = 1 \} \,. \tag{2.13}$$
We have seen in the introduction the role of unitary matrices in relation to symmetries in quantum theories.

[5] $(M^\dagger)_{ij} = (M_{ji})^*$.

5. The **pseudo-unitary** group $U(n,m)$ is defined analogously to $O(n,m)$, with the transpose replaced by the Hermitian conjugate.

6. The **symplectic group** requires us first to define a matrix Ω, such that Ω is a fixed, antisymetric, $2n \times 2n$ matrix. Usually this is taken to be
$$\Omega := \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \tag{2.14}$$
where I_n is the $n \times n$ identity matrix and the 0 should be read as an $n \times n$ matrix of zeros. The symplectic group is formed of matrices M which satisfy $M^\mathsf{T}\Omega M = \Omega$:
$$Sp(2n,\mathbb{R}) := \left\{ M \in GL(2n,\mathbb{R}) \;\middle|\; M^\mathsf{T}\Omega M = \Omega \right\} \,. \tag{2.15}$$
The condition on M, and the fact that $\det \Omega \neq 0$, implies that $\det M = \pm 1$. In fact, one can show $\det M = 1$. The **Pfaffian** of a $2n \times 2n$, antisymmetric matrix A is defined to be
$$\operatorname{Pf} A := \frac{1}{2^n n!} \epsilon_{i_1 i_2 \dots i_{2n}} A_{i_1 i_2} \cdots A_{i_{2n-1} i_{2n}} \tag{2.16}$$
where ϵ is the $2n$-dimensional antisymmetric symbol with $\epsilon_{1\dots 2n} = 1$.[6] The crucial step is to show
$$\operatorname{Pf}(M^\mathsf{T}\Omega M) = \det M \operatorname{Pf}\Omega \,. \tag{2.17}$$
Using $(M^\mathsf{T}\Omega M) = \Omega$ on the left-hand side, then $\operatorname{Pf}\Omega \neq 0$ implies $\det M = 1$. (In fact $\operatorname{Pf}\Omega = (-1)^{n(n-1)/2}$.)

[6] With some work, one can show $(\operatorname{Pf} A)^2 = \det A$. Note that the Pfaffian is only defined for antisymmetric (i.e. skew-symmetric) matrices.

The symplectic group can be defined as above, but over the field of complex numbers
$$Sp(2n,\mathbb{C}) = \left\{ M \in GL(2n,\mathbb{C}) \;\middle|\; M^\mathsf{T}\Omega M = \Omega \right\} \,. \tag{2.18}$$
An alternative definition of the symplectic group replaces the matrix Ω (2.14) in the definition (2.15) by another matrix J, where
$$J = \begin{pmatrix} \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & 0 & \cdots & 0 \\ 0 & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \end{pmatrix} \,. \tag{2.19}$$

It is straightforward to see that $\operatorname{Pf} J = 1$.

Group elements as transformations

One can define **group actions** on set elements. We will primarily be concerned with how group elements act on elements of the same group and with how group elements act on vectors in vector spaces. First we give some definitions for actions by elements of a group G on an abstract set X.

Definition 2.5. The **left action** of G on X is defined by a map $\mathrm{L} : G \times X \to X$ such that $L(e, x) = x$ and $L(g_2, L(g_1, x)) = L(g_2 g_1, x)$ for all $g_1, g_2 \in G$ and all $x \in X$.

Often a shorter notation is used, where the map is written as multiplication, i.e. the action of any $g \in G$ on any $x \in X$ is given by $g : X \to X$ with $g(x) = gx$.

Definition 2.6. The **right action** of G on X is defined by $g : X \to X$ with $g(x) = xg^{-1}$, for all $g \in G$ and $x \in X$.

The right action uses the inverse of g in order to preserve group composition law, $g_2(g_1(x)) = (g_2 g_1)(x)$.

Definition 2.7. **Conjugation** by G on X is the action defined by $g(x) = gxg^{-1}$, for all $g \in G$ and $x \in X$.

Sometimes it is useful to consider orbits of group transformations.

Definition 2.8. Given a group G and a set X, an **orbit** of an element $x \in X$ is the set of elements in X which are in the image of an action of G on x.

For example, if we have a left action of G on X, then the orbit of any $x \in X$ is

$$Gx = \{\, gx \mid g \in G \,\} \,. \tag{2.20}$$

The set of orbits under the action of G partition X. This concept underlies quotient groups, which we will see elsewhere.

Matrices represent linear operations which act as transformations of vectors in particular vector spaces. Physics is often concerned with how vectors or functions of vectors behave under such transformations.

For example, the group $O(n)$ represents rotations and reflections of vectors in $\mathbb{R}^n$. The usual inner product is preserved under $O(n)$ transformations. The usual inner product of two vectors $\vec{v}_1$ and $\vec{v}_2$ in $\mathbb{R}^n$ is defined as $(\vec{v}_2, \vec{v}_1) = \mathbf{v}_2^{\mathsf{T}} \mathbf{v}_1$, where on the right-hand side we change font to denote the vectors as column matrices. It follows then that for $R \in O(n)$

$$(R\vec{v}_2, R\vec{v}_1) = \mathbf{v}_2^{\mathsf{T}} \mathbf{R}^{\mathsf{T}} \mathbf{R} \mathbf{v}_1 = (\vec{v}_2, \vec{v}_1) \,. \tag{2.21}$$

Similarly for vectors in $\vec{v}_1$ and $\vec{v}_2$ in $\mathbb{C}^n$, the inner product $\langle \vec{v}_2 | \vec{v}_1 \rangle := \mathbf{v}_2^\dagger \mathbf{v}_1$ is preserved under $U(n)$ transformations.

Let us consider the Lie group $SO(2)$, the group of rotations of vectors in $\mathbb{R}^2$:

$$SO(2) = \left\{ R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \middle| \, \theta \in [0, 2\pi) \right\} . \tag{2.22}$$

The group properties are easily checked. Closure implies that the product of two rotations is another rotation, and the smooth dependence on the underlying manifold requires that $R(\theta_2)R(\theta_1) = R(\theta_1 + \theta_2)$, which can be shown to hold explicitly. The identity corresponds to $R(0)$, and $R(\theta)^{-1} = R(2\pi - \theta)$. Of course associativity follows from ordinary matrix multiplication.

Rotations of vectors in $\mathbb{R}^3$ are described by the Lie group $SO(3)$. Three-dimensional rotations are specified by a unit vector $\vec{n} \in S^2$ corresponding to the axis of rotation and an angle θ; therefore, $\dim SO(3) = 3$. Note that a rotation of angle $\theta \in [-\pi, 0]$ about $\vec{n}$ is equivalent to a rotation of angle $-\theta$ about $-\vec{n}$, so it suffices to confine $\theta \in [0, \pi]$ with $\vec{n} \in S^2$. We can depict the manifold of $SO(3)$ as a ball of radius π in $\mathbb{R}^3$, each point $\theta\vec{n}$ corresponding to an angle and direction. Antipodal points on the surface of the ball are identified with each other: for any $\vec{n}$ rotations clockwise and anticlockwise by an angle π are equivalent, so we must have $\pi\vec{n} = -\pi\vec{n}$.

Later we will show that elements of $SO(3)$ can be written as

$$R_{ij} = \cos\theta\, \delta_{ij} + (1 - \cos\theta) n_i n_j - \sin\theta\, \epsilon_{ijk} n_k \, . \tag{2.23}$$

The manifolds of $SO(2)$ and $SO(3)$ are examples of compact groups, that is Lie groups which are closed and bounded manifolds.

Definition 2.9. A manifold – in our case a Lie group – is said to be **compact** if it is closed and bounded. Otherwise, the group is **noncompact**.

2025-10-16: Corrected definition, the previous version of which made reference to volume. While it is correct that compact manifolds with Riemannian metrics have finite volume, noncompact manifolds can also have finite volume.

By closed, we mean that every limit point is contained in the manifold. Boundedness requires some notion of distance in order to define. Once we have coordinates for the manifold, boundedness means that the coordinates lie in bounded intervals.

Turning to the pseudo-orthogonal group $SO(n, m)$, these matrices act on vectors in $\mathbb{R}^{n+m}$ and preserve the scalar product $\mathbf{v}_1^\mathsf{T} \eta \mathbf{v}_2$ for vectors $\vec{v}_1$ and $\vec{v}_2$ in the vector space. The introduction of the signs in η does not change the dimensionality of the group manifold compared to the corresponding orthogonal group, so $\dim O(n, m) = \dim O(n + m)$.

Examples familiar from the theory of special relativity are

$$SO(1,1) = \left\{ \begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix} \middle| \, \psi \in \mathbb{R} \right\} \tag{2.24}$$

and $SO(1,3)$ or $SO(3,1)$, depending on your convention for the Minkowski metric. Note these groups are noncompact, as the rapidity ψ is unbounded, e.g. in (2.24).

One application of the symplectic group $Sp(2n,\mathbb{R})$ is in classical dynamics. Consider a system with n coordinates q_i and corresponding canonical momenta p_i, where $i = 1,2,\ldots,n$. The dynamics is given by Hamilton's equations

You are not expected to revise Hamiltonian dynamics for this course's exam.

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}\,. \tag{2.25}$$

The state of the system is given by a $2n$-component vector in phase space $z = (q_1,\ldots,q_n,p_1,\ldots,p_n)^{\mathsf{T}}$. Canonical transformations of variables, $z \mapsto Z$ are those which preserve the form of Hamilton's equations (2.25). This is equivalent to requiring

$$\frac{dZ}{dt} = M(z,t)\frac{dz}{dt} \quad \text{with} \quad M \in Sp(2n,\mathbb{R})\,. \tag{2.26}$$

Parametrization of Lie groups

Through examples, we have exhibited many properties of Lie groups. Here we develop some generic notation, suitable for any Lie group, although we do assume that we are able to work with a global coordinate system. The discussion is much more complicated if we have to worry about the details of navigating multiple charts.

An n-dimensional Lie group G has an associated underlying manifold $\mathcal{M}_G \subseteq \mathbb{R}^n$, so that any element of G we have

$$g(x) \in G\,, \text{ with } x := (x^1,x^2,\ldots,x^n) \in \mathbb{R}^n\,. \tag{2.27}$$

Note that here, this $g(x)$ is different from the discussion of group actions. Here, the x is a coordinate labelling a set of group elements. $g(x)$ is the element of G labelled by the coordinate x. In other words, $g : \mathbb{R}^n \to G$ is the inverse chart map telling us which element of G has coordinate x.

2025-10-16 Notation clarification

Closure under group multiplication states that

$$g(y)g(x) = g(z)\,. \tag{2.28}$$

The smoothness property mentioned earlier is expressed mathematically by the statement that for the product above, the components of z obey

$$z^r = \varphi^r(x,y) \tag{2.29}$$

where the $\varphi^r(x,y)$ are continuously differentiable functions of x and y and $r = 1,\ldots,n$. Usually, one chooses the origin of the manifold such that the corresponding group element is the identity: $g(0) = e$. In this case the property of the group identity implies that

$$\varphi^r(x,0) = x^r \quad \text{and} \quad \varphi^r(0,y) = y^r\,. \tag{2.30}$$

We can similarly use the group inverse to constrain φ. Define $\bar{x}$ such that $g(\bar{x}) = g(x)^{-1}$, then it follows that

$$\varphi^r(\bar{x}, x) = 0 = \varphi^r(x, \bar{x})\,. \tag{2.31}$$

Associativity,

$$g(z)[g(y)g(x)] = [g(z)g(y)]g(x) \tag{2.32}$$

implies

$$\varphi^r(\varphi(x,y), z) = \varphi^r(x, \varphi(y,z))\,. \tag{2.33}$$

The main point is that the group composition laws are given by continuously differentiable functions – up to issues arising with possibly having to have multiple coordinate systems to cover the whole manifold.

2.2 *Lie algebras*

Working with Lie groups can be complicated. For example, if we want to show that two Lie groups are isomophic to each other, then we need to consider both their manifolds and their group structures. Equating two manifolds means showing that all topological invariants are the same for both. Showing the group operations are equivalent means showing the group composition maps are isomorphic – generally these are nonlinear functions, difficult to work with.

Things would be simpler if we could "linearize" Lie groups, examining the behaviour in small neighbourhoods. This is the study of Lie algebras which we introduce in this section.

Lie groups are homogeneous: that is, the neighbourhood about one point in G can be smoothly mapped to a neighbourhood about another point in G. If $h \in G$ is close to $g_1 \in G$, then $g_2 g_1^{-1} h$ is close to $g_2 \in G$ (see Figure 2.1). Therefore, it suffices to linearize near the identity e of a Lie group G.

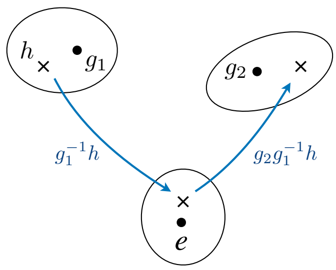


Figure 2.1: Given a Lie group G, imagine small neighbourhoods about the identity e and arbitrary group elements g_1 and g_2 (dots). If we take another group element h in the neighbourhood of g_1, then the smoothness of inversion and group multiplication implies that $g_1^{-1}h$ is close to the identity and hence $g_2 g_1^{-1} e$ is close to g_2 (crosses). In this way, we can construct smooth maps between arbitrary neighbourhoods. This is what we mean by the statement that Lie groups are homogeneous.

Let us define Lie algebras first, without reference to Lie groups, and then make the connection in the next section.

Definition 2.10. A **Lie algebra** is a vector space V which additionally has as a vector product the **Lie bracket** $[\cdot, \cdot] : V \times V \to V$ which must have the following properties. For all $X, Y, Z \in V$:

1. Antisymmetry

$$[X, Y] = -[Y, X]\,. \tag{2.34}$$

2. Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0\,. \tag{2.35}$$

3. Linearity. For $\alpha, \beta \in \mathbb{F}$,

$$[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]\,. \tag{2.36}$$

Note that any vector space which has an associative product $*: V \times V \to V$ can be made into a Lie algebra by defining the Lie bracket to be

$$[X, Y] := X * Y - Y * X\,. \qquad (2.37)$$

Associativity of $*$ is required in order to respect the Jacobi identity. For example, the space of square matrices $\mathrm{Mat}_n(\mathbb{R})$ is a vector space under the operation of matrix addition. Matrix multiplication is a vector product in this space, and the commutator of matrices is a Lie bracket.

We can define basis vectors for the Lie algebra V. Denoted by $\{T_a\}$, with $a = 1, \ldots, \dim V$, we will refer to these as **generators** of the Lie algebra. We define **structure constants** f through the Lie brackets of these basis vectors

$$[T_a, T_b] = f^c{}_{ab}\, T_c\,. \qquad (2.38)$$

Antisymmetry of the bracket implies $f^c{}_{ba} = -f^c{}_{ab}$. The Jacobi identity (2.35), with $X = T_a$, $Y = T_b$, and $Z = T_c$, implies

$$f^e{}_{ad} f^d{}_{bc} + f^e{}_{cd} f^d{}_{ab} + f^e{}_{bd} f^d{}_{ca} = 0\,. \qquad (2.39)$$

We can write a general elements of the Lie algebra as linear combinations of the basis vectors, for example $X = X^a T_a$, where the X^a are coefficients in $\mathbb{F}$. Then

$$[X, Y] = X^a Y^b f^c{}_{ab} T_c\,. \qquad (2.40)$$

2.3 *Lie groups and their Lie algebras*

The goal of this section is to establish that the Lie algebra $L(G)$ of a Lie group G is the tangent space to G at the identity e. It is also common to distinguish a Lie group G from its Lie algebra $\mathfrak{g}$ by a change of case and font, e.g. $L(SU(n)) = \mathfrak{su}(n)$. Below we will see that every Lie group has a corresponding Lie algebra. Several Lie groups can have the same Lie algebra.

Before we give a general presentation, let us examine some simple examples.

Illustrative examples: tangent spaces $T_e(G)$ of some Lie groups G

Recall that $SO(2)$ is 1-dimensional, so we have a single coordinate θ. Group elements may be written as

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad (2.41)$$

where $g(0)$ now corresponds to the identity of $SO(2)$. The tangent vector to the identity for this g is

$$\left.\frac{dg}{d\theta}\right|_0 = \left.\begin{pmatrix} -\sin\theta & -\cos\theta \\ \cos\theta & -\sin\theta \end{pmatrix}\right|_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \qquad (2.42)$$

The 1-dimensional tangent space at the origin is spanned by the single vector

$$\mathrm{e}_\theta := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{2.43}$$

Thus the tangent space of $SO(2)$ at the identity is

$$L(SO(2)) = T_e(SO(2)) = \left\{ \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \middle| \, a \in \mathbb{R} \right\} . \tag{2.44}$$

We are getting ahead of ourselves by claiming that this is a Lie algebra; we still have to show that there is a suitable Lie bracket. This will come later.

Now for general $SO(n)$. Let us consider a curve in $SO(n)$, i.e. a single-parameter family of orthogonal matrices, passing through the identity I: $M(t) \in SO(n)$ such that $M(0) = I$. By the group definition, $M^\mathsf{T}(t)M(t) = I$ for all t. Hence

$$\begin{aligned} 0 &= \frac{d}{dt}\left[M^\mathsf{T}(t)M(t)\right] \\ &= \frac{dM^\mathsf{T}}{dt} M + M^\mathsf{T} \frac{dM}{dt} \\ &= \dot{M}^\mathsf{T} M + M^\mathsf{T} \dot{M} , \end{aligned} \tag{2.45}$$

where we introduce the notation that a dot represents directional differentiation with respect to its parameter. Looking at $t = 0$ and using $M(0) = I$ we infer that

$$\dot{M}^\mathsf{T}\Big|_0 = -\dot{M}\big|_0 , \quad \text{or} \quad \frac{dM^\mathsf{T}}{dt}\bigg|_0 = -\frac{dM}{dt}\bigg|_0 . \tag{2.46}$$

In other words, the tangent space (ultimately the Lie algebra) consists of real anti-symmetric $n \times n$ matrices:

$$L(SO(n)) = T_e(SO(n)) \subseteq \left\{ X \in \mathrm{Mat}_n(\mathbb{R}) \,\middle|\, X^\mathsf{T} = -X \right\} =: \mathrm{Skew}_n . \tag{2.47}$$

We note in passing that real, antisymmetric matrices necessarily have 0 entries along the diagonal, therefore they are traceless.

Note that $L(O(n)) = L(SO(n))$, since a matrix $R \in O(n)$ which is in a neighbourhood with the identity also has $\det R = 1$. The matrices in $O(n)$ with determinant equal to -1 correspond to a disconnected part of the group manifold.

Turning to the unitary groups $SU(n)$, let $M(t)$ be a curve in $SU(n)$ with $M(0) = I$. For small t let us write

$$M(t) = I + tX + O(t^2) \tag{2.48}$$

where X is a matrix of constants. By construction, $X = dM(t)/dt|_0$. Unitarity of M implies

$$I = M^\dagger M \tag{2.49}$$

$$= I + t(X + X^\dagger) + O(t^2) \tag{2.50}$$

from which we conclude $X^\dagger = -X$, i.e. X is antihermitian. Writing out

$$M(t) = \begin{pmatrix} 1+tX_{11} & tX_{12} & tX_{13} & \cdots \\ tX_{21} & 1+tX_{22} & tX_{23} & \cdots \\ tX_{31} & tX_{32} & 1+tX_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.51}$$

we have

$$1 = \det M \tag{2.52}$$

$$= 1 + t \operatorname{Tr} X + O(t^2)\,. \tag{2.53}$$

This shows that X must be traceless.[7] If we recall Jacobi's formula for the derivative of the determinant of an invertible matrix[8]

$$\frac{d}{dt} \det M(t) = \det M(t) \operatorname{Tr}\left[M(t)^{-1} \frac{dM}{dt} \right]$$

$$0 = \operatorname{Tr} X \tag{2.54}$$

where we set $t = 0$ in the second line. Thus we find that the Lie algebra of $SU(n)$ consists of traceless, antihermitian matrices.

$$L(SU(n)) \subseteq \left\{ X \in \mathrm{Mat}_n(\mathbb{C}) \;\middle|\; X^\dagger = -X,\ \operatorname{Tr} X = 0 \right\}. \tag{2.55}$$

For $U(n)$, the determinant can be any phase $e^{i\theta}$, so X need not be traceless and

$$L(U(n)) \subseteq \left\{ X \in \mathrm{Mat}_n(\mathbb{C}) \;\middle|\; X^\dagger = -X \right\}. \tag{2.56}$$

Later we will show that (2.47), (2.55), and (2.56) are all satisfied with equality. Any matrix in the set on the righthand side is also in the corresponding Lie algebra.

[7] Antihermiticity alone does not imply tracelessness. Can you think of an antihermitian 2×2 matrix which has a nonzero trace? How about a 1×1 antihermitian matrix?

[8] See Wikipedia's entry for Jacobi forumla.

Lie algebra of a matrix Lie group

We have examined the tangent spaces of some matrix Lie groups above, sometimes referring them as the corresponding Lie algebras without justification. Now we will show that the group product gives rise to a corresponding Lie bracket for the tangent space. We will only do this for matrix Lie groups; the proof for abstract Lie groups is beyond the scope of this course.

Let X_1 and X_2 be elements of the tangent space $T_e(G)$ (ultimately, the Lie algebra $L(G)$):

$$X_1 = \left.\frac{dg_1(t)}{dt}\right|_0 = \dot{g}_1|_0 \qquad X_2 = \left.\frac{dg_2(t)}{dt}\right|_0 = \dot{g}_2|_0 \tag{2.57}$$

for curves $g_1(t), g_2(t) \in G$ with $g_1(0) = g_2(0) = e$. Taking a group product $g_3(t) = g_2(t)g_1(t)$ the result,

$$\begin{aligned} \dot{g}_3|_0 &= \frac{d}{dt}\left[g_2(t)g_1(t)\right]|_0 \\ &= (g_2\dot{g}_1 + \dot{g}_2 g_1)|_0 \\ &= X_1 + X_2\,, \end{aligned} \tag{2.58}$$

is another vector in the tangent space $T_e(G)$. A multiplication in the group leads to an addition in the tangent space.

We can use the group structure to show how a Lie bracket arises. We will show that the group commutator leads to the familiar algebra commutator, which fulfills the conditions of a Lie bracket.

Definition 2.11. The **group commutator** of two group elements g_1 and $g_2 \in G$ is

$$[g_1, g_2]_G := g_1^{-1} g_2^{-1} g_1 g_2 \,. \tag{2.59}$$

Here, let $g_1(t)$ and $g_2(u)$ be two curves in G passing through the identity e. We could choose an arbitrary relation between the parameters of g_1 and g_2, but for simplicity here, let us choose $u = t$. Taylor expanding about the origin, we have

$$\begin{aligned} g_1(t) &= e + tX_1 + t^2 W_1 + O(t^3) \\ g_2(t) &= e + tX_2 + t^2 W_2 + O(t^3) \,. \end{aligned} \tag{2.60}$$

Then respective products are given by

$$\begin{aligned} g_2(t)g_1(t) &= e + t(X_1 + X_2) + t^2(X_2X_1 + W_1 + W_2) + O(t^3) \\ g_1(t)g_2(t) &= e + t(X_1 + X_2) + t^2(X_1X_2 + W_1 + W_2) + O(t^3) \end{aligned} \tag{2.61}$$

Now examine the group commutator

$$h(t) := [g_2(t)g_1(t)]^{-1} \, g_1(t)g_2(t) \,. \tag{2.62}$$

Taking the inverse of one of these (which is also an element of G) and then the product with the other yields

$$\begin{aligned} h(t) &= [g_2(t)g_1(t)]^{-1} \, g_1(t)g_2(t) \\ &= e + t^2[X_1, X_2] + O(t^3) \end{aligned} \tag{2.63}$$

where $[X_1, X_2] := X_1X_2 - X_2X_1$. Note that the steps above could equally well have been done taking $u = -t$, which would result in a minus sign in front of the t^2 term. Thus we can reparametrize this curve as

$$\tilde{h}(s) = e + s[X_1, X_2] + \dots \,. \tag{2.64}$$

Thus $\tilde{h}$ is a curve in the manifold, passing through the identity, which has with tangent vector at the origin $[X_1, X_2] \in L(G)$. This demonstrates that $L(G)$ is closed under the commutator, which satisfies the defining properties of a Lie bracket (2.34), (2.35), and (2.36). Therefore, the Lie algebra $L(G)$ is the tangent space $T_e(G)$ along with the matrix commutator $[\cdot, \cdot]$ as its Lie bracket.

Note that the commutator is only nonzero if and only if the group G is nonabelian. If G is abelian, then the first line of (2.63) reads $h(t) = e$ for all t, implying the commutator on the second line must vanish.

Tangent space to G at general element g

Let us consider a matrix Lie group $G < GL(n, \mathbb{F})$ and examine the tangent space at a group element $p = g(t_0)$ which sits on a curve $\mathcal{C} = g(t)$. The tangent space is denoted $T_p(G)$. The curve need not pass through the group identity e. Looking at a nearby point along the curve, $g(t_0 + \varepsilon)$, let us expand

$$g(t_0 + \varepsilon) = g(t_0) + \varepsilon \dot{g}(t_0) + O(\varepsilon^2)\,. \tag{2.65}$$

Since $g(t_0 + \varepsilon)$ is a group element and the group operation is smooth, there must be some element of G, near the identity, $h_p(\varepsilon)$ which satisfies

$$g(t_0 + \varepsilon) = g(t_0) h_p(\varepsilon)\,. \tag{2.66}$$

The p subscript reminds us that h will be different if we choose a different point along $\mathcal{C}$. Expanding about small ε we have

$$h_p(\varepsilon) = e + \varepsilon X_p + O(\varepsilon^2) \tag{2.67}$$

for some $X_p \in L(G)$, again dependent on which point p we are examining. $h_p(\varepsilon)$ is the group element which generates the translation $t_0 \mapsto t_0 + \varepsilon$.

Working through $O(\varepsilon)$

$$\begin{aligned} e + \varepsilon X_p = h_p(\varepsilon) &= g(t_0)^{-1} g(t_0 + \varepsilon) \\ &= g(t_0)^{-1} \left[g(t_0) + \varepsilon \dot{g}(t_0) \right] \\ &= e + \varepsilon g(t_0)^{-1} \dot{g}(t_0)\,. \end{aligned} \tag{2.68}$$

We can identify X_p with $g(t_0)^{-1}\dot{g}(t_0)$. In other words, $g(t_0)^{-1}\dot{g}(t_0) \in L(G)$ for any t_0. By a similar argument $\dot{g}(t_0) g(t_0)^{-1} \in L(G)$. This shows that, although vectors $\dot{g}(t_0) \in T_p(G)$ are not in the Lie algebra, one can map them to the Lie algebra by left or right multiplication by $g(t_0)^{-1}$.

Conversely, consider any $X \in L(G)$. Then there exists a curve $g(t)$ satisfying

$$g^{-1}(t)\dot{g}(t) = X\,. \tag{2.69}$$

If we specify an "initial condition", i.e. that the curve go through a point such that $g(t_0) = g_0$, then the curve satisfying (2.69) will be unique. Existence and uniqueness of this curve follow from the usual existence and uniqueness theorem for differential equations, extended to more general manifolds. The solution is given by the exponential map

$$g(t) = g_0(\exp\ tX)\,. \tag{2.70}$$

For matrix Lie algebras, this is just the matrix exponential

$$\exp\ tX = \sum_{k=0}^{\infty} \frac{1}{k!} (tX)^k\,. \tag{2.71}$$

One parameter subgroups

Given an element of the Lie algebra $X \in L(G)$, the curve parametrized by $t \in J$, where J is an interval in $\mathbb{R}$,

$$g_X(t) = e\,(\exp\ tX) \tag{2.72}$$

is an Abelian subgroup of G, said to be **generated** by X. The group properties are easily checked. We have the identity $g_X(0) = e$ and unique inverses $g_X(t)^{-1} = g_X(-t)$. Associativity is manifest. Finally we have closure and commutativity since

$$g_X(t_2)g_X(t_1) = g_X(t_2 + t_1) = g_X(t_1 + t_2) = g_X(t_1)g_X(t_2)\,. \tag{2.73}$$

If there is an additional $g_X(t_0) = e$ where $t_0 \neq 0$, then $J = [0, t_0)$ and the subgroup is isomorphic to S^1 and therefore compact. Otherwise, the subgroup is isomorphic to $\mathbb{R}$ and is noncompact. (There is also the trivial case where $g_X(t) = e$ for all t.)

2.4 *Lie groups from Lie algebras*

Let us apply the exponential map to elements of the Lie algebra $L(G)$. We will see that we can recover some components of group G. We have already seen that several groups can have the same algebra, so we cannot expect to generate any Lie group solely from its algebra.

We use the exponential map $\exp : L(G) \to G$ such that, for all $X \in L(G)$

$$X \mapsto \exp\ X\,. \tag{2.74}$$

For matrix Lie algebras, the exponential map is defined through the series

$$\exp\ X = \sum_{k=0}^{\infty} \frac{X^k}{k!}\,. \tag{2.75}$$

The curve of matrices parametrized by $t \in \mathbb{R}$ as $M(t) = \exp\ tX$ lie in G since

$$\frac{d}{dt} \exp\ tX \bigg|_0 = X \tag{2.76}$$

for all $X \in L(G)$. In fact, for fixed X, this curve is a one-parameter subgroup of G as discussed in the previous section.

Locally, the map is bijective (one-to-one). However, globally the map is generally not one-to-one. For example, consider the Lie group $U(1)$ and its Lie algebra

$$\begin{aligned} U(1) &= \left\{ e^{i\theta} \;\middle|\; \theta \in (-\pi, \pi) \right\} \\ L(U(1)) &= \{\, ix \mid x \in \mathbb{R} \,\}\,. \end{aligned} \tag{2.77}$$

the map $f : L(U(1)) \to U(1)$ given by $f(x) = \exp\ ix$ is not one-to-one since $\exp(2\pi n i) = 1$ for all integer n. The exponential map is surjective (i.e. it is "onto"), but not injective (it is "many-to-one").

Let us look at the exponential map in the case of the orthogonal group $O(n)$. Let $X \in L(O(n))$ and $M = \exp\ tX$. We want to show that $M \in O(n)$.

2025-10-21: Moved example earlier, as lectured

$$\begin{aligned} M^{\mathsf{T}} &= (\exp\ tX)^{\mathsf{T}} \\ &= \exp\ tX^{\mathsf{T}} \\ &= \exp(-tX) = M^{-1} . \end{aligned} \tag{2.78}$$

Therefore, M is orthogonal. Let's check the determinant. Since X is antisymmetric, and real, its diagonal entries are all zero and $\mathrm{Tr}\ X = 0$. More laboriously,

$$\mathrm{Tr}\ X + \mathrm{Tr}\ X^{\mathsf{T}} = \mathrm{Tr}\ X - \mathrm{Tr}\ X = 0 \tag{2.79}$$

while also

$$\mathrm{Tr}\ X + \mathrm{Tr}\ X^{\mathsf{T}} = \mathrm{Tr}(X + X^{\mathsf{T}}) = 2\,\mathrm{Tr}\ X . \tag{2.80}$$

Therefore $\mathrm{Tr}\ X = 0$.

Let X have eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then

$$\begin{aligned} \det M = \det\ \exp\ tX &= e^{t\lambda_1} e^{t\lambda_2} \cdots e^{t\lambda_n} \\ &= \exp(t\,\mathrm{Tr}\ X) = 1 . \end{aligned} \tag{2.81}$$

A ha! We see that $\exp\ X$ is necessarily in $SO(n)$. We cannot generate an element of $O(n)$ whose determinent is -1 from the Lie algebra. The orthogonal group consists of disconnected manifolds, one containing the proper $SO(n)$ rotations and the other containing improper rotations (those with a reflection).

We are now in a position to return to our statement that $L(SO(n))$ contains every element of Skew_n. Let A be any matrix in Skew_n. Let us use the exponential map to define a curve of matrices $\gamma(t)$ on some manifold

$$\gamma(t) := \exp\ tA . \tag{2.82}$$

In fact, by the above derivation,

$$(\gamma(t))^{\mathsf{T}}\ \gamma(t) = I \tag{2.83}$$

and $\det \gamma(t) = 1$. Therefore, $\gamma(t) \in SO(n)$. By construction,

$$\dot{\gamma}(0) = A \tag{2.84}$$

so matrix A is tangent to a curve in $SO(n)$, and is consequently an element of $L(SO(n))$.

Having a map from the algebra to the group, it remains to check that, given the Lie bracket, can we determine the group product? The answer is yes, via the Baker–Campbell–Hausdorff relation:

$$\exp tX\ \exp tY = \exp tZ \tag{2.85}$$

where

$$Z = X + Y + \frac{t}{2}[X, Y] + \frac{t^2}{12}\Big([X, [X, Y]] + [Y, [Y, X]]\Big) + O(t^3) . \tag{2.86}$$

The proof of this proceeds order-by-order in t and quickly becomes tedious. This universal formula (2.85)–(2.86) shows that the group structure of G near the identity e can be determined by the algebra $L(G)$.

Representations

As interesting as Lie groups and their algebras are, what we need in theoretical physics are the actions of group elements on vectors in vector spaces. We want to understand what types of vector spaces admit Lie group (and algebra) representations.

3.1 Lie group representations

Definition 3.1. A **representation** D of a group G is a smooth group homomorphism

$$D : G \to GL(V) \tag{3.1}$$

from G to the group of automorphisms on a vector space V, called the **representation space** associated with D.

That is, for all $g \in G$, $D(g) : V \to V$ is a bijective, linear map such that

$$v \to D(g)v \quad \text{for } v \in V\,. \tag{3.2}$$

Linearity of the map implies

$$D(g)(\alpha v_1 + \beta v_2) = \alpha D(g)v_1 + \beta D(g)v_2 \tag{3.3}$$

for all $\alpha, \beta \in \mathbb{F}$ and $v_1, v_2 \in V$. In order to be a **group homomorphism** the mapping must preserve the group operation, i.e. we must have

$$D(g_2 g_1) = D(g_2)D(g_1)\,. \tag{3.4}$$

This property (3.4) implies that the identity of the group maps to the identity map on V, I_N if we have a matrix representation, and that the inverses map to inverses:

$$D(e) = \mathrm{id}_V \tag{3.5}$$
$$D(g)^{-1} = D(g^{-1})\,. \tag{3.6}$$

Definition 3.2. The **dimension** of a representation is the dimension of its representation space.

If V is finite dimensional, say $\dim V = N$, then $GL(V)$ is isomorphic to $GL(N, \mathbb{F})$. An isomorphism is easily identified by choosing a basis for V.

Definition 3.3. A representation is **faithful** if $D(g) = \mathrm{id}_V$ only for $g = e$, the group identity. That is, the representation is faithful if $\ker D = \{ e \}$.[9]

Failthfulness is enough to imply that D is injective, i.e. $D(g_1) = D(g_2) \implies g_1 = g_2$.[10]

Let's first consider a simple group and some of its representations. Take $G = (\mathbb{R}, +)$. Any representation D of G must satisfy

$$D(\alpha + \beta) = D(\alpha)D(\beta) \quad \text{with}\, \alpha, \beta \in \mathbb{R}\,. \tag{3.7}$$

Each of these are representations of this group:

(a) For some $k \in \mathbb{R}$, $D(\alpha) = e^{k\alpha}$ is a one-dimensional representation acting on $V = \mathbb{R}$, say. This a faithful representation as long as $k \neq 0$. (If $k = 0$, this is the trivial representation, to be defined below.)

(b) For some $k \in \mathbb{R}$, $D(\alpha) = e^{ik\alpha}$ is a one-dimensional representation on $\mathbb{C}^1$. This is not a faithful representation since $\ker D = \left\{ \frac{2\pi}{k} n \;\middle|\; \forall n \in \mathbb{Z} \right\}$.

(c) We can reformulate the previous example as a 2-dimensional representation acting on $\mathbb{R}^2$, with

$$D(\alpha) = \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix}\,. \tag{3.8}$$

This representation is not faithful because of the periodicity of the trigonometric functions.

(d) We can form an infinite dimensional representation by considering V to be the vector space of all real functions $f : \mathbb{R} \to \mathbb{R}$ and letting

$$\left(D(\alpha)f\right)(x) = f(x - \alpha)\,. \tag{3.9}$$

The representation is faithful since $D(\alpha)f(x) = f(x)$ for all f only if $\alpha = 0$, i.e. $\ker D = \{0\}$.

Let us now list three important representations.

Definition 3.4. The **trivial representation** D_0 is where

$$D_0(g) = 1\,, \quad \forall g \in G\,, \tag{3.10}$$

the 1×1 identity, for all $g \in G$. This trivial representation is not faithful: its dimension is 1 and the kernel of the homomorphism D_0 is all of G.

In fact, we can form trivial representations of any dimension M by mapping $D_0(g) = I_M$ for all $g \in G$. When we introduce the notions of reducible and irreducible representations, we see that we can think of this case as M copies of the 1-dimensional trivial representation.

[9] The **kernel** of a map $D : G \to GL(V)$ consists of the elements of G which map to id_V.

[10] Try a short proof by contradiction.

We can have nontrivial representations of 1-dimension. Consider $O(3)$, for example. We can represent the proper rotations by 1 and the improper rotations by -1. That is, for $A \in O(3)$, we can let

$$D(A) = \det A\,. \tag{3.11}$$

Next we give a working definition which applies for matrix Lie groups and which corresponds to widespread usage in physics. Later in the course, we will give a more precise definition of fundamental representations which is broader than the one given here.

Working definition 3.5. If G is a Lie group of matrices of dimension n, i.e. if $G \leqslant GL(n, \mathbb{F})$, then the representation

$$D_f(g) = g\,, \quad \forall g \in G\,, \tag{3.12}$$

is called the **fundamental representation** or **defining representation** of G.

The dimension of this fundamental representation is n. The kernel of the homomorphism D_f is trivial, so the fundamental representation is faithful. The dimension $\dim D_f = n$.

Let G be a matrix Lie group and consider the case where the vector space V is the corresponding matrix Lie algebra $L(G)$.

Definition 3.6. The **adjoint representation** of G is the representation of G on $L(G)$: $\mathrm{Ad} : G \to GL(L(G))$. The adjoint representation plays such a special role that it has a special denotation, otherwise we might write $D^{\mathrm{adj}} = \mathrm{Ad}$. The mapping $\mathrm{Ad}_g : L(G) \to L(G)$ is given, for any $g \in G$ and any $X \in L(G)$, by

$$\mathrm{Ad}_g X := gXg^{-1}\,. \tag{3.13}$$

The dimension of the adjoint representation is equal to the dimension of the Lie algebra $L(G)$. The adjoint map Ad_g is a linearized version of action of G on itself by conjugation, where "linearized" corresponds to acting on the $O(t)$ term in $g(t) = e + tX + \ldots$.

Let us check that Ad is a well-defined representation, and that it preserves the Lie bracket relations of the algebra

- Closure: For a given X, there is some curve in G, $g(t) = I + tX + \ldots$, with tangent X at $t = 0$. For any $h \in G$, we have another curve in G given by $\tilde{g}(t) = hg(t)h^{-1}$. Near the identity

  $$\tilde{g}(t) = I + thXh^{-1} + \ldots\,. \tag{3.14}$$

  Since hXh^{-1} is tangent to the curve at $t = 0$, it is a vector in the tangent space, thus $\mathrm{Ad}_h X = hXh^{-1} \in L(G)$.

- Group operation:

  $$\begin{aligned} \mathrm{Ad}_{g_2 g_1} X &= (g_2 g_1) X (g_2 g_1)^{-1} \\ &= g_2 g_1 X g_1^{-1} g_2^{-1} \\ &= \mathrm{Ad}_{g_2} \mathrm{Ad}_{g_1} X\,. \end{aligned} \tag{3.15}$$

- Lie bracket: For any $X, Y \in L(G)$ and any $g \in G$,

$$[\mathrm{Ad}_g X, \mathrm{Ad}_g Y] = [gXg^{-1}, gYg^{-1}] = g[X,Y]g^{-1} = \mathrm{Ad}_g([X,Y])\,. \tag{3.16}$$

3.2 *Lie algebra representations*

Analyzing Lie group representations can be difficult, due to the need to account for the global structure of the group manifold. We can simplify things greatly, and learn a lot about the group representations, but first studying the representations of the corresponding Lie algebra instead.

Definition 3.7. A **representation** d of a Lie algebra $L(G)$ (also written $\mathfrak{g}$) is a linear map from $L(G)$ to a set of linear maps within $\mathfrak{gl}(V)$,

$$d : L(G) \to \mathfrak{gl}(V) \tag{3.17}$$

where $\mathfrak{gl}(V) = L(GL(V))$ is the Lie algebra of $GL(V)$, and where d preserves the Lie bracket. In other words d is a Lie algebra homomorphism.

That is, for each $X \in L(G)$, we have $d(X) : V \to V$, a linear map, not necessarily invertible, such that for all $v \in V$

$$v \mapsto d(X)v\,. \tag{3.18}$$

Linearity of each $d(X)$, for all $X \in L(G)$ implies

$$d(X)(\alpha Y + \beta Z) = \alpha\, d(X)Y + \beta\, d(X)Z \tag{3.19}$$

for all $\alpha, \beta \in \mathbb{F}$ and all $Y, Z \in L(G)$. For d to be a Lie algebra homomorphism, we must also have linearity of the map d itself

$$d(\alpha X + \beta Y) = \alpha\, d(X) + \beta\, d(Y)\,, \tag{3.20}$$

and the representation must preserve the Lie bracket, i.e.

$$d\left([X,Y]\right) = [d(X), d(Y)]\,. \tag{3.21}$$

As with group representations, the dimension of an algebra representation is the dimension of its corresponding vector space V. If V is finite dimensional, say $N = \dim V$, then we can choose a basis and $d : L(G) \to \mathfrak{gl}(N, \mathbb{F}) \subseteq \mathrm{Mat}_N\mathbb{F}$.

Definition 3.8. The **trivial representation** d_0 simply maps each $X \in L(G)$ to the zero element of $L(G)$

$$d_0(X) = 0\,, \quad \forall\, X \in L(G)\,. \tag{3.22}$$

such that

$$d_0(X)v = \vec{0}\,, \quad \forall\, X \in L(G), \forall\, v \in V\,. \tag{3.23}$$

Definition 3.9. Considering matrix Lie group $G \leqslant GL(n, \mathbb{F})$, the **fundamental representation** of $L(G)$ is $d_f : L(G) \to \text{Mat}_n(\mathbb{F})$ with

$$d_f(X) = X\,, \quad \forall\, X \in L(G)\,. \tag{3.24}$$

The dimension of the fundamental representation is n.

Definition 3.10. The **adjoint representation** of (any) Lie algebra, $\text{ad} : L(G) \to \mathfrak{gl}(L(G))$, is given $\forall X \in L(G)$ by $\text{ad}_X : L(G) \to L(G)$, with

$$\text{ad}_X Y = [X, Y]\,. \tag{3.25}$$

The dimension of ad is $\dim L(G)$. We will discuss the adjoint representation of the Lie algebra, specifically how this action by Lie bracket arises, in the next subsection.

Algebra representations from group representations

Consider tangents to curves in the group manifold which pass through the identity

$$g(t) = e + tX + \ldots \quad \in G\,, \tag{3.26}$$

where $X \in L(G)$. Let D be a representation of G and V be its representation space. Along the curve in the group manifold, we have

$$D(g(t)) = I + td(X) + \ldots\,. \tag{3.27}$$

where I is the identity map on V and (3.27) implicitly defines $d(X)$ in relation to $D(g)$. We can check that these $d(X)$ form a Lie algebra by checking their Lie bracket. Consider two curves

$$\begin{aligned} g_1(t) &= e + tX_1 + t^2 W_1 + \ldots \\ g_2(t) &= e + tX_2 + t^2 W_2 + \ldots \end{aligned} \tag{3.28}$$

and two ways of writing the representation of the product $g_1^{-1} g_2^{-1} g_1 g_2$:

$$D(g_1^{-1} g_2^{-1} g_1 g_2) = D(g_1)^{-1} D(g_2)^{-1} D(g_1) D(g_2)\,. \tag{3.29}$$

On the lefthand side we have

$$\begin{aligned} D(g_1^{-1} g_2^{-1} g_1 g_2) &= D(e + t^2 [X_1, X_2] + \ldots) \\ &= I + t^2 d([X_1, X_2]) + \ldots \end{aligned} \tag{3.30}$$

and on the righthand side

$$D(g_1)^{-1} D(g_2)^{-1} D(g_1) D(g_2) = I + t^2 [d(X_1), d(X_2)] + \ldots\,. \tag{3.31}$$

Therefore, we conclude that

$$d([X_1, X_2]) = [d(X_1), d(X_2)]\,, \tag{3.32}$$

as required.

Let us see this for the case of the adjoint representations of matrix Lie groups and algebras. For $g \in G$ and $X, Y \in L(G)$,

$$\begin{aligned} \mathrm{Ad}_g Y &= gYg^{-1} \\ &= (I + tX)Y(I - tX) + \dots \\ &= Y + t[X, Y] + \dots \\ &= (I + t\,\mathrm{ad}_X + \dots)Y \end{aligned} \tag{3.33}$$

where in the last line we identify $\mathrm{ad}_X Y = [X, Y]$ as before.

Group representations from algebra representations

Given that d is a representation of $L(G)$ and $X \in L(G)$, let $g = \exp X \in G$ and $D(g) = \exp\ d(X)$. We can use the Baker–Campbell–Hausdorff formula to confirm that, for $g_1, g_2 \in G$, D obeys

$$D(g_2 g_1) = D(g_2) D(g_1)\,. \tag{3.34}$$

However, D may not be a representation of G since the exponential map is not guaranteed to be surjective (onto). Generally, it will not be. However,

Proposition 3.11. If G is simply connected, then $\exp\ d$ is guaranteed to be a representation of G.

Definition 3.12. A group G (or a more general manifold) is said to be **simply connected** if

2025-10-27: Corrected last year's version

1. G is path-connected; that is, any two points can be joined by a continuous path,
2. all closed curves in G can be shrunk to a point.

In this case, $D(g) = \exp\ d(X)$ is a representation of G. This is discussed in Hall, §3.6 of the first edition and §5.7 of the second edition.[11]

[11] B C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015. ISBN 978-3319134666. URL https://link.springer.com/book/10.1007/978-0-387-21554-9

3.3 *Useful concepts*

In this section we introduce a number of concepts which are useful in discussing group and algebra representations. Keep in mind that whatever we say about a Lie group representation has implications for its Lie algebra representation, and whatever we say about a Lie algebra's representation informs us about the representations of the groups which share the Lie algebra.

Definition 3.13. Equivalent representations: Representations D_1 and D_2 of G or d_1 and d_2 of $L(G)$ are **equivalent** if there exists a fixed, invertible linear map R or S such that,

$$\begin{aligned} D_2(g) &= R\, D_1(g)\, R^{-1} \quad \forall g \in G \\ \text{or } d_2(X) &= S\, d_1(X)\, S^{-1} \quad \forall X \in L(G)\,. \end{aligned} \tag{3.35}$$

Elaborating, if V_1 and V_2 are the respective representation spaces of the two representations, then R and S are linear isomorphisms $V_1 \to V_2$, such that for all $v \in V_1$,

2025-10-27: Added elaboration

$$D_2(g)Rv = R(D_1(g)v) \quad \forall g \in G$$
$$\text{or } d_2(X)Sv = S(d_1(X)v) \quad \forall X \in L(G)\,. \tag{3.36}$$

Definition 3.14. A representation D of G, or d of $L(G)$ with corresponding vector space V has an **invariant subspace** $W \subset V$ if

$$D(g)w \in W$$
$$\text{or } d(X)w \in W \tag{3.37}$$

for all $w \in W$ and all $g \in G$ or $X \in L(G)$.

All representations have two trivial invariant subspaces, $\{0\}$ and V.

Definition 3.15. An **irreducible representation**, or **irrep**, is a representation with no nontrivial invariant subspaces. A representation which has a nontrivial subspace is **reducible**.

Definition 3.16. The **direct sum** of two vector spaces U and W

$$U \oplus W = \{\, u \oplus w \mid u \in U, w \in W \,\} \tag{3.38}$$

where the direct sum of vectors obeys

$$(u_1 \oplus w_1) + (u_2 \oplus w_2) = (u_1 + u_2) \oplus (w_1 + w_2) \tag{3.39}$$
$$\lambda(u \oplus w) = (\lambda u) \oplus (\lambda w)\,. \tag{3.40}$$

Note that $\dim(U \oplus W) = \dim U + \dim W$.

Definition 3.17. A **totally reducible** representation d of $L(G)$ (or D of G) can be decomposed into irreducible pieces such that V can be written as a direct sum of invariant subspaces $V = W_1 \oplus W_2 \oplus \ldots$, where $d(X)w_i \in W_i$ for any $w_i \in W_i$ and all $X \in L(G)$. A basis for V exists in which each $d(X)$ is block-diagonal

$$d(X) = \begin{pmatrix} d_1(X) & 0 & 0 & \cdots \\ 0 & d_2(X) & 0 & \cdots \\ 0 & 0 & d_3(X) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.41}$$

Let us consider a simple example, again with the group $(\mathbb{R}, +)$. Let V be the space of all 2π-periodic functions $f : \mathbb{R} \to \mathbb{R}$, with $f(x + 2\pi) = f(x)$. We take D as a representation of f such that

$$(D(\alpha)f)(x) = f(x - \alpha)\,. \tag{3.42}$$

Unlike the example given in (3.9), the kernel of this representation is nontrivial; for $\alpha = 2\pi k$ and integer k, we have

$$(D(2\pi k)f)(x) = f(x)\,, \quad \forall f\,. \tag{3.43}$$

Therefore, D is not a faithful representation. D is also reducible. The invariant (one-dimensional) subspaces are given by

$$W_n = \{\, a_n \cos nx + b_n \sin nx \mid \forall a_n, b_n \in \mathbb{R} \,\} \tag{3.44}$$

for each $n \in \mathbb{Z}_{\geqslant 0}$ (b_0 is unnecessary). Invariance is evident since

$$a_n \cos n(x-\alpha) + b_n \sin n(x-\alpha) = a_n' \cos nx + b_n' \sin nx \\ \in W_n\,, \quad \forall \alpha\,, \tag{3.45}$$

where the dashed coefficients are constants which depend on α. [It's tedious, but one can confirm that $D(\beta+\alpha) = D(\beta)D(\alpha)$ holds. Alternatively, one can analytically continue into the complex plane and use complex phases.]

The familiar Fourier decomposition is an illustration of this total reducibility. Any 2π-periodic function can be decomposed into the Fourier sum

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx\right) . \tag{3.46}$$

Thus the vector space V can be written as the infinite direct sum

$$V = W_0 \oplus W_1 \oplus W_2 \oplus \ldots =: \bigoplus_{n=0}^{\infty} W_n \tag{3.47}$$

where we introduce the symbol for a direct sum in the last step. Note that each invariant subspace occurs exactly once in V.

Definition 3.18. Given a Lie group G, a representation D of G whose representation space V has an inner product $\langle \cdot, \cdot \rangle$ is **unitary** if

$$\langle D(g)v, D(g)w \rangle = \langle v, w \rangle$$

for all $g \in G$ and all $v, w \in V$.

This is important in quantum physics, as we expect inner products, especially norms, to be preserved under symmetry transformations.

Proposition 3.19. Given a unitary representation D of G, the corresponding algebra representation d of $L(G)$ is anti-Hermitian (skew-adjoint).

Recall that the adjoint operator $\dagger$ is defined in relation to the inner product, so that $\langle v, A^\dagger w \rangle = \langle Av, w \rangle$ for all $v, w \in V$.

Proposition 3.20. If $\dim V = N$ is finite and D is unitary, then

$$D(g) \in U(N) \quad \forall g \in G \\ d(X) \in L(U(N)) \quad \forall X \in L(G)\,.$$

Let us investigate an important theorem combining many of the concepts introduced above.

Theorem 3.21. (Maschke) A finite-dimensional, unitary representation is either irreducible or totally reducible. We sketch the proof here, leaving more detail for an Examples Sheet question. For each invariant subspace W, one can show that the orthogonal compement $W_\perp$, constructed using the usual inner product and unitarity of the representation, is also an invariant subspace. This implies $V = W \oplus W_\perp$. If W and $W_\perp$ have any nontrivial invariant subspaces, the we repeat the process. Since the representation is finite-dimensional, this process must terminate.

In the case of discrete groups and compact Lie groups, Maschke's theorem can be extended to finite representations which are not elements of $U(N)$. One defines a new group-invariant inner product with respect to which $D(g)$ is unitary. This is known as Weyl's "unitarian trick." The proof is not too complicated for discrete groups, but the derivation for compact Lie groups goes beyond the scope of this course. Nevertheless, we will rely on the applicability of Maschke's theorem to compact Lie groups in later chapters.

Tensor product

We give a brief review of tensor product spaces in Appendix § 10.8.

Given vector spaces V and W, the **tensor product space** $V \otimes W$ is spanned by vectors of the form $v \otimes w \in V \otimes W$, where v and w are elements of V and W, respectively. The **tensor product** of two vectors $v \in V$ and $w \in W$ obeys the following distributive properies

$$\begin{aligned} v \otimes (\lambda_1 w_1 + \lambda_2 w_2) &= \lambda_1 v \otimes w_1 + \lambda_2 v \otimes w_2 \\ (\lambda_1 v_1 + \lambda_2 v_2) \otimes w &= \lambda_1 v_1 \otimes w + \lambda_2 v_2 \otimes w \,. \end{aligned} \tag{3.48}$$

The dimension of $V \otimes W$ is equal to $(\dim V)(\dim W)$. A vector $\Phi \in V \otimes W$ which is equal to the direct product of two vectors $v \in V$ and $w \in W$ can be written as

$$\Phi_A := \Phi_{\alpha a} := v_\alpha w_a \quad (\text{if } \Phi = v \otimes w) \tag{3.49}$$

where the indices $\alpha = 1, 2, \ldots, \dim V$; $a = 1, 2, \ldots, \dim W$; and $A = 1, 2, \ldots, \dim (V \otimes W)$. A simple map between these indices is $A = \alpha(\dim W) + a$.

Note that not all elements of $V \otimes W$ can be written as a direct product of a vector in V with a vector in W. For example, the linear combination

$$\lambda_1 v_1 \otimes w_1 + \lambda_2 v_2 \otimes w_2 \in V \otimes W \tag{3.50}$$

may generally not be written as $\lambda_3 v_3 \otimes w_3$ for some $v_3 \in V$, $w_3 \in W$.

Tensor products allow one to combine representations of Lie groups (and consequently their Lie algebras). Let $D^{(1)}$ and $D^{(2)}$ be representations of G with respective representation spaces V and W. That is, for all $g \in G$

$$\begin{aligned} D^{(1)}(g) : \qquad & v_\alpha \mapsto D^{(1)}(g)_{\alpha\beta}\, v_\beta \,, \quad v \in V \\ D^{(2)}(g) : \qquad & w_a \mapsto D^{(2)}(g)_{ab}\, w_b \,, \quad w \in W \,. \end{aligned} \tag{3.51}$$

The **tensor product representation** $D^{(1)} \otimes D^{(2)}$ acts on vector space $V \otimes W$ such that

$$(D^{(1)} \otimes D^{(2)})(g)_{\alpha a,\beta b} = D^{(1)}(g)_{\alpha\beta}\, D^{(2)}(g)_{ab}\,. \tag{3.52}$$

This acts on vectors in the tensor product space, $\Phi \in V \otimes W$ as

$$\Phi_A = \Phi_{\alpha a} \mapsto D^{(1)}(g)_{\alpha\beta}\, D^{(2)}(g)_{ab}\, \Phi_{\beta b} = \Phi_B\,. \tag{3.53}$$

If we have a product vector, i.e. if we can write $\Phi = v \otimes w$, then

$$(D^{(1)} \otimes D^{(2)})(g)(v \otimes w) := (D^{(1)}(g)v) \otimes (D^{(2)}(g)w)\,. \tag{3.54}$$

Otherwise, Φ is a linear combination of product states and the linear operators can be distributed term-by-term.

In order to find out how tensor products of algebra representations behave, consider a curve of group elements parameterized by t, $g_t \in G$, where $g_0 = e$ and $\dot{g}_0 = X \in L(G)$. We wish to look at the tangent to the curve in the representation space, which we get by looking at the derivative of (3.54) with respect to t

$$\frac{d}{dt}\Big[(D^{(1)} \otimes D^{(2)})(g_t)\Big]_{t=0}(v \otimes w) = \frac{d}{dt}[D^{(1)}(g_t)]_0 v \otimes w \\ + v \otimes \frac{d}{dt}[D^{(2)}(g_t)]_0 w\,. \tag{3.55}$$

Let $d^{(1)}$ and $d^{(2)}$ be the algebra representations corresponding to $D^{(1)}$ and $D^{(2)}$, respectively. We define their tensor product, for $X \in L(G)$, as

$$(d^{(1)} \otimes d^{(2)})(X) = d^{(1)}(X) \otimes \mathrm{id}_W + \mathrm{id}_V \otimes d^{(2)}(X)\,, \tag{3.56}$$

where id_V and id_W are the identity maps on V and W. An examples sheet question will explore the relation between $D^{(1)} \otimes D^{(2)}$ and $d^{(1)} \otimes d^{(2)}$ in more detail.

An important corollary to Maschke's theorem states that finite representations of $d^{(1)} \otimes d^{(2)}$ can be written as the direct sum of irreps of $L(G)$:

$$d^{(1)} \otimes d^{(2)} = \tilde{d}_1 \oplus \tilde{d}_2 \oplus \ldots \oplus \tilde{d}_k = \bigoplus_i \tilde{d}_i\,. \tag{3.57}$$

We will refer to this as **decomposition** of a direct product representation into its constituent irreducible representations.

Angular momentum: SO(3) and SU(2)

$SO(3)$ is the group of rotations in three dimensions. Naturally, representations of $SO(3)$ correspond to physical states with orbital angular momentum, as we will see. In fact, not all angular momentum states can be described by $SO(3)$ representations. Spin angular momentum states of particles with half-integer spin require $SU(2)$ representations. This chapter investigates the relation between these groups and their representations. In an appendix, § 11.2 – § 11.3, we review the treatment of angular momentum as often presented in undergraduate quantum mechanics.

4.1 *Relationship between SO(3) and SU(2)*

The Lie algebra of $SU(2)$ consists of the set of traceless, antihermitian matrices

$$\mathfrak{su}(2) = L(SU(2)) = \left\{ X \in \mathrm{Mat}_2(\mathbb{C}) \;\middle|\; X^\dagger = -X, \mathrm{Tr}\, X = 0 \right\} . \quad (4.1)$$

We can choose a basis for this 3-dimensional vector space to be

$$T_a = -\frac{i}{2}\sigma_a \,, \quad \text{with } a = 1, 2, 3, \quad (4.2)$$

where the σ_a are the Pauli matrices.[12] As mentioned earlier, these basis vectors are often referred to as the **generators** of the Lie algebra.

[12] In most physics literature, the convention is to use Hermitian generators $t_a = iT_a$ so that $\exp(X^a T_a)$ becomes $\exp(-iX^a t_a)$, possibly absorbing the minus sign $-X^a \to X^a$.

Let's look at the Lie brackets of the generators; this will determine the structure constants as in (2.38). Recalling the identity $\sigma_a \sigma_b = I\delta_{ab} + i\epsilon_{abc}\sigma_c$, we have

$$\begin{aligned} [T_a, T_b] &= -\frac{1}{4}(\sigma_a\sigma_b - \sigma_b\sigma_a) \\ &= -\frac{1}{4}(i\epsilon_{abc} - i\epsilon_{bac})\sigma_c \\ &= \epsilon_{abc} T_c \,. \end{aligned} \quad (4.3)$$

We can see that the structure constants for $\mathfrak{su}(2)$ are $f^c{}_{ab} = \epsilon_{abc}$.

The elements of the Lie algebra of $SO(3)$ are the 3×3 antisymmetric, real matrices

$$\mathfrak{so}(3) = L(SO(3)) = \mathrm{Skew}_3 \,. \quad (4.4)$$

A convenient basis is the following,

$$\tilde{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \tilde{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \tilde{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.5}$$

or, more briefly $(\tilde{T}_a)_{bc} = -\epsilon_{abc}$. After using $\epsilon_{acd}\epsilon_{bde} = -\delta_{ab}\delta_{ce} + \delta_{ae}\delta_{bc}$ in one direction and then the other, we see that $[\tilde{T}_a, \tilde{T}_b] = \epsilon_{abc}\tilde{T}_c$, and thus $f^c{}_{ab} = \epsilon_{abc}$, as for the basis vectors of $\mathfrak{su}(2)$. The fact that the two algebras have the same structure constants, and therefore are isomorphic to each other, suggests that there will be similarities between the groups $SU(2)$ and $SO(3)$.

Making contact with quantum mechanics, we note that these generators, (4.2) and (4.5), when multiplied by $-\sqrt{-1}$, are the 2- and 3-dimensional representations of the components of the angular momentum operator $\vec{J} = (J_1, J_2, J_3)$.

To make the connection between the groups complete, let us consider the group manifolds. We discussed $SO(3)$ as an example in § 2.1. The manifold is a 3-ball of radius π, with antipodes identified. An element of $A \in SU(2)$ can be written as

$$A = a_0 I + i\vec{a}\cdot\vec{\sigma} \tag{4.6}$$

with $(a_0, \vec{a})$ real and $a_0^2 + |\vec{a}|^2 = 1$. This manifold is then the unit sphere in $\mathbb{R}^4$, namely S^3.

Recall that the **centre** of a group is the set of all $x \in G$ such that

$$xg = gx \quad \forall g \in G\,. \tag{4.7}$$

The centre $Z(G) \trianglelefteq G$ is a normal subgroup of G (since $gxg^{-1} \in Z(G)$, $\forall g \in G$). The group $SU(2)$ has centre $Z(SU(2)) = \{\, I, -I \,\} \cong \mathbb{Z}_2$. If we look at the coset formed for any $A \in SU(2)$ with the centre, we have

$$A\, Z(SU(2)) = \{\, A, -A \,\}\ . \tag{4.8}$$

The set of all such cosets form a quotient group $SU(2)/\mathbb{Z}_2$ (under coset multiplication) whose manifold is S^3, now with antipodes identified.

We can draw the manifold as the upper half of S^3 (i.e. $a_0 \geqslant 0$) with opposite points on the equator identified. This is just a curved version of the $SO(3)$ manifold. Therefore,

$$SO(3) \cong SU(2)/\mathbb{Z}_2\,. \tag{4.9}$$

We can write down an explicit map $\rho : SU(2) \to SO(3)$ such that, for $A \in SU(2)$

$$\rho(A) = R \quad \text{where } R \text{ has components } R_{ij} = \frac{1}{2}\mathrm{Tr}(\sigma_i A \sigma_j A^\dagger)\,. \tag{4.10}$$

The map is 2-to-1, since $\rho(-A) = \rho(A)$, and is called a **double covering** of $SO(3)$. That is, $SU(2)$ is the **double cover** of $SO(3)$.

There is a theorem which states that every Lie algebra is the Lie algebra of exactly one simply-connected Lie group (see Hall §3.6 in the first edition, §5.7 in the second edition).[13] In the case here, $SU(2)$ is the simply-connected Lie group, while $SO(3)$ is doubly covered by $SU(2)$. To see that $SO(3)$ is not simply connected, we draw a curve from one point on the surface of the manifold, say $\pi\vec{n}$ to its antipode $-\pi\vec{n}$. Because these points are identified – they are coordinates for the same rotation – this is a closed curve. There is no way to shrink this curve to a point, so the manifold is not simply connected.

[13] B C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015. ISBN 978-3319134666. URL `https://link.springer.com/book/10.1007/978-0-387-21554-9`

4.2 *Representations of* $L(SU(2))$

It will be convenient to work with complex vector spaces rather than real ones. Let V be a real vector space and let $\{T_a\}$ be a basis of V:

$$V = \{\, \lambda^a T_a \mid \lambda^a \in \mathbb{R} \,\} \,. \tag{4.11}$$

The **complexification** of V is the complex span of the same basis set:

$$V_{\mathbb{C}} = \{\, \lambda^a T_a \mid \lambda^a \in \mathbb{C} \,\} \,. \tag{4.12}$$

Note that when we talk a vector space being real or complex, we are referring to the field in which the coefficients λ^a live; it has nothing to do with the $\{T_a\}$.

Let $\mathfrak{g} = L(G)$ be a real Lie algebra, and denote its complexification by $\mathfrak{g}_{\mathbb{C}} = L(G)_{\mathbb{C}}$. A representation d of $\mathfrak{g}$ can be extended to $\mathfrak{g}_{\mathbb{C}}$ by imposing

$$d(X + iY) = d(X) + id(Y) \tag{4.13}$$

where $X, Y \in \mathfrak{g}$. Conversely, if we have a representation $d_{\mathbb{C}}$ of $\mathfrak{g}_{\mathbb{C}}$, we can restrict it to d by writing

$$d(X) = d_{\mathbb{C}}(X) \tag{4.14}$$

where $X \in \mathfrak{g} \subset \mathfrak{g}_{\mathbb{C}}$.

A **real form** of a complex Lie algebra $\mathfrak{h}$ is a real Lie algebra $\mathfrak{g}$ whose complexification is $\mathfrak{h}$, i.e. such that

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{h} \,. \tag{4.15}$$

A complex Lie algebra can have multiple nonisomorphic real forms, as we will see later.

Now to our case, $\mathfrak{su}(2)$. Its complexification is

$$\mathfrak{su}(2)_{\mathbb{C}} = \{\, \lambda^a \sigma_a \mid \lambda^a \in \mathbb{C} \,\} \,. \tag{4.16}$$

Note that, while the elements of $\mathfrak{su}(2)$ are the traceless, antihermitian 2×2 matrices, the complexification breaks the antihermiticity property, extending the algebra to all traceless matrices. This is just

$\mathfrak{sl}(2,\mathbb{C})$, the Lie algebra of $SL(2,\mathbb{C})$. In fact this is true for $\mathfrak{su}(n)_\mathbb{C}$, so we have

$$\mathfrak{su}(n)_\mathbb{C} \cong \mathfrak{sl}(n,\mathbb{C})\,. \tag{4.17}$$

Within the complex vector space of $\mathfrak{su}(2)_\mathbb{C}$, we can employ a more useful basis (Cartan–Weyl)

$$\begin{aligned} H &= \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ E_+ &= \frac{1}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ E_- &= \frac{1}{2}(\sigma_1 - i\sigma_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\,. \end{aligned} \tag{4.18}$$

These satisfy

$$\begin{aligned} [H, E_\pm] &= \pm 2E_\pm \\ [E_+, E_-] &= H\,. \end{aligned} \tag{4.19}$$

As we will see below, the motivation for this change of basis will be familiar from quantum mechanics; we can identify $H = 2J_3$, $E_\pm = J_\pm = J_1 \pm J_2$. (Maybe $\vec{S}$ was used rather than $\vec{J}$. We assume natural units, where $\hbar = 1$.) The $J_\pm$ are usually called **raising** and **lowering operators**, respectively. By multiplying J_3 by two, we will simplify the mathematics, avoiding the need to work with both integers and "half-integers," even though this is what the physics ultimately dictates.

Recall that $\mathrm{ad}_X Y = [X, Y]$, then the first relation in (4.19 is equivalent to

$$\mathrm{ad}_H E_\pm = \pm 2E_\pm\,, \tag{4.20}$$

and we also have

$$\mathrm{ad}_H H = [H, H] = 0\,. \tag{4.21}$$

We see that E_-, H, E_+ are eigenvectors of ad_H, with respective eigenvalues $-2, 0, 2$. The nonzero eigenvalues of ad_H, ± 2, are called the **roots** of $\mathfrak{su}(2)$. When we return to this topic later, we will see that roots are generally vectors in a vector space.

Let d be a finite-dimensional irreducible representation of $\mathfrak{su}(2)$ with representation space V. Write an eigenvector of $d(H)$ as v_λ:

$$d(H)v_\lambda = \lambda v_\lambda \tag{4.22}$$

The eigenvalues λ of $d(H)$ are the **weights** of the representation d. The operators $E_\pm$ are called **ladder** or **step** operators. Let's see why. Apply $d(H)$ to the vectors resulting from applying the ladder operators to some v_λ:

$$\begin{aligned} d(H)d(E_\pm)v_\lambda &= \{d(E_\pm)d(H) + [d(H), d(E_\pm)]\}\, v_\lambda \\ &= \{d(E_\pm)\lambda + d\left([H, E_\pm]\right)\}\, v_\lambda \\ &= (\lambda \pm 2)d(E_\pm)v_\lambda\,. \end{aligned} \tag{4.23}$$

Therefore, as long as $d(E_\pm)v_\lambda \neq 0$, we see that $d(E_\pm)v_\lambda$ are eigenvectors of $d(H)$ with eigenvalues $\lambda \pm 2$.

If we have a finite-dimensional representation, then there are only a finite number of eigenvalues. There must be some Λ such that we have

$$d(H)v_\Lambda = \Lambda v_\Lambda \quad \text{and} \quad d(E_+)v_\Lambda = 0 \,. \tag{4.24}$$

The eigenvalue Λ is called the **highest weight**, in the sense that the latter relation above implies that $\Lambda + 2$ is not an eigenvalue. Now apply $d(E_-)$ n times to define

$$v_{\Lambda-2n} = \left(d(E_-)\right)^n v_\Lambda \,. \tag{4.25}$$

The process must terminate for some $n = N$, again since the representation is finite. This suggests we have a basis for the irrep

$$\{v_\Lambda, v_{\Lambda-2}, \ldots, v_{\Lambda-2N}\} \,. \tag{4.26}$$

Applying the raising operator, we have

$$d(H)d(E_+)v_{\Lambda-2n} = (\Lambda - 2n + 2)d(E_+)v_{\Lambda-2n} \,. \tag{4.27}$$

Is this vector that we create with the raising operator (4.27) the same one we generated by repeatedly applying the lowering operator (4.25)? If not, then we would have degenerate eigenvalues and (4.26) would not span the whole vector space. Let's check.

$$\begin{aligned} d(E_+)v_{\Lambda-2n} &= d(E_+)d(E_-)v_{\Lambda-2n+2} \\ &= \{d(E_-)d(E_+) + [d(E_+), d(E_-)]\}\, v_{\Lambda-2n+2} \\ &= d(E_-)d(E_+)v_{\Lambda-2n+2} + (\Lambda - 2n + 2)v_{\Lambda-2n+2} \end{aligned} \tag{4.28}$$

where we have used $[d(E_+), d(E_-)] = d(H)$ between the second and third lines.

Eq. (4.28) is a recursion relation we can solve as follows. First consider $n = 1$

$$d(E_+)v_{\Lambda-2} = 0 + \Lambda v_\Lambda \,. \tag{4.29}$$

Applying the raising operator to $v_{\Lambda-2}$ just gives us a vector parallel to the v_Λ we started with. Let's look at $n = 2$

$$\begin{aligned} d(E_+)v_{\Lambda-4} &= d(E_+)d(E_-)v_{\Lambda-2} + (\Lambda - 2)v_{\Lambda_2} \\ &= (2\Lambda - 2)v_{\Lambda-2} \,, \end{aligned} \tag{4.30}$$

using (4.29) to obtain the last line. Again, the raising operator just undoes the lowering operator up to a multiplicative factor. In general we will have the form

$$d(E_+)v_{\Lambda-2n} = r_n v_{\Lambda-2n+2} \,. \tag{4.31}$$

Substituting this into (4.28) we find

$$r_n = r_{n-1} + \Lambda - 2n + 2 \tag{4.32}$$

with $r_1 = \Lambda$ from (4.29). The solution is[14]

$$r_n = (\Lambda + 1 - n)n\,. \tag{4.33}$$

[14] Try $r_n = A + Bn + Cn^2$.

Recalling that we can only have a finite number of eigenvalues and that we denoted the smallest one $\Lambda - 2N$, we must have $d(E_-)v_{\Lambda-2N} = 0$, which implies

$$r_{N+1} = 0 = (\Lambda - N)(N+1) \tag{4.34}$$

Since $N + 1$ is a positive integer, we have $\Lambda = N$, an integer.

We conclude that the finite-dimensional irreducible representations of $\mathfrak{su}(2)$ are labelled by $\Lambda \in \mathbb{Z}_{\geqslant 0}$ as d_Λ with corresponding weights

$$S_\Lambda = \{ -\Lambda, -\Lambda+2, \ldots, \Lambda-2, \Lambda \}\,. \tag{4.35}$$

The work above showed that these weights are nondegenerate; consequently $\dim d_\Lambda = \Lambda + 1$. Some special cases:

- d_0 is the trivial representation.
- d_1 is the fundamental representation.
- d_2 is the adjoint representation.

Of course we also have higher dimension irreps.

2025-11-07: Added remarks on real form

The irreps of the $\mathfrak{su}(2)$ are just those of $\mathfrak{su}(2)_\mathbb{C}$. It suffices to take as a real form, the real span of $\{ T_a \}$ (4.2), where, from (4.18),

$$\begin{aligned} T_1 &= -\frac{i}{2}(E_+ + E_-) \\ T_2 &= -\frac{1}{2}(E_+ - E_-) \\ T_3 &= -\frac{i}{2}H\,. \end{aligned} \tag{4.36}$$

The discussion of this section parallels treatment of spin angular momentum in quantum mechanics. There one introduces a Hermitian angular momentum operator $\vec{J} = (J_1, J_2, J_3)$ which commutes with the Hamiltonian, and choose one component, say J_3, to complete the set of commuting operators. The eigenvectors are labelled by j and m, where $2j \in \mathbb{Z}_{\geqslant 0}$ and $m \in \{ -j, -j+1, \ldots, j-1, j \}$, with eigenvalues

$$\begin{aligned} J^2 \lvert jm\rangle &= j(j+1)\lvert jm\rangle \\ J_3 \lvert jm\rangle &= m \lvert jm\rangle\,. \end{aligned} \tag{4.37}$$

We identify

$$\begin{aligned} d(H) &= 2J_3 \\ d(E_\pm) &= J_1 \pm iJ_2 \end{aligned} \tag{4.38}$$

along with $\Lambda = 2j$ and $n = 2m$.

4.3 *Representations of $SU(2)$ and $SO(3)$*

In this section, we use what we learned about representations of the Lie algebra $\mathfrak{su}(2)$ to construct representations of the groups $SU(2)$ and $SO(3)$.

$SU(2)$ is simply connected, so an $\mathfrak{su}(2)$ irrep d_Λ yields a representation D_Λ of $SU(2)$ through the exponential map (§ 3.2). Given $g = \exp X$, with $X \in \mathfrak{su}(2)$ and $g \in SU(2)$,

$$D_\Lambda(g) = \exp d_\Lambda(X)\,. \tag{4.39}$$

Turning to $SO(3)$, we recall that $SO(3) \cong SU(2)/\mathbb{Z}_2$; any element in $SO(3)$, say A, corresponds to a pair of elements $\{-A, A\}$ in $SU(2)$. The $SU(2)$ representation D_Λ given above is also a representation of $SO(3)$, $\tilde{D}_\Lambda$, if and only if it respects the identification of pairs A and $-A$

$$\tilde{D}_\Lambda(\mathsf{A}) = D_\Lambda(A) = D_\Lambda(-A)\,. \tag{4.40}$$

It is sufficient to check whether $D_\Lambda(I) = D_\Lambda(-I)$, since (4.40) can then be obtained by multiplying both sides by $D_\Lambda(A)$.

First we note that, after multiplication by $i\pi$, the basis element $H \in \mathfrak{su}(2)$ (4.18) maps to

$$-I = \exp i\pi H \in SU(2)\,. \tag{4.41}$$

Therefore, the $SU(2)$ and $\mathfrak{su}(2)$ representations satisfy

$$D_\Lambda(-I) = \exp i\pi\, d_\Lambda(H)\,. \tag{4.42}$$

We found that the representation d_Λ has weights (i.e. $d_\Lambda(H)$ has eigenvalues) in $\lambda \in \{-\Lambda, -\Lambda+2, \ldots, \Lambda\}$. Note that if Λ is odd (even), then all the λ are odd (even). The eigenvalues of $D_\Lambda(-I)$ are then given by

$$\exp i\pi\lambda = (-1)^\lambda = (-1)^\Lambda\,. \tag{4.43}$$

Therefore, we conclude that $D_\Lambda(-I) = D_\Lambda(I)$ if and only if Λ is even. Only the Λ even representations of $\mathfrak{su}(2)$ (and $SU(2)$) can work as representations of $SO(3)$. These correspond to integer values of spin angular momentum j. Representations of $\mathfrak{su}(2)$ with odd Λ correspond to **spinor representations**, where the spin angular momentum j is a half-integer.

4.4 *Tensor products of $L(SU(2))$ representations*

Let us investigate the tensor product representations of $L(SU(2))$. Recalling the general way in which tensor products work (3.56), we have, for $X \in L(SU(2))$,

$$(d_\Lambda \otimes d_{\Lambda'})(X)(v \otimes v') = (d_\Lambda(X)v) \otimes v' + v \otimes \left(d_{\Lambda'}(X)v'\right) \tag{4.44}$$

Note that $\dim(d_\Lambda \otimes d_{\Lambda'}) = (\dim d_\Lambda)(\dim d_{\Lambda'}) = (\Lambda+1)(\Lambda'+1)$.

We seek a decomposition

$$d_\Lambda \otimes d_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathbb{Z}_{\geqslant 0}} \mathcal{L}^{\Lambda''}_{\Lambda,\Lambda'} d_{\Lambda''} \tag{4.45}$$

where $\mathcal{L}^{\Lambda''}_{\Lambda,\Lambda'}$ are nonnegative integers, *viz* multiplicities, (**Littlewood–Richardson coefficients**), counting the number of times irrep $d_{\Lambda''}$ appears in the decomposition of the tensor product representation.

Given bases for the representation spaces V_Λ and $V_{\Lambda'}$ respectively

$$\begin{aligned} &\{v_\lambda\} : \lambda \in S_\Lambda \qquad &&\text{such that} \qquad &&d_\Lambda(H)v_\lambda = \lambda v_\lambda \\ &\{v'_{\lambda'}\} : \lambda' \in S_{\Lambda'} \qquad &&\text{such that} \qquad &&d_{\Lambda'}(H)v'_{\lambda'} = \lambda' v'_\lambda \end{aligned} \tag{4.46}$$

a basis for tensor product space can be formed

$$\left\{ v_\lambda \otimes v'_{\lambda'} \;\middle|\; \lambda \in S_\Lambda \text{ and } \lambda' \in S_{\Lambda'} \right\} . \tag{4.47}$$

Combining (4.44) and (4.46)

$$(d_\Lambda \otimes d_{\Lambda'})(H)(v_\lambda \otimes v'_{\lambda'}) = (\lambda + \lambda')(v_\lambda \otimes v'_{\lambda'}) \tag{4.48}$$

Therefore, weights add in the tensor product to give the weight set of the tensor product representation as

$$S_{\Lambda,\Lambda'} = \left\{ \lambda + \lambda' \;\middle|\; \lambda \in S_\Lambda \text{ and } \lambda' \in S_{\Lambda'} \right\} \tag{4.49}$$

where we count the weights with multiplicity – the same weight can appear more than once.

The highest weight has multiplicity 1 since there is only one way to obtain $\Lambda + \Lambda'$, so $\mathcal{L}^{\Lambda+\Lambda'}_{\Lambda,\Lambda'} = 1$ and we can partial decompose the tensor product representation into the irrep $d_{\Lambda+\Lambda'}$ and a remainder $\tilde{d}_{\Lambda,\Lambda'}$

$$d_\Lambda \otimes d_{\Lambda'} = d_{\Lambda+\Lambda'} \oplus \tilde{d}_{\Lambda,\Lambda'} . \tag{4.50}$$

The remainder representation has weight set $\tilde{S}_{\Lambda,\Lambda'}$ such that

$$S_{\Lambda,\Lambda'} = S_{\Lambda+\Lambda'} \cup \tilde{S}_{\Lambda,\Lambda'} . \tag{4.51}$$

The highest weight in $\tilde{S}_{\Lambda,\Lambda'}$ is $\Lambda + \Lambda' - 2$. This weight had multiplicity 2 in $S_{\Lambda,\Lambda'}$, but one instance of it is included in $S_{\Lambda+\Lambda'}$, so it occurs only once in $\tilde{S}_{\Lambda,\Lambda'}$. Now the argument repeats itself so that eventually we find the tensor product representation fully decomposed into irreps

$$d_\Lambda \otimes d_{\Lambda'} = d_{\Lambda+\Lambda'} \oplus d_{\Lambda+\Lambda'-2} \oplus \ldots \oplus d_{|\Lambda-\Lambda'|} . \tag{4.52}$$

Take, for example, the case $\Lambda = \Lambda' = 1$, corresponding to the tensor product of $j = j' = \frac{1}{2}$ spinor representations. The weight sets

$$\begin{aligned} S_1 &= \{ -1, 1 \} \\ S_{1,1} &= \{ -2, 0, 0, 2 \} . \end{aligned} \tag{4.53}$$

The highest weight is 2 with multiplicity 1, so (4.51) becomes

$$S_{1,1} = \{ -2, 0, 2 \} \cup \{ 0 \} = S_2 \cup S_0 \,. \tag{4.54}$$

Thus we find

$$d_1 \otimes d_1 = d_2 \oplus d_0 \tag{4.55}$$

or the combination of two spin-$\frac{1}{2}$ states can yields spin-1 and spin-0 states. Sometimes equations like (4.55) are written using the representations' dimensions, e.g.

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1} \,. \tag{4.56}$$

Let's continue the connection of this section to the addition of angular momenta in quantum mechanics. Say we wish to investigate the tensor product of irreps $\Lambda_1 = 2j_1$ and $\Lambda_2 = 2j_2$. Let's focus on a specific one, say $\Lambda_3 = 2J$. From our discussion above, we know $J \leqslant j_1 + j_2$. We label the eigenvectors of H in the respective representation spaces by $|j_1 m_1\rangle$ and $|j_2 m_2\rangle$, where $m_i \in \{-j_i, -j_i + 1, \ldots, j_i - 1, j_i\}$ corresponds to index $\lambda/2$ above. Then states in the tensor product irrep can be written as a linear combination of the product vectors as

$$|JM\rangle = \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = M}} C^{J, j_1, j_2}_{m_1, m_2} \, |j_1 m_1\rangle \otimes |j_2 m_2\rangle \ . \tag{4.57}$$

The assignment $M = m_1 + m_2$ reflects that the eigenvalues of $d(H)$ add as in (4.48). The coefficients $C^{J,j_1,j_2}_{m_1,m_2}$ are called **Clebsch–Gordon coefficients** and can be tedious to determine.

Relativistic symmetries

In the previous chapter, we investigated the implications of rotational symmetry. In this chapter, we wish to combine rotations with relativistic boosts and discrete symmetries. At the end, we will also consider translational invariance. The book by Costa and Fogli is a reasonable reference for this chapter.[15] Chapter 2 of Weinberg's Quantum Field Theory I is a thorough treatment, paying more attention to quantum mechanical details than we can.[16]

[15] G Costa and G Fogli. *Symmetries and Group Theory in Particle Physics*. Springer, 2012. ISBN 978-3-642-15481-2. URL `https://link.springer.com/book/10.1007/978-3-642-15482-9`

[16] S Weinberg. *The Quantum Theory of Fields I*. Cambridge University Press, 1995. ISBN 0-521-55001-7. URL `https://www.cambridge.org/core/books/quantum-theory-of-fields/22986119910BF6A2EFE42684801A3BDF`

5.1 Lorentz group

We write 4-vectors as $x = \{x^\mu\} = (x^0, \{x_i\})$

The Lorentz group consists of the set of transformations $x^\mu \mapsto x'^\mu$ which leave scalar products such as $x^\mu \eta_{\mu\nu} x^\nu$ invariant, where we take the Minkowski metric to be

$$\eta_{\mu\nu} := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{5.1}$$

Let Λ be a linear operator associated with such a transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu . \tag{5.2}$$

Invariance of the scalar product implies

$$\begin{aligned} x^\mu \eta_{\mu\nu} x^\nu &= x^\sigma \Lambda^\mu{}_\sigma \eta_{\mu\nu} \Lambda^\nu{}_\rho x^\rho \\ \implies \eta_{\rho\sigma} &= \Lambda^\mu{}_\sigma \eta_{\mu\nu} \Lambda^\nu{}_\rho \\ \eta &= \Lambda^\mathsf{T} \eta \Lambda . \end{aligned} \tag{5.3}$$

The condition (5.3) is exactly that we gave for elements of the pseudo-orthogonal group $O(1,3)$ in (2.11). Since the metric is symmetric, (5.3) consists of 10 constraints. This implies there are 6 independent parameters determining the 4×4 matrices $\Lambda \in O(1,3)$. (We will also see that the dimension of $L(O(1,3))$ is equal to 6.)

The Lorentz group is disconnected; it consists of 4 disjoint sets depending on the signs of $\det \Lambda$ and $\Lambda^0{}_0$.

$$\begin{aligned} \det(\Lambda^\mathsf{T} \eta \Lambda) &= \det \eta \\ \det \Lambda^\mathsf{T} \det \Lambda &= 1 \\ (\det \Lambda)^2 &= 1 \\ \det \Lambda &= \pm 1 . \end{aligned} \tag{5.4}$$

Also, set $\rho = \sigma = 0$ in (5.3)

$$\begin{aligned} \Lambda^{\mu}{}_{0}\eta_{\mu\nu}\Lambda^{\nu}{}_{0} &= \eta_{00} = 1 \\ (\Lambda^{0}{}_{0})^2 - \sum_i (\Lambda^{i}{}_{0})^2 &= 1 \\ (\Lambda^{0}{}_{0})^2 &\geqslant 1 \end{aligned} \tag{5.5}$$

so either $\Lambda^{0}{}_{0} \geqslant 1$ or $\Lambda^{0}{}_{0} \leqslant -1$.

The case with $\det \Lambda = 1$ and $\Lambda^{0}{}_{0} \geqslant 1$ contains the identity, and forms the subgroup $SO(1,3)^{\uparrow}$, the **proper**, **orthochronous** Lorentz group. The other parts of $O(1,3)$ can be obtained from elements of $SO(1,3)^{\uparrow}$ by using the time reversal and parity reversal operators

$$[T^{\mu}{}_{\nu}] := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } [P^{\mu}{}_{\nu}] := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \tag{5.6}$$

Special cases of elements of $SO(1,3)^{\uparrow}$

1. Rotations

$$[(\Lambda_R)^{\mu}{}_{\nu}] := \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \quad \text{with } R \in SO(3)\,. \tag{5.7}$$

   Note that 3 parameters are required to specify an element of $SO(3)$.

2. Lorentz boosts

$$[(\Lambda_B)^{\mu}{}_{\nu}] := \begin{pmatrix} \cosh\psi & -\mathbf{n}^{\mathsf{T}}\sinh\psi \\ -\mathbf{n}\sinh\psi & I - \mathbf{n}\mathbf{n}^{\mathsf{T}}(\cosh\psi - 1) \end{pmatrix} \tag{5.8}$$

   where $\mathbf{n}$ is the column matrix of a unit 3-vector $\vec{n}$. The boost velocity is $\vec{v} = \vec{n}\tanh\psi$ (in $c = 1$ units). The boosts also require 3 parameters: the rapidity ψ and two components of $\vec{n}$.

While the rotations form a subgroup, $SO(3) < SO(1,3)^{\uparrow}$, the boosts do not. The boosts do not close under composition (at least in any spacetime with more than 1 spatial dimension).

In analogy to $SO(3)$ and $SU(2)$, we seek the covering group to $SO(1,3)^{\uparrow}$. In Examples Sheet 2, you will show that the covering group is $SL(2,\mathbb{C})$. In fact, one finds that

$$SO(1,3)^{\uparrow} \cong \frac{SL(2,\mathbb{C})}{\mathbb{Z}_2}\,; \tag{5.9}$$

$SL(2,\mathbb{C})$ is the double cover of $SO(1,3)^{\uparrow}$. They share the same Lie algebra, i.e.

$$\mathfrak{so}(1,3) \cong \mathfrak{sl}(2,\mathbb{C})\,. \tag{5.10}$$

5.2 Lie algebra of the Lorentz group

We expand $\Lambda \in SO(1,3)^{\uparrow}$ about the identity δ^{μ}_{ν},

$$\Lambda^{\mu}{}_{\nu}(t) = \delta^{\mu}{}_{\nu} + t\omega^{\mu}{}_{\nu} + O(t^2) \tag{5.11}$$

where t is a small parameter. Inserting this expression into $\Lambda^{\mathsf{T}}\eta\Lambda = \eta$ yields

$$\begin{aligned} \eta_{\sigma\rho}(\delta^{\sigma}{}_{\mu} + t\omega^{\sigma}{}_{\mu})(\delta^{\rho}{}_{\nu} + t\omega^{\rho}{}_{\nu}) &= \eta_{\mu\nu} + O(t^2) \\ \eta_{\mu\nu} + t(\omega_{\mu\nu} + \omega_{\nu\mu}) &= \eta_{\mu\nu} + O(t^2)\,. \end{aligned} \tag{5.12}$$

This implies that $\omega_{\mu\nu}$ is antisymmetric, or

$$[\omega_{\mu\nu}] = -[\omega_{\nu\mu}] = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix} \tag{5.13}$$

where a through f are real. Elements of the Lie algebra $L(O(1,3))$ are then of the form

$$[\omega^{\sigma}{}_{\nu}] = [\eta^{\sigma\mu}\omega_{\mu\nu}] = \begin{pmatrix} 0 & a & b & c \\ a & 0 & -d & -e \\ b & d & 0 & -f \\ c & e & f & 0 \end{pmatrix}. \tag{5.14}$$

We can form a basis for $L(O(1,3))$ by setting each of a-f equal to ± 1, and the others equal to 0. A conventional choice is the following (with one upper and one lower index, e.g. $K_i \equiv [(K_i)^{\mu}{}_{\nu}]$)

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$
$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{5.15}$$

Using the exponential map, it is evident that K_i generates a Lorentz boost in the x_i direction, whereas a rotation in the plane normal to x_i is generated by J_i.

Let us introduce some notation which will be useful in the next section. We can package the J_i and K_i as elements of the antisymmetric matrix $M^{\mu\nu}$ with

$$\begin{aligned} M^{0j} &= K_j \\ M^{ij} &= \epsilon_{ijk}J_k\,, \end{aligned} \tag{5.16}$$

or equivalently in terms of matrix elements,

$$(M^{\mu\nu})^{\alpha}_{\beta} = \eta^{\mu\alpha}\delta^{\nu}_{\beta} - \eta^{\nu\alpha}\delta^{\mu}_{\beta}\,. \tag{5.17}$$

Now we can write any element of $SO(1,3)^{\uparrow}$ as

$$\Lambda = \exp \frac{1}{2}\omega_{\mu\nu}M^{\mu\nu} \tag{5.18}$$

with real $\omega_{\mu\nu} = -\omega_{\nu\mu}$. Given that $SO(1,3)^{\uparrow}$ is noncompact, a mathematical proof of the surjectivity of the exponential map here is nontrivial. We will simply appeal to the fact that any proper, orthochronous Lorentz transformation can be composed of a sum of arbitrarily small boosts or rotations.

The Lie brackets of the generators (5.15) are

$$\begin{aligned} [J_i, J_j] &= \epsilon_{ijk}J_k \\ [J_i, K_j] &= \epsilon_{ijk}K_k \\ [K_i, K_j] &= -\epsilon_{ijk}J_k\,. \end{aligned} \tag{5.19}$$

We can write any $\Lambda \in SO(1,3)^{\uparrow}$ as

$$\Lambda = \exp(\theta^i J_i + \psi^j K_j) \tag{5.20}$$

with θ^i and $\psi^j \in \mathbb{R}$.

We can simplify the brackets (5.19) by taking the complex linear combination

$$L_i := \frac{1}{2}(J_i + iK_i)R_i \qquad\qquad := \frac{1}{2}(J_i - iK_i)\,. \tag{5.21}$$

The Lie brackets of these take the familiar form of $\mathfrak{su}(2)$ generators, but with two separate copies:

$$\begin{aligned} [L_i, L_j] &= \epsilon_{ijk}L_k \\ [R_i, R_j] &= \epsilon_{ijk}R_k \\ [L_i, R_j] &= 0\,, \quad \forall_{i,j}\,. \end{aligned} \tag{5.22}$$

Let's look at what we have done. We can write any element of $\mathfrak{so}(1,3)^{\uparrow}$ as the linear combination

$$\theta^i J_i + \psi^i K_i \ \text{ with } \ \theta^i, \psi^i \in \mathbb{R}\,. \tag{5.23}$$

Complexifying to $\mathfrak{so}(1,3)^{\uparrow}_{\mathbb{C}}$, we change basis and write any element as

$$\alpha^i L_i + \beta^i R_i \ \text{ with } \ \alpha^i, \beta^i \in \mathbb{C}\,. \tag{5.24}$$

Since the L_i and R_i commute, we have separated the algebra of $\mathfrak{so}(1,3)^{\uparrow}_{\mathbb{C}}$ into two copies of the algebra $\mathfrak{su}(2)_{\mathbb{C}}$. Recalling that $\mathfrak{su}(2)_{\mathbb{C}} \cong \mathfrak{sl}(2,\mathbb{C})$, we have the result

$$\begin{aligned} \mathfrak{so}(1,3)^{\uparrow}_{\mathbb{C}} &\cong \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}} \\ &\cong \mathfrak{sl}(2,\mathbb{C}) \oplus \mathfrak{sl}(2,\mathbb{C})\,. \end{aligned} \tag{5.25}$$

The latter line should not be surprising. We can view a complexified algebra $\mathfrak{g}_{\mathbb{C}}$, the space spanned by basis vectors times complex coefficients, as two copies of its original real form $\mathfrak{g}$, with real coefficients multiplying the first basis set and imaginary coefficients multiplying the second basis set. We will need different real forms for irreducible representations of the Lorentz algebra, however.

To see how representations of the algebra arise, it may (or may not) be helpful to look first at group representations. Let G be the group which is the image of the exponential map of the generators L and R. Let us abbreviate the sum $\alpha_i L_i$ as αL, and similarly for βR. Let us temporarily take the real form where the α^i and β^i are real. That is, write $\Lambda \in G$ as

$$\begin{aligned}\Lambda &= \exp(\alpha L + \beta R) \\ &= \exp(\alpha L)\exp(\beta R) \quad \text{since } [L_i, R_j] = 0 \\ &= U_L U_R =: (U_L, U_R)\,. \end{aligned} \tag{5.26}$$

Above we implicitly define $U_L := e^{\alpha L}$ and $U_R := e^{\beta R}$, and in the last step introduce some notation which anticipates that G is a direct product group.[17] Note that elements of the form U_L form their own subgroup of G, say $A \leqslant G$, as do elements of the form U_R, say $B \leqslant G$. For any $\Lambda_1, \Lambda_2 \in G$, we can write $\Lambda_j := (U_{jL}, U_{jR})$ and find

[17] See Definition 9.10.

$$\Lambda_2\Lambda_1 = (U_{2L}U_{1L}, U_{2R}U_{1R})\,. \tag{5.27}$$

as expected for a direct product group $G = A \times B$.

We can form a representation of a direct product group by taking the tensor product of representations of the two subgroups

$$D^G((a,b)) = D^A(a) \otimes D^B(b)\,. \tag{5.28}$$

In our case

$$\begin{aligned} A &= \left\{ e^{\alpha^i L_i} \;\middle|\; \alpha^i \in \mathbb{C} \right\} \\ B &= \left\{ e^{\beta^i R_i} \;\middle|\; \beta^i \in \mathbb{C} \right\} \end{aligned} \tag{5.29}$$

so we have

$$D^G((e^{\alpha L}, e^{\beta R})) = D^A(e^{\alpha L}) \otimes D^B(e^{\beta R})\,. \tag{5.30}$$

For the corresponding Lie algebra representation, we use the product rule as before to obtain

$$d^{\mathfrak{g}}(\alpha L + \beta R) = \alpha d^{\mathfrak{a}}(L) \otimes I + \beta I \otimes d^{\mathfrak{b}}(R)\,. \tag{5.31}$$

where the $d^{\mathfrak{a}}$ and $d^{\mathfrak{b}}$ are irreps of $\mathfrak{su}(2)$.

The real form discussed above, with two commuting sets of generators, is not what we have in the real Lorentz algebra. The Lie bracket of two boost generators is not a boost generator (5.19). We need to undo the transformation (5.21): $J_i = L_i + R_i$ and $K_i = -i(L_i - R_i)$. For the Lorentz algebra decomposition (5.25), we can use the irreps of $\mathfrak{su}(2)$ on the righthand side. Change labels to now specify $\mathfrak{su}(2)$ irreps by their highest weights (divided by 2), i.e. let $d^{(j)} := d_\Lambda$ with $\Lambda = 2j$. We find

$$\begin{aligned} d^{(j_1,j_2)}(J_i) &= d^{(j_1)}(T_i) \otimes I + I \otimes d^{(j_2)}(T_i) \\ d^{(j_1,j_2)}(K_i) &= -i\left[d^{(j_1)}(T_i) \otimes I - I \otimes d^{(j_2)}(T_i)\right]\,. \end{aligned} \tag{5.32}$$

Let us consider some examples:

- $(j_1, j_2) = (\frac{1}{2}, 0)$. This is the **fundamental representation** of $\mathfrak{sl}(2,\mathbb{C})$ and is the (spinor) representation of a left-handed Weyl fermion.

- $(j_1, j_2) = (0, \frac{1}{2})$. This representation is conjugate to the fundamental representation, also called the **antifundamental representation** of $\mathfrak{sl}(2,\mathbb{C})$ and is the (spinor) representation of a right-handed Weyl fermion.

- $(j_1, j_2) = (\frac{1}{2}, \frac{1}{2})$. This is the representation of 4-vectors. Under $SO(3)$ rotations, this representation is reducible

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3} \tag{5.33}$$

  corresponding to the 4-vector decomposition $(x^0, \vec{x})$. Under the full Lorentz group, $(\frac{1}{2}, \frac{1}{2})$ is irreducible. The 1-to-1 correspondence between 4-vectors and these $(\frac{1}{2}, \frac{1}{2})$ "bispinors" is explored on Example Sheet 2.

- The Dirac spinor representation is $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$.

The classification of irreducible representations of the Lorentz algebra is at the foundation of quantum field theory. Each irrep corresponds to a distinct type of particle or field, with allowed interactions dictated by the tensor product of these representations. While these Lorentz irreps describe how fields transform at a single spacetime point, in the next section we show that including translations further constrains the properties of single-particle states.

5.3 *Poincaré group and algebra*

The Poincaré group combines Lorentz transformations with spacetime translations. This is the isometry group of Minkowski space $\mathcal{M}_4$: $ISO(1,3)$. (The Poincaré group is also sometimes referred to as the inhomogeneous Lorentz group, with $O(1,3)$ being the homogeneous Lorentz group.) The **isometry group** of a metric space is the set of all distance-preserving maps from the metric space onto itself (under composition). We should be aware that we are taking liberties with terminology. The Lorentz scalar product in Minkowski space is not positive semidefinite, so it does not yield a metric or distance in the strict sense. Nevertheless, the terminology persists, at least in the physics literature.

The Poincaré group is the semidirect product[18] of the Lorentz group and the group of translations

[18] See § 9.3.

$$ISO(1,3) = O(1,3) \ltimes T^{1,3} \, . \tag{5.34}$$

Let us use $T^{1,3}$ to denote the translation group, essentially $\mathbb{R}^{1,3}$ under of vector addition. The action of $(\Lambda^\mu_{\ \nu}, a^\mu) \in ISO(1,3)$ on $x \in \mathcal{M}_4$ is

$$x^\mu \mapsto \Lambda^\mu_{\ \nu} x^\nu + a^\mu \, . \tag{5.35}$$

It can be convenient to use 5-dimensional vectors and matrices to rewrite (5.35) as

$$\begin{pmatrix} x \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} \Lambda x + a \\ 1 \end{pmatrix} . \tag{5.36}$$

Group multiplication, or composition of two successive transformations, then reads

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda' & a' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \Lambda\Lambda' & \Lambda a' + a \\ 0 & 1 \end{pmatrix} . \tag{5.37}$$

In the notation of (9.28), $(n_1, h_1) = (a', \Lambda')$, $(n_2, h_2) = (a, \Lambda)$ and $(n_2, h_2)(n_1, h_1) = (a + \Lambda a', \Lambda\Lambda')$, and we can identify $\phi_{h_2}(n_1)$ as $\Lambda a'$.

One can see that the Lorentz group and translation group are each subgroups of the Poincaré group, but (5.37) shows that they do not commute,[19] hence $ISO(1,3) \neq O(1,3) \times T^{1,3}$. To further see the semidirect product holds, note that the intersection of the subgroups is only the identity, then show that the translations are a normal subgroup while the Lorentz group is not a normal subgroup. That is,

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & a' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix}^{-1} \in T^{1,3} \tag{5.38}$$

for all Λ and a, whereas, for some Λ and a,

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda' & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix}^{-1} \notin O(1,3) . \tag{5.39}$$

[19] Compare (5.37) in the case where $a = 0$ and $\Lambda' = I$ to the case where $a' = 0$ and $\Lambda = I$.

Next we turn to the Poincaré algebra. Recall the elements of of $L(O(1,3))$ were the generators of rotations J_i and boosts K_i (where $i = 1, 2, 3$), which we combined into $M^{\mu\nu}$ (5.16). In the 5-component notation of (5.36) we can write a basis for $L(ISO(1,3))$ as

$$\tilde{M}^{\mu\nu} = \begin{pmatrix} M^{\mu\nu} & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \tilde{P}^{\sigma} = \begin{pmatrix} 0 & P^{\sigma} \\ 0 & 0 \end{pmatrix} , \tag{5.40}$$

where $(P^\sigma)^\beta = \eta^{\sigma\beta}$. We have $6 + 4 = 10$ basis elements for the Lie algebra, so we conclude $\dim L(ISO(1,3)) = 10$.

The Lie brackets are

$$\begin{aligned} [\tilde{M}^{\mu\nu}, \tilde{M}^{\rho\sigma}] &= \eta^{\nu\rho}\tilde{M}^{\mu\sigma} - \eta^{\mu\rho}\tilde{M}^{\nu\sigma} + \eta^{\mu\sigma}\tilde{M}^{\nu\rho} - \eta^{\nu\sigma}\tilde{M}^{\mu\rho} \\ [\tilde{M}^{\mu\nu}, \tilde{P}^{\sigma}] &= \eta^{\nu\sigma}\tilde{P}^{\mu} - \eta^{\mu\sigma}\tilde{P}^{\nu} \\ [\tilde{P}^{\mu}, \tilde{P}^{\nu}] &= 0 . \end{aligned} \tag{5.41}$$

The last expression implies that translations in Minkowski spacetime commute, as we would expect.

In the next chapter, we will give a proper definition of "Casimir elements." Here, let us just use this term to refer to products of the generators which commute with all of the generators. One such Casimir element is

$$P^2 = \tilde{P}_\mu \tilde{P}^\mu \,. \tag{5.42}$$

This is a quadratic Casimir, since it is a product of two generators. Another example involves the Pauli–Lubanski pseudovector

$$W_\mu := \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\tilde{M}^{\nu\rho}\tilde{P}^\sigma \,. \tag{5.43}$$

Its square, $W^2 = W_\mu W^\mu$, is a quartic Casimir. As we will learn later, the Casimirs are useful for labelling representations, in particular for infinite-dimensional representations. In fact we have seen a hint of this in § 4.3, in that we labelled representations by their maximum weight Λ. This was then related to the angular momentum quantum number $j = \Lambda/2$, which is in turn related to the eigenvalue of the angular momentum operator $J^2 = J_1^2 + J_2^2 + J_3^2$. This is also a quadratic Casimir. We will return to discuss this in the next Chapter.

5.4 *Representations of the Poincaré group*

There are finite-dimensional representations of the Poincaré algebra (e.g. the 5-dimensional representation given above); however none of these are unitary as we need for quantum theories. The Poincaré group is not compact, so we cannot use the unitarity trick applicable to compact groups. For quantum theories, we need norms to be preserved, so physical states must then be represented by infinite-dimensional unitary representations. In what follows, we restrict our attention to representations suitable for describing single-particle states.

Let us denote a unitary representation of the Poincaré group by U, so that for any $(\Lambda, a) \in ISO(1,3)$, we have an automorphism on vector space V: $U(\Lambda, a) : V \to V$. Because $ISO(1,3)$ is a semi-direct product group, we write a general Poincaré group element as a product of a Lorentz element Λ and a translation a. In terms of the representation, we write

$$U(\Lambda, a) = T(a)U(\Lambda) \tag{5.44}$$

introducing the shorthand $T(a) := U(I, a)$ and $U(\Lambda) := U(\Lambda, 0)$. The fact we can write $(\Lambda, \Lambda a)$ as either $(\Lambda, 0)(I, a)$ or $(I, \Lambda a)(\Lambda, 0)$ implies

$$U(\Lambda)T(a) = T(\Lambda a)U(\Lambda), . \tag{5.45}$$

The group of translations is 4-dimensional, its algebra generated by the P^σ (5.40),

$$(I, a) = \exp a_\sigma P^\sigma \equiv \exp a \cdot P \,, \tag{5.46}$$

and $T(a)$ is the unitary representation of this group element. Let us denote eigenvectors of translations in our representation space V by $|p,s\rangle$:

$$T(a)\,|p,s\rangle = e^{ia\cdot p}\,|p,s\rangle \tag{5.47}$$

with a 4-vector of imaginary eigenvalues ip^σ for each antihermitian P^σ. The p^σ will correspond to particle 4-momentum. The label s represents any degrees-of-freedom which are internal, that is unaffected by translations; we take these to be discrete, as appropriate for single-particle states.

Let's see what happens when we apply a Lorentz transformation Λ to an eigenvector $|p,s\rangle$; is it still an eigenvector of $T(a)$?

$$\begin{aligned} T(a)\,(U(\Lambda)\,|p,s\rangle) &= U(\Lambda)T(\Lambda^{-1}a)\,|p,s\rangle \\ &= U(\Lambda)\,e^{i(\Lambda^{-1}a)\cdot p}\,|p,s\rangle \\ &= e^{i(\Lambda^{-1}a)\cdot p}\,(U(\Lambda)\,|p,s\rangle) \\ &= e^{ia\cdot(\Lambda p)}\,(U(\Lambda)\,|p,s\rangle)\ . \end{aligned} \tag{5.48}$$

Yes! The Lorentz transformation $U(\Lambda)$ maps eigenvalues $p^\mu \mapsto p'^\mu = \Lambda^\mu{}_\nu p^\nu$. Note that

$$p'^2 = p^2 \tag{5.49}$$

as we must have since Lorentz transformations preserve the Lorentz scalar product (5.3).

For any fixed value of p^2, we have a class of momenta all related by Lorentz transformation. We can choose a "standard" momentum k^μ, such that $k^2 = p^2$ and write

$$p^\mu = L(p)^\mu{}_\nu k^\nu \tag{5.50}$$

where $L(p)$ is a Lorentz tranformation, dependent on p^μ. We can then relate eigenvectors

$$|p,s\rangle = U(L(p))\,|k,s\rangle\ . \tag{5.51}$$

Acting on an eigenvector with an arbitrary Lorentz transformation, we have

$$U(\Lambda)\,|p,s\rangle = U(\Lambda L(p))\,|k,s\rangle\ . \tag{5.52}$$

Inserting the identity $I = L(\Lambda p)\,L^{-1}(\Lambda p)$ and using the fact that U is a group homomorphism, we have

$$U(\Lambda)\,|p,s\rangle = U(L(\Lambda p))\,U(L^{-1}(\Lambda p)\Lambda L(p))\,|k,s\rangle\ . \tag{5.53}$$

This second factor simplifies since $L(p)k = p$, then $L^{-1}(\Lambda p)\Lambda p = k$, so

$$W(\Lambda, p) := L^{-1}(\Lambda p)\Lambda L(p) \tag{5.54}$$

is an element of a subgroup of the Lorentz group which leaves k^μ invariant. Eq. (5.53) becomes

$$U(\Lambda)\,|p,s\rangle = U(L(\Lambda p))\,U(W(\Lambda, p))\,|k,s\rangle\ . \tag{5.55}$$

Let's generically denote such transformations by $W^\mu{}_\nu$:

$$W^\mu{}_\nu k^\nu = k^\mu \,. \tag{5.56}$$

Following Wigner, we call this a **little group**.

Let us assume that we can find a representation of the little group such that

$$U(W)\,|k,s\rangle = \sum_{s'} D_{s',s}(W)\,|k,s'\rangle \tag{5.57}$$

where the coefficients $D_{s',s}(W)$ define the representation. Then we can **induce** a representation on the whole vector space using (5.55)

$$\begin{aligned} U(\Lambda)\,|p,s\rangle &= \sum_{s'} D_{s',s}(W(\Lambda,p))\,U(L(\Lambda p))\,|k,s'\rangle \\ &= \sum_{s'} D_{s',s}(W(\Lambda,p))\,|\Lambda p,s'\rangle \,. \end{aligned} \tag{5.58}$$

In going from (5.55) to (5.58), we implicitly commuted the scalar coefficients $D_{s',s}(W)$ to the left. Eq. (5.58) says that we can form representations of the Poincaré group by working with momentum eigenstates in the eigenvector basis of translations. The Lorentz part of the transformation is composed of the transformation on the standard momentum state (5.57), followed by the transformation of the eigenstates (5.58).

There is still one issue to discuss. What are the possible standard momenta k^μ? Considering that proper, orthochronous, Lorentz transformations preserve the sign of the temporal component p^0 and the Lorentz scalar $p^2 = p_\mu p^\mu$, there are six logical possibilities. Single particle states should have timelike 4-momentum, so we can discount spacelike momenta, where $p^2 < 0$ (in the mostly-minus metric signature), for which we can take $k = (0, \vec{K})$. We are also only interested in positive energy states, so we discount the cases $p^0 < 0$, with $p^2 = 0$ and $p^2 > 0$ being separate cases, respective standard momenta could be $k = (-|\vec{K}|, \vec{K})$ and $(-K, \vec{0})$, with $K > 0$. The case where $p^\mu = k^\mu = 0$ describes the vacuum.

Therefore, there are two types of representations we consider for describing single-particle states, and two different procedures for classifying the infinite-dimensional irreps.

For massive states, we have $p^2 = m^2 > 0$. In this case, we can take the standard momentum to be the particle's rest frame where $k = (m, \vec{0})$, and $m > 0$ corresponds to the rest mass. The little group is the group of three-dimensional rotations, $SO(3)$. Hence we can use the results of the previous chapter and classify particles by spin and mass. Quantum states can be labelled by their momentum and total spin (corresponding to the maximum weight of their $\mathfrak{su}(2)$ representation), and then choosing a basis for this eigenspace, e.g. $|p^\mu, j, j_3\rangle$.

Massless states have $p^2 = 0$. Without loss of generality, we can choose standard momentum to be $k := (\omega, 0, 0, \omega)$, with fixed $\omega > 0$. As you will show on Examples Sheet 3, the little group is the isotropy group of $\mathbb{R}^2$, $ISO(2) = SO(2) \ltimes T^2$, the semi-direct product group of rotations and translations in a plane. One shows this by considering the Lorentz transformations which leave k invariant. One can then form generators of the little group from $E_1 := K_1 - J_2$ and $E_2 := K_2 + J_1$ so that the Lie brackets of the corresponding algebra are

$$\begin{aligned}
[J_3, E_1] &= E_2 \\
[E_2, J_3] &= E_1 \\
[E_1, E_2] &= 0\,.
\end{aligned} \tag{5.59}$$

This is the algebra of $ISO(2)$. The analysis of the little group irreps proceeds similarly to what we have already seen, so we will not go into detail here. After some work, one can show that there exist only 2 “helicity” states $h = \pm j$ for total spin j. Quantum states can be labelled by $|p^\mu, j, h\rangle$.

Classification of Lie algebras

In this chapter we extend what we did for $\mathfrak{su}(2)$ to a broader class of Lie algebras. Along the way, we will introduce the notion of roots as geometric objects in a space isomorphic to $\mathbb{R}^n$ and derive conditions which restrict and characterize Lie algebras.

6.1 Definitions

Here we define a few key terms required in the rest of the chapter.

Definition 6.1. A **subalgebra** of a Lie algebra $\mathfrak{g}$ is a vector subspace which is also a Lie algebra under the Lie bracket.

Definition 6.2. An **ideal** (or **invariant subalgebra**) of a Lie algebra $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$ such that $[X, Y] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and all $Y \in \mathfrak{h}$.

An ideal is to an algebra what a normal subgroup is to a group. Every algebra $\mathfrak{g}$ has two trivial ideals, $\{0\}$ and $\mathfrak{g}$.

Definition 6.3. The **derived algebra** of a Lie algebra $\mathfrak{g}$ is

$$\mathfrak{i}(\mathfrak{g}) = [\mathfrak{g}, \mathfrak{g}] := \operatorname{span}_{\mathbb{F}} \{ [X, Y] \mid X, Y \in \mathfrak{g} \} . \tag{6.1}$$

This is an ideal of $\mathfrak{g}$.

Definition 6.4. The **centre** of $\mathfrak{g}$ is

$$J(\mathfrak{g}) := \{ X \in \mathfrak{g} \mid [X, Y] = 0, \ \ \forall Y \in \mathfrak{g} \} . \tag{6.2}$$

This is also an ideal of $\mathfrak{g}$, which can be seen by applying the Jacobi identity.

Definition 6.5. A Lie algebra is **Abelian** if $[X, Y] = 0$ for all $X, Y \in \mathfrak{g}$. In this case $J(\mathfrak{g}) = \mathfrak{g}$ and $\mathfrak{i}(\mathfrak{g}) = \{0\}$.

Definition 6.6. A Lie algebra is **simple** if it is nonabelian and has no nontrivial ideals.

If a Lie algebra $\mathfrak{g}$ is simple, then we have $\mathfrak{i}(\mathfrak{g}) = \mathfrak{g}$ and $J(\mathfrak{g}) = \{0\}$.

2025-11-12: It was a mistake to write on the board that the implication runs the other way too.

Definition 6.7. A Lie algebra is **semisimple** if it has no Abelian, nontrivial ideals.

Proposition 6.8. A semisimple Lie algebra can be decomposed into the direct sum of simple Lie algebras.

The proof of this requires material yet to come, e.g. the Killing form.

6.2 Killing form

The goal of this section is to define an invariant bilinear form on a Lie algebra, and to use this to define a 1-to-1 map from a vector space to its dual. This will prepare us to generalize our analysis of $\mathfrak{su}(2)$ to general Lie algebras. First let us review a few mathematical definitions on which we will rely.

Definition 6.9. An **inner product** is a map on a vector space V over field $\mathbb{F}$ (which we assume is either $\mathbb{C}$ or $\mathbb{R}$)

$$i : V \times V \to \mathbb{F}\,. \tag{6.3}$$

which satisfies the following three properties.

Letting $u, v, w \in V$ and $\alpha, \beta \in \mathbb{F}$,

1. Conjugate symmetry. $(i(u,v))^* = i(v,u)$.
2. **Sesquilinearity** in its arguments. That is, linearity in one of the arguments. In theoretical physics, where $i(u,v)$ is written $\langle u \,|\, v\rangle$, we choose to impose linearity in the second argument: $\langle u \,|\, \alpha v + \beta w\rangle \;=\; \alpha \langle u \,|\, v\rangle + \beta \langle u \,|\, w\rangle$. In mathematics, one usually chooses to impose linearity on the first argument. By conjugate symmetry, the inner product is conjugate-linear in the other argument, e.g. in physics conventions $\langle \alpha u + \beta v \,|\, w\rangle = \alpha^* \langle u \,|\, w\rangle + \beta^* \langle v \,|\, w\rangle$.
3. Positive-definiteness. For $v \neq 0$, the inner product $i(v,v) > 0$.

If $\mathbb{F} = \mathbb{R}$, then conjugate symmetry is just symmetry and sesquilinearity become bilinearity.

Now let us define a similar object

Definition 6.10. A **bilinear form** is a map

$$B : V \times V \to \mathbb{F} \tag{6.4}$$

which is linear in both its arguments, even for $\mathbb{F} = \mathbb{C}$:

$$\begin{aligned} B(u, \alpha v + \beta w) &= \alpha B(u,v) + \beta B(u,w) \\ B(\alpha u + \beta v, w) &= \alpha B(u,w) + \beta B(v,w)\,. \end{aligned} \tag{6.5}$$

Definition 6.11. A **symmetric, bilinear form** satisfies $B(u,v) = B(v,u)$ for all $u, v \in V$. Note that there is no condition on $B(v,v)$.

Definition 6.12. A bilinear form is said to be **negative semidefinite** if $B(v,v) \leqslant 0$ for all $v \in V$. Furthermore, it is said to be **negative definite** if $B(v,v) < 0$ for all $v \in V$.

Definition 6.13. A bilinear form B is said to be **nondegenerate** if, for all nonzero $v \in V$, there exists some $w \in V$ such that

$$B(v,w) \neq 0\,. \tag{6.6}$$

We will see that the degeneracy or nondegeneracy of the bilinear form we are about to define tells us about the type of Lie algebra we are considering. We will be able to use a nondegenerate bilinear form as a kind of generalized inner product.[20]

[20] The scalar product of two 4-vectors in Minkowski space is a familiar example of a nondegenerate, symmetric, bilinear form.

Definition 6.14. The **Killing form** of a Lie algebra $\mathfrak{g}$ is the symmetric, bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$ such that

$$\kappa(X,Y) = \frac{1}{\mathcal{N}} \mathrm{Tr}\left(\mathrm{ad}_X \circ \mathrm{ad}_Y\right) , \tag{6.7}$$

where $\mathcal{N}$ is a normalization factor. We will take $\mathcal{N} = 1$ for the rest of the section.

Let's find a simpler expression for this. Recall that the adjoint representation of an element $X \in \mathfrak{g}$ is a map $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ with $\mathrm{ad}_X Z = [X, Z]$ (3.25). Then the composition acting on Z gives

$$\begin{aligned} (\mathrm{ad}_X \circ \mathrm{ad}_Y) Z &= [X, [Y, Z]] \\ &= X^a Y^b Z^c [T_a, [T_b, T_c]] \\ &= X^a Y^b Z^c f^d{}_{bc} [T_a, T_d] \\ &= X^a Y^b Z^c f^d{}_{bc} f^e{}_{ad} T_e \\ &= M(X,Y)^e_c Z^c T_e \end{aligned} \tag{6.8}$$

defining $M(X,Y)^e_c := X^a Y^b f^e{}_{ad} f^d{}_{bc}$. The Killing form is the trace

$$\kappa(X,Y) = \mathrm{Tr}\, M(X,Y) = X^a Y^b f^c{}_{ad} f^d{}_{bc} = \kappa_{ab} X^a Y^b , \tag{6.9}$$

defining

$$\kappa_{ab} := f^c{}_{ad} f^d{}_{bc} \in \mathbb{F} . \tag{6.10}$$

Note that $\kappa_{ba} = \kappa_{ab}$. Bilinearity of $\kappa(X,Y)$ follows from bilinearity of the Lie bracket, and symmetry is evident from the trace, so the Killing form is a symmetric, bilinear form.

If $\mathfrak{g}$ is finite-dimensional, we can associate a matrix with the linear maps ad_X, for example

$$\mathrm{ad}_{T_a} T_b = [T_a, T_b] = f^c{}_{ab} T_c \tag{6.11}$$

implies matrix elements

$$(\mathrm{ad}_{T_a})^c{}_b = f^c{}_{ab} . \tag{6.12}$$

Thus,

$$\kappa(T_a, T_b) = \mathrm{Tr}[(\mathrm{ad}_{T_a})^e{}_d (\mathrm{ad}_{T_b})^d{}_c] = \mathrm{Tr}(f^e{}_{ad} f^d{}_{bc}) = \kappa_{ab} . \tag{6.13}$$

Next we want to show that the Killing form is group-invariant. Group transformations acting on elements of the Lie algebra transform under the adjoint representation of G, $\mathrm{Ad}_g X = g X g^{-1}$ for $g \in G$ (3.13). What we want to show is that

$$\kappa(\mathrm{Ad}_g X, \mathrm{Ad}_g Y) = \kappa(X,Y) \quad \forall g \in G . \tag{6.14}$$

While this can be done generally,[21] we are most interested in consequences for small group transformation. That is, let $g = e + tZ + O(t^2)$ and neglect subleading terms in t, so that

$$gXg^{-1} = X + t\,\mathrm{ad}_Z X\,. \tag{6.15}$$

The effect on the Killing form, for an infinitesimal transformation, dropping terms $O(t^2)$,

$$\begin{aligned}\kappa(\mathrm{Ad}_g X, \mathrm{Ad}_g Y) &= \kappa(X + t\,\mathrm{ad}_Z X, Y + t\,\mathrm{ad}_Z Y)\\ &= \kappa(X,Y) + t\left[\kappa(\mathrm{ad}_Z X, Y) + \kappa(X, \mathrm{ad}_Z Y)\right]\,. \end{aligned}\tag{6.16}$$

Invariance of the Killing form implies that the $O(t)$ term vanishes, i.e. that

$$\begin{aligned}\kappa(\mathrm{ad}_Z X, Y) + \kappa(X, \mathrm{ad}_Z Y) &= 0\\ \kappa([Z,X],Y) + \kappa(X,[Z,Y]) &= 0\\ \text{or}\quad \kappa(Y,[Z,X]) &= \kappa([Y,Z],X)\,. \end{aligned}\tag{6.17}$$

We will show this invariance property of the Killing form explicitly next.

First, observe for $Z, W, U \in \mathfrak{g}$, by applying the Jacobi identity and antisymmetry, we have, omitting the composition symbol for brevity,

$$\begin{aligned}\mathrm{ad}_{[Z,W]}U &= [[Z,W],U]\\ &= -[[W,U],Z] - [[U,Z],W]\\ &= [Z,[W,U]] - [W,[Z,U]]\\ &= (\mathrm{ad}_Z\mathrm{ad}_W - \mathrm{ad}_W\mathrm{ad}_Z)U\,. \end{aligned}\tag{6.18}$$

Thus,

$$\begin{aligned}\kappa([Z,X],Y) &= \mathrm{Tr}(\mathrm{ad}_{[Z,X]}\mathrm{ad}_Y)\\ &= \mathrm{Tr}(\mathrm{ad}_Z\mathrm{ad}_X\mathrm{ad}_Y) - \mathrm{Tr}(\mathrm{ad}_X\mathrm{ad}_Z\mathrm{ad}_Y) \end{aligned}\tag{6.19}$$

$$\kappa(X,[Z,Y]) = \mathrm{Tr}(\mathrm{ad}_X\mathrm{ad}_Z\mathrm{ad}_Y) - \mathrm{Tr}(\mathrm{ad}_X\mathrm{ad}_Y\mathrm{ad}_Z)\,. \tag{6.20}$$

The sum of the two equations above gives 0, in accordance with (6.17).

Next we will put the Killing form to use.

Theorem 6.15. (Cartan): The Killing form of a Lie algebra $\mathfrak{g}$ is nondegenerate if and only if $\mathfrak{g}$ is semisimple.

We will only prove this in one direction, namely that κ nondegenerate implies that $\mathfrak{g}$ is semisimple. Suppose $\mathfrak{g}$ is not semisimple. There there exists a nontrivial, Abelian ideal $\mathfrak{a} \subset \mathfrak{g}$. That is, for all $A \in \mathfrak{a}$ and all $X \in \mathfrak{g}$

$$[X, A] \in \mathfrak{a}\,. \tag{6.21}$$

We will show that this implies the Killing form is degenerate. Let us choose a basis $\{T_i\}$ for $\mathfrak{a}$ and extend it to $\mathfrak{g}$:

$$\{T_B\} = \{\, T_i \mid i = 1, \ldots, \dim\mathfrak{a}\,\} \cup \{\, T_\alpha \mid \alpha = 1, \ldots, \dim\mathfrak{g} - \dim\mathfrak{a}\,\} \tag{6.22}$$

[21] The key steps are to show that $\mathrm{ad}_{\mathrm{Ad}_g X} = \mathrm{Ad}_g \circ \mathrm{ad}_X \circ \mathrm{Ad}_{g^{-1}}$ and that $\mathrm{Ad}_{g^{-1}} = (\mathrm{Ad}_g)^{-1}$.

B is an index running $1, \ldots, \dim \mathfrak{g}$. The fact that $\mathfrak{a}$ is Abelian implies

$$[T_i, T_j] = 0 \implies f^B{}_{ij} = 0\,. \tag{6.23}$$

The fact that $\mathfrak{a}$ is an ideal implies

$$[T_i, T_\alpha] \in \mathfrak{a} \implies f^\beta{}_{i\alpha} = 0\,. \tag{6.24}$$

In other words, given any i and α, only $f^j{}_{i\alpha}$ can be nonzero. If we look at the structure constant in (6.23) restricting the superscript to the basis vectors not in $\mathfrak{a}$, i.e. $B = \beta$, then the combination with (6.24) implies

$$f^\beta{}_{iB} = 0 = f^\beta{}_{Bi}\,. \tag{6.25}$$

Now we look at the Killing form for $A \in \mathfrak{a}$ and $X \in \mathfrak{g}$: $\kappa(X, A) = \kappa_{Bi} X^B A^i$ with

$$\begin{aligned}\kappa_{Bi} &= f^C{}_{BD} f^D{}_{iC} \\ &= f^\alpha{}_{BD} f^D{}_{i\alpha} + f^j{}_{BD} f^D{}_{ij} \\ &= f^\alpha{}_{B\beta} f^\beta{}_{i\alpha} + f^\alpha{}_{Bj} f^j{}_{i\alpha} = 0\,.\end{aligned} \tag{6.26}$$

In going from the first line to the second, we split index $\{C\} = \{\alpha\} \cup \{j\}$. In the second line $f^D{}_{ij}$ vanishes by (6.23). We split index $\{D\} = \{\beta\} \cup \{j\}$ to arrive at the third line, where $f^\beta{}_{i\alpha} = 0$ by (6.24) and $f^\alpha{}_{Bj} = 0$ by (6.25). Therefore, we see that if $\mathfrak{g}$ is not semisimple, then $\kappa(X, A) = 0$ for $A \in \mathfrak{a}$ and all $X \in \mathfrak{g}$, so the Killing form must be degenerate.

Semisimplicity is discussed in detail in Fulton and Harris at level higher than we can reach.[22]

[22] W Fulton and J Harris. *Representation Theory: A First Course*. Springer-Verlag, 1991. ISBN 0-387-97527-6. URL https://link.springer.com/book/10.1007/978-1-4612-0979-9

Looking at $\mathfrak{su}(2)$, for example, the structure constants $f^c_{ab} = \epsilon_{abc}$. Therefore the coefficients of the Killing form are

$$\kappa_{ab} = f^c{}_{ad} f^d{}_{bc} = \epsilon_{adc}\epsilon_{bcd} = -2\delta_{ab}\,, \tag{6.27}$$

so $\kappa(X, Y) = -2X^a Y^b \delta_{ab}$ which is not generally zero. Since the Killing form is nondegenerate, we've proved that $\mathfrak{su}(2)$ is semisimple. (In fact, it is simple.)

Next, we state a few facts about the Killing form which we will not prove.

Proposition 6.16. If κ is nondegenerate, then it is invertible and, for κ_{ab} we can find a $(\kappa^{-1})^{bc}$ such that

$$\kappa_{ab}(\kappa^{-1})^{bc} = \delta_a{}^c\,. \tag{6.28}$$

Referring to groups:

Definition 6.17. A Lie group is **semisimple** if its Lie algebra is semisimple (see Definition 6.7).

Some authors also only use this term to describe connected Lie groups, so some care is needed.[23]

[23] A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL https://link.springer.com/book/10.1007/978-1-4757-2453-0

Proposition 6.18. If the Killing form of a real Lie algebra $\mathfrak{g} = L(G)$ is negative definite, G is compact and $\mathfrak{g}$ is said to be of **compact type**.

As an aside, we note the following proposition, but it our focus will be on semisimple groups.

Proposition 6.19. A compact group which is *not* semisimple has an algebra with negative *semidefinite* Killing form (but not negative definite).

There are noncompact groups which also have negative semidefinite Killing forms. (See Knapp[24] for more details.)

[24] A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL https://link.springer.com/book/10.1007/978-1-4757-2453-0

Proposition 6.20. Every semisimple, complex, finite-dimensional lie algebra $L(G)_{\mathbb{C}}$ has a real form with

$$\kappa_{ab} = -\kappa\delta_{ab}\,, \quad \kappa \in \mathbb{R}^+\,. \tag{6.29}$$

Propositions (6.16), (6.18), and (6.20) are the most relevant to the next few chapters.

Definition 6.21. Any basis for which $\kappa_{ab} \propto \delta_{ab}$ (if such exists) is called an **adapted basis**.

In an adapted basis $\{ T_a \}$ we have the following

$$\begin{aligned} \kappa([T_a, T_c], T_b) &= f^d{}_{ac}\,\kappa(T_d, T_b) = -\kappa f^d{}_{ac}\delta_{db} = -\kappa f^b{}_{ac} \\ \kappa(T_a, [T_c, T_b]) &= f^d{}_{cb}\,\kappa(T_a, T_d) = -\kappa f^d{}_{cb}\delta_{ad} = -\kappa f^a{}_{cb}\,. \end{aligned} \tag{6.30}$$

By the invariance of the Killing form (6.17) the two expressions above are equal; therefore

$$f^b{}_{ac} = f^a{}_{cb} = -f^a{}_{bc}\,. \tag{6.31}$$

Comparing the first and last expressions above, we conclude that, in and adapted basis, the structure constants are antisymmetric under exchange of any pair of indices, including the upper index with a lower index. In this case it is common to write all the indices as either upper or lower indices, e.g. defining

$$f_{abc} := f^c{}_{ab} = f^a{}_{bc} = f^b{}_{ca}\,. \tag{6.32}$$

6.3 Casimir elements

This section is a little outside the theme of this Chapter, but, now that we have introduced adapted bases, this is a good opportunity to introduce a few concepts which will be useful in the Standard Model and Advanced Quantum Field Theory courses.

Given an Lie algebra $\mathfrak{g}$, the objects we will define below, namely Casimir elements, examples of which we saw in § 5.3, are not generally elements of $\mathfrak{g}$ itself, but of the universal enveloping algebra

(UEA) of the Lie algebra. The **universal enveloping algebra** of $\mathfrak{g}$ is the formal span of

2025-11-16: Clarification that the set starts with the field $\mathbb{F}$ of the algebra

$$\{\, \mathbb{F},\ \mathfrak{g},\ \mathfrak{g}\otimes\mathfrak{g},\ \mathfrak{g}\otimes\mathfrak{g}\otimes\mathfrak{g},\ \ldots \} \tag{6.33}$$

subject to the following multiplication rule for $X, Y \in \mathfrak{g}$

$$XY - YX = [X, Y] \tag{6.34}$$

where the righthand side is the Lie bracket. For matrix Lie algebras, this multiplication is just matrix multiplication. That is, generally $XY \notin \mathfrak{g}$, instead $XY = X^a Y^b T_a T_b \in \mathfrak{g}\otimes\mathfrak{g}$. This leads to new identities such as

$$[X, YZ] = [X, Y]Z + Y[X, Z]\,, \tag{6.35}$$

which will be familiar from commutator relations in Quantum Mechanics courses.

A **Casimir element** (or **Casimir operator** or **Casimir invariant**) is a polynomial function of elements of a Lie algebra $\mathfrak{g}$ which commutes with all elements of $\mathfrak{g}$. In what follows, let $\mathfrak{g}$ be the real Lie algebra of a simple, compact group. In the case we can choose an adapted basis such that the Killing form can be written

$$\kappa_{ab} = -\delta_{ab} \tag{6.36}$$

where we have assumed a normalization of the basis $\{\, T_a \,\}$ such that the coefficient $\kappa = 1$ in (6.29).

The **universal, quadratic Casimir element** in the UEA of a Lie algebra is

$$C := T_b T_b\,, \quad \text{(implied sum over } b)\,. \tag{6.37}$$

We can see that C commutes with any basis element T_a

$$\begin{aligned}[T_a, C] &= [T_a, T_b T_b] \\ &= T_b[T_a, T_b] + [T_a, T_b]T_b \\ &= f_{abc} T_b T_c + f_{abc} T_c T_b \\ &= f_{abc} T_b T_c - f_{acb} T_c T_b \\ &= 0\,. \end{aligned} \tag{6.38}$$

Therefore, C commutes with all $X = X^a T_a \in \mathfrak{g}$.

Some Lie algebras have higher order Casimirs (i.e. higher order polynomial functions of elements of $\mathfrak{g}$), but these are specific to the algebra so are not universal.

Consider the quadratic Casimir in a representation d of $L(G) = \mathfrak{g}$

$$C_d := \sum_a d(T_a) d(T_a)\,. \tag{6.39}$$

As above,

$$[d(X), C_d] = 0 \ \ \forall X \in \mathfrak{g}\,. \tag{6.40}$$

Let $D = \exp d$ be the group representation corresponding to d. If d is irreducible then so is D, and by Schur's lemma (see the Examples sheet)

$$C_d D(g) = D(g) C_d \;\; \forall g \in G \quad \Longrightarrow \quad C_d = c_d I \;\; \text{with } c_d \in \mathbb{R} \qquad (6.41)$$

that is, the quadratic Casimir is proportional to the identity.

Take for example $\mathfrak{su}(2)$. In an adapted basis normalized according to (6.36)

$$T_a = -\frac{i}{2\sqrt{2}}\sigma_a \,. \qquad (6.42)$$

Consider the fundamental representation, where $j = \frac{1}{2}$ ($\Lambda = 1$). Then

$$C_{\frac{1}{2}} = -\frac{1}{8}\sigma_a \sigma_a = -\frac{3}{8} I \,. \qquad (6.43)$$

Thus, $c = -3/8$ is the quadratic Casimir of the fundamental representation. For general j where the generators are normalized to be $\frac{i}{\sqrt{2}} J_a$

$$C_j = -\frac{1}{2} J_a J_a = -\frac{1}{2} j(j+1) I \,. \qquad (6.44)$$

6.4 *Cartan–Weyl basis*

We begin with some definitions.

Definition 6.22. Let $\mathfrak{g}$ be a Lie algebra. Element $X \in \mathfrak{g}$ is **ad-diagonalizable** if map $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable. In the Cartan–Weyl basis (4.18) of $\mathfrak{su}(2)_\mathbb{C}$ for example, H is ad-diagonalizable, while $E_\pm$ are not.

Definition 6.23. A **maximal** Abelian subalgebra $\mathfrak{h}$ is not contained in any larger Abelian subalgebra.

Definition 6.24. If $\mathfrak{g}$ is a complex, semisimple Lie algebra, then a **Cartan subalgebra** (CSA) $\mathfrak{h}$ of $\mathfrak{g}$ is a complex space such that

1. $\mathfrak{h}$ is Abelian: for all H_1 and $H_2 \in \mathfrak{h}$, their Lie bracket vanishes: $[H_1, H_2] = 0$.
2. $\mathfrak{h}$ is a maximal Abelian subalgebra: For any $X \in g$, if $[H, X] = 0$ for all $H \in \mathfrak{h}$, then $X \in \mathfrak{h}$.
3. All $H \in \mathfrak{h}$ are ad-diagonalizable.

It can be proved that semisimple Lie algebras have Cartan subalgebras. In fact, one can extend the definition of Cartan subalgebras to all Lie algebras, but then the CSA is not necessarily Abelian. In this Chapter and next, our interest is restricted to semisimple Lie algebras and Abelian CSAs. (The interested reader is referred to Chapter II of Knapp.[25])

[25] A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL https://link.springer.com/book/10.1007/978-1-4757-2453-0

Definition 6.25. The dimension of an algebra's CSA is called the **rank** of the Lie algebra: $\operatorname{rank}\mathfrak{g} = \dim\mathfrak{h}$.

As an example, take $\mathfrak{su}(2)_{\mathbb{C}}$. The $H = \sigma_3$ introduced in (4.18) gives a 1-dimensional CSA: $\mathfrak{h} = \operatorname{span}_{\mathbb{C}}\{H\}$. We could equally well have chosen $H = \sigma_1$ or $H = \sigma_2$. In any case $\dim\mathfrak{h} = 1$. Therefore, $\operatorname{rank}(\mathfrak{su}(2)_{\mathbb{C}}) = 1$.

Let us extend our example to $\mathfrak{su}(n)_{\mathbb{C}}$, the algebra of traceless complex $n \times n$ matrices, that is,

$$\mathfrak{su}(n)_{\mathbb{C}} = \{ X \in \mathrm{Mat}_n(\mathbb{C}) \mid \operatorname{Tr} X = 0 \} . \tag{6.45}$$

We know that diagonal matrices commute with each other, so let us take the set of diagonal, traceless matrices to be our CSA $\mathfrak{h}$. You can convince yourself that if a matrix X satisfies $[H, X] = 0$ for all $H \in \mathfrak{h}$, then X must also be diagonal. We can choose a basis for $\mathfrak{h}$ as

$$(H_i)_{\alpha\beta} = \delta_{\alpha i}\delta_{\beta i} - \delta_{\alpha(i+1)}\delta_{\beta(i+1)} \tag{6.46}$$

where $i = 1, \ldots, n-1$ and $\alpha, \beta = 1, \ldots n$. We note the rank of $\mathfrak{su}(n)_{\mathbb{C}}$ is then $n-1$.

Eigenvectors of ad_H

As we did for $\mathfrak{su}(2)$ in (4.21), we wish to study the eigenvectors of the adjoint representation of $\mathfrak{h}$. For any $H, H' \in \mathfrak{h}$

$$\mathrm{ad}_H H' = [H, H'] = 0 \tag{6.47}$$

which implies

$$[\mathrm{ad}_H, \mathrm{ad}_{H'}] = 0 . \tag{6.48}$$

All the adjoint maps communte and are therefore simultaneously diagonalizable. By the spectral decomposition theorem, this means that the Lie algebra $\mathfrak{g}$ is spanned by the set of simultaneous eigenvectors of the $\{\mathrm{ad}_H\}$. The elements of the CSA, $\mathfrak{h}$, are the zero-eigenvectors. Within this subspace, let us choose a basis

$$\{ H_i \mid i = 1, \ldots, r \} \tag{6.49}$$

where $r = \dim\mathfrak{h} = \operatorname{rank}\mathfrak{g}$. Since the CSA is maximal, no other eigenvectors of $\{\mathrm{ad}_H\}$ can have eigenvalue zero. Denote the rest of the eigenvectors E_α, where the subscript α labels each eigenvector and is called a **root**. The corresponding eigenvalues $\{ \alpha_i \mid i = 1, \ldots, r \}$ will generally depend on our choice of basis (6.49)

$$\mathrm{ad}_{H_i} E_\alpha = [H_i, E_\alpha] = \alpha_i E_\alpha \tag{6.50}$$

where $\alpha_i \in \mathbb{C}$.

Claim 6.26. The nonzero simultaneous eigenvectors of $\{\mathrm{ad}_H\}$ are nondegenerate, unique up to a normalization.

The proof is beyond the scope of this course, but the interested reader is referred to Sections 6.4 (1st ed)/7.3 (2nd ed) of Hall or Proposition 2.21 of Knapp.[26] We will call these eigenvectors **ladder** or **step operators**. The collection of possible roots α is called the **root set** and denoted Φ.

[26] B C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015. ISBN 978-3319134666. URL https://link.springer.com/book/10.1007/978-0-387-21554-9; and A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL https://link.springer.com/book/10.1007/978-1-4757-2453-0

We wish to show that the roots are vectors in the dual vector space $\mathfrak{h}^*$ (see § 10.4). Any element $H \in \mathfrak{h}$ can be written as a linear combination of basis vectors (6.49) as

$$H = \rho^i H_i\,, \quad \text{with } \rho^i \in \mathbb{C}\,. \tag{6.51}$$

Then the action of this general ad_H gives

$$[H, E_\alpha] = \rho^i [H_i, E_\alpha] = \rho^i \alpha_i E_\alpha =: \alpha(H) E_\alpha \tag{6.52}$$

defining $\alpha(H) := \rho^i \alpha_i \in \mathbb{C}$. We see that $\alpha : \mathfrak{h} \to \mathbb{C}$, but is it linear? Yes. Here is the proof. Take any H and $H' \in \mathfrak{h}$. Then

2025-12-05 Typo fixed.

$$\begin{aligned}\alpha(H + H')E_\alpha &= [H + H', E_\alpha] \\ &= [H, E_\alpha] + [H', E_\alpha] \\ &= \big(\alpha(H) + \alpha(H')\big)\, E_\alpha\,.\end{aligned} \tag{6.53}$$

Therefore the roots $\alpha \in \Phi$ are vectors in the dual vector space $\mathfrak{h}^*$. In fact, they are *nonzero* vectors, otherwise the corresponding E_α would have to be included in the CSA. We will have a lot more to say about $\mathfrak{h}^*$ in the next section.

Definition 6.27. The **Cartan–Weyl basis** for $\mathfrak{g}$ is given by

$$\{\, H_i \mid i = 1, \ldots, r \,\} \cup \{\, E_\alpha \mid \alpha \in \Phi \,\}\,. \tag{6.54}$$

Given Claim 6.26, we must have $\dim \mathfrak{g} - \dim \mathfrak{h}$ roots in order for (6.54) to be a basis. Noting that $\dim \mathfrak{h}^* = \dim \mathfrak{h}$, we will soon see that only a subset of roots are linearly independent vectors in $\mathfrak{h}^*$.

We now wish to determine the Lie brackets for the Cartan–Weyl basis elements and determine some properties of the roots. In what follows, we establish the Killing form (6.7) as a means to define an inner product in the dual space where roots are vectors. We also make rely on Cartan's theorem that the Killing form is nondegenerate for semisimple Lie algebras.

First let us state and prove four lemmas involving the Killing form.

Lemma 6.28. $\kappa(H, E_\alpha) = 0$ for all $H \in \mathfrak{h}$ and all $\alpha \in \Phi$.

Since α is a nonzero vector in $\mathfrak{h}^*$, there exists an $H' \in \mathfrak{h}$ such that $\alpha(H') \neq 0$. Then

$$\begin{aligned}\alpha(H')\kappa(H, E_\alpha) &= \kappa(H, \alpha(H')E_\alpha) && \text{by linearity} \\ &= \kappa(H, [H', E_\alpha]) && \text{by (6.52)} \\ &= \kappa([H, H'], E_\alpha) && \text{by invariance (6.17)} \\ &= \kappa(0, E_\alpha) = 0\,.\end{aligned} \tag{6.55}$$

Since $\alpha(H') \neq 0$, we have $\kappa(H, E_\alpha) = 0$.

Lemma 6.29. $\kappa(E_\alpha, E_\beta) = 0$ for all $\alpha, \beta \in \Phi$ such that $\alpha + \beta \neq 0$.

For all $H \in \mathfrak{h}$, we have

$$\begin{aligned}(\alpha(H) + \beta(H))\,\kappa(E_\alpha, E_\beta) &= \kappa([H, E_\alpha], E_\beta) + \kappa(E_\alpha, [H, E_\beta]) \\ &= 0 \quad \text{by invariance.}\end{aligned} \tag{6.56}$$

Lemma 6.30. If $\alpha \in \Phi$ then $-\alpha \in \Phi$ and $\kappa(E_\alpha, E_{-\alpha}) \neq 0$.

Lemmas 6.28 and 6.29 together give us

$$\begin{aligned}\kappa(E_\alpha, H) &= 0\,, \quad \forall H \in \mathfrak{h} \\ \kappa(E_\alpha, E_\beta) &= 0\,, \quad \forall \beta \in \Phi \text{ with } \beta \neq -\alpha.\end{aligned} \tag{6.57}$$

We know, however, that κ cannot be degenerate – otherwise $\mathfrak{g}$ would not be semisimple – so there must be some $X \in \mathfrak{g}$ such that $\kappa(E_\alpha, X) \neq 0$. By process of elimination, the only possibility is that $-\alpha \in \Phi$ and that $\kappa(E_\alpha, E_{-\alpha}) \neq 0$.

Lemma 6.31. For all $H \in \mathfrak{h}$, there exists an $H' \in \mathfrak{h}$ such that $\kappa(H, H') \neq 0$.

Suppose the contrapositive, that $\exists H \in \mathfrak{h}$ such that $\kappa(H, H') = 0, \forall H' \in \mathfrak{h}$. From (6.28), we also have $\kappa(H, E_\alpha) = 0$ for all $\alpha \in \Phi$. Then the Killing form $\kappa(H, X) = 0$ for all $X \in \mathfrak{g}$ and the κ would be degenerate. However, this contradicts our assertion that we are working with a semisimple Lie algebra $\mathfrak{g}$.

The consequence of Lemma 6.31, of nondegeneracy really, is that κ is nondegenerate and can be inverted within $\mathfrak{h}$. Using coordinates ρ^i and basis vectors H_i as before, we can write

$$\kappa(H, H') = \kappa(\rho^i H_i, \rho'^j H_j) = \kappa_{ij}\rho^i \rho'^j\,. \tag{6.58}$$

Then we can find a matrix κ^{-1} which has components satisfying

$$(\kappa^{-1})^{ik}\kappa_{kj} = \delta^i{}_j\,. \tag{6.59}$$

This can be used to give a nondegenerate inner product on $\mathfrak{h}^*$:

$$(\alpha, \beta) := (\kappa^{-1})^{ij}\alpha_i\beta_j\,. \tag{6.60}$$

We might also write the inner product in the usual way

$$(\alpha, \beta) = \alpha_i\beta^i = \alpha^j\beta_j \tag{6.61}$$

where we define what we mean with an upper index for vectors in $\mathfrak{h}^*$: $\alpha^i := (\kappa^{-1})^{ij}\alpha_j$ and $\beta^i := (\kappa^{-1})^{ij}\beta_j$. We defer until the next section proof that this is indeed an inner product. If you like, you can just think of (6.60) as a symmetric, bilinear form until we get to the end of § 6.5.

To recap, so far we have established that

$$\begin{aligned}[H_i, H_j] &= 0 \\ [H_i, E_\alpha] &= \alpha_i E_\alpha\,.\end{aligned}$$

We still need to find an expression for $[E_\alpha, E_\beta]$.[27]

For any $H \in \mathfrak{h}$ and $\alpha, \beta \in \Phi$ we have

$$\begin{aligned} \mathrm{ad}_H[E_\alpha, E_\beta] &= [H, [E_\alpha, E_\beta]] \\ &= -[E_\beta, [H, E_\alpha]] - [E_\alpha, [E_\beta, H]] \quad \text{by Jacobi (2.35)} \\ &= (\alpha(H) + \beta(H))[E_\alpha, E_\beta]\,. \end{aligned} \tag{6.62}$$

Eq. (6.62) says that $\alpha(H) + \beta(H)$ is an eigenvalue of ad_H with eigenvector $[E_\alpha, E_\beta]$. If $\alpha + \beta \neq 0$ and $\alpha + \beta \in \Phi$, then that eigenvector must be proportional to $E_{\alpha+\beta}$. On the other hand, if $\alpha + \beta \notin \Phi$, then $[E_\alpha, E_\beta] = 0$. Note that (6.62) implies that 2α is not a root, since $[E_\alpha, E_\alpha] = 0$.

If $\alpha + \beta = 0$, then (6.62) becomes $\mathrm{ad}_H[E_\alpha, E_{-\alpha}] = 0$, which says that $[E_\alpha, E_{-\alpha}] \in \mathfrak{h}$. We can use the Killing form to write this element of $\mathfrak{h}$:

$$\begin{aligned} \kappa([E_\alpha, E_{-\alpha}], H) &= \kappa(E_\alpha, [E_{-\alpha}, H]) \\ &= \alpha(H)\kappa(E_\alpha, E_{-\alpha})\,. \end{aligned} \tag{6.63}$$

Define a normalized element of $\mathfrak{h}$

$$H_\alpha = \frac{[E_\alpha, E_{-\alpha}]}{\kappa(E_\alpha, E_{-\alpha})} \tag{6.64}$$

such that $\kappa(H_\alpha, H) = \alpha(H)$. In components $H_\alpha = \rho^i_\alpha H_i$ and $H = \rho^i H_i$, so

$$\kappa_{ij}\rho^i_\alpha \rho^j = \alpha_i \rho^i\,. \tag{6.65}$$

The coefficients ρ^i are arbitrary (since the statements above are for any $H \in \mathfrak{h}$) so

$$\begin{aligned} \rho^i_\alpha &= (\kappa^{-1})^{ij}\alpha_j \\ \implies H_\alpha &= (\kappa^{-1})^{ij}\alpha_j H_i\,. \end{aligned} \tag{6.66}$$

Summarizing the above, we finally have

$$[H_i, H_j] = 0 \tag{6.67}$$

$$[H_i, E_\alpha] = \alpha_i E_\alpha \tag{6.68}$$

$$[E_\alpha, E_\beta] = \begin{cases} N_{\alpha,\beta}\, E_{\alpha+\beta} & \alpha + \beta \in \Phi \\ \kappa(E_\alpha, E_\beta)\, H_\alpha & \alpha + \beta = 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.69}$$

where $N_{\alpha,\beta}$ is a normalization factor.

Note that there are special elements in $\mathfrak{h}$, the $\{ H_\alpha \}$, associated with the roots $\{ \alpha \}$. These have Lie brackets

$$\begin{aligned} [H_\alpha, E_\beta] &= (\kappa^{-1})^{ij}\alpha_j [H_i, E_\beta] \\ &= (\kappa^{-1})^{ij}\alpha_j \beta_i E_\beta \\ &= (\alpha, \beta) E_\beta \end{aligned} \tag{6.70}$$

using (6.60).

[27] Just as a reminder, for $\mathfrak{su}(2)_\mathbb{C}$ we had $[H, E_\pm] = \pm 2E_\pm$ and $[E_+, E_-] = H$ (4.19).

$\mathfrak{sl}(2,\mathbb{C})$ subalgebras

Here we show that there is an $\mathfrak{sl}(2,\mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}$ subalgebra associated with each root α. It is convenient to choose another normalization such that

$$\begin{aligned} h_\alpha &:= \frac{2}{(\alpha,\alpha)} H_\alpha \\ e_\alpha &:= \sqrt{\frac{2}{(\alpha,\alpha)\,\kappa(E_\alpha,E_{-\alpha})}}\, E_\alpha\,. \end{aligned} \tag{6.71}$$

The Lie brackets (6.69) in basis (6.71) are then

$$\begin{aligned} [h_\alpha, h_\beta] &= 0 \\ [h_\alpha, e_\beta] &= \frac{2(\alpha,\beta)}{(\alpha,\alpha)} e_\beta \\ [e_\alpha, e_\beta] &= \begin{cases} n_{\alpha,\beta}\, e_{\alpha+\beta} & \alpha+\beta \in \Phi \\ h_\alpha & \alpha+\beta = 0 \\ 0 & \text{otherwise} \end{cases} . \end{aligned} \tag{6.72}$$

Note that for each $\alpha \in \Phi$, there is an $\mathfrak{su}(2)_{\mathbb{C}}$ (equivialently, $\mathfrak{sl}(2,\mathbb{C})$) subalgebra with basis $\{ h_\alpha, e_\alpha, e_{-\alpha} \}$ where

$$\begin{aligned} [h_\alpha, e_{\pm\alpha}] &= \pm 2 e_{\pm\alpha} \\ [e_\alpha, e_{-\alpha}] &= h_\alpha\,. \end{aligned} \tag{6.73}$$

We will write the subalgebra corresponding to root α as $\mathfrak{sl}(2)_\alpha$. These will help us in the next section.

6.5 *Geometry of roots*

The previous section has introduced the roots as the nonzero eigenvalues of the maps ad_H, along with the corresponding vector space $\mathfrak{h}^*$. In this section we develop a geometry useful for studying the roots of a particular Lie algebra. This geometry will become useful in classifying the possible representations of a Lie algebra.

There are several steps leading to the geometry. In brief, our plan is to show the following

(i) The inner products of roots, (α,β), defined in (6.60), are real.

(ii) $\mathfrak{h}^*$ is spanned by the root set Φ.

(iii) There is a real vector space $\mathfrak{h}^*_{\mathbb{R}} \subseteq \mathfrak{h}^*$ spanned by Φ, and $\mathfrak{h}^*_{\mathbb{R}}$ contains all $\alpha \in \Phi$.

(iv) $(\alpha,\alpha) \geqslant 0$, so a length can be defined as $|\alpha| = \sqrt{(\alpha,\alpha)}$.

(v) We can determine lengths and angles between roots.

We begin by defining a root string.

Definition 6.32. Let $\alpha \in \Phi$ and let $\beta \in \Phi$ with $\beta \neq \alpha$. The α-**root string** (or α-string, for short) passing through β is the set

$$S_{\alpha,\beta} = \{ \beta + \rho\alpha \in \Phi \mid \rho \in [n_-, n_+] \text{ and } n_-, n_+ \in \mathbb{Z} \}\,. \tag{6.74}$$

This builds on the final equation in (6.72), which can be iterated to read $[e_\alpha, e_{\alpha+\beta}] = n_{\alpha,\alpha+\beta}\, e_{2\alpha+\beta}$ if $2\alpha + \beta \in \Phi$, and so on. Note that allowing $\beta = \alpha$ above would give a root strong consisting only of α. As can be seen from $[E_\alpha, E_\alpha] = 0$ and (6.69), $2\alpha \notin \Phi$ and, of course, nor is 0.

Claim 6.33. We have a "quantization condition" for the inner product of roots which states

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = -(n_+ + n_-) \in \mathbb{Z}\,. \tag{6.75}$$

To prove this, consider the vector space

$$V_{\alpha,\beta} = \operatorname{span}_{\mathbb{C}} \left\{ e_{\beta+\rho\alpha} \;\middle|\; \beta + \rho\alpha \in S_{\alpha,\beta} \right\} \tag{6.76}$$

where $\beta \in \Phi$ and $\beta \neq \pm\alpha$. The action of $\mathfrak{sl}(2)_\alpha$ on $V_{\alpha,\beta}$ is given by

$$\begin{aligned} \mathrm{ad}_{h_\alpha} e_{\beta+\rho\alpha} = [h_\alpha, e_{\beta+\rho\alpha}] &= \frac{2(\alpha, \beta+\rho\alpha)}{(\alpha,\alpha)}\, e_{\beta+\rho\alpha} \\ &= \left[\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho\right] e_{\beta+\rho\alpha} \end{aligned} \tag{6.77}$$

and

$$\mathrm{ad}_{e_{\pm\alpha}} e_{\beta+\rho\alpha} = [e_{\pm\alpha}, e_{\beta+\rho\alpha}] \propto \begin{cases} e_{\beta+(\rho\pm1)\alpha} & \beta + (\rho\pm1)\alpha \in \Phi \\ 0 & \text{otherwise} \end{cases}\,. \tag{6.78}$$

Hence $V_{\alpha,\beta}$ is closed under $\mathfrak{sl}(2)_\alpha$ and is the representation space of some $\mathfrak{sl}(2,\mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}$ representation Λ. The weight set of this representation

$$S_\Lambda = \left\{ \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho \;\middle|\; \beta + \rho\alpha \in \Phi \right\} \tag{6.79}$$

must be equivalent to $\{ -\Lambda, -\Lambda+2, \dots, \Lambda-2, \Lambda \}$. This implies

$$\begin{aligned} \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_- &= -\Lambda & \qquad \text{for } \rho = n_- \\ \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_+ &= \Lambda & \qquad \text{for } \rho = n_+\,. \end{aligned} \tag{6.80}$$

Summing the two equations, we have proved our claim that

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = -n_+ - n_- \in \mathbb{Z}\,. \tag{6.81}$$

This identifies roots strings with irreps of $\mathfrak{sl}(2,\mathbb{C})$.

Definition 6.34. A root string $S_{\alpha,\beta}$ (6.74) has **length**

$$\ell_{\alpha,\beta} = n_+ - n_- + 1 \tag{6.82}$$

A root strong of length $\ell_{\alpha,\beta}$ corresponds to an $\mathfrak{sl}(2,\mathbb{C})$ representation of dimension $\ell_{\alpha,\beta}$. We will use this "quantization condition" (6.81) for the ratio $2(\alpha,\beta)/(\alpha,\alpha)$ soon.

One other useful identity involving the inner product is the following

Claim 6.35. The inner product of roots $\alpha, \beta \in \Phi$ can be expressed

$$(\alpha, \beta) = \sum_{\gamma \in \Phi} (\alpha, \gamma)(\gamma, \beta)\,. \tag{6.83}$$

This follows from the Killing form (6.7). In the Cartan–Weyl basis we have

$$[H_i, E_\gamma] = \gamma_i E_\gamma \tag{6.84}$$

for all $i = 1, \ldots, r$ and $\gamma \in \Phi$. Using the fact that the $\{\, \mathrm{ad}_{H_i} \,\}$ are diagonal in the Cartan–Weyl basis, with diagonal entries either 0 or $\{\, \gamma_i \,\}$, the Killing form of the CSA basis elements can be written as[28]

[28] For simplicity we set the normalization constant $\mathcal{N} = 1$ in what follows.

$$\begin{aligned} \kappa_{ij} = \kappa(H_i, H_j) &= \mathrm{Tr}(\mathrm{ad}_{H_i} \circ \mathrm{ad}_{H_j}) \\ &= \sum_{\gamma \in \Phi} \gamma_i \gamma_j \end{aligned} \tag{6.85}$$

where $\mathcal{N}$ is some real normalization factor. Then we can write

$$\begin{aligned} (\alpha, \beta) = \alpha_i \beta_j (\kappa^{-1})^{ij} &= \alpha^i \beta^j \kappa_{ij} \\ &= \sum_{\gamma \in \Phi} \alpha^i \gamma_i \gamma_j \beta^j \\ &= \sum_{\gamma \in \Phi} (\alpha, \gamma)(\gamma, \beta)\,, \end{aligned} \tag{6.86}$$

where we define what we mean with an upper index for vectors in $\mathfrak{h}^*$: $\alpha^i := (\kappa^{-1})^{ij} \alpha_j$ and $\beta^j := (\kappa^{-1})^{jk} \beta_k$.

Now we are ready to justify the claims enumerated at the start of this section

Claim 6.36. Inner products of roots, (α, β), are real.

Dividing both sides of (6.83) by $(\alpha, \alpha)(\beta, \beta)/4$ we have

$$\frac{2}{(\beta, \beta)} \underbrace{\frac{2(\alpha, \beta)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} = \sum_{\gamma \in \Phi} \underbrace{\frac{2(\alpha, \gamma)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} \underbrace{\frac{2(\beta, \gamma)}{(\beta, \beta)}}_{\in \mathbb{Z}}\,. \tag{6.87}$$

The three terms indicated above are integers by our quantization condition (6.81). Either $(\alpha, \beta) = 0$ (and hence is real), or the equation above implies that $(\beta, \beta) \in \mathbb{R} \setminus \{\, 0 \,\}$ and $(\alpha, \beta) \in \mathbb{R}$.

Claim 6.37. $\mathfrak{h}^*$ is spanned by the root set Φ.

We know the roots are elements of the dual vector space $\mathfrak{h}^*$. Generally, there are more roots than the dimension of $\mathfrak{h}^*$: $|\Phi| \geqslant \dim \mathfrak{h}^*$. In fact the roots span $\mathfrak{h}^*$. We can prove this by contradiction.

Suppose the roots do not span $\mathfrak{h}^*$. In that case, there must be an orthogonal subspace and there must exist a vector $\lambda \in \mathfrak{h}^*$ such that $(\lambda, \alpha) = 0$ for all $\alpha \in \Phi$. The corresponding vector in $\mathfrak{h}$ is

$$H_\lambda = \lambda^i H_i \in \mathfrak{h}\,. \tag{6.88}$$

Since $\mathfrak{h}$ is the Cartan subalgebra,

$$[H_\lambda, H] = 0 \;\; \forall H \in \mathfrak{h}\,. \tag{6.89}$$

We also know that

$$\begin{aligned}[H_\lambda, E_\alpha] &= \lambda^i [H_i, E_\alpha] \\ &= \lambda^i \alpha_i E_\alpha \\ &= (\lambda, \alpha) E_\alpha = 0\,. \end{aligned} \tag{6.90}$$

This H_λ commutes with all $X \in \mathfrak{g}$, implying that $\mathrm{span}_\mathbb{C} \{ H_\lambda \}$ is a nontrivial (Abelian) ideal and consequently $\mathfrak{g}$ is not semisimple. This contradicts our assumptions. Therefore the roots do span $\mathfrak{h}^*$.

Let us denote a specific basis of $\mathfrak{h}^*$ by

$$\mathfrak{h}^* = \mathrm{span}_\mathbb{C} \left\{ \alpha_{(i)} \,\middle|\, \alpha_{(i)} \in \Phi, i = 1, \dots, r \right\}. \tag{6.91}$$

Note that we use parentheses in the subscripts to distinguish basis elements from other vectors in $\mathfrak{h}^*$. In § 6.6 we will introduce the simple roots as such a basis, but for the time being, we do not need any special details of the basis.

Claim 6.38. Let $\mathfrak{h}^*_\mathbb{R} \subseteq \mathfrak{h}^*$ be the real vector space

$$\mathfrak{h}^*_\mathbb{R} = \mathrm{span}_\mathbb{R} \left\{ \alpha_{(i)} \right\}. \tag{6.92}$$

Then $\mathfrak{h}^*_\mathbb{R}$ contains all roots.

To prove this, take some $\beta \in \Phi$ and write it as a linear combination of basis roots

$$\beta = c^i \alpha_{(i)}\,. \tag{6.93}$$

Are the coefficients c^i real? They could be complex if we chose any vector in $\mathfrak{h}^*$, but we are insisting that β be a root. Take the inner product with any $\alpha_{(j)}$

$$(\beta, \alpha_{(j)}) = c^i \left(\alpha_{(i)}, \alpha_{(j)}\right). \tag{6.94}$$

By Claim 6.36 both inner products are real, and the equations are nondegenerate since the $\{ \alpha_{(i)} \}$ form a basis, therefore we have a system of linear equations which we can solve to find the c^i, which we see to be real.

Claim 6.39. For all $\lambda \in \mathfrak{h}^*_\mathbb{R}$

$$(\lambda, \lambda) \geqslant 0 \tag{6.95}$$

with $(\lambda, \lambda) \iff \lambda = 0$.

As we did with roots (6.83), we can write this inner product as

$$(\lambda, \lambda) = \sum_{\gamma \in \Phi} (\lambda, \gamma)^2 \geqslant 0\,. \tag{6.96}$$

Furthermore, $(\lambda, \lambda) = 0$ if and only if $(\lambda, \gamma) = 0$ for all $\gamma \in \Phi$. Since the roots span $\mathfrak{h}^*_\mathbb{R}$, this can only happen for $\lambda = 0$.

We can now establish a geometry of the roots.

Definition 6.40. Define $|\alpha| := \sqrt{(\alpha,\alpha)}$ as the **norm** or **length** of a root $\alpha \in \Phi$ (or any vector in $\mathfrak{h}^*_{\mathbb{R}}$).

Claim 6.41. For any pair of roots, α and β, there is an angle θ such that

$$(\alpha,\beta) = |\alpha||\beta|\cos\theta\,. \tag{6.97}$$

This follows from the Cauchy–Schwartz inequality $(\alpha,\beta) \leq |\alpha||\beta|$ phrased in terms of our inner product on $\mathfrak{h}^*_{\mathbb{R}}$.

The quantization condition (6.81) implies

$$\begin{aligned} \frac{2(\alpha,\beta)}{(\alpha,\alpha)} &= 2\frac{|\beta|}{|\alpha|}\cos\theta \in \mathbb{Z} \\ \frac{2(\alpha,\beta)}{(\beta,\beta)} &= 2\frac{|\alpha|}{|\beta|}\cos\theta \in \mathbb{Z}\,. \end{aligned} \tag{6.98}$$

Multiplying these together gives

$$4\cos^2\theta \in \mathbb{Z}\,. \tag{6.99}$$

Since cosine is bounded in $[-1,1]$, the angle between roots is restricted to be in the set

$$|\theta| = \left\{ 0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}, \pi \right\}\,. \tag{6.100}$$

6.6 *Simple roots*

We have more roots than are needed to span $\mathfrak{h}^*_{\mathbb{R}}$. We can pick a hyperplane of dimension $r-1$ to divide the roots into two halves, one we arbitrarily call the positive half, the other negative. The hyperplane should go through 0 and not contain any roots in it. This is possible since there are a finite number of roots. If a root α is in the positive half of $\mathfrak{h}^*_{\mathbb{R}}$, we say α is a positive root and $-\alpha$ is a negative root. All roots are then divided into equal size root sets

$$\Phi = \Phi_+ \cup \Phi_-\,. \tag{6.101}$$

We have two properties (one stated above)

$$\alpha \in \Phi_+ \implies -\alpha \in \Phi_- \tag{6.102}$$

$$\alpha,\beta \in \Phi_+ \implies \alpha+\beta \in \Phi_+\,. \tag{6.103}$$

Definition 6.42. A **simple root** is a positive root which cannot be written as the sum of two positive roots. The set of simple roots is denoted Φ_S.

We will derive a few properties of the simple roots.

Claim 6.43. If α and β are simple roots, then $\alpha-\beta$ is not a root.

The proof is by contradiction. Suppose $\alpha-\beta \in \Phi_+$. Then

$$\alpha = (\alpha-\beta)+\beta \tag{6.104}$$

which implies α is not a simple root, as it is the sum of two positive roots. Otherwise suppose $\alpha - \beta \in \Phi_-$. Then $\beta - \alpha \in \Phi_+$ and

$$\beta = (\beta - \alpha) + \alpha \tag{6.105}$$

which contradicts $\beta \in \Phi_S$. Therefore $\alpha - \beta$ cannot be a root. This property will be useful when we write down the Lie brackets using simple roots.

Claim 6.44. For $\alpha, \beta \in \Phi_S$, the α-root string through β has length

$$\ell_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}^+ . \tag{6.106}$$

Since the difference of two simple roots is not a root (Claim 6.43), α and β being simple implies that $n_- = 0$ in (6.74). Then (6.81) implies

$$n_+ = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}_{\geqslant 0} . \tag{6.107}$$

We add 1 to obtain the length

$$\ell_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}^+ . \tag{6.108}$$

The length property will be useful in the next section, and it has a useful corollary, which is the next claim.

Claim 6.45. $(\alpha,\beta) \leqslant 0$ for $\alpha, \beta \in \Phi_S$ and $\beta \neq \alpha$. This follows immediately from Claim 6.44 and the fact that $(\alpha,\alpha) > 0$ (Claim 6.39).

Claim 6.46. Any positive root, $\beta \in \Phi_+$, can be written as a linear combination of simple roots with positive integer coefficients.

If $\beta \in \Phi_S$, we are done. If, on the other hand, $\beta = \beta_1 + \beta_2$, then we consider whether both β_1 and β_2 are simple. If so, we are done, otherwise we iterate until all the summands are simple.

The next three properties show that the simple roots form a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

Claim 6.47. All roots $\alpha \in \Phi$ can be written as a linear combination

$$\alpha = \sum_i p_i \alpha_{(i)} \tag{6.109}$$

with $p_i \in \mathbb{Z}$.

This follows immediately from Claim 6.46 for $\alpha \in \Phi_+$. If α is a negative root, then $-\alpha$ is a positive root and we apply Claim 6.46. Note that, by Claim 6.43, the nonzero coefficients p_i are either all positive, or all negative. Since $\mathfrak{h}^*_{\mathbb{R}}$ is spanned by the roots, we now see that the simple roots are sufficient to span the space.

Claim 6.48. The simple roots are linearly independent.

Combining Claims 6.37 and 6.47, all dual vectors $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$ can be written as

$$\lambda = \sum_i c_i \alpha_{(i)}, \quad \text{with } c_i \in \mathbb{R}\,. \tag{6.110}$$

We now prove the simple roots are linear independent by showing

$$\lambda = 0 \iff c_i = 0 \ \ \forall_i\,. \tag{6.111}$$

Let us assume that one or more $c_i \neq 0$ and separate the positive and negative coefficients into two different sets, with the corresponding indices belonging to either J_+ or J_-, where

2025-11-24: Clarified parts of the argument.

$$J_\pm := \{\, i \mid c \gtrless 0 \,\}\ . \tag{6.112}$$

Let us further define

$$\lambda_+ := \sum_{i\in J_+} c_i \alpha_{(i)} \quad \text{and}$$
$$\lambda_- := -\sum_{i\in J_-} c_i \alpha_{(i)} = \sum_{i\in J_-} b_i \alpha_{(i)}, \tag{6.113}$$

where we define positive coefficients $b_i = -c_i$ for $i \in J_-$. Then

$$\lambda = \lambda_+ - \lambda_- = \sum_{i\in J_+} c_i \alpha_{(i)} - \sum_{i\in J_-} b_i \alpha_{(i)}\,. \tag{6.114}$$

Using the inner product on $\mathfrak{h}^*_{\mathbb{R}}$ we have

$$(\lambda,\lambda) = (\lambda_+,\lambda_+) + (\lambda_-,\lambda_-) - 2(\lambda_+,\lambda_-) \tag{6.115}$$

Since λ_+ (λ_-) is the sum of simple roots with nonnegative coefficients, it is a nonzero vectors unless all of the c_i (b_i) are zero. Therefore at least one of (λ_+,λ_+) and (λ_-,λ_-) is positive, and we can write the strict inequality

$$(\lambda,\lambda) > -2(\lambda_+,\lambda_-) = -2\sum_{i\in J_+}\sum_{j\in J_-} c_i b_j (\alpha_{(i)},\alpha_{(j)})$$
$$(\lambda,\lambda) > 0 \tag{6.116}$$

since $c_i b_j > 0$ by construction and $(\alpha_{(i)},\alpha_{(j)}) \leqslant 0$ by Claim 6.45. Therefore, nonzero c_i implies $(\lambda,\lambda) > 0$. Conversely, having $(\lambda,\lambda) = 0$ implies that all the $c_i = 0$, running the argument the other way. Therefore, the only way $\lambda = 0$ is if all coefficients c_i vanish; consequently, the simple roots are linearly independent. Furthermore, since they also span the space, they form a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

Here is a more efficient proof. Suppose that $\lambda = 0$. Then (6.114) rearranges to

2025-11-26 Added a streamlined proof

$$\sum_{i\in J_+} c_i \alpha_{(i)} = \sum_{i\in J_-} b_i \alpha_{(i)} =: \mu \tag{6.117}$$

defining $\mu \in \mathfrak{h}^*_{\mathbb{R}}$. Following Claim 6.39 we know the inner product (μ,μ) is positive for nonzero μ, so

$$0 \leqslant (\mu,\mu) = \left(\sum_{i\in J_+} c_i \alpha_{(i)}, \sum_{j\in J_-} b_j \alpha_{(j)}\right)$$
$$= \sum_{i\in J_+}\sum_{j\in J_-} c_i b_j (\alpha_{(i)},\alpha_{(j)}) \leqslant 0\,. \tag{6.118}$$

In the last step we used that $(\alpha_{(i)}, \alpha_{(j)}) \leqslant 0$ for unequal simple roots (Claim 6.45) and that the c_i and b_j in (6.118) are both positive by construction. We conclude that $(\mu, \mu) = 0$ and hence $\mu = 0$, which can only occur if all coefficients $c_i = 0 = b_j$.

Claim 6.49. There are $r = \dim \mathfrak{h}^*_{\mathbb{R}}$ simple roots, i.e.

$$|\Phi_S| = r\,, \tag{6.119}$$

and they form a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

This follows immediately from Claims 6.47 and 6.48.

Using the properties above, one can construct the full root set from knowledge of the simple roots. On Examples Sheet problems, you will need to construct an algorithm by which you take pairs of roots and see if the corresponding root strings give any new roots.

6.7 *Classification*

In the previous section we showed that the simple roots form a basis for $\mathfrak{h}^*_{\mathbb{R}}$, so the next step is to write the basis for the Lie algebra $\mathfrak{g}$ using the simple roots. This basis for $\mathfrak{g}$ is called the **Chevalley basis**.

We define the $r \times r$ **Cartan matrix** A, where $r = \dim \mathfrak{h}^*_{\mathbb{R}}$, to have matrix elements

$$A_{ji} = \frac{2(\alpha_{(j)}, \alpha_{(i)})}{(\alpha_{(i)}, \alpha_{(i)})}\,. \tag{6.120}$$

Note that A is not generally symmetric. We know from (6.81) that $A_{ji} \in \mathbb{Z}$.

Now using (6.72) and focussing on the simple roots, we have

$$\begin{aligned}
[h_{\alpha(i)}, h_{\alpha(j)}] &= 0 \\
[h_{\alpha(i)}, e_{\pm\alpha(j)}] &= \pm A_{ji}\, e_{\pm\alpha(j)} \\
[e_{\alpha(i)}, e_{-\alpha(j)}] &= \delta_{ij}\, h_i\,,
\end{aligned} \tag{6.121}$$

with no sums over indices above. The first and second relations are straightforward transcriptions of (6.72). The third one is clear in the case where $i = j$. For $i \neq j$ we use the fact that $\alpha(i) - \alpha(j) \notin \Phi$.

From (6.72) we have

$$[e_{\alpha(i)}, e_{\alpha(j)}] = \mathrm{ad}_{e_{\alpha(i)}}(e_{\alpha(j)}) \propto e_{\alpha(i)+\alpha(j)} \tag{6.122}$$

as long as $\alpha(i) + \alpha(j) \in \Phi$. In this case, then $\alpha(i) + \alpha(j)$ is part of a root string which we can write without loss of generality as $n\alpha(i) + \alpha(j)$ whose length we know (6.106), and corresponds to $n = 0, \ldots, -A_{ji}$. (We know $n_- = 0$ by Claim 6.43, the difference of simple roots is not a root.) For $\ell_{ij} = 1 - A_{ji}$ applications of $\mathrm{ad}_{e_{\alpha(i)}}$, then the string ends, so we obtain the **Serre relation**

$$(\mathrm{ad}_{e_{\alpha(i)}})^{1-A_{ji}}(e_{\alpha(j)}) = 0\,. \tag{6.123}$$

The relations (6.121) and (6.123) completely characterize the Lie algebra, generating basis vectors, the r zero-eigenvectors $\{h_{\alpha(i)}\}$ and the $|\Phi|$ nonzero eigenvectors generated by applying $\mathrm{ad}_{e_{\alpha(i)}}$ to the $e_{\alpha(j)}$.

It can further be proven that any finite-dimensional, simple, complex Lie algebra is uniquely determined by its Cartan matrix. We proceed by classifying the possible Cartan matrices, and showing how to reconstruct the corresponding Lie algebras.

Claim 6.50. All diagonal entries, A_{ii} (no sum), are equal to 2.

This is evident from the definition (6.120).

Claim 6.51. $A_{ji} = 0 \iff A_{ij} = 0$

This follows from the definition (6.120) and the symmetry of the inner product.

Claim 6.52. All off-diagonal entries are non-positive: $A_{ji} \in \mathbb{Z}_{\leqslant 0}$ for all $i \neq j$.

This follows from the inner product relation $(\alpha_{(i)}, \alpha_{(j)}) \leqslant 0$ for $i \neq j$.

Claim 6.53. $\det A > 0$.

2025-11-26: Correction and elaboration from last year's version

Proof: [To avoid confusing the i indices we used to index an arbitrary basis $\{ H_i \}$ of $\mathfrak{h}$, we use a and b indices for the simple roots below.] Write $A = \kappa D$, where $\kappa^{ab} = \kappa(H_{\alpha_{(a)}}, H_{\alpha_{(b)}}) = (\alpha_{(a)}, \alpha_{(b)})$ [29] and D is the diagonal matrix given by

[29] Show this using (6.60) and (6.66).

$$D^a{}_b = \frac{2}{(\alpha_{(a)}, \alpha_{(b)})} \delta^a{}_b \,. \tag{6.124}$$

Clearly $\det D > 0$. From Claim 6.39, we know that the inner product on $\mathfrak{h}^*_{\mathbb{R}}$ is positive definite, i.e. for any nonnegative λ, [30]

$$(\lambda, \lambda) = \lambda_a (\kappa)^{ab} \lambda_b > 0 \,. \tag{6.125}$$

[30] Short proof from linear algebra: Let v_ρ be an eigenvector of κ with eigenvalue ρ. Then

$$0 < (v_\rho, v_\rho) = B_{ab} v^a_\rho v^b_\rho = \rho \sum_a (v^a_\rho)^2 \,.$$

This implies that all the eigenvalues $\rho > 0$, and hence $\det \kappa > 0$.

Therefore κ has a positive determinant and $\det A = \det \kappa \det D > 0$.

Proposition 6.54. For simple Lie algebras, A is irreducible, i.e. cannot be made block-triangular (or block-diagonal) by a permutation transformation of the form PAP^{-1}, where P is a permutation matrix.[31] If A is reducible, then the Lie algebra is semisimple, but not simple.

[31] This is the standard definition of (ir)reducibility for *integer*-valued matrices.

A complete proof is difficult and beyond this course. We will see examples below.

Claim 6.55. The product $A_{ij}A_{ji} \in \{ 0, 1, 2, 3 \}$ for $i \neq j$ (no sum).

Write $(\alpha_{(i)}, \alpha_{(j)}) = |\alpha_{(i)}||\alpha_{(j)}| \cos\phi_{ij}$, where ϕ_{ij} is the angle between simple roots $\alpha_{(i)}$ and $\alpha_{(j)}$. Then

$$\begin{aligned} A_{ij}A_{ji} &= \left(2\frac{|\alpha_{(i)}|}{|\alpha_{(j)}|} \cos\phi_{ij}\right) \left(2\frac{|\alpha_{(j)}|}{|\alpha_{(i)}|} \cos\phi_{ij}\right) \\ &= 4\cos^2\phi_{ij} \,. \end{aligned} \tag{6.126}$$

We found allowed values for ϕ_{ij} earlier (6.100). In this case we know ϕ_{ij} cannot be an integer multiple of π, since $\alpha_{(j)} \neq \pm\alpha_{(i)}$ for $j \neq i$. Thus we have $A_{ij}A_{ji} \in \{0, 1, 2, 3\}$ for $i \neq j$.

Example: Let $r = 2$. Then the Cartan matrix must be of the form

$$A = \begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix} \tag{6.127}$$

where m and n are positive integers and $\det A = 4 - mn > 0$. The only valid possibilities are

$$(m, n) \in \{ (1,1), (1,2), (1,3), (2,1), (3,1) \} . \tag{6.128}$$

(Having either m or n equal to 0 would give a reducible matrix A.)

6.8 Dynkin diagrams

There is a very enlightening connection between the Cartan matrices corresponding to simple Lie algebras and graphs known as Dynkin diagrams.

Definition 6.56. Given a Cartan matrix A, its corresponding **Dynkin diagram** is defined through the following procedure:

1. Draw a node for each simple root $\alpha_{(i)}$.

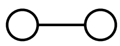
2. Join the nodes representing two simple roots $\alpha_{(i)}$ and $\alpha_{(j)}$ with k lines, where

$$k = \max\left(|A_{ij}|, |A_{ji}|\right) \in \{0, 1, 2, 3\} . \tag{6.129}$$

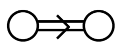
3. If more than one line connects two nodes, draw an arrowhead pointing from the longer root $\alpha_{(\ell)}$ to the shorter root $\alpha_{(s)}$, where $|\alpha_{(\ell)}| > |\alpha_{(s)}|$.

Let's illustrate this with case where $r = 2$.

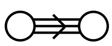
$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \tag{6.130}$$



$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \tag{6.131}$$



$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \tag{6.132}$$



Definition 6.57. A Dynkin diagram is said to be **admissible** if the corresponding Cartan matrix satisfies the constraints derived in previous sections, that is, if the diagram is connected and corresponds to a system of independent vectors (simple roots) such that the angles between any two are

$$\phi_{ij} \in \left\{ \tfrac{\pi}{2}, \tfrac{2\pi}{3}, \tfrac{3\pi}{4}, \tfrac{5\pi}{6} \right\} . \tag{6.133}$$

We now elaborate on the connections between Cartan matrix properties and the constraints these imply for admissible Dynkin diagrams. These points are discussed in §21.2 of Fulton & Harris and §II.7 of Knapp.[32]

[32] W Fulton and J Harris. *Representation Theory: A First Course*. Springer-Verlag, 1991. ISBN 0-387-97527-6. URL `https://link.springer.com/book/10.1007/978-1-4612-0979-9`; and A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL `https://link.springer.com/book/10.1007/978-1-4757-2453-0`

Claim 6.58. A diagram must be connected, otherwise the corresponding Cartan matrix A is reducible (see Claim 6.54).

For example, consider a disconnected diagram consisting of two connected parts corresponding to two sets of simple roots

$$\left\{ \alpha_{(1)}, \ldots, \alpha_{(k)} \right\} \text{ and } \left\{ \alpha_{(k+1)}, \ldots, \alpha_{(r)} \right\} \tag{6.134}$$

with $1 \leqslant k < r$. Having a disconnect between the two sets of nodes implies that every vector in the first set is orthogonal (has zero inner product) with every vector in the second set. Thus the spaces spanned by the two sets of vectors are orthogonal, and A can be reduced to block-diagonal form. This contradicts our premise that we are examining a simple Lie algebra. Thus, disconnected diagrams are not admissible.

Claim 6.59. Given an admissible diagram, any subdiagram, obtained by removing some node(s) and all lines to it (them), is still admissible as long as it remains connected.

This is the case because, given a set of linearly independent vectors, removing one or more from the set still leaves a linearly independent set. As long as the remaining nodes are connected, they will correspond to an irreducible Cartan matrix.

Claim 6.60. For a diagram with r nodes, there are at most $r-1$ pairs of nodes which are connected by lines.

To see this, let us denote rescaled, unit vectors by $\hat{\alpha}_{(i)} = \alpha_{(i)}/|\alpha_{(i)}|$. If two roots labelled by i and j are connected, then by (6.133) we have $\phi_{ij} \geqslant \frac{2\pi}{3}$ and $(\hat{\alpha}_{(i)}, \hat{\alpha}_{(i)}) \leqslant -\frac{1}{2}$.

$$0 < \left(\sum_{i=1}^{r} \hat{\alpha}_{(i)}, \sum_{i=1}^{r} \hat{\alpha}_{(i)} \right) = r + 2 \sum_{i<j} \left(\hat{\alpha}_{(i)}, \hat{\alpha}_{(j)} \right) \leqslant r - p \tag{6.135}$$

where p is the number of connected pairs of roots. Therefore $p < r$, or $p \leqslant r - 1$ since p and r are integers.

Claim 6.61. A corollary of the previous claim is: an admissible diagram contains no loops, or cycles.

A diagram which is a cycle has $p = r$ (as in Fig. 6.1). By Claim 6.59, no admissible diagram can have a cycle as a subdiagram.

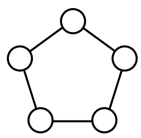


Figure 6.1: A cycle, inadmissible by Claim 6.61.

Claim 6.62. No node has more than 3 lines attached to it.

Consider a diagram where the node corresponding to root $\alpha_{(1)}$ is attached to nodes for $\alpha_{(2)}, \ldots \alpha_{(k)}$ and there are no lines attached between any of the other roots. First argue that

$$1 = (\hat{\alpha}_{(1)}, \hat{\alpha}_{(1)}) > \sum_{i=2}^{k} (\hat{\alpha}_{(1)}, \hat{\alpha}_{(i)})^2, \tag{6.136}$$

based on the fact that $\hat{\alpha}_{(1)}$ is not in the span of $\left\{ \hat{\alpha}_{(i)} \,\middle|\, i = 2, \ldots, k \right\}$ – (6.136) is an application of Bessel's inequality. Then show why this implies at most 3 lines are attached to $\alpha_{(1)}$. By Claim 6.59 this must then hold for all diagrams.

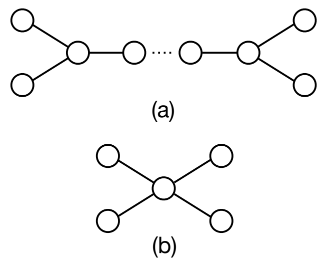


Figure 6.2: Diagrams like (a) are inadmissible by Claim 6.63, as we can replace the inner chain of nodes by a single node, resulting in (b) which has a node connected to 4 others, violating Claim 6.62.
2025-12-02: Clarifying that the ends of the linear chain are also removed, as shown in Fig. 6.2.

Claim 6.63. A linear chain of nodes (including its end nodes) in an admissible diagram can be shrunk to a single node to give another admissible diagram.

Let $\left\{ \hat{\alpha}_{(1)}, \ldots, \hat{\alpha}_{(k)} \right\}$ be the unit vectors of a linear chain of nodes. Show that $\hat{\alpha}' := \hat{\alpha}_{(1)} + \ldots + \hat{\alpha}_{(k)}$ is also a unit vector. Then argue that the inner product of $\hat{\alpha}'$ with any node not in the chain, $\hat{\alpha}_{(j)}$, is equal to either $(\hat{\alpha}_{(1)}, \hat{\alpha}_{(j)})$ or $(\hat{\alpha}_{(k)}, \hat{\alpha}_{(j)})$.

Claim 6.63 is most useful for ruling out larger diagrams based on smaller inadmissible diagrams, as in see Fig. 6.2.

Claim 6.64. A diagram with a node connected to 3 linear chains of lengths m, n, and p is admissible only if

$$\frac{1}{m} + \frac{1}{n} + \frac{1}{p} > 1\,. \tag{6.137}$$

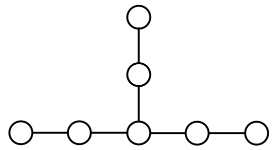


Figure 6.3: A diagram inadmissible by Claim 6.64 ($m = n = p = 3$).

See Fig. 6.3 for an inadmissible example. We leave the proof as an exercise.

Claim 6.65. A diagram which consists of $m + n$ nodes, where $\hat{\alpha}_{(1)}$ through $\hat{\alpha}_{(m)}$ are connected in a linear chain of single lines, $\hat{\alpha}_{(m+1)}$ through $\hat{\alpha}_{(m+n)}$ are connected in a linear chain of single lines, and $\hat{\alpha}_{(m)}$ is connected to $\hat{\alpha}_{(m+1)}$ by a double line (see Fig. 6.4) is only admissible if

$$(m-1)(n-1) < 2\,. \tag{6.138}$$

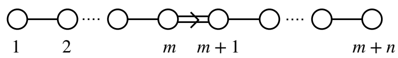


Figure 6.4: The class of diagrams discussed in Claim 6.65.

We leave this as an exercise.

These many constraints listed above on what constitutes an admissible Dynkin diagram lead to the following theorem, known as **Cartan's classification**. The theorem states that any admissible diagram is one of those depicted in Figure 6.5. There are four infinite families of diagrams, A_r, B_r, C_r, and D_r, where r is the rank of the algebra. These correspond to the complex Lie algebras (of so-called **classical** Lie groups) as tabulated in Table 6.1. Note that for small rank, some of the diagrams are equivalent, for example $C_2 = B_2$ and $D_3 = A_3$, which correspond to isomorphisms between corresponding algebras.

Table 6.1: Correspondence between Cartan's classification of Lie algebras and the Lie algebras of classical Lie groups.

A_r	$\mathfrak{su}(r+1)_{\mathbb{C}}$
B_r	$\mathfrak{so}(2r+1)_{\mathbb{C}}$
C_r	$\mathfrak{sp}(2r)_{\mathbb{C}}$
D_r	$\mathfrak{so}(2r)_{\mathbb{C}}$

There are 5 **exceptional** algebras: E_6, E_7, E_8, F_4 and G_2. As an exercise, check to see what constraint is violated as one tries to extend these algebras to larger rank.

6.9 *Reconstruction*

In this section, we show how a given Dynkin diagram completely specifies a Lie algebra, which can be reconstructed using the following procedure.

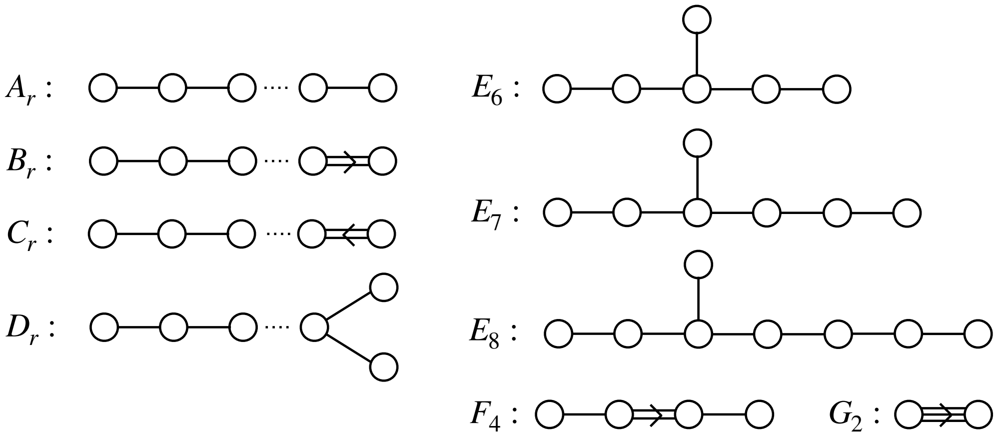


Figure 6.5: The admissible Dynkin diagrams corresponding to the possible Lie algebras.

1. As was evident from the preceding section, there is a one-to-one correspondence between a Dynkin diagram and a Cartan matrix A.

2. The Cartan matrix A determines the relative lengths and angles of the simple roots in $\mathfrak{h}^*_{\mathbb{R}}$, up to global rotations and reflections.

3. From the simple roots, the other roots can be constructed using root strings through pairs of known roots. (See the discussion around (6.106.) **TODO: Elaborate**

4. The roots determine the Cartan–Weyl basis (6.54), which is a basis for the Lie algebra $\mathfrak{g}$.

5. The full set of Lie brackets can be generated from the brackets of the simple roots using the Jacobi identity (2.35).

TODO: Example $\mathfrak{su}(3)_{\mathbb{C}}$

Classification of representations

Having thoroughly analyzed Lie algebras, we return to a discussion of representations, generalizing what was done for $\mathfrak{su}(2)$ in Chapter 4.

7.1 Basics

Let d be an N-dimensional representation of a Lie algebra $\mathfrak{g}$, and let $\{ H_i, E_\alpha \}$ be a Cartan–Weyl basis for $\mathfrak{g}$. Since d is a representation, then we must have

$$[d(H_i), d(H_j)] = d([H_i, H_j]) = 0 \tag{7.1}$$

where the latter equality follows from the $\{H_i\}$ being in the Cartan subalgebra. The full equation (7.1) implies that the $\{ d(H_i) \}$ are simultaneously diagonalizable and the representation space is spanned by the simultaneous eigenvectors of $\{ d(H_i) \}$.

Definition 7.1. Let V_λ be a vector space spanned by eigenvectors with eigenvalue λ,

$$V_\lambda := \{ v \mid d(H_i)v = \lambda_i v;\ \lambda = (\lambda_1, \ldots, \lambda_r),\ \lambda_i \in \mathbb{C} \} . \tag{7.2}$$

V_λ is said to be the eigenspace of λ, and λ is a **weight** of representation d.

Definition 7.2. Let S_d be the set of weights of representation d. We call S_d the **weight set** of d.

The full representation space V is the direct sum of the eigenspaces

$$V = \bigoplus_{\lambda \in S_d} V_\lambda . \tag{7.3}$$

Note: weights can in general have nontrivial multiplicity; any V_λ could appear more than once in the direct sum.

Claim 7.3. Let $v \in V_\lambda$. Then $d(E_\alpha)v \in V_{\lambda+\alpha}$ if $\lambda + \alpha \in S_d$, otherwise $d(E_\alpha) = 0$.

Proof:

$$\begin{aligned} d(H_i)d(E_\alpha)v &= d(E_\alpha)d(H_i)v + [d(H_i), d(E_\alpha)]v \\ &= \lambda_i d(E_\alpha)v + d([H_i, E_\alpha])v \\ &= (\lambda_i + \alpha_i)d(E_\alpha)v \end{aligned} \tag{7.4}$$

having used $[H_i, E_\alpha] = \alpha_i E_\alpha$. We have shown that either $\lambda_i + \alpha_i$ is an eigenvalue for each H_i or $d(E_\alpha)v$ is the zero vector.

Consider the action of the $\mathfrak{sl}(2)_\alpha$ generators $\{ d(h_\alpha), d(e_\alpha), d(e_{-\alpha}) \}$ on V. Each defines a linear map $V \to V$, so V is a valid representation space of $\mathfrak{sl}(2)_\alpha$.

$$\begin{aligned} d(h_\alpha)v &= \frac{2}{(\alpha,\alpha)}(\kappa^{-1})^{ij}\alpha_i\, d(H_j)v \\ &= \frac{2}{(\alpha,\alpha)}(\kappa^{-1})^{ij}\alpha_i\, \lambda_j v \\ &= \frac{2(\alpha,\lambda)}{(\alpha,\alpha)} v\,. \end{aligned} \tag{7.5}$$

The first equation follows from (6.66) and (6.71). Hence $\frac{2(\alpha,\lambda)}{(\alpha,\alpha)}$ is a weight, and as a weight of $\mathfrak{sl}(2,\mathbb{C})$, is an integer. We might refer to

$$\frac{2(\alpha,\lambda)}{(\alpha,\alpha)} \in \mathbb{Z} \tag{7.6}$$

as our **second quantization condition**.

Note that the weights are elements of the vector space which is dual to the Cartan subalgebra, $\mathfrak{h}^*$. This is the space of linear functionals which act on representation spaces of a Lie algebra. The roots are a special case, where the representation is the adjoint representation and its representation space is the Lie algebra itself.

7.2 *Root and weight lattices*

We have already seen much about the geometry of the roots of Lie algebras. We continue a little further by defining the root lattice.

Definition 7.4. The **root lattice** $\mathcal{L}[\mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$, whose simple roots are $\{\alpha_{(i)}\}$, is defined by

$$\mathcal{L}[\mathfrak{g}] := \operatorname{span}_{\mathbb{Z}} \left\{ \alpha_{(i)} \right\}. \tag{7.7}$$

That is, the root lattice is the discrete set of vectors in $\mathfrak{h}^*_{\mathbb{R}}$ which are linear combinations of the simple roots, with integer coefficients.

We can simplify some notation below by defining rescaled simple roots, i.e. the simple **co-roots** $\{\breve{\alpha}_{(i)}\}$ are simple roots normalized as follows:

$$\breve{\alpha}_{(i)} := \frac{2}{(\alpha_{(i)},\alpha_{(i)})}\alpha_{(i)}\,. \tag{7.8}$$

Definition 7.5. Correspondingly, the **co-root lattice** is

$$\breve{\mathcal{L}}[\mathfrak{g}] := \operatorname{span}_{\mathbb{Z}} \left\{ \breve{\alpha}_{(i)} \right\}. \tag{7.9}$$

Definition 7.6. The **weight lattice** is dual to the co-root lattice; it consists of all the points which have an integer inner product with elements of the co-root lattice:

$$\mathcal{L}_W[\mathfrak{g}] := \breve{\mathcal{L}}^*[\mathfrak{g}] := \left\{ \lambda \in \mathfrak{h}^*_{\mathbb{R}} \;\middle|\; (\lambda,\breve{\alpha}) \in \mathbb{Z},\, \breve{\alpha} \in \breve{\mathcal{L}}[\mathfrak{g}] \right\}. \tag{7.10}$$

Writing $\breve{\alpha} = n^i \breve{\alpha}_{(i)}$ with integer n^i, the condition

$$(\lambda, \breve{\alpha}_{(i)}) = \frac{2(\lambda, \alpha_{(i)})}{(\alpha_{(i)}, \alpha_{(i)})} \in \mathbb{Z} \tag{7.11}$$

is the same as the second quantization condition (7.6). All weights lie on the weight lattice. Note that, as the weights of the adjoint representation, the roots of $\mathfrak{g}$, lie on the weight lattice.[33] For any representation d of $\mathfrak{g}$, the weight set is a subset of the weight lattice

[33] Short exercise: Show this!

$$S_d \subset \mathcal{L}_W[\mathfrak{g}] \,. \tag{7.12}$$

As the co-roots form a basis for $\breve{\mathcal{L}}[\mathfrak{g}]$, we define the dual basis $\{\omega_{(i)}\}$ of $\mathfrak{g}$ through the requirement that

$$(\breve{\alpha}_{(i)}, \omega_{(j)}) = \delta_{ij} \,. \tag{7.13}$$

Definition 7.7. The elements $\omega_{(i)}$ of the dual basis are called the **fundamental weights**.

Since the simple roots span $\mathfrak{h}^*_{\mathbb{R}}$, we can write

$$\omega_{(j)} = \sum_{k=1}^{r} B_{jk}\, \alpha_{(k)} \tag{7.14}$$

with $B_{jk} \in \mathbb{R}$. Then (7.13) implies

$$\sum_{k=1}^{r} B_{jk} \frac{2(\alpha_{(i)}, \alpha_{(k)})}{(\alpha_{(i)}, \alpha_{(i)})} = \sum_{k=1}^{r} B_{jk} A_{ki} = \delta_{ij} \,, \tag{7.15}$$

from which we can conclude that

$$B = A^{-1} \tag{7.16}$$

and

$$\alpha_{(i)} = \sum_j A_{ij}\, \omega_{(j)} \,. \tag{7.17}$$

The fundamental weights are mapped into the simple roots by the Cartan matrix.

Let us consider for example, $\mathfrak{g} = A_2 = \mathfrak{su}(3)_{\mathbb{C}}$. The corresponding Cartan matrix is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \,, \tag{7.18}$$

from which we infer $|\alpha_{(1)}| = |\alpha_{(2)}|$ and $\cos\phi_{12} = -\frac{1}{2}$, or $\phi_{12} = \frac{2\pi}{3}$. We could choose coordinates such that $\alpha_{(1)} = (1, 0)$ and $\alpha_{(2)} = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$. Inverting A gives, after a little algebra

$$\begin{aligned} \omega_{(1)} &= \frac{1}{3}(2\alpha_{(1)} + \alpha_{(2)}) = \frac{1}{2}\left(1, \frac{1}{\sqrt{3}}\right) \\ \omega_{(2)} &= \frac{1}{3}(\alpha_{(1)} + 2\alpha_{(2)}) = \left(0, \frac{1}{\sqrt{3}}\right) \,. \end{aligned} \tag{7.19}$$

We will continue with this example in the next section. First, we return to our general discussion.

Definition 7.8. Any weight $\lambda \in S_d \subset \mathcal{L}_W[\mathfrak{g}]$ can be written as a linear combination of fundamental weights

$$\lambda = \sum_{i=1}^{r} \lambda^i \omega_{(i)} \tag{7.20}$$

with the set of $\lambda^i \in \mathbb{Z}$ called the **Dynkin labels** of weight λ.

Dynkin labels are a useful and common way (at least in quantum physics) to denote specific eigenvectors in a representation space. A vector $v_\lambda \in V_\Lambda$ can be written

$$\left|\lambda^1, \lambda^2, \ldots, \lambda^r\right\rangle := v_\lambda \tag{7.21}$$

where the components in the ket are the Dynkin labels of λ. For later convenience, let us also use square brackets to denote a weight's Dynkin labels

$$\lambda = [\lambda^1, \lambda^2, \ldots, \lambda^r] \, . \tag{7.22}$$

This will be useful when we work through specific cases.

Thinking of a specific representation d of $\mathfrak{g}$, we can make the following statements.

Definition 7.9. Every finite-dimensional representation d has at least one **highest weight** $\Lambda \in S_d$ such that $d(E_\alpha)v = 0$ for all positive $\alpha \in \Phi_+$ and vectors $v \in V_\Lambda$.

Definition 7.10. The **Dynkin labels of a representation** are the Dynkin labels of its highest weight(s), $\{\Lambda^i\}$, where

$$\Lambda = \sum_{i=1}^{r} \Lambda^i \omega_{(i)} \, . \tag{7.23}$$

Proposition 7.11. If Λ is a highest weight, then all $\Lambda^i \geqslant 0$.

Proposition 7.12. If d is irreducible, Λ is unique and the space V_Λ is one-dimensional, i.e. Λ is a nondegenerate weight.

We will not prove the statements above. The interested reader may refer to §7.2 (1st ed)/§9.1 (2nd ed) of Hall.[34]

[34] B C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015. ISBN 978-3319134666. URL https://link.springer.com/book/10.1007/978-0-387-21554-9

Now thinking about all possible representations, we make the following definition and statements

Definition 7.13. A weight λ is **dominant** if all its Dynkin labels are nonnegative, i.e. if $\lambda^i \geqslant 0$.

Proposition 7.14. (Highest Weight Theorem): For any dominant weight $\lambda \in \mathcal{L}_W[\mathfrak{g}]$, there exists a unique, irreducible, finite-dimensional representation d_λ with highest weight λ.

The proof of the Highest Weight Theorem is beyond the scope of these lectues; e.g. see Hall's Theorem 7.15 (1st ed).

Given an irrep d_Λ with highest weight Λ, for $v_\Lambda \in V_\Lambda$ we can find vectors in other eigenspaces by using lowering operators, for example

$$v_\lambda = d_\Lambda(E_{-\alpha_\ell}) \cdots d_\Lambda(E_{-\alpha_2})\, d_\Lambda(E_{-\alpha_1})\, v_\Lambda \tag{7.24}$$

where all the $\alpha_i \in \Phi_+$. Thus, every weight of d_Λ can be written as $\lambda = \Lambda - \mu$ with $\mu = \sum_{i=1}^r \mu^i \alpha_{(i)}$, and $\mu^i \in \mathbb{Z}_{\geqslant 0}$.

Lemma 7.15. ("No holes" lemma): For any finite-dimensional representation d, if $\lambda = \sum_i \lambda^i \omega_{(i)} \in S_d$, then $\lambda - m_{(i)}\alpha_{(i)} \in S_d$, where $\alpha_{(i)} \in \Phi_S$ and $m_{(i)} \in \{0, 1, \ldots, \lambda^i\}$.

Lemma 7.15 provide the basis for an algorithm for generating the weight set for representation d from its highest weight Λ. For each $\Lambda^i > 0$ one subtracts $\alpha_{(i)}$ successively $1, \ldots, \Lambda^i$ times, with each result giving another weight in S_d. One then repeats the process for each resulting weight, subtracting roots corresponding to any positive Dynkin labels λ^i. We will see how this works by example in the next section.

7.3 *Representations of* $\mathfrak{su}(3)_\mathbb{C}$

Returning to the example of $\mathfrak{su}(3)$, let us now investigate a few dominant weights. Writing

$$\Lambda = \sum_{i=1}^{2} \Lambda^i \omega_{(i)} \tag{7.25}$$

we can denote each irrep by $d_{[\Lambda^1, \Lambda^2]}$.

In following the algorithm outlined above, we will start by writing out the weights and roots in their usual notation, but also make use of the kets consisting of the weight's Dynkin labels. For reference, recall that the simple roots have Dynkin labels $\alpha_{(1)} = [2, -1]$ and $\alpha_{(2)} = [-1, 2]$.

- $d_{[0,0]}$. $\Lambda = 0$, $v_\Lambda = |0, 0\rangle$: This is the **trivial representation**, which exists for every lie algebra. While it is trivial in the mathematical sense, it certainly is not in physical applications. The trivial irrep is applicable for particle states or fields which are invariant under symmetry transformations. We say the particle or field is a scalar, or transforms as a singlet under symmetry transformations.

- $d_{[1,0]}$. $\Lambda = \omega_{(1)} = [1, 0]$: This is the **fundamental representation**. In order to find the other weights, we note that $\Lambda^1 = 1$ and $\Lambda^2 = 0$, so the only thing we can do is to subtract $\alpha_{(1)}$ to find a new weight

$$\begin{aligned} \lambda &= \Lambda - \alpha_{(1)} = [1, 0] - [2, -1] \\ &= \omega_{(2)} - \omega_{(1)} = [-1, 1] \,. \end{aligned} \tag{7.26}$$

  Now $\lambda^1 \leqslant 0$ and $\lambda^2 > 0$, so we can only subtract $\alpha_{(2)}$ to find

$$\lambda - \alpha_{(2)} = [-1, 1] - [-1, 2] = -\omega_{(2)} = [0, -1] \,. \tag{7.27}$$

  Both Dynkin labels are nonpositive, so we are done. The eigenvectors spanning the representation space of the irrep $\Lambda = \omega_{(1)}$

are $|1,0\rangle$, $|-1,1\rangle$, $|0,-1\rangle$; consequently, this is a 3-dimensional irrep.

Exercise: Show the weights of the **antifundamental representation**, $d_{[0,1]}$, are $\omega_{(2)}$, $\omega_{(1)} - \omega_{(2)}$, and $-\omega_{(1)}$ (corresponding to $|0,1\rangle$, $|1,-1\rangle$, $|-1,0\rangle$).[35]

[35] N.b. the choice of which of these irreps we call fundamental or antifundamental is arbitrary.

- $d_{[1,1]}$. $\Lambda = \omega_{(1)} + \omega_{(2)} = [1,1]$. We will see that this corresponds to the adjoint representation. For brevity's sake, we only write the weights using the bracket notation. Starting with $\Lambda = [1,1]$, we can obtain a weight by subtracting either root:

$$\Lambda - \alpha_{(1)} = [1,1] - [2,-1] = [-1,2] = \alpha_{(2)} \tag{7.28}$$

$$\Lambda - \alpha_{(2)} = [1,1] - [-1,2] = [2,-1] = \alpha_{(1)} \,. \tag{7.29}$$

In both expressions above, we have a Dynkin index equal to 2, meaning that we should subtract the corresponding root twice, obtaining two weights in each case. Starting from (7.28), we obtain

$$(\Lambda - \alpha_{(1)}) - \alpha_{(2)} = [0,0] = 0 \tag{7.30}$$

$$(\Lambda - \alpha_{(1)}) - 2\alpha_{(2)} = [1,-2] = -\alpha_{(2)} \,. \tag{7.31}$$

From (7.29), we find

$$(\Lambda - \alpha_{(2)}) - \alpha_{(1)} = [0,0] = 0 \tag{7.32}$$

$$(\Lambda - \alpha_{(2)}) - 2\alpha_{(1)} = [-2,1] = -\alpha_{(1)} \,. \tag{7.33}$$

Finally, we can take $\alpha_{(1)}$ from (7.31) or $\alpha_{(2)}$ from (7.33) to obtain

$$\Lambda - 2\alpha_{(1)} - 2\alpha_{(2)} = [-1,-1] = -\alpha_{(1)} - \alpha_{(2)} \,. \tag{7.34}$$

We see that this procedure, also depicted in Fig. 7.1, has generated a set of weights coinciding with the full root set of $\mathfrak{su}(3)$.

One thing this procedure does not tell us is about any nontrivial multiplicities. We obtained the weight $[0,0]$ in two ways. Should this be counted as a single or double weight? It could be either following this procedure for a general irrep. In the case of the adjoint representation of $\mathfrak{su}(3)$ we have additional information. We see that there are 6 nontrivial roots, and we know that the rank of the algebra is 2, i.e. the dimensionality of the Cartan subalgrebra is 2. We know that the dimension of the algebra is the sum of these, so this is an 8-dimensional representation.

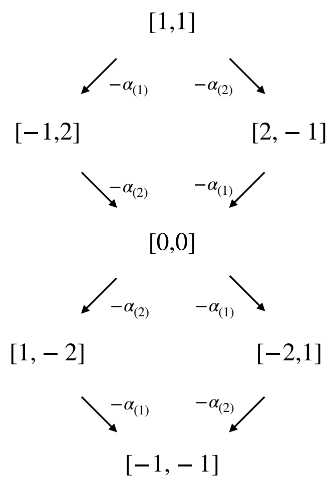


Figure 7.1: Generating the full weight set for $d_{(1,1)}$.

7.4 *Decomposition of tensor products*

Here we generalize what we did in § 4.4 for $\mathfrak{su}(2)$. Let d_Λ and $d_{\Lambda'}$ be irreducible representations of $\mathfrak{g}$ with corresponding representation spaces $V^{(\Lambda)}$ and $V^{(\Lambda')}$. These may be written as the direct sum of subspaces, each corresponding to different weights in the weight sets S_Λ and $S_{\Lambda'}$

$$V^{(\Lambda)} = \bigoplus_{\lambda \in S_\Lambda} V_\lambda \quad \text{and} \quad V^{(\Lambda')} = \bigoplus_{\lambda' \in S_{\Lambda'}} V_{\lambda'} \,. \tag{7.35}$$

If $v_\lambda \in V_\lambda$ and $v_{\lambda'} \in V_{\lambda'}$ then for $H \in \mathfrak{h}$

$$\begin{aligned}(d_\Lambda \otimes d_{\Lambda'})(H)(v_\lambda \otimes v_{\lambda'}) &= d_\Lambda(H)v_\lambda \otimes v_{\lambda'} + v_\lambda \otimes d_{\Lambda'}(H)v_{\lambda'} \\ &= (\lambda + \lambda')(H)(v_\lambda \otimes v_{\lambda'})\,. \end{aligned} \tag{7.36}$$

[Reminder: $H = \rho^i H_i \in \mathfrak{h}$ and $\lambda(H) = \rho^i \lambda_i \in \mathfrak{h}^*_{\mathbb{R}}$.]. Thus, we see from (7.36) that we have a weight $\lambda + \lambda'$ in the tensor product representation. Since the tensor product space is spanned by vectors of the form $v_\lambda \otimes v_{\lambda'}$, the whole weight set for the tensor product representation $d_\Lambda \otimes d_{\Lambda'}$ is

$$S_{\Lambda \otimes \Lambda'} = \{\, \lambda + \lambda' \mid \lambda \in S_\Lambda\,, \lambda' \in S_{\Lambda'} \,\} \tag{7.37}$$

keeping track of multiplicities. To carry out the decomposition of the tensor product representation into irreps of $\mathfrak{g}$, we follow the same procedure as in § 4.4:

1. Find a highest weight of the tensor product. This corresponds to a unique irrep.
2. Subtract out the corresponding weights from the weight set.
3. Repeat the previous two steps with the remainder of the weights until the representation has been complete decomposed into a direct sum of irreps.

For example, let us take $\mathfrak{g} = \mathfrak{su}(3)_{\mathbb{C}}$ and look at the tensor product of the fundamental representation with itself: $d_{[1,0]} \otimes d_{[1,0]}$. We have the weight set

$$S_{[1,0]} = \left\{ \omega_{(1)}, -\omega_{(1)} + \omega_{(2)}, -\omega_{(2)} \right\} = \{\, [1,0], [-1,1], [0,-1] \,\}\,. \tag{7.38}$$

where we use the Dynkin labels for convenience. Using (7.37), we find the weight set of the tensor product representation to be

$$S_{[1,0]\otimes[1,0]} = \{\, \underbrace{[2,0], [-2,2], [0,-2]}_{\text{multiplicity 1}}, \underbrace{[0,1], [1,-1], [-1,0]}_{\text{multiplicity 2}} \,\}\,. \tag{7.39}$$

(See Fig. 7.2.) Noting that $[2,0]$ is the highest weight for $d_{[2,0]}$, we can deduce that

$$S_{[1,0]\otimes[1,0]} = S_{[2,0]} \cup S_{[0,1]}\,. \tag{7.40}$$

In other words, the tensor product representation can be decomposed as

$$d_{[1,0]} \otimes d_{[1,0]} = d_{[2,0]} \oplus d_{[0,1]}$$

or, equivalently,

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}\,. \tag{7.41}$$

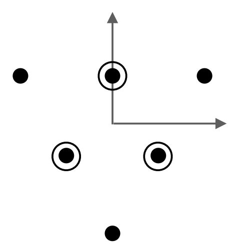


Figure 7.2: The weights of $d_{(1,0)} \otimes d_{(1,0)}$. The weights which have multiplicity 2 are circled.

7.5 $SU(3)$ flavour

The example of $\mathfrak{su}(3)$ has an important physical application in particle physics. The up (u), down (d), and strange (s) quarks are all light compared to the masses of the proton, neutron, and strange baryons. In the (theoretical) limit where the quark masses are equal and electrodynamic interactions can be neglected, we can think of u, d, and s as components of a 3-dimensional vector $q = (u, d, s)$ in **flavour** space. In this limit, quark flavour would be indistinguishable. The QCD Lagrangian becomes invariant under $SU(3)$ transformations of q and $\bar{q}$, where q transforms in the fundamental representation and $\bar{q}$ transforms in the antifundamental representation. To distinguish this approximate, global symmetry of QCD, we denote flavour symmetry by $SU(3)_F$.

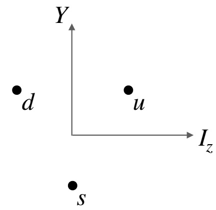


Figure 7.3: The weight diagram of the $SU(3)_F$ quark triplet.

The 2-dimensional Cartan subalgebara is spanned by quantum numbers I_z and Y, which we shall define now. First, we define baryon and strangeness numbers as differences of numbers of quarks and antiquarks as follows

$$\text{Baryon number: } B := \frac{1}{3}(n_q - n_{\bar{q}})$$
$$\text{Strangeness: } S := -n_s + n_{\bar{s}}$$

and then

$$\text{Hypercharge: } Y := B + S\,. \tag{7.42}$$

Finally, we need to consider the isospin subgroup of $SU(2)_I < SU(3)_F$, where we ignore the strange quark and consider u and d to be 2 components of an $SU(2)$ doublet $\begin{pmatrix} u \\ d \end{pmatrix}$, analogous to a 2-component spinor. We define u (d) to have z-component of isospin to be $I_z = +\frac{1}{2}$ $(-\frac{1}{2})$. It turns out that electromagnetic charge can be inferred from

$$Q := I_z + \frac{1}{2}Y\,. \tag{7.43}$$

Table 7.2: The spin-$\frac{1}{2}$ baryons transforming as an $SU(3)_F$ octet.

Baryon	content	I_z	Y
n	udd	$-\frac{1}{2}$	1
p	uud	$\frac{1}{2}$	1
Λ^0	$uds(I=0)$	0	0
Σ^-	dds	-1	0
Σ^0	$uds(I=1)$	0	0
Σ^+	uus	1	0
Ξ^-	dss	$-\frac{1}{2}$	-1
Ξ^0	uss	$\frac{1}{2}$	-1

We now look at weights of various representations on the weight lattice of $\mathfrak{h}^*_{\mathbb{R}}$. Figure 7.3 shows the $SU(3)_F$-triplet of quarks in the fundamental representation at $(I_z, Y) = (\frac{1}{2}, \frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$, and $(0, -1)$, corresponding to the u, d, and s components of the triplet. The triplet of antiquarks in the antifundamental representation have weights at $(I_z, Y) = (-\frac{1}{2}, -\frac{1}{2})$, $(\frac{1}{2}, -\frac{1}{2})$, and $(0, 1)$, corresponding to $\bar{u}$, $\bar{d}$, and $\bar{s}$.

Bound states of quarks, called hadrons, must also respect $SU(3)_F$ in the symmetry limit. They form themselves into irreducible representations which make up tensor products of quark representations. For mesons, quark – antiquark states, these are the irreps of $\mathbf{3} \otimes \bar{\mathbf{3}}$. Baryons, composed of 3 quarks, lie in the irreps of $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$. The lightest spin-$\frac{1}{2}$ baryons listed in Table 7.2 transform as an 8-dimensional representation. Their weights are shown in Fig. 7.4. Note that the $I_z = 0 = Y$ state is degenerate; however, the Casimir

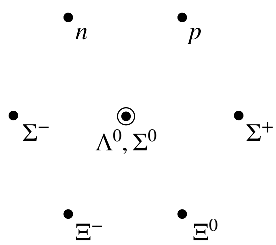


Figure 7.4: The weight diagram of the baryon octet. As can be inferred from Table 7.2, the horizontal axis is I_z and the vertical axis is Y.

operator I^2 distinguishes these. The Λ^0 baryon is an $I = 0$ singlet, while the three Σ baryons transform as an $I = 1$ triplet under $SU(2)_I$.

Gauge theories

8.1 Electrodynamics

This section reviews classical and quantum electrodynamics, which we expect you have seen elsewhere.

Maxwell's equations read

$$\nabla \cdot \vec{E} = 4\pi\rho\,, \qquad \nabla \times \vec{B} = 4\pi\vec{J} + \frac{\partial \vec{E}}{\partial t}\,,$$
$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}\,, \qquad \nabla \cdot \vec{B} = 0 \tag{8.1}$$

The latter two equations imply that $\vec{E}$ and $\vec{B}$ can be written in terms of scalar and vector potentials

$$\vec{E} = -\nabla\Phi + \frac{\partial \vec{A}}{\partial t}\,, \qquad \vec{B} = \nabla \times \vec{A}\,. \tag{8.2}$$

The electric and magnetic fields are invariant under gauge transformations

$$\Phi \mapsto \Phi + \frac{\partial \alpha}{\partial t}\,, \qquad \vec{A} \mapsto \vec{A} + \nabla\alpha \tag{8.3}$$

where $\alpha = \alpha(t, \vec{x})$ is a real function of each spacetime point $\alpha : \mathbb{R}^{1,3} \to \mathbb{R}$. The potentials are the key quantities in formulating quantum theories, but we have to keep in mind that there is an unphysical ambiguity, or redundancy, in Φ and $\vec{A}_\mu$.

In a relativistic treatment one combines the scalar. potential as the temporal component of the 4-vector potential

$$a_\mu = \begin{pmatrix} \Phi \\ A_i \end{pmatrix} \tag{8.4}$$

The gauge tranformation (8.3) then reads

$$a_\mu \mapsto a_\mu + \partial_\mu \alpha\,. \tag{8.5}$$

The electric and magnetic fields are components of the field strength tensor

$$f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & -B_1 & 0 \end{pmatrix} \tag{8.6}$$

or more succinctly, $f_{0i} = E_i$ and $f_{ij} = -\epsilon_{ijk} B_k$. The electromagnetic Lagrangian is

$$\mathcal{L}_{\text{EM}} = \frac{1}{2e^2}\left(|\vec{E}|^2 - |\vec{B}|^2\right) = -\frac{1}{4e^2} f_{\mu\nu} f^{\mu\nu} . \tag{8.7}$$

We assume the mostly minus Minkowski metric, and we have introduced the electromagnetic coupling e in what perhaps seems to be an unusual place – in the next section we will rescale the fields to move the coupling to a more usual place.[36] In physics, we normally work with this a_μ and $f_{\mu\nu}$, capitalizing the letters. For consistency with this course, we prefer to use

$$\begin{aligned} A_\mu &:= -i a_\mu \in i\mathbb{R} = L(U(1)) \\ F_{\mu\nu} &:= -i f_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu . \end{aligned} \tag{8.8}$$

[36] In a few words, we can rescale the fields to move the coupling so that it appears in the gauge-covariant derivative instead of in the coefficients in (8.7). That approach might seem more natural in a physics context, but the present treatment will allow more consistency with previous lectures.

In this notation the gauge transformation is

$$A_\mu \mapsto A_\mu - i\partial_\mu \alpha . \tag{8.9}$$

Scalar electrodynamics

Consider a complex scalar field $\phi : \mathbb{R}^{1,3} \to \mathbb{C}$ with Lagrangian (density)

$$\mathcal{L}_\phi = \partial_\mu \phi^* \, \partial^\mu \phi - W(\phi^* \phi) . \tag{8.10}$$

The first term represents the kinetic energy of the field and the second is some potential. The Lagrangian (8.10) is invariant under a global (spacetime-independent) $U(1)$ tranformation of ϕ. Let $g = e^{i\theta} \in U(1)$ with constant $\theta \in [0, 2\pi)$, and transform

$$\phi \mapsto g\phi \quad \text{and} \quad \phi^* \mapsto \phi^* g^{-1} . \tag{8.11}$$

In order to make a connection to what we learned from Cartan's classification, we will want to consider small transformations, where g is near the identity. Let $0 < \varepsilon \ll 1$ and expand

$$g = \exp \varepsilon X \approx 1 + \varepsilon X \tag{8.12}$$

where $X \in i\mathbb{R} = L(U(1))$. The we write the transformation in the small ε limit as

$$\phi \mapsto \phi + \delta_X \phi , \quad \text{with} \quad \delta_X \phi = \varepsilon X \phi . \tag{8.13}$$

In this notation, we have

$$\delta_X(\phi^* \phi) = 0 \quad \text{and} \quad \delta_X \mathcal{L} = 0 . \tag{8.14}$$

This is fine for a theory of scalar particles. However, if we want to describe scalars interacting with an electromagnetic field, we would like to couple ϕ to the vector field A_μ in a way which accounts for the gauge ambiguity (8.9) of the 4-vector potential. The way to do this is to ”gauge” the global $U(1)$ symmetry. That is, we find a Lagrangian which is invariant under the "local" transformation

$$\phi(x) \mapsto g(x)\phi(x) \quad \text{and} \quad \phi^*(x) \mapsto \phi^*(x)g(x)^{-1} \tag{8.15}$$

where now $g : \mathbb{R}^{1,3} \to U(1)$ associates a different group element to each point in spacetime.

It is clear that the kinetic term in $\mathcal{L}_\phi$ is no longer invariant. Considering infinitesimal transformations

$$g(x) \approx 1 + \varepsilon X(x) \tag{8.16}$$

we have

$$\delta_X \phi = \varepsilon X \phi \tag{8.17}$$

$$\delta_X(\partial_\mu \phi) = \varepsilon X \, \partial_\mu \phi + \varepsilon(\partial_\mu X)\phi \, . \tag{8.18}$$

The last term above is uncancelled in the transformation of the kinetic term.

A small miracle occurs, though. This extra term is of the same form as the ambiguity in A_μ (8.9). We can insist that under (8.15)

$$\delta_X A_\mu = -\varepsilon \, \partial_\mu X \, . \tag{8.19}$$

This is just as in (8.9). Then we can form a gauge-invariant Lagrangian by replacing all partial derivatives of the scalar field with a **gauge-covariant derivative**

$$D_\mu := \partial_\mu + A_\mu \, . \tag{8.20}$$

With this change, $D_\mu \phi$ transforms just like ϕ:

$$\begin{aligned} \delta_X(D_\mu \phi) &= \partial_\mu(\delta_X \phi) + A_\mu \delta_X \phi + (\delta_X A_\mu)\phi \\ &= \varepsilon \left[\partial_\mu(X\phi) + A_\mu X \phi - (\partial_\mu X)\phi \right] \\ &= \varepsilon X \, D_\mu \phi \, . \end{aligned} \tag{8.21}$$

Similarly, $(D_\mu \phi)^*$ transforms like ϕ^* and we have a gauge-invariant Lagrangian for scalar electrodynamics in

$$\mathcal{L} = \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^*(D^\mu \phi) - W(\phi^* \phi) \, . \tag{8.22}$$

8.2 *Nonabelian gauge theory*

In this section, we generalize the ideas of the preceding section to construct theories where matter couples to theories where the gauge fields map to elements of more general Lie algebras. We start with a scalar field theory with a global symmetry group G, which we take to be compact. Say we have an N-component field ϕ

$$\phi : \mathbb{R}^{1,3} \to V \, . \tag{8.23}$$

where V is $\mathbb{R}^N$ or $\mathbb{C}^N$. We also assume the usual inner product on V, so that we can write the Lagrangian as

$$\mathcal{L}_\phi = (\partial_\mu \phi, \partial^\mu \phi) - W((\phi, \phi)) \, . \tag{8.24}$$

Say that $\mathcal{L}_\phi$ is invariant under transformations of some group G. More explicitly, what we mean is that, for an N-dimensional representation Δ of G, the field transforms as

$$\phi \mapsto \Delta(g)\phi \tag{8.25}$$

for any $g \in G$. Note that we change from using D as a group representation to Δ in order to avoid clashing notation with the gauge-covariant derivative.

We will assume that G is a compact group, so Δ is a unitary representation (or is equivalent to one). That is, we can chose a representation so that the inner product is group-invariant:

$$(\Delta(g)\phi, \Delta(g)\phi) = (\phi, \phi) \quad \forall g \in G \tag{8.26}$$

since $\Delta(g)^\dagger = \Delta(g)^{-1}$.

Notice that so far we have assumed that the field $\phi(x)$ is transformed by the same $\Delta(g)$ at each spacetime point.

Definition 8.1. A field transformation is **global** if the action on the field is independent of spacetime coordinate. If the Lagrangian is invariant under such a transformation, the associated symmetry is said to be a **global symmetry**.

Definition 8.2. A field transformation is **local** if the action on the field is generally different for each spacetime coordinate. If the Lagrangian is invariant under such a transformation, the associated symmetry is said to be a **local symmetry** or a **gauge symmetry**.

Say we wish to **gauge** the scalar field theory, that is, we wish to construct one which is invariant under local transformations

$$\phi(x) \mapsto (\Delta(g(x))\phi(x)\,, \quad g(x) \in G\,. \tag{8.27}$$

It will be useful to consider infinitesimal transformations, where, introducing a small parameter ε, we can write

$$\begin{aligned} g(x) &= \exp \varepsilon X(x) \approx e + \varepsilon X(x) \\ \Delta(g(x)) &= \exp \varepsilon\, d(X(x)) \approx I + \varepsilon\, d(X(x))\,, \end{aligned} \tag{8.28}$$

where d is the representation of $L(G)$ corresponding to Δ. When we make this a local transformation, the change in the field becomes $\phi \mapsto \phi + \delta_X\phi$ with

$$\delta_X\phi(x) = \varepsilon\, d(X(x))\phi(x)\,. \tag{8.29}$$

To couple the field ϕ to gauge degrees-of-freedom and give a gauge-invariant Lagrangian, we define the gauge-covariant derivative

$$D_\mu\phi(x) = \partial_\mu\phi(x) + d(A_\mu(x))\phi(x) \tag{8.30}$$

where $A_\mu : \mathbb{R}^{1,3} \to L(G)$ is the gauge field. In order for $D_\mu\phi$ to transform just as ϕ does, i.e.

$$\begin{aligned} &D_\mu\phi \mapsto \Delta(g)D_\mu\phi \\ \text{or} \quad &\delta_X(D_\mu\phi) = \varepsilon\, d(X)D_\mu\phi\,. \end{aligned} \tag{8.31}$$

We must work out how A_μ should transform so that (8.31) holds. Let's say that $A_\mu \mapsto A'_\mu$. Then we require

$$\begin{aligned} D_\mu\phi &\mapsto (\partial_\mu + d(A'_\mu)\Delta(g)\phi = \Delta(g)D_\mu\phi \\ \Delta(g)^{-1}(\partial_\mu &+ d(A'_\mu)\Delta(g)\phi = (\partial_\mu + d(A_\mu))\phi \\ &\implies \; d(A'_\mu) = \Delta(g)d(A_\mu)\Delta(g)^{-1} - (\partial_\mu\Delta(g))\Delta(g)^{-1} \\ &\quad\;\; \text{or} \;\; A'_\mu = gA_\mu g^{-1} - (\partial_\mu g)g^{-1} \end{aligned} \tag{8.32}$$

Writing $\Delta(g) = I + \varepsilon\, d(X)$, the infinitesimal version of (8.31) implies

$$\delta_X A_\mu = -\varepsilon\,\partial_\mu X + \varepsilon[X, A_\mu]\,. \tag{8.33}$$

Let's check that this works:

$$\begin{aligned} \delta_X(D_\mu\phi) &= \delta_X\left(\partial_\mu\phi + d(A_\mu)\phi\right) \\ &= \partial_\mu(\delta_X\phi) + d(A_\mu)\delta_X\phi + d(\delta_X A_\mu)\phi \\ &= \varepsilon\Big\{\underbrace{\partial_\mu(d(X)\phi)}_{(1)} + \underbrace{d(A_\mu)d(X)\phi}_{(2)} - d(\partial_\mu X)\phi + \varepsilon d([X, A_\mu])\phi\Big\} \\ &= \varepsilon\Big\{\underbrace{d(X)\partial_\mu\phi + d(\partial_\mu X)\phi}_{(1)} + \underbrace{d(X)d(A_\mu)\phi + d([A_\mu, X])\phi}_{(2)} \\ &\qquad - d(\partial_\mu X)\phi - d([A_\mu, X])\phi\Big\} \\ &= \varepsilon\, d(X)D_\mu\phi\,. \end{aligned} \tag{8.34}$$

Note that we implicitly used the linearity of $d(X)$: for $A_\mu \to A_\mu + \delta_X A_\mu$, we have $d(A_\mu) \to d(A_\mu + \delta_X A_\mu) = d(A_\mu) + d(\delta_X A_\mu)$.

Note that Δ being unitary implies that

$$\begin{aligned} \Delta(g)^\dagger &= \Delta(g)^{-1} \\ I + \varepsilon d(X)^\dagger &= I - \varepsilon d(X) \\ d(X)^\dagger &= -d(X)\,. \end{aligned} \tag{8.35}$$

Therefore the inner product $(D_\mu\phi, D^\mu\phi)$ is gauge-invariant

$$\begin{aligned} \delta_X(D_\mu\phi, D^\mu\phi) &= \varepsilon\left[(d(X)D_\mu\phi, D^\mu\phi) + (D_\mu\phi, d(X)D^\mu\phi)\right] \\ &= \varepsilon\left[(D_\mu\phi, d(X)^\dagger D^\mu\phi) + (D_\mu\phi, d(X)D^\mu\phi)\right] \\ &= 0\,. \end{aligned} \tag{8.36}$$

In physics notation, context is enough to distinguish between group or algebra elements and representations thereof. Now that we have set the scene with careful notation, in the following we simply write g instead of $\Delta(g)$, A_μ instead of $d(A_\mu)$, and X instead of $d(X)$.

The field strength tensor is defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \in L(G)\,. \tag{8.37}$$

Note that

$$\begin{aligned} [D_\mu, D_\nu]\phi &= [\partial_\mu + A_\mu, \partial_\nu + A_\nu]\phi \\ &= \partial_\mu(A_\nu\phi) + A_\mu\partial_\nu\phi - \partial_\nu(A_\mu\phi) - A_\nu\partial_\mu\phi + [A_\mu, A_\nu]\phi \\ &= F_{\mu\nu}\phi\,. \end{aligned} \tag{8.38}$$

Along with (8.31), this tells us that

$$F_{\mu\nu} \mapsto g F_{\mu\nu} g^{-1} \,. \tag{8.39}$$

We see that $F_{\mu\nu}$ transforms in the adjoint representation of G. Infinitesimally, this transforms as

$$\begin{aligned}
\delta_X F_{\mu\nu} &= \underbrace{\partial_\mu(\delta_X A_\nu)}_{(1)} - \underbrace{\partial_\nu(\delta_X A_\mu)}_{(2)} + \underbrace{[\delta_X A_\mu, A_\nu]}_{(3)} + \underbrace{[A_\mu, \delta_X A_\nu]}_{(4)} \\
&= \varepsilon\Big\{ -\underbrace{\partial_\mu\partial_\nu X + \partial_\mu[X, A_\nu]}_{(1)} + \underbrace{\partial_\nu\partial_\mu X - \partial_\nu[X, A_\mu]}_{(2)} \\
&\qquad - \underbrace{[\partial_\mu X, A_\nu] + [[X, A_\mu], A_\nu]}_{(3)} - \underbrace{[A_\mu, \partial_\nu X] + [A_\mu, [X, A_\nu]]}_{(4)} \Big\} \\
&= \varepsilon\Big\{ \underbrace{[X, \partial_\mu A_\nu]}_{(1)} - \underbrace{[X, \partial_\nu A_\mu]}_{(2)} - \underbrace{[A_\nu, [X, A_\mu]]}_{(3)} - \underbrace{[A_\mu, [A_\nu, X]}_{(4)} \Big\} \\
&= \varepsilon[X, F_{\mu\nu}]
\end{aligned} \tag{8.40}$$

where we used the Jacobi identity (2.35) in the obtaining last line. Again we see that $F_{\mu\nu}$ transforms in the adjoint representation of G.

In order to construct a Lorentz scalar for the Lagrangian, we again look to $F_{\mu\nu}F^{\mu\nu}$, however this is the product of two Lie algebra elements and is not gauge-invariant. We need an inner product, which we have with the Killing form. Therefore we arrive at the **Yang–Mills Lagrangian** for the gauge degrees-of-freedom

$$\mathcal{L}_{\text{YM}} = \frac{1}{g_s^2} \sum_{\mu,\nu} \kappa(F_{\mu\nu}, F^{\mu\nu}) \,. \tag{8.41}$$

Because we started with a compact group, we can choose a basis in which the Killing form is diagonal

$$\kappa_{ab} = \kappa(T_a, T_b) = -\kappa\delta_{ab} \,. \tag{8.42}$$

The requirement to have the canonical normalization for the kinetic energy term of the gauge field fixes κ, and so we have

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2g_s^2} \operatorname{Tr} F_{\mu\nu}F^{\mu\nu} \,. \tag{8.43}$$

In (8.43), the trace is carried out in the representation space V appropriate for ϕ.

Like in the case of QED, we introduced a coupling constant g_s into the Lagrangian. We can rescale the gauge field to move the coupling constant to a more natural place, at least from the point of view of perturbation theory and Feynman diagrams. Let $\hat{A}_\mu = A_\mu / g_s$, which results in

$$\mathcal{L}_{\text{YM}} = -\frac{1}{2} \operatorname{Tr} \hat{F}_{\mu\nu}\hat{F}^{\mu\nu} \,, \tag{8.44}$$

where

$$\hat{F}_{\mu\nu} = \partial_\mu \hat{A}_\nu - \partial_\nu \hat{A}_\mu + g_s[\hat{A}_\mu, \hat{A}_\nu] \,. \tag{8.45}$$

The covariant derivative becomes $D_\mu = \partial_\mu + g_s \hat{A}_\mu$ which enters $\mathcal{L}_\phi$ through the usual $(D_\mu \phi, D^\mu \phi)$ term. Now the role of g_s as a coupling becomes clear when generating the perturbative series of Feynman diagrams (see AQFT next term).

The Standard Model

A prime example of a gauge theory, is the Standard Model of elementary particle physics. There is full lecture course next term on the Standard Model, so this section serves to connect what we have developed from the side of symmetry groups to basic ingredients of the Standard Model. The gauge group is a direct product of three compact Lie groups

$$G = SU(3)_c \times SU(2)_L \times U(1)_Y \,. \tag{8.46}$$

The subscripts are labels to denote physics related to the different gauge sectors, as we will explain. Briefly, the "c" stands for "colour"; $SU(3)_c$ is the gauge symmetry of the strong interactions. The "L" in $SU(2)_L$ refers to the fact that the weak interactions only involve fermions with left-handed chirality. The remaining $U(1)_Y$ is the **weak hypercharge** part of the Standard Model; this hypercharge is completely different physics from the one introduced in the context of the approximate flavour symmetry $SU(3)_F$ (§ 7.5). Together, $SU(2)_L \times U(1)_Y$ is called the **electroweak** sector of the Standard Model. This gauge symmetry is said to be spontaneously broken down to a remnant $U(1)_{\text{em}}$, i.e. electromagnetism. This is a fascinating story which will be told properly next term.

The algebra of the Standard Model is thus

$$\mathfrak{g} = \mathfrak{su}(3) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1) \,. \tag{8.47}$$

For each algebra, we introduce a gauge field which allows us to construct covariant derivatives acting on any matter fields. For $\mathfrak{su}(3)$, we have the 8-dimensional gluon field $A^a_\mu(x)$, where $a = 1, \ldots, 8$. For $\mathfrak{su}(2) \times \mathfrak{u}(1)$, we have a triplet gauge field $\mathcal{W}^k_\mu(x)$, where $k = 1, 2, 3$, and a single $\mathfrak{u}(1)$ gauge field $B_\mu(x)$. Before symmetry breaking, all gauge fields are massless. However, the spontaneous breaking via the (Anderson–. . . –) Higgs mechanism mixes the electroweak gauge fields, giving mass to 3 of them; we have the massive vector bosons $W^+_\mu(x)$, $W^0_\mu(x)$, $W^-_\mu(x)$ and the massless photon $A_\mu(x)$.

Having specified the gauge group of the Standard Model, we now consider its matter content. From this course, we know how to classify the possible representation spaces in which fields *could* lie. Here we describe the fields which are necessary to understand experimental observations.

Only quark fields interact nontrivially with the gluon field. As seen in § 7.5, we wish describe the low-lying baryon spectrum by triplets of quark fields. The way we do this is to describe each quark field in the fundamental, 3-dimensional representation **3** of

$SU(3)_c$; antiquarks then are in the antifundamental representation, $\bar{\mathbf{3}}$. Other particles, which do not feel the strong force, are described by fields which transform trivially under $SU(3)_c$, i.e. in representation $\mathbf{1}$.

Recalling the notion of flavour isospin introduced in § 7.5, the approximate global symmetry of hadrons formed from up and down quarks, we extend the idea to **weak isospin** for representations of $SU(2)_L$ and we denote weak isospin by T. Fields which do not interact weakly have $T = 0$. The nontrivial representation needed for the electroweak interactions is $T = \frac{1}{2}$. We can choose some axis in the corresponding 3-dimensional representation space, say T_3, as the basis for the $\mathfrak{su}(2)$ Cartan subalgebra, Then we have $T_3 = \pm\frac{1}{2}$ for the $SU(2)_L$-doublets of the Standard Model.

Finally, we use the weak hypercharge Y to fully classify the Standard Model fields. After symmetry breaking, the electromagnetic charge of the fields is given by

$$Q = T_3 + \frac{Y}{2}\,. \tag{8.48}$$

Now we are ready to list the Standard Model matter content. We use the notation $(SU(3)\text{ irrep}, SU(2)\text{ irrep})_Y$.

- Scalar Higgs: We have a single field whose Standard Model representation is

  $$(\mathbf{1},\mathbf{2})_1\,. \tag{8.49}$$

  This is an $SU(2)$-doublet which is a consequence of the Higgs mechanism. (A lot of work in constructing theories and searching for experimental hints for physics beyond the Standard Model seeks to explain this scalar field in terms of some even more fundamental dynamics.)

- Leptons: The left-handed components of electrons and electron neutrinos should be understood as the two components of an $SU(2)_L$ doublet, while the right-handed part of the electron field does not interact weakly. That is, we have the fields

  $$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L\,,\quad e_R \tag{8.50}$$

  where ν_e has $T_3 = +\frac{1}{2}$ and e_L has $T_3 = -\frac{1}{2}$. (Recall, e_R has $T = 0$ and hence $T_3 = 0$.) The corresponding lepton representations are

  $$(\mathbf{1},\mathbf{2})_{-1} \oplus (\mathbf{1},\mathbf{1})_{-2}\,. \tag{8.51}$$

  In addition to the electron and electron neutrino, there are 2 more **generations** of lepton fields: the μ and τ leptons and their partner neutrinos. Note that there are no right-handed neutrinos in the Standard Model. There is also no mixing between neutrinos of different generations in the Standard Model, something we now know happens in Nature. The Standard Model extensions required to describe neutrino mixing are not as elaborate as attempted to "explain" electroweak symmetry breaking, but they have the virtue of having real experimental evidence of physics beyond the Standard Model.

- Quarks: These are the only particles to feel all fundamental forces. The $SU(2)$-doublets couple pairs of left-handed quark fields: up–down, charm–strange, top–bottom. The right-handed parts are $SU(2)$-singlets:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L , \quad u_R , \quad d_R . \tag{8.52}$$

  Once again, the left-handed $SU(2)$-doublets and the right-handed $SU(2)$-singlets are in separate irreps. Together, the Standard Model irrep for any pair of quark fields is

$$(\mathbf{3}, \mathbf{2})_{\frac{1}{3}} \oplus (\mathbf{3}, \mathbf{1})_{\frac{4}{3}} \oplus (\mathbf{3}, \mathbf{1})_{-\frac{2}{3}} \tag{8.53}$$

  corresponding to (8.52).

Of course, there is a lot more physics in the Standard Model than a taxonomy of its content. You have 24 lectures to look forward to next term!

Appendix: Discrete groups

In this appendix we introduce a number of concepts in the theory of discrete groups. Many of these will generalize straightforwardly to Lie groups.

The texts by Jones[37] and Ramond[38] have nice introductory chapters on discrete groups, going into more detail than we can.

[37] H F Jones. *Groups, Representations, and Physics*. IOP Publishing, 1990. ISBN 0-85274-029-8

[38] P Ramond. *Group Theory: A Physicist's Survey*. Cambridge University Press, 2010. ISBN 978-0-521-89603-0. URL https://www.cambridge.org/core/books/group-theory/8BAC137A9F0C43D65448E6420248D841

9.1 *Basic definitions*

Definition 9.1. A **group** is a set G of objects or elements together with a binary operation, here denoted by $\cdot$, which we will usually call a **product**, where the four group axioms discussed below hold.

The group may have a finite number of elements n, in which case the **order** $|G|$ of the **finite group** is n, or it may have an infinite number of elements, in which case we say that it is an **infinite group**, or a group of infinite order. This appendix reviews a few concepts within the context of **discrete groups**, those groups with a countable number of elements

$$G = \{g_1, g_2, \ldots, g_k, \ldots\}\,. \tag{9.1}$$

The following axioms must hold

(i) **Closure**: The product of any pair of group elements must also be an element of the group. We use the notation

$$\forall g_1, g_2 \in G\,, \quad g_2 \cdot g_1 \in G \tag{9.2}$$

which reads "for all elements g_1 and g_2 in G, their product $g_2 \cdot g_1$ is also an element of G."

(ii) **Associativity**: The group operation must be associative, that is we must have

$$g_3 \cdot (g_2 \cdot g_1) = (g_3 \cdot g_2) \cdot g_1 \tag{9.3}$$

$\forall g_1, g_2, g_3 \in G$.

(iii) **Identity**: One, and only one, element of G, denoted e, must satisfy the following for all $g \in G$

$$e \cdot g = g \cdot e = g\,. \tag{9.4}$$

You can prove by contradiction that the identity must be unique.

(iv) **Inverse**: Each element $g \in G$ must have a unique inverse $g^{-1} \in G$, such that

$$g \cdot g^{-1} = e = g^{-1} \cdot g\,. \tag{9.5}$$

You can also use proof by contradiction to show that for every g, g^{-1} is unique.

Note that we have not insisted that the group operation be commutative. If the following property holds for all $g_1, g_2 \in G$

$$g_2 \cdot g_1 = g_1 \cdot g_2 \tag{9.6}$$

then we say that the group is **commutative** or **Abelian**, otherwise we say the group is **noncommutative** or **nonabelian**.

Examples

Some examples of infinite groups

- The integers $\mathbb{Z}$ form a group under addition. Let us denote any three integers by a, b, c. Then closure holds since $a + b \in \mathbb{Z}$. Integer addition is associative: $(a + b) + c = a + (b + c)$. The identity is 0. The inverse of a is $(-a)$: $a + (-a) = 0 = (-a) + a$. The group is commutative since $a + b = b + a$.
- The rational numbers excluding 0 form a group under multiplication. Check for yourselves that the group axioms hold and that the group is commutative. Why must 0 be excluded?

9.2 *Groups and symmetries*

Let g_1 represent a transformation of a physical system which leaves physical laws invariant, and let g_2 represent another such transformation. Then the combination, or composition, of these two transformations, written

$$g_2 \circ g_1 = g_2 g_1 \tag{9.7}$$

and read right-to-left, also leaves the physical system invariant. The collection of all symmetry transformations $\{g_1, g_2, \ldots, g_k, \ldots\} = G$ forms a group under composition. The group axioms are all satisfied. By construction, closure is satisfied. Composition of these transformations is associative. The identity is the "do nothing" transformation. And we assume that any transformation can be undone, so the corresponding group element has its inverse in G.

Let's focus on the symmetries of a two-sided equilateral triangle. The triangle is invariant under rotation by an angle $2\pi/3$ and under reflection about any of the three lines connecting a vertex to the triangle's centre (see Fig. 9.1). The symmetry group is the **dihedral group** D_3.

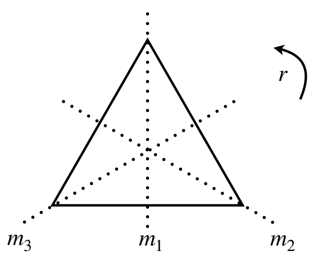


Figure 9.1: Actions of D_3 on an equilateral triangle.

Let us denote the anticlockwise rotation by $2\pi/3$ by r, and a reflection about the vertical axis by m. The key relation for the

	⟵ g_1 ⟶					
	e	r	r^2	m_1	m_2	m_3
	r	r^2	e	m_2	m_3	m_1
↑	r^2	e	r	m_3	m_1	m_2
g_2	m_1	m_3	m_2	e	r^2	r
↓	m_2	m_1	m_3	r	e	r^2
	m_3	m_2	m_3	r^2	r	e

Table 9.3: Multiplication table for products $g_2 g_1$ in D_3, where $m_1 = m$, $m_2 = mr^2$, and $m_3 = mr$.

dihedral group is that the product mr, an anticlockwise rotation followed by a reflection, is equivalent to $r^{-1}m$, a reflection followed by a clockwise rotation. With this we can then determine all the group products and fill in the group's **multiplication table** (Table 9.3). We can generate all symmetry rotations and reflections using only r and m; therefore, we call r and m **generators** of D_3.

Note that each row in the multiplication table is a distinct permutation of the group elements. The same fact holds for the columns. Convince yourself that if this were not the case, a contradiction would arise.

There is a very convenient way to specify a group using its generators. This is called a **presentation** and takes the form $G = \langle \text{generators} \,|\, \text{rules} \rangle$. The rules specify how the generators are related to the identity and to each other. For example

$$D_3 = \left\langle r, m \,\middle|\, r^3 = m^2 = e\,;\; mr = r^{-1}m \right\rangle . \tag{9.8}$$

In fact, the dihedral group describing the symmetries of an n-gon is the generalization

$$D_n = \left\langle r, m \,\middle|\, r^n = m^2 = e\,;\; mr = r^{-1}m \right\rangle . \tag{9.9}$$

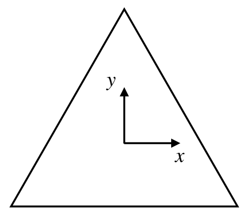


Figure 9.2: Mapping of the equilateral triangle onto $\mathbb{R}^2$.

Another concept which will be a major component of this course later is that of a group **representation**. Let's introduce it here informally, in the context of D_3. Imagine drawing x- and y-axes on the triangle, with the origin at its centre (Fig. 9.2). Then vectors in that 2-dimensional vector space transform under r and m according to left-multiplication by the following matrices

$$\begin{aligned} r = R\left(\tfrac{2\pi}{3}\right) &= \begin{pmatrix} \cos\frac{2\pi}{3} & -\sin\frac{2\pi}{3} \\ \sin\frac{2\pi}{3} & \cos\frac{2\pi}{3} \end{pmatrix} \\ m &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} . \end{aligned} \tag{9.10}$$

Matrix representations for the other group elements can be obtained using the multiplication table. Denoting the matrix representation of any group element g by $\mathbf{D}(g)$, any vector in the 2-dimensional vector space we defined when we drew the axes transforms as $\mathbf{v} \mapsto \mathbf{D}(g)\mathbf{v}$.

Since all the matrices are distinct, this representation is said to be **faithful**.

9.3 *Useful concepts*

In this section, let us recall some concepts which will be useful later

Conjugacy:

Definition 9.2. Two elements of G, say g_1 and g_2, are said to be **conjugate**, written $g_1 \sim g_2$, if there exists any $g \in G$ such that

$$g_2 = g g_1 g^{-1} \,. \tag{9.11}$$

Conjugacy is an equivalence relation, which partitions G into **conjugacy classes**. For example, in D_3 the conjugacy classes are

$$\{e\}, \quad \{r, r^2\}, \quad \{m_1, m_2, m_3\} \,, \tag{9.12}$$

that is, the identity, the rotations, and the reflections.

Definition 9.3. A subset $H \subseteq G$ is a **subgroup** if H is a group in its own right. We write $H \leqslant G$. Every group has two **improper subgroups**, the trivial group $\{e\}$ and the group G itself. All other subgroups are **proper subgroups**, and the notation $H < G$ is used.

The proper subgroups of D_3 are

$$\{e, r, r^2\}, \quad \{e, m_1\}, \quad \{e, m_2\}, \quad \{e, m_3\} \,. \tag{9.13}$$

Lagrange's theorem states that if G is finite and $H \leqslant G$, then $|H|$ divides $|G|$.

Definition 9.4. A mapping from group $(H, \cdot)$ to group $(K, *)$, $\phi : H \to K$, is a **group homomorphism** if it preserves the group structure. That is, we require $\phi(h_1 \cdot h_2) = \phi(h_1) * \phi(h_2)$, for all $h_1, h_2 \in H$.

Definition 9.5. The **kernel** of ϕ, ker ϕ, consists of all elements of H which map to the identity of K.

A special case of homomorphism is as follows.

Definition 9.6. Two groups H and K are said to be **isomorphic**, denoted $H \cong K$, if there is a bijective (one-to-one) map (an **isomorphism**), $\phi : H \to K$ which preserves the group structure.

It follows that $|H| = |K|$ and the kernel of ϕ is solely the identity in H.

Definition 9.7. A subgroup $H < G$ is **normal**, written $H \lhd G$, if H is a union of conjugacy classes.

In our example $\{e, r, r^2\} \lhd D_3$, but the order-2 subgroups $\{e, m_i\}$ are not normal. If we wish to consider normal subgroups which may not be proper, then we write $H \trianglelefteq G$.

Definition 9.8. Let $H < G$ and $g \in G$. Then the **left coset** of H with g, written gH, is the set formed by the left product of g with each element H:

$$gH = \{\, gh \mid h \in H \,\} \,. \tag{9.14}$$

Similarly, the **right coset** of H with g, written Hg is

$$Hg = \{\, hg \mid h \in H \,\} \,. \tag{9.15}$$

It is a theorem that, for all $g \in G$, the left and right cosets are equal if and only if H is normal:

$$gH = Hg\,, \;\; \forall g \in G \iff H \lhd G\,. \tag{9.16}$$

(The reader is invited to prove this theorem as an exercise.) For example, take $G = D_3$ and $H = \{e, r, r^2\}$. Then

$$gH = Hg = \begin{cases} \{e, r, r^2\}\,, & g \in H \\ \{m_1, m_2, m_3\}\,, & g \notin H \end{cases} . \tag{9.17}$$

If we took H to be one of the order-2 subgroups, none of which are normal, then we would find $gH \neq Hg$ for some $g \in D_3$.

Coset product: Let us define a product of two cosets to be the set of the distinct products of all elements. Continuing with our example

$$\{e, r, r^2\} \cdot \{m_1, m_2, m_3\} = \{m_1, m_2, m_3\} \tag{9.18}$$

where we have noted that $\{em_1, em_2, em_3\}$ is the same set as $\{rm_1, rm_2, rm_3\}$ and $\{r^2m_1, r^2.m_2, r^2m_3\}$. Let us label the two cosets

$$I = \{e, r, r^2\} \;\text{ and }\; a = \{m_1, m_2, m_3\}\,. \tag{9.19}$$

These two cosets form a group under coset multiplication, in fact it is the cyclic group of order 2,

$$\{I, a\} \cong \langle a | a^2 = e \rangle = C_2 \tag{9.20}$$

where I acts as the identity. More generally the **cyclic group** of order n is

$$C_n = \langle a \,|\, a^n = e \rangle\ . \tag{9.21}$$

We observe that $I = \{e, r, r^2\}$ is isomorphic to C_3.

Definition 9.9. Generalizing the example above, the cosets of any normal subgroup $H \trianglelefteq G$ form a group under the coset product, called the **quotient group** (or **factor group**) G/H, with H as its identity.

We omit the nontrivial proof that the quotient group is, in fact, a group, as described. The example above shows that the quotient group $D_3/C_3 = C_2$.

One way we combine groups in theoretical physics is the direct product.

Definition 9.10. A group G is a **direct product** of groups A and B, that is

$$G = A \times B = \{\ (a, b) \mid a \in A,\ b \in B\ \} \tag{9.22}$$

and the group product is defined as

$$gg' = (aa', bb') \tag{9.23}$$

for any $g = (a, b)$ and $g' = (a', b')$ in G.

In high energy physics, this definition and this ordered pair notation is probably the most usual way we encounter direct product groups. We think about forming a larger group by taking the direct product of two smaller groups.

From another perspective, we can consider subgroups of G, namely $\tilde{A}$ and $\tilde{B}$ defined via

$$\begin{aligned} \tilde{A} &= \{ (a, e_B) \mid a \in A \} \\ \tilde{B} &= \{ (e_A, b) \mid b \in B \} \end{aligned} \tag{9.24}$$

where $e_A(e_B)$ is the identity of $A(B)$. These groups are clearly isomorphic to A and B, respectively. In this sense, we can say that a group G is the direct product of subgroups $\tilde{A} \times \tilde{B}$ where

(i) $\tilde{A} \cap \tilde{B} = (e_A, e_B) =: e_G$

(ii) $\tilde{a}\tilde{b} = \tilde{b}\tilde{a}$ for all $\tilde{a} \in \tilde{A}$ and $\tilde{b} \in \tilde{B}$,

(iii) All $g \in G$ can be expressed uniquely as some $g = \tilde{a}\tilde{b}$.

These conditions are necessary and sufficient to say that a group is the direct product of two of its subgroups.

For example, the dihedral group $D_2 = \langle r, m | r^2 = m^2 = e, mr = rm \rangle$ is the direct product $D_2 = C_2^{(r)} \times C_2^{(m)}$ (the superscripts distinguish between the cyclic groups with different generators). On the other hand, despite the fact that $D_3/C_3 = C_2$, D_3 is not a direct product group, $D_3 \neq C_3 \times C_2$.

This definition extends straightforwardly to Lie groups, and plays a role in high energy physics. In the Standard Model of particle physics, the gauge group is the direct product $SU(3)_c \times SU(2)_L \times U(1)_Y$, and the approximate flavour symmetry of the light hadrons, $SU(3)_F$ is the unbroken subgroup of $SU(3)_L \times SU(3)_R$ chiral symmetry (separate $SU(3)$ transformations for the left- and right-handed components of the quark spinors).

In addition to direct product groups, we will also encounter semidirect product groups. Let G be a group with normal subgroup $N \triangleleft G$ and another subgroup $H < G$ which is not necessarily normal.

Definition 9.11. G is the **semidirect product** group $G = H \ltimes N$ or $G = N \rtimes H$,[39] provided that the following equivalent statements to hold

[39] Note the open end of the symbol is next to the normal subgroup.

(i) N and H have a trivial intersection, and G is equal to the subgroup product NH (or HN).

(ii) Every element of G can be uniquely written as a product nh (or hn) for some $n \in N$ and $h \in H$.

Let us elaborate. Let ϕ be the group homomorphism

$$\phi : H \to \text{Aut}(N) \tag{9.25}$$

such that

$$\phi_h(n) := \phi(h)(n) = hnh^{-1} \in N \tag{9.26}$$

for all $h \in H$ and all $n \in N$. That is $\phi_h : N \to N$, which is what we mean by $\phi_h \in \mathrm{Aut}(N)$. Recall that $gng^{-1} \in N$ for all $g \in G$ is equivalent to the defining property of a normal subgroup.

Now form a group G' consisting of pairs

$$G' = \{ (n, h) \mid n \in N, h \in H \} \tag{9.27}$$

under the operation "$\cdot$", where

$$(n_2, h_2) \cdot (n_1, h_1) = (n_2\phi_{h_2}(n_1), h_2h_1)\,. \tag{9.28}$$

The groups G and G' are isomorphic. We find an isomorphism in the map $f : G \to G'$, with $f(g) = f(nh) = (n, h)$. Note that a direct product group is a special case of a semidirect product group, one where ϕ_h is the identity map for all $h \in H$. Then $G = N \times H$ and $H \lhd G$.

We close this chapter by listing the groups of order 8. In addition to the cyclic group C_8 and the dihedral group D_4, we have the quaternions, defined succinctly by the presentation

$$Q = \left\langle i, j \,\middle|\, i^4 = e, i^2 = j^2, ij = j^{-1}i \right\rangle \tag{9.29}$$

or more familiarly as the set

$$Q = \{\pm 1, \pm i, \pm j, \pm k\} \tag{9.30}$$

with the rules $ij = k$ and $i^2 = j^2 = k^2 = ijk = -1$. We also have the direct product groups $C_4 \times C_2$ and $C_2^{(a)} \times C_2^{(b)} \times C_2^{(c)}$ (superscript labels just distinguish the different copies of C_2).

Appendix: Vector spaces

In this Appendix, we collect a few facts from linear algebra which will be used in this course. A few key concepts are defined, and relevant facts are used. Where it is simple to do so, we give proofs of some claims. Otherwise, the proofs are left for the curious to reproduce.

10.1 Basic notions

Definition 10.1. A **vector space** V over a field $\mathbb{F}$ is an Abelian group under vector addition ($u + v \in V$ for every u and $v \in V$) which also is closed under scalar multiplication ($\alpha v \in V$ for every $\alpha \in \mathbb{F}$ and $v \in V$), and for which axioms of associativity and distributivity hold, as described below.

The group operation is vector addition, with the zero vector $\vec{0}$ as the identity and the inverse $v^{-1} = -v$ for every $v \in V$.[40] In this course, we will only be concerned with the cases where $\mathbb{F}$ is the field of real or complex numbers, $\mathbb{R}$ or $\mathbb{C}$. The axioms which must be satisfied are

[40] We refrain from using special notation for vectors unless we need to avoid ambiguity.

(i) For all $\alpha, \beta \in \mathbb{F}$ and $v \in V$, we have the associativity property that

$$\alpha(\beta v) = (\alpha\beta)v\,. \tag{10.1}$$

(ii) For all $\alpha, \beta \in \mathbb{F}$ and $v \in V$, we have distributivity in $\mathbb{F}$:

$$(\alpha + \beta)v = \alpha v + \beta v\,. \tag{10.2}$$

(iii) For all $\alpha \in \mathbb{F}$ and $u, v \in V$, we have distributivity in V:

$$\alpha(u + v) = \alpha u + \alpha v\,. \tag{10.3}$$

(iv) Given that 1 is the multiplicative identity of $\mathbb{F}$, we have for all $v \in V$

$$1v = v\,. \tag{10.4}$$

Definition 10.2. A subset $U \subseteq V$ is a **subspace** of V, $U \leqslant V$, if it is itself a vector space, with the same operations as V.

Definition 10.3. A **linear combination** of elements $\{\, v_1, \ldots, v_k \,\} \in V$ is an element of V which can be written as

$$\alpha_1 v_1 + \ldots + \alpha_k v_k \tag{10.5}$$

where $\{\, \alpha_1, \ldots, \alpha_k \,\} \in \mathbb{F}$.

Definition 10.4. Let V be a $\mathbb{F}$ vector space and let $U \subseteq V$. The **span** of U is defined to be the smallest subspace of V containing S, with

$$\langle U \rangle = \left\{ \sum_{i=1}^{n} \alpha_i u_i \;\middle|\; \alpha_i \in \mathbb{F},\ u_i \in U,\ n \geqslant 0 \right\}. \tag{10.6}$$

Definition 10.5. If $\langle U \rangle = V$, we say that U **spans** V.

Definition 10.6. A set of vectors $\{ u_1, \ldots, u_k \}$ is **linearly independent** if none of the vectors can be written as a linear combination of the others. Equivalently, the set is linearly independent if and only if

$$\alpha_1 u_1 + \ldots + \alpha_k u_k = 0 \implies \alpha_1 = \ldots = \alpha_k = 0\,. \tag{10.7}$$

Definition 10.7. Let V be a vector space over $\mathbb{F}$ and $U \subseteq V$. U is a **basis** for V if U is linearly independent and spans V.

Definition 10.8. The **dimension** of a vector space V is the number of vectors in any basis for V.

We refer to linear algebra texts for proof that the dimension of a vector space is uniquely defined.

10.2 *Direct sum*

In this section we explore ways of summing vector spaces. The general sum will not be very useful for this course, but the special case of the direct sum will be important.

Definition 10.9. Let U and W be be subspaces of some $\mathbb{F}$ vector space V. The **sum** of U and W is

$$U + W = \{ u + w \mid u \in U,\ w \in W \}\ . \tag{10.8}$$

It can easily be shown that $U + W$ is also a subspace of V. The intersection space $U \cap W$ is also a subspace of V. By choosing bases for these spaces, one can show that

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)\,. \tag{10.9}$$

The direct sum of vector spaces is a special case.

Definition 10.10. (Internal definition) Let U and W be be subspaces of some $\mathbb{F}$ vector space V. V is the **direct sum** of U and W, written $V = U \oplus W$, if (i) $V = U + W$ and (ii) $U \cap W = \vec{0}$.

This is equivalent to saying that every element $v \in V$ can be expressed as $v = u + w$ for some $u \in U$ and $w \in W$.

We can define the direct sum of vector spaces without first referring to the larger vector space.

Definition 10.11. (External definition) Let U and W be $\mathbb{F}$ vector spaces. The vector space formed from the **direct sum** of these is

$$U \oplus W = \{ (u, w) \mid u \in U,\ w \in W \} \tag{10.10}$$

with straightforward extensions of the definitions of vector addition

$$(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2) \tag{10.11}$$

and scalar multiplication by $\alpha \in \mathbb{F}$

$$\alpha(u, w) = (\alpha u, \alpha w) . \tag{10.12}$$

Multiple vector spaces can be combined via the direct sum, which we can write as

$$U_1 \oplus U_2 \oplus \ldots \oplus U_n = \bigoplus_{i=1}^{n} U_i = \{ (u_1, \ldots, u_n) \mid u_i \in U_i \} . \tag{10.13}$$

10.3 Linear maps

Definition 10.12. A **linear map** f from F vector space U to $\mathbb{F}$ vector space V is a map which satisfies

$$f(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 f(u_1) + \alpha_2 f(u_2) \tag{10.14}$$

where $\alpha_1, \alpha_2 \in \mathbb{F}$ and $u_1, u_2 \in U$.

Definition 10.13. A linear map $f : U \to V$ is an **isomorphism** if the map is a bijection, that is, if it is both one-to-one and onto. Equivalently, f is an isomorphism if there exists a linear map $g : V \to U$ such that $f \circ g = \mathrm{id}_V$ and $g \circ f = \mathrm{id}_U$, where id_U is the identity map for $U \to U$, and similarly for id_V.

Elaborating on the notation above, we mean that, for all $u \in U$ and $v \in V$

$$\begin{aligned} (f \circ g)(v) &= f(g(v)) & \mathrm{id}_V(v) &= v \\ (g \circ g)(u) &= g(f(u)) & \mathrm{id}_U(u) &= u . \end{aligned} \tag{10.15}$$

Claim 10.14. If $f : U \to V$ is a linear map from between finite-dimensional $\mathbb{F}$ vector spaces, then, after choosing bases for each vector space, then f can be represented by an $n \times m$ matrix, where $m = \dim U$ and $n = \dim V$.

Proving this claim involves several intermediate steps, so we simply refer the reader to Linear Algebra notes or texts.

Here is a little terminology we probably will not need, except immediately below.

Definition 10.15. A mapping $f : V \to V$, that is a linear map where both the source and the target spaces are the same, is generically called an **endomorphism** and written as $\mathrm{End}(V)$.

Generally, we will only be interested in a subset of endomorphisms, namely automorphisms.

Definition 10.16. An **automorphism** is an endomorphism which is also an isomorphism. That is, $\mathrm{Aut}(V)$ consists of all bijections, or invertible linear maps, $f : V \to V$.

The set of automorphisms $V \to V$, $\text{Aut}(V)$, form a group $GL(V)$ under the composition operation.

- **Closure:** If f and g are both linear maps, then so is $f \circ g : V \to V$, with $(f \circ g)(v) = f(g(v))$ for all $v \in V$.
- **Associativity:** $f \circ (g \circ h) = (f \circ g) \circ h$.
- **Identity:** We define the identity map $\text{id}_V : V \to V$ such that $\text{id}_V(v) = v$ for all $v \in V$.
- **Inverse:** By definition, an automorphism is an isomorphism, so each mapping f has a corresponding inverse mapping f^{-1}.

As we have seen, for a finite, n-dimensional, F vector spaces, we can associate an invertible $n \times n$ matrix with any given linear map once we have chosen a basis. In this context, we usually write the group of matrices corresponding to $GL(V)$ as $GL(n, \mathbb{F})$. We can say that $GL(V)$ and $GL(n, \mathbb{F})$ are isomorphic.

10.4 Dual vector space

In the last section, we saw that the set of invertible linear maps $V \to V$ form a group $GL(V)$. Here we show that they also form a vector space which we denote by V^*.

Definition 10.17. Given a vector space V over a field $\mathbb{F}$, the **dual vector space** V^* is the space of linear functions, or linear maps, $f : V \to \mathbb{F}$.

Claim 10.18. The dual vector space is itself a vector space.

Proof. For any $v \in V$, $\alpha \in \mathbb{F}$, and $f, g \in V^*$,

$$(\alpha\beta f)(v) = \alpha[\beta f(v)] = \alpha\beta[f(v)] \tag{10.16}$$

$$[(\alpha + \beta)f](v) = (\alpha f)(v) + (\beta f)(v) = (\alpha + \beta)f(v) \tag{10.17}$$

$$(f + g)(v) = f(v) + g(v) \tag{10.18}$$

□

Claim 10.19. If V is finite-dimensional, $\dim V^* = \dim V$.

Given a basis $\{v_i\}$ for V, we can find a dual basis in V^*, $\{v_i^*\}$ such that $v_i^*(v_j) = \delta_{ij}$.

10.5 Quantum notation

In quantum mechanics, Dirac's *bra-ket* notation is often used. Vectors in the relevant vector space are represented by *kets*, $|\Psi\rangle$ and dual vectors are represented by *bras*, $\langle\Phi|$.

Definition 10.20. The **Hermitian adjoint** operator is a map $\dagger : V \to V^*$ such that, for any $|\Psi\rangle \in V$,

$$|\Psi\rangle^\dagger = \langle\Psi| \in V^* . \tag{10.19}$$

10.6 Inner product

Definition 10.21. An **inner product** is a map on a vector space V over field $\mathbb{F}$

$$i : V \times V \to \mathbb{F}\,. \tag{10.20}$$

which satisfies conditions of conjugate symmetry, sesquilinearity in its arguments, and positive-definiteness.

Let us elaborate on these properties. Letting $u, v, w \in V$ and $\alpha, \beta \in \mathbb{F}$,

1. **Conjugate symmetry.** $(i(u,v))^* = i(v,u)$.
2. **Sesquilinearity** in its arguments. That is, linearity in one of the arguments. In quantum notation, $i(u,v)$ is written $\langle u \,|\, v\rangle$ and we choose to impose linearity in the second argument: $\langle u \,|\, \alpha v + \beta w\rangle = \alpha \langle u \,|\, v\rangle + \beta \langle u \,|\, w\rangle$. In mathematics, one usually chooses to impose linearity on the first argument. By conjugate symmetry, the inner product is conjugate-linear in the other argument, e.g. in physics conventions $\langle \alpha u + \beta v \,|\, w\rangle = \alpha^* \langle u \,|\, w\rangle + \beta^* \langle v \,|\, w\rangle$.
3. **Positive-definiteness.** For $v \neq 0$, the inner product $i(v,v) > 0$.

If $\mathbb{F} = \mathbb{R}$, then conjugate symmetry is just symmetry and sesquilinearity become bilinearity.

The following definitions are not used in the course, but provided here for reference.

Definition 10.22. An **inner product space** is a vector space with an inner product.

Definition 10.23. A **metric space** is a set along with a notion of a distance between two elements.

Definition 10.24. A metric space is called **complete** if there are no points missing. That is, it is complete if every Cauchy sequence of points has a limit which is also in the space.

Definition 10.25. A **Hilbert space** is an inner product space which is also a complete metric space, with the distance defined using the inner product.

10.7 Bilinear forms

A bilinear form is similar to an inner product.

Definition 10.26. A **bilinear form** is a map

$$B : V \times V \to \mathbb{F} \tag{10.21}$$

which is linear in both its arguments (even for $\mathbb{F} = \mathbb{C}$):

$$\begin{aligned} B(u, \alpha v + \beta w) &= \alpha B(u,v) + \beta B(u,w) \\ B(\alpha u + \beta v, w) &= \alpha B(u,w) + \beta B(v,w)\,. \end{aligned} \tag{10.22}$$

Definition 10.27. A **symmetric, bilinear form** satisfies $B(u, v) = B(v, u)$ for all $u, v \in V$. Note that there is no condition on $B(v, v)$.

Definition 10.28. A bilinear form B is said to be **nondegenerate** if, for all nonzero $v \in V$, there exists some $w \in V$ such that

$$B(v, w) \neq 0 \,. \tag{10.23}$$

We will see that the degeneracy or nondegeneracy of the bilinear form we are about to define tells us about the type of Lie algebra we are considering. We will be able to use a nondegenerate bilinear form as a lind of generalized inner product.[41]

A bilinear form can also be an inner product, e.g.if $\mathbb{F} = \mathbb{R}$ and the bilinear form satisfies positive-definiteness.

[41] The scalar product of two 4-vectors in Minkowski space is a familiar example of a nondegenerate, symmetric, bilinear form.

10.8 Tensor product spaces

Let V and W be vector spaces over $\mathbb{F}$, and that both are finite dimensional.

Definition 10.29. The **tensor product space** $V \otimes W$ is the $\mathbb{F}$ vector space consisting of the linear combinations of elements $v \otimes w$, with $v, v_1, v_2 \in V$ and $w, w_1, w_2 \in W$, satisfying the following relations:

(i) $(\alpha v) \otimes w = v \otimes (\alpha w) = \alpha(v \otimes w)$

(ii) $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$

(iii) $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2 \,.$

If $\{ v_i \}$ and $\{ w_j \}$ are bases of V and W, respectively, then $\{ v_i \otimes w_j \}$ is a basis of $V \otimes W$. Consequently

$$\dim(V \otimes W) = \dim V \dim W \,. \tag{10.24}$$

Tensor product spaces are at the core of quantum computing. A **qubit** $|\psi\rangle$ is a normalized vector in $\mathbb{C}^2$ which, after choosing orthonormal basis vectors $|0\rangle$ and $|1\rangle$, can be written as

$$|\psi\rangle = \alpha \,|0\rangle + \beta \,|1\rangle \ , \tag{10.25}$$

with $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$. A two-qubit state is a normalized vector in the tensor product space $\mathbb{C}^2 \otimes \mathbb{C}^2$, spanned by the product states $\{|0\rangle \otimes |0\rangle \,, |0\rangle \otimes |1\rangle \,, |1\rangle \otimes |0\rangle \,, |1\rangle \otimes |1\rangle\} \equiv \{|00\rangle \,, |01\rangle \,, |10\rangle \,, |11\rangle\}$, introducing some abbreviated notation.

Definition 10.30. A **product state** or **product vector** is an element of a tensor product space $V \otimes W$ which can be written as $v \otimes w$ for some $v \in V$ and $w \in W$.

For example, with $|\psi\rangle$ given in (10.25) and

$$|\phi\rangle = \gamma \,|0\rangle + \delta \,|1\rangle \,, \tag{10.26}$$

we can form the product state

$$|\psi\rangle \otimes |\phi\rangle = \alpha\gamma \,|00\rangle + \alpha\delta \,|01\rangle + \beta\gamma \,|10\rangle + \beta\delta \,|11\rangle \ . \tag{10.27}$$

Definition 10.31. An **entangled state** is an element of a tensor product state which is not a product state.

For example,

$$\frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) \tag{10.28}$$

is entangled; it cannot be expressed as a product state (10.27). It can be shown that the state $|\Psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$

$$|\Psi\rangle = a\,|00\rangle + b\,|01\rangle + c\,|10\rangle + d\,|11\rangle \tag{10.29}$$

is entangled if and only if $ad - bc \neq 0$. (Try it!)

Appendix: Quantum mechanics review

Undergraduate quantum mechanics is a prerequisite for many of the Part III courses in theoretical physics, including this course. Nevertheless, below we attempt to summarize the most relevant ideas from quantum mechanics. Those wishing to read further could consult the following texts: Gasiorowicz[42] is a standard undergraduate reference, the books by Binney and Skinner [43] and by Weinberg [44] are approporate reference for Part III theoretical physicists, and one of my pure mathematician colleagues recommends Faddeev and Yakubovskiĭ[45] as an introduction to quantum mechanics aimed at mathematicians.

[42] S Gasiorowicz. *Quantum Physics*. Wiley, 1974

[43] J Binney and D Skinner. *The Physics of Quantum Mechanics*. Oxford University Press, 2014. ISBN 978-0-19-968857-9

[44] S Weinberg. *Lectures on Quantum Mechanics*. Cambridge University Press, 2013. ISBN 978-0-107-02872-2. URL https://www.cambridge.org/core/books/lectures-on-quantum-mechanics/F739B9577D2473995024FA5E9ABA9B6C

[45] L D Faddeev and O A Yakubovskiĭ. *Lectures on Quantum Mechanics for Mathematics Students*. American Mathematical Society, 2009. ISBN 978-0-8218-4699-5

11.1 Basics

In quantum mechanics, physical states are represented by rays in a Hilbert space $\mathcal{H}$. $\mathcal{H}$ is often infinite-dimensional.

Definition 11.1. A **ray** is a family of vectors in a Hilbert space $\mathcal{H}$ which are equal up to a multiplicative complex factor, i.e. for any $|\Psi\rangle \in \mathcal{H}$,

$$|\Psi\rangle_{\text{ray}} = \{\, \lambda\,|\Psi\rangle \mid \lambda \in \mathbb{C} \,\} \,. \tag{11.1}$$

Since these overall factors do not influence physical observables, we usually do not distinguish between the ray representing a physical state and any particular vector in the ray. We are usually interested in states which can be normalized such that $\langle \Psi \,|\, \Psi \rangle = 1$. Once normalized, the allowed vectors in the ray differ by a multiplicative phase $\lambda = \exp i\theta$, with $\theta \in [0, 2\pi)$.

In the Schrödinger picture, where we assign a time-dependence to the vector $|\Psi(t)\rangle$, the evolution in time of a system is governed by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\,|\Psi\rangle = \hat{H}\,|\Psi\rangle \tag{11.2}$$

where $\hat{H}$ is an automophism on $\mathcal{H}$ representing the Hamiltonian of the system. The constant $\hbar$ is the reduced Planck constant; below we will work in units where $\hbar = 1$. We can formally write the solution to (11.2) as

$$|\Psi(t)\rangle = e^{-i\hat{H}t}\,|\Psi(0)\rangle \,, \tag{11.3}$$

where the operator $e^{-i\hat{H}t}$ is called the time-evolution operator.

Suppose that we are interested in the quantum dynamics of a single particle of mass M. The Hamiltonian consists of two terms

$$\hat{H} = \frac{1}{2M}|\hat{\vec{p}}|^2 + V(\hat{\vec{x}})\,. \tag{11.4}$$

The first term is the kinetic energy term, consisting of the momentum operator squared. The second term is the potential energy term, where we assume that the potential energy is a function of the position operator only. Solving (11.3) amounts to determining the eigenstates of $\hat{H}$ and writing $|\Psi(0)\rangle$ in this basis.

The eigenstates of the Hamiltonian are those with definite energy E:

$$\hat{H}\,|\Psi(0)\rangle = E\,|\Psi(0)\rangle \implies |\Psi(t)\rangle = e^{-i\hat{E}t}\,|\Psi(0)\rangle\ . \tag{11.5}$$

That is, the time evolution of such an "energy eigenstate" or "stationary state" is just multiplication by an overall phase, yielding vectors belonging to the same ray.

Let $|\psi\rangle$ be an energy eigenstate. In this case Schrödinger's equation (11.2) reduces to the time-independent Schrödinger equation

$$\hat{H}\,|\psi\rangle = \left[\frac{1}{2M}|\hat{\vec{p}}|^2 + V(\hat{\vec{x}})\right]|\psi\rangle = E\,|\psi\rangle\ . \tag{11.6}$$

11.2 *Orbital angular momentum*

Let us assume that we have a central potential, *viz* $V = V(|\hat{\vec{x}}|)$. In this case, it is useful to look at the angular momentum operator

$$\hat{\vec{L}} := \hat{\vec{x}} \times \hat{\vec{p}}\,. \tag{11.7}$$

From the canonical commutation relations

$$[\hat{x}_j, \hat{p}_k] = i\delta_{jk} \tag{11.8}$$

it is straightforward to derive the commutation relations

$$[\hat{L}_j, \hat{L}_k] = i\epsilon_{jk\ell}\hat{L}_\ell \quad \text{and} \quad [|\hat{\vec{L}}|^2, \hat{L}_k] = 0\,. \tag{11.9}$$

With some further work, one can conclude that, for central potentials, the Hamiltonian commutes with $|\hat{\vec{L}}|^2$ and any individual component $\hat{L}_k$. Without loss of generality, usually one chooses $\hat{L}_z$. Then the three relations

$$[\hat{H}, |\hat{\vec{L}}|^2] = 0\,,\quad [\hat{H}, \hat{L}_z] = 0\,,\quad [|\hat{\vec{L}}|^2, \hat{L}_z] = 0\,, \tag{11.10}$$

imply that we can find simultaneous eigenvectors of the three operators $\{\,\hat{H}, |\hat{\vec{L}}|^2, \hat{L}_z\,\}$, by the spectral decomposition theorem.

Our study of the finite-dimensional representations of $\mathfrak{su}(2)$ in § 4.2 is the modern way to solve the angular part of these problems using symmetry arguments. Below we briefly outline the more historical treatment using the position representation. This treatment may be familiar from undergraduate lectures. It may be instructive to see how the two approaches fit together.

The position representation uses eigenstates of the position operator, $|x\rangle$, such that $\hat{\vec{x}}\,|\vec{x}\rangle = \vec{x}\,|\vec{x}\rangle$. Any state – here we focus on stationary states – can be written in the position representation as

$$|\psi\rangle = \int d^3x\, \psi(\vec{x})\, |\vec{x}\rangle \ . \tag{11.11}$$

The complex-valued function $\psi(\vec{x}) = \langle \vec{x} \,|\, \psi \rangle$ is called the **wavefunction**.

The derivation you may have encountered from a wave mechanics approach would follow the steps below. The level of detail here is meant to be just enough to remind one how the story goes from this point of view and to provide a familiar application of the algebraic approach in the text. A detailed revision of this material is not needed for this course.

1. Recalling that in the position representation, the momentum operator is $p_k = -i\frac{\partial}{\partial x_k}$, the operator $L^2 := |\hat{\vec{L}}|^2$ can be expressed as

$$L^2 = -r^2\nabla^2 + \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} \, . \tag{11.12}$$

   where $r = |\vec{x}|^2$. Schrödinger's time-independent equation (11.6) becomes

$$-\frac{1}{2Mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{L^2}{2Mr^2}\,\psi + V(r)\,\psi = E\,\psi \, . \tag{11.13}$$

2. We will work in spherical polar coordinate $x_1 = r\cos\phi\sin\theta$, $x_2 = r\sin\phi\sin\theta$, $x_3 = r\cos\theta$, where

$$L^2 = -\frac{1}{\sin^2\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{\partial^2}{\partial\phi^2}\right] \, . \tag{11.14}$$

3. Consider the wavefunction $\psi(\vec{x})$ in the limit that we approach the origin. As long as $V(r)$ is not too singular – we will see what this means soon – $\psi(\vec{r})$ can be written as a power series in r, where each coefficient is a function of θ and ϕ. Let the integer $\ell \geqslant 0$ be the power of the dominant term in the power series as $r \to 0$, i.e. the smallest power, such that as $r \to 0$ we can write

$$\psi(\vec{x}) \sim r^\ell Y(\theta,\phi)\,, \tag{11.15}$$

   where all angular dependence is contained in the function $Y(\theta,\phi)$. Rearranging (11.13) as

$$L^2\psi(\vec{r}) = \frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + 2Mr^2[E - V(r)]\psi(\vec{r})\,, \tag{11.16}$$

   and taking the limit $r \to 0$ and using (11.15), we have

$$L^2\psi = \ell(\ell+1)\psi + 2Mr^2[E - V(r)]\psi \to \ell(\ell+1)\psi \, . \tag{11.17}$$

   The second term above vanishes in the limit $r \to 0$ as long as $V(r)$ is less singular than r^{-2}.

4. For general r, we use (11.15) as inspiration to look for separable solutions, where we can factor out the angular dependence from all powers of r:

$$\psi(\vec{x}) = R(r)\, Y(\theta, \phi)\,. \tag{11.18}$$

   The L^2 operator (11.14) only involves derivatives with respect to angular variables, so separable solutions (11.18) are eigenfunctions of L^2

$$\begin{aligned} L^2\psi &= R(r)\, L^2\, Y(\theta,\phi) \\ L^2\, Y(\theta,\phi) &= \ell(\ell+1) Y(\theta,\phi)\,. \end{aligned} \tag{11.19}$$

5. Since any component of $\hat{\vec{L}}$ commutes with $\hat{H}$ and $|\hat{\vec{L}}|^2$, we can find a basis of functions of the form $Y(\theta,\phi)$ which are eigenfunctions of $L_3 = -i\frac{\partial}{\partial\phi}$:

$$\begin{aligned} L_3 Y(\theta,\phi) &= m Y(\theta,\phi) \\ \frac{\partial Y}{\partial \phi} &= im Y(\theta,\phi) \end{aligned} \tag{11.20}$$

   The differential equation implies we can write $Y(\theta,\phi) = Q(\theta)e^{im\phi}$ for some function $Q(\theta)$. Since we must have $Y(\theta,\phi+2\pi) = Y(\theta,\phi)$, m must be an integer. The $Y(\theta,\phi)$ are characterized by the two integers ℓ and m, one often writes them as $Y_\ell^m(\theta,\phi)$. These special functions are called **spherical harmonics**.

6. Applying L^2 as in (11.14), using $\frac{\partial^2 Y}{\partial\phi^2} = -m^2 Y$, and letting $u = \cos\theta$ (hence $\sin\theta\frac{d}{d\theta} = -(1-u^2)\frac{d}{du}$), and $P(u) = Q(\theta)$, we find

$$\begin{aligned} L^2 P(u)\, e^{im\phi} &= -\left\{ \frac{d}{du}\left[(1-u^2)\frac{dP}{du}\right] + \frac{m^2}{1-u^2}P(u)\right\} e^{im\phi} \\ &= \ell(\ell+1) P(u)\, e^{im\phi}\,. \end{aligned} \tag{11.21}$$

   Thus $P(u)$ satisfies the ordinary differential equation

$$(1-u^2)\frac{d^2P}{du^2} - 2u\frac{dP}{du} + \left[\ell(\ell+1) - \frac{m^2}{1-u^2}\right] P = 0\,. \tag{11.22}$$

   This is the general Legendre equation. It has nonsingular solutions for integer m with $-\ell \leqslant m \leqslant \ell$. (We need solutions to be nonsingular if they are to be normalizable.) These are the associated Legendre polynomials and are denoted as $P_\ell^m(u)$, with

$$P_\ell^m(u) = \frac{1}{2^\ell \ell!}(1-u^2)^{m/2}\, \frac{d^{\ell+m}}{du^{\ell+m}}\, (u^2-1)^\ell\,. \tag{11.23}$$

   The spherical harmonics

$$Y_\ell^m(\theta,\phi) = P_\ell^m(\cos\theta)\, e^{im\phi} \tag{11.24}$$

form a complete basis of orthonormal functions. It can be shown that

$$\int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta d\theta\; [Y_\ell^m(\theta,\phi)]^*\, Y_{\ell'}^{m'}(\theta,\phi) = \delta_{\ell\ell'}\delta_{mm'} \tag{11.25}$$

assuming that the Y_ℓ^m have been suitably normalized.

For fixed ℓ, we have a set of $2\ell+1$ orthonormal functions. The $2\ell+1$ spherical harmonics Y_ℓ^m are related. As in the main text, one can form ladder operators.

$$\begin{aligned}\hat{L}_\pm &= \hat{L}_x \pm i\hat{L}_y \\ &= e^{\pm i\phi}\left(\pm\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\phi}\right) .\end{aligned} \tag{11.26}$$

Using known properties of the spherical harmonics, one can verify that

$$\hat{L}_+ Y_\ell^m = \begin{cases} \sqrt{\ell(\ell+1)-m(m+1)}\, Y_\ell^{m+1} & -\ell \leqslant m < \ell \\ 0 & m = \ell \end{cases} \tag{11.27}$$

and

$$\hat{L}_- Y_\ell^m = \begin{cases} \sqrt{\ell(\ell+1)-m(m-1)}\, Y_\ell^{m-1} & -\ell < m \leqslant \ell \\ 0 & m = -\ell \end{cases} . \tag{11.28}$$

11.3 Spin angular momentum

Many quantum particles interact with magnetic fields. A classical rotating charge acts like a magnetic dipole, with magnetic (dipole) moment $\vec{\mu}$. In a magnetic field $\vec{B}$, the interaction energy contributing to the Hamiltonian is $-\vec{\mu}\cdot\vec{B}$. Even though we cannot think about elementary particles spinning in the same way as classical particles, many of them interact similarly with magnetic fields.

The spin (angular momentum) operator $\vec{S}$ obeys the same commutation relations as $\vec{L}$, see (11.9). If the Hamiltonian commutes with $\vec{S}$, then we can find simultaneous eigenstates of $\{\,\hat{H}, |\hat{\vec{S}}|^2, \hat{S}_z\,\}$ denoted by $|n,s,s_z\rangle$ with

$$\begin{aligned}\hat{H}\,|n,s,s_z\rangle &= E_n\,|n,s,s_z\rangle \ , \\ |\hat{\vec{S}}|^2\,|n,s,s_z\rangle &= s(s+1)\,|n,s,s_z\rangle \ , \\ \hat{S}_z\,|n,s,s_z\rangle &= s_z\,|n,s,s_z\rangle \ .\end{aligned} \tag{11.29}$$

Experiments reveal that s takes on half-integer values $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. Accordingly the component of spin along, say, the z-axis $s_z \in [-s, -s+1, \ldots s-1, s]$.

The results of this Appendix follow from the derivation of $\mathfrak{su}(2)$ irreps given in § 4.2.

Bibliography

J Binney and D Skinner. *The Physics of Quantum Mechanics*. Oxford University Press, 2014. ISBN 978-0-19-968857-9.

G Costa and G Fogli. *Symmetries and Group Theory in Particle Physics*. Springer, 2012. ISBN 978-3-642-15481-2. URL https://link.springer.com/book/10.1007/978-3-642-15482-9.

L D Faddeev and O A Yakubovskiĭ. *Lectures on Quantum Mechanics for Mathematics Students*. American Mathematical Society, 2009. ISBN 978-0-8218-4699-5.

W Fulton and J Harris. *Representation Theory: A First Course*. Springer-Verlag, 1991. ISBN 0-387-97527-6. URL https://link.springer.com/book/10.1007/978-1-4612-0979-9.

S Gasiorowicz. *Quantum Physics*. Wiley, 1974.

B C Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2015. ISBN 978-3319134666. URL https://link.springer.com/book/10.1007/978-0-387-21554-9.

H F Jones. *Groups, Representations, and Physics*. IOP Publishing, 1990. ISBN 0-85274-029-8.

A W Knapp. *Lie Groups Beyond an Introduction*. Springer Science, 1996. ISBN 978-1-4757-2455-4. URL https://link.springer.com/book/10.1007/978-1-4757-2453-0.

P Ramond. *Group Theory: A Physicist's Survey*. Cambridge University Press, 2010. ISBN 978-0-521-89603-0. URL https://www.cambridge.org/core/books/group-theory/8BAC137A9F0C43D65448E6420248D841.

B Schutz. *Geometrical Methods of Mathematical Physics*. Cambridge University Press, 1980. ISBN 0-521-29887-3. URL https://doi.org/10.1017/CBO9781139171540.

S Weinberg. *The Quantum Theory of Fields I*. Cambridge University Press, 1995. ISBN 0-521-55001-7. URL https://www.cambridge.org/core/books/quantum-theory-of-fields/22986119910BF6A2EFE42684801A3BDF.

S Weinberg. *Lectures on Quantum Mechanics*. Cambridge University Press, 2013. ISBN 978-0-107-02872-2. URL https://www.

cambridge.org/core/books/lectures-on-quantum-mechanics/
F739B9577D2473995024FA5E9ABA9B6C.