Symmetries, Fields and Particles

Part III Michaelmas 2019

Lectures by Nick Dorey

Report typos to: uco21@cam.ac.uk
More notes at: uco21.user.srcf.net

Version: May 27, 2020

Contents

1 Introduction

In this course, we will cover Lie groups G, Lie algebras $\mathscr{L}(G)$, and their representations.

Resources

- Notes (online): Manton, Osborn, Gutowski (Cartan classification)
- Book: Fuchs & Schweigert (Ch 1-7) "Symmetries, Lie Algebras and Representations"

1.1 Symmetry and Groups

Definition 1: A *symmetry* is a transformation of dynamical variables that leaves the form of physical laws invariant.

Example 1.1.1 (Rotation)**:** Consider a vector $\mathbf{x} \in \mathbb{R}^3$. A rotation of this vector is a transformation

$$\mathbf{x} \to \mathbf{x}' = M \cdot \mathbf{x}, \tag{1.1}$$

where M is a 3×3 matrix, which is *orthogonal* ($MM^T = 1_3$) and *special* ($\det M = 1$). Applying a rotation to Newton's laws gives:

$$\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} \tag{1.2}$$

$$\to \mathbf{F}' = M \cdot \mathbf{F}' = m\frac{\mathrm{d}^2\mathbf{x}'}{\mathrm{d}t^2}, \tag{1.3}$$

so rotations of the coordinates are symmetries of Newtonian physics.

If two transformations individually leave a set of physical laws invariant, their combination also will. This property of symmetry transformations is captured in saying that they have a *group* structure:

Definition 2 (Group)**:** A group is a (finite or infinite) set G with the following properties:

1. Closure: $g_1, g_2 \in G \;\Rightarrow\; g_1 g_2 \in G$
2. Unit: $\exists e \in G : eg = ge = g \forall g \in G$
3. Inverse: $\forall g \in G, \exists g^{-1} \in G : g^{-1}g = g^{-1} = e$
4. Associativity: $\forall g_1, g_2, g_3 \in G : (g_1 g_2) g_3 = g_1 (g_2 g_3)$

In physics, we will in general deal with unconstrained, infinite groups. In general, say for rotations in three dimensions, $G = SO(3)$, the order in which we apply the symmetry transformations makes a difference, but there are some groups for which it does not.

Definition 3 (Abelian)**:** A commutative group G, where $g_1 g_2 = g_2 g_1$, is called *Abelian*.

Exercise 1.1: Show that the set $SO(3)$ of all 3×3 special orthogonal matrices, representing rotations, form a group under matrix multiplication.

Remark: A rotation in $\mathbb{R}^3$ depends continuously on 3 parameters $\hat{n} \in S^2$ and $\theta \in [0, \pi]$.

1.2 Lie Groups and Lie Algebras

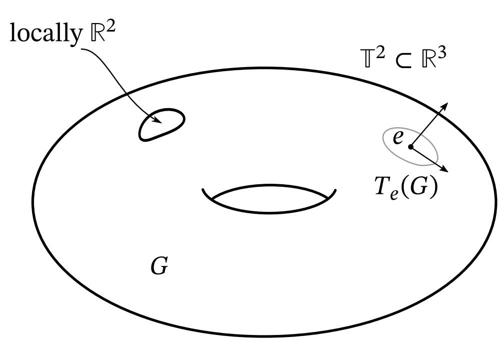


Figure 1.1: Locally, the torus $\mathbb{T}^2$ looks like $\mathbb{R}^2$

A *Lie group G* is a group which is also a *smooth manifold*. Due to the algebraic properties of the group each $g \in G$ defines a map from the group to itself

$$
\begin{aligned}
L_g : G &\to G \\
h &\longmapsto gh
\end{aligned}
\tag{1.4}
$$

Roughly speaking, this definition will require that group and manifold structures must be compatible (e.g. L_g must be smooth). As we will see in Chapter 4, this map will allow us to move around the manifold. As a result, G is almost completely determined by its behaviour "near" the identity e. In other words, the Lie group is almost completely defined by infinitesimal symmetry transformations.

Rather than the whole of the manifold, we will only think about the tangent space $T_e(G)$ of the identity.

Tangent vectors at e, $\mathbf{v}_1, \mathbf{v}_2 \in T_e(G)$, equipped with a bracket $[,] : T_e(G) \times T_e(G) \to T_e(G)$, define a *Lie algebra* $\mathscr{L}(G)$. Lie groups are (almost) determined by their Lie algebra.

1.3 Key Result: Cartan Classification (1895)

The Cartan Classification arises from the question of whether we can have a finite dimensional Lie algebra. The problem reduces to analysing *simple* Lie algebras, which all Lie algebras can be built from.

Theorem 1 (Cartan Classification)**:** All finite-dimensional semi-simple Lie algebras (over $\mathbb{C}$) belong to one of

- four infinite families A_n, B_n, C_n, D_n, where $n \in \mathbb{N}$,
- five exceptional cases E_6, E_7, E_8, G_2, F_4.

This basically lists the allowed gauge theories, since the gauge groups must correspond to the Lie algebras in this list. All physical groups come from the low-n values of the infinite families. In more modern theoretical physics, exceptional Lie groups also show up. For example, in String theory certain anomalies only cancel with combinations like $E_8 \oplus E_8$ or $\mathfrak{so}(32) = D_8$.

1.4 Symmetries in Quantum Physics

By Noether's theorem, symmetries imply the existence of conserved quantities.

Example 1.4.1: The invariance rotational symmetry in $\mathbb{R}^3$ implies the conservation of angular momentum $\mathbf{L} = (L_1, L_2, L_3)$.

Instead of the classical phase space, quantum mechanical

- states are vectors $|\psi\rangle$ in a (potentially infinite dimensional) Hilbert space $\mathcal{H}$,
- observables are linear operators $\hat{O} : \mathcal{H} \to \mathcal{H}$. These do not commute $\hat{O}_1\hat{O}_2 \neq \hat{O}_2\hat{O}_1$.

1.4.1 Angular Momentum

The angular momentum operators $\hat{L}_1, \hat{L}_2\hat{L}_3$ have the structure of the $\mathscr{L}(SO(3))$ lie algebra, where the commutator plays the role of the lie bracket:

$$[\hat{L}_i, \hat{L}_j] = i\hbar\varepsilon_{ijk}\hat{L}_k. \tag{1.5}$$

These operators form a basis of the tangent space $T_e(SO(3))$.

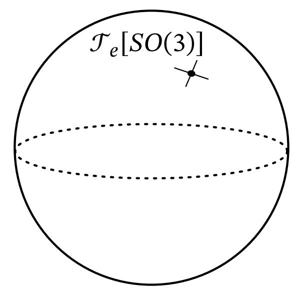


Figure 1.2: The angular momentum operators form a basis of the tangent space of $SO(3)$.

A particle with a definite spin defines a vector space with a certain dimension. The angular momentum operators in QM often act on finite dimensional vector spaces.

Definition 4 (Representation)**:** A *representation* of a Lie algebra $\mathscr{L}(G)$ is a map

$$R: \mathscr{L}(G) \to \mathrm{Mat}_n(\mathbb{C}) \tag{1.6}$$

which preserves the bracket of the Lie algebra.

> **Remark:** In the context of QM systems, we need to know both about the Lie algebra, but also about its representation in terms of finite dimensional matrices. Lie groups are largely determined by their Lie algebra. As seen in the last lecture, Lie algebras itself can be classified. One of the goals of this course will also be the classification of their representations.

Example 1.4.2 ($\mathbb{C}^2$)**:** The Hilbert space $\mathbb{C}^2$ is spanned by two vectors

$$|\uparrow\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \qquad |\downarrow\rangle = \begin{pmatrix}0\\1\end{pmatrix}. \tag{1.7}$$

On this representation space, we can build a two-dimensional representation of $\mathscr{L}(SO(3))$: In place of the angular momentum operators $\hat{L}_i$, we take three 2×2 matrices Σ_i, with $i = 1, 2, 3$, which obey the same commutation relations

$$[\Sigma_i, \Sigma_i] = i\hbar\Sigma_k, \tag{1.8}$$

analogous to (1.5). The Lie algebra is appearing again in a way in which its generators (basis vectors) are represented by the Pauli matrices σ_i:

$$\Sigma_i = \frac{1}{2}\hbar\sigma_i. \tag{1.9}$$

Recall that the Pauli matrices are

$$\sigma_x = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \qquad \sigma_y = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix}, \qquad \sigma_z = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}. \tag{1.10}$$

1.4.2 Degeneracy of the Spectrum

In Quantum mechanics, a rotational symmetry manifests itself in the commutation relation

$$\left[\hat{H}, \hat{L}_i\right] = 0 \tag{1.11}$$

with the Hamiltonian $\hat{H}$. The fact that the angular momentum generators, which move you around in the Hilbert space, commute with the Hamiltonian means that states in any representation of $\mathscr{L}(SO(3))$ have the same energy.

Example 1.4.3: The spin vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ have the same energy in a rotationally invariant system.

The degeneracies in the spectrum of a quantum system are effectively determined by the representations of the (global) symmetry group. This can be seen as a tool for which we do not yet know the underlying symmetry and the Lagrangian.

Example 1.4.4 (Approximate Symmetry of Hadrons)**:** Strongly interacting particles have an observed degeneracy which led Gell-Mann to postulate the approximate symmetry $G = SU(3)$ of 3×3 complex, unitary matrices with unit determinant. This is where group theory really took off in physics.

It was observed that there were sets of particles in the accelerators with approximately the same energy. Their interaction was characterised by conserved charges which could be assigned to integer values. The particular pattern is illustrated in Fig 1.3.

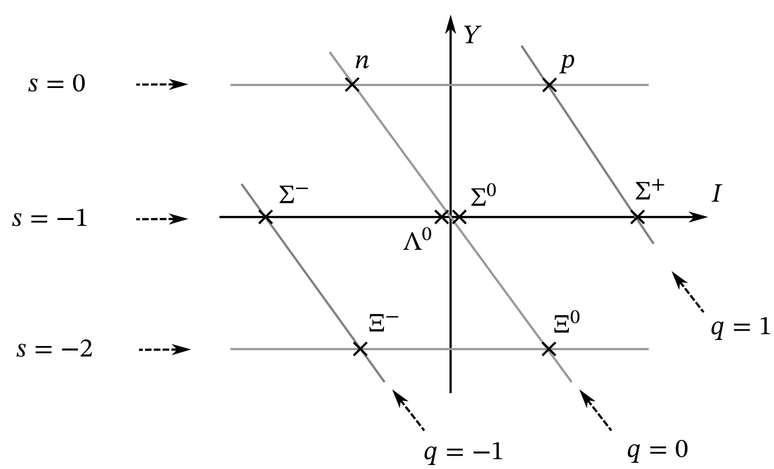


Figure 1.3: The *eightfold way* showing the approximate octet formed by the spin-$1/2$ baryons of strangeness s, electric charge q, hypercharge Y, and isospin I.

This turned out to be precisely explained by the mathematics of Lie groups and Lie algebras, and gives us the main tool of organising structures in particle physics.

1.4.3 Types of Symmetries

Global Symmetries: Global symmetries can be understood as operators in some Hilbert space which commute with the Hamiltonian. These include the spacetime symmetries, which fit into the pattern of non-Abelian Lie-groups: Rotations are modelled by $SO(3)$, whereas Lorentz transformations form the group $SO(3,1)$. Together with translations, these form the Poincaré group. However, the Poincaré group is not a simple Lie group.

Internal Symmetries: Moreover, we also have the *internal* symmetries like the $SU(3)$ *flavour symmetry*, which is the approximate symmetry of strong interactions. Advancements in particle physics have often hinged on the idea of enlarging the symmetry group. This is where the interaction between mathematics and physics also really took off: There are powerful theorems which prevent the combination of the global and internal symmetry groups, in the context of ordinary Lie algebras. Physicists then started to relax the constraints of Lie algebras, which led to *supersymmetry*.

Gauge Symmetries: The other topic which we will talk about is *Gauge symmetry*. This is not really a symmetry since it does not obey the definition in the first lecture. It is actually a *redundancy* in the mathematical description of the physics. Examples of gauge symmetries are

- phase of the wavefunction: $\psi \to e^{i\delta}\psi$
- electromagnetism: $\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\chi$, where χ is some arbitrary scalar function.

These transformations, constituting the Gauge group, do not affect any physical quantities.

> **Remark:** In QFT, as far as we know, only Gauge theories can describe in a consistent, renormalisable way the interaction of spin-1 particles.

The Standard Model is a particular type of Gauge theory with $G_{SM} = SO(3) \times SU(2) \times U(1)$.

2 Lie Groups

2.1 Manifold Structure and Coordinates

Definition 5 (Manifold)**:** A manifold $\mathcal{M}$ is a space that locally looks like Euclidean space. For each coordinate patch $\mathcal{P}$, there is a bijective map $\phi_{\mathcal{P}} : \mathcal{P} \leftrightarrow \mathbb{R}^n$. Moreover, the transition functions be-

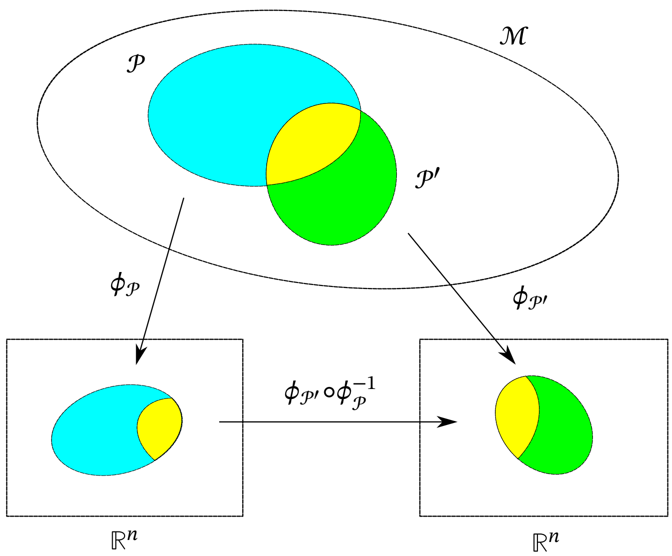


Figure 2.1: An illustration of the concept of a manifold.

tween the coordinates have to be *smooth*.

Definition 6 (Lie groups)**:** A *Lie group* G is a group that is also a manifold. The group operations must define *smooth maps* on the manifold. The *dimension* of the Lie group is the dimension of the manifold.

The definition of the manifold allows us to introduce coordinates $\{\theta^i\}$, $i = 1, \dots, D = \dim(G)$. In the patch $\mathcal{P}$, the group elements depend continuously on the coordinates $\{\theta^i\}$. WLOG, we can choose coordinates in which identity element lies at the origin: $g(0) = e$.

2.1.1 Compatibility of Group and Manifold Structures

Multiplication: The group operation of multiplication will define a map on the manifold. Since this map has to be smooth, it will have to be composed of continuous differentiable functions in the coordinate systems.

By closure of the group, and assuming that multiplication gives us another element in the same patch:

$$g(\theta)g(\theta') = g(\varphi) \in G \tag{2.1}$$

> **Remark:** In general it is not necessarily the case that the new group element will be in the same coordinate patch $\mathcal{P}$. In that case, we will have to make use of the transition functions.

This defines a map $G \times G \to G$ from a pair of group elements to a third. We can express this map in coordinates as $\varphi^i = \varphi^i(\theta, \theta') : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$. In terms of these coordinates, the condition of *compatibility* between a group operation and a manifold structure simply means that these maps φ^i have to be continuous and differentiable.

Inversion: Group inversion also defines a smooth map, this time from G to itself. For all elements $g(\theta) \in G$, where θ describes the position in the coordinate patch $\mathcal{P}$, the group axioms imply that there exists an inverse element $g^{-1}(\theta)$. Assuming that this is still in the same coordinate patch, we can write this as $g^{-1}(\theta) = g(\widetilde{\theta})$. If it is not, we simply apply the relevant transition function.

$$g(\theta)g(\widetilde{\theta}) = g(\widetilde{\theta})g(\theta) = e \tag{2.2}$$

In coordinates, $\widetilde{\theta}^i = \widetilde{\theta}^i(\theta)$ have to be continuous and differentiable.

Example 2.1.1: The simplest possible example of a Lie group is $G = (\mathbb{R}^D, +)$.

- operation: $\mathbf{x}'' = \mathbf{x} + \mathbf{x}'$ for all $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^D$
- inversion: $\mathbf{x}^{-1} = -\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^D$

This is an *Abelian* group.

2.1.2 Embedded Submanifolds

We can define properties of manifold in an intrinsic way. However, for some of these properties it is significantly easier to define manifolds as subspaces embedded in real space.

Theorem 2 (Embedding Theorem)**:** Consider the subspace of $\mathbb{R}^m$ defined as

$$\mathcal{M} = \{\mathbf{x} \in \mathbb{R}^m \mid \mathcal{F}_\alpha(\mathbf{x}) = 0\} \tag{2.3}$$

where the constraint $\mathcal{F}_\alpha : \mathbb{R}^m \to \mathbb{R}$, with $\alpha = 1, \cdots, l$ is a smooth map. Then $\mathcal{M}$ is a manifold of dimension $D = m - l$ if and only if the Jacobian matrix

$$(\mathcal{J})_{\alpha,i} = \frac{\partial \mathcal{F}_\alpha}{\partial x_i} \tag{2.4}$$

has maximal rank l everywhere on $\mathcal{M}$.

Example 2.1.2 (S^2)**:** We can realise the two-sphere S^2 as a manifold embedded in three dimensional Euclidean space

$$\mathcal{M} = \left\{\mathbf{x} \in \mathbb{R}^3 \mid |\mathbf{x}| = R\right\} \tag{2.5}$$

The solution space of $x^2 + y^2 + z^2 - R^2 = F$ defines a manifold. In this case, the Jacobian

$$\mathcal{J} = \left(\frac{\partial \mathcal{F}}{\partial x}, \frac{\partial \mathcal{F}}{\partial y}, \frac{\partial \mathcal{F}}{\partial z}\right) = 2(x, y, z) \tag{2.6}$$

has rank 1 except at $x = y = z = 0$, but that point is not contained in $\mathcal{M}$, so that is allowed by the theorem.

Definition 7 (connected)**:** A manifold is said to be *connected* if there is a smooth path between any two points on the manifold.

Definition 8 (simply connected)**:** A manifold is said to be *simply connected* if all loops are 'trivial', in the sense that they can be continuously contracted to a point.

Example 2.1.3: The spherical surface S^2 is simply connected, while the torus T^2 is not.

Definition 9 (compact)**:** An embedded manifold is *compact* if it is closed and bounded.[1]

Example 2.1.4: The sphere is a compact space, while the hyperboloid is not.

Definition 10 (submanifold)**:** A *submanifold* is a subspace of a manifold that is also a manifold.

Definition 11 (Lie subgroup)**:** A *Lie subgroup* is a subset of a Lie group which is also a Lie group.

2.2 Matrix Lie Groups

Matrix multiplication is closed, associative, and there exists a unit element $e = 1_n \in \mathrm{Mat}_n(F)$. Here, F can be any field, but we will mostly be working with $F = \mathbb{R}$ or $\mathbb{C}$. However, the set of matrices $\mathrm{Mat}_n(F)$ is not a group under matrix multiplication since not all matrices are *invertible*.

[1] This is the Heine–Borel theorem. On a general manifold we can define compactness from the underlying topology.

Definition 12 (general linear group)**:** The general linear group of dimension n is the set of $n \times n$ matrices with non-vanishing determinant

$$GL(n, F) = \{M \in \mathrm{Mat}_n(F) \mid \det M \neq 0\} \tag{2.7}$$

guaranteeing invertibility.

Definition 13 (special linear group)**:** The special linear group has unit determinant

$$SL(n, F) = \{M \in GL(n, F) \mid \det M = 1\} \tag{2.8}$$

These conditions are enough to guarantee that these are groups. In particular, the closure property follows from

$$\det(M_1 M_2) = \det(M_1)\det(M_2) \qquad \forall M_1, M_2 \in \mathrm{Mat}_n(F). \tag{2.9}$$

This is connected to the embedding theorem in the following way. Taking $SL(n, \mathbb{R})$ we apply the embedding theorem with $m = n^2$. The number of constraints is $l = 1$ due to the determinant being constraint to unity:

$$F_1(M) = \det M - 1 \tag{2.10}$$

To apply the embedding theorem we have to calculate the Jacobian. It is useful to recall the definition of a minor:

Definition 14 (minor)**:** Let $M \in \mathrm{Mat}_n(\mathbb{R})$ be an $n \times n$ matrix with real entries. We define the *minor* $\hat{M}^{(ij)}$ of each element of the matrix as the $(n-1)\times(n-1)$ matrix with i^{th} row and j^{th} column deleted.

Using this definition, we can differentiate (2.10) as follows

$$\frac{\partial F_1}{\partial M_{ij}} = \pm \det\big(\hat{M}^{(ij)}\big). \tag{2.11}$$

Therefore, the Jacobian $\frac{\partial F_1}{\partial M_{ij}}$ has rank 1, unless all the determinants of the minors vanish. This is equivalent to the determinant of the matrix vanishing

$$\det\big(\hat{M}^{(ij)}\big) = 0 \iff \det(M) = 0 \neq 1 \tag{2.12}$$

This shows that $SL(n, \mathbb{R})$ is a smooth manifold of dimension $n^2 - 1$. We could do this with $SL(n, \mathbb{C})$ by splitting the coordinates and then the detminant condition in its real and imaginary parts. For the general linear groups, we define them by removing a condition.

Exercise 2.1: Complete the proof that $SL(n, \mathbb{R})$ is a Lie group. For this, you have to convince yourself that matrix multiplication, considered element by element, provides a smooth map.

This gives us four families of matrix Lie groups. We have established that

$$\dim(SL(n,\mathbb{R})) = n^2 - 1. \tag{2.13}$$

A similar argument allows us to find that

$$\dim(SL(n,\mathbb{C})) = 2n^2 - 2. \tag{2.14}$$

Moreover, for the general linear group we can find that

$$\dim(GL(n,\mathbb{R})) = n^2, \qquad \dim(GL(2,\mathbb{C})) = 2n^2. \tag{2.15}$$

> **Remark:** We talk here about the *real dimension*, i.e. the dimension of the Lie group as embedded into the real manifold $\mathbb{R}^n$.

2.2.1 Compact Subgroups of $GL(n,\mathbb{R})$

The Orthogonal Groups

Definition 15 (orthogonal groups)**:** The *orthogonal groups* are defined to be those elements of the general linear groups which satisfy the orthogonality condition

$$O(n) = \{M \in GL(n,\mathbb{R}) \mid MM^T = 1_n\}. \tag{2.16}$$

Definition 16 (orthogonal transformations)**:** Orthogonal transformations are of the form

$$\mathbf{v} \in \mathbb{R}^n \to \mathbf{v}' = M \cdot \mathbf{v} \in \mathbb{R}^n \tag{2.17}$$

where $M \in O(n)$ is an orthogonal matrix.

Orthogonal transformations can be thought of as linear transformations which preserve the length of vectors since

$$\left|\mathbf{b}'\right| = \mathbf{v}'^T \cdot \mathbf{v}' = \mathbf{v} \cdot M^T M \cdot \mathbf{v} = \mathbf{v}^T \cdot \mathbf{v} = |\mathbf{v}|^2. \tag{2.18}$$

If we have an orthogonal real matrix, we have

$$\det(MM^T) = \det(M)^2 = 1, \tag{2.19}$$

so its determinant is $\det(M) = \pm 1$.

From this, by continuity, we can tell that $O(n)$ is not a connected manifold. Indeed, $O(n)$ must have two *connected components*. Moreover, only one of these can contain the identity element. As such, we can consider the space which only includes the connected components of the identity:

Definition 17 (special orthogonal group)**:** The *special orthogonal groups* $SO(n)$ are defined to be the subsets of $O(n)$ with positive unit determinant:

$$SO(n) := \{M \in O(n) \mid \det(M) = 1\}. \tag{2.20}$$

How do we distinguish between matrices which have $\det(M) = \pm 1$? Given a frame $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\} \in \mathbb{R}^n$, an orthogonal transformation

$$\mathbf{v}_a \in \mathbb{R}^n \to \mathbf{v}'_a = M \cdot \mathbf{v}_a \in \mathbb{R}^n, \qquad M \in O(n) \tag{2.21}$$

preserves the *orientation* of a frame, i.e. the sign of the volume element

$$\Omega = \varepsilon^{i_1,\dots,i_n} v_1^{i_1} v_2^{i_2} \cdots v_n^{i_n}, \tag{2.22}$$

where ε is the n-dimensional alternating tensor, only if $M \in SO(n)$. Elements of $SO(n)$ correspond to *rotations*, whereas elements of $O(n)$ with $\det(M) = -1$ correspond to some mixture of *rotation* and *reflection*.

Exercise 2.2: Use the embedding theorem to check that $O(n)$ is a manifold and that its dimension is

$$\dim[O(n)] = \dim[SO(n)] = \frac{1}{2}n(n-1). \tag{2.23}$$

Remember to show that the Jacobian matrix has maximal rank.

Claim 1: Consider the matrix eigenvalue equation

$$M\mathbf{v}_\lambda = \lambda \mathbf{v}_\lambda, \tag{2.24}$$

where $M \in O(n)$, the two defining properties of the orthogonal group are then that

1. If λ is an eigenvalue, then its complex conjugate λ^* is an eigenvalue as well.
2. The eigenvalues are normalised such that $|\lambda|^2 = 1$.

Proof. 1. Complex conjugating both sides of (2.24) gives

$$M\mathbf{v}_\lambda^* = \lambda^* \mathbf{v}_\lambda^* \tag{2.25}$$

2. First, note that we have $(M\mathbf{v}^*)^T \cdot M\mathbf{v} = \mathbf{v}^\dagger M^T M \mathbf{v} = \mathbf{v}^\dagger \cdot \mathbf{v}$. Then, if $\mathbf{v} = \mathbf{v}_\lambda$:

$$(M\mathbf{v}_\lambda^*)^T \cdot M\mathbf{v}_\lambda = |\lambda|^2 \mathbf{v}_\lambda^\dagger \mathbf{v}_\lambda = \mathbf{v}_\lambda^\dagger \mathbf{v}_\lambda \;\Rightarrow\; |\lambda|^2 = 1. \tag{2.26}$$

□

Example 2.2.1 ($SO(2)$)**:** Let M be a matrix in $SO(2)$. Then M has eigenvalues $\lambda = e^{i\theta}, e^{-i\theta}$ for small $\theta \in \mathbb{R}$, with the identification $\theta \sim \theta + 2\pi$. In a matrix representation, we write

$$M = M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{2.27}$$

Although this is a real matrix, its eigenvalues are complex. Provided that we made the identification $\theta \sim \theta + 2\pi$, the matrix is uniquely specified by θ. Therefore, the manifold of this Lie group is $M(SO(2)) \cong S^1$. Moreover, since the matrices are commutative, $M(\theta_1)M(\theta_2) = M(\theta_2)M(\theta_1) = M(\theta_1 + \theta_2)$, this is an Abelian Lie group.

> **Remark:** This is in fact the simplest compact Lie group.

Example 2.2.2 ($SO(3)$)**:** We consider now matrices M in the three-dimensional special orthogonal group $SO(3)$. The eigenvalues are $\lambda = e^{+i\theta}, e^{-i\theta}, +1$, where we again have made the identification $\theta \sim \theta + 2\pi$. To parametrise a rotation matrix in three dimensions, consider the normalised eigenvector corresponding to the $\lambda = +1$ eigenvalue:

$$\hat{\mathbf{n}} \in \mathbb{R}^3, \qquad M\hat{\mathbf{n}} = \hat{\mathbf{n}}, \qquad \hat{\mathbf{n}} \cdot \hat{\mathbf{n}} = 1. \tag{2.28}$$

The direction of $\hat{\mathbf{n}}$ parametrises the axis, and θ parametrises the angle of rotation.

Exercise 2.3: One can write a general group element of SO(3) as

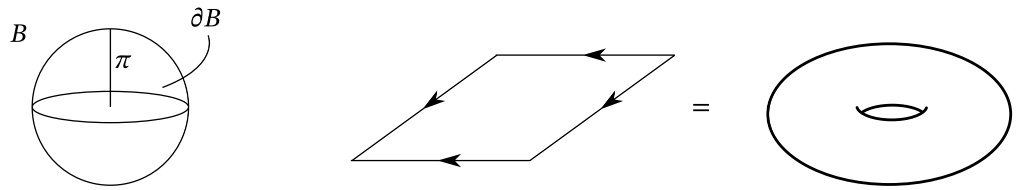
$$M(\hat{\mathbf{n}}, \theta)_{ij} = \cos\theta\delta_{ij} + (1 - \cos\theta)n_i n_j - \sin\theta\varepsilon_{ijk}n_k. \tag{2.29}$$

We want to specify the elements uniquely. Above, one needs to be careful about the uniqueness, due to two issues:

1. Identification: $M(\hat{\mathbf{n}}, 2\pi - \theta) = M(-\hat{\mathbf{n}}, \theta)$
2. If $\theta = 0$, then for all directions $\hat{\mathbf{n}}$, we have $M(\hat{\mathbf{n}}, 0) = I_3$,

To be precise, we need to identify these rotations. To get a better parametrisation, define the parameter $\boldsymbol{\omega} = \theta\hat{\mathbf{n}}$. Consider the ball $B_3 \subset \mathbb{R}^3 = \{\boldsymbol{\omega} \in \mathbb{R}^3 \mid |\boldsymbol{\omega}| \leqslant \pi\}$. The group manifold associated with SO(3) is obtained by taking B_3 and identifying antipodal points on the boundary.

> **Remark:** In general, freely acting quotients give a manifold. Here, the group we quotiented out is the group of inversion $\mathbb{R}_2$.



(a) The ball B_3 with radius π.

(b) Identifying opposing edges of a sheet of paper gives a space topologically equivalent to a torus.

Figure 2.2: The group manifold associated with SO(3) is obtained by identifying antipodal points on the boundary ∂B of the ball B_3.

The resulting manifold is connected, but not simply connected. This is because loops that come out and back via the identification cannot be contracted to a point; antipodal points are always antipodal. This is illustrated in Figure 2.3. As such, we have $\pi_1(SO(3)) \neq \{0\}$. In fact, the fundamental group is

$$\pi_1(SO(3)) \simeq \mathbb{Z}_2 = \{+1, -1\}\,. \tag{2.30}$$

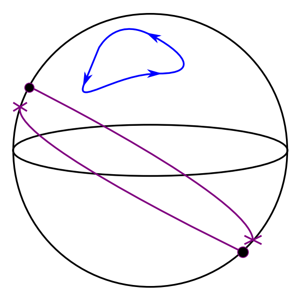


Figure 2.3: Loops passing through the identification of antipodal points cannot be contracted to a point. Note that the purple loop is constructible. This can be seen by "rotating" one half of the loop until it matches up with the other.

The Unitary Groups

Definition 18 (unitary group)**:** The *unitary group* $U(n)$ is the set of invertible complex matrices obeying $U^\dagger = U^{-1}$:

$$U(n) := \left\{U \in GL(n, \mathbb{C}) \mid U^\dagger U = I_n\right\}. \tag{2.31}$$

Claim 2: Let $U \in U(n)$ be a matrix in the unitary group. Under such a unitary transformation, the length of a vector $\mathbf{v}$ is unchanged.

Proof. The vector $\mathbf{v}$ transforms as $\mathbf{v} \in \mathbb{C}^n \to \mathbf{v}' = U\mathbf{v} \in \mathbb{C}^n$. Using the property $UU^\dagger = 1_n$ of unitary matrices, we have

$$\left|\mathbf{v}'\right|^2 = \mathbf{v}'^\dagger \cdot \mathbf{v}' = (\mathbf{v}^\dagger U^\dagger) \cdot (U\mathbf{v}) = \mathbf{v}^\dagger \cdot \mathbf{v} = |\mathbf{v}|^2 \tag{2.32}$$

□

Claim 3: Let $U \in U(n)$ be an element of the group of unitary $n \times n$ matrices. Then $\det U = e^{i\delta}$, where $\delta \in \mathbb{R}$.

Proof. $U^\dagger U = 1_n \;\Rightarrow\; |\det U|^2 = 1 \;\Rightarrow\; \det U = e^{i\delta}, \delta \in \mathbb{R}$. □

> **Remark:** Since $\delta \in \mathbb{R}$ is able to vary continuously, $U(n)$ is connected, whereas $O(n)$ was not.

Definition 19 (special unitary group)**:** The *special unitary group $SU(n)$* is the subset of $U(n)$ with unit determinant:

$$SU(n) = \{U \in U(n) \mid \det U = 1\} \tag{2.33}$$

Claim 4: The groups $U(n)$ and $SU(n)$ are indeed Lie groups $\subset GL(n, \mathbb{C})$. Their dimensions are given by

$$\dim(U(n)) = 2n^2 - n^2 = n^2, \qquad \dim(SU(n)) = n^2 - 1. \tag{2.34}$$

Proof. To see this, take any matrix $M \in \mathrm{Mat}(n, \mathbb{C}) \leftrightarrow \mathbb{R}^{2n^2}$ (real and complex components) and apply the embedding theorem. The constraint for $U(n)$ is $\mathcal{F} = UU^\dagger - I = 0$. This gives quadratic constraints in the matrix elements. Since $H = UU^\dagger$ is Hermitian, these are actually n^2 constraints instead of $2n^2$. For $SU(n)$, we require $\mathcal{F} = \det U - 1 = 0$. Since $\det U = e^{i\varphi}$, this is only one additional constraint. □

2.2.2 Non-Compact Subgroups of $GL(n, \mathbb{R})$

Orthogonal matrices obey $MM^T = I_n$. We can also read this as $MI_nM^T = I_n$; orthogonal transformations preserve the Euclidean metric $g = \mathrm{diag}\underbrace{(1, \cdots, 1)}_{n \text{ times}}$ on $\mathbb{R}^n$.

We can generalise this by defining $O(p, q)$ to be the set of transformations that preserve the metric of signature p, q :

$$O(p, q) = \left\{M \in GL(n, \mathbb{R}) \mid M^T\eta M = \eta, \text{where } \eta = \mathrm{diag}(\underbrace{-1, \dots, -1}_{p \text{ times}}, \underbrace{+1, \dots, +1}_{q \text{ times}})\right\}. \tag{2.35}$$

In general, the manifolds of $O(p, q)$ are not compact.

Example 2.2.3: The elements of $SO(2)$ are matrices of the form $\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$, with the identification $\theta \sim \theta + 2\pi$. This identification means that as manifolds, $SO(2) \simeq S^1$, which is compact.

In contrast to this, elements of $SO(1,1)$ are matrices of the form $\begin{pmatrix}\cosh\varphi & -\sinh\varphi\\ \sinh\varphi & \cosh\varphi\end{pmatrix}$, where φ is free to take any value $\varphi \in \mathbb{R}$. This means that the manifold of $SO(1,1)$ is $\mathbb{R}$, which is evidently non-compact.

2.3 Isomorphisms

In this course, we are interested in classifying Lie groups and Lie algebras *up to isomorphism*. These isomorphisms are maps which preserve both the group and manifold structure respectively.

Definition 20 (isomorphism)**:** Two Lie groups G and G' are *isomorphic* ($G \simeq G'$) if there exists a one-to-one smooth map $J : G \to G'$ such that for all $g_1, g_2 \in G$, we have $J(g_1 g_2) = J(g_1)J(g_2)$.

Let us look at some low-dimensional examples of unitary groups:

Example 2.3.1 ($G = U(1)$)**:** Let $U(1) = \{z \in \mathbb{C} \mid |z| = 1\}$. A general element $z = e^{i\theta} \in G = U(1)$ is parametrised by $\theta \in \mathbb{R}$ with identification $\theta \sim \theta + 2\pi$. Now we define the map

$$\begin{aligned} J:\ U(1) &\ \to\ SO(2) \\ z(\theta) &\ \longmapsto\ M(\theta), \end{aligned} \tag{2.36}$$

where

$$M(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \in G' = SO(2). \tag{2.37}$$

The map J is one-to-one and

$$J[z(\theta_1)z(\theta_2)] = M(\theta_1 + \theta_2) = M(\theta_1)M(\theta_2) = J(z(\theta_1))J[z(\theta_2)]. \tag{2.38}$$

This means that J is an isomorphism and

$$U(1) \simeq SO(2). \tag{2.39}$$

Example 2.3.2 ($G = SU(2)$)**:** This is a three-dimensional group. The matrix parametrised as $U = a_0 I_2 + i\mathbf{a}\cdot\boldsymbol{\sigma}$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ and $a_0 \in \mathbb{R}$, $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$, is an element of $SU(2)$ provided that

$$a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1. \tag{2.40}$$

This implies that the underlying manifold of $SU(2)$ is $S^3 \subset \mathbb{R}^4$. The fundamental group of this manifold is

$$\pi_1(SU(2)) = \{1\} \tag{2.41}$$

whereas we had $\pi_1(SO(3)) = \mathbb{Z}_2$ from (2.30). This means that the two Lie groups cannot be isomorphic:

$$SU(2) \not\cong SO(3). \tag{2.42}$$

3 Lie Algebras

Definition 21 (Lie algebra)**:** A *Lie algebra* $\mathfrak{g}$ is a vector space over a field $F = \mathbb{R}, \mathbb{C}$ with a bracket,

$$[\bullet, \bullet]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{3.1}$$

with the following properties for all $X, Y, Z \in \mathfrak{g}$, and for all coefficients $\alpha, \beta \in F$:

1. Anti-symmetry: $[X, Y] = -[Y, X]$,
2. (Bi-)linearity: $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$,
3. Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Let V be a vector space that has a product $*: V \times V \to V$, which is associative, meaning that for all $X, Y, Z \in V$,

$$(X * Y) * Z = X * (Y * Z). \tag{3.2}$$

Moreover, the product is distributive over the field

$$Z * (\alpha X + \beta Y) = \alpha Z * X + \beta Z * Y. \tag{3.3}$$

Then, we obtain a Lie algebra by defining the Lie bracket to be the commutator

$$[X, Y] = X * Y - Y * X. \tag{3.4}$$

One example we have in mind is the case where V is the vector space of matrices and $*$ is matrix multiplication.

> **Remark:** Compare this to the Lie algebra of differential operators/vectors in differential geometry.

Definition 22 (dimension)**:** The dimension of a Lie algebra $\mathfrak{g}$ is the dimension of its underlying vector space.

Choose a basis B for $\mathfrak{g}$

$$B = \{T^a, a = 1, \dots, n = \dim(\mathfrak{g})\}. \tag{3.5}$$

Then any $X \in \mathfrak{g}$ can be written as

$$X = X_a T^a := \sum_{a=1}^{n} X_a T^a \tag{3.6}$$

where $X_a \in F$. The bracket of two Lie algebra elements $X, Y \in \mathfrak{g}$ can then be written as

$$[X, Y] = X_a Y_b [T^a, T^b]. \tag{3.7}$$

Therefore, knowing the brackets of basis elements allows us to construct the full Lie algebra. We write

$$[T^a, T^b] = f^{ab}{}_c T^c. \tag{3.8}$$

The *structure constants* $f^{ab}{}_c \in F$ therefore define the Lie algebra, and two Lie algebras are isomorphic if they have the same structure constants. Note however, that structure constants are basis dependent. We will eventually want to find a way to classify Lie algebras that is independent of our choice of basis. Let $f^{ab}{}_c \in F$, $a, b, c = 1, \dots, \dim(\mathfrak{g})$ be structure constants of a Lie algebra $\mathfrak{g}$. The axioms of Lie algebras then imply

Property 1. $\Rightarrow$ $f^{ab}{}_c = -f^{ba}{}_c$

Property 3. $\Rightarrow$ $f^{ab}{}_c f^{cd}{}_e + f^{da}{}_c f^{cb}{}_e + f^{bd}{}_c f^{ca}{}_e = 0$

Definition 23 (isomorphism)**:** Two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ are said to be *isomorphic*, denoted $\mathfrak{g} \simeq \mathfrak{g}'$, if there exists a linear, one-to-one map $f : \mathfrak{g} \to \mathfrak{g}'$ such that $\forall X, Y \in \mathfrak{g}$,

$$[f(X), f(Y)] = f([X, Y]). \tag{3.9}$$

We will be interested in classifying Lie algebras *up to isomorphism*.

Definition 24 (subalgebra)**:** A *subalgebra* $\mathfrak{h} \subset \mathfrak{g}$ is a vector subspace of $\mathfrak{g}$ that is also a Lie algebra.

Definition 25 (ideal)**:** An *ideal* of $\mathfrak{g}$ is a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ such that $\forall X \in \mathfrak{g}$ and $\forall Y \in \mathfrak{h}$,

$$[X, Y] \in \mathfrak{h}. \tag{3.10}$$

The notion of ideal roughly corresponds to the concept of a normal subgroup.

Example 3.0.1: Every Lie algebra $\mathfrak{g}$ has two *'trivial'* ideals:

$$\mathfrak{h} = \{0\} \qquad \text{and} \qquad \mathfrak{h} = \mathfrak{g}. \tag{3.11}$$

Example 3.0.2: The *derived algebra*

$$i = [\mathfrak{g}, \mathfrak{g}] \coloneqq \mathrm{Span}_F \{[X, Y] \mid X, Y \in \mathfrak{g}\} \tag{3.12}$$

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

Example 3.0.3: The *centre* $\xi(\mathfrak{g})$ of $\mathfrak{g}$, defined by

$$\xi(\mathfrak{g}) = \{X \in \mathfrak{g} \mid \forall Y \in \mathfrak{g},\ [X, Y] = 0\}, \tag{3.13}$$

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

Definition 26: An *Abelian* Lie algebra is such that all brackets vanish:

$$[X, Y] = 0 \qquad \forall X, Y \in \mathfrak{g} \tag{3.14}$$

For an Abelian Lie algebra, the Lie algebra is equal to its own centre $\mathfrak{g} = \xi(\mathfrak{g})$ and the derived algebra is trivial $i(g) = \{0\}$.

Definition 27 (simple)**:** A Lie algebra $\mathfrak{g}$ is said to be *simple* if it is non-Abelian and it has no non-trivial ideals.

This implies that for simple Lie algebras, $\xi(\mathfrak{g}) = \{0\}$ and $i(\mathfrak{g}) = \mathfrak{g}$.

The main theorem that we will work up to is the *Cartan classification*, which will allow us to classify all finite-dimensional, simple, complex Lie algebras $\mathfrak{g}$.

4 Lie Algebras from Lie Groups

So far, we have introduced the concepts of Lie group and Lie algebras. These appear to be very different objects. We will now relate these two concepts by showing that every Lie group gives rise to a Lie algebra, by considering the tangent space near its identity element.

4.1 Preliminaries

Let M be a smooth manifold of dimension D. Pick a point $p \in M$. Because of the manifold structure, we can introduce a coordinate chart $\{x^i \in \mathbb{R}\}$, with $i = i, \ldots, D$, in some region $\mathcal{P} \subset M$. Without loss of generality, we can choose the coordinates such that the point p corresponds to the origin $x^i = 0$, $\forall i$.

Definition 28: The *tangent space* T_pM to M at p is a D-dimensional vector space spanned by differential operators $\{\partial/\partial x^j\}$ acting on functions $f : M \to \mathbb{R}$. An element of the tangent space is called a *tangent vector*. We can expand every tangent vector V as a linear combination of the differential operators

$$V = v^i \frac{\partial}{\partial x^i} \in T_pM, \tag{4.1}$$

where the real coefficients $v^i \in \mathbb{R}$ are called the *components* of V.

This decomposition tells us that tangent vectors act on a function $f = f(x)$ as

$$V \cdot f = v^i \left. \frac{\partial f(x)}{\partial x^i} \right|_{x=0} \tag{4.2}$$

where the evaluation at $x = 0$ signifies that we are dealing with a tangent space at point p, which is at the origin of our coordinates.

Definition 29 (smooth curve)**:** A *smooth curve* C on a manifold M is a smooth map from an interval $I \in \mathbb{R}$ to M. In coordinates, we can use a parameter $t \in I$ to parametrise the curve as

$$C : t \longmapsto x^i(t). \tag{4.3}$$

We will work with curves in which the coordinates $\{x^i(t)\}$ are continuous and differentiable at least once. Moreover, we choose coordinates in which $x^i(0) = 0$, $\forall i = 1, \ldots, D$.

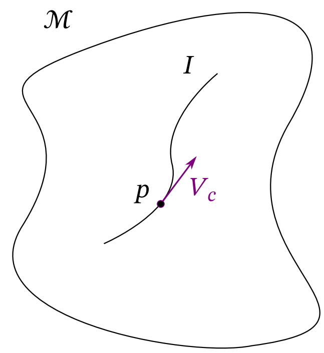


Figure 4.1: Tangent vector V_c to curve $C : I \to M$ at point $p \in M$.

Now consider a curve $C : I \to M$ passing through the point p at $x^i(0) = 0$. The tangent vector to C at p is

$$V_c = \dot{x}^i(0)\frac{\partial}{\partial x^i} \in T_pM, \tag{4.4}$$

where the derivatives with respect to the parameter t are denoted $\dot{x}^i(t) = \mathrm{d}x^i(t)\big/\mathrm{d}t$. Intuitively, in a physical picture where C is the trajectory of the particle, parametrised by a time coordinate t, the tangent vector $V_c \in T_pM$ corresponds to the velocity vector at $t = 0$. Every smooth curve has a tangent vector at every point it passes through. This construction will also work for the end points of the parameter interval.

4.2 The Lie Algebra $\mathscr{L}(G) = \mathfrak{g}$

Let G be a Lie group of dimension D. Recall that Lie groups are also manfolds; we will apply the definitions above to examine certain tangent spaces on the manifold structure of G. Let $\{\theta^i\}$, with $i = 1, \ldots, D$, be a set of coordinates in some coordinate patch $\mathcal{P}$ containing the identity $e \in \mathcal{P} \subset G$. These coordinates allow us to parametrise any element $g \in \mathcal{P}$ as $g = g(\theta) \in G$. The statement that the coordinates are taken to be centered at the identity e translates to $g(0) = e$.

Claim 5: Let G be a Lie group and consider the tangent space T_eG at the identity element $e \in G$. We can define a Lie bracket operation

$$[\cdot,\cdot] : T_eG \times T_eG \to T_eG \tag{4.5}$$

so that the tangent space becomes a Lie algebra $\mathscr{L}(G) = \{T_eG, [\cdot,\cdot]\}$ when we equip it with this bracket structure.

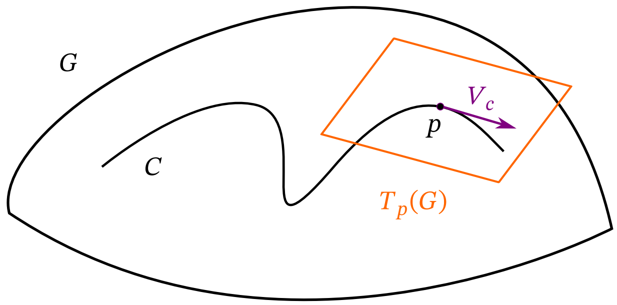


Figure 4.2: Tangent space T_pG at point p in the manifold G.

Proof. It is evident that the tangent space T_eG at the identity is a D-dimensional vector space. We have to find a suitable bracket that satisfies Definition 21 and makes this tangent space into a Lie algebra. Showing this is easiest for matrix Lie groups. Let $G \subset \mathrm{Mat}_n(F)$ be a subspace of $n \in \mathbb{N}$ dimensional matrices over the field $F = \mathbb{R}$ or $\mathbb{C}$. We can map tangent vectors to matrices by constructing the following map:

$$\begin{aligned} f : T_eG &\to \mathrm{Mat}_n(F) \\ v^i \frac{\partial}{\partial \theta^i} &\longmapsto v^i \left.\frac{\partial g(\theta)}{\partial \theta^i}\right|_{\theta=0} . \end{aligned} \tag{4.6}$$

The map f is injective and linear. This allows us to identify the tangent space T_eG with the subset of $\mathrm{Mat}_n(F)$ spanned by $\left\{ \partial g(\theta)/\partial\theta^i \big|_{\theta=0} \,\middle|\, i = 1, \dots, D \right\}$. Due to this identification with matrices, the *matrix commutator*, defined for any two $X, Y \in T_eG$ as

$$[X, Y] := XY - YX, \tag{4.7}$$

provides an obvious candidate for the bracket. It is easy to check that the defining properties of Definition 21 of the Lie algebra hold. Note in particular that the Jacobi identity follows from the associativity of matrix multiplication.

It remains to show *closure*: For any two tangent vectors X, Y, we need to show that their commutator $[X, Y]$ is itself also a tangent vector. To achieve this, we use the correspondence between tangent vectors and *curves* on a manifold. Let $C : I \to G$ be a smooth curve passing through the identity e.

$$\begin{aligned} C : I \subset \mathbb{R} &\to G \\ t &\longmapsto g(t) \end{aligned} \tag{4.8}$$

The parameter is chosen in such a way that $t = 0$ parametrises the identity matrix: $g(0) = I_N$. By the chain rule, we can differentiate a group element g as

$$\frac{\mathrm{d}g(t)}{\mathrm{d}t} = \frac{\mathrm{d}\theta^i(t)}{\mathrm{d}t} \frac{\partial g(\theta)}{\partial \theta^i}. \tag{4.9}$$

Consider the derivative at the origin:

$$\dot{g}(0) = \left.\frac{\mathrm{d}g(t)}{\mathrm{d}t}\right|_{t=0} = \dot{\theta}^i(0) \left.\frac{\partial g(\theta)}{\partial \theta^i}\right|_{\theta=0} \in T_e G. \tag{4.10}$$

This is the tangent vector to the curve C at e. We also have an explicit representation $\dot{g}(\theta) \in \mathrm{Mat}_n(F)$ for the matrix Lie group. However, they are not in general elements of the Lie group G.

Near $t = 0$, we have a Taylor expansion

$$g(t) = I_n + Xt + O(t^2) \tag{4.11}$$

where the term X appearing in the first order expansion is a tangent vector $X = \dot{g}(0) \in \mathscr{L}(G)$.

Given two elements $X_1, X_2 \in \mathscr{L}(G)$, we can find smooth curves $C_1 : t \longmapsto g_1(t) \in G$ and $C_2 : t \longmapsto g_2(t) \in G$, passing through the origin $g_1(0) = g_2(0) = I_n$, such that $\dot{g}_1(0) = X_1$ and $\dot{g}_2(0) = X_2$.

Near $t = 0$, another way of saying what we just said is that

$$g_1(t) = I_n + X_1 t + W_1 t^2 + O(t^3), \qquad g_2(t) = I_n + X_2 t + W_2 t^2 + O(t^3) \tag{4.12}$$

for some $W_1, W_2 \in \mathrm{Mat}_n(F)$.

Define a new curve

$$h(t) = g_1^{-1}(t) g_2^{-1}(t) g_1(t) g_2(t) \in G. \tag{4.13}$$

Since this is a composition of smooth maps, h is itself smooth. Equivalently, we have $\forall t \in I$

$$g_1(t) g_2(t) = g_2(t) g_1(t) h(t). \tag{4.14}$$

Expanding this near $t = 0$ in steps,

$$g_1(t) g_2(t) = I_n + (X_1 + X_2)t + (X_1 X_2 + W_1 + W_2)t^2 + O(t^3) \tag{4.15}$$

$$g_2(t) g_1(t) = I_n + (X_1 + X_2)t + (X_2 X_1 + W_1 + W_2)t^2 + O(t^3). \tag{4.16}$$

Here we can already see the terms $X_1 X_2$ and $X_2 X_1$ that we will want to isolate for our Lie bracket.

Now $h(t)$ has a Taylor series of the form

$$h(t) = I_n + h_1 t + h_2 t^2 + O(t^3). \tag{4.17}$$

Plugging this into (4.14) and using the above expansions for $g_1 g_2$ and $g_2 g_1$, we find that the coefficients in the Taylor series expansion are given by

$$h_1 = 0 \qquad h_2 = (X_1 X_2 - X_2 X_1) = [X_1, X_2]. \tag{4.18}$$

Thus, the curve h_3 is written in its Taylor series expansion as

$$h(t) = I_n + t^2[X_1, X_2] + O(t^3). \tag{4.19}$$

However, a curve should have its tangent vector in the linear term of its parameter. To have a sneeky way around this issue, we redefine our parameter $s = t^2 \geq 0$, where e lies at the end point. This defines a new curve $C_3 : s \longmapsto g_3(s) = h(\sqrt{s})$, which near $s = 0$ has the form

$$g_3(s) = I_n + s[X_1, X_2] + O(s^{3/2}). \tag{4.20}$$

This is a good curve in C^1, provided that $s \geq 0$. Finally, we can isolate the commutator by taking the derivative and evaluating it at $s = 0$:

$$[X_1, X_2] = \left.\frac{\mathrm{d}g_3(s)}{\mathrm{d}s}\right|_{s=0} = \dot{g}_3(0) \in \mathscr{L}(G) \tag{4.21}$$

This demonstrates the closure property of the Lie algebra, since the bracket of two tangent vectors gives another. □

Note that the second derivative of g_3 will give us a negative power, and as $s \to 0$, the second derivative would blow up. So $\ddot{g}_3(0)$ does not exist. As slightly longer proof can construct good C^n curves for any n, but $n = 1$ is sufficient for this proof.

Example 4.2.1 (G = SO(n))**:** Let $g(t) = R(t) \in SO(n)$ for all $t \in \mathcal{I} \subset \mathbb{R}$, with $R(0) = I_n$ be a curve on the group manifold $G = SO(n)$. This implies that

$$R^T(t)R(t) = I_n \quad \forall t \in \mathcal{I}. \tag{4.22}$$

Differentiating with respect to t, we have

$$\dot{R}^T(t)R(t) + R^T(t)\dot{R}(t) = 0, \qquad \forall t \in \mathcal{I}. \tag{4.23}$$

If $x_i = \dot{R}(0)$, then at $t = 0$ we get $x^T + x = 0$. Since continuity automatically implies $\det R(t) = +1$, there is no further constraint from imposing this. Therefore,

$$\mathfrak{o}(n) \simeq \mathfrak{so}(n) = \{X \in \mathrm{Mat}_n(\mathbb{R}) \mid X^T = -X\}. \tag{4.24}$$

Counting the parameters in these matrices, we then have

$$\dim(\mathfrak{so}(n)) = \frac{1}{2}n(n-1) = \dim(SO(n)). \tag{4.25}$$

Example 4.2.2 (G = SU(n))**:** Let $g(t) = U(t) \in SU(n)$ be a curve passing through the origin, with $U(0) = I_n$. As in the previous example, we differentiate the unitarity property $U^\dagger(t)U(t) = I_n$ with respect to t and evaluate at $t = 0$ to give

$$Z^\dagger + Z = 0 \tag{4.26}$$

for the derivatives $Z = \dot{U}(0) = \mathfrak{su}(n)$. Here, in contrast to the previous example, we *will* get an additional constraint from $\det U(t) = 1 \; \forall t \in \mathbb{R}$.

Exercise 4.1: Show that $\det U(t) = 1 + t\,\mathrm{Tr}\,Z + O(t^2)$

Therefore, the constraint on the determinant imposes that the traces $\mathrm{Tr}\,Z$ vanish. The higher order terms in the Taylor series will not impose any additional constraints. Hence, we can use the same Lie algebra as $SO(n)$, except that we have to constrain the matrices be traceless:

$$\mathfrak{su}(n) = \{Z \in \mathrm{Mat}_n(\mathbb{C}) \mid Z^\dagger = -Z,\ \mathrm{Tr}\,Z = 0\} \tag{4.27}$$

From this, we find that $\dim(\mathfrak{su}(n)) = 2n^2 - n^2 - 1 = n^2 - 1$; the dimension of the Lie algebra is the same as the dimension of the Lie group $\dim(SU(n))$.

Example 4.2.3:

$$\dim SU(2) = \dim SO(3) = 3 \tag{4.28}$$

Let $G = SU(2)$, then from our considerations above we know that the Lie algebra $\mathfrak{su}(2)$ is the set of 2×2 traceless anti-hermitian matrices. The basis for the Lie algebra of $SU(2)$ can be built from the hermitian Pauli matrices: for $a = 1, 2, 3$, they satisfy

$$\sigma_a = \alpha_a^\dagger \qquad \text{and} \qquad \mathrm{Tr}\,\sigma_a = 0. \tag{4.29}$$

A basis element in the Lie algebra of $SU(2)$ can then be written

$$T^a = -\frac{1}{2} i\alpha_a, \tag{4.30}$$

where the i guarantees that T^a is anti-hermitian. At the moment, the upper or lower placement of indices is purely notational at the moment, but will gain significance later. In order to compute the structure constants, we recall that the Pauli matrices obey the following identity:

$$\sigma_a \sigma_b = \delta_{ab} I_2 + i\varepsilon_{abc}\sigma_c. \tag{4.31}$$

Using this, the structure constants $f^{ab}{}_c$ are, according to their definition, obtained from the commutator

$$[T^a, T^b] = -\frac{1}{4}[\sigma_a, \sigma_b] = -\frac{1}{2} i\varepsilon_{abc}\sigma_c = f^{ab}{}_c T^c. \tag{4.32}$$

Therefore, we simply read off

$$\boxed{f^{ab}{}_c = \varepsilon_{abc}} \qquad a, b, c = 1, 2, 3. \tag{4.33}$$

Example 4.2.4 (G = SO(3))**:** From our previous analysis, we know that the Lie algebra of $SO(3)$ consists of

$$\mathfrak{so}(3) = \{3 \times 3 \text{ real anti-symmetric matrices}\}. \tag{4.34}$$

We can write down a convenient basis as

$$\widetilde{T}^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \widetilde{T}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \widetilde{T}^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.35}$$

This can be used to find the structure constants as

$$[\widetilde{T}^a, \widetilde{T}^b] = f^{ab}{}_c \widetilde{T}^c \tag{4.36}$$

with $f^{ab}{}_c = \varepsilon_{abc}$ with $a, b, c = 1, 2, 3$. We find that the Lie algebras are the same

$$\mathfrak{so}(3) \simeq \mathfrak{su}(2) \tag{4.37}$$

However, the Lie groups themselves are not isomorphic:

$$SO(3) \not\simeq SU(2). \tag{4.38}$$

We learn that different Lie groups can lead to the same Lie algebra. However, we will see later that this degeneracy will be easy to deal with. In fact, there is a relation between these two Lie groups:

$$SO(3) \simeq \frac{SU(2)}{\mathbb{Z}_2}. \tag{4.39}$$

> **Remark:** This is because there is a *double cover*, or two-to-one, surjective homomorphism
>
> $$\phi: \; SU(2) \to SO(3), \qquad \ker(\phi) = \mathbb{Z}_2 = \{\pm 1\}. \tag{4.40}$$

4.3 Diffeomorphisms

A Lie group is a very special type of manifold. In particular, for each element h in the Lie group G, we have two smooth maps from the group to itself that come from left and right group multiplication,

$$L_h: \; G \to G \qquad g \in G \longmapsto hg \in G, \tag{4.41}$$

$$R_h: \; G \to G \qquad g \in G \longmapsto gh \in G, \tag{4.42}$$

known as *left-* and *right-translations* respectively.

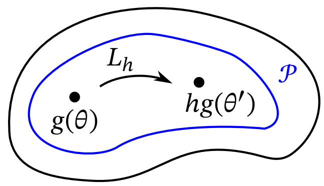
Claim 6: These maps are *surjective*, which means that for every $g' \in G$, there is an element $g \in G$, such that $L_h(g) = g'$.

Proof. To see this, set $g = h^{-1}g'$. □

Claim 7: These maps are also *injective*, which means that for all $g, g' \in G$, we have that if $L_h(g) = L_h(g)$, then we must have $g = g'$.

Proof. $L_h(g) = L_h(g') \;\Rightarrow\; gh = gh' \;\Rightarrow\; g = g'$. □

Definition 30: Since L_h and R_h are both surjective and injective—they are said to be *one-to-one*, or *diffeomorphisms* of $\mathcal{M}(G)$—the *inverse map* $(L_h)^{-1} = L_{h^{-1}}$ exists and is smooth.



Now introduce coordinates $\{\theta^i\}$, $i = 1, \dots, D$ in some region containing the identity element. We can then parametrise every $g = g(\theta) \in G$, and we can choose the parametrisation such that $g(0) = e$. Assuming that we map to an element g' that is in the same coordinate patch as g, let $g' = g(\theta') = L_h(g) = hg(\theta)$. In coordinates, L_h is specified by D real functions $\theta'^i = \theta'^i(\theta)$, $i = 1, \dots, D$. Since L_h is a diffeomorphism, the Jacobian matrix

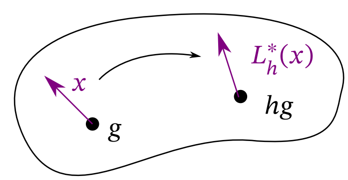
$$J^i{}_j(\theta) = \frac{\partial \theta'^i}{\partial \theta^j} \tag{4.43}$$

exists and is invertible, which is equivalent to saying that $\det J \neq 0$. We have this family of invertible diffeomorphisms. Such maps would not normally exist on a generic manifold. Let us see what the consequences of the existence of these diffeomorphisms are.

Consequences

The map $L_h : G \to G$ induces a map $L_{h*} : T_g G \to T_{hg} G$. This map, called the *differential*[1] of L_h, is in coordinate $\{\theta^i\}$ defined as

$$\begin{aligned} L_{h*}: \quad & T_g(G) && \to && T_{hg}(G) \\ & v = v^i \, \partial/\partial\theta^i && \longmapsto && J^i{}_j(\theta) v^j \, \partial/\partial\theta'^i \end{aligned} \tag{4.44}$$



Definition 31: A *(tangent) vector field* X on G specifies a tangent vector $X(g) \in T_g G$ at each point g in the manifold G.

Given a set of coordinates $\{\theta^i\}$, we can expand any vector in terms of the coordinate basis $\left\{\frac{\partial}{\partial\theta}\right\}$

$$X(\theta) = v^i(\theta) \frac{\partial}{\partial \theta^i} \in T_{g(\theta)}(G), \qquad i = 1, \dots, D, \tag{4.45}$$

where we employ a slight abuse of notation and write $X(\theta)$ when we really mean $X(g(\theta))$.

Definition 32: A vector field X is said to be *smooth* if the functions $v^i(\theta)$ are smooth.

As soon as we have a left multiplication map, we can shift tangent vectors across the manifold and define a smooth tangent vector field: Starting from a tangent vector at the identity $\omega \in T_e G$, we define a vector field $X(g)$ for each element $g \in G$ as

$$X(g) = L_{g*}(\omega) \in T_g(G). \tag{4.46}$$

[1] In the *General Relativity* course, we have met this as the *pushforward*.

However, we also have some more restrictive properties that come with this map. As L_{g*} is smooth and invertible, $X(g)$ is smooth and *non-vanishing*. Starting from a basis $\{\omega_a\}$, $a = 1, \dots, D$, for T_eG, we actually get D independent, non-vanishing vector fields

$$X_g(g) = L_{g*}(\omega_a). \tag{4.47}$$

These are called the *left-invariant vector fields*; these exist on any Lie group manifold.

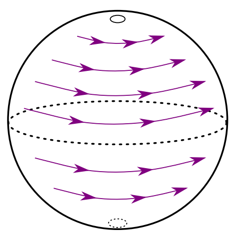


Figure 4.3

Example 4.3.1 (hairy ball theorem)**:** On a two sphere $\mathcal{M}(G) \simeq S^2$, you can never have a smooth, non-vanishing vector field. Informally, you can never comb flat the hair on a spherical doll; there will always be at least two zeros where the hair parts.

In general, the number of zeros a manifold has is related to the Euler character. Out of the two dimensional compact manifolds, the only allowed manifold is the torus $\mathcal{M}(G) = T^2$, where $G = U(1) \times U(1)$.

We can also use the technology of left-invariant vector fields to define the Lie algebra of vector fields without resorting to matrix representations of Lie groups.

4.4 Matrix Lie Groups

Let us go back to the more restricted context of matrix Lie groups. This means that we can realise things more explicitly since we know how to multiply matrices.

Claim 8: Let $G = \mathrm{Mat}_N(F)$, $n \in \mathbb{N}$, $F = \mathbb{R}$ or $\mathbb{C}$. Then, $\forall h \in G$, and $\forall X \in \mathfrak{g}$, we can define the left multiplication map $L_{h*}(X) = hX \in T_h(G)$ simply as the multiplication of two matrices.

> **Remark:** It is highly non-obvious that we can do this, since h and X are different objects; one is an element of the Lie group, the other of the Lie algebra.

Proof. Consider the Taylor series of a point g lying on a curve C in the Lie group G, as depicted in Figure 4.4.

$$g(t) = g(0) + \dot{g}(0)t + O(t^2). \tag{4.48}$$

Then $\dot{g}(0) \in T_{g(0)}G$. Consider the curve $C : t \in I \subset R \longmapsto g(t) \in G$. Let $g(0) = e = \mathbb{1}_n$ and $\dot{g}(0) = X \in \mathfrak{g} \simeq T_e(G)$. We can then define a new curve, $C' : t \longmapsto h(t) = h \cdot g(t) \in G$. Near $t = 0$,

$$h(t) \simeq h + thX + O(t^2). \tag{4.49}$$

Therefore, $hX \in T_h G$. $\square$

Equivalently, consider the smooth curve $C : t \longmapsto g(t) \in G$. Then

$$\dot{g}(t) T_{g(t)} G \implies g^{-1}(t)\dot{g}(t) = L_{g^{-1}(t)*}(\dot{g}(t)) \in T_e G = \mathfrak{g}. \tag{4.50}$$

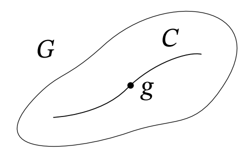


Figure 4.4

This is a very general story that we will meet again when studying gauge theories. For matrix Lie groups, we can always define an element of the Lie algebra by a multiplication of the form $g^{-1}\dot{g}$.

5 Lie Groups from Lie Algebras

5.1 The Exponential Map

Let us now work backwards; imagine that we are given an element $X \in \mathfrak{g}$ of the Lie algebra. Since there is a correspondence between curves and tangent vectors, we can explicitly reconstruct a curve $C : I \in \mathbb{R} \to G$ on the Lie group by solving the ODE

$$\mathrm{g}^{-1}(t)\frac{\mathrm{dg}(t)}{\mathrm{d}t} = X. \tag{5.1}$$

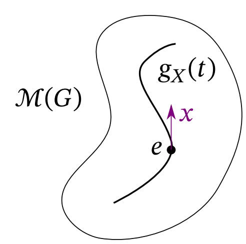


Figure 5.1

This is a curve whose tangent vector at the identity is X. We can enforce this to have a unique solution by specifying the boundary condition $\mathrm{g}(0) = \mathbb{1}_n$. To solve (5.1), we define the exponential of a matrix:

Definition 33: The exponential of a matrix $M \in \mathrm{Mat}_n(F)$ is defined to be

$$\mathrm{Exp}(M) := \sum_{l=0}^{\infty} \frac{1}{l!} M^l \in \mathrm{Mat}_n(F). \tag{5.2}$$

Claim 9: We solve (5.1) by setting $g(t) = \mathrm{Exp}(tX)$, $\forall t \in \mathbb{R}$.

Proof. Note that $\mathrm{g}(0) = \mathrm{Exp}(0) = \mathbb{1}_n$.

$$\frac{\mathrm{dg}(t)}{\mathrm{d}t} = \sum_{l=1}^{\infty} \frac{1}{(l-1)!} t^{l-1} X^l = \mathrm{Exp}(tX) \cdot X = \mathrm{g}(t) X. \tag{5.3}$$

□

The exponential map takes us from the Lie algebra to the Lie group. It must be the case that

$$\mathrm{Exp}(tX) \in G, \qquad \forall t \in \mathbb{R},\ \forall X \in \mathfrak{g}. \tag{5.4}$$

> **Remark:** As previously mentioned, we can also define the exponential map without reference to the matrix representation of Lie groups.

Exercise 5.1: If $X \in \mathfrak{su}(N)$, check that $\mathrm{Exp}(tX) \in SU(N)$, $\forall t \in \mathbb{R}$.

5.2 Baker-Campell-Hausdorff Identity

Setting $t = 1$, we have a map

$$\mathrm{Exp}: \mathfrak{g} \to G. \tag{5.5}$$

Claim 10: The exponential map is one-to-one in some neighbourhood of the identity $e \in G$.

Given elements of the Lie algebra $X, Y \in \mathfrak{g}$, we can construct Lie group elements $g_X = \mathrm{Exp}(X) \in G$ and $g_Y = \mathrm{Exp}(Y) \in G$. Group multiplication can then be recovered by requiring

$$g_X g_Y = g_Z = \mathrm{Exp}(Z) \in G. \tag{5.6}$$

By comparing the left and hight hand side, we get the *Baker-Campbell-Hausdorff* (BCH) formula

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] - [Y, [X, Y]) + \dots, \tag{5.7}$$

where $(\dots)$ reflects the terms that are quartic and higher in the matrices. This series allows us to reconstruct from the bracket of the Lie algebra close to the identity the multiplication law of the group.

> **Remark:** The formula is given purely in terms of sums of nested commutators of matrices. This reflects the fact that this is a formula in the Lie algebra. In fact, the same formula holds and makes sense even if we do not deal with matrix Lie groups. Moreover, this series is not unique since we can always shift around the higher order terms with the Jacobi identity.

Therefore, $\mathfrak{g}$ determines G in some neighbourhood of the identity (where the BCH series converges). More generally, Exp is not globally one-to-one. In particular,

Claim 11: The map Exp is not surjective when G is not connected.

Example 5.2.1 ($G = O(n)$)**:** The Lie algebra of $O(n)$ is the set of real anti-symmetric $n \times n$ matrices:

$$\mathscr{L}(O(n)) = \left\{x \in \mathrm{Mat}_n(\mathbb{R}) \mid X + X^T = 0\right\} \tag{5.8}$$

Recall that the matrices $X \in \mathscr{L}(O(n))$ are traceless $\mathrm{Tr}X = 0$. To find out which connected component $\mathrm{Exp}(X)$ is in, we use the identity

$$\det(\mathrm{Exp}(X)) = \exp(\mathrm{Tr}X) = +1. \tag{5.9}$$

Therefore, for any element X of the Lie algebra, $\mathrm{Exp}(X) \in SO(n)$. Since the image of the exponential map is not the whole of $O(n)$, but only $SO(n) \subset O(n)$, this is not surjective.

Claim 12: In general, for compact Lie groups G, the image of $\mathfrak{g}$ under Exp is the connected component of the identity.

Claim 13: The exponential map is not injective when G has a $U(1)$ subgroup.

Example 5.2.2 ($G = U(1)$)**:** The Lie algebra of $U(1)$ is $\mathscr{L}(U(1)) = \{X = ix \in \mathbb{C} \mid x \in \mathbb{R}\}$. Then $g = \mathrm{Exp}(X) = \exp(ix) \in U(1)$. Lie algebra elements ix and $ix + 2\pi i$ yield the same group element. The inverse of Exp is multi-valued so it is not injective.

5.3 SU(2) vs SO(3)

We know that the Lie algebras have the same structure constants, meaning that they are isomorphic:

$$\mathscr{L}(SU(2)) \simeq \mathscr{L}(SO(3)). \tag{5.10}$$

Although we cannot construct an isomorphism $SU(2) \not\simeq SO(3)$, we can construct a *double-covering*, that is a globally 2-to-one map

$$\begin{aligned} d : SU(2) &\to SO(3) \\ A &\longmapsto d(A), \end{aligned} \tag{5.11}$$

where $d(A)_{ij} = \frac{1}{2}\,\mathrm{tr}_2(\sigma_i A \sigma_j A^\dagger)$. Note that $d(A) = d(-A)$, $\forall A \in SU(2)$. This map provides an isomorphism between groups

$$SO(3) \simeq SU(2)/\mathbb{Z}_2, \qquad \mathbb{Z}_2 = \{\mathbb{1}_2, -\mathbb{1}_2\} \text{ centre of SU(2).} \tag{5.12}$$

However, in this case this is more than merely an isomorphism between groups: Lie groups are also manifolds. The manifold of $SU(2)$ is $\mathcal{M}(SU(2)) \simeq S^3$.

$$S^3 \simeq \left\{\mathbf{x} \in \mathbb{R}^4 \mid |\mathbf{x}|^2 = 1\right\}. \tag{5.13}$$

Moreover, the manifold of $SO(3)$ is S^3 with antipodal points identified. Quotienting out $\mathbb{Z}_2$ is the same as taking S^3 and identifying antipodal points $\mathbf{x} \sim -\mathbf{x}$. This is written as

$$\mathcal{M}(SO(3)) \simeq S^3_+ \cup \{\text{equator with antipodal points identified}\}, \tag{5.14}$$

where S^3_+ is the upper hemisphere $x_3 \geq 0$. Note that $S_3^+ \simeq B_3$.

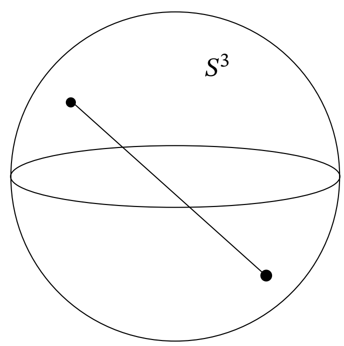


Figure 5.2: S^3 with antipodal points identified.

6 Representation Theory

Definition 34: A *representation* of a group G is a map $D\colon\ G \to \mathrm{Mat}_n(F)$, $n \in \mathbb{N}$, $F = \mathbb{R}$ or $\mathbb{C}$, that preserves the structure of the group. In particular, for all $g_1, g_2 \in G$ we have

$$D(g_1)D(g_2) = D(g_1 g_2). \tag{6.1}$$

Definition 35: A representation D is *faithful* if it is injective; in other words, distinct group elements have distinct representations.

Definition 36: For a *Lie group* G, a representation D is a group representation in the sense above, but the map D must be smooth.

Definition 37: For a *Lie algebra* $\mathfrak{g}$, a *representation* R is a map $R\ :\ \mathfrak{g} \to \mathrm{Mat}_n(F)$ that preserves the structure of the Lie algebra. In particular, for all $X_1, X_2 \in \mathfrak{g}$ and $\alpha, \beta \in F$ we have

1. $[R(X_1), R(X_2)] = R([X_1, X_2])$
2. $R(\alpha X_1 + \beta X_2) = \alpha R(X_1) + \beta R(X_2)$

> **Remark:** Again, representations of Lie groups need not be faithful.

Definition 38: The *dimension* of a representation is the dimension n of the matrices.

Definition 39: The representation matrices act on a vector space $V \simeq F^n$ known as the *representation space* or $\mathfrak{g}$-*module*.

> **Remark:** In a physical context, this will be the Hilbert space of states on which the matrix operator act.

There is a direct relation between the representations of Lie groups and Lie algebras. Take a representation D of a matrix Lie group G. Note that in general, $\dim D \neq \dim G$. We will construct a corresponding representation of the Lie algebra. To do this, we use again the correspondence between tangent vectors and curves. For $X \in \mathfrak{g}$, define a curve $C_X : I \subset \mathbb{R} \to G$ that maps $t \longmapsto g_X(t)$. We then expand $g_X(t) \simeq \mathbb{1}_n + tX + O(t^2)$ and map the curve onto $D(g_X(t)) \in GL(n, F) \subset \mathrm{Mat}_n(F)$ in the space of matrices. We then Taylor expand this around the origin to give

$$D(g(t)) = D(\mathbb{1}_m) + \left.\frac{\mathrm{d}D(g(t))}{\mathrm{d}t}\right|_{t=0} t + O(t^2). \tag{6.2}$$

We have of course $D(\mathbb{1}_m) = \mathbb{1}_n$.

Claim 14: For a representation D of G, we then obtain a representation R of $\mathfrak{g}$ as

$$R(X) := \left.\frac{\mathrm{d}D(g(t))}{\mathrm{d}t}\right|_{t=0}, \qquad \forall X \in \mathfrak{g}. \tag{6.3}$$

Proof. The linearity of the representation follows automatically from the definition, so we only need to show the homomorphism property $[R(X_1), R(X_2)] \overset{!}{=} R([X_1, X_2])$. For any pair $X_i \in \mathfrak{g}$, $i = 1, 2$, we construct two curves $g_i : I \to G$, starting at the identity $g_1(0) = g_2(0) = \mathbb{1}_m$, such that its tangent vectors at the identity are $\dot{g}_i(0) = X_i$. We then define a third curve $h : I \to G$ by

$$h(t) = g_1^{-1}(t) g_2^{-1}(t) g_1(t) g_2(t). \tag{6.4}$$

Expanding $g_i \approx \mathbb{1}_n + tX_i + \dots$, the coefficient of t in the expansion of $h(t)$ vanishes. To second order in t, we have

$$h(t) \approx \mathbb{1}_m + t^2[X_1, X_2] + \dots. \tag{6.5}$$

Since the group representation D is by definition linear, we have

$$D(h(t)) = D(\mathbb{1}_m + t^2[X_1, X_2] + O(t^3)) \tag{6.6}$$

$$= \mathbb{1}_n + R([X_1, X_2])t^2 + O(t^3) \tag{6.7}$$

$$= D(\mathbb{1}_m) + t^2 \left.\frac{\mathrm{d}^2 D(h(t))}{\mathrm{d}t^2}\right|_{t=0} + O(t^3) \tag{6.8}$$

But we can also use the homomorphism property of the group representation:

$$D(h) = D(g_1^{-1} g_2^{-1} g_1 g_2) = D(g_1)^{-1} D(g_2)^{-1} D(g_1) D(g_2). \tag{6.9}$$

We then plug in the Taylor expansions for g_i and compare (6.9) and (6.7) at $O(t^2)$ to give

$$R([X_1, X_2]) = [R(X_1), R(X_2)]. \tag{6.10}$$

□

Exercise 6.1: Let R be a representation of $\mathscr{L}(G)$. For all elements of G of the form $g = \mathrm{Exp}(X)$, with $X \in \mathscr{L}(G)$. Then define

$$D(g) = \mathrm{D}(\mathrm{Exp}\, X) := \mathrm{Exp}\, R(X). \tag{6.11}$$

Under what circumstances does this give a good representation of G?

6.1 Representations of Lie Algebras

For any matrix Lie algebra $\mathfrak{g} = \mathscr{L}(G)$ of some matrix Lie group $G \subset \mathrm{Mat}_n(F)$, we can always write down three canonical representations: For all $X \in \mathfrak{g}$, we define the

trivial representation by mapping everything to the trivial element, $R_0(X) = 0$. The trivial representation has dimension $\dim R_0 = 1$.

fundamental representation by mapping the matrices X to themselves, $R_F(X) = X$. For $n \times n$ matrices, this has dimension $\dim R_F = n$.

> **Remark:** The trivial and adjoint representations exist for all Lie algebras, whereas the fundamental representation only makes sense for matrix Lie algebras.

adjoint representation by mapping each element X to a particular linear map, $R_A(X) = \mathrm{ad}_X$, which is defined as

$$\begin{aligned} \mathrm{ad}_X : \mathfrak{g} &\to \mathfrak{g} \\ Y &\longmapsto [X, Y]. \end{aligned} \tag{6.12}$$

Since the underlying vector space is the Lie algebra $\mathfrak{g}$ itself, the adjoint representation has dimension $\dim d_A = \dim \mathfrak{g}$.

The map ad_X is equivalent to a $D \times D$ matrix: For a choice of basis $B = \{T^a\}$, for $a = 1, \ldots, D$, we can expand two elements $X, Y \in \mathfrak{g}$ as

$$X = X_a T^a \qquad Y = Y_a T^a. \tag{6.13}$$

In particular, as we have seen before, the bracket of X, Y is determined by the bracket of the generators T^a

$$\mathrm{ad}_X(Y) = [X, Y] = X_a Y_b [T^a, T^b] = X_a Y_b f^{ab}{}_c T^c. \tag{6.14}$$

Therefore, the c^{th} component of $\mathrm{ad}_X(Y)$ is

$$[\mathrm{ad}_X(Y)]_c = (R_X)^b{}_c Y_b \quad \Rightarrow \quad \boxed{(R_X)^b{}_c = X_a f^{ab}{}_c}. \tag{6.15}$$

More concretely, we can define it in a particular basis as $[R_{\mathrm{Adj}}(X)]^b{}_c = (R_X)^b{}_c$.

Claim 15: We want to show that the adjoint representation is indeed a representation. Let us check the defining properties of a representation. For all $X, Y \in \mathfrak{g}$, we must have

$$[R_{\mathrm{Adj}}(X), R_{\mathrm{Adj}}(Y)] = R_{\mathrm{Adj}}([X, Y]) \tag{6.16}$$

Proof. We have $R_{\text{Adj}}(X) = ad_X$ and $R_{\text{Adj}}(Y) = ad_Y$. Hence, for all elements $Z \in \mathfrak{g}$ of the Lie algebra, we have

$$\left(R_{\text{Adj}}(X) \circ R_{\text{Adj}}(Y)\right)(Z) = [X, [Y, Z]] \tag{6.17}$$
$$\left(R_{\text{Adj}}(Z) \circ R_{\text{Adj}}(Y)\right)(X) = [Z, [Y, X]] \tag{6.18}$$
$$\tag{6.19}$$

Subtracting these two, we get

$$[R_{\text{Adj}}(X), R_{\text{Adj}}(Y)](Z) = [X, [Y, Z]] - [Y, [X, Z]] \tag{6.20}$$

This is the left hand side of (6.16). Subtracting the right hand side we get

$$(\text{LHS} - \text{RHS})(Z) = [X, [Y, Z]] - [Y, [X, Z]] - [[X, Y], Z] \tag{6.21}$$
$$= [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0 \tag{6.22}$$

where we used the Jacobi identity. We see that Jacobi identity of the Lie algebra therefore ensures that the Lie algebra has an adjoint representation on itself. Property two of the Lie algebra is satisfied since ad_X is linear in X. □

Remark: Let $g = \exp(tX)$ for some $X \in \mathfrak{g}$. Let $v \in V$ be a vector. Then the representation

$$D(g)v = gvg^{-1} \tag{6.23}$$

of the Lie group G induces the adjoint representation of the Lie algebra $\mathfrak{g}$ on V; differentiating both sides with respect to t and evaluating at $t = 0$ gives

$$R(X)v = [X, v]. \tag{6.24}$$

If we change the basis of the representation, the matrices change, but we still have the same representation. This motivates the following definition:

Definition 40: Two representations R_1 and R_2 of $\mathfrak{g}$ are *isomorphic*, $R_1 \simeq R_2$, if there exists an invertible matrix S such that for all $X \in \mathfrak{g}$

$$R_2(X) = SR_1(X)S^{-1}. \tag{6.25}$$

Definition 41: A representation R over a vector space V has an *invariant subspace* $U \in V$ if $\forall X \in \mathfrak{g}, u \in U$,

$$R(X)u \in U. \tag{6.26}$$

Example 6.1.1: Any representation has two 'trivial' invariant subspaces

$$U = \{0\} \quad \text{and} \quad U = V. \tag{6.27}$$

Definition 42: An *irreducible* representation R of $\mathfrak{g}$ has no non-trivial invariant subspaces.

6.2 Representation Theory of $\mathscr{L}(SU(2))$

6.2.1 Cartan Basis

Take the real basis $T^a = -\frac{1}{2}i\sigma^a$ with $a = 1, 2, 3$, which we have already met before. Then

$$\mathscr{L}(SU(2)) = \text{Span}_{\mathbb{R}} \{T^a \mid a = 1, 2, 3\}. \tag{6.28}$$

Now, consider what happens when we change to a new (complex) basis:

$$H = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad E_+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad E_- = \frac{1}{2}(\sigma_1 - i\sigma_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{6.29}$$

This is called the *Cartan basis*. Note that for $X \in \mathscr{L}(SU(2))$, we have $X = -X^\dagger$. Thus if $X = X_H H + H_+ E^+ + X_- E^-$, then $X_H = i\mathbb{R}$ and $X_+ = -(X_-)^*$.

Let us now allow not only real but also complex coefficients of the generators T^a. This defines the *complexification* of the Lie algebra:

$$\mathscr{L}_{\mathbb{C}}(SU(2)) = \text{Span}_{\mathbb{C}} \{T^a \mid a = 1, 2, 3\}. \tag{6.30}$$

The $\{H, E^+, E^-\}$ are a basis for $\mathscr{L}_{\mathbb{C}}(SU(2))$. Often we refer to H as the *Cartan element* of the basis, while $E_\pm$ are the step operators. These matrices satisfy the same commutation relations as we are used to from angular momentum:

$$[H, E_\pm] = \pm 2E_\pm, \qquad [E_+, E_-] = H. \tag{6.31}$$

The relation $[H, E_\pm]$ implies that

$$ad_H(E_\pm) = \pm 2E_\pm, \tag{6.32}$$

while $ad_H(H) = 0$ corresponds to the statement that $[H, H] = 0$. Thus, in the Cartan basis, $\{H, E_+, E_-\}$ are eigenvectors of

$$\text{ad}_H : \mathscr{L}(SU(2)) \to \mathscr{L}(SU(2)) \tag{6.33}$$

with eigenvectors $\{0, +2, -2\}$.

We refer to the non-zero eigenvalues as *roots* of the Lie algebra $\mathscr{L}(SU(2))$.

> **Remark:** From now on, we will write $\mathscr{L}(SU(2)) = \mathfrak{su}(2)$ and $\mathscr{L}_{\mathbb{C}}(SU(2)) = \mathfrak{su}_{\mathbb{C}}(2)$.

Consider a representation R of $\mathfrak{su}(2)$ with representation space V. H is diagonal, assume $R(H)$ is diagonalisable. This is equivalent to saying that the representation space V is spanned by the eigenvectors of $R(H)$. We will introduce those eigenvectors by their eigenvalues $\lambda \in \mathbb{C}$:

$$R(H)v_\lambda = \lambda v_\lambda. \tag{6.34}$$

These eigenvalues $\{\lambda\}$ of $R(H)$ are called the *weights* of the representation R.

> **Remark:** Weights belong to a representation, whereas roots belong to the Lie algebra itself.

6.2.2 Step Operators

As we have casually mentioned before, $E_\pm$ are known as *step operators*. In the following discussion, the reasons for this will become clear. We can change the order of two representation matrices by introducing the commutator:

$$R(H)R(E_\pm)v_\lambda = (R(E_\pm)R(H) + [R(H), R(E_\pm)])v_\lambda \tag{6.35}$$

By definition, the commutator is $[R(H), R(E_\pm)] = \pm 2R(E_\pm)$. Therefore, we get

$$\cdots = (\lambda \pm 2)R(E_\pm)v_\lambda. \tag{6.36}$$

Therefore, when we act on the basis with a representation of the step operator, we will move up or down to different values of the weight. In a finite-dimensional representation, we need to have a finite basis, and therefore a finite number of different weights. A finite-dimensional representation R of $\mathfrak{su}(2)$ must therefore have a highest weight $\Lambda \in \mathbb{C}$ and a highest weight vector v_Λ with

$$R(H)v_\Lambda = \Lambda v_\Lambda, \qquad R(E_+)v_\Lambda = 0. \tag{6.37}$$

Later on, we will also use the fact there needs to be a lowest weight as well for a finite-dimensional representation.

Starting from v_Λ, we can generate a new basis vector with weight $\Lambda - 2$ by acting on it with $R(E_-)$. If we repeat this process, which is illustrated in 6.1, do we eventually get all of the basis vectors of the representation? If R is irreducible, this is indeed the case.

Claim 16: If R is irreducible, the remaining basis vectors are strings of $R(H)$, $R(E_+)$, and $R(E_-)$ acting on v_Λ.

> **Remark:** These strings can also be reordered using the commutators.

Proof. Let us define $v_{\Lambda-2n} = (R(E_-))^n v_\Lambda$ for $n \in \mathbb{N}$. Consider the action of $R(E_+)$ on $v_{\Lambda-2n}$. Then, we recall that we can obtain this by acting on the step down operator

$$R(E_+)v_{\Lambda-2n} = R(E_+)R(E_-)\Lambda_{\Lambda-2n+2}. \tag{6.38}$$

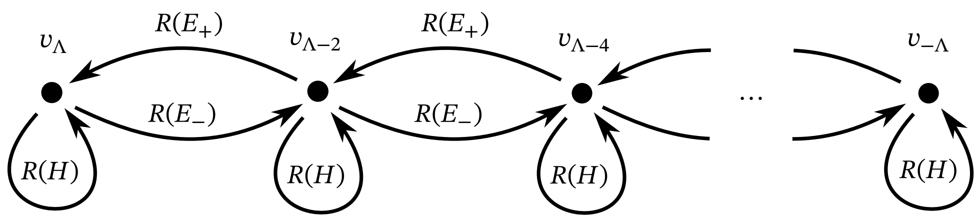


Figure 6.1: The action of the $\mathfrak{su}_{\mathbb{C}}(2)$ representation matrices on basis vectors v_λ of weight λ in the representation space V.

Using the commutator, we can then reverse the order to give

$$\dots = \big(R(E_-)R(E_+) + [R(E_+), R(E_-)]\big)v_{\Lambda-2n+2} \tag{6.39}$$

$$= R(E_-)R(E_+)v_{\Lambda-2n+2} + (\Lambda - 2n + 2)v_{\Lambda-2n+2} \tag{6.40}$$

Let us now set $n = 1$ in equation (6.40). The relation then boils down to

$$R(E_+)v_{\Lambda-2} = \Lambda v_\Lambda. \tag{6.41}$$

Thus, we do not get a new linearly independent eigenvector, but instead a multiple of the highest weight vector. Similarly, considering the $n = 2$ case yields

$$R(E_+)v_{\Lambda-4} = R(E_-)R(E_+)v_{\Lambda-2} + (\Lambda - 2)v_{\Lambda-2}. \tag{6.42}$$

Using the result (6.41), we get

$$\cdots = \Lambda R(E_-)v_\Lambda + (\Lambda - 2)v_{\Lambda-2} = (2\Lambda - 2)v_{\Lambda-2}. \tag{6.43}$$

Proceeding by induction, we find that for all n,

$$R(E_+)v_{\Lambda-2n} \propto v_{\Lambda-2n+2}. \tag{6.44}$$

The constants of proportionality can be obtained by substitution into (6.40). In particular, we will set the constants of proportionality to

$$R(E_+)v_{\Lambda-2n} = r_n v_{\Lambda-2n+2}. \tag{6.45}$$

Equation (6.40) implies the following first order recurrence relation for these coefficients

$$r_n = r_{n-1} + \Lambda - 2n + 2 \tag{6.46}$$

In addition to this, we know that $R(E_+)v_\Lambda = 0$. This sets the boundary condition $r_0 = 0$. We can solve (6.46) to get

$$r_n = (\Lambda + 1 - n)n. \tag{6.47}$$

Finally, we consider the fact that the finite dimension of the representation R implies the existence of a lowest weight $\Lambda - 2N$ for some $N \in \mathbb{N}_0$. By definition there exists some lowest weight vector $v_{\Lambda-2N} \neq 0$, which is annihilated by the lowering operator

$$R(E_-)v_{\Lambda-2N} = 0 \Rightarrow v_{\Lambda-2N-2} = 0 \tag{6.48}$$

Using $n = N + 1$ in Equation (6.45), we deduce that

$$R(E_+)v_{\Lambda-2N-2} = r_{N+1}v_{\Lambda-2N} = 0. \tag{6.49}$$

Therefore, since $v_{\Lambda-2N} \neq 0$, we must have $r_{N+1} = 0$. In particular,

$$r_{N+1} = (\Lambda - N)(N + 1) = 0. \tag{6.50}$$

Therefore, we find that the highest weight, rather than being an arbitrary number, is equal to an integer $\Lambda = N$.

Conclusion: The finite dimensional irreps R_Λ of $\mathfrak{su}(2)$ are labeled by the *heighest weight* $\Lambda \in \mathbb{N}_0$. The remaining weights in R_1 are

$$S_\Lambda = \{-\Lambda, -\Lambda + 2, \dots, \Lambda - 2, \Lambda\} \subset \mathbb{Z}. \tag{6.51}$$

Hence, the dimension of the irrep labelled by Λ is $\dim(R_\Lambda) = \Lambda + 1$.

In perfoming this proof, we determined the structure of the irreducible representation of $\mathfrak{su}(2)$. □

Hence, we have unique candidates for the irreps of dimensions in Table 6.1.

R_Λ	Type	Dimension
$R_0 = d_0$	trivial	1
$R_1 = d_f$	fundamental	2
$R_2 = d_{adj}$	adjoint	3

Table 6.1: Irreducible representations of $\mathfrak{su}(2)$.

6.2.3 Angular Momentum in Quantum Mechanics

What we have here described is analogous to the theory of angular momentum in quantum mechanics. Recall that the total angular momentum consists of orbital and spin angular momenta. We have a Hermitian operator $\mathbf{J} = (J_1, J_2, J_3)$ of total angular momentum. Its corresponding eigenstates are labelled by $j \in \mathbb{Z}/2$, $j \geq 0$, and $m \in \{-j, -j+1, \ldots, j-1, j\}$:

$$J^2 |j,m\rangle = \hbar j(j+1) |j,m\rangle \tag{6.52}$$

$$J_3 |j,m\rangle = \hbar m |j,m\rangle, \tag{6.53}$$

where $J^2 = J_1^2 + J_2^2 + J_3^2$. In the Cartan representation, we have the correspondance

$$J_3 = \frac{1}{2}R(H), \qquad J_\pm = J_1 \pm iJ_2 = R(E_\pm). \tag{6.54}$$

The highest weight is $\Lambda = 2j \in \mathbb{Z}^+$ and the other weights are $\lambda = 2m \in \mathbb{Z}$. The associated states are

$$v_\Lambda \sim |j,j\rangle, \qquad v_\lambda \sim |j,m\rangle. \tag{6.55}$$

6.2.4 $SU(2)$ Representations from $\mathfrak{su}(2)$ Representations

Locally, we can parametrise group elements $A \in SU(2)$ as $A = \text{Exp}(X)$, where X is an element of the corresponding Lie algebra $\mathfrak{su}(2)$. Starting from an irrep R_Λ of $\mathfrak{su}(2)$, we have a representation

$$D_\Lambda(A) = \text{Exp}\big(R_\Lambda(X)\big). \tag{6.56}$$

As before, $\Lambda \in \mathbb{N}_0$. We will now use the relationship between $SU(2)$ and $SO(3)$ that we established. In particular, you can think of $SO(3)$ as the quotient

$$SO(3) \simeq \frac{SU(3)}{\mathbb{Z}_2}, \tag{6.57}$$

corresponding to the identification of antipodal points $A \sim -A$. This condition implies that the representation matrices of A and $-A$ must be the same. In other words, for a representation of $SO(3)$, we require that $\forall A \in SU(3)$

$$D_\Lambda(-A) = D_\Lambda(A). \tag{6.58}$$

Using the representation property, we can see that this is actually equivalent to the simpler condition

$$D_\Lambda(-\mathbb{1}_2) = D_\Lambda(\mathbb{1}_2). \tag{6.59}$$

Now if we represent $-\mathbb{1}_2 = \text{Exp}(i\pi H)$ with

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \text{Exp}(i\pi H) = \begin{pmatrix} e^{i\pi} & \\ & e^{-i\pi} \end{pmatrix} = -\mathbb{1}. \tag{6.60}$$

Therefore, the representation is

$$D_\Lambda(-\mathbb{1}_2) = \text{Exp}(i\pi R(\Lambda)(H)) \tag{6.61}$$

$$= \text{Exp}\begin{pmatrix} e^{i\pi\Lambda} & & & \\ & e^{i\pi(\Lambda-2)} & & \\ & & \ddots & \\ & & & e^{-i\pi\Lambda} \end{pmatrix}. \tag{6.62}$$

with $R_\Lambda(H) = \text{diag}\{\Lambda, \Lambda - 2, \dots, -\Lambda + 2, -\Lambda\}$. We find that $D_\Lambda(-\mathbb{1}_2)$ has eigenvalue $\exp(i\pi\lambda) = (-1)^\lambda = (-1)^\Lambda$. Hence,

$$D_\Lambda(-\mathbb{1}_2) = D_\Lambda(+\mathbb{1}_2) = \mathbb{1}_{\Lambda+1} \quad \iff \quad \Lambda \in 2\mathbb{Z}. \tag{6.63}$$

Now we have two cases:

- $\Lambda \in 2\mathbb{Z} \Rightarrow D_\Lambda$ is a representation of $SU(2)$ and $SO(3)$.
- $\Lambda \in 2\mathbb{Z} + 1 \Rightarrow D_\Lambda$ is a representation of $SU(2)$, but *not* of $SO(3)$.

6.3 New representations from old

Definition 43 (conjugate rep)**:** If R is a representation of a *real* Lie algebra $\mathfrak{g}$, we define a *conjugate representation* by $\overline{R}(X) = R(X)^*$ for all $X \in \mathfrak{g}$.

Sometimes, we find that $\overline{R} \simeq R$.

> **Remark:** In physics, if a particle transforms under a representation R, its anti-particles will transform under $\overline{R}$.

Definition 44 (direct sum)**:** Suppose we are given representations R_1 and R_2 of any $\mathfrak{g}$ (not necessarily real) with representation spaces V_1 and V_2 of dimensions d_1 and d_2. We then define a *direct sum* $R_1 \oplus R_2$ as a new representation that acts on the representation space

$$V_1 \oplus V_2 = \{v_1 \oplus v_2 \mid v_1 \in V_1, v_2 \in V_2\} \tag{6.64}$$

as

$$(R_1 \oplus R_2)(X)(v_1 \oplus v_2) = (R_1(X)v_1) \oplus (R_2(X)v_2) \in V_1 \oplus V_2 \tag{6.65}$$

for all $X \in \mathfrak{g}$. The matrix corresponding to the linear map $(R_1 \oplus R_2)(X)$ is in block matrix notation

$$(R_1 \oplus R_2)(X) = \begin{pmatrix} R_1(X) & 0 \\ 0 & R_2(X) \end{pmatrix}. \tag{6.66}$$

6.3.1 The Tensor Product Representation

Two Particle States

Let us start with some motivation. Assume we have two particles, each occupying a two-particle Hilbert space $\mathcal{H}_i = \{|\uparrow\rangle_i, |\downarrow\rangle_i\}$, $i = 1, 2$. The two particle states are states like $|\uparrow\rangle_1 |\downarrow\rangle_2$ or $|\uparrow\rangle_1 |\downarrow\rangle_2 -$

$|\downarrow\rangle_1 |\uparrow\rangle_2 \in \mathcal{H}_{12}$. To analyse this we will have to think more about how to combine two representations.

Definition 45: Given two vector spaces V_1 and V_2, we define the *(tensor) product space*

$$V_1 \otimes V_2 = \mathrm{Span}_F \{v_1 \otimes v_2 \mid v_1 \in V_1, v_2 \in V_2\} \tag{6.67}$$

where the tensor product satisfies $\forall v_i \in V_i, \alpha \in F$:

- $(v_1 + w_1) \otimes (v_2 + w_2) = v_1 \otimes v_2 + w_1 \otimes v_2 + v_1 \otimes w_2 + w_1 \otimes w_2$
- $\alpha(v_1 \otimes v_2) = (\alpha v_1) \otimes v_2 = v_1 \otimes (\alpha v_2)$

Definition 46: Given two linear maps $M_i : V_i \to V_i$, $i = 1, 2$, we define the *tensor product map*

$$\begin{aligned}(M_1 \otimes M_2) : V_1 \otimes V_2 &\to V_1 \otimes V_2 \\ v_1 \otimes v_2 &\longmapsto (M_1 v_1) \otimes (M_2 v_2).\end{aligned} \tag{6.68}$$

We extend this action on the basis $v_1 \otimes v_2$ by linearity to the full tensor product space $V_1 \otimes V_2$.

Definition 47: Given representations R_i, $i = 1, 2$ of $\mathfrak{g}$ with corresponding representation spaces V_i, we have, for each element X in the Lie algebra $\mathfrak{g}$, a map $R_i(X) : V_i \to V_i$. We now define the *tensor product representation* $R_1 \otimes R_2$ with representation space $V_1 \otimes V_2$ which acts on all $X \in \mathfrak{g}$ as

$$(R_1 \otimes R_2)(X) = R_1(X) \otimes I_2 + I_1 \otimes R_2(X). \tag{6.69}$$

> **Remark:** Note that his is not the same as $R_1(X) \otimes R_2(X)$.

Let us choose a particular set of basis vectors for the two vector spaces:

$$B_1 = \left\{V_1^j \mid j = 1, \dots, d_1\right\} \qquad B_2 = \left\{V_2^\alpha \mid \alpha = 1, \dots, d_2\right\}. \tag{6.70}$$

We can then represent $R_1 \otimes R_2$ as a 'matrix' with $i, j = 1, \dots, d_1$ and $\alpha, \beta = 1, \dots, d_2$ as

$$(R_1 \otimes R_2)(X) = R_1(X)_{ij} \mathbb{1}_{ij} + \mathbb{1}_{ij} R_2(X)_{\alpha\beta}. \tag{6.71}$$

The representation $R_1 \times R_2$ has dimension $d_1 \times d_2$.

Exercise 6.2: Check that $R_1 \otimes R_2$ as defined is a representation of $\mathfrak{g}$.

> **Remark:** In physics terms, the motivation for this construction come from the laws of quantum mechanics: multi-particle states live in the tensor product space of the individual one-particle states. The definition of the addition of angular momentum is described by this definition of the tensor product representation.

> **Remark:** Up to isomorphism, the tensor product is associative: $(R_1 \otimes R_2) \otimes R_2 = R_1 \otimes (R_2 \otimes R_3)$.

6.4 Reducibility

Definition 48: Let us recall that a representation R with representation space V has an *invariant subspace* $U \subset V$ if $R(X)u \in U$ for all $X \in \mathfrak{g}$ and all $u \in U$.

We already know that an *irreducible* representation (irrep) has no non-trivial invariant subspace.

Definition 49: A *fully reducible* representation can be expressed as a direct sum of irreps.

Let us look at what this means for a particular basis. If R has a non-trivial invariant subspace, then we can find a basis that makes this manifest. In other words, we can find a basis such that $R(X)$ has block diagonal form:

$$R(X) = \begin{pmatrix} A(X) & B(X) \\ 0 & C(X) \end{pmatrix}. \tag{6.72}$$

In this case, elements of U correspond to vectors $\begin{pmatrix} u \\ 0 \end{pmatrix}$ for all elements X of the Lie algebra $\mathfrak{g}$.

If R is fully reducible, then $R = R_1 \oplus R_2 \oplus \cdots \oplus R_l$. Hence, we have a basis, where $R(X)$ is block diagonal, where the blocks are the individual irreps

$$R(X) = \begin{pmatrix} R_1(X) & & & \\ & R_2(X) & & \\ & & \ddots & \\ & & & R_l(X) \end{pmatrix}. \tag{6.73}$$

This is useful for us due to the following theorem, which we will not prove in this course:

Theorem 3: Let R_i, $i = 1, \dots, m$, be finite-dimensional irreps of a simple Lie algebra $\mathfrak{g}$. The tensor product $(R_1 \otimes \cdots \otimes R_m)$ is fully reducible. In other words,

$$R_1 \otimes R_2 \cdots \otimes R_m \simeq \widetilde{R}_1 \oplus \widetilde{R}_2 \oplus \cdots \oplus \widetilde{R}_{m'}. \tag{6.74}$$

> **Remark:** Another more mathematical way to say this is that the tensor representation forms a ring.

6.5 Tensor Product of $\mathfrak{su}(2)$ Representations

Let R_Λ and $R_{\Lambda'}$ be irreps of $\mathfrak{su}(2)$ with highest weights $\Lambda, \Lambda' \in \mathbb{N}_0$. Previously, we have found that $\dim(R_\Lambda) = \Lambda + 1$ and $\dim(R_{\Lambda'}) = \Lambda' + 1$. Let V_Λ and $V_{\Lambda'}$ be the respective representation spaces. We then form the tensor product representation $R_\Lambda \otimes R_{\Lambda'}$ with the complex representation space $V_\Lambda \otimes V_{\Lambda'}$, by representing each $X \in \mathfrak{su}(2)$ by a matrix $(R_\Lambda \otimes R_{\Lambda'})(X)$ that acts as

$$(R_\Lambda \otimes R_{\Lambda'})(X)(v \otimes v') = (R_\Lambda(X)v) \otimes v' + v \otimes (R_{\Lambda'}(X)v'). \tag{6.75}$$

We know that the dimension is simply the product of the two representation dimensions $\dim(R_\Lambda \otimes R_{\Lambda'}) = (\Lambda + 1)(\Lambda' + 1)$. Moreover, we know that this has to be a fully reducible representation of $\mathfrak{su}(2)$, which means that we can write

$$R_\Lambda \otimes R_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathbb{N}_0} \mathcal{L}^{\Lambda''}_{\Lambda,\Lambda'} R_{\Lambda''} \tag{6.76}$$

where the *multiplicities* $\mathcal{L}^{\Lambda''}_{\Lambda,\Lambda'} \in \mathbb{N}_0$ are sometimes called *Littlewood coefficents*.

> **Remark:** In *Foundations of Quantum Mechanics* and *Group Theory*, we have met the Clebsch-Gordan coefficients. Clebsch-Gordan coefficients are matrix elements that are not always integers! So they are not the same as these multiplicities.

Let us try to calculate these coefficients. V_Λ has a basis $\{v^\lambda\}$ that are eigenvectors of $R_\Lambda(H)$ with eigenvalues $\lambda \in S_\Lambda = \{-\Lambda, -\Lambda + 2, \dots, +\Lambda\}$. Similarly, we have the same for $V_{\Lambda'}$. For the Lie algebra of $SU(2)$, we have already describe the full set of irreducible representations R_Λ with $\dim(R) = \Lambda + 1$, where $\Lambda \in \mathbb{N}_0$. Let us consider the tensor product representations $R_\Lambda \otimes R_{\Lambda'}$. Because $\mathfrak{su}(2)$ is simple, it is fully reducible. In other words, we can decompose the product representation into a sum over the irreps $R_\Lambda \otimes R_{\Lambda'} \simeq \bigoplus_{\Lambda'' \in \mathbb{N}_0} \mathcal{L}^{\Lambda''}_{\Lambda,\Lambda'} R_{\Lambda''}$. Our task is now to find the multiplicities $\mathscr{L}$, which describe how often each irrep enters the decomposition. Let us use the Cartan basis. The vector space V_Λ has basis $\{v_\lambda\}$, $\lambda \in S_\Lambda = \{-\Lambda, -\Lambda + 2, \dots, +\Lambda\}$. The eigenvectors $R_\Lambda(H)v_\lambda = \lambda v_\lambda$. Similarly, we have a second vector space $V_{\lambda'}$ with basis $\left\{v'_{\lambda'}\right\}$. Construct a basis for $V_\Lambda \otimes V_{\Lambda'}$ as $B = \left\{v_\lambda \otimes v'_\lambda \mid \lambda \in S_\Lambda, \lambda' \in S_{\Lambda'}\right\}$. We now want to understand the weights of the new representation. To work out this weight, we act on a general basis vector with the representation of the Cartan element H. By definition, the tensor product representation acts as

$$\begin{aligned}(R_\Lambda \otimes R_{\Lambda'})(H)(v_\lambda \otimes v_{\lambda'}) &= (R_\Lambda(H)v_\lambda) \otimes v'_{\lambda'} + v_\lambda \otimes (R_{\Lambda'}(H)v'_{\lambda'}) &(6.77)\\ &= (\lambda + \lambda')(v_\lambda \otimes v_{\lambda'}). &(6.78)\end{aligned}$$

The eigenvalues are just the sums of the eigenvalues of the basis elements. Therefore, we can deduce that $R_\Lambda \otimes R_{\Lambda'}$ has weight set $S_{\Lambda,\Lambda'} = \left\{\lambda + \lambda' \mid \lambda \in S_\Lambda, \lambda' \in S_{\Lambda'}\right\}$. Sometimes, this sum adds up the same weight in n different cases. We then have to keep track of these n multiplicities. We can then construct the representation, which is the sum of irreducibles, explicitly. The irreps are uniquely determined by identification of their weight set. We must take this weight set and decompose it into unions of the weight sets associated with the irreducibles; we will see that there is a unique way to do this. Consider first the highest weight. This will only come when $\lambda = \Lambda$ and $\lambda' = \Lambda'$. Since there is only one way to obtain the sum $\Lambda + \Lambda'$, the irrep $R_{\Lambda+\Lambda'}$ must appear in the decomposition with multiplicity one! In other words, $\mathcal{L}^{\Lambda+\Lambda'}_{\Lambda,\Lambda'} = 1$. We can decompose the tensor product as

$$R_\Lambda \otimes R_{\Lambda'} = R_{\Lambda+\Lambda'} \oplus \widetilde{R}_{\Lambda,\Lambda'}. \tag{6.79}$$

The problem is reduced to finding the remainder $\widetilde{R}_{\Lambda,\Lambda'}$, which will have weight set $\widetilde{S}_{\Lambda,\Lambda'}$ where $S_{\Lambda,\Lambda'} = S_{\Lambda+\Lambda'} \cup \widetilde{S}_{\Lambda,\Lambda'}$. We remove the weight set

$$S_{\Lambda+\Lambda'} = \left\{-\Lambda - \Lambda', \dots, +\Lambda + \Lambda'\right\}, \tag{6.80}$$

and find the highest weight of the remainder $\widetilde{R}_{\Lambda,\Lambda'}$ and keep repeating until the remainder is empty.

Example 6.5.1 ($\Lambda = \Lambda' = 1$)**:** In this case, the weight set of both representations consists of two elements

$$S_1 = \{-1, +1\}. \tag{6.81}$$

This is the representation a spin-$\frac{1}{2}$ particle carries. Then the weight set of the tensor product is

$$S_{1,1} = \{-1, +1\} + \{-1, +1\} = \{-2, 0, 0, +2\}. \tag{6.82}$$

$$= \{-2, 0, +2\} \cup \{0\}, \tag{6.83}$$

where $\{-2, 0, +2\}$ is the weight set of the highest weight representation R_2. We find that

$$R_1 \otimes R_1 = R_2 \oplus R_0. \tag{6.84}$$

In quantum mechanics, this means that two spin-$\frac{1}{2}$ particles form states of spin-1 (triplet) and spin-0 (singlet). Note that the dimensions match as well.

Exercise 6.3 (Sheet 2, Qu 8)**:** Consider two irreps of $SU(2)$: $\Lambda' = M$ and $\Lambda = N$. Then show, by applying the above algorithm, that $R_N \otimes R_M = R_{|M-N|} \oplus R_{|M-N|+2} \oplus \cdots \oplus R_{N+M}$. For $SU(2)$ this is essentially the whole story.

7 The Cartan Classification

Our ultimate aim of this chapter will be to classify all finite-dimensional simple complex Lie algebras $\mathfrak{g}$ as was done by Cartan in 1894. In order to do this, we will have to build up some mathematical machinery.

7.1 The Killing Form

We can specify an inner product on $\mathbb{R}^3$, by mapping $(\mathbf{u},\mathbf{v}) \to \mathbf{u}\cdot\mathbf{v}$. Similarly, we might specify this scalar product by giving a metric such as $(u_i, v_j) \to \delta^{ij} u_i v_j$.

Definition 50 (Inner Product)**:** Given a vector space V over $F = \mathbb{R}$ or $\mathbb{C}$, an *inner product* is a bilinear, symmetric map $i: V \times V \to F$.

Definition 51 (non-degenerate)**:** We say that an inner product i is *non-degenerate* if for all non-zero $v \in V$ there exists some $w \in V$ such that the inner product $i(v,w) \neq 0$.

> **Remark:** This amounts to saying that there is no vector which is orthogonal to itself and all other vectors in the vector space.

There a 'natural' inner product that we can write down on a Lie algebra $\mathfrak{g}$ called the *Killing form*.

Definition 52: The *Killing form* κ of a Lie algebra $\mathfrak{g}$ over a field F is the map

$$\begin{aligned} \kappa: \quad \mathfrak{g}\times\mathfrak{g} &\quad\to\quad F \\ (X,Y) &\quad\longmapsto\quad \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y) \end{aligned} \tag{7.1}$$

Let us find out what this means explicitly in components. The action of the inner ad-map composition on an element $Z \in \mathfrak{g}$ is

$$(ad_X \circ ad_Y)(Z) = [X,[Y,Z]]. \tag{7.2}$$

Now choose a basis $\{T^a\}$, where $a = 1, \dots, D = \dim(\mathfrak{g})$, for the Lie algebra $\mathfrak{g}$. We then know that by definition of the structure constants $f^{ab}{}_c$, we have $[T^a, T^b] = f^{ab}{}_c T^c$. Expanding the elements X, Y, and Z of the Lie algebra in terms of this basis, we see that

$$[X,[Y,Z]] = X_a Y_b Z_c [T^a, [T^b, T^c]] \tag{7.3}$$

$$= X_a Y_b Z_c f^{ad}{}_e f^{bc}{}_d T^e \tag{7.4}$$

$$= M(X,Y)^a{}_e Z_c T^e \tag{7.5}$$

with $M(X,Y)^c{}_e = X_a Y_b f^{ad}{}_e f^{bc}{}_d$. This is the explicit form of the inner ad-map composition. To find the explicit map defined by the Killing form, we take the trace of this:

$$\kappa(X,Y) = \mathrm{Tr}_D[M(X,Y)] = \kappa^{ab} X_a Y_b, \tag{7.6}$$

which means that the components of the Killing form are

$$\boxed{\kappa^{ab} = f^{ad}_c f^{bc}_d} \tag{7.7}$$

This is manifestly symmetric.

Now what does 'natural' mean?

Claim 17: The Killing form κ remains invariant under the adjoint action of the Lie algebra $\mathfrak{g}$, meaning that $\forall X, Y, Z \in \mathfrak{g}$

$$\kappa(X,[Y,Z]) = \kappa([X,Y],Z) \tag{7.8}$$

Proof. By the defining property of the adjoint representation, we have

$$ad_{[Z,X]} = ad_Z \circ ad_X - ad_X \circ ad_Z. \tag{7.9}$$

By cyclic invariance of the trace $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB) = \mathrm{Tr}(BCA)$, we find that

$$\kappa(X,[Y,Z]) = \mathrm{Tr}\big(ad_X \circ ad_{[Y,Z]}\big) \tag{7.10}$$

$$= \mathrm{Tr}(ad_X \circ ad_Y \circ ad_Z) - \mathrm{Tr}(ad_X \circ ad_Z \circ ad_Y) \tag{7.11}$$

$$= \mathrm{Tr}(ad_X \circ ad_Y \circ ad_Z) - \mathrm{Tr}(ad_Y \circ ad_X \circ ad_Z) \tag{7.12}$$

$$= \mathrm{Tr}\big(ad_{[X,Y]} \circ ad_Z\big) \tag{7.13}$$

$$= \kappa([X,Y],Z). \tag{7.14}$$

$\square$

In fact, for simple Lie algebras, the Killing form κ is the unique invariant inner product, up to an overall scalar multiple.

Definition 53: A Lie algebra is *semi-simple* if it has no Abelian ideals.

Claim 18: A finite-dimensional semi-simple Lie algebra can be written as the direct sum of a finite number of simple Lie algebras.

Proof. Sheet 2 Qu 9b. □

> **Note:** The underlying set for the vector space direct sum $\mathfrak{g}_1 \oplus \mathfrak{g}_2$ is
>
> $$\mathfrak{g} \times \mathfrak{g}' = \left\{(x, x') \,\middle|\, x \in \mathfrak{g},\ x' \in \mathfrak{g}'\right\}. \tag{7.15}$$
>
> where scalar multiplication and addition is defined as
>
> $$c(x, x') := (cx, cx') \tag{7.16}$$
> $$(x, x') + (y, y') := (x + y, x' + y'). \tag{7.17}$$
>
> In order to make this a Lie algebra, we define the commutator to be
>
> $$[(x, x'), (y, y')] := ([x, y], [x', y']). \tag{7.18}$$
>
> Note that the subsets $\mathfrak{g}, \mathfrak{g}'$ commute with each other: $[(x, 0), (0, x')] = 0$.

We will now try to find out under what conditions the Killing form is non-degenerate.

Theorem 4 (Cartan)**:** The Killing form κ is non-degenerate if and only if the associated Lie algebra $\mathfrak{g}$ is semi-simple.

Proof of forward direction only. Assume that κ is non-degenerate. Suppose for contradiction that $\mathfrak{g}$ is not semi-simple. This means that $\mathfrak{g}$ has an Abelian ideal $\mathfrak{j}$. Denote $\dim(\mathfrak{g}) = D$ and $\dim(\mathfrak{j}) = d$. Choose a basis

$$B = \{T^a\} = \left\{T^i \mid i = 1, \dots, d\right\} \cup \{T^\alpha \mid \alpha = 1, \dots, D - d\}, \tag{7.19}$$

where $\left\{T^i\right\}$ span $\mathfrak{j}$. As $\mathfrak{j}$ is Abelian, we must have $[T^i, T^j] = 0$, $\forall i, j$. Moreover, as $\mathfrak{j}$ is an ideal, $[T^\alpha, T^j] = f^{\alpha j}{}_\kappa T^k \in \mathfrak{j}$ and therefore $f^{ij}_a = 0$ and $f^{\alpha j}_\beta = 0$. For $X = X_a T^a \in \mathfrak{g}$ and $Y = Y_j T^j \in \mathfrak{j}$, we have $\kappa[X, Y] = \kappa^{ai} X_a Y_i$ with

$$\kappa^{ai} = f^{ad}_c f^{ic}_d = f^{aj}_\alpha f^{i\alpha}_j = 0. \tag{7.20}$$

Therefore, we have $K[X, Y] = 0$ for all $X \in \mathfrak{j}$ and all $X \in \mathfrak{g}$. In other words, κ is degenerate, which contradicts the assumption. Hence, $\mathfrak{g}$ is semi-simple. □

Definition 54 (compact type)**:** A real Lie algebra $\mathfrak{g}_\mathbb{R}$ is of *compact type* if there is a basis for which the Killing form is negative definite

$$\kappa^{ab} = -\kappa\delta^{ab}, \qquad \kappa \in \mathbb{R}^+ \tag{7.21}$$

Claim 19: If $\mathscr{L}(G)$ is of compact type, then G is compact (as a topological manifold).

Theorem 5: Every complex semi-simple Lie algebra $\mathfrak{g}$ of finite dimension has a real form $\mathfrak{g}_{\mathbb{R}}$ of compact type.

Exercise 7.1: Show that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n$, where $\mathfrak{g}_i$ for $i = 1, \ldots, n$ are simple.

7.2 Complexification

Given a *real* Lie algebra $\mathfrak{g}$, we can find a basis $\{T^a\}$, $a = 1, \ldots, \dim \mathfrak{g}$, with real structure constants

$$[T^a, T^b] = f^{ab}_c T^c, \qquad f^{ab}_c \in \mathbb{R}. \tag{7.22}$$

Definition 55: Given a real Lie algebra $\mathfrak{g} = \mathrm{Span}_{\mathbb{R}} \{T^a\}$, its *complexification* is $\mathfrak{g}_{\mathbb{C}} = \mathrm{Span}_{\mathbb{C}} \{T^a\}$. We say that $\mathfrak{g}$ is a *real form* of $\mathfrak{g}_{\mathbb{C}}$.

Together with the bracket (7.22), $\mathfrak{g}_{\mathbb{C}}$ is a *complex Lie algebra.* A complex Lie algebra can have multiple real forms.

Example 7.2.1: Consider the Lie algebra of $SU(2)$:

$$\mathscr{L}(SU(2)) = \mathfrak{su}(2) = \mathrm{span}_{\mathbb{R}} \left\{ T^a = -\frac{i\sigma_a}{2} \;\middle|\; a = 1, 2, 3 \right\}. \tag{7.23}$$

This is the set of 2×2 traceless anti-Hermitian matrices. Its complexification is

$$\mathscr{L}_{\mathbb{C}}(SU(2)) = \mathfrak{su}_{\mathbb{C}}(2) = \mathrm{Span}_{\mathbb{C}} \left\{ T^a = -\frac{i\sigma_a}{2} \;\middle|\; a = 1, 2, 3 \right\}, \tag{7.24}$$

which is the set of 2×2 traceless complex matrices.

Definition 56: The *Cartan-Weyl basis* for $\mathfrak{su}_{\mathbb{C}}(2)$ is given by the *Cartan element* $H = 2iT^3$ and the two elements $E_\pm = iT^1 \pm T^2$ with brackets

$$[H, E_\pm] = \pm 2E_\pm, \qquad [E_+, E_-] = H. \tag{7.25}$$

7.2.1 Recovering a Real Representation

Let us star from a representation R of $\mathfrak{su}_{\mathbb{C}}(2)$ with

$$[R(H), R(E_\pm)] = \pm 2R(E_\pm) \tag{7.26a}$$

$$[R(E_+), R(E_-)] = R(H). \tag{7.26b}$$

We can pass back to original basis via

$$R(T^1) = \frac{1}{2i}\big(R(E_+) + R(E_-)\big), \tag{7.27a}$$

$$R(T^2) = \frac{1}{2i}\big(R(E_+) - R(E_-)\big), \tag{7.27b}$$

$$R(T^3) = \frac{1}{2i}R(H). \tag{7.27c}$$

For all $X \in \mathfrak{su}(2)$, we can expand $X = X_a T^a$ for some $X_a \in \mathbb{R}$. Set $R(X) = X_a R(T^a)$ to get a representation of $\mathfrak{su}(2)$.

7.3 Cartan Subalgebra

Definition 57 (ad-diagonalisable)**:** We say that an element $X \in \mathfrak{g}$ is *ad-diagonalisable* (or *semisimple*), if the map ad_X is diagonalisable, meaning that there exists a basis $B = \{T^a\}$ of $\mathfrak{g}$ such that $[X, T^a]$ is proportional to T^a for any element of B.

Definition 58 (Cartan subalgebra)**:** A *Cartan subalgebra* $\mathfrak{h}$ of $\mathfrak{g}$ is a maximal Abelian subalgebra consisting entirely of ad-diagonalisable elements:

1. ad-diagonalisable: $H \in \mathfrak{h} \implies H$ is ad-diagonalisable,
2. Abelian: $H, H' \in \mathfrak{h} \implies [H, H'] = 0$,
3. maximal: $X \in \mathfrak{g}$ with $[X, H] = 0$ for all $H \in \mathfrak{h} \implies X \in \mathfrak{h}$.

Claim 20: All possible Cartan subalgebras of $\mathfrak{g}$ are isomorphic and have the same dimension.

This motivates the following definition:

Definition 59: The *rank* of a Lie algebra $\mathfrak{g}$ is the dimension of a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$.

$$\operatorname{rank} \mathfrak{g} := \dim \mathfrak{h}. \tag{7.28}$$

> In physics, $\operatorname{rank} \mathfrak{g}$ provides the maximal number of quantum numbers which can be used to label (at least partially) the states of a physical system that has $\mathfrak{g}$ as its symmetry algebra. [F&S]

Example 7.3.1: Let $\mathfrak{g} = \mathfrak{su}_{\mathbb{C}}(2) = \operatorname{span}_{\mathbb{C}}\{H, E_\pm\}$. The Cartan element H is ad-diagonalisable, but $E_\pm$ are not, so we can choose $\mathfrak{h} = \operatorname{Span}_{\mathbb{C}}\{H\}$ to be a Cartan subalgebra. Hence $\operatorname{rank}(\mathfrak{g}) = 1$.

Given a Cartan subalgebra $\mathfrak{h}$, we can choose a basis $\{H^i\}$, $i = 1, \dots, r$, where $[H^i, H^j] = 0$.

Example 7.3.2: Consider the set of traceless complex $n \times n$-matrices, denoted $\mathfrak{sl}(n, \mathbb{C})$. Choose $(H^i)_{\alpha\beta} = \delta_{\alpha i}\delta_{\beta i} - \delta_{\alpha i+1}\delta_{\beta i+1}$. Then

$$\operatorname{rank}[\mathfrak{sl}(n, \mathbb{C})] = n - 1. \tag{7.29}$$

7.4 Roots

The fact that the Cartan subalgebra is Abelian implies

- $\Rightarrow$ by the defining property of the adjoint rep.: $[\mathrm{ad}_{H^i}, \mathrm{ad}_{H^j}] = \mathrm{ad}_{[H^i, H^j]} = 0$,
- $\Rightarrow$ the r linear maps $\mathrm{ad}_{H^i} : \mathfrak{g} \to \mathfrak{g}$, with $i = d, \dots, r$ are *simultaneously diagonalisable*,
- $\Rightarrow$ the Lie algebra $\mathfrak{g}$ is spanned by simultaneous eigenvectors of ad_{H^i},

The eigenvectors of ad_{H^i} within the Cartan subalgebra can be distinguished by their eigenvectors:

Zero eigenvalues correspond to the *Cartan elements* H^i:

$$\mathrm{ad}_{H^i}(H^j) = [H^i, H^j] = 0 \tag{7.30}$$

Non-zero eigenvalues correspond to *step-operators* E^α:

$$\mathrm{ad}_{H^i}(E^\alpha) = [H^i, E^\alpha] = \alpha^i E^\alpha \tag{7.31}$$

By construction, $\alpha^i \in \mathbb{C}$ are not all zero.

For a general element $H \in \mathfrak{h}$, we can write $H = c_i H^i$ with $c_i \in \mathbb{C}$. By (7.31), we have

$$[H, E^\alpha] = \alpha(H) E^\alpha, \tag{7.32}$$

with $\alpha(H) = c_i \alpha^i \in \mathbb{C}$.

> When comparing the eigenvalue equation (7.32) to the general expansion $[H, T^a] = \sum_b \xi(H)^a{}_b T^b$, it follows that the eigenvalues of ad_H are the roots of the characteristic equation $\det(\xi(H) - \alpha(H)\mathbb{1}) = 0$.

Definition 60 (roots)**:** The non-zero eigenvalues $\alpha(H)$ of ad_H are called the *roots* of the Lie algebra.

> It is a matter of convention that only the non-zero eigenvalues are called roots.

For any fixed element $H \in \mathfrak{h}$, the eigenvalue $\alpha(H)$ of E^α is some complex number which depends linearly on H. Each root therefore gives a linear function $\alpha : \mathfrak{h} \to \mathbb{C}$, which means that α is an element of the dual space $\mathfrak{h}^*$.

As $\mathfrak{g}$ is spanned by the eigenvectors of ad_H, it can be written as a direct sum of vector spaces γ_α according to

$$\mathfrak{g} = \bigoplus_\alpha \mathfrak{g}_\alpha, \qquad \mathfrak{g}_\alpha = \{x \in \mathfrak{g} \mid [H, x] = \alpha(H) \cdot x \text{ for all } H \in \mathfrak{h}\}. \tag{7.33}$$

Separating out the Cartan subalgebra $\mathfrak{h} = \mathfrak{g}_0$, we obtain the *root space decomposition* of $\mathfrak{g}$ relative to $\mathfrak{h}$:

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \neq 0} \mathfrak{g}_\alpha. \tag{7.34}$$

Remark: In Yang–Mills theory, gauge bosons carry the adjoint representation of some simple Lie algebra. Their quantum numbers are zero if they correspond to generators of the Cartan subalgebra and are otherwise given by the non-zero roots, in which case the bosons are charged and self-interact. In electro-weak theory, the $W^\pm$ correspond to the roots of an $\mathfrak{sl}(2)$ algebra and hence are charged, whereas a linear combination of the photon and Z boson, which are uncharged, correspond to the Cartan subalgebra. In Abelian gauge theories like electrodynamics there are no roots and gauge bosons do not self-interact.

In general, we may have a problem of degenerate eigenvalues. In a full proof of the Cartan construction, one can prove the following[1]:

Claim 21: The roots are non-degenerate if $\mathfrak{g}$ is simple; each root α corresponds to the one-dimensional subspace spanned by the step-generator E^α.

Corollary: This means that the set Φ of roots of the Lie algebra $\mathfrak{g}$ must consist of $(\dim \mathfrak{g} - \operatorname{rank} \mathfrak{g})$ distinct elements of $\mathfrak{h}^*$.

Definition 61: The *Cartan-Weyl basis* is $B = \left\{H^i \mid i = 1, \dots, r\right\} \cup \{E^\alpha \mid \alpha \in \Phi\}$.

7.4.1 Inner Product on the Root Space

To analyse the root system, we will define an inner product on the space $\mathfrak{h}^*$ of roots via the Killing form on $\mathfrak{g}$.

By Cartan's theorem, Theorem 4, the Killing form with normalisation N,

$$\kappa(X, Y) = \frac{1}{N} \operatorname{Tr}[ad_X \circ ad_Y], \tag{7.35}$$

is non-degenerate if $\mathfrak{g}$ is simple. We will always assume $\mathfrak{g}$ is simple from now on. Let us evaluate κ in the Cartan-Weyl basis.

[1] A more complete story is given by Fuchs and Schweigert.

Claim 22: 1. $\forall H \in \mathfrak{h}, \forall \alpha \in \Phi$, we have $\kappa(H, E^\alpha) = 0$

2. $\forall \alpha, \beta \in \Phi$, such that $\alpha + \beta \neq 0$, we have $\kappa(E^\alpha, E^\beta) = 0$

3. non-degeneracy of κ then implies that $\forall H \in \mathfrak{h}$, $\exists H' \in \mathfrak{h}$ such that $\kappa(H, H') \neq 0$.

Proof of 1. For any $H' \in \mathfrak{h}$, we use the relation (7.32) to write

$$\alpha(H')\kappa(H, E^\alpha) = \kappa(H, [H', E^\alpha]) \tag{7.36}$$

Using the invariance (7.8) of the Killing form, this becomes

$$\alpha(H')\kappa(H, E^\alpha) = \kappa([H, H'], E^\alpha) \tag{7.37}$$

$$= \kappa(0, E^\alpha) = 0 \tag{7.38}$$

$$\alpha(H') \neq 0 \quad \Rightarrow \quad \kappa(H, E^\alpha) = 0. \tag{7.39}$$

□

Proof of 2. For all $H' \in \mathfrak{h}$, we have

$$[\alpha(H') + \beta(H')]\kappa(E^\alpha, E^\beta) \overset{(7.32)}{=} \kappa([H', E^\alpha], E^\beta) + \kappa(E^\alpha, [H', E^\beta]) \tag{7.40}$$

$$\overset{(7.8)}{=} 0 \tag{7.41}$$

Nor for all $\alpha, \beta \in \Phi$ such that $\alpha + \beta \neq 0$, we have

$$\alpha(H') + \beta(H') \neq 0 \text{ for some } H' \quad \Rightarrow \quad \kappa(E^\alpha, E^\beta) = 0. \tag{7.42}$$

□

Proof of 3. Take any $H \in \mathfrak{h}$ and assume for contradiction that $\kappa(H, H') = 0$, $\forall H' \in \mathfrak{h}$. From 1, $\kappa(H, E^\alpha) = 0$ for all $\alpha \in \Phi$. This means that $\forall X \in \mathfrak{g}$, we have $\kappa(H, X) = 0$. In other words, κ is degenerate, which is a contradiction. □

The third statement implies that κ is a non-degenerate inner product on $\mathfrak{h}$

$$\kappa(H, H') = \kappa^{ij} c_i c'_j, \tag{7.43}$$

where $H = c_i H^i$ and $H' = c'_i H^i$. This means that $\kappa^{ij} = \kappa(H^i, H^j)$ is an invertible $r \times r$ matrix. Explicitly, there exists an inverse matrix κ^{-1} such that $(\kappa^{-1})_{ij}\kappa^{jk} = \delta_i^k$. This inverse matrix κ^{-1} defines a non-degenerate inner product on the dual space $\mathfrak{h}^*$. This is where the roots live! Suppose we have roots $\alpha, \beta \in \Phi$ such that

$$[H^i, E^\alpha] = \alpha^i E^\alpha, \qquad [H^i, E^\beta] = \beta^i E^\beta. \tag{7.44}$$

We then define an inner product

$$\boxed{(\alpha, \beta) := (\kappa^{-1})_{ij}\alpha^i \beta^j} \tag{7.45}$$

Claim 23: If $\alpha \in \Phi$ is a root, then $-\alpha \in \Phi$ is too and $\kappa(E^\alpha, E^{-\alpha}) \neq 0$.

Proof. From Claim 22.1, $\kappa(E^\alpha, H) = 0$ holds for all $H \in \mathfrak{h}$. From Claim 22.2, $\kappa(E^\alpha, E^\beta) = 0$ holds for all roots $\beta \in \Phi$ with $\alpha \neq -\beta$. Hence, unless $-\alpha \in \Phi$ and $\kappa(E^\alpha, E^{-\alpha}) \neq 0$, we have $\kappa(E^\alpha, X) = 0$ for all elements $X \in \mathfrak{g}$, which would mean that κ is degenerate, being a contradiction. □

7.5 Algebra of the Cartan-Weyl Basis

So far, we have the following brackets:

$$[H^i, H^j] = 0, \qquad \forall i, j = 1, \dots, r, \tag{7.46a}$$

$$[H^i, E^\alpha] = \alpha^i E^\alpha, \qquad \alpha \in \Phi. \tag{7.46b}$$

It remains to evaluate $[E^\alpha, E^\beta]$ for the roots $\alpha, \beta \in \Phi$. We can use the Jacobi identity

$$[H^i, [E^\alpha, E^\beta]] = [E^\alpha, [E^\beta, H^i]] - [E^\beta, [H^i, E^\alpha]] \tag{7.47}$$

$$\stackrel{(7.32)}{=} (\alpha^i + \beta^i)[E^\alpha, E^\beta] \tag{7.48}$$

Therefore, if $\alpha + \beta \in \Phi$, then

$$[E^\alpha, E^\beta] = N_{\alpha,\beta} E^{\alpha+\beta}, \tag{7.49}$$

where $N_{\alpha,\beta}$ is a proportionality constant. On the other hand, if $\alpha + \beta = 0$, then

$$[H^i, [E^\alpha, E^{-a}]] = 0, \quad \forall i = 1, \dots, r. \tag{7.50}$$

Since the Cartan subalgebra is maximal, this means that $[E^\alpha, E^{-\alpha}] \in \mathfrak{h}$.

> Since the restriction of κ to the Cartan subalgebra $\mathfrak{h}$ is non-degenerate as well, we can use it to identify the spaces $\mathfrak{h}$ and $\mathfrak{h}^*$. In particular, we will associate to any root α an element H^α of $\mathfrak{h}$, which up to normalisation c_α is unique, such that
>
> $$\alpha(H) = c_\alpha \kappa(H^\alpha, H), \tag{7.51}$$
>
> for all $H \in \mathfrak{h}$. The non-degenerate inner product on $\mathfrak{h}^*$ is then defined
>
> $$(\alpha, \beta) := c_\alpha c_\beta \kappa(H^\alpha, H^\beta) = c_\beta \alpha(H^\beta). \tag{7.52}$$

We will fix the normalisation by defining

$$H^\alpha := \frac{[E^\alpha, E^{-\alpha}]}{\kappa(E^\alpha, E^{-\alpha})}. \tag{7.53}$$

Consider the Killing form $\kappa(H^\alpha, H)$ for all $H \in \mathfrak{h}$. This is

$$\kappa(H^\alpha, H) = \frac{1}{\kappa(E^\alpha, E^{-\alpha})}\kappa([E^\alpha, E^{-\alpha}], H) \tag{7.54}$$

$$\overset{(7.8)}{=} \frac{1}{\kappa(E^\alpha, E^{-\alpha})}\kappa(E^\alpha, [E^{-\alpha}, H]) \tag{7.55}$$

$$\overset{(7.32)}{=} \alpha(H)\frac{\cancel{\kappa(E^\alpha, E^{-\alpha})}}{\cancel{\kappa(E^\alpha, E^{-\alpha})}}. \tag{7.56}$$

▌ Thus, this element solves (7.51) with normalisation $c_\alpha = 1$.

In components, $H^\alpha = c^\alpha{}_i H^i$, where $H = c_i H^i \in \mathfrak{h}$. Using (7.56), we have that $\kappa^{ij} c^\alpha{}_i c_j = \alpha^j c_j$ $\Rightarrow \kappa^{ij} c^\alpha{}_i = \alpha^j$. Therefore we find that the components are $c^\alpha{}_i = (\kappa^{-1})_{ij}\alpha^j$ and

$$H^\alpha = c^\alpha{}_i H^i = (\kappa^{-1})_i \alpha^j H^i. \tag{7.57}$$

In summary, the Cartan-Weyl basis is $\{H^i \mid i = 1 \dots, r\} \cup \{E^\alpha \mid \alpha \in \Phi \subset \mathfrak{h}^*\}$, with Lie brackets

$$[H^i, H^j] = 0, \tag{7.58a}$$

$$[H^i, E^\alpha] = \alpha^i E^\alpha, \tag{7.58b}$$

$$[E^\alpha, E^\beta] = \begin{cases} N_{\alpha,\beta} E^{\alpha+\beta} & \alpha + \beta \in \Phi \\ \kappa(E^\alpha, E^{-\alpha}) H^\alpha & \alpha + \beta = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{7.58c}$$

7.5.1 The Chevalley Basis

To simplify this further, consider the brackets of $H^\alpha \in \mathfrak{h}$ defined above. For all $\alpha, \beta \in \Phi$,

$$[H^\alpha, E^\beta] \overset{(7.57)}{=} (\kappa^{-1})_{ij}\alpha^j [H^j, E^\beta] \tag{7.59}$$

$$\overset{(7.31)}{=} (\kappa^{-1})_{ij}\alpha^i \beta^j E^\beta \tag{7.60}$$

$$\overset{(7.45)}{=} (\alpha, \beta) E^\beta. \tag{7.61}$$

Now $\forall \alpha \in \Phi$, we can define a new normalisation

$$e^\alpha := \sqrt{\frac{2}{(\alpha,\alpha)\kappa(E^\alpha, E^{-\alpha})}} E^\alpha, \qquad h^\alpha := \frac{2}{(\alpha,\alpha)} H^\alpha. \tag{7.62}$$

Remark: We will show in Sec. 7.7 that $(\alpha, \alpha) \neq 0$ for any root α, so that division by (α, α) does not cause any problems.

Now we have the algebra in its final form: The *Chevalley basis* is the set $\{e^\alpha, h^\alpha \mid \}$, together with the following brackets:

$$[h^\alpha, h^\beta] = 0, \tag{7.63a}$$

$$[h^\alpha, e^\beta] = \frac{2(\alpha,\beta)e^\beta}{(\alpha,\alpha)}, \tag{7.63b}$$

$$[e^\alpha, e^\beta] = \begin{cases} n_{\alpha,\beta}e^{\alpha+\beta} & \alpha+\beta\in\Phi \\ h^\alpha & \alpha+\beta = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{7.63c}$$

The inner product is $(\alpha,\beta) = (\kappa^{-1})_{ij}\alpha^i\beta^j$, and $n_{\alpha\beta}$ are undetermined constants.

7.6 $\mathfrak{sl}(2)_\alpha$ Subalgebras

We know from Claim 23 that α is a root if and only if $-\alpha$ is also a root. Hence for each pair $\pm\alpha\in\Phi$, we have a subalgebra of $\mathfrak{g}$ with basis $\{h^\alpha, e^\alpha, e^{-\alpha}\}$. From (7.63), we find the brackets

$$[h^\alpha, e^{\pm\alpha}] = \pm 2e^{\pm\alpha}, \qquad [e^\alpha, e^{-\alpha}] = h\alpha, \tag{7.64}$$

which are the same relations as (6.31) for $\mathfrak{su}_\mathbb{C}(2) \simeq \mathfrak{sl}(2,\mathbb{C})$. This subalgebra is called $\mathfrak{sl}(2)_\alpha$. We have one such $\mathfrak{su}_\mathbb{C}(2)$ subalgebra for each root.

7.6.1 Consequences: 'Root Strings'

Definition 62: For $\alpha,\beta\in\Phi$, where $\alpha\neq\beta$, the *α-string passing through β* is the set of all roots of the form $\beta + l\alpha$ for some $l\in\mathbb{Z}$:

$$S_{\alpha,\beta} = \{\beta + l\alpha \in \Phi \mid l \in \mathbb{Z}\} \tag{7.65}$$

The corresponding vector subspace of $\mathfrak{g}$ is $V_{\alpha,\beta} = \mathrm{Span}_\mathbb{C}\left\{e^{\beta+l\alpha} \mid \beta + l\alpha \in S_{\alpha,\beta}\right\}$. Now consider the action of $\mathfrak{sl}(2)_\alpha$ on $V_{\alpha,\beta}$.

$$[h^\alpha, e^{\beta+l\alpha}] \overset{(7.63)}{=} \frac{2(\alpha,\beta+l\alpha)}{(\alpha,\alpha)}\, e^{\beta+l\alpha} = \left(\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2l\right) e^{\beta+l\alpha} \tag{7.66}$$

$$\text{and} \qquad [e^{\pm\alpha}, e^{\beta+l\alpha}] \begin{cases} \propto e^{\beta+(l\pm1)\alpha} & \text{if } \beta+(l\pm1)\alpha\in\Phi \\ = 0 & \text{otherwise} \end{cases} \tag{7.67}$$

The step operators carry us up or down the root string, but we always stay inside $V_{\alpha,\beta}$. This means that the vector space $V_{\alpha,\beta}$ is *closed* or *invariant* under the ajoint action of the subalgebra $\mathfrak{sl}(2)_\alpha$.

Therefore, $V_{\alpha,\beta}$ is the representation space for some representation R of $\mathfrak{sl}(2)_\alpha$. Moreover, (7.66) tells us that the weight set of R is

$$S_R = \left\{ \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2l \,\middle|\, \beta + l\alpha \in \Phi \right\} \tag{7.68}$$

As roots of $\mathfrak{g}$ are finite in number and non-degenerate, R is finite-dimensional and irreducible. Therefore, we must have $R \simeq R_\Lambda$ for some $\Lambda \in \mathbb{Z}_{\geq 0}$. The associated weight set is

$$S_R = \{-\Lambda, -\Lambda + 2, \dots, +\Lambda\}. \tag{7.69}$$

Comparing (7.68) and (7.69), we find two things. Firstly, $l = n \in \mathbb{Z}$ takes on all the values $n_- \leq n \leq n_+$ for some $n_\pm \in \mathbb{Z}$. In other words, the set of allowed roots forms an unbroken 'string' $S_{\alpha,\beta} = \{\beta + n\alpha \mid n_- \leq n \in \mathbb{Z} \leq n_+\}$ of length $n_+ - n_- + 1$. Secondly, comparing the upper and lower boundaries gives $(-\Lambda) = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_-$ and $+\Lambda = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_+$. Adding these, we find that the weight vectors or roots have angles such that their inner products are quantised in integer units:

$$\boxed{\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 2\frac{|\beta|}{|\alpha|}\cos(\theta_{\alpha,\beta}) = -(n_+ + n_-) \in \mathbb{Z}} \tag{7.70}$$

Claim 24: The inner product of any two roots $\alpha, \beta \in \Phi$ is real:

$$(\alpha,\beta) \in \mathbb{R}. \tag{7.71}$$

Proof. In the Cartan-Weyl basis, the Killing form components are

$$\kappa^{ij} = \kappa(H^i, H^j) = \frac{1}{\mathcal{N}}\operatorname{Tr}[ad_{H^i} \circ ad_{H^j}] \tag{7.72}$$

$$= \frac{1}{\mathcal{N}} \sum_{\delta\in\Phi} \delta^i \delta^j, \tag{7.73}$$

where $\mathcal{N} \in \mathbb{R}$ is a real normalisation constant. Hence for $\alpha, \beta \in \Phi$, where we raise and lower indices with the Killing form,

$$(\alpha,\beta) = \alpha^i \beta^j (\kappa^{-1})_{ij} = \alpha_i \beta^i = \alpha_i \beta_j \kappa^{ij} \tag{7.74}$$

$$= \frac{1}{\mathcal{N}} \sum_{\delta\in\Phi} \alpha_i \delta^i \delta^j \beta_j = \frac{1}{\mathcal{N}} \sum_{\delta\in\Phi} (\alpha,\delta)(\beta,\delta) \tag{7.75}$$

From (7.70), we have integers

$$m_{\alpha,\beta} = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z} \subset \mathbb{R}. \tag{7.76}$$

Therefore, dividing (7.75) by $4/[(\alpha,\alpha)(\beta,\beta)]$ we get

$$\frac{2m_{\alpha,\beta}}{(\beta,\beta)} = \frac{1}{\mathcal{N}} \sum_{\delta\in\Phi} m_{\alpha,\delta} R_{\beta,\delta} \in \mathbb{R}. \tag{7.77}$$

Therefore, $(\beta,\beta) \in \mathbb{R}$ for all roots $\beta \in \Phi$. So from (7.70), we also find that $(\alpha,\beta) \in \mathbb{R}$. □

7.7 Real Geometry of Roots

Recall that roots $\alpha \in \Phi$ are elements of $\mathfrak{h}^*$, the dual of the Cartan subalgebra $\mathfrak{h} \in \mathfrak{g}$.

Claim 25: In fact, the roots span $\mathfrak{h}^*$:

$$\mathfrak{h}^* = \mathrm{Span}_{\mathbb{C}} \{\alpha \in \Phi\}. \tag{7.78}$$

Proof. Assume for contradiction that the roots $\alpha \in \Phi$ do not span $\mathfrak{h}^*$. This means $\exists \lambda \in \mathfrak{h}^*$ such that $\forall \alpha \in \Phi$,

$$(\lambda, \alpha) = (\kappa^{-1})_{ij}\lambda^i \alpha^j = \kappa^{ij}\lambda_i \alpha_j = 0. \tag{7.79}$$

Then we can build $H_\lambda = \lambda_i H^i \in \mathfrak{h}$ such that $[H_\lambda, H] = 0$ and $[H_\lambda, E^\alpha] = (\lambda, \alpha)E^\alpha = 0$ for all $H \in \mathfrak{h}$ and $\alpha \in \Phi$. Since these form a basis of $\mathfrak{g}$, we find that $[H_\lambda, X] = 0$ for all $X \in \mathfrak{g}$. In other words, $\mathfrak{g}$ has a non-trivial ideal $\mathfrak{i} = \mathrm{Span}_{\mathbb{C}} \{H_\lambda\}$, which contradicts our assumption that $\mathfrak{g}$ is simple. Thus, $\mathfrak{h}^* = \mathrm{Span}_{\mathbb{C}} \{\alpha \in \Phi\}$. □

Corollary: Since $\dim(\mathfrak{h}^*) = \dim(\mathfrak{h}) = r = \mathrm{rank}(\mathfrak{g})$, we can therefore find r roots, $\{\alpha_{(i)} \in \Phi \mid i = 1, \ldots, r\}$, which provide a basis for the complex vector space $\mathfrak{h}^*$.

As such, we can write any root $\beta \in \Phi$ as $\beta = \sum_{i=1}^r \beta^i \alpha_{(i)}$, where $\beta^i \in \mathbb{C}$ can in general be complex.

Definition 63: We define a real subspace $\mathfrak{h}^*_{\mathbb{R}} \subset \mathfrak{h}^*$ as

$$\mathfrak{h}^*_{\mathbb{R}} := \mathrm{Span}_{\mathbb{R}} \{\alpha_{(i)}, i = 1, \ldots, r\}. \tag{7.80}$$

The coefficients β^i solve

$$(\beta, \alpha_{(j)}) = \sum_{i=1}^r \beta^i (\alpha_{(i)}, \alpha_{(j)}). \tag{7.81}$$

From the fact that $(\alpha, \beta) \in \mathbb{R}$, we can immediately infer that $\beta^i \in \mathbb{R}$. This is the same as saying that the root β actually lives in the real subspace $\mathfrak{h}^*_{\mathbb{R}}$:

$$\boxed{\beta \in \mathfrak{h}^*_{\mathbb{R}}, \qquad \forall \beta \in \Phi} \tag{7.82}$$

Consider now the inner product of two arbitrary elements $\lambda, \mu \in \mathfrak{h}^*_{\mathbb{R}}$. By linearity, this is related to the inner product of the basis,

$$(\lambda, \mu) = \sum_{i,j=1}^r \lambda^i \mu^j (\alpha_{(i)}, \alpha_{(j)}) \in \mathbb{R}. \tag{7.83}$$

Hence, the inner product of any two roots is real. From (7.75), we have the length of a vector as

$$(\lambda, \lambda) = \frac{1}{\mathcal{N}} \sum_{\delta \in \Phi} \lambda_i \delta^i \delta^j \lambda_j = \frac{1}{\mathcal{N}} \sum_{\delta \in \Phi} (\lambda, \delta)^2 \geq 0. \tag{7.84}$$

n	sign	possible angles θ	root relation
0	$\pm$	$\pi/2$	$(\alpha,\beta)=0$
4	$+$	0	$\alpha=\beta$
	$-$	π	$\alpha=-\beta$
1, 2, 3	$+$	$\pi/6, \pi/4, \pi/3$	$(\alpha,\beta)>0$
	$-$	$2\pi/3, 3\pi/4, 5\pi/6$	$(\alpha,\beta)<0$

Table 7.1: Possible angles and signs of the inner product between two roots α and β according to (7.88).

As a result of the non-degeneracy of the inner product κ, we have thus

$$|\lambda|^2 := (\lambda,\lambda) = 0 \iff (\lambda,\delta) = 0 \quad \forall\delta\in\Phi \iff \lambda = 0. \tag{7.85}$$

In summary, the roots $\alpha\in\Phi$ live in the *real vector space* $\mathfrak{h}^*_{\mathbb{R}} \simeq \mathbb{R}^r$, where $r = \mathrm{Rank}(\mathfrak{g})$, equipped with a *Euclidean inner product*: For all vectors $\lambda,\mu\in\mathfrak{h}^*_{\mathbb{R}}$,

1. $(\lambda,\mu)\in\mathbb{R}$
2. $(\lambda,\lambda)\geq 0$
3. $(\lambda,\lambda)=0 \iff \lambda=0$

Definition 64: As $(\alpha,\alpha)>0 \quad \forall\alpha\in\Phi$, we can define a *length* $|\alpha| := +(\alpha,\alpha)^{1/2} > 0$.

The inner product of two vectors α,β takes the standard form

$$(\alpha,\beta) = |\alpha||\beta|\cos\theta_{\alpha,\beta} \tag{7.86}$$

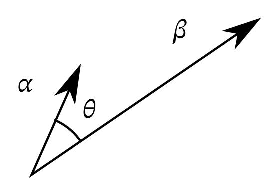


Figure 7.1

The *angle* $\theta_{\alpha,\beta}\in[0,\pi]$ is constrained by quantisation (7.70):

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = \frac{2|\beta|}{|\alpha|}\cos\theta_{\alpha,\beta}\in\mathbb{Z} \tag{7.87a}$$

$$\frac{2(\beta,\alpha)}{(\beta,\beta)} = \frac{2|\alpha|}{|\beta|}\cos\theta_{\alpha,\beta}\in\mathbb{Z} \tag{7.87b}$$

Multiplying (7.87a) $\times$ (7.87b), we find that $4\cos^2\theta_{\alpha,\beta}\in\mathbb{Z}$, which implies that

$$\cos\theta_{\alpha,\beta} = \pm\frac{1}{2}\sqrt{n}. \tag{7.88}$$

Solutions to this can only be found for $n\in\{0,1,2,3,4\}$ and are enumerated in Table 7.1.

Exercise 7.2: Show that if $\alpha\in\Phi$, then $k\alpha\in\Phi$ only for $k=\pm1$.

7.8 Simple Roots

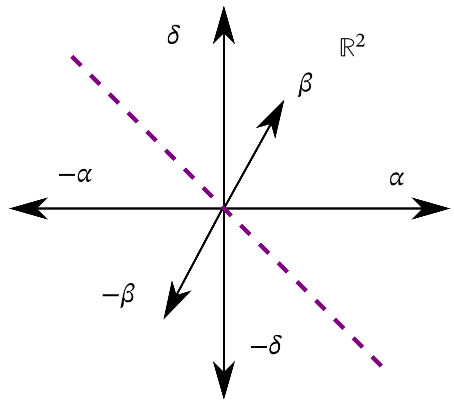


Figure 7.2: In $\mathbb{R}^2$ we can choose a line to separate the positive from the negative roots.

The root system is some finite set of root vectors living in $\mathbb{R}^r$. If $\beta \in \Phi$, then so is $-\beta \in \Phi$. This motivates the following definition.

Definition 65 (positive roots)**:** We choose some plane in $\mathfrak{h}^*_{\mathbb{R}}$ of dimension $\mathbb{R}^{r-1}$ in which none of the root vectors lie, such as illustrated in 7.2 for the case of $\mathbb{R}^2$. We can always find such a plane since there are finitely many roots, but a continuous infinity of choices of such a plane. This divides the roots into positive and negative roots

$$\Phi = \Phi_+ \cup \Phi_-. \tag{7.89}$$

Positive and negative roots satisfy the following relations:

1. $\alpha \in \Phi_+ \iff -\alpha \in \Phi_-$,
2. If $\alpha + \beta \in \Phi$ and $\alpha, \beta \in \Phi_\pm$, then $\alpha + \beta \in \Phi_\pm$.

Definition 66 (simple root)**:** A *simple root* $\delta \in \Phi_S$ is a positive root that cannot be written as a sum of two positive roots:

$$\delta \in \Phi_S \qquad \iff \qquad \delta \in \Phi_+ \quad \wedge \quad \delta \neq \alpha + \beta \quad \forall \alpha, \beta \in \Phi_+ \tag{7.90}$$

We will see that these simple roots form a particularly simple version of the Cartan-Weyl basis, which we will use to completely determine the Lie algebra.

Definition 67 (height)**:** The *height* of a root is the perpendicular distance from the hyperplane, see Fig. 7.2.

Claim 26: The difference of two simple roots is not a root.

Proof. Let $\alpha, \beta \in \Phi_S$. Suppose for contradiction that $\alpha - \beta \in \Phi$.

Either $\alpha - \beta \in \Phi_+$ then $\alpha = \alpha - \beta + \beta$ is not simple ※

Or $\beta - \alpha \in \Phi_+$ then $\beta = \beta - \alpha + \alpha$ is not simple ※

□

Claim 27: If $\alpha, \beta \in \Phi_S$, then the α-string through β has length

$$l_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}. \tag{7.91}$$

Proof. The α-string passing through β is $S_{\alpha,\beta} = \{\beta + n\alpha \in \Phi,\ n \in \mathbb{Z},\ n_- \leq n \leq n_+\}$ with $n_+ \geq 0$ and $n_- \leq 0$. Now Eq. (7.70) implies that $(n_+ + n_-) = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$. From Claim 26, the difference $\beta - \alpha \notin \Phi$ is not a root. Therefore the value $n = -1$ is not allowed, and since the root string is unbroken, we must have $n_- = 0$. Again from (7.70), we have

$$n_+ = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}_0 \tag{7.92}$$

The length of the string is the number of roots in the root string:

$$l_{\alpha,\beta} = n_+ - n_- + 1 = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}^+ \tag{7.93}$$

□

Corollary: From (7.92), we find that the inner product of simple roots is zero or negative[1]

$$(\alpha,\beta) \leq 0 \qquad \forall \alpha,\beta \in \Phi_S \tag{7.94}$$

Claim 28: Any positive root $\beta \in \Phi_+$ can be written as

$$\beta = \sum_i c_i \alpha_{(i)}, \tag{7.95}$$

an integer combination of simple roots, where $\alpha_{(i)} \in \Phi_S$ and $c_i \in \mathbb{Z}_{\geq 0}$

Proof. The statement is true if $\beta \in \Phi_S$ is a simple root. If $\beta \notin \Phi_S$, then we must have that $\beta = \beta_1 + \beta_2$ for $\beta_1, \beta_2 \in \Phi_+$. Now if $\beta_1, \beta_2 \in \Phi_S$, then we are done. If not, then we subdivide again. We repeat this; this process can only terminate on a linear combination of simple roots as in the claim. Since the space is finite dimensional, and there is a highest root, then this process must indeed terminate. □

Remark: Similarly, we can decompose negative roots in terms of negative simple roots.

[1] This means that their angle $\theta_{\alpha,\beta} \geq \pi/2$ is right or obtuse.

This is the nice property of the simple roots! Any other roots can be written as a linear combination of them. To show that they are the basis, we still need to show the following:

Claim 29: Simple roots are linearly independent.

Proof. Let $\lambda = \sum_{i\in I} c_i\alpha_{(i)}$, with $c_i \in \mathbb{R}\setminus\{0\}$, $\alpha_{(i)} \in \Phi_S$ and the index set is $I \subset \Phi_S$. Then simple roots being linearly independent means that $\lambda \neq 0$ unless $c_i = 0$ for all i. Define $\lambda_+ = \sum_{i\in I_+} c_i\alpha_i$, and $\lambda_- = -\sum_{i\in I_+} c_i\alpha_{(i)}$, with $I = I_+ \cup I_-$ and $I_+ = \{i \in I \mid c_i > 0\}$, $I_- = \{i \in I \mid c_i < 0\}$. Now $\lambda = \lambda_+ - \lambda_-$ where λ_+, λ_- not both zero

$$
\begin{aligned}
(\lambda,\lambda) &= (\lambda_+,\lambda_+) + (\lambda_-,\lambda_-) - 2(\lambda_+,\lambda_-) && (7.96)\\
&> -2(\lambda_+,\lambda_-) && (7.97)\\
&= 2\sum_{i\in I_+}\sum_{j\in I_-} c_i c_j(\alpha_{(i)},\alpha_{(j)}) && (7.98)
\end{aligned}
$$

Now $(\alpha_{(i)},\alpha_{(j)}) \leq 0$. Moreover, $c_i > 0$, $c_j > 0$. Therefore, we have that $(\lambda,\lambda) > 0$. Since we have a non-degenerate inner product, we must have $\lambda \neq 0$. □

Finally, to form a basis, we need to make sure that there are enough of them:

Claim 30: There are exactly $|\Phi_S| = r = \text{Rank}[\mathfrak{g}]$ simple roots.

Proof. By Claim 29, simple roots are linearly independent and therefore $|\Phi_S| \leq r$. Suppose for contradiction that $|\Phi_S| < r$. Then there would exist a vector $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$, which is orthogonal to all simple roots:

$$
(\lambda,\alpha) = 0 \qquad \forall\alpha \in \Phi_S. \tag{7.99}
$$

But all roots are either positive or negative. And by Claim 28, we can write all roots as linear combinations of simple roots. Therefore, λ is orthogonal to all roots

$$
(\lambda,\alpha) = 0 \qquad \forall\alpha \in \Phi. \tag{7.100}
$$

Then $H_\lambda = \lambda_i H^i \in \mathfrak{h}$

$$
\begin{aligned}
[H_\lambda, H] &= 0 \qquad \forall H \in \mathfrak{h} && (7.101)\\
[H_\lambda E^\alpha] &= (\lambda,\alpha)E^\alpha = 0 \qquad \forall\alpha \in \Phi && (7.102)
\end{aligned}
$$

This would imply that $\mathfrak{g}$ has a non-trivial ideal $\mathfrak{i} = \text{Span}_{\mathbb{C}}\{H_\lambda\}$, contradicting the assumption that $\mathfrak{g}$ is simple. □

Corollary: We can now choose the simple roots as a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

$$
B = \{\alpha \in \Phi_S\} = \left\{\alpha_{(i)} \,\middle|\, i = 1,\ldots,r\right\}. \tag{7.103}
$$

7.9 The Cartan Matrix

Definition 68: The inner products between the simple roots are encoded in the *Cartan matrix*

$$A^{ij} = 2\frac{(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})}. \tag{7.104}$$

Now we know from our discussion at the beginning of Section 7.6 that for each simple root $\alpha_{(i)}$ we have an associated $\mathfrak{su}_{\mathbb{C}}(2)$ subalgebra with generators

$$\{h^i = h^{\alpha_{(i)}}, \quad e^i_\pm = e^{\pm\alpha_{(i)}}\}, \qquad i = 1, \dots, r. \tag{7.105}$$

The corresponding brackets (7.64) become

$$\boxed{\begin{aligned} &[h^i, h^j] = 0 && && (7.106)\\ &[h^i, e^j_\pm] = \pm A^{ji} e^j_\pm && \text{(no summation)} && (7.107)\\ &[e^i_+, e^j_-] = \delta^{ij} h^i && && (7.108)\end{aligned}}$$

This is the *Chevalley basis*. We will see that we can restrain the Cartan matrix very heavily, and that any finite-dimensional simple Lie algebra is *uniquely* determined by its Cartan matrix.

There are no arbitrary structure constants in this set of brackets. In fact, this is not yet a basis (despite us naming it that); we only have $2 \times \mathrm{rank}(\mathfrak{g})$ step generators associated with the simple roots. In general, the number of step generators is $\dim(\mathfrak{g}) - \mathrm{rank}(\mathfrak{g})$, which is typically greater than $2r$.

The full algebra is generated by repeated brackets subject to the *Chevally-Serre relations*:

$$\boxed{(\mathrm{ad}_{e^i_\pm})^{1-A^{ji}} e^j_\pm = 0} \qquad \forall i, j = 1, \dots, r. \tag{7.109}$$

This encodes the fact that the '$\alpha_{(i)}$ string through $\alpha_{(j)}$' has length

$$l_{ij} = 1 - \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(i)}, \alpha_{(i)})} = 1 - A^{ji}, \tag{7.110}$$

which follows from Claim 27.

Let us see how the full algebra is generated. Consider first

$$[e^i_+, e^j_+] = [e^{+\alpha_{(i)}}, e^{+\alpha_{(j)}}] \propto \begin{cases} e^{\alpha_{(j)}+\alpha_{(i)}} & \alpha_{(i)} + \alpha_{(j)} \in \Phi \\ 0 & \text{else} \end{cases} \tag{7.111}$$

Applying $\mathrm{ad}_{e_+^i}$ again, we get

$$[e_+^i, [e_+^i, e_+^j]] \propto e^{\alpha_{(j)}+2\alpha_{(i)}} \qquad \text{if } \alpha_{(j)} + 2\alpha_{(i)} \in \Phi \tag{7.112}$$

Repeatedly applying the bracket, we find

$$(\mathrm{ad}_{e_+^i})^n e_+^j = \underbrace{[e_+^i, \dots [e_+^i, [e_+^i,}_{n \text{ times}} e_+^j]]] \propto e^{\alpha_{(j)}+n\alpha_{(i)}} \qquad \text{if } \alpha_{(j)} + n\alpha_{(i)} \in \Phi \tag{7.113}$$

Hence, by the Chevally-Serre relations, $\alpha_{(j)} + n\alpha_{(i)} \in \Phi$ if and only if $n < 1 - A^{ji} = l_{ij}$.

7.10 Reconstructing $\mathfrak{g}$ from A^{ij}

We will now proceed with the Cartan classification in multiple parts. Starting with generators corresponding with the simple roots, we generate other roots with the root strings. Their brackets are given by iterating the Jacobi identities and the Serre-Chevally relations.

Claim 31: A finite-dimensional simple complex Lie algebra is uniquely determined by its Cartan matrix.

If this is true, then to classify Lie algebras, we want to find out about these Cartan matrices.

7.10.1 Constraints on the Cartan Matrix

The Cartan matrix $A^{ij} = 2\frac{(\alpha_{(i)},\alpha_{(j)})}{(\alpha_{(j)},\alpha_{(j)})} \in \mathbb{Z}$ has the following properties:

a) $A^{ii} = 2$ for $i = 1, \dots, r$

b) $A^{ij} = 0 \iff A^{ji} = 0$; zeros on the off-diagonal come in pairs

c) Claim 7.8 tells us that $A^{ij} \in \mathbb{Z}_{\leq 0}$ for $i \neq j$

d) $\det A > 0$

e) $\mathfrak{g}$ is simple $\Rightarrow$ A is irreducible, meaning it is not block diagonal:

$$A \neq \begin{pmatrix} A^{(1)} & 0 \\ 0 & A^{(2)} \end{pmatrix} \tag{7.114}$$

Proof of d). Let $\lambda, \mu \in \mathfrak{h}^*_{\mathbb{R}}$. We employ Einstein summation convention $\lambda = \sum_{i=1}^r \lambda^i \alpha_{(i)} = \lambda^i a_{(i)}$. Then $(\lambda, \mu) = (\kappa^{-1})_{ij} \kappa^i \mu^j$ with $(\kappa^{-1})^{ij} = (\alpha_{(i)}, \alpha_{(j)})$, with $i, j = 1, \dots, r$. The inverse Killing form κ^{-1} is a real symmetric matrix. This means we can diagonalise it with some orthogonal matrix O

$$O(\kappa^{-1})O^T = \mathrm{diag}(l_1, \dots, l_r), \qquad l_i \in \mathbb{R}. \tag{7.115}$$

For each eigenvalue l, there is an eigenvector $v_l = v_l^i \alpha_{(i)}$ which satisfies

$$(\kappa^{-1})_{ij} v_l^j = l \delta_{ij} v_l^j \tag{7.116}$$

Its length squared is therefore

$$(v_l, v_l) = (\kappa^{-1})_{ij} v_l^i v_l^j = l \delta_{ij} v_l^i v_l^j = |v_l|^2 > 0 \tag{7.117}$$

Hence, every eigenvalue $l > 0$ is positive, and $\det(\kappa^{-1}) > 0$. Now the Cartan matrix is $A^{ij} = S^{ik} D_k{}^j$, where $S^{ik} = (\alpha_{(i)}, \alpha_{(j)}) = (\kappa^{-1})_{ik}$ and $D_k{}^j = \frac{2}{(\alpha_{(j)}, \alpha_{(j)})} \delta_k^j$. Hence $\det A = \det(D) \det(S) > 0$. □

Proof of e). Suppose that $A = \begin{pmatrix} A^{(1)} & 0 \\ 0 & A^{(2)} \end{pmatrix}$ is reducible. The simple root set then decomposes into $\Phi_S = \Phi^{(1)} \cup \Phi^{(2)}$ with $(\alpha, \beta) = 0$ whenever $\alpha \in \Phi^{(1)}$ and $\beta \in \Phi^{(2)}$. As the simple roots span the root space, we have $\mathfrak{h}^*_{\mathbb{R}} = R_1 \oplus R_2$, where $R_i = \mathrm{Span}_{\mathbb{R}} \{\alpha \in \Phi^{(i)}\}$. As the inner product is *Euclidean*, these are orthogonal $R_2 = (R_1)_\perp$ and their intersection $R_1 \cap R_2 = \emptyset$ is empty. Now consider

$$\mathfrak{g}_1 = \mathrm{Span}_{\mathbb{C}} \left\{ h^\alpha, e^\alpha \;\middle|\; \alpha \in \Phi^{(1)} \right\}. \tag{7.118}$$

Claim 32: $\mathfrak{g}_1$ is an ideal of $\mathfrak{g}$.

> *Proof.* Check $[X, Y] \in \mathfrak{g}_1$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}_1$. Any non-trivial case $X = e^{\alpha'}$, where $\alpha' \in R_2 = (R_1)_\perp$ gives
>
> $$[e^{\alpha'}, h^\alpha] = -2 \frac{(\alpha, \alpha')}{(\alpha, \alpha)} e^{\alpha'} = 0 \qquad \forall \alpha \in R_1 \tag{7.119}$$
>
> $$[e^{\alpha'}, e^\alpha] \begin{cases} \propto e^{\alpha' + \alpha} & \alpha + \alpha' \in \Phi \\ = 0 & \text{otherwise.} \end{cases} \tag{7.120}$$
>
> □

Reconstructing the root system with root strings we find for $\alpha \in \Phi^{(1)}$ and $\beta \in \Phi^{(2)}$

$$l_{\alpha,\beta} = 1 - 2 \frac{(\alpha, \beta)}{(\alpha, \alpha)} = 1, \tag{7.121}$$

which means that all roots $\alpha \in \Phi$ belong to either $\alpha \in R_1$ or $\alpha \in R_2$. Then consider $\mathfrak{g}_1 = \mathrm{Span}_{\mathbb{C}} \{h^\alpha, e^\alpha \mid \alpha \in R_1\}$.

Claim 33: $\mathfrak{g}_1$ is a non-trivial ideal of $\mathfrak{g}$[1].

Proof. Check $[X, Y] \in \mathfrak{g}_1$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}_1$. The only non-trivial bracket is $X = e^\beta$ for $\beta \in R_2$ and $Y = e^\alpha$ for $\alpha \in R_1$

$$[e^\beta, e^\alpha] \begin{cases} \propto e^{\beta+\alpha} & \beta + \alpha \in \Phi \\ = 0 & \text{otherwise} \end{cases} \tag{7.122}$$

Now if $\alpha \in R_1$ and $\beta \in R_2$ then $\alpha + \beta \neq R_1 \cup R_2$ or $R_1 \cap R_2$ is non-empty. Hence $[e^\beta, e^\alpha] = 0$. □

Therefore, A being reducible implies that $\mathfrak{g}$ is not simple. Hence, $\mathfrak{g}$ being simple implies that A is irreducible. □

For $r = \text{rank}(\mathfrak{g}) = 1$ we simply have $A = 2$. However, for $r = 2$ we have a Cartan matrix of the form

$$\begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix} \qquad m, n \in \mathbb{N}_0 \tag{7.123}$$

From property d), we get that $4 - mn > 0$ and up to exchange of m and n we therefore have only the pairs $(1, 1), (1, 2), (1, 3)$. We want to represent the data of the Cartan matrix in a way that is invariant under this relabelling.

Exercise 7.3: Show that $A^{ij}A^{ji} = \{0, 1, 2, 3\}$ for $i \neq j$ and enumerate the possible solutions for A^{ij}.

Exercise 7.4: Show that a simple Lie algebra $\mathfrak{g}$ has simple roots of at most two different lengths.

7.11 Dynkin Diagrams

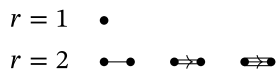
Dynkin diagrams the data contained in the Cartan matrix in diagrammatic form.

1. Draw a node for each $\alpha_{(i)} \in \Phi_S$, $i = 1, \ldots, r$.
2. Join nodes $\alpha_{(i)}$ and $\alpha_{(j)}$ with $\max(|A|^{ij}, |A|^{ji}) \in \{0, 1, 2, 3\}$ lines.
3. If roots have different lengths, then draw an array from long to short[2].

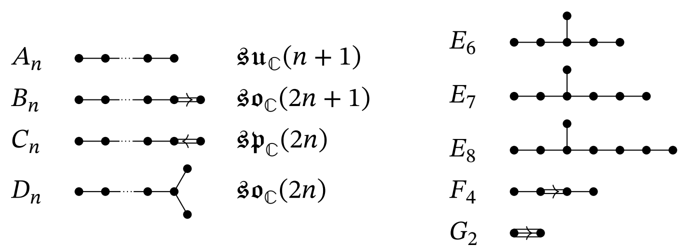
[1] See Gutowski Proposition 6.
[2] This distinction makes sense since there can only be two different simple root lengths.

If we relabel the points in the diagram, we get the same diagram. The process of enumerating the roots of the Lie algebra comes down to enumerating Dynkin diagrams. The classification of Cartan matrices / Dynkin diagrams satisfying constraints a) - e) is a purely combinatoric problem.

Let us translate what we learned at the level of the Cartan matrix for ranks $r = 1, 2$ into diagrams:



In 1894, Cartan showed that the diagrams are classified into infinite families labelled by rank$(\mathfrak{g}) = n \in \mathbb{N}^+$:



> **Note:** We have not talked about the symplectic group $Sp(2N)$ in this course.

For $n = 1$, we have $A_1 \simeq B_1 \simeq C_1$, corresponding to the fact that $\mathfrak{su}_{\mathbb{C}}(2) \simeq \mathfrak{so}_{\mathbb{C}}(3) \simeq \mathfrak{sp}_{\mathbb{C}}(2)$. For $n = 2$, we have $B_2 \simeq C_2$ and thus $\mathfrak{so}_{\mathbb{C}}(5) \simeq \mathfrak{sp}_{\mathbb{C}}(4)$. For $n = 3$, $D_3 \simeq A_3$, meaning that $\mathfrak{so}_{\mathbb{C}}(6) \simeq \mathfrak{su}_{\mathbb{C}}(4)$.

Definition 69 (level)**:** The *level* of a root is the weight in its root string expansion.

7.12 Reconstructing $\mathfrak{g}$ from A^{ij}

The Cartan matrix determines[1] the simple roots $\alpha_{(i)}$, $i = 1, \dots, r = \text{rank}(\mathfrak{g})$ via

$$A^{ij} = \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})} = \frac{2\left|\alpha_{(i)}\right|}{\left|\alpha_{(j)}\right|} \cos(\varphi_{ij}) \tag{7.124}$$

We have the same picture as in Fig. 7.1. The remaining roots are found by constructing root strings. As we know from Claim 27, the length of the $\alpha_{(i)}$-string through $\alpha_{(j)}$ is $l_{ij} = 1 - A_{ji} \in \mathbb{N}^+$.

[1] up to choice of one vector (such as $\alpha_{(1)}$) in $\mathbb{R}^r$

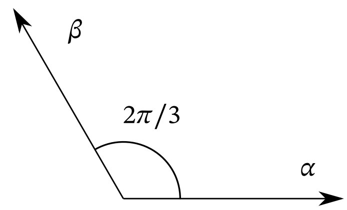


Figure 7.3: The simple roots of A_2 obey $|\alpha| = |\beta|$.

7.12.1 Root System for A_2

Let $\mathfrak{g} = A_2 = \mathfrak{su}_{\mathbb{C}}(3)$. The Dynkin diagram for A_2 is $\bullet\!-\!\bullet$, which encodes that the Cartan matrix is $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. In other words, A_2 has two simple roots α, β with

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = \frac{2(\beta,\alpha)}{(\beta,\beta)} = -1. \tag{7.125}$$

Therefore, the roots have the same length $|\alpha| = |\beta|$ and $\cos(\theta_{\alpha,\beta}) = -1/2$, which means that the angle between the two simple roots is $\theta = 2\pi/3$. This geometry is illustrated in Fig. 7.3.

To find the remaining roots we use the following facts about roots and root strings that we proved earlier:

1. For $\alpha, \beta \in \Phi_S$ the difference $\pm(\alpha - \beta) \notin \Phi$.
2. The length of the α-string through β is
$$l_{\alpha,\beta} = 1 - 2\frac{(\alpha,\beta)}{(\alpha,\alpha)} = 2 \tag{7.126}$$
and the length of the β-string through α is
$$l_{\beta,\alpha} = 1 - 2\frac{(\beta,\alpha)}{(\beta,\beta)} = 2, \tag{7.127}$$
so $\beta + n\alpha$ and $\alpha + n'\beta$ are roots for $n, n' \in \{0,1\}$. So we find that other than α and β, we also have $\alpha + \beta \in \Phi$.
3. If $\alpha \in \Phi$, then $-\alpha \in \Phi$. This means that we also have $-\alpha, -\beta, -(\alpha+\beta) \in \Phi$.
4. The α- and β-strings through $\alpha + \beta$ (and vice-versa) yield no additional roots.

The new root $\alpha + \beta$ has squared length

$$\begin{aligned}(\alpha+\beta, \alpha+\beta) &= (\alpha,\alpha) + (\beta,\beta) + 2(\alpha,\beta) && (7.128)\\ &= (\alpha,\alpha)[1+1-1] = (\alpha,\alpha) && (7.129)\end{aligned}$$

In summary, the root set is

$$\Phi = \{\pm\alpha, \pm\beta, \pm(\alpha+\beta)\}, \tag{7.130}$$

where all the roots have the same length $|\alpha| = |\beta| = |\alpha+\beta|$. This is illustrated in 7.4. The CW

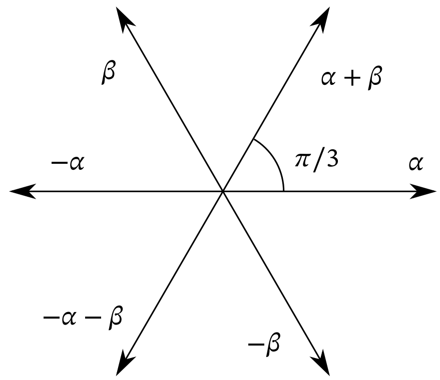


Figure 7.4: The root system of $A_2 \simeq \mathfrak{su}_{\mathbb{C}}(3)$.

basis for A_2 is $\{H^1, H^2, E^{\alpha}, E^{-\alpha}, E^{\beta}, E^{-\beta}, E^{\alpha+\beta}, E^{-(\alpha+\beta)}\}$. Since $\mathfrak{g} = A_2 \simeq \mathfrak{su}_{\mathbb{C}}(3)$, we have $\dim\mathfrak{g} = 3^2 - 1 = 8$, as expected.

Exercise 7.5: Find the root system of B_2.

8 Representation Theory II

To find the finite-dimensional rep of $\mathfrak{su}_{\mathbb{C}}(2)$, the idea was to work in the CW basis and work with step operators. Because we have now generalised the CW basis to general simple Lie algebras, we can use the same idea.

8.1 Weights

Let R be a representation of $\mathfrak{g}$ with dimension N. For example,

$$\begin{matrix} H^i \\ E^\alpha \end{matrix} \quad \longmapsto \quad \begin{matrix} R(H^i) \\ R(E^\alpha) \end{matrix} \quad \in \mathrm{Mat}_N(\mathbb{C}) \qquad i = 1, \dots, r \quad \alpha \in \Phi \tag{8.1}$$

The H^i are *diagonalisable* in the sense that $[H^i, H^j] = 0$. The representation matrices also commute since $[R(H^i), R(H^j)] = R([H^i, H^j]) = 0$. So if one is diagonalisable, then all of them are. We will assume that $R(H^i)$ are diagonalisable for $i = 1, \dots, r$. The fact that they commute with each other means that $\{R(H^i)\}$ are *simultaneously diagonalisable*.

This means that the vector space $V \simeq \mathbb{C}^N$ on which they act is spanned by the simultaneous eigenvectors of $\{R(H^i)\}$. Just as for the $\mathfrak{su}_{\mathbb{C}}(2)$ case, we can write our representation space V as

$$V = \bigoplus_{\lambda \in S_R} V_\lambda \tag{8.2}$$

a decomposition into a direct sum of linear subspaces associated with the eigenvalues λ, where $\forall v \in V_\lambda$, we have

$$R(H^i)v = \lambda^i v \qquad \begin{matrix} \lambda^i \in \mathbb{C} \\ i = 1, \dots, r. \end{matrix} \tag{8.3}$$

For $\mathfrak{su}_{\mathbb{C}}(2)$, we called the eigenvalue λ the *weights*. In this case, they are vectors rather than just numbers.

Definition 70: The eigenvalue $\lambda \in \mathfrak{h}^*$ is a *weight* of the representation R. We denote by S_R the associated weight set.

Remark: Weights $\lambda \in S_R \subset \mathfrak{h}^*$ have multiplicities $m_\lambda = \dim V_\lambda \geq 1$.

This motivates the following rephrasing of Def. 60:

Definition 71: Roots $\alpha \in \Phi$ are the weights of the adjoint representation $R(X) = ad_X$.

The key point for $\mathfrak{su}_{\mathbb{C}}(2)$ was that we generated the other weight vectors from the highest weight by applying the step operators. We likewise consider the action of step operators $R(E^\alpha)$, with root $\lambda \in \Phi$, on vectors v belonging to a particular subspace V_λ with weight λ.

$$\begin{aligned} R(H^i)R(E^\alpha)v &= R(E^\alpha)R(H^i)v + [R(H^i), R(E^\alpha)]v && (8.4) \\ &= (\lambda^i + \alpha^i)R(E^\alpha)v && (8.5) \end{aligned}$$

> **Note:** This is similar to how we solve the quantum harmonic oscillator with step operators, showing that $a\,|n\rangle$ is an eigenvalue of H with energy $n-1$.

Thus, for all vectors $v \in V_\lambda$, we find that

$$R(E^\alpha)v \begin{cases} \in V_{\lambda+\alpha}, & \text{if } \lambda + \alpha \in S_R \\ = 0, & \text{otherwise} \end{cases} \tag{8.6}$$

Recall that for $\mathfrak{su}_{\mathbb{C}}(2)$, the weights were all integers. Here, we can use again the fact that the Lie algebra contains a lot of $\mathfrak{su}_{\mathbb{C}}(2)$ subalgebras. This allows us to say something precise about the weights: Consider the action of $\mathfrak{sl}(2)_\alpha$ of a particular $\mathfrak{su}_{\mathbb{C}}(2)$ subalgebra with generators $R(h^\alpha)$ and $R(e^{\pm\alpha})$ on V. Representation matrices obey the $\mathfrak{su}_{\mathbb{C}}(2)$ commutation relations

$$\begin{aligned} [R(h^\alpha), R(e^{\pm\alpha})] &= \pm 2R(e^{\pm\alpha}) && (8.7a) \\ [R(e^{+\alpha}), R(e^{-\alpha})] &= R(h^\alpha). && (8.7b) \end{aligned}$$

Each generator defines a linear map $V \to V$. This means that V is a representation space for some representation of $\mathfrak{sl}(2)_\alpha$. (The representation R_α is finite-dimensional but not necessarily irreducible). Recall the definition

$$h^\alpha = \frac{2}{(\alpha,\alpha)}H^\alpha, \qquad H^\alpha = (\kappa^{-1})_{ij}\alpha^i H^j \tag{8.8}$$

Using linearity, we have that $\forall v \in V_\lambda$,

$$R(h^\alpha)v = \frac{2}{(\alpha,\alpha)}(\kappa^{-1})_{ij}\alpha^j R(H^j)v \tag{8.9}$$

$$= \left(\frac{2}{(\alpha,\alpha)}(\kappa^{-1})_{ij}\alpha^i\lambda^j\right)v \tag{8.10}$$

$$= \frac{2(\alpha,\lambda)}{(\alpha,\alpha)}\cdot v \tag{8.11}$$

We learned something new; the weights λ have to obey the following quantisation condition:

$$\boxed{\frac{2(\alpha,\lambda)}{(\alpha,\alpha)} \in \mathbb{Z} \qquad \begin{array}{r}\forall\lambda \in S_R\\ \alpha \in \Phi\end{array}} \tag{8.12}$$

8.2 Root, Coroots, Weights, and their Lattices

Representation spaces are spanned by a basis of weight-vectors, which are simultaneous eigenvectors of the Cartan generators. We have established that for any R any weight $\lambda \in S_R$ has to satisfy (8.12). This means that we have a nice picture of the allowed weights: they lie on a lattice.

Definition 72: For each simple root $\alpha_{(i)}$, the *dual root* or *coroot* $\alpha^\vee_{(i)}$ is defined[1]

$$\alpha^\vee_{(i)} = \frac{2\alpha_{(i)}}{(\alpha_{(i)},\alpha_{(i)})}, \qquad i = 1,\dots,r \tag{8.13}$$

Using coroots, the defining equation (7.104) for the Cartan matrix becomes

$$A^{ij} := (\alpha_{(i)},\alpha^\vee_{(j)}) \tag{8.14}$$

When one considers the root space as consisting of real linear combinations of roots, then the dual space of the root space is called the *weight space*, and its elements are the *weights* of $\mathfrak{g}$. This turns out to be equivalent to our previous the definition of weights as eigenvalues of (representations of) the Cartan subalgebra generators. [F&S, Sec. 13.2]

We can choose the simple coroots $\left\{\alpha^\vee_{(i)}\right\}$ as a basis B of the root space.

Definition 73 (Dynkin basis)**:** The dual basis B^*, often referred to as the *Dynkin basis*, for the weight space $\mathcal{L}^{\vee *}[\mathfrak{g}] = \mathcal{L}_W[\mathfrak{g}]$ is

$$B^* = \left\{\omega_{(i)} \;\middle|\; i = 1,2,\dots,r\right\}, \tag{8.15}$$

where the basis vectors $\omega_{(i)}$ are known as *fundamental weights* and obey

$$(\alpha^\vee_{(i)},\omega_{(j)}) = \frac{2(\alpha_{(i)},\omega_{(j)})}{(\alpha_{(i)},\alpha_{(i)})} = \delta_{ij}. \tag{8.16}$$

[1] the upside-down wedge $\vee$ is \vee in LaTeX

Definition 74 (Dynkin labels)**:** For any weight $\lambda \in S_R \subset \mathcal{L}_W[\mathfrak{g}]$,

$$\lambda = \sum_{i=1}^{r} \lambda^i \omega_{(i)} \qquad \lambda^i \in \mathbb{Z}, \quad i = 1, \dots, r \tag{8.17}$$

The components λ^i of a weight λ in the Dynkin basis are called *Dynkin labels*.

Definition 75 (lattice)**:** The *lattice* associated to some discrete subset V_0 (without accumulation points) of a vector space is the set of all linear combinations of elements of V_0 with integral coefficients.

From Claim 28, we know that all roots $\beta \in \Phi$ are linear combinations of the simple roots, $\beta = \sum_{i=1}^{r} \beta^i \alpha_{(i)}$, where $\beta^i \in \mathbb{Z}$ are integral coefficients. This means that all roots lie in the *root lattice*

$$\mathcal{L}[\mathfrak{g}] := \operatorname{span}_{\mathbb{Z}} \{\Phi_s\} = \operatorname{span}_{\mathbb{Z}} \{\Phi\}. \tag{8.18}$$

Similarly, the integer span of the simple coroots is the *co-root lattice* $\mathcal{L}^\vee[\mathfrak{g}]$ and the integer span of the fundamental weights the *weight lattice* $\mathcal{L}_w[\mathfrak{g}]$. The weight-lattice is the lattice dual (over $\mathbb{Z}$) to the co-root lattice:

$$\mathcal{L}_w[\mathfrak{g}] := (\mathcal{L}^\vee[\mathfrak{g}])^* = \left\{\lambda \in \mathfrak{h}^*_{\mathbb{R}} \,\middle|\, (\lambda, \alpha^\vee) \in \mathbb{Z},\ \forall \alpha^\vee \in \mathcal{L}^\vee[\mathfrak{g}]\right\} \tag{8.19}$$

Example 8.2.1: In solid state physics, the momenta lie in the dual (or reciprocal) lattice of the position lattice.

By comparison with (8.12), we can see that if $\lambda \in \mathcal{L}_W[\mathfrak{g}]$, then $\frac{2(\alpha_{(i)},\lambda)}{(\alpha_{(i)},\alpha_{(i)})} \in \mathbb{Z}$ for $i = 1, \dots, r$. The weights of any finite dimensional representation R of $\mathfrak{g}$ lie in $\mathcal{L}_W[\mathfrak{g}]$.

As the simple roots span $\mathfrak{h}^*_{\mathbb{R}}$, there is some matrix B such that

$$\omega_{(i)} = \sum_{j=1}^{r} B_{ij} \alpha_{(j)}. \tag{8.20}$$

Substituting this into (8.16) gives $\sum_{k=1}^{r} \frac{2(\alpha_{(i)},\alpha_{(k)})}{(\alpha_{(i)},\alpha_{(i)})} B_{jk} = \delta_{ij}$. In other words, $\sum_{k=1}^{r} B_{jk} A^{ki} = \delta^i_j$, so B is the inverse of the Cartan matrix A .

Remark: There is no significance here to raising and lowering of indices, so we might as well write them all in a lowered way. However, we first introduced A with indices upstairs so we will keep it that way for now.

The inverse of the decomposition (8.20) is therefore given by

$$\boxed{\alpha_{(i)} = \sum_{j}^{r} A^{ij} \omega_{(j)}} \tag{8.21}$$

The Cartan matrix has integer entries. This means that the simple roots $\alpha_{(i)}$, and therefore all of the roots, actually also lie on the weight lattice $\alpha_{(i)} \in \mathcal{L}_\omega[\mathfrak{g}]$.

Example 8.2.2: Let $\mathfrak{g} = A_2$. From $A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ we can read off that our simple roots are given by

$$\alpha = \alpha_{(1)} = 2\omega_{(1)} - \omega_{(2)}, \qquad \beta = \alpha_{(2)} = -\omega_{(1)} + 2\omega_{(2)}. \tag{8.22}$$

Equivalently, we may write

$$\omega_{(1)} = \frac{1}{3}(2\alpha + \beta), \qquad \omega_{(2)} = \frac{1}{3}(\alpha + 2\beta). \tag{8.23}$$

This geometry is illustrated in Fig. 8.1.

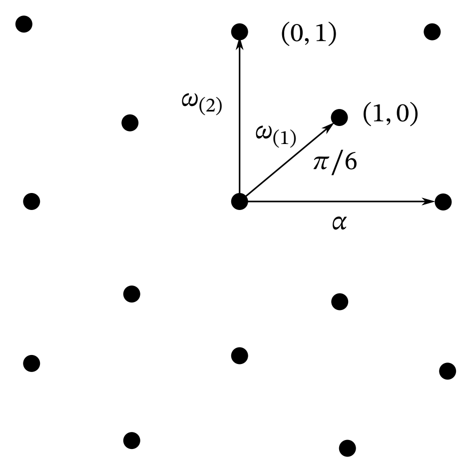


Figure 8.1: The weight lattice of A_2. (It's supposed to be hexagonal.)

8.3 Algorithm: Finding the Weight Set

Highest weight: Every finite-dimensional representation R of $\mathfrak{g}$ has a *highest weight*

$$\Lambda = \sum_{i=1}^{r} \Lambda^i \omega_{(i)} \in S_R \qquad \begin{array}{l} \Lambda^i \in \mathbb{Z} \\ \Lambda^i \geq 0 \end{array} \tag{8.24}$$

such that the eigenvector $v_\Lambda \in V$, which satisfies $R(H^i)v_\Lambda = \Lambda^i v_\Lambda$, has to be annihilated by all the step generators of the positive roots. This is because these increase the weight by taking it further away from the hyperplane that we use to split positive and negative roots.

$$R(E^\alpha)v_\Lambda = 0 \qquad \forall \alpha \in \Phi_+. \tag{8.25}$$

Definition 76: The Λ^i for the highest weight are the *Dynkin labels* of the representation R.

Remaining Weights: The remaining weights $\lambda \in S_R$ are generated by the *lowering operators*, which are the representatives of the step-down generators $R(E^{-\alpha})$ for $\alpha \in \Phi_+$. All remaining weights have the form

$$\lambda = \Lambda - \mu, \qquad \text{where } \mu = \sum_{i=1}^{r} \mu^i \alpha_{(i)}, \quad 0 \leq \mu^i \in \mathbb{N} \leq \lambda^i \tag{8.26}$$

For any finite dimensional representation of $\mathfrak{g}$. If $\lambda = \sum \lambda^i \omega_{(i)}$ is a weight, then $\lambda - m_i \alpha_{(i)}$ (no sum) is a weight as well, provided that $0 \leq m_i \in \mathbb{Z} \leq \lambda^i$, i.e. only if the associated Dynkin label is positive:

$$\boxed{\lambda = \lambda^i \omega_{(i)} \in S_R \;\Rightarrow\; \lambda - m_i \alpha_{(i)} \in S_R \quad \forall m_i, \lambda^i \in \mathbb{N} \text{ s.t. } 0 \leq m_i \leq \lambda^i.} \tag{8.27}$$

Starting from the highest weight Λ, we can use this to iteratively find all the remaining weights.

Example 8.3.1: For the *fundamental representation* $R_{(1,0)}$ of $A_2 \simeq \mathfrak{su}_{\mathbb{C}}(3)$, the Dynkin labels are $(\Lambda^1, \Lambda^2) = (1, 0)$.

- The highest weight is $\Lambda = \omega_{(1)} \in S_{(1,0)}$.
- Applying the algorithm, we find that

$$\omega_{(1)} \in S_{(1,0)} \;\Rightarrow\; \omega_{(1)} - \alpha_{(1)} = \omega_{(1)} - (2\omega_{(1)} - \omega_{(2)}) \tag{8.28}$$
$$= -\omega_{(1)} + \omega_{(2)} \in S_{(1,0)} \tag{8.29}$$

- We recursively repeat the algorithm with this new root (note that only the positive coefficients λ^i incur additional roots)

$$-\omega_{(1)} + \omega_{(2)} \in S_{(1,0)} \;\Rightarrow\; (-\omega_{(1)} + \omega_{(2)}) - \alpha_{(2)} = -\omega_{(1)} + \omega_{(2)} - (2\omega_{(2)} - \omega_{(1)}) \tag{8.30}$$
$$= -\omega_{(2)} \in S_{(1,0)} \tag{8.31}$$

- Since this root only has a negative coefficient $\lambda^2 = -1$, the algorithm terminates here

8.4 Irreducible Representations of $A_2 \simeq \mathfrak{su}_{\mathbb{C}}(3)$

Recall that the highest weight of an irreducible $\mathfrak{sl}(2)$ representation determines the dimension of that representation. The same is true for representations of A_2.

Claim 34: The representations $R_{(\Lambda_1,\Lambda_2)}$ of A_2 are labelled by the highest weight $\Lambda = \Lambda_1\omega_{(1)} + \Lambda_2\omega_{(2)}$, where $(\Lambda_1, \Lambda_2) \in \mathbb{N}_0^2$ are the Dynkin labels of the representation. Their dimensions are

$$\dim(R_{(\Lambda_1,\Lambda_2)}) = \frac{1}{2}(\Lambda_1 + 1)(\Lambda_2 + 1)(\Lambda_1 + \Lambda_2 + 2). \tag{8.32}$$

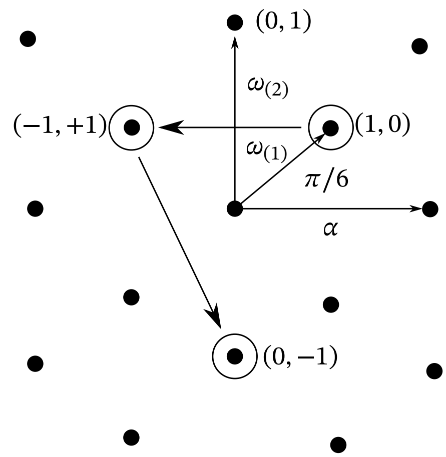


Figure 8.2: The weights of the fundamental representation $R_{(1,0)} = \mathbf{3}$ of $A_2 \simeq \mathfrak{su}_{\mathbb{C}}(3)$, generated by successively applying the algorithm to the highest weight $\Lambda = \omega_{(1)}$.

Corollary: This formula is symmetric under interchange of $\Lambda_1 \leftrightarrow \Lambda_2$. This means that whenever $\Lambda_1 \neq \Lambda_2$, then we get pairs of representations

$$R_{(\Lambda_2,\Lambda_1)} = \overline{R}_{(\Lambda_1,\Lambda_2)}. \tag{8.33}$$

Using the formula to compute the A_2 irreducible representations of lowest dimension, we obtain Table 8.1. Commonly, the irreducible representations are referred to by their dimension, written in a bold font. For instance, we denote the fundamental and anti-fundamental representations of A_2 as $R_{(1,0)} = \mathbf{3}$ and $R_{(0,1)} = \overline{\mathbf{3}}$ respectively.

A_2 irrep	dim	name
$R_{(0,0)}$	$\mathbf{1}$	trivial
$R_{(1,0)}$	$\mathbf{3}$	fundamental
$R_{(0,1)} = \overline{R}_{(1,0)}$	$\overline{\mathbf{3}}$	anti-fundamental
$R_{(2,0)}$	$\mathbf{6}$	
$R_{(0,2)} = \overline{R}_{(2,0)}$	$\overline{\mathbf{6}}$	
$R_{(1,1)}$	$\mathbf{8}$	adjoint
$R_{(3,0)}$	$\mathbf{10}$	
$R_{(0,3)} = \overline{R}_{(3,0)}$	$\overline{\mathbf{10}}$	

Table 8.1: Irreducible representations (irreps) of A_2 with lowest dimensions.

Since $\dim(A_2) = \dim(\mathfrak{su}_{\mathbb{C}}(3)) = 8$, the eight dimensional representation $R_{(1,1)} = \mathbf{8}$ is the natural candidate to be the adjoint representation. Starting from the highest weight vector $\omega_1 + \omega_2$ and successively applying the step down operator algorithm, we recover the other roots. This is shown

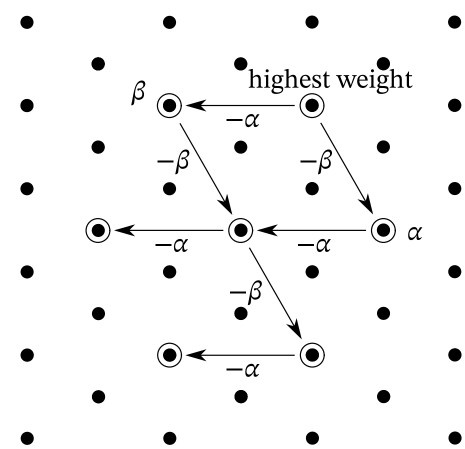


Figure 8.3: The weight lattice for the $R_{(1,1)}$ representation of A_2. Applying the step-down operators to the highest weight $\omega_1 + \omega_2 = \alpha + \beta$ gives the $\left|S_{(1,1)}\right| = 8$ weights.

in Figure 8.3. The weights of the adjoint representation are the roots. From the diagram, we see that the weight set, expressed in terms of simple roots α and β, is

$$S_{(1,1)} = \{\alpha + \beta, \alpha, \beta, 0, -\alpha, -\beta, -(\alpha + \beta)\}. \tag{8.34}$$

The representation space is just the Lie algebra itself, generated by the step operators and two Cartan elements

$$A_2 = \text{span}\left\{E^\alpha, E^{-\alpha}, E^\beta, E^\beta, E^{\alpha+\beta}, E^{-(\alpha+\beta)}, H^1, H^2\right\}. \tag{8.35}$$

Since $\dim R_{(1,1)} = 8$, and the 6 roots are non-degenerate by Claim 21, the weight $(0,0)$ must have multiplicity two, corresponding to the two Cartan elements H^1 and H^2. The weights (8.34) correspond to the generators of the Lie algebra, so $R_{(1,1)}$ is indeed the adjoint representation.

8.5 Tensor Products Revisited

The representation of a composite system, say of two particles, is given by the tensor product of the individual particle representations.

Let R_Λ and $R_{\Lambda'}$ be irreducible representations of the simple complex Lie algebra $\mathfrak{g}$. The corresponding representation spaces are V_Λ and $V_{\Lambda'}$, where Λ and Λ' denote the highest weight of the weight sets $S_\Lambda, S_{\Lambda'} \subset \mathfrak{h}^*_{\mathbb{R}}$. The representation space is a direct sum of subspaces corresponding to each weight

$$V_\Lambda = \bigoplus_{\lambda \in S_\Lambda} V_\lambda. \tag{8.36}$$

In Sec. 6.3.1 we defined the tensor product of two arbitrary representations.

Claim 35: If $\lambda \in S_\Lambda$ and $\lambda' \in S_{\Lambda'}$, then $\lambda+\lambda' \in \mathcal{L}_\omega[\mathfrak{g}]$ is a weight of the tensor product representation $R_\Lambda \oplus R_{\Lambda'}$.

Proof. Let $v_\lambda \in V_\lambda$ be a weight vector. Recall that this satisfies the eigenvalue equation

$$R_\Lambda(H^i)v_\lambda = \lambda^i v_\lambda. \tag{8.37}$$

And similarly for $v_{\lambda'} \in V_{\lambda'}$. Basis vectors for $V_{R_\Lambda \oplus R_{\Lambda'}} = V_\Lambda \oplus V_{\Lambda'}$ are tensor products $v_\lambda \oplus v_{\lambda'}$ of pairs of weight vectors from the original representation. Then by Definition 47 of the tensor product representation, we have

$$\begin{align}(R_\Lambda \otimes R_{\Lambda'})(H^i)(v_\lambda \otimes v_{\lambda'}) &= R_\Lambda(H^i)v_\lambda \otimes v_{\lambda'} + v_\lambda \otimes R_{\Lambda'} \otimes R_{\Lambda'}(H^i)v'_\lambda \tag{8.38}\\ &= (\lambda + \lambda')(v_\lambda \otimes v_{\lambda'}) \tag{8.39}\end{align}$$

□

For finite, simple, complex $\mathfrak{g}$, the tensor product representation is reducible:

$$R_\Lambda \otimes R_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathcal{L}_\omega} \mathcal{N}^{\Lambda''}_{\Lambda,\Lambda'} R_{\Lambda''}. \tag{8.40}$$

Example 8.5.1 ($\mathfrak{g} = A_2$)**:** Have $R_{(1,0)} \otimes R_{(1,0)}$. The weight set is $S_{(1,0)} = \{\omega_1, \omega_2 - \omega_1, -\omega_2\}$. Keeping track of the degeneracy of weights, here in particular ω_2 and $-\omega_1$, we first find the sum of the weight sets and then decompose it:

$$\begin{align}S_{(1,0)\otimes(1,0)} &= \{2\omega_1, \omega_2, \omega_2, \omega_1 - \omega_2, \omega_1 - \omega_2, -\omega_1, -\omega_1, -2\omega_1 - 2\omega_2, -2\omega_2\} \tag{8.41}\\ &= \{2\omega_1, \omega_2, \omega_1 - \omega_2, -\omega_1, -2\omega_1 - 2\omega_2, -2\omega_2\} \cup \{\omega_2, \omega_1 - \omega_2, -\omega_1\} \tag{8.42}\\ &= S_{(2,0)} \cup S_{(0,1)} \tag{8.43}\end{align}$$

This decomposition of the weight set implies that $R_{(1,0)} \otimes R_{(1,0)} = R_{(2,0)} \oplus R_{(0,1)}$. In the dimensional notation of Table 8.1, this translates into $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}$, which makes it obvious that the dimensions work out.

8.5.1 Strong Interactions, Quarks and the Eightfold Way

Not lectured. Taken from Prof. Manton's notes.

This is where our discussion comes back to the observation of the *eightfold way* from Sec. 1.4.2. Each of the weights corresponds to a particle, which transform according to the $SU(3)_{\text{flavour}}$ approximate symmetry. Since the rank of A_2 is 2, these particles are simultaneous eigenstates of two quantum numbers, which we choose to be *isospin* $I = H^1/2$ and *hypercharge* $Y = (H^1 + 2H^2)/3$. Electric charge is related to these by $Q = I + Y/2$. Particles in irreducible representations have approximately the same mass, and related strong interactions. The particles transforming under the fundamental representation **3** of flavour $SU(3)$ are the quarks q, while antiquarks $\overline{q}$ are in $\overline{\mathbf{3}}$. This is shown in Fig. 8.4.

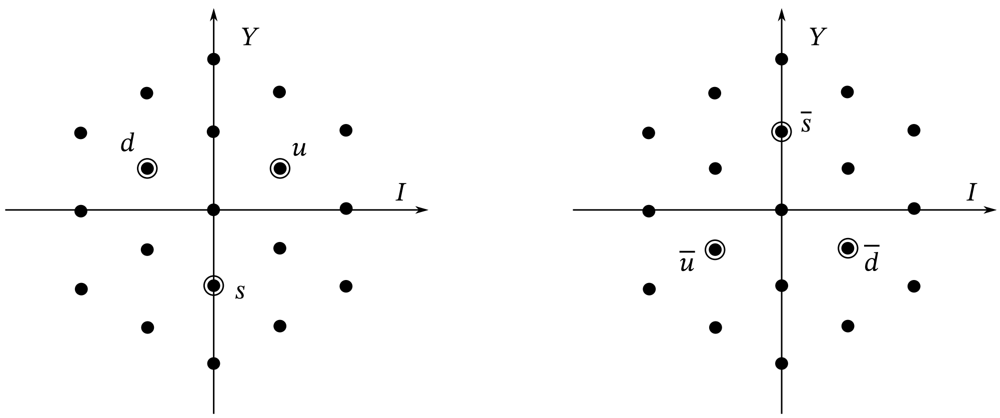


Figure 8.4: The quarks q transform under **3** and anti-quarks $\overline{q}$ under $\overline{\mathbf{3}}$ of flavour $SU(3)$.

All other hadrons, such as the mesons $q\overline{q}$ and the baryons qqq, are multi-quark states and can be obtained by tensor products. In particular, the particles in the *eightfold way* of Fig. 1.3 are the baryon octet **8**, which arises in the decomposition of $qqq \simeq \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$, alongside the baryon decuplet **10**. The $SO(3)$ gauge invariance of physical states in QCD explains why only colour singlets $\mathbf{3} \otimes \overline{\mathbf{3}}$ $(q\overline{q})$, $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ (qqq), and $\overline{\mathbf{3}} \otimes \overline{\mathbf{3}} \otimes \overline{\mathbf{3}}$ $(\overline{qqq})$ are observed, but does not tell us why colour singlets of $qq\overline{qq}$ or $qqqqqq$ states are not observed.

9 Gauge Theory

We will assume that most of the basics of Abelian gauge theory have been covered in other courses. We will develop these quickly and then move on to non-Abelian gauge theories.

Gauge theories have other kinds of symmetries than those that we have been dealing with; they are redundancies in the way we describe our system.

9.1 Electromagnetism

The classic example is the gauge invariance of electromagnetism. Consider a field

$$\mathcal{A}_\mu = \begin{cases} \Phi, & \mu = 0 \\ A_i, & \mu = i = 1,2,3 \end{cases} \tag{9.1}$$

Gauge transformations are $\mathcal{A}_\mu \to \mathcal{A}_\mu + \partial_\mu \alpha$, where $\alpha = \alpha(\mathbf{x}, t)$ and leave the fields invariant. In particular, they leave the field strength tensor $\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$ invariant.

The Lagrangian is

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4g^2} \mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu} \tag{9.2}$$

We will work with a redefined gauge potential, which is imaginary, and field strength tensor

$$A_\mu = -i\mathcal{A}_\mu \qquad F_{\mu\nu} = -i\mathcal{F}_{\mu\nu}. \tag{9.3}$$

We will use this to reconcile it with the fact that the gauge field naturally lives in the Lie algebra of the symmetry group, which we represented by complex numbers $e^{i\theta} \in \mathbb{C}$.

9.1.1 Coupling to Matter

Define a complex scalar field $\phi : \mathbb{R}^3 \to \mathbb{C}$ with the standard Lagrangian

$$\mathcal{L}_\phi = \partial_\mu \phi^* \partial^\mu \phi - W(\phi\phi^*). \tag{9.4}$$

Choosing the potential W to only depend on the modulus makes sure that this Lagrangian is invariant under a global $U(1)$ symmetry:

$$\begin{aligned} \phi &\to g\phi \\ \phi^* &\to g^{-1}\phi^* \end{aligned} \qquad g = \exp(i\delta) \in U(1), \quad \delta \in [0, 2\pi]. \tag{9.5}$$

Consider now the infinitesimal symmetry generator

$$g = \exp(\epsilon X) \simeq 1 + \epsilon X + O(\epsilon^2). \tag{9.6}$$

Working close to the identity element means taking $\epsilon \ll 1$. Moreover, $X \in \mathfrak{u}(1) \simeq i\mathbb{R}$. We write the infinitesimal symmetry transformation as

$$\begin{aligned} \phi &\to \phi + \delta_X \phi & \delta_X \phi &= \epsilon X \phi \\ \phi^* &\to \phi^* + \delta_X \phi^* & \delta_X \phi^* &= -\epsilon X \phi^*. \end{aligned} \tag{9.7}$$

We say that this transformation is a *symmetry* if the variation $\delta_X \mathcal{L}$ of the Lagrangian vanishes.

Now the symmetry condition $\delta_X \mathcal{L} = 0$ on the Lagrangian implies that we have a conserved Noether charge. To couple scalar field to electromagnetism, we gauge the $U(1)$: Previously, g was simply an element of $U(1)$. Now, we promote g to a field that depends on spacetime

$$g: \mathbb{R}^{3,1} \to U(1) \qquad \begin{aligned} \phi &\to g(x)\phi \\ \phi^* &\to g^{-1}(x)\phi^* \end{aligned} \tag{9.8}$$

Consider again the infinitesimal transformation $g = \exp(\epsilon X)$

$$\delta_X \phi = \epsilon X(x) \phi. \tag{9.9}$$

Once we gauge the symmetry, $\mathcal{L}_\phi$ is no longer invariant. To restore gauge invariance, we introduce a *covariant derivative* $D_\mu = \partial_\mu + A_\mu$, where we introduced the gauge field $A_\mu : \mathbb{R}^{3,1} \to \mathfrak{u}(1) \simeq i\mathbb{R}$, which lives in the Lie algebra of the gauge group. The gauge field transforms as

$$\delta_X A_\mu = -\epsilon \partial_\mu X. \tag{9.10}$$

We find a consistent coupling by taking the invariant Lagrangian

$$\mathcal{L} = \frac{1}{4g^2} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^* (D^\mu \phi) - W(\phi^* \phi), \tag{9.11}$$

where we replaced the ordinary derivatives by covariant ones. In fact, we know that this is the correct Lagrangian since the variation with respect to the gauge fields gives the correct classical equations of motion. As far as we know, this is the unique way of getting renormalisable spin-1 interacting theories for $U(1)$. We want to generalise this successful Lagrangian to other symmetry groups.

9.2 Non-Abelian Gauge Theory

Consider now a general gauge symmetry γ based on a non-Abelian Lie group G. As in electromagnetism, we start with a global symmetry: We obtain a global symmetry by choosing a representation D of the Lie group G of dimension N. In general, we will think of a complex representation; in this case we can identify the representation space with $V \simeq \mathbb{C}^N$. We will define the standard Hermitian inner product $(u, v) := \mathbf{u}^\dagger \cdot \mathbf{v}, \forall u, v \in V$. If we then have a field $\phi : \mathbb{R}^{3,1} \to V$, then the Lagrangian $\mathcal{L}_\phi = (\partial_\mu \phi, \partial^\mu \phi) - W[(\phi, \phi)]$ generalises the analysis of the $U(1)$ case. The transformation of the field

$$\phi \to D(g)\phi \qquad \forall g \in G \tag{9.12}$$

is defined to be a *global non-Abelian symmetry transformation* if it preserves $\mathcal{L}_\phi$. We will see that this places a restriction on the representation; the $D(g)$ have to be unitary matrices, i.e. $D(g)^\dagger D(g) = \mathbb{1}_N$. It is therefore quite important that G is a compact group, because only compact Lie groups have unitary representations of finite dimension.

As in the Abelian case, we want to focus on the infinitesimal transformations. In a Lie group, we do this by expanding near the identity:

$$g = \mathrm{Exp}(\epsilon X) \qquad \begin{array}{l} \epsilon \ll 1 \\ X \in \mathfrak{g} \end{array} \tag{9.13}$$

From our earlier analysis, we know that the representation $D : G \to \mathrm{Mat}_N(\mathbb{C})$ of the Lie group induces a representation $R : \mathfrak{g} \to \mathrm{Mat}_N(\mathbb{C})$ of the Lie algebra by

$$D(g) = \mathrm{Exp}(\epsilon R(X)). \tag{9.14}$$

Definition 77: A *unitary* representation $R(X)$ of the Lie algebra $\mathfrak{g}$ is anti-hermitian:

$$R(X) = -R(X)^\dagger, \qquad \forall X \in \mathfrak{g}. \tag{9.15}$$

Near the identity ($\epsilon \ll 1$), we have $D(g) \approx \mathbb{1}_N + \epsilon R(X) + O(\epsilon^2)$. This means we can write down an infinitesimal version of the symmetry transformation (9.12) as $\phi \to \phi + \delta_X \phi$ with $\delta_X \phi = \epsilon R(X)\phi \in V$.

We want to imitate the procedure of the Abelian case in electromagnetism to gauge the symmetry. We upgrade X to be spacetime dependent:

$$X : \mathbb{R}^{3,1} \to \mathfrak{g} \tag{9.16}$$

In particular, the symmetry transformation becomes

$$\delta_X \phi = \epsilon R(X(x))\phi \in V. \tag{9.17}$$

Under this transformation, $\mathcal{L}_\phi$ is no longer invariant. As before we introduce a gauge field $A_\mu : \mathbb{R}^{3,1} \to \mathfrak{g}$, living in the Lie algebra of the gauge group.

> The gauge transformation for a matrix Lie group $g \in G$ is
>
> $$A_\mu \to A'_\mu = g(A_\mu - \partial_\mu \ln g)g^{-1}. \tag{9.18}$$
>
> Note that inverting $g = \exp(\epsilon X)$ gives $\ln g = \epsilon X \in \mathfrak{g}$, so it makes sense to add the gauge field A_μ and the logarithm $\ln g$ (and its derivative). Moreover, if $X \in \mathfrak{g}$, then $gXg^{-1} \in \mathfrak{g}$, so that the transformed gauge field A'_μ also takes values in the Lie algebra.

The infinitesimal gauge transformation is

$$\delta_X A_\mu = -\epsilon\partial_\mu X + \epsilon[X, A_\mu] \tag{9.19}$$

The first term is the same as the Abelian gauge transformation. Indeed it is quite important that we recover Abelian gauge theory in the limit of our couplings going to zero. However, the symmetries also allow us to add the unique additional term (up to coefficients) on the right. It is an element of the Lie algebra, has the right dimension, Lorentz indices, etc…

We need to add this term because we want to define a covariant derivative.

Claim 36: The covariant derivative is

$$\boxed{D_\mu\phi = \partial_\mu\phi + R(A_\mu)\phi}. \tag{9.20}$$

In other words, for all spacetime dependent $X \in \mathfrak{g}$, we have

$$\delta_X(D_\mu\phi) = \epsilon R(X) D_\mu\phi. \tag{9.21}$$

Proof. We need to compute the variation in the covariant derivative. Variation is a linear operator satisfying the Leibniz rule:

$$\begin{aligned} \delta_X(D_\mu\phi) &= \delta_X(\partial_\mu\phi + R(A_\mu)\phi) && (9.22)\\ &= \partial_\mu(\delta_X\phi) + R(A_\mu)\phi_X\phi + R(\delta_X A_\mu)\phi. && (9.23) \end{aligned}$$

We have used the fact that derivatives and representations are linear operations, meaning that we can take the variation inside them.

Exercise 9.1: Verify this property!

Applying (9.19) to the latter terms, we have

$$\begin{aligned}
\ldots &= \partial_\mu(\epsilon R(X)\phi) + \epsilon R(A_\mu)R(X)\phi - \epsilon R(\partial_\mu X)\phi + R([X, A_\mu])\phi. && (9.24)\\
&= \cancel{\epsilon R(\partial_\mu X)\phi} + \epsilon R(X)\partial_\mu\phi + \epsilon R(X)R(A_\mu)\phi + \cancel{\epsilon[R(A_\mu, R(X))]\phi} - \cancel{\epsilon R(\partial_\mu X)\phi} + \cancel{\epsilon[R(X), R(A_\mu)]\phi} && (9.25)\\
&= \epsilon R(X)\partial_\mu\phi + \epsilon R(X)R(A_\mu)\phi && (9.26)\\
&= \epsilon R(X)D_\mu\phi. && (9.27)
\end{aligned}$$

□

This means that we have a good analogue of the covariant derivative of the Abelian case. Now the variation of the inner product vanishes:

$$\delta_X(D^\mu\phi, D_\mu\phi) = \epsilon(R(X)D_\mu\phi, D^\mu\phi) + \epsilon(D_\mu\phi, R(X)D^\mu\phi) = 0, \tag{9.28}$$

using the fact that for a unitary representation we have $R(X)^\dagger = -R(X)$. This is why we need to restrict to unitary representations and therefore compact Lie groups.

9.2.1 Action for the Gauge Field

The last thing we have to do is to find the appropriate generalisation of the Maxwell action for the non-Abelian gauge field $A_\mu : \mathbb{R}^{3,1} \to \mathfrak{g}$; again, gauge invariance is going to do most of the work for us. We define the field strength tensor to be

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]. \tag{9.29}$$

Claim 37: Only with these definitions does the field strength tensor transform with the adjoint representation, meaning that

$$\delta_X(F_{\mu\nu}) = \epsilon[X, F_{\mu\nu}] \in \mathfrak{g}. \tag{9.30}$$

Proof. The variation obeys the Leibniz rule and can be commuted with derivatives

$$\delta_X(F_{\mu\nu}) = \partial_\mu(\delta_X A_\nu) - \partial_\nu(\delta_X A_\mu) + [\delta_X A_\mu, A_\nu] + [A_\mu, \delta_X A_\nu]. \tag{9.31}$$

As we have four terms and each term gets a gauge variation with two terms, we will end up with an expression of eight terms; however, these will reduce as we will see now.

$$\begin{aligned}
\ldots = &\cancel{-\epsilon\partial_\mu\partial_\nu X} + \epsilon\partial_\mu([X, A_\nu]) + \cancel{\epsilon\partial_\nu\partial_\mu X} - \epsilon\partial_\nu(\cancel{[X, A_\mu]})\\
&- \epsilon\cancel{[\partial_\mu X, A_\nu]} - \epsilon\cancel{[A_\mu, \partial_\nu X]} + \epsilon[[X, A_\mu]A_\nu] + \epsilon[A_\mu, [X, A_\nu]] && (9.32)\\
= &\ \epsilon[X, \partial_\mu A_\nu] - \epsilon[X, \partial_\nu A_\mu] - \epsilon[A_\nu, [X, A_\mu]] + \epsilon[A_\mu, [A_\nu, X]] && (9.33)\\
= &\ \epsilon[X, \partial_\mu A_\nu - \partial_\nu A_\mu] + \epsilon[X, [A_\mu, A_\nu]] && (9.34)\\
= &\ \epsilon[X, F_{\mu\nu}], && (9.35)
\end{aligned}$$

where we have made use of the Jacobi identity. □

Corollary: In an Abelian Lie algebra, all the brackets will vanish and $F_{\mu\nu}$ is gauge invariant.

Definition 78: The *Yang-Mills Lagrangian* $\mathcal{L}_{\text{YM}}$ is defined by using the Killing form (which is the unique invariant inner product in a simple Lie algebra)

$$\mathcal{L}_{\text{YM}} = \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}), \tag{9.36}$$

where g is a coupling constant.

Claim 38: The Yang–Mills Lagrangian $\mathcal{L}_{\text{YM}}$ is gauge invariant.

Proof. We use the covariance property of Claim 37. We have $\forall X \in \mathfrak{g}$ that

$$\delta_X \mathcal{L}_{\text{YM}} = \frac{1}{g^2}\kappa(\delta_X F_{\mu\nu}, F^{\mu\nu}) + \frac{1}{g^2}\kappa(F^{\mu\nu}, \delta_X F^{\mu\nu}) \tag{9.37}$$

$$= \frac{1}{g^2}[\kappa([X, F_{\mu\nu}], F^{\mu\nu}) + \kappa(F_{\mu\nu}, [X, F^{\mu\nu}])] = 0. \tag{9.38}$$

□

Remark: You can show that this is the only Lagrangian quadratic in the field strength that you can write down.

For $\mathfrak{g} = \mathscr{L}_{\mathbb{C}}(G)$ with representation R_Λ, we can finally write down the full Yang–Mills Lagrangian coupled to matter

$$\mathcal{L} = \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) + \sum_{\Lambda\in S}(D_\mu\phi_\Lambda, D^\mu\phi_\Lambda) - W[(\phi_\Lambda, \phi_\Lambda)] \tag{9.39}$$

Bibliography

[F&S] J. Fuchs and C. Schweigert, *Symmetries, Lie Algebras, and Representations*, Cambridge University Press (1997).

[1] N. Manton, *Symmetries, Fields and Particles*, Part III Lecture Notes, 2014.