Symmetries, Fields and Particles

Cian Luke Martin

2024-11-26

Contents

Lecture 1
11/10/2024

1 Introduction

Symmetries are hidden throughout undergraduate physics. Lagrangian mechanics relies on the principle of least action, where the action S is given by

$$S = \int_{t_1}^{t_2} \mathrm{d}t \, L\left(q\left(t\right), \dot{q}\left(t\right); t\right). \tag{1}$$

Classical trajectories minimise S which gives us the Euler Lagrange equation,

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0. \tag{2}$$

Theorem 1.1 (Noether's Theorem): Invariance of L under some transformation implies an associated conserved quantity.

Example. Take a particle in a 3-dimensional potential which has Lagrangian

$$L = \frac{1}{2} m \left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - U\left(x, y, z\right). \tag{3}$$

There are a few notable symmetries here

1. L is independent of time t, i.e. under $t \mapsto t + \delta t$.

   Claim. The Hamiltonian $H = T + U$ is conserved.

   In general $H\left(x_i, p_i\right)$ is a function of $x_i = (x, y, z)$ and the conjugate momenta $p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i$ and is written in terms of the Lagrangian through Legendre transform as

$$H\left(x_i, p_i; t\right) = \sum_i \dot{x}_i \frac{\partial L}{\partial \dot{x}_i} - L. \tag{4}$$

Therefore, if L does not depend on time one has

$$\frac{\mathrm{d}H}{\mathrm{d}t} = 0 - \frac{\partial L}{\partial t} = 0, \tag{5}$$

where we have used the Euler Lagrange equations to make the first term vanish.

2. If L is invariant under $x \mapsto x + \delta x$,

$$\frac{\partial L}{\partial x} = 0 \overset{\text{EL}}{\Rightarrow} \frac{\partial L}{\partial \dot{x}} = p_x \text{ is constant.} \tag{6}$$

3. If L is invariant under rotations about the z axis then the z-component of angular momentum $L_z = xp_y - yp_x$ is constant.

   Similarly, in cylindrical coordinates $x = \rho\cos\theta$, $y = \rho\sin\theta$ and the Lagrangian becomes

$$L = \frac{1}{2}\left(m\dot{\rho}^2 + \rho^2\dot{\theta} + \dot{z}^2\right) - U(\rho, z). \tag{7}$$

   Therefore, $\frac{\partial L}{\partial \theta} = 0 \Rightarrow \frac{\partial L}{\partial \dot{\theta}} = m\rho^2\dot{\theta} = xp_y - yp_x = \text{constant}$.

1.1 Symmetry in Quantum Mechanics

Given a system whose states are elements of a Hilbert space $\mathcal{H}$. Here, symmetry implies there exists some invertible operator $U : \mathcal{H} \to \mathcal{H}$ which preserves inner products, up to an overall phase $e^{i\phi}$ (e.g. expectation values, transition amplitudes).

Definition 1.1: Let $|\Phi\rangle, |\Psi\rangle$ be any normalised vectors in $\mathcal{H}$. Denote $|U\Psi\rangle = U|\Psi\rangle$. U is a **symmetry transformation operator** if

$$|\langle U\Phi|U\Psi\rangle| = |\langle\Phi|\Psi\rangle|. \tag{8}$$

Proposition 1.1(Wigner's theorem): Symmetry transformation operators are either

a) linear and unitary, or

b) anti-linear and anti-unitary, meaning for $\alpha, \beta \in \mathbb{C}$,

$$U(\alpha|\Psi\rangle + \beta|\Phi\rangle) = \alpha^* U|\Psi\rangle + \beta^*|\Phi\rangle, \tag{9}$$

   and

$$\langle U\Phi|U\Psi\rangle = \langle\Phi|\Psi\rangle^*, \tag{10}$$

   respectively.

   Most symmetries fall into the former category, but a notable exception is time-reversal symmetry, falling into the latter.

Suppose we have a system with time independent Hamiltonian H. We can write down the time evolution of operators in the Schrödinger picture (where the states depend on time and the operators are static) as

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle. \tag{11}$$

Let's look at applying a symmetry operator U in each of the cases above.

a)

$$\langle U\Phi(t)|U\Psi(t)\rangle = \langle\Phi(t)|\Psi(t)\rangle \tag{12}$$
$$= \langle\Phi(t)|\, e^{-iHt}\, |\Psi(0)\rangle\,. \tag{13}$$

We should find the same result by transforming $|\Psi(0)\rangle$ before the evolution

$$|U\Psi(t)\rangle = e^{-iHt}\,|U\Psi(0)\rangle\,, \tag{14}$$

which implies

$$\langle U\Phi(t)|U\Psi(t)\rangle = \langle U\Phi(t)|\, e^{-iHt}\, |U\Psi(0)\rangle \tag{15}$$
$$= \langle\Phi(t)|\, U^\dagger e^{-iHt} U\, |\Psi(0)\rangle\,. \tag{16}$$

By comparing this to Eq. (13) we find that

$$U^\dagger e^{-iHt} U = e^{-iHt}. \tag{17}$$

Therefore U commutes with the Hamiltonian, $[U, H] = 0$.

Examples.

1) If H commutes with p, H cannot depend on x as $[x_i, p_j] = i\delta_{ij} \neq 0$. Therefore H is invariant under translations $x \to x + a$. One can construct a unitary operator that generates translations with $U = \exp(i\mathbf{p}\cdot\mathbf{a})$.
2) If H is rotationally symmetric the angular momentum operator commutes with H.

2 Lie Groups and algebras

2.1 Lie Groups

Definition 2.1: A **group** is a set G together with a binary operation $\circ$ such that the following properties hold

i) Closure: $g_2 \circ g_1 \in G$, $\forall g_1, g_2 \in G$,

ii) Associativity: $g_3 \circ (g_2 \circ g_1) = (g_3 \circ g_2) \circ g_1$, $\forall g_1, g_2, g_3 \in G$,

iii) Identity: $\exists e \in G$ such that $g \circ e = e \circ g = g$, $\forall g \in G$,

iv) Inverse: $\forall g \in G$, $\exists g^{-1} \in G$ such that $g \circ g^{-1} = e = g^{-1} \circ g$.

The identity e and inverse of g are unique.

Proof. Assume there exists e_1, e_2 which are both identities. Then we have that $e_1 \circ e_2 = e_1$ but also $e_1 \circ e_2 = e_2$ thus $e_1 = e_2$ and we have uniqueness.

For inverses, suppose g has two inverses h and j. One has that

$$g \circ h = e \text{ and } g \circ j = e. \tag{18}$$

Left multiplying by j and h respectively we see that

$$j \circ g \circ h = j \circ e \text{ and } h \circ g \circ j = h \circ e, \tag{19}$$

where simplifying (as we can by associativity) the left operation, we see

$$e \circ h = j \text{ and } e \circ j = h, \tag{20}$$

both of which imply $h = j$ and thus we have uniqueness. □

Definition 2.2: A group $(G, \circ)$ is **commutative** (**abelian**) if

$$g_1 \circ g_2 = g_2 \circ g_1, \tag{21}$$

$\forall g_1, g_2 \in G$. Otherwise G is **non-commutative** (**non-abelian**).

Lecture 2
15/10/2024

Definition 2.3: A **manifold** is a space which looks like Euclidean space ($\mathbb{R}^n$) locally. A **differentiable manifold** is one which satisfies certain smoothness conditions.

Definition 2.4: A **Lie group** consists of a differentiable manifold G along with a binary operation $\bullet$ such that the group axioms hold and that the operations $\left(\bullet, \cdot^{-1}\right)$ are smooth operations.

2.2 Matrix Lie Groups

The general linear group $GL\left(, \mathbb{F}\right)$ is the group of invertible $n \times n$ matrices over the field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Namely,

$$GL\left(n, \mathbb{F}\right) = \{M \in \mathrm{Mat}_n\left(\mathbb{F}\right) \mid \det M \neq 0\}. \tag{22}$$

The group operation is matrix multiplication and inverses are defined as $\det M \neq 0$.

The dimension of $GL\left(n, \mathbb{R}\right)$ is n^2, and thus we have n^2 free parameters.

For $GL\left(n, \mathbb{C}\right)$, the real dimension is $2n^2$ and the complex dimension is n^2.

There are a number of important subgroups of $GL\left(n, \mathbb{F}\right)$.

1. The *special linear group*, denoted $SL\left(n, \mathbb{F}\right) = \{M \in GL\left(n, \mathbb{F}\right) \mid \det M = 1\}$, where the constraint leaves us with a dimension of $n^2 - 1$.
2. The *orthogonal group*, denoted $O\left(n\right) = \{M \in GL\left(n, \mathbb{R}\right) \mid M^T M = I\}$. Notice that
$$M^T M = I \Rightarrow \det M = \pm 1. \tag{23}$$
3. The *special orthogonal group*, denoted $SO\left(n\right) = \{M \in O\left(n\right) \mid \det M = 1\}$
4. The *pseudo-orthogonal group*, where we define an $(n + m) \times (n + m)$ (metric) matrix by
$$\eta \equiv \begin{pmatrix} I_n & 0 \\ 0 & -I_m \end{pmatrix}. \tag{24}$$
This group is denoted
$$O\left(n, m\right) = \{M \in GL\left(n + m, \mathbb{R}\right) \mid M^T \eta M = \eta\}. \tag{25}$$
Similarly, there is a *special* subset of this group denoted $SO\left(n, m\right) \Rightarrow \det M = 1$.

5. The *unitary* matrices, which are denoted

$$U\left(n\right)=\{M\in GL\left(n,\mathbb{C}\right)\mid M^{T}M=I\}. \tag{26}$$

As before, we also have $SU\left(n\right)$ which restricts to matrices with $\det M=1$.

6. The *pseudo-unitary* group, given by

$$U\left(n,m\right)=\{M\in GL\left(n,\mathbb{C}\right)\mid M^{T}\eta M=\eta\}. \tag{27}$$

7. The *symplectic group*, for which we define a fixed, antisymmetric $2n\times 2n$ matrix, such as

$$\Omega\equiv\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}. \tag{28}$$

The symplectic group is then

$$\mathrm{Sp}\left(2n,\mathbb{R}\right)=\{M=GL\left(2n,\mathbb{R}\right)\mid M^{T}\Omega M=\Omega\}. \tag{29}$$

One can show that $M\in\mathrm{Sp}\left(2n,\mathbb{R}\right)$ satisfies $\det M=1$.

Definition 2.5: Given a $2n\times 2n$ antisymmetric matrix A, its **Pfaffian** is given by

$$\mathrm{Pf}A\equiv\frac{1}{2^n n!}\varepsilon_{i_1 i_2\cdots i_{2n}}A^{i_1 i_2}A^{i_3 i_4}\cdots A^{i_{2n-1}i_{2n}}, \tag{30}$$

where $\varepsilon_{i_1 i_2\cdots i_n}$ is the totally antisymmetric symbol $\varepsilon_{i_1 i_2\cdots i_n}=-\varepsilon_{i_2 i_1\cdots i_n}$.

2.3 Group elements as transformations

We can define actions of group elements $g\in G$ on a set X. X might be G itself, but could also be a vector space (i.e. rotation matrices acting on vectors in $\mathbb{R}^3$).

Definition 2.6: The **left action** of G on X is a map $L:G\times X\to X$ such that for $x\in X$

- $L\left(e,x\right)=x$, for e, the identity of G,
- $L\left(g_2,L\left(g_1 x\right)\right)=L\left(g_2 g_1,x\right)$, $\forall x\in X$, $\forall g_1,g_2\in G$.

The more usual notation is that $\forall g\in G$, we associate a map $g:X\to X$ such that $g\left(x\right)=gx$, however this is slightly less clear.

Definition 2.7: The **right action** of G on X is defined by $gX\to X$ such that $g\left(x\right)=xg^{-1}$, $\forall x\in X$ and $g\in G$.

The inverse preserves group composition. Namely,

$$g_2\left(g_1\left(x\right)\right)=x\underbrace{g_1^{-1}g_2^{-1}}_{\left(g_2 g_1\right)^{-1}}=\left(g_2 g_1\right)\left(x\right). \tag{31}$$

Definition 2.8: Conjugation by G on X is the action defined by

$$g(x) = gxg^{-1}, \tag{32}$$

$\forall g \in G_1$, $x \in X$.

Another definition worth making, even if it won't see immediate use is that of an *orbit*.

Definition 2.9: Given a group G and set X, an **orbit** of an element $x \in X$ is the set of elements of X which are in the image of an action of G on x.

Example. If the action is left, the orbit of $x \in X$ is written $Gx = \{gx \mid g \in G\}$.

It can be shown that the set of orbits under G 'partition' X as we will see.

2.4 Orthogonal groups

The orthogonal group, $O(n)$ in particular, represent rotations and reflections on $\mathbb{R}^n$. This preserves inner products such that

$$\langle \mathbf{v}_2, \mathbf{v}_1 \rangle = \mathbf{v}_2^T \mathbf{v}_1, \tag{33}$$

given $R \in O(n)$,

$$\langle R\mathbf{v}_2, R\mathbf{v}_1 \rangle = \mathbf{v}_2^T \underbrace{\left(R^T R\right)}_{I} \mathbf{v}_1 = \langle \mathbf{v}_2, \mathbf{v}_1 \rangle . \tag{34}$$

This is similar for $U(n)$.

Consider

$$SO(2) = \left\{ R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \middle| \theta \in (0, 2\pi) \right\}. \tag{35}$$

As cos and sin are smooth functions, this is a differentiable manifold. One can also show that $R(\theta_2) R(\theta_1) = R(\theta_1 + \theta_2)$.

Similarly, $SO(3)$ can represent rotations of vectors in $\mathbb{R}^3$ where the axis of the rotation is given by a unit vector $\mathbf{n} \in S^2$ and we rotate by an angle θ. Note that rotation by $\theta \in [-\pi, 0]$ about $\mathbf{n}$ is equivalent to a rotation by $-\theta$ about $-\mathbf{n}$ so we confine to $\theta \in [0, \pi]$.

Therefore we can depict the manifold of $SO(3)$ as a ball of radius π in $\mathbb{R}^3$, where the direction is specified by $\mathbf{n}$ and the distance from the origin is specified by $\theta \in [0, \pi]$. Antipodal points are identified such that $\pi\mathbf{n} = -\pi\mathbf{n}$.

Lecture 3
17/10/2024

3 Lie Algebras

3.1 Pseudo orthogonal group

$SO(n, m)$ act on vectors in $\mathbb{R}^{n+m}$ and preserve the scalar product

$$v_2^T \eta v_1, \tag{36}$$

for $v_1, v_2 \in \mathbb{R}^{n+m}$. For example, $SO(1,1)$ parametrise Lorentz boosts in one dimension and can be written in terms of the *rapidity* η as

$$SO(1,1) = \left\{ \begin{pmatrix} \cosh\eta & \sinh\eta \\ \sinh\eta & \cosh\eta \end{pmatrix} \middle| \eta \in \mathbb{R} \right\}. \tag{37}$$

As η is unbounded, $SO(1,1)$ is clearly noncompact.

3.2 Parametrization of Lie Groups

At least in small neighbourhoods, we can assign coordinates on an n-dimensional manifold to be

$$x := \left(x^1, \cdots, x^n\right) \in \mathbb{R}^n. \tag{38}$$

This allows us to label elements $g(x) \in G$. Closure provides

$$g(y)\, g(x) = g(z). \tag{39}$$

Smoothness gives us that the components of z are continuously differentiable functions of x and y such that for $i \in 1, \cdots, n$,

$$z^i = \phi^i(x, y). \tag{40}$$

We choose the coordinate origin such that $g(0) = e$. Identity gives us that

$$g(0)\, g(x) = g(x) \Rightarrow \phi^i(x, 0) = x^r \text{ and } \phi^i(0, y) = y^i. \tag{41}$$

Similarly, for inverses, we have that there exists some $\widetilde{x}$ such that $g(\widetilde{x}) = g(x)^{-1}$ and thus

$$\phi^i(\widetilde{x}, x) = 0 = \phi^i(x, \widetilde{x}). \tag{42}$$

Lastly, associativity gives us

$$g(z)\,(g(y)\, g(x)) = (g(z)\, g(y))\, g(x) \Rightarrow \phi^i(\phi(x, y), z) = \phi^i(x, \phi(y, z)). \tag{43}$$

This appears like a Leibniz rule/Jacobi identity as we will see.

3.3 Lie Algebras

A Lie group is homogeneous. Any neighbourhood ‘looks like’ (or in a more formal sense, can be mapped to) any other neighbourhood.

For example, for $\varepsilon \in G$ close to g_1, $g_2 g^{-1} \varepsilon$ is close to g_2.

Thus no neighbourhood in particular is special. The natural choice of the representative neighbourhood to study is the one centered at the identity of G. We will linearize near the identity of G.

Definition 3.1: A Lie Algebra is a vector space V, which additionally has a vector product, the **Lie bracket**, $[\cdot,\cdot] : V \times V \to V$ satisfying the following properties for $X, Y, Z \in V$.

1) It is antisymmetric, $[X,Y] = -[Y,X]$,
2) It satisfies the Jacobi identity, $[X,[Y,Z]] + [Y,[X,Z]] + [Z,[X,Y]] = 0$,
3) It is linear such that for $\alpha, \beta \in \mathbb{F}$, $[X, \alpha Y + \beta Z] = \alpha [X,Y] + \beta [X,Z]$.

Note. Any vector space which has a vector product $\star : V \times V \to V$ can be made into a Lie Algebra with its Lie bracket given by

$$[X,Y] = X \star Y - Y \star X. \tag{44}$$

Definition 3.2: Let's choose a basis for V, given by $\{T_a\}$ for $a = 1, \cdots, n = \dim V$. We call these basis vectors **generators** of the Lie algebra, and we write their Lie brackets as

$$[T_a, T_b] = f^c{}_{abc} T_c, \tag{45}$$

where $f^c{}_{ab} \in \mathbb{F}$ are called **structure constants**.

Antisymmetry implies $f^c{}_{ba} = -f^c{}_{ab}$ and the Jacobi identity implies

$$f^e{}_{ad} f^d{}_{bc} + f^e{}_{cd} f^d{}_{ab} + f^e{}_{bd} f^d{}_{ca} = 0. \tag{46}$$

The general element of a Lie algebra can be written as a linear combination of $\{T_a\}$ as

$$X \in V \Rightarrow X = X^a T_a \text{ with } x^a \in \mathbb{F}, \tag{47}$$

which gives us the bracket of any two elements in terms of structure constants with

$$[X,Y] = X^a Y^b f^c{}_{abc} T_c. \tag{48}$$

3.4 Lie Groups and their Lie Algebras

Take $g(\theta) \in SO(2)$ to be

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \tag{49}$$

where $e = I_2 = g(0)$. Points near the identity have $\theta \ll 1$ and thus Taylor expanding the components of $g(\theta)$ we see

$$g(\theta) = I_2 + \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \theta_2^2 I_2 + \mathcal{O}(\theta^3) \tag{50}$$

$$= e + \underbrace{\theta \frac{\mathrm{d}g}{\mathrm{d}\theta}\bigg|_{g=0}}_{\text{tangent vector}} + \frac{\mathrm{d}^2 g}{\mathrm{d}\theta^2} + \mathcal{O}(\theta^2), \tag{51}$$

where the linear term is tangent to the manifold. Here there is a one dimensional tangent space at e given by

$$T_e(SO(2)) = \left\{ \begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} \middle| a \in \mathbb{R} \right\}. \tag{52}$$

This is the Lie algebra of $SO(2)$,

$$\mathfrak{so}(2) := L(SO(2)) := T_e(SO(2)). \tag{53}$$

It remains to show this.

Proof. Notice that

$$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}\begin{pmatrix} 0 & -b \\ b & 0 \end{pmatrix} = \begin{pmatrix} -ab & 0 \\ 0 & -ab \end{pmatrix} = -abI, \tag{54}$$

and thus for any two elements (matrices) of the Lie algebra, they commute (which is trivially antisymmetric and satisfying of Jacobi). Linearity similarly follows immediately by inspection. □

Similarly, one can show $\dim(SO(n)) = \frac{1}{2}n(n-1) \equiv d$, so we have coordinates $x_1 \cdots, x_d$. Consider a single-parameter family of $SO(n)$ elements,

$$M(t) := M(\mathbf{x}(t)) \in SO(n), \tag{55}$$

such that $M(0) = I_n$. Orthogonality ($M^T M = I$) implies

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(M^T(t)\, M(t)\right) \tag{56}$$

$$= \frac{\mathrm{d}M^T}{\mathrm{d}t} + M^T \frac{\mathrm{d}M}{\mathrm{d}t}, \tag{57}$$

where looking at $t = 0$, as $M(0) = I_n$ we see

$$\frac{\mathrm{d}M^T}{\mathrm{d}t} = -\frac{\mathrm{d}M}{\mathrm{d}t}, \tag{58}$$

which implies matrices in the tangent space of $SO(n)$ are antisymmetric (and thus traceless as well).

We have

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \sum_i \frac{\partial M}{\partial x_i}\frac{\mathrm{d}x_i}{\mathrm{d}t}. \tag{59}$$

Lecture 4
19/10/2024

4 The Exponential Map

Observe that

$$T_e(\mathcal{O}(n)) = T_e(SO(n)), \tag{60}$$

as $\det I = 1$, so all curves passing through I have $\det M = 1$.

4.1 Unitary Groups

Let $M(t)$ be a curve in $SU(n)$ with $M(0) = I$. For small t, write $M(t) = I + tX + \mathcal{O}\left(t^2\right)$, where $X = \left.\frac{\mathrm{d}M}{\mathrm{d}t}\right|_{t=0}$.

Unitarity of M provides that for all t,

$$I = M^\dagger M \tag{61}$$
$$U = I + t\left(X + X^\dagger\right) + \mathcal{O}\left(t^2\right), \tag{62}$$

which implies $X^\dagger = -X$, namely, elements of the tangent space are *anti-Hermitian*.

Claim. $\operatorname{tr} X = 0$ for $X \in L(SU(n))$ or $M \in SU(n)$

Proof. Look at

$$M(t) = \begin{pmatrix} 1 + tX_{11} & tX_{12} & \cdots & tX_{1n} \\ tX_{21} & t + tX_{22} & \cdots & tX_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ tX_{n1} & tX_{n2} & \cdots & 1 + tX_{nn} \end{pmatrix}. \tag{63}$$

Notice that

$$1 = \det M = 1 + \underbrace{t \operatorname{tr} X}_{0} + \mathcal{O}\left(t^2\right), \tag{64}$$

where the underbraced term (and higher order ones) must vanish. □

For $U(n)$, X can have non-zero trace.

4.2 Lie algebra of a matrix Lie group

Consider two curves $g_1(x(t))$ and $g_2(x(t))$ through the identity e of some Lie group G. We define

$$X_1 := \left.\dot{g}_1\right|_{t=0}, \qquad X_2 := \left.\dot{g}_2\right|_{t=0}. \tag{65}$$

One can define a product

$$g_3(z(t)) = g_2(y(t))\, g_1(x(t)) \in G, \tag{66}$$

satisfying

$$\left.\dot{g}_3\right|_{t=0} = \left.\left(\dot{g}_2 g_1 + g_2 \dot{g}_1\right)\right|_{t=0} \tag{67}$$
$$= X_2 + X_1 \in T_e(G), \tag{68}$$

another vector in the tangent space.

The Lie bracket arises from the *group commutator*.

Definition 4.1: The **group commutator** of $g_1, g_2 \in G$, is

$$[g_1, g_2]_G := g_1^{-1} g_2^{-1} g_1 g_2 := h \in G. \tag{69}$$

Returning to our two curves through the identity e, $g_i(t)$ for $i \in \{1, 2\}$, we can expand

$$g_i(t) = e + tX_i + t^2 W_i + \mathcal{O}\left(t^3\right). \tag{70}$$

We have that

$$g_1(t)\, g_2(t) = e + t(X_1 + X_2) + t^2 (X_1 X_2 + W_1 + W_2) + \mathcal{O}\left(t^3\right), \tag{71}$$

and

$$g_2(t)\, g_1(t) = e + t(X_1 + X_2) + t^2 (X_2 X_1 + W_1 + W_2) + \mathcal{O}\left(t^3\right). \tag{72}$$

If we then look at

$$h(t) = [g_2(t)\, g_1(t)]^{-1} g_1(t)\, g_2(t) = e + t^2 \underbrace{(X_1 X_2 - X_2 X_1)}_{[X_1, X_2]} + \cdots, \tag{73}$$

and thus the group commutator induces the Lie bracket in the algebra. As $h(t) \in G$, the tangent to $h(t)$ at e is $[X_1, X_2] \in L(G)$, and thus we have closure under the Lie bracket.

- We write the tangent space to a matrix Lie group $G \overset{\text{subgroup}}{<} GL(n, \mathbb{F})$ at a general element p as $T_p(G)$. Let $g(t)$ be a curve in the manifold through p with $g(t_0) = p$, and thus

$$g(t + \varepsilon) = g(t_0) + \dot{\varepsilon}(t_0) + \mathcal{O}\left(\varepsilon^2\right). \tag{74}$$

  As both $g(t_0), g(t_0 + \varepsilon) \in G$, there exists $h_p(\varepsilon) \in G$ such that

$$g(t_0 + \varepsilon) = g(t_0)\, h_p(\varepsilon), \tag{75}$$

  and as $\varepsilon \to 0$, $h_p(\varepsilon) \to e$. For small ε,

$$h_p(\varepsilon) = e + \varepsilon X_p + \mathcal{O}\left(\varepsilon^2\right), \tag{76}$$

  for some $X_p \in L(G) = T_e(G)$. Neglecting $\mathcal{O}\left(\varepsilon^2\right)$,

$$\begin{aligned} e + \varepsilon X_p = h_p(\varepsilon) &= g^{-1}(t_0)\, g(t_0 + \varepsilon) && (77) \\ &= g^{-1}(t_0)\left[g(t_0) + \varepsilon \dot{g}(t_0)\right] && (78) \\ &= e + \varepsilon \underbrace{g^{-1}(t_0)\, \dot{g}(t_0)}_{X_p}. && (79) \end{aligned}$$

Claim. Conversely, for any $X \in L(G)$, there exists a unique curve $g(t)$ with $g^{-1}(t)\, \dot{g}(t) = X$ and $g(0) = g_0$.

Proof. This is a consequence of existence and uniqueness of solutions of ODEs. The solution of this ODE is

$$g(t) = g_0 \exp(tX), \tag{80}$$

where

$$\exp tX := \sum_{k=0}^{\infty} \frac{(tX)^k}{k!}. \tag{81}$$

□

4.3 One parameter subgroups

Given an $X \in L(G)$, the curve

$$g_X(t) = \exp tX, \tag{82}$$

forms an *abelian* subgroup of G, *generated* by X.

Notice that $g_X(t)$ is isomorphic to the group of real numbers under addition $(\mathbb{R}, +)$ if only $g_X(0) = e$. If there exist other $t_0 \neq 0$ such that $g_X(t_0) = 0$, then we have periodic structure and then $g_X(t)$ is isomorphic to the circle S^1. Lecture 5 22/10/2024

4.4 Lie Groups from Lie Algebras

Definition 4.2: Given a Lie algebra $L(G)$ of a Lie group G, we can define the **exponential map**:

$$\exp : L(G) \to G, \tag{83}$$

which for matrix Lie groups, is

$$X \mapsto \exp X = \sum_{k=0}^{\infty} \frac{X^k}{k!}. \tag{84}$$

Locally, the map is bijective (one to one). For the proof see Hall Section 2.7. Globally, the map is generally, not.

Example. For example, $U(1) = \{e^{i\theta} \mid \theta \in [0, 2\pi)\}$ and we have

$$L(U(1)) = \{ix \mid x \in \mathbb{R}\}, \tag{85}$$

where clearly $\exp(ix)$ is not one to one globally since $e^{2\pi ni} = 1$, $\forall n \in \mathbb{Z}$.

Example. $G = O(n)$. Let $X \in L(O(n)) \subset \mathrm{Skew}_n(\mathbb{R})$. Let $M = \exp tX$, and observe that as X is antisymmetric, $M^T = [\exp X]^T = \exp(-tX)$. Therefore,

$$MM^T = I = M^T M, \tag{86}$$

and thus we recover $M \in O(n)$.

Note. $\operatorname{tr} X = 0$. Let $\lambda_1, \cdots, \lambda_n$ be eigenvalues of X, and observe that

$$\begin{aligned} \det M &= \det(\exp tX) && (87)\\ &= \exp(\operatorname{tr} tX) && (88)\\ &= \exp(0) && (89)\\ &= 1. && (90) \end{aligned}$$

and thus $M \in SO(n)$. Thus elements of $O(n)$ with determinant -1 are not in the image of the exponential map.

Therefore, $O(n)$ is a disconnected manifold. One can think of $O(n)$ as two disconnected islands, one with $\det M = 1$ containing the identity called *proper rotations*, and another containing elements with $\det M = -1$ called *improper rotations* as they contain a reflection.

One can show that $A \in \mathrm{Skew}_n(\mathbb{R})$ implies $A \in L(SO(n))$ or $L(O(n))$.

Define $\gamma(t) := \exp tA$ to be a curve of matrices on some manifold. By above, we see that

$$(\gamma(t))^T(\gamma(t)) = I, \tag{91}$$

and thus $\det \gamma(t) = 1$ which implies $\gamma(t) \in SO(n)$. By construction, $A = \dot{\gamma}(t)\Big|_{t=0}$ and thus is tangent to the curve at the identity of $SO(n)$ suggesting $A \in L(SO(n))$. Therefore

$$\dim SO(n) = \dim L(SO(n)) = \dim(\mathrm{Skew}_n(\mathbb{R})) = \frac{n(n-1)}{2}. \tag{92}$$

4.5 Group product from Lie bracket

Recall the Baker-Campbell-Haussdorff (BCH) formula, namely that for $X, Y \in L(G)$, we have

$$\exp(tX)\exp(tY) = \exp(tZ), \tag{93}$$

where

$$Z = X + Y + \frac{t}{2}[X,Y] + \frac{t^2}{12}([X,[X,Y]] + [Y,[X,Y]]) + \mathcal{O}(t^3). \tag{94}$$

One can show this order by order in t. As $L(G)$ is closed under the Lie bracket, $Z \in L(G)$ and thus $\exp tZ \in G$.

5 Representation Theory

Groups and their elements represent transformations under which a system or object is invariant. Representations of groups tell us how the action of the group transforms vectors in a vector space.

We saw $GL(n,\mathbb{F})$ as a group of invertible matrices. These matrices are equivalently linear maps (automorphisms) on the vector space $\mathbb{F}^n$ with

$$GL(n,\mathbb{F}) : \mathbb{F}^n \to \mathbb{F}^n. \tag{95}$$

We generalize this notation to act on any vector space V such that

$$GL(V) : V \to V. \tag{96}$$

If V is finite dimensional, we can choose a basis and recover the original definition.

5.1 Lie group representations

Definition 5.1: A **representation** D of a group G is a smooth group homomorphism

$$D : G \to GL(V), \tag{97}$$

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

That is, $\forall g \in G$, $D(g) : V \to V$ is an invertible, linear map such that for a vector $v \in V$,

$$v \mapsto D(g)\, v. \tag{98}$$

This map is linear such that

$$D(g)(\alpha v_1 + \beta v_2) = \alpha D(g)\, v_1 + \beta D(g)\, v_2, \tag{99}$$

$\forall \alpha, \beta \in \mathbb{F}$, $v_1, v_2 \in V$. Further, the group homomorphism holds such that we have

$$D(g_2 g_1) = D(g_2)\, D(g_1), \tag{100}$$

$\forall g_1, g_2 \in G$. This group homomorphism property implies that

$$D(e) = \mathrm{id}_V, \tag{101}$$

and by an identical argument,

$$D(g)^{-1} = D\left(g^{-1}\right). \tag{102}$$

Definition 5.2: The **dimension** of a representation D is the dimension of the representation space V on which it acts.

If V is finite dimensional, say $\dim V = N$, then $GL(V)$ is isomorphic to $GL(N, \mathbb{F})$. Lecture 6 24/10/2024

6 The Adjoint Representation

Definition 6.1: The **kernel** of a map $D : G \to GL(V)$ consists of all elements of G which map to the identity, $\mathrm{id}_V = I$.

Definition 6.2: A representation D is said to be **faithful** if $D(g) = \mathrm{id}_V$ only for $g = e$. Namely, if $\ker D = \{e\}$.

Faithfulness implies that D is injective, i.e. $D(g_1) = D(g_2) \Rightarrow g_1 = g_2$.

Proof. Assume D is faithful and that $D(g_1) = D(g_2)$. Then,

$$\begin{aligned} D\left(g_1^{-1}\right) D(g_1) &= D\left(g_1^{-1}\right) D(g_2) && (103)\\ D\left(g_1^{-1} g_1\right) &= D\left(g_1^{-1} g_2\right) && (104)\\ D(e) &= D\left(g_1^{-1} g_2\right) && (105)\\ \mathrm{id}_V &= D\left(g_1^{-1} g_2\right), && (106) \end{aligned}$$

where as D is faithful, $g_1^{-1} g_2 = e \Rightarrow g_1 = g_2$. □

Examples. We look at $G = (\mathbb{R}, +)$.

1) For some fixed, $k \in \mathbb{R}$, $D(\alpha) = e^{k\alpha}$, $\forall \alpha \in G$ is a one-dimensional representation.

   One can check that this is a representation, namely, that it respects the group multiplication through a homomorphism
$$D(\alpha) D(\beta) = e^{k\alpha} e^{k\beta} = e^{k(\alpha+\beta)} = D(\alpha + \beta). \tag{107}$$
   For $k \neq 0$, this is a faithful representation as $D(\alpha) = 1 \Rightarrow \alpha = 0$ and thus $\ker D = \{0 \equiv \mathrm{id}_G\}$.

2) For $k = 0$, $D(\alpha) = 1 \forall \alpha$, and thus $\ker D = G$. This is not faithful and is called the *trivial representation.*

3) We can similarly define $D(\alpha) = e^{ik\alpha}$, for $k \in \mathbb{R}$. This is not faithful as $\ker D = \{\frac{2\pi n}{k} \mid n \in \mathbb{Z}\}$. Here $V = \mathbb{C}$.

4) A two dimensional representation can also be defined with
$$D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}, \tag{108}$$
   where $V = \mathbb{R}^2$.

5) Lastly, one can define an infinite dimensional representation. Let
$$V = \{\text{space of all real functions } f(x)\}, \tag{109}$$
   and let
$$D(\alpha) f(x) = f(x - \alpha). \tag{110}$$
   We see that $D(\alpha) f = f$, $\forall f \in V \Rightarrow \alpha = 0$, and thus the representation is faithful.

Definition 6.3: The **trivial representation** D_0 is where
$$D_0(g) = 1, \tag{111}$$
$\forall g \in G$. This is not faithful as $\ker D = G$ and the dimension of D_0 is 1.

Quantities which are invariant under group transformations, transform in the trivial representation. In physics, we call these **singlets**.

Note. One can form a trivial representation of any dimension M such that $D(g) = I_m$, $\forall g \in G$. This representation is *reducible* (as we will define) and can be thought of as m copies of the dimension one trivial representation.

Definition 6.4: If G is a matrix Lie group, then the **fundamental** or **defining representation** D_f is given by
$$D_f(g) = g, \tag{112}$$
$\forall g \in G$.

Only $D_f(e) = e$ thus it is faithful. If $G \subset GL(n, \mathbb{F})$, then $\dim D_f = n$.

Let G be a matrix Lie group and consider its Lie algebra as a vector space $V = L(G)$.

Definition 6.5: The **adjoint representation** $D^{\text{adj}} \equiv \text{Ad}$ is the map

$$\text{Ad} : G \to GL(L(G)), \tag{113}$$

such that $\forall g \in G$,

$$\text{Ad}_g : L(G) \to L(G), \tag{114}$$

with

$$\text{Ad}_g X = gXg^{-1}, \tag{115}$$

$\forall X \in L(G)$. This is action by conjugation.

Let's check that this is a representation.

- *Closure*: For $X \in L(G)$, there is a curve in G such that

$$g(t) = e + tX + \cdots . \tag{116}$$

For any $h \in G$, we have another curve

$$\widetilde{g}(t) = hg(t)h^{-1} \tag{117}$$

$$= e + t\underbrace{hXh^{-1}}_{\in L(G)} + \cdots . \tag{118}$$

Therefore $\text{Ad}_h X = hXh^{-1} \in L(G)$ and thus we have closure.

- *Group homomorphism:* The group operation is preserved as

$$(\text{Ad}_{g_2 g_1}) X = (g_2 g_1) X (g_2 g_1)^{-1} \tag{119}$$

$$= g_2 \left(g_1 X g_1^{-1}\right) g_2^{-1} \tag{120}$$

$$= \text{Ad}_{g_2}(\text{Ad}_{g_1} X) \tag{121}$$

$$= (\text{Ad}_{g_2})(\text{Ad}_{g_1})(X). \tag{122}$$

- *The Lie bracket:* The Lie bracket is preserved as well as

$$\text{Ad}_g([X, Y]) = g[X, Y]g^{-1} \tag{123}$$

$$= \left[gXg^{-1}, gYg^{-1}\right] \tag{124}$$

$$= [\text{Ad}_g X, \text{Ad}_g Y]. \tag{125}$$

6.1 Lie algebra representations

Definition 6.6: A **representation**, d, of a Lie algebra $L(G)$ is a map from $L(G)$ to a set of linear maps with $\mathfrak{gl}(V) = L(GL(V))$, where the Lie bracket is preserved (instead of the group operation).

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

$$v \mapsto d(X)\, v, \tag{126}$$

$\forall v \in V$.

Linearity implies that for $X, Y \in L(G)$, we have $d(\alpha X + \beta Y) = \alpha d(X) + \beta d(Y)$. As we also want to preserve the bracket, we need

$$d([X, Y]) = [d(X), d(Y)], \tag{127}$$

$\forall X, Y \in L(G)$.

Definition 6.7: The **dimension** of $d = \dim V$.

Lecture 7
26/10/2024

The Lie algebra also admits a trivial representation,

$$d_0(X) = 0 \in V, \tag{128}$$

$\forall X \in L(G)$.

The fundamental representation also follows identically and we have

$$d_f(X) = X \in V, \tag{129}$$

$\forall X \in L(G)$.

Lastly we rewrite the adjoint representation. Recall that it can be thought of as the action of the Lie algebra on itself.

Definition 6.8: The **adjoint representation** of a Lie algebra can be written

$$\mathrm{ad} : L(G) \to \mathfrak{gl}(L(G)). \tag{130}$$

Then, for $X \in L(G)$,

$$\mathrm{ad}_X : L(G) \to L(G), \tag{131}$$

such that

$$\mathrm{ad}_X Y = [X, Y], \tag{132}$$

$\forall Y \in L(G)$.

6.2 From The Lie Group Reps to the Lie Algebra Reps

As before, consider tangent curves in G

$$g(t) = e + tX + \cdots. \tag{133}$$

We expand the corresponding elements of the representation D of G as

$$D(g(t)) = \mathrm{id}_V + t d(X) + \cdots. \tag{134}$$

We use this expansion to define d from D and we can check that the Lie bracket is preserved. Namely,

$$D\left(g_1^{-1}g_2^{-1}g_1g_2\right) = D\left(g_1\right)^{-1} D\left(g_1\right) D\left(g_2\right), \tag{135}$$

where expanding the left hand side we see

$$\begin{aligned} D\left(g_1^{-1}g_2^{-1}g_1g_2\right) &= D\left(e + t^2\left[X_1, X_2\right] + \cdots\right) && (136)\\ &= \mathrm{id}_V + t^2 d\left(\left[X_1, X_2\right]\right). && (137)\end{aligned}$$

Expanding $g_i\left(t\right) = e + tX_i + \cdots$, we see that the right hand side of Eq. (135) then becomes

$$D\left(g_1\right)^{-1} D\left(g_1\right) D\left(g_2\right) = \mathrm{id}_V + t^2\left[d\left(X_1\right), d\left(X_2\right)\right], \tag{138}$$

and thus equating the two sides, we arrive at

$$d\left(\left[X_1, X_2\right]\right) = \left[d\left(X_1\right), d\left(X_2\right)\right], \tag{139}$$

is a Lie algebra homomorphism.

Example. The adjoint representation ad_X can be obtained from Ad_g. Namely, given $Y \in L\left(G\right)$,

$$\begin{aligned} \mathrm{Ad}_g Y &= gYg^{-1} && (140)\\ &= \left(I + tX\right) Y \left(I - tX\right) && (141)\\ &= Y + t\left[X, Y\right] && (142)\\ &= \left(I + t\mathrm{ad}_X\right) Y, && (143)\end{aligned}$$

and thus $\mathrm{ad}_X Y = \left[X, Y\right]$ as expected.

6.3 Useful concepts

Definition 6.9: Representations D_1 and D_2 of G (or d_1 and d_2 of $L\left(G\right)$) are **equivalent** if there exists an invertible linear maps R, such that

$$D_2\left(g\right) = RD_1\left(g\right) R^{-1}, \tag{144}$$

$\forall g \in G$ (or $X \in L\left(G\right)$).

Definition 6.10: A representation d of $L\left(G\right)$ with representation space V has an **invariant subspace** $W \subseteq V$ if $\forall w \in W$ and $X \in L\left(G\right)$,

$$d\left(X\right) w \in W. \tag{145}$$

Example. If all $d\left(X\right)$ are all upper triangular matrices, $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, then there is an invariant subspace

$$W = \left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} \right\}. \tag{146}$$

Definition 6.11: An **irreducible representation** ("*irrep*") is a representation with no nontrivial invariant subspaces.

Otherwise, the representation is **reducible**.

Definition 6.12: A **direct sum** of vector spaces U and V is written

$$U \oplus W = \{(u, w) \mid u \in U, w \in W\}\,, \tag{147}$$

where $(u_1, w_1) + (u_2, w_2) = (u_1 + u_2, w_1 + w_2)$ and $\alpha\,(u, w) = (\alpha u, \alpha w)$. Note that

$$\dim U \oplus W = \dim U + \dim W. \tag{148}$$

Definition 6.13: A **totally reducible** representation d of $L\,(G)$ (or $D\,(G)$) can be decomposed into irreducible pieces. Namely, it's representation spaces can be written as a direct sum of irreducible representation spaces,

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_k, \tag{149}$$

such that $d\,(X)\,w_i \in W_i$ for all $X \in L\,(G)$ and $w_i \in W_i$. Then, there exists some basis where $d\,(X)$ becomes block diagonal such that

$$d\,(X) = \begin{pmatrix} d_1\,(X) & 0 & \cdots & 0 \\ 0 & d_2\,(X) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_k\,(X) \end{pmatrix}. \tag{150}$$

We often write $d = \widetilde{d}_1 \oplus \cdots \oplus \widetilde{d}_k$.

Definition 6.14: An N-dimensional representation (for N finite) D is **unitary** if $D\,(g) = U\,(N)$, $\forall g \in G$.

Identically d is unitary if $d\,(X)$ if $d\,(X) \in L\,(U\,(N))$, $\forall X \in L\,(G)$.

If all $D\,(g)$ are real, then $D\,(g) \in O\,(N)$ then D is said to be orthogonal. Most of these claims rely on d being finite dimensional.

Theorem 6.1 (Maschke): A finite-dimensional unitary representation is either irreducible or totally reducible.

Proof. (Sketch) For each invariant subspace W, the orthogonal component $W_\perp$ is also invariant. This implies we can separate the representation space into

$$V = W \oplus W_\perp. \tag{151}$$

Then similarly we can decompose W and $W_\perp$ into any further invariant spaces if they exist (and repeat until there are no more invariant subspaces). If V is finite dimensional then this process must terminate. □

Note. There are a few things of note after this definition storm. Maschke's theorem can be extended to

- all finite representations of discrete groups
- all finite representations of compact Lie groups

Example. Take $V = \{$ all 2 π periodic functions $f : \mathbb{R} \to \mathbb{R}$, $f(x + 2\pi) = f(x)\}$. Take the representation to be

$$(D(\alpha) f)(x = f(x - a)). \tag{152}$$

Recall that this is not faithful. We have invariant subspaces given by

$$W_n = \{f(x) = a_n \cos nx + b_n \sin nx \mid a_n, b_n \in \mathbb{R}\}, \tag{153}$$

which are one dimensional. One can then write

$$V = W_0 \oplus W_1 \oplus W_2 \oplus \cdots = \bigoplus_{n=0}^{\infty} W_n, \tag{154}$$

which is a direct sum of invariant subspaces, each occurring once. Lecture 8 29/10/2024

W_n is invariant as

$$a_n \cos n(x - \alpha) + b_n \sin n(x - \alpha) = a'_n \cos(nx) + b'_n \sin nx, \tag{155}$$

for some $a'_n, b'_n \in \mathbb{R}$. Recall that the Fourier decomposition of any 2π periodic function can be written

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \tag{156}$$

Definition 6.15: Let V and W be vector spaces. The **tensor product space** $V \otimes W$ is spanned by elements, **product vectors**, $v \otimes w$ with $v \in V$ and $w \in W$ satisfying

- linearity, such that $v \otimes (\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 v \otimes w_1 + \lambda_2 v \otimes w_2$, and identically in the first component.
- $\dim(V \otimes W) = (\dim V)(\dim W)$

With a product state $\Phi = v \otimes w$, we write

$$\Phi_A = \Phi_{\alpha a} = v_\alpha w_a, \tag{157}$$

where $\alpha = 1, \cdots, \dim V$, $a = 1, \cdots, \dim W$ and $A = 1, \cdots, \dim V \otimes W$.

Not all elements of $V \otimes W$ are product states (as they can be linear combinations).

Definition 6.16: Let $D^{(1)}$ and $D^{(2)}$ be representations of a group G with representation spaces V and W. These satisfy

$$D^{(1)}(g) : v_\alpha \mapsto D^{(1)}(g)_{\alpha\beta} v_\beta, \ v \in V, \tag{158}$$

$$D^{(2)}(g) : w_a \mapsto D^{(2)}(g)_{ab} w_b, \ w \in W. \tag{159}$$

The **tensor product representation** $D^{(1)} \otimes D^{(2)}$ is

$$\left(D^{(1)} \otimes D^{(2)}\right)(g)(v \otimes w) = \left(D^{(1)}(g)\, v\right) \otimes \left(D^{(2)}(g)\, w\right). \tag{160}$$

Let $g_t \in G$ be a curve in the Lie group G with $g_0 = e$ and $\dot{g}_0 = X \in L(G)$. Then,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\left(D^{(1)} \otimes D^{(2)}\right)(g_t)(v \otimes w)\right] = \left[\frac{\mathrm{d}}{\mathrm{d}t} D^{(1)}(g_t)\, v\right]_{t=0} \otimes D^{(2)}(g_0)\, w + D^{(1)}(g_0)\, v \otimes \left[\frac{\mathrm{d}}{\mathrm{d}t} D(g_t)\, w\right]_{t=0}. \tag{161}$$

Let $d^{(1)}$ and $d^{(2)}$ be Lie algebra representations corresponding to $D^{(1)}$ and $D^{(2)}$. Their tensor product is given by

$$\left(d^{(1)} \otimes d^{(2)}\right)(X) = d^{(1)}(X) \otimes \mathrm{id}_W + \mathrm{id}_V \otimes d^{(2)}(X). \tag{162}$$

There is an important corollary to Maschke's theorem.

Corollary 6.1: Representations of $d^{(1)} \otimes d^{(2)}$ can be, if finite, be written as the direct sum of irreducible representations of $L(G)$, $\widetilde{d}_i$ such that

$$d^{(1)} \otimes d^{(2)} = \widetilde{d}_1 \oplus \cdots \oplus \widetilde{d}_k = \bigoplus_{i=1}^{k} \widetilde{d}_i. \tag{163}$$

The is the desired decomposition into irreducible representations.

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

$SO(3)$ describes rotations in 3 dimensions and appears when studying the quantization of angular momentum in quantum mechanics. When studying spin angular momentum, we find half integer quantum numbers which lead to $SU(2)$ representations.

The Lie algebra of $SU(2)$ is given by

$$\mathfrak{su}(2) = L(SU(2)) \tag{164}$$

$$= \{\ 2 \times 2 \text{ traceless, anti-hermitian matrices }\} \tag{165}$$

$$= \left\{X \in \mathrm{Mat}_2(\mathbb{C}) \mid X^\dagger = -X, \mathrm{tr}\, X = 0\right\}. \tag{166}$$

We can choose as a basis $t_a = -\frac{i}{2}\sigma_a$, where $a = 1, 2, 3$ and σ_a are the Pauli matrices. Recall that

$$\sigma_a \sigma_b = I\delta_{ab} + i\varepsilon_{abc}\sigma_c, \tag{167}$$

which implies

$$[T_a, T_b] = \varepsilon_{abc} T_c, \tag{168}$$

and thus the structure constants of $SU(2)$ are $f^c{}_{ab} = \varepsilon_{abc}$.

Similarly, for $SO(3)$, we see that

$$\mathfrak{so}(3) = L(SO(3)) = \text{Skew}_3. \tag{169}$$

We have a basis of the form

$$\widetilde{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \widetilde{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \widetilde{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{170}$$

namely, such that

$$\left(\widetilde{T}_a\right)_{bc} = -\varepsilon_{abc} \widetilde{T}_c, \tag{171}$$

and thus

$$\left[\widetilde{T}_a, \widetilde{T}_b\right] = \varepsilon_{abc} \widetilde{T}_c, \tag{172}$$

and thus $SO(3)$ has the same structure constants as $SU(2)$.

To show that these algebras are isomorphic, we would need an isomorphism

$$\phi : \mathfrak{g} \to \mathfrak{h}, \tag{173}$$

such that

$$\phi([X, Y]) = [\phi(X), \phi(Y)], \tag{174}$$

$\forall X, Y \in \mathfrak{g}$.

While, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$ are (as their structure constants are the same, $SU(2)$ and $SO(3)$ are in fact not isomorphic, as we will see.

When we discussed $SO(3)$ earlier, we were picturing it as a 3-ball of radius π spanned by a unit vector $\mathbf{n}$ and an angle $0 \le \theta \le \pi$ with antipodes identified.

For $SU(2)$, take $U \in SU(2)$ we can write it as

$$U = a_0 I + i \mathbf{a} \cdot \boldsymbol{\sigma}, \tag{175}$$

with $(a_0, \mathbf{a}) \in \mathbb{R}^4$ and $a_0^2 + |\mathbf{a}|^2 = 1$. Therefore $SU(2)$ as a manifold is a unit sphere in $\mathbb{R}^4$, S^3.

Definition 6.17: Let H be a subgroup of G. For any $g \in G$, we can form a **left coset** of H as

$$gH = \{gh \mid h \in H\}, \tag{176}$$

and a right coset given by

$$Hg = \{hg \mid h \in H\}. \tag{177}$$

Definition 6.18: If $H \overset{\text{subgroup}}{<} G$ is a **normal subgroup** of G, $H \triangleleft G$ if $gH = Hg$, $\forall g \in G$.

Definition 6.19: Define a set G/H to be

$$G/H = \{gH \mid g \in G\}. \tag{178}$$

We define coset multiplication by

$$(g_2 H)(g_1 H) = (g_2 g_1) H. \tag{179}$$

Theorem 6.2: For $H \triangleleft G$, G/H is a group under coset multiplication, with $H = eH$ as the identity element.

Definition 6.20: Such a group G/H is called a **quotient group** or **factor group**.

Next, we will show that

$$SO(3) \simeq SU(2)/\mathbb{Z}_2, \tag{180}$$

with $\mathbb{Z}_2 = (I_2, -I_2)$. Lecture 9 31/10/2024

Definition 6.21: The center of a group is the set of all $x \in G$ which satisfy $xg = gx$, $\forall g \in G$.

Theorem 6.3: The center $Z(G) \trianglelefteq G$ is a normal subgroup of G.

Proof. □

$SU(2)$ has centre $Z(SU(2)) = \{I_2, -I_2\} \cong \mathbb{Z}_2 = \{1, -1\}$.

We then look at cosets of the form $UZ(SU(2))$ for $U \in SU(2)$ and see

$$UZ(SU(2)) = \{U, -U\}. \tag{181}$$

The set of all such cosets forms the quotient group $SU(2)/\mathbb{Z}_2$ whose manifold is S^3 with antipodes identified, or equivalently just the upper half of S^3 ($a_0 \geq 0$) with opposite points on the equator identified.

One can see that this is just a curved picture of the $SO(3)$ manifold, as we claim

$$SO(3) \cong SU(2)/\mathbb{Z}_2. \tag{182}$$

We desire an explicit map to show this isomorphism.

One can define the map $\rho : SU(2) \to SO(3)$. For $A \in SU(2)$, $\rho(A) = R$ with components

$$R_{ij} = \frac{1}{2} \operatorname{tr}\left(\sigma_i A \sigma_j A^\dagger\right), \tag{183}$$

for $i = 1, 2, 3$. This is a 2 to 1 map as both $A, -A \mapsto \rho(A) = \rho(-A)$. This is called a **double covering** of $SO(3)$.

One also says that $SU(2)$ is the *double cover* of $SO(3)$.

Proposition 6.1: Every Lie algebra is the Lie algebra of exactly one **simply-connected** Lie group.

Definition 6.22: A manifold is **simply connected** if it is path connected and any closed loop can be smoothly contracted to a point.

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

Observe that $T_a = -i\frac{\sigma_a}{2}$ are generators of the algebra. It is convenient to enlarge this real vector space to the field $\mathbb{C}$. Given a real vector space V,

$$V := \{\lambda^a T_a \mid \lambda^a \in \mathbb{R}\} = \text{span}_{\mathbb{R}}\{T_a\}, \tag{184}$$

the *complexification* of V is

$$V_{\mathbb{C}} = \text{span}_{\mathbb{C}}\{T_a\}. \tag{185}$$

For example, we have

$$\mathfrak{su}(n) = \{X \in \text{Mat}_n(\mathbb{C}) \mid X^\dagger = -X, \text{Tr}\, X = 0\}, \tag{186}$$

becomes

$$\mathfrak{su}_{\mathbb{C}}(n) = \{X \in \text{Mat}_n(\mathbb{C}) \mid \text{tr}\, X = 0\} \cong \mathfrak{sl}(n, \mathbb{C}). \tag{187}$$

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

$$d(X + iY) = d(X) + id(Y), \tag{188}$$

where $X, Y \in L(G)$ and $X + iY \in L(G)_{\mathbb{C}}$.

Conversely, if we have a representation $d_{\mathbb{C}}$ of $L(G)_{\mathbb{C}}$ we can restrict it to the representation d of $L(G)$ by writing

$$d(X) = d_{\mathbb{C}}(X), \tag{189}$$

for $X \in L(G) \subset L(G)_{\mathbb{C}}$.

Definition 6.23: A **real form** of a complex Lie algebra $\mathfrak{h}$ is a real Lie algebra $\mathfrak{g}$ whose complexification is $\mathfrak{h}$, $\mathfrak{g}_{\mathbb{C}} = \mathfrak{h}$.

In general a complex Lie algebra can have multiple non-isomorphic real forms.

Now moving to $\mathfrak{su}(2)$, we see

$$\mathfrak{su}(2)_{\mathbb{C}} = \text{span}_{\mathbb{C}}\{\sigma_a \mid a = 1, 2, 3\}. \tag{190}$$

There exists a more convenient basis (Cartan-Weyl basis), with

$$H = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{191}$$

$$E_+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{192}$$

$$E_0 = \frac{1}{2}(\sigma_1 - i\sigma_2) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{193}$$

Observe that we have

$$[H, E_\pm] = \pm 2E_\pm, \qquad [E_+, E_-] = H. \tag{194}$$

Recall that $\mathrm{ad}_X Y = [X, Y]$ and thus

$$[H, E_\pm] = \mathrm{ad}_H E_\pm = \pm 2E_\pm. \tag{195}$$

We also have

$$[H, H] = \mathrm{ad}_H H = 0. \tag{196}$$

We see that E_-, H and E_+ are eigenvectors of ad_H with eigenvalues of $-2, 0, 2$. These eigenvalues are called the **roots** of $\mathfrak{su}(2)$.

Let d be a finite dimensional irreducible representation ("irrep") of $\mathfrak{su}(2)$ with representation space V. We write an eigenvector of $d(H) = v_\lambda$ where

$$d(H)\, v_\lambda = \lambda v_\lambda. \tag{197}$$

Definition 6.24: The eigenvalues of $d(H)$ are called the **weights** of the representation d.

Note. Roots are the weights of the adjoint representation.

The operators $d(E_\pm)$ are called **ladder** operators as

$$\begin{aligned} d(H)\,(d(E_\pm)\, v_\lambda) &= \left\{ d(E_\pm)\, d(H) + \underbrace{[d(H), d(E_\pm)]}_{d([H, E_\pm])} \right\} v_\lambda & (198) \\ &= (\lambda \pm 2)\,(d(E_\pm)\, v_\lambda), & (199) \end{aligned}$$

and thus $d(E_\pm)\, v_\lambda$ is also an eigenvector of $d(H)$ with eigenvalue $\lambda \pm 2$, *or*, $d(E_\pm)\, v_\lambda = 0$. Lecture 10 02/11/2024

If d is a finite ((n) −dimensional) representation, there must be a finite number of eigenvalues. We take d to be irreducible here. There must be some Λ such that

$$d(H)\, v_\Lambda = \Lambda v_\Lambda \text{ and } d(E_+)\, v_\Lambda = 0. \tag{200}$$

Such a Λ is called a **highest weight**.

Applying $d(E_-)$, n times, we see

$$v_{\Lambda-2n} = (d(E_-))^n v_\Lambda. \tag{201}$$

This process must terminate for some integer N as d is finite dimensional. This implies that we have a basis of eigenvectors for this representation, $\{v_\Lambda, v_{\Lambda-2}, \cdots, v_{\Lambda-2N}\}$.

We have that for $1 \leq n \leq N$,

$$d(H)\,d(E_+)\,v_{\Lambda-2n} = (\Lambda - 2n + 2)\,d(E_+)\,v_{\Lambda-2n}. \tag{202}$$

Seeking to show that this is the set of all possible eigenvectors, we check if $d(E_+)\,v_{\Lambda-2n} \propto V_{\Lambda-2n+2}$. Observe that

$$\begin{aligned}
d(E_+)\,v_{\Lambda-2n} &= d(E_+)\,d(E_-)\,v_{\Lambda-2n+2} &(203)\\
&= \left(d(E_-)\,d(E_+) + \underbrace{[d(E_+), d(E_-)]}_{d(H)}\right) v_{\Lambda-2n+2} &(204)\\
&= d(E_-)\,d(E_+)\,v_{\Lambda-2n+2} + (\Lambda - 2n + 2)\,v_{\Lambda-2n+2}. &(205)
\end{aligned}$$

This is a recursion relation. Consider $n = 1$, for which we would have $d(E_-)\,d(E_+)\,v_\Lambda = d(E_-)(0) = 0$ and thus

$$d(E_+)\,v_{\Lambda-2} = 0 + \Lambda v_\Lambda. \tag{206}$$

For $n = 2$, observe that

$$\begin{aligned}
d(E_+)\,v_{\Lambda-4} &= d(E_-)\underbrace{d(E_+)\,v_{\Lambda-2}}_{\Lambda v_\Lambda} + (\Lambda-2)\,v_{\Lambda-2} &(207)\\
&= \Lambda d(E_-)\,v_\Lambda + (\Lambda-2)\,v_{\Lambda-2} &(208)\\
&= (2\Lambda-2)\,v_{\Lambda-2}. &(209)
\end{aligned}$$

In general, we have

$$d(E_+)\,v_{\Lambda-2n} = r_n v_{\Lambda-2n+2}. \tag{210}$$

Plugging this into Eq. (205), we find

$$r_n = r_{n-1} + \Lambda - 2n + 2, \tag{211}$$

with $r_1 = \Lambda$ from Eq. (206). This has solution

$$r_n = (\Lambda + 1 - n)\,n. \tag{212}$$

As established, a finite number of eigenvalues implies that for $n = N$, $d(E_-)\,v_{\Lambda-2N} = 0$. This implies that

$$\begin{aligned}
&r_{N+1} \overset{!}{=} 0 = [(\Lambda+1) - (N+1)(N+1)] = (\Lambda - N)(N+1) = 0 &(213)\\
&\Rightarrow \Lambda = N. &(214)
\end{aligned}$$

Note. From this we can infer that the highest weights Λ have to be non-negative integers N. We will use these highest weights to classify/label the irreducible representations.

Namely, the finite-dimensional irreducible representations of $L(SU(2)) = \mathfrak{su}(2)$ are labelled by $\Lambda \in \mathbb{Z}_{\geq 0}$, d_Λ with weights

$$S_\Lambda = \{-\Lambda, -\Lambda+2, \cdots, \Lambda-2, \Lambda\}. \tag{215}$$

S_Λ is called the **weight set** of d_Λ. The weights are non-degenerate and thus $\dim d_\Lambda = \Lambda + 1$.

- d_0 is the trivial representation with $\dim d_0 = 1$
- d_1 is the fundamental/defining representation with $\dim d_1 = 2$.
- d_2 is the adjoint representation and has $\dim d_2 = 3$.

This discussion appears in quantum mechanics when discussing angular momentum. In that context, the angular momentum operators $\mathbf{J} = (J_1, J_2, J_3)$ have eigenstates

$$\mathbf{J}\cdot\mathbf{J}\,|j,m\rangle = j(j+1)\,|jm\rangle \tag{216}$$

$$J_3\,|j,m\rangle = m\,|jm\rangle\,, \tag{217}$$

with $2j \in \mathbb{Z}_{\geq 0}$, $2m \in \mathbb{Z}$ with $-j \leq m \leq j$.

Then we can translate between these domains with

$$d(H) = 2H_3 \tag{218}$$

$$\Lambda = 2j \tag{219}$$

$$d(E_\pm) = J_1 \pm iJ_2, \tag{220}$$

and the eigenvalues are $\lambda = 2m$.

6.6 Representations of $SU(2)$ and $SU(3)$

$SU(2)$ is simply connected (while $SO(3)$ is not), so a representation d_Λ of $\mathfrak{su}(2)$ gives a representation D_Λ of $SU(2)$ via the exponential map.

For $SO(3)$ recall that $SO(3) \cong SU(2)/\mathbb{Z}_2$. Namely, an element in $SO(3)$ corresponds to a pair of elements in $SU(2)$. $\{-A, A\}$, $A \in SU(2)$. The representation has to respect this.

D_Λ is a representation of $SO(3)$ iff it respects the identification of A with $-A$, namely,

$$D_\Lambda(-A) = D_\Lambda(A)\,. \tag{221}$$

It is sufficient to check whether $D_\Lambda(-I) = D_\Lambda(I)$.

For $H = \sigma_3$, we have $-I = \exp(i\pi H) \in SU(2)$ and thus

$$D_\Lambda(-I) = \exp\left(i\pi d_{\Lambda(H)}\right)\,. \tag{222}$$

As we established that $d_\Lambda(H)$ has eigenvalues $\lambda \in \{-\Lambda, -\Lambda+2, \cdots, \Lambda-2, \Lambda\}$, the eigenvalues of $D_\Lambda(-I)$ are

$$e^{i\pi\lambda} = (-1)^\lambda = (-1)^\Lambda\,, \tag{223}$$

as λ all have the same parity as Λ. Therefore

- for Λ even, we get suitable irreducible representations of both $SU(2)$ and $SO(3)$,
- for Λ odd, they are suitable only for $SU(2)$. These are *spinor representations.*

6.7 Tensor products of $\mathfrak{su}(2)$ irreducible representations

Given an arbitrary tensor product of representations, we want to decompose it into the direct sum of irreducible representations. Take irreps d_Λ and $d_{\Lambda'}$ with $\Lambda, \Lambda' \in \mathbb{Z}_{\geq 0}$ and the spaces V_Λ and V'_Λ (decomposed from $V_\Lambda \otimes V_{\Lambda'}$).

For $X \in \mathfrak{su}(2)$,

$$(d_\Lambda \otimes d_{\Lambda'})(X)(v \otimes v') = (d_\Lambda(X)v) \otimes v' + v \otimes (d_{\Lambda'}(X)v'), \tag{224}$$

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

Such a decomposition implies we can write

$$d_\Lambda \otimes d_{\Lambda'} = \bigoplus_{\Lambda'' \in \mathbb{Z}_{>0}} \mathscr{L}^{\Lambda''}_{\Lambda,\Lambda'} d_{\Lambda''}, \tag{225}$$

where $\mathscr{L}$ are called *Littlewood-Richardson coefficients* (or multiplicities).

We have bases for $V_\Lambda, V_{\Lambda'}$ given by $\{v_\lambda\}$ with $\lambda \in S_\Lambda = \{-\Lambda, \cdots, \Lambda\}$ and identically $\{v_{\lambda'}\}$ with $\lambda' \in S_{\Lambda'}$. The basis for $V_\Lambda \otimes V_{\Lambda'}$ is

$$\{v_\lambda \otimes v_{\lambda'} \mid \lambda \in S_\Lambda,\ \lambda' \in S_{\Lambda'}\}. \tag{226}$$

As a result, lets look at the action on the diagonal element H. We have

$$\begin{aligned}(d_\Lambda \otimes d_{\Lambda'})(H)(v_\lambda \otimes v_{\lambda'}) &= \lambda v_\lambda \otimes v_{\lambda'} + v_\lambda \otimes (\lambda' v_{\lambda'}) &(227)\\ &= (\lambda + \lambda') v_\lambda \otimes v_{\lambda'}, &(228)\end{aligned}$$

and thus the weights add. The weight set for the tensor product rep is then

$$S_{\Lambda,\Lambda'} = \{\lambda + \lambda' \mid \lambda \in S_\Lambda,\ \lambda' \in S_{\Lambda'}\}, \tag{229}$$

noting the multiplicities. Lecture 11 05/11/2024

The highest weight $\lambda + \lambda'$ has multiplicity one as there is only one way to form it, so $\mathscr{L}^{\Lambda+\Lambda'}_{\Lambda,\Lambda'} = 1$. Then,

$$d_\Lambda \otimes d_{\Lambda'} = d_{\Lambda+\Lambda'} \oplus \widetilde{d}_{\Lambda,\Lambda'}, \tag{230}$$

where $\widetilde{d}_{\Lambda,\Lambda'}$ is the remainder, which has a weight set such that

$$S_{\Lambda,\Lambda'} = S_{\Lambda+\Lambda'} \cup \widetilde{S}_{\Lambda,\Lambda'}. \tag{231}$$

The highest weight in $\widetilde{S}_{\Lambda,\Lambda'}$ is $\Lambda + \Lambda' - 2$ with multiplicity 1. Repeating this process above, we find

$$d_\Lambda \otimes d_{\Lambda'} = d_{\Lambda+\Lambda'} \oplus d_{\Lambda+\Lambda'-2} \oplus \cdots \oplus d_{|\Lambda-\Lambda'|}. \tag{232}$$

Example. Take $\Lambda = \Lambda' = 1$, or in physics language, $j = j' = \frac{1}{2}$. We then have weight sets

$$S_1 = \{-1, 1\}, \tag{233}$$

and thus

$$S_{1,1} = \{-2, 0, 0, 2\}. \tag{234}$$

The highest weight is 2 and thus we see

$$\begin{aligned} S_{1,1} &= \{-2, 0, 2\} \cup \{0\} && (235)\\ &= S_2 \cup S_0 && (236)\\ \Rightarrow d_1 \otimes d_2 &= d_2 \oplus d_0, && (237) \end{aligned}$$

which is also sometimes denoted

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1}, \tag{238}$$

where $\mathbf{n}$ denotes the space by its dimension. This gets more complicated when we have inequivalent representations of the same dimension.

This may be familiar if one has studied Clebsch-Gordon coefficients for the addition of angular momenta. Namely, given two irreducible representations of highest weights $\Lambda_1 = 2j_1$ and $\Lambda_2 = 2j_2$, the representation $\Lambda_3 = 2J$ with $J \leq j_1 + j_2$, has eigenvectors given by

$$|JM\rangle = \sum_{\substack{m_1, m_2 \\ m_1 + m_2 = M}} C^{J, j_1, j_2}_{m_1, m_2} |j_1 m_2\rangle \otimes |j_2 m_2\rangle, \tag{239}$$

where $C^{J,j_1,j_2}_{m_1,m_2}$ is a Clebsch-Gordon coefficient.

7 Relativistic Symmetries

7.1 Lorentz group

Lorentz transformations leave the scalar product invariant such that

$$x^\mu \eta_{\mu\nu} x^\nu = x^\sigma \Lambda^\mu{}_\sigma \eta_{\mu\nu} \Lambda^\nu{}_\rho x^\nu, \tag{240}$$

which implies

$$\eta_{\rho\sigma} = \Lambda^\mu{}_\sigma \eta_{\mu\nu} \Lambda^\nu{}_\rho, \tag{241}$$

and thus $\Lambda \in O(1,3)$. One can count degrees of freedom and thus while η contains 16 real degrees of freedom, invariance under $O(1,3)$ subtracts 10 leaving 6 degrees of freedom.

The Lorentz group consists of four disjoint sets, depending on $\det \Lambda$ and $\Lambda^0{}_0 > 0$. Observe that

$$\begin{aligned} \det \Lambda^T \eta \Lambda &= \det \eta && (242)\\ \det \Lambda^T \det \Lambda &= 1 && (243)\\ \Rightarrow \det \Lambda &= \pm 1. && (244) \end{aligned}$$

Similarly, set $\rho = \sigma = 0$ in $\eta = \Lambda^T \eta \Lambda$, then

$$\Lambda^\mu{}_0 \eta_{\mu\nu} \Lambda^\nu{}_0 = \eta_{00} = 1 \tag{245}$$

$$\Rightarrow \left(\Lambda^0{}_0\right)^2 - \sum_i \left(\Lambda^1{}_0\right)^2 = 1, \tag{246}$$

and thus $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$.

The set with $\det \Lambda = 1$ and $\Lambda^0{}_0 \geq 1$ contains the identity and forms a subgroup $SO(1,3)^+$ called the **proper** orthochronous Lorentz group. The disconnected parts of $O(1,3)$ can be obtained by composing elements of this subgroup with time reversal or parity transformations,

$$T = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}. \tag{247}$$

The elements in $SO(1,3)^+$ can be further categorized.

1. Rotations are of the form

$$\left[\left(\Lambda_R\right)^\mu{}_\nu\right] : \begin{pmatrix} 1 & \mathbf{0}_{1\times 3} \\ \mathbf{0}_{3\times 1} & R \end{pmatrix}, \tag{248}$$

   with $R \in SO(3)$.

2. Boosts are of the form

$$\left[\left(\Lambda_B\right)^\mu{}_\nu\right] = \begin{pmatrix} \cosh\psi & -\mathbf{n}^T \sinh\psi \\ -\mathbf{n}\sinh\psi & \mathbf{I} - \mathbf{n}\mathbf{n}^T \end{pmatrix} \left(\cosh\psi - 1\right), \tag{249}$$

   with $\mathbf{n}$ a unit vector and the rapidity $\psi \in \mathbb{R}$. The boost velocity is $\mathbf{v} = \mathbf{n}\tanh\psi$.

Thus we have all six degrees of freedom accounted for.

Note. $SO(3)$ is a subgroup of $SO(1,3)^+$ but boosts are not.

Exercise 1: Show that $SO(1,3)^+ \cong SL(2,\mathbb{C})/\mathbb{Z}_2$.

Proof.

□

Thus, $SL(2,\mathbb{C})$ is the double cover of $SO(1,3)^+$.

7.2 Lie algebra of the Lorentz group

Recall that $\mathfrak{sl}(2,\mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}$. The plan is to write $\mathfrak{sl}(2,\mathbb{C})$ irreducible representations as $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ irreducible representations.

We begin by expanding $\Lambda \in SO(1,3)^+$ near the identity about $\Lambda^\mu{}_\nu = \delta^\mu_\nu$ such that

$$\Lambda^\mu{}_\nu(t) = \delta^\mu_\nu + t\omega^\mu{}_\nu + \mathcal{O}\left(t^2\right). \tag{250}$$

Inserting this into $\eta = \Lambda^T \eta \Lambda$ implies

$$\eta_{\sigma\rho}\left(\delta^\sigma_\mu + t\omega^\sigma{}_\mu\right)\left(\delta^\rho_\nu + t\omega^\rho{}_\nu\right) = \eta_{\mu\nu} + \mathcal{O}\left(t^2\right) \tag{251}$$

$$\eta_{\mu\nu} + t\left(\omega_{\mu\nu} + \omega_{\nu\mu}\right) = \eta_{\mu\nu} + \mathcal{O}\left(t^2\right) \Rightarrow \omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{252}$$

Thus we can construct a basis for $\mathfrak{so}\left(1,3\right)^+$ with

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{253}$$

and

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \; J_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \; J_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{254}$$

These have Lie brackets

$$[J_i, J_j] = \varepsilon_{ijk} J_k \qquad [J_i, K_j] = \varepsilon_{ijk} K_k \qquad [K_i, K_j] = -\varepsilon_{ijk} J_k. \tag{255}$$

Note. We could use $\widetilde{K}_k = -K_k$ and the commutation relations would be the same. The same is not true about the J matrices.

Lecture 12
07/11/2024

We can equivalently write these matrices as

$$M^{0j} = K_j \qquad \varepsilon_{ijk} J_k, \tag{256}$$

where

$$\left(M^{\mu\nu}\right)^{\alpha}{}_{\beta} = \eta^{\mu\alpha}\delta^{\nu}_{\beta} - \eta^{\nu\alpha}\delta^{\mu}_{\beta}. \tag{257}$$

Then $\Lambda \in SO\left(1,3\right)^+$ can be written

$$\Lambda = \exp\left(\frac{1}{2}\omega_{\mu\nu} M^{\mu\nu}\right) = \exp\left(\theta^i J_i + \psi^i K_i\right), \tag{258}$$

where $\theta^i, \psi^i \in \mathbb{R}$ and $\omega_{\mu\nu} = -\omega_{\nu\mu}$ as before.

We can simplify these brackets by complexifying such that

$$L_i \equiv \frac{1}{2}\left(J_i + iK_i\right) \in \mathfrak{sl}\left(2,\mathbb{C}\right)_{\mathbb{C}} \tag{259}$$

$$R_i \equiv \frac{1}{2}\left(J_i - iK_i\right) \in \mathfrak{sl}\left(2,\mathbb{C}\right)_{\mathbb{C}}. \tag{260}$$

Then, we see

$$[L_i, L_j] = \varepsilon_{ijk} L_k \qquad [R_i, R_j] = \varepsilon_{ijk} R_k, \tag{261}$$

where for all i, j

$$[L_i, L_j] = 0. \tag{262}$$

Then given a generic linear combination of $\theta^i J_i + \psi^i K_i \in \mathfrak{sl}(2, \mathbb{C})$ where $\theta^i, \psi^i \in \mathbb{R}$, complexifying gives us $\alpha^i L_i + \beta^i R_i \in \mathbf{sl}(2, \mathbb{C})_{\mathbb{C}}$ where $\alpha, \beta \in \mathbb{C}$.

As they commute, we see that $\alpha^i L_i$ alone are elements of $\mathfrak{su}(2)_{\mathbb{C}}$ and identically for $\beta^i R_i$. Thus we see that

$$\begin{aligned} \mathbf{sl}(2, \mathbb{C})_{\mathbb{C}} &\cong \mathfrak{su}(2)_{\mathbb{C}} \oplus \mathfrak{su}(2)_{\mathbb{C}} && (263)\\ &\cong \mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C}). && (264) \end{aligned}$$

Given $\Lambda \in G = \{\exp\left(x^i L_i + \beta^i R_i\right) \mid x^i, \beta^i \in \mathbb{C}\}$, we have as $[L_i, R_i] = 0$

$$\begin{aligned} \Lambda &= \exp\left(\alpha^i L_i\right) \exp\left(\beta^i R_i\right) && (265)\\ &= U_L U_R =: (U_L, U_R). && (266) \end{aligned}$$

This forms a direct product group.

Note. A **direct product group** $G = A \times B$ is formed from groups A and B given by

$$G = \{(a, b) \mid a \in A, b \in B\}, \tag{267}$$

with operation $g, g' \in G$ given by

$$gg' = (aa', bb'). \tag{268}$$

A and B are normal subgroups of the direct product group.

We can see that this applies to $\Lambda_1 \Lambda_2 = (U_{1L} U_{2L}, U_{1R} U_{2R})$.

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

$$D^G((a, b)) = D^A(a) \otimes D^B(b). \tag{269}$$

Hence $A = \{\exp\left(\alpha^i L_i\right) \mid \alpha^i \in \mathbb{C}\}$ and analogously for B,

$$D^G\left(e^{\alpha^i L_i} e^{\beta^i R_i}\right) = D^A\left(e^{\alpha^i L_i}\right) \otimes D^B\left(e^{\beta^i R_i}\right). \tag{270}$$

For the Lie algebra, we then have

$$d^{L(G)}(\alpha L + \beta R) = \alpha d^{L(A)}(L) \otimes I + \beta I \otimes d^{(B)}(R), \tag{271}$$

which are representations of $\mathfrak{sl}(2, \mathbb{C})_{\mathbb{C}}$. Specifically for

$$J_i = L_i + R_i \qquad\qquad K_i == i(L_i - R_i). \tag{272}$$

For the Lorentz algebra, it is more useful to label our representations with

$$d^{(j_1, j_2)}(J_i) = d^{(j_1)}(T_i) \otimes I + I \otimes d^{(j_2)}(T_i), \tag{273}$$

where $\{T_i\}$ is a basis for $\mathfrak{su}(2)_{\mathbb{C}}$. Identically for the boosts, we see

$$d^{(j_1, j_2)}(K_i) == i\left(d^{(j_1)}(T_i) \otimes I - I \otimes d^{(j_2)}(T_i)\right). \tag{274}$$

$2j_1$ and $2j_2$ are the highest weights of some $\mathfrak{su}(2)$ irreducible representations.

Examples.

a) $(0,0)$ is the trivial representation corresponding to scalars.

b) $\left(\frac{1}{2},0\right)$ is a spinor: The *fundamental representation* of $\mathfrak{sl}(2,\mathbb{C})$. Also a **Weyl spinor** and is left-handed.

c) $\left(0,\frac{1}{2}\right)$ is also a (Weyl) spinor and is conjugate to the fundamental representation. It is right handed.

d) $\left(\frac{1}{2},\frac{1}{2}\right)$ gives us a 4-vector representation. Under $SO(3)$ rotations, this irreducible representation is reducible such that $\mathbf{2}\otimes\mathbf{2}=\mathbf{1}\oplus\mathbf{3}$. This can be thought of as $\left(x^0,x^i\right)$ where x^0 is trivial under such rotations and x^i obviously mixes. This is not the case for boosts as $\left(\frac{1}{2},\frac{1}{2}\right)$ is irreducible.

Note. A Dirac spinor is given by $\left(\frac{1}{2},0\right)\oplus\left(-,\frac{1}{2}\right)$.

7.3 Poincare group and algebra

The Poincare group is the Lorentz group (rotations and boosts) with spacetime translations as well.

This is an example of an isometry group of Minkowski space, i.e. one which preserves distances in some sense (Lorentz scalar product). We write it as $ISO(1,3)$.

This is a semi direct product group. Namely,

$$ISO(1,3)\cong O(1,3)\ltimes T^{1,3}, \tag{275}$$

where $T^{1,3}\cong\left(\mathbb{R}^{1,3},+\right)$ is the spacetime translation group. $T^{1,3}$ is a normal subgroup, but $O(1,3)$ is not.

Let G be a group with a normal subgroup $N\trianglelefteq G$ and another subgroup $H\leq G$, not necessarily normal. Let ϕ be the group homomorphism $\phi: H\to\mathrm{Aut}(N)$. Namely, for $h\in H$, $n\in N$, $\exists\phi_h: N\to N$ such that

$$\phi_h(n)=hnh^{-1}\in N, \tag{276}$$

since N is normal. The group G' consists of pairs $G'=\{(n,h)\mid n\in N,\ h\in H\}=H\ltimes N$ with group operation

$$(n_2,h_2)(n_1,h_1)=(n_2,\phi_{h_2}(n_1),h_2h_1). \tag{277}$$

We have $G'\cong G$ where G is the semi-direct product group. Lecture 13 09/11/2024

Definition 7.1: G is the **semi-direct** product group $H\ltimes N$ provided that the two equivalent statements hold:

i) N and H have trivial intersection, $N\cap H=I$. Then $G=NH$.

ii) Every element of G can be written uniquely as nh for $n\in N$ and $h\in H$.

The action of $(\Lambda^\mu{}_\nu,a^\mu)$ on $xx\in M_4$ can be written

$$x^\mu\mapsto\Lambda^\mu{}_\nu x^\nu+a^\mu, \tag{278}$$

or equivalently,

$$\begin{pmatrix} x \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} \Lambda x + a \\ 1 \end{pmatrix}, \tag{279}$$

and then composition becomes

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda' & a' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \Lambda\Lambda' & \Lambda a' + a \\ 0 & 1 \end{pmatrix}. \tag{280}$$

Recall $N \triangleleft G \Leftrightarrow gN = Ng, \forall g \in G \Leftrightarrow gng^{-1} \in N, \forall n \in N, g \in G$. Translations ratify this and thus form a normal subgroup

$$\begin{pmatrix} \Lambda a & \\ 0 & 1 \end{pmatrix} \begin{pmatrix} I & a' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix}^{-1} \in T^{1,3}. \tag{281}$$

But there exists Λ and a such that

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix}^{-1} \notin O(1,3). \tag{282}$$

Note. Observe that

$$\begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} I & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & 1 \end{pmatrix}. \tag{283}$$

Recall that $O(1,3)$ has basis elements given by rotations J_i and boosts K_i such that

$$M^{0j} = K^j, \qquad M^{ij} = \varepsilon^{ijk} J^k. \tag{284}$$

The generators for translations are P^σ which in our 5×5 notation can be written

$$\widetilde{M}^{\mu\nu} = \begin{pmatrix} M^{\mu\nu} & 0 \\ 0 & 0 \end{pmatrix} \qquad \widetilde{P}^\sigma = \begin{pmatrix} 0 & P^\sigma \\ 0 & 0 \end{pmatrix}, \tag{285}$$

with $(P^\sigma)^\beta = \eta^{\sigma\beta}$. With 6 generators from $O(1,3)$ and 4 from $T^{1,3}$ we have a 10 dimensional algebra.

The group elements can then be written using the exponential map with

$$(\Lambda, a) = \exp(a_\sigma P^\sigma) = \exp(a_\sigma P^\sigma) \exp\left(\frac{1}{2}\omega_{\mu\nu} M^{\mu\nu}\right). \tag{286}$$

We have Lie brackets

$$[M^{\mu\nu}, M^{\rho\sigma}] = \eta^{\nu\rho} M^{\mu\sigma} - M^{\nu\sigma} + \eta^{\mu\sigma} M^{\nu\rho} - \eta^{\nu\sigma} M^{\mu\rho} \tag{287}$$
$$[M^{\mu\nu}, P^\sigma] = \eta^{\nu\sigma} P^\mu - \eta^{\mu\sigma} P^\nu \tag{288}$$
$$[P^\mu, P^\nu] = 0. \tag{289}$$

Later we will treat Casimir elements (those objects which commute with all algebra generators) more properly. Here observe that $P^2 = P_\sigma P^\sigma$ commutes with $M^{\mu\nu}$ and P^σ. This is a quadratic Casimir element.

We also have the Pauli-Lubowski pseudovector given by

$$W_\mu = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma} M^{\nu\rho} P^\sigma. \tag{290}$$

$W_\mu W^\mu$ is a quartic Casimir element.

7.4 Representations of the Poincare group

There are no finite dimensional representations of the Poincare group which are unitary.

We focus here on representations useful for describing single particle states.

We denote a unitary representation of the Poincare group by by U. Then for any $(\Lambda, a) \in ISO(1,3)$, we have $U(\Lambda, a) : V \to V$ with representation space V.

As this is a semidirect product, we can write

$$U(\Lambda, a) = T(a) U(\Lambda), \tag{291}$$

where $T(a) = U(I, a)$ and $U(\Lambda) = (U\Lambda, 0)$.

Note. We can write $(\Lambda, \Lambda a) = (\Lambda, 0)(I, a) = (I, \Lambda a)(\Lambda, 0)$. Thus

$$U(\Lambda) T(a) = T(\Lambda a) U(\Lambda). \tag{292}$$

o

Translations are generated by P^σ and thus

$$(0, a) = \exp(a_\sigma P^\sigma) = e^{a \cdot P}, \tag{293}$$

where $T(a)$ is the unitary representation of this group element. Let $|p, s\rangle$ be an eigenvector of $T(a)$ in a vector space V, appropriate for a single particle state. We then have

$$T(a) |p, s\rangle = e^{i\mathbf{a}\cdot\mathbf{p}}, \tag{294}$$

where ip^σ is an eigenvalue of P^σ. p^σ correspond to particle momenta and s is an internal discrete degree of freedom.

Lorentz transforming an eigenvector, we see

$$T(a)(U(\Lambda)|p, s\rangle) = U(\Lambda) T\left(\Lambda^{-1} a\right) |p, s\rangle \tag{295}$$

$$= e^{i\left(\Lambda^{-1} a\right)\cdot p} U(\Lambda |p, s\rangle) \tag{296}$$

$$= e^{ia\cdot(\Lambda p)} U(\Lambda) |p, s\rangle. \tag{297}$$

This is still an eigenvector with eigenvalue

$$p^\mu \to p'^\mu = \Lambda^\mu{}_\nu p^\nu. \tag{298}$$

Note that $(p')^2 = p^2$ as the scalar product is preserved under Lorentz transform. This and the above, make use of $\Lambda^T = \Lambda^{-1}$.

For any fixed p^2, we have an equivalence class of momentum eigenvectors all related by Lorentz transformation.

We choose some "standard" or "reference" momentum k such that $k^2 = p^2$. Then $p^\mu = L(p)^\mu{}_\nu k^\nu$ where $L(p)$ is a Lorentz transformation that takes us to the k frame.

Then we write eigenvectors as

$$|p, s\rangle = U(L(p)) |k, s\rangle . \tag{299}$$

Lecture 14
12/11/2024

Acting with an arbitrary Lorentz transformation Λ, we have that

$$\begin{aligned} U(\Lambda)|p, s\rangle &= U(\Lambda) U(L(p)) |k, s\rangle & (300)\\ &= U(\Lambda L(p)) |k, s\rangle . & (301) \end{aligned}$$

Inserting the identity $I = U\left(L(\Lambda p) L^{-1}(\Lambda p)\right)$,

$$U(\Lambda)|p, s\rangle = U(L(\Lambda p)) U\left(L^{-1}(\Lambda p) \Lambda L(p)\right) . \tag{302}$$

Recall that $L(p) k = p$ and $L(\Lambda p) k = \Lambda p$ which has inverse $L^{-1}(\Lambda p)(\Lambda p) = k$.

Claim. We claim

$$W(\Lambda, p) = L^{-1}(\Lambda p) \Lambda L(p) , \tag{303}$$

is an element of $O(1,3)$ which leaves k invariant.

Proof. Observe that

$$L^{-1}(\Lambda p) \Lambda \underbrace{L(p) k}_{p} = L^{-1}(\Lambda p)(\Lambda p) = k, \tag{304}$$

as desired. ☐

We sometimes use the shorthand $W^\mu{}_\nu k^\nu = k^\mu$. Such elements form a subgroup of $O(1,3)$ called **little groups**.

We assume that we know the representations of the little group. We then use these to *induce* a representation on the whole Poincare group.

Say $D(W)$ is a representation such that

$$U(W) = |k, s\rangle = \sum_{s'} D_{s,s'}(W) |k, s'\rangle . \tag{305}$$

We use Eq. (302) to induce a representation on the whole group with

$$U(\Lambda)|p, s\rangle = U(L(\Lambda p)) U(W(\Lambda, p)) |k, s\rangle \tag{306}$$

$$= \sum_{s'} D_{s,s'}(W)\, U(L(\Lambda p))\, |k, s'\rangle \tag{307}$$
$$= \sum_{s'} D_{s,s'}(W)\, |\Lambda p, s'\rangle\,. \tag{308}$$

There are 6 possibilities for k^μ, 4 of which do not correspond to single particle states.

- For spacelike 4-momenta, $p^2 < 0$ e.g. $k^\mu = (0, \mathbf{k})$.
- For negative energy states $p^0 < 0$, $k^2 = 0 \Rightarrow k^\mu = (-|\mathbf{k}|, \mathbf{k})$ and $k^2 > 0 \Rightarrow k = (-k, 0)$.
- The vacuum state has $p^\mu = k^\mu = 0$.

The two interesting cases are $p^2 \geq 0$ and $p^0 > 0$.

- For massive states, $p^2 = m^2 > 0$. Let $k^\mu = (m, 0, 0, 0)$. This is a particle at rest. k^μ is invariant under 3 dimensional rotations, so the little group is $SO(3)$. The corresponding irreducible representations are those of $\mathfrak{su}(2)$. Such states are characterized by three numbers: p^μ, j and j_3.
- For massless states, $p^2 = 0$. We rotate to a frame where $k^\mu = (\omega, 0, 0, \omega)$ for $\omega > 0$.

  Claim. The little group is $ISO(2) \cong SO(2) \ltimes T^2$, namely rotations and translations in a plane.

  Proof. Start with Lorentz generators J_3, K_1 and K_2 which leave k^μ invariant. Form the linear combination

  $$E_1 := K_1 - J_2 \qquad\qquad E_2 := K_2 + J_1, \tag{309}$$

  then we have

  $$[J_3, E_1] = E_2, \qquad [E_2, J_3] = E_1, \qquad [E_1, E_2] = 0. \tag{310}$$

  □

8 Cartan's Classification of Lie Algebras

8.1 Definitions

Definition 8.1: A **subalgebra** $\mathfrak{h}$ of an algebra $\mathfrak{g}$ is a vector subspace which is also an algebra itself with the same composition rule. For a Lie subalgebra, this is the Lie bracket.

Definition 8.2: A subalgebra $\mathfrak{h}$ is called an **ideal** or *invariant* subalgebra of a Lie algebra $\mathfrak{g}$ if

$$[X, Y] \in \mathfrak{h}, \tag{311}$$

for $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$.

Every algebra has two trivial ideals, $\{0\}$ and $\mathfrak{g}$. The phrase non-trivial ideals excludes these.

Definition 8.3: The **derived algebra** $\mathfrak{i}$ of a Lie algebra $\mathfrak{g}$ is

$$\mathfrak{i} = [\mathfrak{g}, \mathfrak{g}] = \{[X, Y] \mid X, Y \in \mathfrak{g}\} . \tag{312}$$

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

Definition 8.4: The **center** of $\mathfrak{g}$ is

$$J = \{X \in \mathfrak{g} | \, [X, Y] = 0, \; \forall Y \in \mathfrak{g}\} . \tag{313}$$

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

Definition 8.5: A Lie algebra $\mathfrak{g}$ is **abelian** or *commutative* if $[X, Y] = 0$, $\forall X, Y \in \mathfrak{g}$.

This is equivalent to $J = \mathfrak{g}$.

Definition 8.6: A Lie algebra is **simple** if it is *non-abelian* and has no nontrivial ideals.

Definition 8.7: A Lie algebra is **semi-simple** if it is non-abelian and has no nontrivial *abelian* ideals.

Proposition 8.1: If $\mathfrak{g}$ is semi-simple, then it can be written as a direct sum of simple Lie algebras, $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n$.

Lecture 15
14/11/2024

8.2 The Killing form

Definition 8.8: A **bilinear form** $B : V \times V \to \mathbb{F}$ is linear in both arguments, even for $\mathbb{F} = \mathbb{C}$ such that

$$B(u, \alpha v + \beta w) = \alpha B(u, v) + \beta B(u, w) , \tag{314}$$
$$B(\alpha u + \beta v, w) = \alpha B(u, w) + \beta B(v, w) . \tag{315}$$

A *symmetric* bilinear form satisfies $B(u, v) = B(v, u)$.

Definition 8.9: A bilinear form is **nondegenerate** if $\forall v \in V$, $(v \neq 0)$, $\exists w \in V$ such that

$$B(v, w) \neq 0. \tag{316}$$

Definition 8.10: The **Killing form** of a Lie algebra $\mathfrak{g}$ is the symmetric bilinear form $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$ such that

$$\kappa(X, Y) := \frac{1}{\mathcal{N}} \operatorname{Tr}(\operatorname{ad}_X \circ \operatorname{ad}_Y) , \tag{317}$$

for $X, Y \in \mathfrak{g}$. $\mathcal{N}$ is a normalization that we take to be $\mathcal{N} = 1$ here.

Notice a few properties of interest.

- ad_X is linear and thus κ is bilinear.
- As the trace is cyclic, κ is symmetric.

- For us, usually $\mathbb{F} = \mathbb{R}$, $\kappa(X, Y) \in \mathbb{R}$.

Let $\{T_a\}$ be a basis for $\mathfrak{g}$ and recall that

$$\mathrm{ad}_{T_a} T_b = [T_a, T_b] = f^c{}_{ab} T_c. \tag{318}$$

Thus we have matrix element

$$(\mathrm{ad}_{T_a})^c{}_b = f^c{}_{ab}. \tag{319}$$

For example, for $\mathfrak{su}(2)$, $[T_a, T_b] = \varepsilon_{abc} T_c$ gives

$$(\mathrm{ad}_{T_1}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \tag{320}$$

and so on. For a generic element, we have

$$X = X^a T_a \Rightarrow \mathrm{ad}_X = X^a \mathrm{ad}_{T_a}. \tag{321}$$

The Killing form can then be written

$$\kappa(T_a, T_b) = \mathrm{tr}\left[(\mathrm{ad}_{T_a})^e{}_d (\mathrm{ad}_{T_b})^d{}_c\right] \tag{322}$$

$$= f^c{}_{ad} f^d{}_{bc} \tag{323}$$

$$= \kappa_{ab}. \tag{324}$$

For general $X, Y \in \mathfrak{g}$, we have

$$\kappa(X, Y) = X^a Y^b \kappa(T_a, T_b) \tag{325}$$

$$= X^a Y^b \kappa_{ab}. \tag{326}$$

8.3 Invariance of the Killing form

Let $\mathfrak{g} = L(G)$.

Claim. For any $g \in G$,

$$\kappa(\mathrm{Ad}_g X, \mathrm{Ad}_g Y) = \kappa(X, Y). \tag{327}$$

Proof. Recall that $\mathrm{Ad}_g X = gXg^{-1}$. The key step is to show that

$$\mathrm{Ad}_{gXg^{-1}} = \mathrm{Ad}_g \circ \mathrm{ad}_X \circ \mathrm{Ad}_{g^{-1}}. \tag{328}$$

□

Let $g = e + tZ + \mathcal{O}(t^2)$. We then have

$$\kappa(\mathrm{Ad}_g X, \mathrm{Ad}_g Y) = \kappa(X + t\mathrm{ad}_Z X, Y + t\mathrm{ad}_Z Y) \tag{329}$$

$$= \kappa(X,Y) + t(\kappa(\mathrm{ad}_Z X, Y) + \kappa(X, \mathrm{ad}_Z Y)). \tag{330}$$

Invariance implies the t term vanishes giving

$$\kappa(\mathrm{ad}_Z X, Y) = -\kappa(X, \mathrm{ad}_Z Y) \tag{331}$$
$$\kappa([Z,X], Y) = -\kappa(X, [Z,Y]) \tag{332}$$
$$\kappa([X,Z], Y) = \kappa(X, [Z,Y]), \tag{333}$$

Theorem 8.1 (Cartan): The Killing form of a Lie algebra $\mathfrak{g}$ is nondegenerate if and only if $\mathfrak{g}$ is semisimple:

Proof. We will one prove one direction (the forwards one). We proceed by contradiction.

Suppose that $\mathfrak{g}$ is not semisimple, then there exists a nontrivial abelian ideal $\mathfrak{a} \in \mathfrak{g}$. That is,

$$[X, A] \in \mathfrak{a}, \tag{334}$$

$\forall A \in \mathfrak{a}$ and $X \in \mathfrak{g}$.

Take a basis for $\mathfrak{a}$ to be $\{T_i \mid i = 1, \cdots, \dim \mathfrak{a}\}$ and extend to a basis for the rest of $\mathfrak{g}$ such that $\{T_a \mid \alpha = 1, \cdots, \dim \mathfrak{g} - \dim \mathfrak{a}\}$. Thus a basis for $\mathfrak{g}$ is

$$\{T_B\} = \{T_i\} \cup \{T_\alpha\}. \tag{335}$$

As $\mathfrak{a}$ is abelian, $[T_i, T_j] = 0$ and thus $f^B{}_{ij} = 0$.

Further, as $\mathfrak{a}$ is an ideal, $[T_i, T_\alpha] \in \mathfrak{a}$ and thus $f^\beta{}_{i\alpha} = 0$.

Together, these give us that $f^\beta{}_{iB} = f^\beta{}_{Bi} = 0$.

Consider $\kappa(X, A) = \kappa_{Bi} X^B A^i$. As $\kappa_{Bi} = f^C{}_{BD} f^D{}_{iC}$, As $\{C\} = \{j\} \cup \{\alpha\}$,

$$\kappa_{Bi} = f^\alpha{}_{BD} f^D{}_{i\alpha} + f^j{}_{BD} \underbrace{f^D{}_{ij}}_{0}. \tag{336}$$

Then letting $\{D\} = \{j\} \cup \{\beta\}$, this becomes

$$\kappa_{Bi} = f^\alpha{}_{B\beta} \underbrace{f^\beta{}_{i\alpha}}_{0} + \underbrace{f^\alpha{}_{Bj}}_{0} f^j{}_{i\alpha} \tag{337}$$
$$= 0, \tag{338}$$

and thus κ is degenerate.

□

We divert for a moment with some definitions and facts.

Proposition 8.2: If κ is nondegenerate, then it is invertible. We can find (κ^{-1}) such that

$$\kappa_{ab} \left(\kappa^{-1}\right)^{bc} = \delta_a^c. \tag{339}$$

Proof. Linear algebra, following from nondegeneracy of the Killing form. □

Definition 8.11: A Lie group is **semisimple** if its Lie algebra is semisimple.

Note. Some authors only use this term for connected Lie groups.

Proposition 8.3: If the Killing form of a real Lie algebra $\mathfrak{g}$ is negative definite, such that $\kappa(X, X) < 0$, $\forall X \in g\mathfrak{g}$, then G is compact and $\mathfrak{g}$ is said to be of **compact type**.

Proposition 8.4: A compact group which is not semisimple has an algebra with *negative, semidefinite* ($\kappa(X, X) \leq 0$) Killing form (but not negative definite).

Note. There are also noncompact groups which have negative semi definite Killing forms.

Proposition 8.5: Every semisimple, complex Lie algebra $L(G)_{\mathbb{C}}$ has a real form with

$$\kappa_{ab} = -\kappa_{ab}, \tag{340}$$

$\kappa \in \mathbb{R}^{+}$.

By above, G is compact and its real form is called a **compact real form**.

Definition 8.12: Any basis for which $\kappa_{ab} \propto \delta_{ab}$, if it exists, is called an **adapted basis**.

In an adapted basis, $\{T_a\}$,

$$\begin{aligned} \kappa([T_c, T_a], T_b) &= f^{d}{}_{ca}\, \kappa(T_d, T_b) && (341)\\ &= -\kappa f^{b}{}_{ca} && (342)\\ &= \kappa(T_c, [T_a, T_b]) = f^{d}{}_{ab}\, \kappa(T_c, T_C) && (343)\\ &= -\kappa f^{c}{}_{ab}. && (344) \end{aligned}$$

Thus invariance implies $f^{b}{}_{ca} = f^{c}{}_{ab} = -f^{c}{}_{ba}$, namely that f in an adapted basis is now totally antisymmetric.

8.4 Casimir elements

Definition 8.13: A **Casimir element** is a polynomial function of elements of the Lie algebra which commutes with all elements of the Lie algebra.

Namely, the Casimir is an element of the *universal enveloping algebra* of $\mathfrak{g}$. The UEA is the span of $\{I, \mathfrak{g}, \mathfrak{g} \otimes \mathfrak{g}, \cdots\}$ subject to the rule $X \otimes Y - Y \otimes X = [X, Y]$.

Note. $XY \equiv X \otimes Y$ is generally not in $\mathfrak{g}$ is generically not in $\mathfrak{g}$ but in the universal enveloping algebra.

This gives us new identities such as

$$[X, YZ] = [X, Y]\, Z + Y\, [X, Z]. \tag{345}$$

Lecture 16
16/11/2024

In what follows, let $\mathfrak{g}$ be the real Lie algebra of a simple, compact group and use an adapted basis such that the Killing form is $\kappa = -\delta_{ab}$. This can come about from normalizing the generators $\{T_a\}$.

The universal quadratic Casimir element in the universal enveloping algebra of $\mathfrak{g}$

$$C := T_b T_b, \tag{346}$$

where there is implicit summation over b and we do not raise the index as there is no notion of covariance here.

One can check this with

$$[T_a, C] = [T_a, T_b T_b] = T_b [T_a, T_b] + [T_a, T_b] T_b = f_{abc} T_b T_c + f_{abc} T_c T_b = 0. \tag{347}$$

Therefore, $[X, C] = 0$, $\forall X \in \mathfrak{g}$.

Some algebras have higher order (polynomial) Casimirs.

Consider the quadratic Casimir in a representation, d, of $\mathfrak{g} = L(G)$,

$$C_d = d(T_a)\, d(T_a). \tag{348}$$

As above,

$$[d(X), C_d] = 0, \tag{349}$$

$\forall X \in \mathfrak{g}$.

Let $D(g) = \exp d(X)$ be a representation of G and $g \in G$.

If d and therefore D are irreducible, by Schur's lemma,

$$C_d D(g) = D(g)\, C_d, \tag{350}$$

$\forall g \in G$, implies

$$C_d = c_d I, \tag{351}$$

where $c_d \in \mathbb{R}$ as our Lie algebra is real.

Example. Take $SU(2)$ in adapted basis $T_a = -\frac{i}{2\sqrt{2}}\sigma_a$ such that $\kappa_{ab} = -\delta_{ab}$. Let $j = \frac{1}{2}$ (eq. $\Lambda = 1$). We then see

$$C_{\frac{1}{2}} = -\frac{1}{8}\sigma_a\sigma_a = \frac{3}{8}I, \tag{352}$$

and thus $c_{\frac{1}{2}} = -\frac{3}{8}$. This is the quadratic Casimir of the fundamental representation of $\mathfrak{su}(2)$.

More generally, in $SU(2)$, with adapted basis $\{\frac{i}{\sqrt{2}}J_a\}$

$$C_j = \frac{1}{2}|\mathbf{J}|^2 = -\frac{1}{2}j(j+1)\, I, \tag{353}$$

with eigenvalues of $|\mathbf{J}|^2$ on $|jm\rangle$.

8.5 Cartan-Weyl basis

Let $\mathfrak{g}$ be a Lie algebra.

Definition 8.14: Element $X \in \mathfrak{g}$ is **ad-diagonalizable** if the map ad_X is diagonalizable.

Definition 8.15: A **maximal** (abelian) subalgebra $\mathfrak{h}$ is not contained in any larger, non-trivial (abelian) subalgebra.

Definition 8.16: If $\mathfrak{g}$ is complex, semisimple Lie algebra, then a **Cartan subalgebra** $\mathfrak{h}$ of $\mathfrak{g}$ is a complex subspace such that

- $\forall H_1, H_2 \in \mathfrak{h}$,
$$[H_1, H_2] = 0 \text{ (abelian) }, \tag{354}$$
- $\forall X \in \mathfrak{g}$, if $[H, X] = 0$, $\forall H \in \mathfrak{h}$, then $X \in \mathfrak{h}$ (maximal),
- and $\forall H \in \mathfrak{h}$, ad_H is diagonalizable.

Namely, it is the **maximal abelian ad-diagonalizable subalgebra**.

Proposition 8.6: Every complex semisimple Lie algebra has a Cartan subalgebra.

Example. Take $\mathfrak{su}(2)_{\mathbb{C}}$. The H of the Cartan-Weyl basis $\{H, E_+, E_-\}$ gives a 1-dimensional Cartan subalgebra $\mathfrak{h} = \mathrm{span}\mathbb{C}\{H\}$. $H = \sigma_3$ is not unique. We could have chosen σ_1 or σ_2 and found the same result in a different basis. The dimension of the Cartan subalgebra would be the same regardless.

Definition 8.17: The **rank** of a Lie algebra is the dimension of its Cartan subalgebra.

Example. $\mathrm{rank}\left(\mathfrak{su}(2)\right) = 1$.

Example. Consider $L\left(SU(n)\right)_{\mathbb{C}} = \{X \in \mathrm{Mat}_n(\mathbb{C}) \mid \mathrm{tr}\, X = 0\}$.

For the Cartan subalgebra, one can choose its basis to be the diagonal (traceless) elements

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 0 & \\ & & & \ddots \end{pmatrix}, \cdots, \begin{pmatrix} \ddots & & & \\ & 0 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}. \tag{355}$$

$$\left(H_i\right)_{\alpha\beta} = \delta_{\alpha i}\delta_{\beta i} - \delta_{\alpha,i+1}\delta_{\beta,i+1}. \tag{356}$$

Clearly, $\mathrm{rank}\left(\mathfrak{su}(n)_{\mathbb{C}}\right) = n - 1$. Lecture 17 19/11/2024

Notice that elements of the Cartan subalgebra are eigenvectors of the $\{\mathrm{ad}_H\}$ as $[H, H'] = 0$ $\forall H, H' \in \mathfrak{h}$. The adjoint map is a representation and thus observe

$$[\mathrm{ad}_H, \mathrm{ad}_{H'}] = 0. \tag{357}$$

All the ad maps commute which implies they are simultaneously diagonalizable. So all the $H \in \mathfrak{h}$ are simultaneously ad-diagonalizable. By the spectral theorem, $\mathfrak{g}$ is spanned by the simultaneous eigenvectors of the $\{\mathrm{ad}_H\}$.

Every element of $\mathfrak{h}$ is a zero eigenvector

$$\mathrm{ad}_H H' = [H, H'] = 0. \tag{358}$$

Choose a basis for $\mathfrak{h}$ to be $\{H_i | i = 1, \cdots, r = \mathrm{rank}\,(\mathfrak{g})\}$, where $\mathrm{rank}\,(\mathfrak{g}) = \dim \mathfrak{h}$.

As the Cartan subalgebra is maximal, no other eigenvectors of ad_H can have eigenvalue zero for all $H \in \mathfrak{h}$. Label the rest of the vectors by their eigenvalues: $\{\alpha_i \mid i = 1, \cdots, r\}$.

For each H, in this basis,

$$\mathrm{ad}_{H_i} E_\alpha = \alpha_i E_\alpha. \tag{359}$$

We call α_i a component of a **root**. The collection of α_i across all H_i in the basis set for a fixed E_α is called the **root**.

Definition 8.18: The set of all nonzero roots, α, of a Lie algebra, is called its root set.

Proposition 8.7: The nonzero simultaneous eigenvectors of $\mathfrak{h}$ are nondegenerate.

Proof. Omitted. See Knapp Prop 2.2.1 □

A general element of the CSA $H \in \mathfrak{h}$ is a linear combination of the H_i, $H = \rho^i H_i$ for $\rho^i \in \mathbb{C}$. Then

$$\mathrm{ad}_H E_\alpha = [H, E_\alpha] = \rho^i [H_i, E_\alpha] = \rho^i \alpha_i E_\alpha, \tag{360}$$

where one can define $\alpha(H) \equiv \rho^i \alpha_i$. This tells us that α_i are dual to H.

Definition 8.19: Given a vector space V over a field $\mathbb{F}$, the **dual vector space** V^* is the vector space of linear functions $f : V \to$
$mathbbF$. In finite dimensions, $\dim V^* = \dim V$.

Given a basis for V, $\{v_i\}$ we have a basis of V^*, $\{v_i^*\}$, such that

$$v_i^* (v_j) = \delta_{ij}. \tag{361}$$

Claim. The roots α are vectors in the space dual to $\mathfrak{h}$, denoted $\mathfrak{h}^*$.

One can check

$$\alpha(H + H') E_\alpha = [H + H', E_\alpha] = [H, E_\alpha] + [H', E_\alpha] = (\alpha(H) + \alpha(H')) E_\alpha. \tag{362}$$

Definition 8.20: The **Cartan-Weyl basis** for $\mathfrak{g}$ is given by

$$\{H_i \mid i = 1, \cdots, r\} \cup \{E_\alpha \mid \alpha \in \Phi\}. \tag{363}$$

Claim. It is a basis.

Proof. $\dim \mathfrak{h} = r$ and therefore the size of the root set must $|\Phi| = \dim \mathfrak{g} - \dim \mathfrak{h}$. □

As $|\phi| > r$ and $\dim \mathfrak{h}^* = \dim \mathfrak{h} = r$, not all α are linearly independent in $\mathfrak{h}^*$. Note that this does not make any statement about E_α's.

Our next task is to use the Killing form $\kappa(X, Y) = \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y)$ and its nondegeneracy for semisimple Lie algebras to define an inner product on the roots in $\mathfrak{h}^*$.

We now present four lemmas.

Lemma 8.1: $\kappa(H, E_\alpha) = 0$, $\forall H \in \mathfrak{h}$ and $\forall \alpha \in \Phi$.

Proof. Given some $\alpha \in \Phi$, $\exists H' \in \mathfrak{h}$ such that $\alpha(H') \neq 0$. Then by bilinearity,

$$\alpha(H')\,\kappa(H, E_\alpha) = \kappa(H\alpha(H')\,E_\alpha) \tag{364}$$

$$= \kappa(H, [H', E_\alpha]) \tag{365}$$

$$= \kappa\left(\underbrace{[H, H']}_{0}, E_\alpha\right) = 0. \tag{366}$$

□

Lemma 8.2: $\kappa(E_\alpha, E_\beta) = 0$, $\forall \alpha, \beta$ given $\alpha + \beta \neq 0$.

Proof. $\forall H \in \mathfrak{h}$,

$$(\alpha(H) + \beta(H))\,\kappa(E_\alpha, E_\beta) = \kappa([H, E_\alpha], E_\beta) + \kappa(E_\alpha, [H, E_\beta]) \tag{367}$$

$$= -\kappa([E_\alpha, H], E_\beta) + \kappa(E_\alpha, [H, E_\beta]) = 0. \tag{368}$$

□

Lemma 8.3: If $\alpha \in \Phi$, then $-\alpha \in \Phi$ and $\kappa(E_\alpha, E_{-\alpha}) \neq 0$.

Proof. Using the two previous lemmas, $\kappa(E_\alpha, H) = 0$ and $\kappa(E_\alpha, E_\beta) = 0$ for $\beta \neq -\alpha$. As $\mathfrak{g}$ is semisimple, κ is nondegenerate, and thus there must exist some $X \in \mathfrak{g}$ such that $\kappa(E_\alpha, X) \neq 0$. As we have excluded other elements this must be $X = E_{-\alpha}$. □

Lemma 8.4: $\forall H \in \mathfrak{h}$, there exists some $H' \in \mathfrak{h}$ such that $\kappa(H, H') \neq 0$.

Proof. Suppose there exists $H \in \mathfrak{h}$ such that $\kappa(H, H') = 0$, $\forall H' \in \mathfrak{h}$. Then κ would be degenerate as $\kappa(H, E_\alpha) = 0$, $\forall \alpha \in \Phi$. □

As a consequence of this last lemma, κ can be inverted within $\mathfrak{h}$. Choose a basis $\{H_i\}$. As any $H, H' \in \mathfrak{h}$ can be written in terms of $H = \rho^i H_i$, $H' = \rho'^j H_j$, write

$$\kappa(H, H') = \kappa\left(\rho^i H_i, \rho'^j J_j\right) = \kappa_{ij}\rho^i\rho'^j. \tag{369}$$

As this is nondegenerate, there exists some matrix κ^{-1} such that

$$\left(\kappa^{-1}\right)^{ik}\kappa_{kj} = \delta^i_j. \tag{370}$$

Given any $\alpha, \beta \in \Phi$, define an inner product

$$(\alpha, \beta) = \left(\kappa^{-1}\right)^{ij}\alpha_i\beta_j. \tag{371}$$

Later, we will show that $(\alpha, \beta) \in \mathbb{R}$ and $(\alpha, \alpha) > 0$. This will allow us to build a geometry of roots.

Note. If $\left(\kappa^{-1}\right)^{ij}$ is diagonal (amounting to a choice of basis), sometimes one writes $(\alpha, \beta) = \alpha \cdot \beta$. Lecture 18 22/11/2024

Recall that we have $[H_i, H_j] = 0$ and $[H_i, E_\alpha] = \alpha_i E_\alpha$, however we still need $[E_\alpha, E_\beta]$. Recall that for $\mathfrak{su}(2)_\mathbb{C}$, $[E_+, E_-] = H$, for any $H \in \mathfrak{h}$, α, $\beta \in \Phi$, we have

$$\begin{aligned}
\mathrm{ad}_H [E_\alpha, E_\beta] &= [H, [E_\alpha, E_\beta]] &(372)\\
&= -[E_\beta, [H, E_\alpha]] - [E_\alpha, [E_\beta, H]] &(373)\\
&= -\alpha(H)[E_\beta, E_\alpha] + \beta(H)[E_\alpha, E_\beta] &(374)\\
&= (\alpha(H) + \beta(H))[E_\alpha, E_\beta]. &(375)
\end{aligned}$$

Thus if $\alpha - \beta \neq 0$, then either $[E_\alpha, E_\beta] = 0$ or $[E_\alpha, E_\beta] = N_{\alpha,\beta}E_{\alpha+\beta}$ for constant $N_{\alpha,\beta}$ as this is an eigenvector of ad_H with eigenvalue $\alpha + \beta$ and thus $\alpha + \beta \in \Phi$.

If $\alpha + \beta = 0$, then $\mathrm{ad}_H [E_\alpha, E_{-\alpha}] = 0$ which implies $[E_\alpha, E_{-\alpha}] \in \mathfrak{h}$.

Proof. Using the Killing form,

$$\kappa([E_\alpha, E_{-\alpha}], H) = \kappa(E_\alpha, [E_{-\alpha}, H]) = \alpha(H)\kappa(E_\alpha, E_{-\alpha}), \tag{376}$$

where $\kappa(E_\alpha, E_{-\alpha})$ is nonzero generically. □

Define a normalized element of $\mathfrak{h}$ to be

$$H_\alpha = \frac{[E_\alpha, E_{-\alpha}]}{\kappa(E_\alpha, E_{-\alpha})}, \tag{377}$$

such that we have $\kappa(H_\alpha, H) = \alpha(H)$.

In components, $H_\alpha = \rho^i_\alpha H_i$ and $H = \rho^i H_i$. Then, $\kappa_{ij}\rho^i_\alpha\rho^j = \alpha_i\rho^i$, $\forall H \in \mathfrak{h}$. As this holds for all ρ^i, we have

$$\rho^i_\alpha = \left(\kappa^{-1}\right)^{ij}\alpha_j \qquad\qquad H_\alpha = \left(\kappa^{-1}\right)^{ij}\alpha_i H_j. \tag{378}$$

Note.

$$[H_\alpha, E_\beta] = \left(\kappa^{-1}\right)^{ij} \alpha_j [H_i, E_\beta] = \left(\kappa^{-1}\right)^{ij} \alpha_j \beta_i E_\beta = (\alpha, \beta) E_\beta. \tag{379}$$

We can choose a new normalization

$$e_\alpha = \frac{2}{\sqrt{(\alpha, \alpha) \kappa (E_\alpha, E_{-\alpha})}} E_\alpha \qquad h_\alpha = \frac{2}{(\alpha, \alpha)} H_\alpha. \tag{380}$$

Then,

$$[h_\alpha, h_\beta] = 0 \tag{381}$$

$$[h_\alpha, E_\beta] = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} H_\alpha \tag{382}$$

$$[e_\alpha, e_\beta] = \begin{cases} n_{\alpha,\beta} e_{\alpha+\beta}, & \alpha + \beta \in \Phi, \\ h_\alpha, & \alpha + \beta = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{383}$$

Note. For each $\alpha \in \Phi$, there is an $\mathfrak{sl}(2, \mathbb{C}) \cong \mathfrak{su}(2)_{\mathbb{C}}$ subalgebra with basis $\{h_\alpha, e_\alpha, e_{-\alpha}\}$ and closed commutation relations $[h_\alpha, e_{\pm\alpha}] = \pm 2 e_{\pm\alpha}$ and $[e_\alpha, e_{-\alpha}] = h_\alpha$. We denote these subalgebras as $\mathfrak{sl}(2)_\alpha$ or $\mathfrak{su}(2)_\alpha$

8.6 Geometry of Roots

With commutation relations established, we would like to construct a geometry of the root space, namely defining lengths and angles within it. This requires some machinery which we will develop in the following order.

i) For $\alpha, \beta \in \Phi$, $(\alpha, \beta) \in \mathbb{R}$.

ii) $\mathfrak{h}^*$ is spanned by the root set Φ.

iii) There is a real vector space $\mathfrak{h}^*_{\mathbb{R}}$ spanned by Φ and $\mathfrak{h}^*_{\mathbb{R}}$ contains all $\alpha \in \Phi$.

iv) $(\alpha, \alpha) \geq 0$, so a length can be defined as $|\alpha| = \sqrt{(\alpha, \alpha)}$.

v) We can then determine angles between roots.

We begin by defining a useful notion.

Definition 8.21: Let $\alpha, \beta \in \Phi$. The **α-root string** passing through β is defined to be the set of roots $S_{\alpha,\beta} = \{\beta + \rho\alpha \mid \rho \in \mathbb{Z}\}$.

Claim. The allowed values for ρ are $\rho = n_-, n_- + 1, \cdots, n_+$ where these bounds satisfy

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = (n_+ - n_-) \in \mathbb{Z}. \tag{384}$$

Proof. The vector space spanned by $S_{\alpha,\beta}$ is

$$V_{\alpha,\beta} = \text{span}_{\mathbb{C}}\{e_{\beta+\rho\alpha} \mid \beta+\rho\alpha \in S_{\alpha,\beta}\}. \tag{385}$$

The action of $\mathfrak{sl}(2)_\alpha$ on $V_{\alpha,\beta}$ is

$$\text{ad}_H e_{\beta+\rho\alpha} = [h_\alpha, e_{\beta+\rho\alpha}] = \frac{2(\alpha,\beta+\rho\alpha)}{(\alpha,\alpha)} e_{\beta+\rho\alpha} = \left(\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho\right) e_{\beta+\rho\alpha}. \tag{386}$$

Also

$$[e_{\pm\alpha}, e_{\beta\pm\rho\alpha}] \propto \begin{cases} e_{\beta+(\rho\pm 1)\alpha}, & \text{if } \beta+(\rho\pm1)\alpha \in \Phi, \\ 0, & \text{otherwise.} \end{cases} \tag{387}$$

So $V_{\alpha,\beta}$ is closed under $\mathfrak{sl}(2)_\alpha$ and thus $V_{\alpha\beta}$ is a representation space for some representation of $\mathfrak{sl}(2)_\alpha$.

The weight space for representations of $\mathfrak{su}(2,\mathbb{C}) \cong \mathfrak{sl}(2)_{\mathbb{C}}$ are characterized by weight Λ such that

$$S_\Lambda = \left\{\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2\rho \mid \beta+\rho\alpha \in \Phi, \rho \in \mathbb{Z}\right\} = \{-\Lambda, \cdots, \Lambda-2, \Lambda\}. \tag{388}$$

For some integers n_- and n_+ the extrema of this set are satisfied such that for minimal and maximal ρ respectively,

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_- = -\Lambda \qquad \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_+ = \Lambda. \tag{389}$$

Adding these two equations we see

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = n_+ - n_- \in \mathbb{Z}. \tag{390}$$

Thus we have the desired result, sometimes referred to as a *'quantization' condition.* □

Claim. For $\alpha,\beta \in \Phi$, we can write $(\alpha,\beta) = \frac{1}{\mathcal{N}}\sum_{\gamma\in\Phi}(\alpha,\gamma)(\gamma,\beta)$. We set $\mathcal{N}=1$ in this course.

Proof. Recall $[H_i,H_j]=0$ and $[H_i,E_\gamma] = \gamma_i E_\gamma$ where $\{\text{ad}_{H_i}\}$ are diagonalizable with entries either 0 or $\{\gamma_i\}$. So

$$\kappa_{ij} = \kappa(H_i,H_j) = \frac{1}{\mathcal{N}} \text{tr}\left(\text{ad}_{H_i} \circ \text{ad}_{H_j}\right) = \frac{1}{\mathcal{N}} \sum_{\gamma\in\Phi} \gamma_i\gamma_j. \tag{391}$$

Define upper indices such that $\alpha^i = \left(\kappa^{-1}\right)^{ij}\alpha_j$ and $\beta^j = \left(\kappa^{-1}\right)^{jk}\beta_k$ then

$$(\alpha,\beta) = \alpha_i\beta_j\left(\kappa^{-1}\right)^{ij} \tag{392}$$
$$= \alpha^i\beta^j\kappa_{ij} \tag{393}$$

$$= \frac{1}{\mathcal{N}} \sum_{\gamma \in \Phi} (\alpha, \gamma)(\gamma, \beta). \tag{394}$$

□

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

Proof. Divide the above by $(\alpha, \alpha)(\beta, \beta)/4$ and see that

$$\frac{2}{(\beta, \beta)} \underbrace{\frac{2(\alpha, \beta)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} = \frac{1}{\mathcal{N}} \sum_{\gamma \in \Phi} \underbrace{\frac{2(\alpha, \gamma)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} \underbrace{\frac{2(\gamma, \beta)}{(\beta, \beta)}}_{\in \mathbb{Z}}. \tag{395}$$

Either $(\alpha, \beta) = 0 \in \mathbb{R}$ or $(\beta, \beta) \in \mathbb{R} \setminus \{0\}$ and $(\alpha, \beta) \in \mathbb{R}$. □

Lecture 19
23/11/2024

Claim. The roots in Φ span $\mathfrak{h}^*$. Recall $|\Phi| > \dim h^*$.

Proof. Suppose that the roots do not span $\mathfrak{h}^*$. Then there exists some orthogonal subspace, i.e. $\exists \lambda \in \mathfrak{h}^*$ such that $(\lambda, \alpha) = 0$, $\forall \alpha \in \Phi$.

The corresponding vector in $\mathfrak{h}$ is

$$H_\lambda \equiv \lambda^i H_i \in \mathfrak{h}. \tag{396}$$

As $\mathfrak{h}$ is the Cartan subalgebra, $[H_\lambda, H] = 0$, $\forall H \in \mathfrak{h}$. We also know that $[H_\lambda, E_\alpha] = \lambda^i [H_i, E_\alpha] = \lambda^i \alpha_i E_\alpha = (\lambda, \alpha) E_\alpha = 0$, $\forall \alpha \in \Phi$.

So H_λ commutes with all $x \in \mathfrak{g}$ and therefore $\text{span}_{\mathbb{C}}\{H_\lambda\}$ is a nontrivial abelian ideal which contradicts $\mathfrak{g}$ being semisimple. Thus the roots must span $\mathfrak{h}^*$. □

Choose a basis $\{\alpha_{(i)} \in \Phi \mid i = 1, \cdots, r\}$ where the parentheses indicate we are discussing a basis. We than write

$$\mathfrak{h}^* = \text{span}_{\mathbb{C}}\{\alpha_{(i)}\}. \tag{397}$$

We call these **simple roots**.

Let $\mathfrak{h}^*_{\mathbb{R}} \subset \mathfrak{h}^*$ be the real vector space

$$\mathfrak{h}^*_{\mathbb{R}} = \text{span}_{\mathbb{R}}\{\alpha_{(i)}\}, \tag{398}$$

for $i = 1, \cdots, r$.

Claim. $\mathfrak{h}^*_{\mathbb{R}}$ contains all roots.

Proof. Let $\beta \in \Phi$. We can find coefficients such that $\beta = \beta^i \alpha_{(i)}$. Taking an inner product, we see

$$\left(\beta, \alpha_{(j)}\right) = \beta^i \left(\alpha_{(i)}, \alpha_{(j)}\right). \tag{399}$$

As we established these inner products are real, and the equations above are non-degenerate, we must have real coefficients $\beta^i \in \mathbb{R}$. Thus $\beta \in \mathfrak{h}^*_{\mathbb{R}}$. □

Claim. $\forall \lambda \in \mathfrak{h}^*_{\mathbb{R}}$, $(\lambda, \lambda) \geq 0$, with $(\lambda, \lambda) = 0 \Leftrightarrow \lambda = 0$

Proof. Recall

$$(\lambda, \lambda) = \frac{1}{\mathcal{N}} \sum_{\gamma \in \Phi} (\lambda, \gamma)(\gamma, \lambda) = \frac{1}{\mathcal{N}} \sum_{\gamma \in \Phi} (\lambda, \gamma)^2 \geq 0. \tag{400}$$

Further, $(\lambda, \lambda) = 0 \Leftrightarrow (\lambda, \gamma) = 0$, $\forall \gamma \in \Phi$, including $\{\alpha_{(i)}\}$. If the coefficients are zero for the basis set then we must have $\lambda = 0$. □

Definition 8.22: Define $|\alpha| = \sqrt{(\alpha, \alpha)}$ as the **length** or **norm** of a root $\alpha \in \Phi$.

This also holds identically for any $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$.

For any $\alpha, \beta \in \Phi$, there is an **angle** θ such that

$$(\alpha, \beta) = |\alpha| \, |\beta| \cos\theta. \tag{401}$$

Then the quantization condition which gave us $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$ gives us that

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 2\frac{|\beta|}{|\alpha|} \cos\theta \in \mathbb{Z} \qquad \frac{2(\alpha, \beta)}{(\beta, \beta)} = 2\frac{|\alpha|}{|\beta|} \cos\theta \in \mathbb{Z}. \tag{402}$$

Multiplying these together, we see

$$4\cos^2\theta \in \mathbb{Z}, \tag{403}$$

and thus as $0 \leq \cos^2\theta \leq 1$, this must be 0, 1, 2 3, or 4. Therefore with $\cos\theta = \pm\sqrt{n}/2$, for $n \in Z$, the angle is restricted to

$$|\theta| = \left\{0, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{2}, 2\frac{\pi}{3} 3\frac{\pi}{4}, 5\frac{\pi}{6}, \pi\right\}. \tag{404}$$

8.7 Simple roots

Recall that $|\Phi| > \dim \mathfrak{h}^*_{\mathbb{R}} = r$. We pick a hyperplane of dimension $r = 1$ to partition $\mathfrak{h}^*_{\mathbb{R}}$ in two. The hyperplane includes the origin and excludes any roots.

The roots are divided into two sets of equal number. Call one set positive roots and the other negative such that

$$\Phi = \Phi_+ \cup \Phi_-, \tag{405}$$

where $\alpha \in \Phi_+ \Leftrightarrow -\alpha \in \Phi_-$ and $\alpha, \beta \in \Phi_+ \Rightarrow \alpha + \beta \in \Phi_+$.

Definition 8.23: A **simple root** is a positive root which cannot be written as a sum of other positive roots. The set of simple roots is denoted by Φ_S.

We will see that Φ_S forms a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

Claim. If $\alpha, \beta \in \Phi_S$ then $\alpha - \beta \notin \Phi$.

Proof. We proceed by contradiction. Suppose $\alpha - \beta \in \Phi$.

a) If $\alpha - \beta \in \Phi_+$, then $\alpha = (\alpha - \beta) + \beta$ which implies α is not a simple root which is a contradiction.

b) If $\alpha - \beta \in \Phi_-$, then $\beta - \alpha \in \Phi_+$ and thus $\beta = (\beta - \alpha) + \alpha$ and thus β is not a simple root.

□

Recall the α-string through β

$$S_{\alpha,\beta} = \left\{\beta + \rho\alpha \mid \rho \in \mathbb{Z}, n_- \leq \rho \leq n_+\right\}, \tag{406}$$

with $n_+ + n_- = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in Z$.

Definition 8.24: The **length** of a root string is

$$\ell_{\alpha,\beta} = n_+ - n_- + 1. \tag{407}$$

For $\alpha, \beta \in \Phi_S$, the α-root string through β has length $\ell_{\alpha,\beta} = n_+ + 1 = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 1$.

By the above claim, $n_- = 0$.

Claim. $(\alpha, \beta) \leq 0$ for $\alpha, \beta \in \Phi_S$ and $\beta \neq \alpha$.

Proof. This follows from $\ell_{\alpha,\beta} \geq 1$ and $(\alpha, \alpha) > 0$. □

Claim. Any $\beta \in \Phi_+$ can be written as a linear combination of simple roots with integer coefficients.

Proof. If $\beta \in \Phi_S$, we are done. If $\beta \notin \Phi_S$, then $\beta = \beta_1 + \beta_2$. If β_1 and β_2 are simple roots, we are done. Otherwise one iterates and thus terminates as $\dim \mathfrak{h}^*_{\mathbb{R}} = r$ is finite. □

Claim. All roots $\alpha \in \Phi$ can be written

$$\alpha = \sum_i p_i \alpha_{(i)}, \tag{408}$$

with $p_i \in \mathbb{Z}$ and $\alpha_{(i)} \in \Phi_S$.

By the properties above, all p_i have the same sign or are 0.

Claim. The simple roots are linearly independent.

Lecture 20
26/11/2024

Proof. All $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$ can be written as

$$\lambda = \sum_i c_i \alpha_{(i)}, \tag{409}$$

for $c_i \in \mathbb{R}$. Linear independence states that $\lambda = 0 \Leftrightarrow c_i = 0$, $\forall i$. Assume $\exists c_i \neq 0$. Let $J_\pm = \{i \mid c_i \gtrless 0\}$ and define

$$\lambda_+ = \sum_{j \in J_+} c_i \alpha_{(i)}, \qquad \lambda_- = -\sum_{j \in J_-} c_j \alpha_{(j)} = \sum_{j \in J_-} b_j \alpha_{(j)}, \tag{410}$$

with $b_j = -c_j > 0$. Then one can write

$$\lambda = \lambda_+ - \lambda_- = \sum_{j \in J_+} c_i \alpha_{(i)} - \sum_{j \in J_-} b_j \alpha_{(j)}. \tag{411}$$

Then observe that the inner product

$$(\lambda, \lambda) = (\lambda_+, \lambda_+) + (\lambda_-, \lambda_-) - 2(\lambda_+, \lambda_-) > -2(\lambda_+, \lambda_-), \tag{412}$$

since at least one of (λ_+, λ_+) and (λ_-, λ_-) are greater than 0. Further

$$(\lambda_+, \lambda_-) = \sum_{i \in J_+} \sum_{j \in J_-} c_i b_j \left(\alpha_{(i)}, \alpha_{(j)}\right) \leq 0, \tag{413}$$

as $\left(\alpha_{(i)}, \alpha_{(j)}\right) \leq 0$ for distinct simple roots and $c_i b_j > 0$ by construction. Therefore $(\lambda, \lambda) > 0$ and thus the simple roots are linearly independent. □

There are $r = \dim \mathfrak{h}^*_{\mathbb{R}}$ simple roots, $|\Phi_S| = r$, as they minimally (i.e. linearly independently) span $\mathfrak{h}^*_{\mathbb{R}}$.

8.8 Classification

As the simple roots form a basis for $\mathfrak{h}^*_{\mathbb{R}}$ we want extend this to write a basis for $\mathfrak{g}$ using them as well. This is called a **Chevelley basis**.

Definition 8.25: Define the $r \times r$ **Cartan matrix** A such that

$$A_{ji} = \frac{2\left(\alpha_{(j)}, \alpha_{(i)}\right)}{\left(\alpha_{(i)}, \alpha_{(i)}\right)}, \tag{414}$$

which is notably not symmetric. However, A_{ji} are integers due to the quantization condition and in fact, $A_{ji} \in \mathbb{Z}_{\leq 0}$ for $i \neq j$. For $i = j$, $A_{ji} = 2$.

We had commutation relations for h_α and e_β. For $\alpha_{(i)}, \alpha_{(j)} \in \Phi_S$,

$$\left[h_{\alpha_{(i)}}, h_{\alpha_{(j)}}\right] = 0 \tag{415}$$

$$\left[h_{\alpha_{(i)}}, e_{\pm\alpha_{(j)}}\right] = \pm A_{ji} e_{\pm\alpha_{(i)}}, \tag{416}$$

$$\left[e_{\alpha_{(i)}}, e_{\alpha_{(j)}}\right] = \delta_{ij} h_{\alpha_{(i)}}. \tag{417}$$

One can then see the relation

$$\left[e_{\alpha_{(i)}}, e_{\alpha_{(j)}}\right] = \mathrm{ad}_{e_{\alpha_{(i)}}}\left(e_{\alpha_{(j)}}\right) \propto e_{\alpha_{(i)}+e_{\alpha_{(j)}}}, \tag{418}$$

as long as $\alpha_{(i)} + \alpha_{(j)} \in \Phi$. In this case, $\alpha_{(i)} + \alpha_{(j)}$ is part of a root string which we can write as $n\alpha_{(i)} + \alpha_{(j)}$ and the string has length $\ell = 1 - A_{ji}$. For ℓ applications of ad, as the string ends we have

$$\left(\mathrm{ad}e_{\alpha_{(i)}}\right)^{1-A_{ji}}\left(e_{\alpha_{(j)}}\right) = 0. \tag{419}$$

This is the **Serre relation**.

The Cartan matrix is quite heavily constrained by the algebra. In fact, any finite dimensional, complex, Lie algebra is *uniquely determined* by its Cartan matrix.

As such, the entries are of particular interest and we can identify general properties:

i) $\forall i$ fixed, $A_{ii} = 2$ on the diagonal.

ii) $A_{ji} = 0 \Leftrightarrow A_{ij} = 0$ by symmetry of the inner product.

iii) $A_{ji} \in \mathbb{Z}_{\leq 0}$ $\forall i \neq j$ from $\left(\alpha_{(i)}, \alpha_{(j)}\right) \leq 0$.

iv) $\det A > 0$.

> ***Proof.*** Write $A = \kappa^{-1}D$ where $D^j_k = \frac{2}{\left(\alpha_{(j)},\alpha_{(j)}\right)}\delta^j_k$ with no summation which implies $\det D > 0$. Recall $(\lambda, \lambda) = \lambda_i \left(\kappa^{-1}\right)^{ij}\lambda_j > 0$ and thus κ^{-1} is positive definite and thus $\det A = \det \kappa^{-1} \det D > 0$ as desired. □

Proposition 8.8: For simple Lie, algebras, A is irreducible, i.e. cannot be made block triangular by a permutation transformation, where P is a permutation matrix. Otherwise $\mathfrak{g}$ is semisimple.

Claim. $A_{ij}A_{ji} \in \{0, 1, 2, 3, \}$ for $i \neq j$ without summation.

> ***Proof.*** Recall
>
> $$\left(\alpha_{(i)}, \alpha_{(j)}\right) = \left|\alpha_{(i)}\right|\left|\alpha_{(j)}\right| \cos\phi_{ij}, \tag{420}$$
>
> which implies
>
> $$A_{ij}A_{ji} = 4\cos^2\phi_{ij} \in [0, 4]\,. \tag{421}$$
>
> If $\phi = \{0, \pi, \cdots, n\pi\}$, then $\alpha_{(i)} = \pm\alpha_{(j)}$ which is not possible for simple roots and thus $A_{ij}A_{ji} \in \{0, 1, 2, 3\}$. □

This implies for simple complex Lie algebras with rank $r = 2$ that one has

$$A = \begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix}, \tag{422}$$

with $m, n \in \mathbb{Z}_{\geq 0}$ and $\det A = 4 - mn > 0$. This implies we have

$$(m, n) = \{(1,1), (1,2), (2,1), (1,3), (3,1)\}\,. \tag{423}$$

Note. $(m, n) = (0, 0)$ gives a diagonal matrix which is reducible and thus not semisimple.

Note that the Cartan matrix completely specifies and defines the simple Lie algebra (up to relabelling roots) and thus differences like $(1, 2)$ and $(2, 1)$ are trivial permutations that correspond to the same algebra.

8.9 Dynkin diagrams

Definition 8.26: Given a Cartan matrix A, its corresponding **Dynkin diagram** is defined through the procedure as follows:

i) Draw a node for each simple root $\alpha_{(i)}$.

ii) Join the nodes representing $\alpha_{(i)}$ and $\alpha_{(j)}$ with $\max\left(|A_{ij}|, |A_{ji}|\right) \in \{0, 1, 2, 3\}$ lines.

iii) If the number of lines is greater than 1, add an arrow pointing from the longer root $\alpha_{(\ell)}$ to the shorter root, $\alpha_{(s)}$, namely, with $\left|\alpha_{(\ell)}\right| > \left|\alpha_{(s)}\right|$

Example. For $r = 2$ with $(m, n) = (1, 1)$ we have

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{424}$$

This has Dynkin diagram $\bullet\!-\!\bullet$.

If we take $(m, n) = (2, 1)$ then

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}. \tag{425}$$

This has Dynkin diagram $\bullet\!\Rightarrow\!\bullet$ where the arrow is pointing towards $\alpha_{(2)}$ as $\alpha_{(1)}$ is larger (as we will show).