Part III: Symmetries, Fields and Particles - Revision

Lectures by Nick Dorey, notes by James Moore

1 Lie groups

1.1 Definitions

Definition: A *Lie group* is a group which is also a manifold. The group operations are *smooth maps*. That is, if we work in some coordinate patch P, with coordinates $\{\theta^i\}_{i=1...D}$, then group multiplication:

$$g(\boldsymbol{\theta})g(\boldsymbol{\theta'}) = g(\boldsymbol{\phi})$$

gives a smooth map of the coordinates $\phi^i = \phi^i(\boldsymbol{\theta}, \boldsymbol{\theta'})$. Also, group inversion

$$g(\boldsymbol{\theta})g(\tilde{\boldsymbol{\theta}}) = g(\tilde{\boldsymbol{\theta}})g(\boldsymbol{\theta}) = e.$$

must give a smooth map of the coordinates $\tilde{\theta}^i = \tilde{\theta}^i(\boldsymbol{\theta})$.

Definition: The *dimension* of a Lie group G, $\dim(G)$, is the dimension of its group manifold.

Example: $G = (\mathbb{R}^D, +)$ is a Lie group. The group multiplication $\mathbf{x}'' = \mathbf{x} + \mathbf{x}'$ is obviously smooth, and the group inverse $\mathbf{x}^{-1} = -\mathbf{x}$ is again obviously smooth.

1.2 Proving objects are manifolds

Theorem (Implicit Function): Let $\mathbf{x} = (x_1, ..., x_{n+m}) \in \mathbb{R}^{n+m}$. Suppose for $\alpha = 1, ..., m$ we have differentiable functions

$$F^\alpha : \mathbb{R}^{n+m} \to \mathbb{R}.$$

Define

$$M = \{\mathbf{x} : F^\alpha(\mathbf{x}) = 0, \alpha = 1, ..., m\}.$$

If the Jacobian matrix

$$J^\alpha{}_i = \frac{\partial F^\alpha}{\partial x_i}$$

has rank m, then M is a manifold of dimension n.

Example: Consider the 2-sphere S^2. Let $m = 1$ and $n = 2$. Define $F^1(\mathbf{x}) = x^2 + y^2 + z^2 - r^2$. Then

$$J = \begin{pmatrix} \frac{\partial F^1}{\partial x} & \frac{\partial F^1}{\partial y} & \frac{\partial F^1}{\partial z} \end{pmatrix} = 2\begin{pmatrix} x & y & z \end{pmatrix}.$$

Since (x, y, z) spans a space of dimension 1 whenever $(x, y, z) \neq (0, 0, 0)$ we're done (since $(0, 0, 0)$ not on sphere). It follows that $x^2 + y^2 + z^2 = r^2$ is a manifold.

Theorem (Open Subset): An open subset of a manifold is a manifold.

Theorem (Closed Subgroup): A closed subgroup of a Lie group is also a Lie group.

1.3 Matrix groups

Definition: Write $\text{Mat}_n(\mathbb{F})$ for the set of $n \times n$ matrices over the field $\mathbb{F}$.

Theorem: $\text{Mat}_n(\mathbb{F})$ is a manifold for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

Proof: Trivially, $\text{Mat}_n(\mathbb{R}) = \mathbb{R}^{n^2}$ which is obviously a manifold, and $\text{Mat}_n(\mathbb{C}) = \mathbb{R}^{2n^2}$, which is obviously a manifold. □

Note, however, that $\text{Mat}_n(\mathbb{F})$ is not a Lie group, because not all matrices are invertible.

Definition: The *general linear group* is defined by

$$GL(n, \mathbb{F}) = \{M \in \text{Mat}_n(\mathbb{F}) : \det(M) \neq 0\},$$

and the *special linear group* is defined by

$$SL(n, \mathbb{F}) = \{M \in \text{Mat}_n(\mathbb{F}) : \det(M) = 1\}.$$

Theorem: $GL(n, \mathbb{F})$ and $SL(n, \mathbb{F})$ are Lie groups (for $\mathbb{F} = \mathbb{R}, \mathbb{C}$).

Proof: $GL(n, \mathbb{F})$ is a manifold by the open subset theorem. It is a Lie group because matrix multiplication and inversion are clearly smooth functions of the coordinates. Then $SL(n, \mathbb{F})$ is a closed subgroup of $GL(n, \mathbb{F})$, so is a Lie group by the closed subgroup theorem. □

Theorem: The general/special linear groups have dimensions:

$$\dim(GL(n, \mathbb{R})) = n^2, \qquad \dim(GL(n, \mathbb{C})) = 2n^2,$$
$$\dim(SL(n, \mathbb{R})) = n^2 - 1, \qquad \dim(SL(n, \mathbb{C})) = 2n^2 - 2.$$

Proof: $GL(n, \mathbb{F})$ is just an open subset of $\text{Mat}_n(\mathbb{F})$ so has the same dimension as $\text{Mat}_n(\mathbb{F})$. Special linear groups have the additional constraint $\det(A) = 1$. This takes one degree of freedom in real case, 2 in complex case (need both real and imaginary parts to match). □

1.4 Orthogonal groups

Definition: The *orthogonal group* is defined by $O(n) = \{M \in GL(n, \mathbb{R}) : M^T M = I\}$.

Since $\det(M^T M) = 1 \Rightarrow \det(M)^2 = 1 \Rightarrow \det(M) = \pm 1$, $O(n)$ has *two connected components*.

Definition: The *special orthogonal group* is defined by $SO(n) = \{M \in O(n) : \det(M) = 1\}$.

$SO(n)$ has only one connected component. Both $O(n)$ and $SO(n)$ are obviously Lie groups by the closed subgroup theorem.

Theorem: Orthogonal transformations preserve lengths of vectors.

Proof: Let $\mathbf{v}' = M\mathbf{v}$ with $M \in O(n)$. Then

$$|\mathbf{v}'|^2 = \mathbf{v}'^T \mathbf{v}' = \mathbf{v}^T M^T M \mathbf{v} = \mathbf{v}^T \mathbf{v} = |\mathbf{v}|^2. \qquad \square$$

Definition: Given a frame $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$, we define the *volume element* by

$$\Omega = \epsilon_{i_1 i_2 ... i_n} v_1^{i_1} ... v_n^{i_n}.$$

A transformation $\mathbf{v}' = M\mathbf{v}$ *preserves volume* if $|\Omega'| = |\Omega|$. A transformation *preserves orientation* if Ω' has the same sign as Ω.

Theorem: Elements of $O(n)$ preserve volume. Elements of $SO(n)$ preserve volume and orientation.

Proof: We have:

$$\begin{aligned} \Omega' = \epsilon_{i_1 i_2 ... i_n} v_1'^{i_1} ... v_n'^{i_n} &= \epsilon_{i_1 i_2 ... i_n} M^{i_1}{}_{j_1} v_1^{j_1} ... M^{i_n}{}_{j_n} v_n^{j_n} \\ &= \det(M) \epsilon_{j_1 ... j_n} v_1^{j_1} ... v_n^{j_n}, \end{aligned}$$

by a standard property of $\det$. Result is clear from here. $\square$

In particular, this theorem implies that elements of $O(n)$ with $\det(M) = +1$ are *rotations* and elements of $O(n)$ with $\det(M) = -1$ are the compositions of a rotation, followed by a reflection.

Theorem: The dimensions of the orthogonal groups are:

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

Proof: Orthogonal matrices have orthonormal columns: $\mathbf{m}_i^T \mathbf{m}_j = \delta_{ij}$. Consider building up $M \in O(n)$ column by column.

First column is any unit vector: $\mathbf{m}_1^T \mathbf{m}_1 = 1$. So $n-1$ degrees of freedom.

Second column is a unit vector obeying $\mathbf{m}_2^T \mathbf{m}_1$. So $n-2$ degrees of freedom. Similarly, third column is a unit vector obeying two orthogonality relations, so $n-3$ degrees of freedom.

Hence total dimension is: $n-1+n-2+...+1 = \frac{1}{2}n(n-1)$. Since $SO(n)$ is just some connected component of $O(n)$, it has the same dimension as the manifold of $O(n)$. $\square$

Theorem: The eigenvalues of an orthogonal matrix obey: (i) λ is an eigenvalue $\Rightarrow$ λ^* is an eigenvalue; (ii) $|\lambda| = 1$.

Proof: (i) Let $M\mathbf{v} = \lambda \mathbf{v}$. Then $M\mathbf{v}^* = \lambda^* \mathbf{v}^*$ since M is a real matrix. (ii) Evaluate $(M\mathbf{v}^*)^T M\mathbf{v}$ in two ways. We have:

$$\begin{aligned} (M\mathbf{v}^*)^T M\mathbf{v} &= \mathbf{v}^\dagger M^T M \mathbf{v} = \mathbf{v}^\dagger \mathbf{v}; \\ (M\mathbf{v}^*)^T M\mathbf{v} &= (\lambda^* \mathbf{v}^*)^T \lambda \mathbf{v} = |\lambda|^2 \mathbf{v}^\dagger \mathbf{v}. \end{aligned}$$

The result follows. $\square$

1.5 Orthogonal groups of small dimension

Example: Consider $G = SO(2)$. A general element is:

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

This is uniquely specified by some $\theta \in [0, 2\pi]$. Hence the group manifold is $\mathcal{M}(SO(2)) \cong S^1$, the circle.

Example: Consider $G = SO(3)$. As discussed before, these are the 3D rotation matrices. A general element may be specified by a unit vector $\hat{\mathbf{n}}$, the axis of rotation, and $\theta \in [0, 2\pi]$, the angle of rotation; write $M(\hat{\mathbf{n}}, \theta)$ to mean the associated element in G. However, we must identify:

$$M(\hat{\mathbf{n}}, 2\pi - \theta) = M(-\hat{\mathbf{n}}, \theta) \qquad \text{(see diagram)}.$$

Thus we restrict $\theta \in [0, \pi]$, and quotient the coordinates $(\hat{\mathbf{n}}, \theta)$ by the equivalence relation $(\hat{\mathbf{n}}, \pi) \sim (-\hat{\mathbf{n}}, \pi)$ (see diagram).

This identification allows us to write the coordinates as $\boldsymbol{\omega} = \theta\hat{\mathbf{n}}$, so that the full set of coordinates is:

$$B_3 = \{\boldsymbol{\omega} \in \mathbb{R}^3 : |\boldsymbol{\omega}| \leq \pi\}.$$

The group manifold is B_3, together with the proviso that whenever we leave the ball B_3, we reappear at the antipodal point to the point at which we left (see diagram).

The group manifold $\mathcal{M}(SO(3))$ has the properties that it is (i) compact; (ii) has no boundary; (iii) is *path connected*, but is not *simply connected* (recall simply connected means all loops are homotopic to a point).

The proof of (iii) is geometric: consider a loop that connects two antipodal points as shown in the diagram. Then we can never contract the loop, because we can never bring the endpoints together (try moving one, the other will move in the opposite direction):

We can identify the fundamental group of the manifold as follows. The only non-trivial loops are the ones that connect antipodal points. If we do travel along one of these loops, then another, then we can contract to a point by moving one of the loops round:

Hence the fundamental group must be $\pi_1(SO(3)) \cong \mathbb{Z}_2$.

1.6 Non-compact matrix groups

Consider matrices preserving more general metrics.

Definition: The matrix group $O(p,q)$ are defined by:

$$O(p,q) = \{M \in GL(n,\mathbb{R}) : M^T\eta M = \eta\},$$

where $\eta = \text{diag}(\underbrace{+1,+1,...,+1}_{p\text{ times}},\underbrace{-1,-1,...,-1}_{q\text{ times}})$.

Example: The Lorentz group is $O(3,1)$, which consists of matrices preserving the Minkowski metric.

Example: A general element of the group $SO(1,1)$ (with obvious meaning) can be written as:

$$M(\phi) = \begin{pmatrix} \cosh(\phi) & \sinh(\phi) \\ \sinh(\phi) & \cosh(\phi) \end{pmatrix}, \qquad \phi \in \mathbb{R}.$$

The group manifold is $\mathcal{M}(SO(1,1)) = \mathbb{R}$; this is non-compact as it is not bounded.

Definition: We say a Lie group is *compact* if its manifold is compact (i.e. closed and bounded). Otherwise a Lie group is *non-compact*.

Example: The matrices

$$U = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}$$

with $|\alpha|^2 - |\beta|^2 = 1$ form a group. Writing $\alpha = x_1 + ix_2$ and $\beta = x_3 + ix_4$, we see the determinant condition translates to $x_1^2 + x_2^2 - x_3^2 - x_4^2 = 1$. So this is a manifold by the implicit function theorem, and it is also non-compact.

1.7 Subgroups of $GL(n,\mathbb{C})$

Definition: The *unitary group* is defined by

$$U(n) = \{U \in GL(n,\mathbb{C}) : U^\dagger U = I\}.$$

Unitary matrices are analogous to rotations in the complex plane. They preserve the length of complex vectors, $|\mathbf{v}|^2 = \mathbf{v}^\dagger\mathbf{v}$. In fact, we can directly relate unitary matrices to real rotations in double the dimensions. We have:

Lemma: $U(n)$ is path-connected.

Proof: Let $A \in U(n)$. Then A is diagonalisable via a unitary matrix: $A = B\text{diag}(e^{i\theta_1}...e^{i\theta_n})B^\dagger$. Notice any matrix of the form $B\text{diag}(e^{it\theta_1}...e^{it\theta_n})B^\dagger$ is unitary. Then $0 \leq t \leq 1$ is a path from the identity to A. So given any $V \in U(n)$, construct the path from V to the identity, then the path from the identity to A. So path-connected. □

Theorem: $U(n)$ is isomorphic to a subgroup of $SO(2n)$.

*Proof:*Let a general element of $U(n)$ be $A + iB$, where A, B are real. The action of $A + iB$ on the complex vector $\mathbf{x} + i\mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is given by:

$$(A + iB)(\mathbf{x} + i\mathbf{y}) = A\mathbf{x} - B\mathbf{y} + i(B\mathbf{x} + A\mathbf{y}).$$

This suggests defining a map $\psi : U(n) \to SO(2n)$ by:

$$\psi(A + iB) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}.$$

To show $U(n)$ is isomorphic to a subgroup of $SO(2n)$, need (i) ψ maps into $SO(2n)$; (ii) ψ injective; (iii) ψ a homomorphism. Then $U(n)$ is isomorphic to im(ψ).

(ii) ψ is injective, since $\psi(A + iB) = \psi(C + iD) \Rightarrow A + iB = C + iD$, by comparing the matrix expansions.

(iii) ψ is a homomorphism, as can be verified by a short calculation.

Finally, for (i), we use $A + iB \in U(n) \Rightarrow$

$$(A+iB)^\dagger(A+iB) = I \;\Rightarrow\; A^T A + B^T B = I,\; A^T B - B^T A = 0.$$

These conditions allow us to verify $\psi(A + iB) \in O(2n)$ by a calculation. Finally, notice ψ is continuous, $\det$ is continuous and $U(n)$ is path-connected. So image of $\det \circ \psi$ is path-connected, so is equal to $\{+1\}$ or $\{-1\}$. Now $\psi(I_n) = I_{2n}$, hence $\det \circ \psi = +1$, and we're done. □

Theorem: For $U \in U(n)$, $|\det(U)| = 1$.

Proof: $\det(U^\dagger U) = 1 \Rightarrow \det(U)^* \det U = 1$. □

Definition: The *special unitary group* is defined by

$$SU(n) = \{U \in U(n) : \det(U) = 1\}.$$

Both $U(n)$ and $SU(n)$ are trivially Lie groups, by the closed subgroup theorem.

Theorem: $\dim(U(n)) = n^2$, $\dim(SU(n)) = n^2 - 1$.

Proof: Similar to $O(n)$ proof. The columns of a unitary matrix are orthonormal: $\mathbf{u}_i^\dagger \mathbf{u}_j = \delta_{ij}$. So choosing $\mathbf{u}_1$, need $\mathbf{u}_1^\dagger \mathbf{u}_1 = 1$. But $\mathbf{u}_1^\dagger \mathbf{u}_1 \in \mathbb{R}$, so $\mathrm{Im}(\mathbf{u}_1^\dagger \mathbf{u}_1) = 0$ is automatically satisfied. So there are $2n-1$ real degrees of freedom here.

For $\mathbf{u}_2$, lose 1 degree of freedom from normalisation as above, but orthogonality $\mathbf{u}_1^\dagger \mathbf{u}_2 = 0$ needs *both* the real and imaginary parts to vanish. Hence $2n - 3$ degrees of freedom here. So dimension is: $2n-1+2n-3+...+1 = n^2$.

For $SU(n)$, we lose one additional degree of freedom from $\det(U) = 1$, since $\det(U)$ is already constrained to be a phase, so is dependent only on some θ. □

1.8 Unitary groups of small dimension

Definition: Two Lie groups G and G' are *isomorphic* if there exists a smooth bijection with smooth inverse $f : G \to G'$ such that for all $g_1, g_2 \in G$, we have $f(g_1 g_2) = f(g_1) f(g_2)$.

Example: Consider $G = U(1) = \{z \in \mathbb{C} : |z| = 1\}$. This looks like a circle so we might expect $U(1)$ to be isomorphic to $SO(2)$. Define $f : U(1) \to SO(2)$ by $f(e^{i\theta}) = M(\theta)$ where $\theta \in [0, 2\pi)$, and as before:

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

Clearly, this map is invertible, smooth and has smooth inverse. Also $f(e^{i\theta} e^{i\phi}) = M(\theta + \phi) = M(\theta) M(\phi) = f(e^{i\theta}) f(e^{i\phi})$, using $M(\theta + \phi) = M(\theta) M(\phi)$, which follows from the trigonometric addition formulae.

It follows that indeed $U(1) \cong SO(2)$.

Example: Consider $G = SU(2)$. Any element of G can be written uniquely in the form:

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix},$$

where $|\alpha|^2 + |\beta|^2 = 1$. It's then easy to check that any element of G can be written uniquely in the alternative form:

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

where $a_0 \in \mathbb{R}$, $\mathbf{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ and $a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1$. Here, $\boldsymbol{\sigma}$ is the vector of Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \; \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The condition $a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1$ shows the group manifold is $\mathcal{M}(SU(2)) \cong S^3$, the 3-sphere. In particular, the 3-sphere is simply-connected. So $SU(2) \ncong SO(3)$. They are distinct Lie groups.

More on this later; these Lie groups are in fact related by a quotient.

The expression $U = a_0 I + i\mathbf{a} \cdot \boldsymbol{\sigma}$ makes these matrices easier to calculate with. The product of two $SU(2)$ matrices $U = a_0 I + i\mathbf{a} \cdot \boldsymbol{\sigma}$ and $V = b_0 I + i\mathbf{b} \cdot \boldsymbol{\sigma}$ is:

$$UV = a_0 b_0 I + i a_0 \mathbf{b} \cdot \boldsymbol{\sigma} + i b_0 \mathbf{a} \cdot \boldsymbol{\sigma} - (\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}).$$

Now notice:

$$(\mathbf{a} \cdot \boldsymbol{\sigma})(\mathbf{b} \cdot \boldsymbol{\sigma}) = a_i \sigma_i b_j \sigma_j = a_i b_j (\sigma_i \sigma_j) = a_i b_j (I \delta_{ij} + i \epsilon_{ijk} \sigma_k).$$

Hence:

$$UV = (a_0 b_0 - \mathbf{a} \cdot \mathbf{b}) I + i(a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{a} \times \mathbf{b}) \cdot \boldsymbol{\sigma}.$$

2 Lie algebras

2.1 Definitions

Definition: A *Lie algebra* $\mathfrak{g}$ is a vector space with bilinear product (called the *Lie bracket*) $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ which is antisymmetric, and obeys the Jacobi identity:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0.$$

Theorem: A vector space V with associative product $*$ is a Lie algebra with Lie bracket:

$$[X,Y] = X * Y - Y * X.$$

Proof: Tedious checking. □

Definition: The *dimension* of a Lie algebra is the dimension of its underlying vector space.

2.2 Calculations in a basis

Since Lie algebras are vector spaces, we can work in a basis: $B = \{T^a\}$. We can write any element as $X = X_a T^a$.

Definition: For a choice of basis $\{T^a\}$, we call T^a the *generators* of the Lie algebra.

In a basis, we clearly have: $[X,Y] = X_a Y_b [T^a, T^b]$. So knowing the Lie brackets of the generators determines the Lie algebra.

Definition: Write $[T^a, T^b] = f^{ab}{}_c T^c$. We call $f^{ab}{}_c$ the *structure constants* of the Lie algebra in basis $\{T^a\}$.

Theorem: The structure constants obey:

$$f^{ba}{}_c = -f^{ab}{}_c, \qquad f^{ab}{}_c f^{cd}{}_e + f^{da}{}_c f^{cb}{}_e + f^{bd}{}_c f^{ca}{}_e = 0.$$

Proof: Follows from antisymmetry and Jacobi. □

2.3 More definitions

Definition: Two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ are *isomorphic* if there exists a bijective linear map $f : \mathfrak{g} \to \mathfrak{g}'$ such that $[f(X), f(Y)] = f([X,Y])$ for all $X, Y \in \mathfrak{g}$.

Definition: A *subalgebra* $\mathfrak{h} \subseteq \mathfrak{g}$ is a subset of $\mathfrak{g}$ which is also a Lie algebra.

Definition: An *ideal* is a subalgebra which satisifes *strong closure*: $[X,Y] \in \mathfrak{h}$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{h}$.

Example: Every Lie algebra has two *trivial* ideals, $\mathfrak{h} = \{0\}$ and $\mathfrak{h} = \mathfrak{g}$.

Definition: The *derived algebra* is defined by

$$\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}] = \text{span}_{\mathbb{R}}\{[X,Y] : X, Y \in \mathfrak{g}\}.$$

Definition: The *centre* $Z(\mathfrak{g})$ is defined by

$$Z(\mathfrak{g}) = \{X \in \mathfrak{g} : [X,Y] = 0, \forall Y \in \mathfrak{g}\}.$$

A Lie algebra is called Abelian if $Z(\mathfrak{g}) = \mathfrak{g}$ (i.e. all brackets are zero).

Theorem: The derived algebra and the centre are both ideals of a Lie algebra.

Proof: Consider the derived algebra first. This is a real span so is trivially a subalgebra. For any $X \in \mathfrak{g}$ and $Y \in \mathfrak{g}' \subseteq \mathfrak{g}$, $[X,Y]$ is a commutator of elements in $\mathfrak{g}$, so $[X,Y] \in \mathfrak{g}'$.

Now consider the centre. Let $X, Y \in Z(\mathfrak{g})$. Then $[\alpha X + \beta Y, Z] = \alpha[X,Z] + \beta[Y,Z] = 0$ since $[X,Z] = [Y,Z] = 0$ for all $Z \in \mathfrak{g}$. So the centre is a subalgebra. For any $Z \in \mathfrak{g}$ and $X \in Z(\mathfrak{g})$, we have $[Z,X] = 0 \in Z(\mathfrak{g})$. Hence this is an ideal. □

Definition: $\mathfrak{g}$ is *simple* if it is non-Abelian and has no non-trivial ideals.

---

3 Lie algebras from Lie groups

3.1 Tangent spaces

Definition: Let M be a manifold. For $p \in M$, introduce coordinates $\{x^i\}$ in a patch P around p, with $x^i = 0$ at p. The *tangent space* $T_p(M)$ is defined to be the vector space spanned by the differential operators:

$$\left\{ \frac{\partial}{\partial x^i}\bigg|_{x=0} \right\},$$

which act on functions $f(x) : M \to \mathbb{R}$ (when written in terms of their coordinates on M). A *tangent vector* is an element of the tangent space; the most general form is:

$$V = v^i \frac{\partial}{\partial x^i}\bigg|_{x=0}, \qquad v^i \in \mathbb{R}.$$

The action on functions $f(x) : M \to \mathbb{R}$ (written in terms of their coordinates on M) is given by:

$$Vf = v^i \frac{\partial f(x)}{\partial x^i}\bigg|_{x=0}.$$

There is a natural correspondence between tangent vectors at a point p and tangent vectors to curves through a point p on a manifold.

Theorem: There is a one-to-one correspondence between tangent vectors to smooth curves $C : \mathbb{R} \to M$ through p and tangent vectors V at p.

Proof: Let's begin with a curve C and construct the tangent vector $V_C \in T_p(M)$. Write the points in $C : \mathbb{R} \to M$ as $g(t)$, with $g(0) = p$. By the chain rule, the tangent to the curve may be written

$$\frac{dg(t)}{dt} = \frac{dx^i(t)}{dt} \cdot \frac{\partial g(\mathbf{x})}{\partial x^i}.$$

where $g(\mathbf{x})$ is the point on the manifold with coordinates $\mathbf{x}$. Since the curve is smooth, the $x^i(t)$ are smooth functions, so this derivative is allowed. Hence:

$$\dot{g}(0) = \dot{x}^i(0) \frac{\partial g(\mathbf{x})}{\partial x^i}\bigg|_{\mathbf{x}=\mathbf{0}}.$$

So the natural corresponding member of the tangent space is:

$$V_C = \dot{x}^i(0) \frac{\partial}{\partial x^i}\bigg|_{x=0}.$$

Conversely, given a vector

$$V = v^i \frac{\partial}{\partial x^i}\bigg|_{x=0},$$

we can choose any curve C with Taylor series $x^i(t) = x^i(0) + t\dot{x}^i(0) + ... = tv^i + ...$ at $t = 0$ (we can construct an explicit curve using the *exponential map*, see Section 3.8 later. □

From the above Theorem, it's clear that for a curve C, the corresponding tangent vector V_C is the *derivative along* C. That is, given a function f, we have:

$$V_C f = \dot{x}^i(0) \frac{\partial f(x)}{\partial x^i}\bigg|_{x=0} = \frac{df}{dt}\bigg|_{t=0},$$

by the chain rule.

3.2 The Lie algebra of a matrix Lie group

Definition: The *Lie algebra* of a Lie group G, denoted $\mathcal{L}(G)$, is the tangent space to the identity of G:

$$\mathcal{L}(G) = T_e(G).$$

We now show that $\mathcal{L}(G)$ is indeed a Lie algebra in the sense of Section 2 by constructing a bracket $[\cdot, \cdot]$. We'll do this for matrix Lie groups here.

Theorem: Suppose $\{\boldsymbol{\theta}\}$ are coordinates on the matrix Lie group near the identity. Let $g(\boldsymbol{\theta})$ be the group elements near the identity, and let $g(\mathbf{0}) = e$. Then $T_e(G)$ can be identified with the subspace of Mat$_n(\mathbb{F})$ spanned by:

$$\left\{ \frac{\partial g(\boldsymbol{\theta})}{\partial \theta^i}\bigg|_{\boldsymbol{\theta}=\mathbf{0}} \right\}.$$

Proof: Define the map $\rho : T_e(G) \to$ Mat$_n(\mathbb{F})$ by:

$$\rho\left(\frac{\partial}{\partial \theta^i}\right) = \frac{\partial g(\boldsymbol{\theta})}{\partial \theta^i}\bigg|_{\boldsymbol{\theta}=\mathbf{0}},$$

and extend by linearity. This is injective since:

$$\rho\left(v^i \frac{\partial}{\partial \theta^i}\right) - \rho\left(w^i \frac{\partial}{\partial \theta^i}\right) = \rho\left((v^i - w^i)\frac{\partial}{\partial \theta^i}\right)$$
$$(v^i - w^i)\frac{\partial g(\boldsymbol{\theta})}{\partial \theta^i}\bigg|_{\boldsymbol{\theta}=\mathbf{0}}.$$

Provided none of the coordinates are redundant, none of the $\partial g(\boldsymbol{\theta})/\partial \theta^i$ can be zero (else we could describe things without that coordinate - it just remains constant), this is zero iff $v^i = w^i$. So ρ is a bijection between a subspace of Mat$_n(\mathbb{F})$ and $\mathcal{L}(G)$. □

From now on then, identify $\mathcal{L}(G)$ with the subspace of Mat$_n(\mathbb{F})$. There is then an obvious bracket, the matrix commutator:

$$[X, Y] = XY - YX.$$

By some tedious checking, this can be seen to be bilinear, antisymmetric and satisfy the Jacobi identity. We just need to show that $[X, Y] \in \mathcal{L}(G)$.

Theorem: $[X, Y] \in \mathcal{L}(G)$.

Since X and Y are in the Lie algebra, there exist curves through e such that

$$g_1(t) = I + Xt + W_1 t^2 + O(t^3), \qquad g_2(t) = I + Yt + W_2 t^2 + O(t^3).$$

We need to product a curve with tangent $[X, Y]$. Define:

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

This is in G because multiplication and inversion keep us in G, and h is smooth because multiplication and inversion are smooth maps in a Lie group. Then:

$$g_1(t) g_2(t) = g_2(t) g_1(t) h(t).$$

Write $h(t) = I + h_1 t + h_2 t^2 + O(t^3)$. Expanding both sides:

$$g_1(t)g_2(t) = I + t(X + Y) + t^2(XY + W_1 + W_2) + O(t^3),$$
$$g_2(t)g_1(t)h(t) = (I + t(X + Y) + t^2(YX + W_1 + W_2) + O(t^3))h(t).$$

Comparing powers of t, we see $h_1 = 0$, and $h_2 = [X, Y]$. Hence

$$h(t) = I + t^2[X, Y] + O(t^3).$$

So we're close but still a bit off. Simply define $g_3(s) = h(\sqrt{s})$ for $s > 0$, to find the one-sided tangent vector $[X, Y]$ for $s > 0$.

For $s < 0$, instead start with $\tilde{h}(t) = g_2^{-1}(t) g_1^{-1}(t) g_2(t) g_1(t)$, to eventually define $g_3(s) = \tilde{h}(\sqrt{-s})$ for $s < 0$. This gives the other side of the tangent, again equal to $[X, Y]$. □

3.3 Examples of matrix Lie algebras

Example: Consider $G = SO(2)$. A general curve at the identity is:

$$g(t) = M(\theta(t)) = \begin{pmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{pmatrix}$$

with $\theta(0) = 0$. Then

$$\dot{g}(0) = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix} \dot{\theta}(0).$$

Hence $\mathcal{L}(SO(2))$ consists of 2×2 antisymmetric matrices.

Example: Consider $G = SO(n)$. A general curve is $g(t) = R(t) \in SO(n)$, with $R(0) = I$. We have:

$$R^T(t)R(t) = I \qquad \Rightarrow \qquad \dot{R}^T(t)R(t) + R^T(t)\dot{R}(t) = 0.$$

So $\dot{R}^T(0) + \dot{R}(0) = 0$ (since $R(0) = I$). There are no further constraints on the matrices $\dot{R}(0)$ from $\det(R) = 1$, since all orthogonal matrices in a neighbourhood of the identity obey this automatically (recall it splits into two connected components). Thus

$$\mathcal{L}(O(n)) = \mathcal{L}(SO(n)) = \{n \times n \text{ antisymmetric matrices}\}.$$

Thus we can verify:

$$\dim(\mathcal{L}(SO(n))) = \frac{1}{2}n(n-1),$$

as should be the case (Lie algebra must have same dimensions as Lie group).

Lemma: The derivative of $\det(U(t))$ is:

$$\det{}'(U(t)) = \det(U(t))\mathrm{tr}\left(U^{-1}(t)U'(t)\right).$$

Proof: We have:

$$\det{}'(U(t)) = \lim_{h\to 0}\left(\frac{\det(U(t+h)) - \det(U(t))}{h}\right)$$

$$= \det(U(t))\lim_{h\to 0}\left(\frac{\det(I + hU^{-1}(t)U'(t) + O(h^2)) - 1}{h}\right).$$

Consider the Taylor expansion of $\det(I + hM + O(h^2))$. At linear order, this is $\det(I + hM)$, which when expanded is the characteristic polynomial of M; the coefficient of h is the trace of M, and the constant term is $\det(I) = 1$. The formula follows. □

Example: Consider $G = SU(n)$. If $g(t) = U(t)$ is a general curve with $U(0) = I$, then

$$U^\dagger(t)U(t) = I \qquad \Rightarrow \qquad \dot{U}^\dagger(0) + \dot{U}(0) = 0,$$

similar to $O(n)$ and $SO(n)$. There is now an extra condition from determinant. Expanding in a Taylor series about $t = 0$, we have:

$$\underbrace{1 = \det(U(t))}_{\text{condition}} = 1 + t\cdot\mathrm{tr}\left(\dot{U}(0)\right) + O(t^2).$$

So we need $\dot{U}(0)$ to be traceless too. Hence

$$\mathcal{L}(SU(n)) = \{n \times n \text{ traceless, anti-Hermitian } \mathbb{C} \text{ matrices}\},$$
$$\mathcal{L}(U(n)) = \{n \times n \text{ anti-Hermitian } \mathbb{C} \text{ matrices}\}.$$

Again, we can verify the dimensions are correct. For $\mathcal{L}(U(n))$, there are $2\cdot\frac{1}{2}n(n-1)$ degrees of freedom for off-diagonal elements (2 for complex, $\frac{1}{2}n(n-1)$ above-diagonal elements). There are n degrees of freedom from the on-diagonal elements, since anti-Hermitian $\Rightarrow$ the diagonal elements can be any *real* numbers. Total is:

$$2\cdot\frac{1}{2}n(n-1) + n = n^2,$$

as expected. For $\mathcal{L}(SU(n))$, we lose a degree of freedom from the trace condition.

Example: Consider $G = SU(2)$. From above,

$$\mathcal{L}(SU(2)) = \{2 \times 2 \text{ traceless, anti-Hermitian matrices}\}.$$

A basis is $T^a = -\frac{1}{2}i\sigma_a$, where σ_a are the Pauli matrices. The structure constants can be computed from:

$$[T^a, T^b] = -\frac{1}{4}[\sigma_a, \sigma_b] = -\frac{1}{4}(\sigma_a\sigma_b - \sigma_b\sigma_a)$$
$$= -\frac{1}{4}(\delta_{ab}I + i\epsilon_{abc}\sigma_c - \delta_{ab}I - i\epsilon_{bac}\sigma_c) = -\frac{1}{2}\epsilon_{abc}i\sigma_c.$$

Hence the structure constants are: ${f^{ab}}_c = \epsilon_{abc}$.

Example: Consider $G = SO(3)$. From above,

$$\mathcal{L}(SO(3)) = \{3 \times 3 \text{ real antisymmetric matrices}\}$$

Use the basis:

$$\tilde{T}^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tilde{T}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tilde{T}^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This can be written as $(\tilde{T}^a)_{bc} = -\epsilon_{abc}$. By a short calculation,

$$[\tilde{T}^a, \tilde{T}^b] = \epsilon_{abc}\tilde{T}^c.$$

So this Lie algebra has the same structure constants as $\mathcal{L}(SU(2))$. This implies:

$$\mathcal{L}(SO(3)) \cong \mathcal{L}(SU(2)),$$

even though $SO(3) \ncong SU(2)$, as we saw earlier.

3.4 Translations of a Lie group

Definition: The *left* and *right translations* of a Lie group G by an element $h \in G$ are:

$$\begin{aligned} L_h : G \to G, \qquad & L_h(g) = hg, \\ R_h : G \to G, \qquad & R_h(g) = gh. \end{aligned}$$

These maps are both smooth (as this is a Lie group, multiplication of elements is smooth). We also have:

Theorem: Translations are bijective.

Proof: Surjective: For any $g' \in G$, let $g = h^{-1}g'$. Then $L_h(g) = g'$. Similarly for right translation.

Injective: If $L_h(g) = L_h(g')$, then $hg = hg' \Rightarrow g = g'$. Similarly for R_h.□

Theorem: $(L_h)^{-1} = L_{h^{-1}}$.

Proof: We have: $L_{h^{-1}}(L_h(g)) = L_{h^{-1}}(hg) = h^{-1}hg = g$. Similarly $L_h(L_{h^{-1}}(g)) = g$. □

Hence L_h is (i) a bijection, (ii) the inverse is a smooth map since it is of translation form. So L_h is a *diffeomorphism* of G. Similarly R_h is a diffeomorphism.

Introduce coordinates $\{\theta^i\}$ in some region containing the identity element. Then for group elements g, g', with $g' = hg$, near the identity, we have:

$$g' = g(\boldsymbol{\theta}') = L_h(g) = h \cdot g(\boldsymbol{\theta}).$$

So translation induces a map between coordinates $\theta'^i = \theta'^i(\boldsymbol{\theta})$. Since L_h is a diffeomorphism, the Jacobian matrix

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

of this transformation is invertible. We can use this feature to define a map from the tangent space at g to the tangent space at $hg = L_h(g)$:

Definition: The *differential* of the translation L_h is the map $L_h^* : T_g(G) \to T_{hg}(G)$ defined by

$$L_h^*\left(v^j \frac{\partial}{\partial \theta^j}\right) = J^i{}_j(\boldsymbol{\theta}) v^j \frac{\partial}{\partial \theta'^i} =: v'^i \frac{\partial}{\partial \theta'^i}.$$

This is supposed to mirror the chain rule:

$$v^j \frac{\partial}{\partial \theta^j} = v^j \frac{\partial \theta'^i}{\partial \theta^j} \frac{\partial}{\partial \theta'^i}.$$

3.5 Vector fields

Definition: A *vector field* V is a map on a Lie group G, assigning to each point $g \in G$ a tangent vector $V(g) \in T_g(G)$. In coordinates:

$$V(g(\boldsymbol{\theta})) = V(\boldsymbol{\theta}) = v^i(\boldsymbol{\theta}) \frac{\partial}{\partial \theta^i}.$$

A vector field is *smooth* if the components $v^i(\boldsymbol{\theta})$ are smooth functions.

The differential of the translation L_g defines a vector field: given any tangent vector $\omega \in T_e(G)$, define:

$$V(g) = L_g^*(\omega) = J^i{}_j(\boldsymbol{\theta}) \omega^j \frac{\partial}{\partial \theta'^i}.$$

As the Jacobian matrix is smooth and non-singular, and $\omega \neq 0$, $V(g)$ is smooth and *non-vanishing*.

In particular, starting for a basis $\{\omega_a\}$ of the tangent space $T_e(G)$ of dimension D, we get D *independent, nowhere-vanishing* vector fields on G: $V_a(g) = L_g^*(\omega_a)$.

Definition: $\{V_a(g)\}_{a=1\ldots D}$ are called *left-invariant vector fields* on G, since they obey:

$$L_h^*(V_a(g)) = L_h^*(L_g^*(\omega_a)) = L_{hg}^*(\omega_a) = V_a(hg).$$

Using the Hairy Ball Theorem, and the above work, we can deduce that no Lie group has S^2 as a manifold. We have:

Theorem (Hairy Ball): Any smooth vector field on S^2 has at least two zeros (or one double zero).

If S^2 were a Lie group manifold, differentials of translations would give us two nowhere-vanishing vector fields, as above.

3.6 Left translation of matrix Lie groups

For a matrix Lie group, $L_h^*: T_e(G) \to T_h(G)$ is such that:

$$L_h^*(X) = hX \in T_h(G).$$

Here, $h \in G$, $X \in \mathcal{L}(G)$. It's not immediately obvious the result is in $T_h(G)$; we will prove it is.

Theorem: For $h \in G$, $X \in \mathcal{L}(G)$, we have $hX \in T_h(G)$.

Proof: Since $X \in \mathcal{L}(G)$, there exists a curve $g(t) \in G$ with $g(0) = e$, $\dot{g}(0) = X$. So near $t = 0$,

$$g(t) = I + tX + O(t^2).$$

Define a new curve $h(t) = h \cdot g(t) \in G$. Then near $t = 0$,

$$h(t) = h + thX + O(t^2),$$

so $hX \in T_h(G)$. □

3.7 Curves and the exponential map

Theorem: For any smooth curve $g(t)$ with $g(0) = e$, we have $g^{-1}(t)\dot{g}(t) \in \mathcal{L}(G)$ for all t.

Proof: For any t_0, note that:

$$\dot{g}(t_0) \in T_{g(t_0)}(G),$$

since near $t = t_0$, $g(t) = g(t_0) + (t - t_0)\dot{g}(t_0) + O(|t - t_0|^2)$. Multiplying this Taylor expansion by $g^{-1}(t_0)$, we see that:

$$g^{-1}(t_0)\dot{g}(t_0) \in \mathcal{L}(G)$$

for all t_0. Since this holds for all t_0, relabel $t_0 \mapsto t$. □

Theorem (Converse of previous theorem): For any $X \in \mathcal{L}(G)$, we can find a unique curve $g(t)$ with $g^{-1}(t)\dot{g}(t) = X$, and $g(0) = e$.

Proof: Uniqueness/existence Theorem for ODEs. □

Definition: The solution to $g^{-1}(t)\dot{g}(t) = X$, $g(0) = e$, is called the *exponential*, and is written $g(t) = \mathrm{Exp}(tX)$.

For matrix Lie algebras, we can construct this explicitly:

$$\dot{g}(t) = g(t)X, g(0) = I \qquad \Rightarrow \qquad g(t) = \exp(tX),$$

where $\exp$ is the *matrix exponential*:

$$\exp(M) = \sum_{l=0}^{\infty} \frac{1}{l!} M^l.$$

3.8 Application: one-parameter subgroups

Definition: A *one-parameter subgroup* of a Lie group G is a subgroup of the form:

$$S_{X,J} = \{g(t) = \mathrm{Exp}(tX) : t \in J\},$$

for $X \in \mathcal{L}(G)$ and $J \subseteq \mathbb{R}$ some interval.

Theorem: $S_{X,\mathbb{R}}$ is a Lie subgroup (for a matrix Lie group G).

Proof: Identity is $g(0) = I$. Associativity is inherited from G. Need to show closure and inverses. Note:

$$g(t_1)g(t_2) = \sum_{l,r=0}^{\infty} \frac{t_1^l}{l!}\frac{t_2^r}{r!} X^l X^r = \sum_{r=0}^{\infty} \frac{1}{r!}\left(\sum_{i=0}^{r} \frac{r!}{i!(r-i)!} t_1^i t_2^{r-i}\right) X^r$$
$$= \sum_{r=0}^{\infty} \frac{(t_1+t_2)^r}{r!} X^r = g(t_1 + t_2).$$

Hence closed. This also shows $g^{-1}(t) = g(-t)$.

This is a Lie group because it is a manifold with the single chart $\phi(g(t)) = t$. □

From this Theorem, we can identify possible one-parameter subgroups. There are two cases:

(i) There exists $t_1 \neq 0$ such that $g(t_1) = I$. Then $J = [0, t_1]$ gives a one-parameter subgroup, with $S_{X,J} \cong U(1)$. This is compact.

(ii) There does not exist such a t_1. Then $J = \mathbb{R}$ is the only possibility. $S_{X,\mathbb{R}}$ is non-compact.

3.9 Reconstructing G from $\mathcal{L}(G)$

Recall $g(t) = \mathrm{Exp}(tX)$ solves $g^{-1}(t)\dot{g}(t) = X$, $g(0) = e$. Taking $t = 1$, and allowing X to vary, we get a map

$$\mathrm{Exp} : \mathcal{L}(G) \to G.$$

Definition: This map is called the *exponential map*.

Theorem: The exponential map is a bijection between $\mathcal{L}(G)$ and some neighbourhood of $e \in G$.

Proof: Too difficult for this course. □

The Lie algebra can then be used to recover multiplication in the Lie group. This is achieved using the *Baker-Campbell-Hausdorff formula*:

Theorem: Let $g_X = \mathrm{Exp}(X)$, $g_Y = \mathrm{Exp}(Y)$. Then $g_X g_Y = \mathrm{Exp}(Z)$ for some $Z \in \mathcal{L}(G)$, given by

$$Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{2}\left([X,[X,Y]] - [Y,[X,Y]]\right) + ...$$

Proof: Too difficult for this course. □

Exp can fail to be a *global* bijection (between $\mathcal{L}(G)$ and G) in only two ways:

(1) G is not connected. Then Exp is not surjective, and maps only to the connected component containing the identity.

(2) G contains a $U(1)$ subgroup. Then Exp is not injective.

Example: Consider $G = O(n)$. The Lie algebra is the set of antisymmetric matrices. But if X is antisymmetric,

$$\mathrm{tr}(X) + \mathrm{tr}(X^T) = 0 \Rightarrow 2\mathrm{tr}(X) = 0.$$

So X is traceless too. Let X have eigenvalues $\lambda_1, ..., \lambda_n$. Then:

$$\det(\mathrm{Exp}(X)) = e^{\lambda_1}...e^{\lambda_n} = e^{\lambda_1 + ... + \lambda_n} = \exp(\mathrm{tr}(X)) = 1.$$

Hence Exp is not surjective: its image is $SO(n)$.

Example: Consider $G = U(1)$. The Lie algebra is the 1D anti-Hermitian matrices, i.e.

$$\mathcal{L}(U(1)) = \{X = ix \in \mathbb{C} : x \in \mathbb{R}\}.$$

Then $g = \mathrm{Exp}(X) = e^{ix}$. So Exp is not injective, since ix and $2\pi i + ix$ give the same Lie group element.

3.10 The relation between $SU(2)$ and $SO(3)$

We can construct a global 2-to-1 function, called a *double-covering*, from $d : SU(2) \to SO(3)$, finally relating these two Lie groups. Define d by:

$$d(A)_{ij} = \frac{1}{2}\mathrm{tr}\left(\sigma_i A \sigma_j A^\dagger\right)$$

Here, it's clear to see that $d(A) = d(-A)$. So two points in $SU(2)$ get sent to the same point in $SO(3)$. Similarly, the inverse of this map is:

$$d^{-1}(R)_{ij} = \pm\frac{(I_2 + \sigma_i R_{ij} \sigma_j)}{2\sqrt{1 + \mathrm{tr}(R)}},$$

where $R \in SO(3)$.

Theorem: (i) $d(A) \in SO(3)$; (ii) $d^{-1}(R)$ is really the inverse of d; (iii) $SO(3) \cong SU(2)/\mathbb{Z}_2$.

Proof: (i) is an exercise in index manipulation; need to show $d(A)_{ij}(d(A))^T_{jk} = \delta_{ik}$. Use the Pauli matrix identity:

$$\sum_{i=1}^{3}(\sigma_i)_{\alpha\beta}(\sigma_i)_{\gamma\delta} = 2\delta_{\alpha\delta}\delta_{\gamma\beta} - \delta_{\alpha\beta}\delta_{\gamma\delta}.$$

Also note that the map is continuous, $SU(2)$ is connected so the image is connected. Since $d(I) = I$, $d(A) \in SO(3)$.

For (ii), let $R = d(A)$. Then

$$\begin{aligned}\mathrm{tr}(d(A)) &= \frac{1}{2}(\sigma_i)_{ab}A_{bc}(\sigma_j)_{cd}A^\dagger_{da}\\ &= (\delta_{ad}\delta_{bc} - \frac{1}{2}\delta_{ab}\delta_{cd})A_{bc}A^\dagger_{da}\\ &= \mathrm{tr}(A)\mathrm{tr}(A^\dagger) - 1.\end{aligned}$$

Also:

$$\begin{aligned}&(\sigma_i R_{ij} \sigma_j)_{ab} = (\sigma_i)_{ac}(\sigma_j)_{cb}(\sigma_i)_{de}A_{ef}(\sigma_j)_{fg}A^\dagger_{gd}\\ &= \frac{1}{2}(2\delta_{ae}\delta_{cd} - \delta_{ac}\delta_{de})(2\delta_{cg}\delta_{bf} - \delta_{cb}\delta_{fg})A_{ef}A^\dagger_{gd}\\ &= 2A^\dagger_{ab}\mathrm{tr}(A^\dagger) - A^\dagger_{ae}A_{eb} - A_{ag}A^\dagger_{gb} + \frac{1}{2}\delta_{ab}A_{df}A^\dagger_{fd}\\ &= 2A_{ab}\mathrm{tr}(A^\dagger) - \delta_{ab}.\end{aligned}$$

Putting all this together,

$$d^{-1}(R)_{ab} = \pm\frac{A_{ab}\mathrm{tr}(A^\dagger)}{\sqrt{\mathrm{tr}(A)\mathrm{tr}(A^\dagger)}}.$$

Now if $\mathrm{tr}(A) = \lambda_1 + \lambda_2$, then $\mathrm{tr}(A^\dagger) = \mathrm{tr}(A^{-1}) = \frac{1}{\lambda_1} + \frac{1}{\lambda_2}$. Hence

$$\mathrm{tr}(A^\dagger) = \frac{\lambda_2 + \lambda_1}{\lambda_2\lambda_1} = \lambda_2 + \lambda_1 = \mathrm{tr}(A),$$

since $\det(A) = 1$. Hence $d^{-1}(R) = d^{-1}(d(A)) = \pm A_{ab}$.

(iii) It's possible to show the map is a homomorphism too, by similar index manipulation. Since the map is 2 to 1, the kernel is precisely $\{I, -I\}$, since $I_3 = d(I_2) = d(-I_2)$. The result follows by the first isomorphism theorem. □

Slogan:

$$SO(3) \cong \frac{SU(2)}{\mathbb{Z}_2}.$$

4 Representations

4.1 Definitions

Definition: A *representation* of a Lie group G is a smooth homomorphism $D : G \to GL(n, \mathbb{F})$ for some n, $\mathbb{F} = \mathbb{R}, \mathbb{C}$. The homomorphism property means we need:

$$D(g_1)D(g_2) = D(g_1 g_2).$$

Definition: A *representation* of a Lie algebra $\mathfrak{g}$ is a linear map $d : \mathfrak{g} \to \mathsf{Mat}_n(\mathbb{F})$, for some n, such that

$$[d(X_1), d(X_2)] = d([X_1, X_2]).$$

Definition: The *dimension* of a representation is the dimension n of the corresponding matrices in each case.

Definition: The vector space $V = \mathbb{F}^n$ on which the matrices of the representation act is called the *representation space*.

4.2 Lie algebra reps and Lie group reps

Theorem: Let D be a representation of dimension n of a matrix Lie group G. Let $X_1 \in \mathcal{L}(G)$, and let $g_1(t) \in G$ be a curve with $g_1(0) = e$, $\dot{g}_1(0) = X_1$. Then

$$d(X_1) = \frac{d}{dt}\left(D(g_1(t))\right)\bigg|_{t=0}$$

is a representation of $\mathcal{L}(G)$.

Proof: Let $g_2(t) \in G$ be a curve with $g_2(0) = e$, $\dot{g}_2(0) = X_2$. Define, as usual, $h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t) \in G$. This has Taylor expansion: $h(t) = I + t^2[X_1, X_2] + O(t^3)$ as before. Now apply D to $h(t)$:

$$\begin{aligned} D(h(t)) &= D(I + t^2[X_1, X_2] + O(t^3)) \\ &= D(I) + t^2 \frac{d}{d(t^2)}(D(h(t)))\bigg|_{t=0} + \ldots \\ &= I + t^2 d([X_1, X_2]) + \ldots \end{aligned}$$

by Taylor expansion. Applying D to $g_1^{-1}g_2^{-1}g_1g_2$, we have:

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

similarly. Hence $d([X_1, X_2]) = [d(X_1), d(X_2)]$.

Note this is also a linear map, since the derivative of $D : G \to GL(V)$ at the identity *is* a linear map $d : T_e(G) \to T_I(GL(V))$, i.e. $d : \mathfrak{g} \to \mathfrak{gl}(V)$. □

Theorem: Conversely, suppose d is a representation of $\mathcal{L}(G)$. Then if $g = \mathrm{Exp}(X)$, $D(g) = \mathrm{Exp}(d(X))$ obeys:

$$D(g_1 g_2) = D(g_1)D(g_2).$$

Proof: Just use BCH formula. □

Note: This is not necessarily a representation of the Lie group, since Exp can fail to be surjective.

4.3 Important representations

Definition: The *trivial representation* d_0 of a Lie algebra is defined by

$$d_0(X) = 0$$

for all $X \in \mathfrak{g}$. Its dimension is $\dim(d_0) = 1$.

Definition: The *fundamental representation* d_f of a matrix Lie algebra is defined by

$$d_f(X) = X.$$

Its dimension is $\dim(d_f) = n$.

Definition: The *adjoint representation* d_{adj} of a D-dimensional Lie algebra is defined by

$$d_{\mathsf{adj}}(X) = \mathrm{ad}_X,$$

where $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is defined by $\mathrm{ad}_X(Y) = [X, Y]$. Since this is a linear map between vector spaces, ad_X can be identified with a $D \times D$ matrix, so that $\dim(d_{\mathsf{adj}}) = D$.

Theorem: The matrix of ad_X is ${(R_X)^b}_c = X_a {f^{ab}}_c$.

Proof: Let $X = X_a T^a$ and $Y = Y_a T^a$. Then

$$\mathrm{ad}_X(Y) = [X, Y] = X_a Y_b {f^{ab}}_c. \qquad \square$$

Theorem: The adjoint representation d_{adj} is indeed a representation.

Proof: First verify preservation of bracket. We have:

$$\begin{aligned} \left(d_{\mathsf{adj}}(X) \circ d_{\mathsf{adj}}(Y)\right)(Z) &= [X, [Y, Z]], \\ \left(d_{\mathsf{adj}}(Y) \circ d_{\mathsf{adj}}(X)\right)(Z) &= [Y, [X, Z]], \end{aligned}$$

hence

$$\begin{aligned} [d_{\mathsf{adj}}(X), d_{\mathsf{adj}}(Y)](Z) &= [X, [Y, Z]] - [Y, [X, Z]] \\ &= [[X, Y], Z] = d_{\mathsf{adj}}([X, Y]) \end{aligned}$$

by Jacobi identity and antisymmetry. Hence $[d_{\mathsf{adj}}(X), d_{\mathsf{adj}}(Y)] = d_{\mathsf{adj}}([X, Y])$. Finally, linearity is trivial by linearity of ad_X. □

4.4 Isomorphic and irreducible reps

Definition: Two representations R_1 and R_2 of a Lie algebra are *isomorphic* or *equivalent* if there exists a non-singular matrix S such that $R_2(X) = SR_1(X)S^{-1}$ for all X. (That is, we can just change bases in the representation's codomain.)

Definition: A representation R with representation space V has an *invariant subspace* $U \subseteq V$ if

$$R(X)u \in U$$

for all $X \in \mathfrak{g}$, $u \in U$. (*Slogan:* We're stuck in U!)

We note that if U is an invariant subspace of a representation R, then $R(X)$ takes a block upper-triangular structure for each $X \in \mathfrak{g}$:

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

Example: All representations have two trivial invariant subspaces: $U = \{0\}$ and $U = V$.

Definition: An *irreducible representation* (abbreviated to *irrep*) is a representation with no non-trivial invariant subspaces.

4.5 Rep theory of $\mathcal{L}_{\mathbb{C}}(SU(2))$

When studying $\mathcal{L}(SU(2))$, we worked with the basis $\{T^a = -\frac{1}{2}i\sigma_a\}$ with $[T^a, T^b] = \epsilon_{abc}T^c$. However, when studying rep theory it is best to *complexify* the Lie algebra first.

Definition: The *complexification* of $\mathcal{L}(SU(2))$ is:

$$\mathcal{L}_{\mathbb{C}}(SU(2)) = \text{span}_{\mathbb{C}} \left\{ T^a = -\frac{1}{2}i\sigma_a \right\}.$$

Theorem: $\mathcal{L}_{\mathbb{C}}(SU(2)) = \{$traceless 2×2 matrices$\}$.

Proof: Clearly $\mathcal{L}_{\mathbb{C}}(SU(2)) = \{\lambda_1\sigma_1 + \lambda_2\sigma_2 + \lambda_3\sigma_3 : \lambda_i \in \mathbb{C}\}$. Note $\text{tr}(\sum_j \lambda_j\sigma_j) = 0$, so all matrices in $\mathcal{L}_{\mathbb{C}}(SU(2))$ are traceless.

Conversely, let a, b and c be any complex numbers. Then

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = \left(\frac{b+c}{2}\right)\sigma_1 + \left(\frac{b-c}{2}\right)i\sigma_2 + a\sigma_3.$$

So any complex traceless matrix is in $\mathcal{L}_{\mathbb{C}}(SU(2))$. □

Working with $\mathcal{L}_{\mathbb{C}}(SU(2))$ allows us to introduce a complex basis:

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

Definition: This basis is called the *Cartan-Weyl basis*. H is called the *Cartan element*. $E_\pm$ are called *step operators*.

Any element of $\mathcal{L}_{\mathbb{C}}(SU(2))$ can be written as:

$$X = X_H + X_+E_+ + X_-E_- = \begin{pmatrix} X_H & X_+ \\ X_- & -X_H \end{pmatrix}.$$

To recover the real Lie algebra, we note we can still use this expansion, we just need to impose the conditions $X_+ = X_-^*$.

Theorem: The Cartan-Weyl basis has commutators:

$$[H, E_\pm] = \pm 2E_\pm, \quad [E_+, E_-] = H.$$

Proof: Tedious checking. □

Theorem: ad_H is diagonalisable.

Proof: Simply note that $\text{ad}_H(E_\pm) = \pm 2E_\pm$ and $\text{ad}_H(H) = 0$. So the Cartan-Weyl basis diagonalises ad_H. □

Definition: The eigenvalues of ad_H are called the *roots* of the Lie algebra.

Now consider representations R of $\mathcal{L}_{\mathbb{C}}(SU(2))$. **We assume that $R(H)$ is diagonalisable.**

Definition: The eigenvectors of $R(H)$ obey $R(H)v_\lambda = \lambda v_\lambda$. We call the eigenvalues $\{\lambda\}$ the *weights* of the representation.

Note: Roots $\equiv$ weights of adjoint representation.

Theorem: Irreducible representations of $\mathcal{L}_{\mathbb{C}}(SU(2))$ are uniquely determined by a greatest weight $\Lambda \in \mathbb{Z}_{\geq 0}$. This weight is called the *highest weight*. The weight set for the representation R_Λ (i.e. the rep with highest weight Λ) is:

$$S_{R_\Lambda} = \{-\Lambda, -\Lambda + 2, ..., \Lambda - 2, \Lambda\},$$

with each weight appearing with multiplicity one.

Proof: First note:

$$R(H)R(E_\pm)v_\lambda = (\lambda \pm 2)R(E_\pm)v_\lambda.$$

For a finite-dimensional rep R, there can only be finitely many weights, so there must exist a weight $\Lambda \in \mathbb{C}$ such that

$$R(H)v_\Lambda = \Lambda v_\Lambda, \qquad R(E_+)v_\Lambda = 0.$$

Assuming R is irreducible, we must be able to apply $R(X)$'s to v_Λ to get all other basis vectors. Since $R(H)$, $R(E_+)$ don't help us, the remaining basis vectors must come from acting with $R(E_-)$. Define:

$$v_{\Lambda - 2n} = (R(E_-))^n v_\Lambda.$$

Acting with $R(H)$ on these new vectors gives us back the same new vectors, with eigenvalues

$$R(H)v_{\Lambda-2n} = (\Lambda - 2n)v_{\Lambda-2n}.$$

So these vectors are independent for all n, and we've produced the entire basis this way.

To finish, we need to consider acting with $R(E_+)$:

$$\begin{aligned} R(E_+)v_{\Lambda-2n} &= R(E_+)R(E_-)v_{\lambda-2n+2} \\ &= R(E_-)R(E_+)v_{\Lambda-2n+2} + (\Lambda - 2n + 2)v_{\Lambda-2n+2}. \end{aligned}$$

When $n = 1$, get $R(E_+)v_{\Lambda-2} = \Lambda v_\Lambda$. When $n = 2$, get $R(E_+)v_{\Lambda-4} = (2\Lambda - 2)v_{\Lambda-2}$. By induction, we find

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

where

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

Also $R(E_+)v_\Lambda = 0 \Rightarrow r_0 = 0$. This allows us to solve the recurrence relation for r_n to find

$$r_n = (\Lambda + 1 - n)n.$$

Finally, since R is finite dimensional, there is also a lowest weight, say $\Lambda - 2N$. We must have $R(E_-)v_{\Lambda-2N} = v_{\Lambda-2N-2} = 0$, which implies $r_{N+1} = 0$. Hence:

$$(\Lambda - N)(N + 1) = 0,$$

i.e. $\Lambda = N$. This produces the weight set given in the Theorem, shows that Λ is a non-negative integer, and determines the representation uniquely (since we've said how all elements act on a basis). □

Definition: Let R_Λ be the representation of $\mathcal{L}_{\mathbb{C}}(SU(2))$ with highest weight Λ.

Theorem: $\dim(R_\Lambda) = \Lambda + 1$.

Proof: Follows from classification, since $R(H)$ has $\Lambda + 1$ distinct eigenvalues, each of multiplicity one. □

Example: R_0 is the trivial representation, R_1 is the fundamental representation and R_2 is the adjoint representation.

4.6 Application: Reps of real $\mathcal{L}(SU(2))$

To get a representation of the real Lie algebra $\mathcal{L}(SU(2))$, we pass back to the generators:

$$T^a = -\frac{1}{2}i\sigma_a,$$

so that:

$$T^1 = \frac{1}{2i}(E_+ + E_-),\ \ T^2 = \frac{1}{2}(E_+ - E_-),\ \ T^3 = \frac{1}{2i},$$

Then for a given representation R of $\mathcal{L}_{\mathbb{C}}(SU(2))$, we have:

$$\begin{aligned} R(T^1) &= \frac{1}{2i}\left(R(E_+) + R(E_-)\right), \\ R(T^2) &= \frac{1}{2}\left(R(E_+) - R(E_-)\right), \\ R(T^3) &= \frac{1}{2i}R(H). \end{aligned}$$

For any $X \in \mathcal{L}(SU(2))$, we have $X = X_a T^a$ for $X_a \in \mathbb{R}$. The resulting representation of $\mathcal{L}(SU(2))$ acts as

$$R(X) = X_a R(T^a).$$

4.7 Application: Reps of $SU(2)$, $SO(3)$ from $\mathcal{L}_{\mathbb{C}}(SU(2))$

$SU(2)$ is connected, so for any $A \in SU(2)$, we can write

$$A = \mathrm{Exp}(X), \qquad \mathcal{L}(SU(2)).$$

Starting from the irrep R_Λ, we can define the representation:

Definition: $D_\Lambda(A) = \mathrm{Exp}(R_\Lambda(X))$.

This is a valid rep of $SU(2)$. For a rep of $SO(3) \cong SU(2)/\mathbb{Z}_2$, we must also have $D_\Lambda(-I) = D_\Lambda(+I)$, since then

$$D_\Lambda(-A) = D_\Lambda(+A),$$

for all $A \in SU(2)$.

Theorem: If Λ is even, D_Λ is a rep of $SO(3)$ and $SU(2)$. If Λ is odd, it is only a rep of $SU(2)$.

Proof: Notice

$$-I = \mathrm{Exp}(i\pi H) = \begin{pmatrix} e^{i\pi} & 0 \\ 0 & e^{-i\pi} \end{pmatrix}.$$

Hence $D_\Lambda(-I) = \mathrm{Exp}(i\pi R_\Lambda(H))$. Now $R_\Lambda(H)$ has eigenvalues $\{\lambda\}$ in the weight set:

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

so $D_\Lambda(-I)$ has eigenvalues $\exp(i\pi\lambda) = (-1)^\lambda = (-1)^\Lambda$ (last equality since all λ's differ by 2). To be a valid rep of $SO(3)$, we need $D_\Lambda(-I) = D_\Lambda(+I) = I$. So all eigenvalues must be 1. It follows that Λ must be even. □

Definition: If D_Λ is a rep of $SU(2)$ only, it is called a *spinor representation*.

4.8 New representations from old

Definition: If R is a rep of a real Lie algebra $\mathfrak{g}$, then the *conjugate representation* $\overline{R}$ is defined by:

$$\overline{R}(X) = R(X)^*.$$

Definition: Let R_1, R_2 be representations, with representation spaces V_1, V_2. Define their *direct sum* to be the representation $R_1 \oplus R_2$ acting on $V_1 \oplus V_2$ by:

$$(R_1 \oplus R_2)(X)(v_1 \oplus v_2) = (R_1(X)v_1) \oplus (R_2(X)v_2).$$

The matrix of $(R_1 \oplus R_2)(X)$ has the block diagonal form:

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

Hence thus is a reducible representation. Also note $\dim(R_1 \oplus R_2) = \dim(R_1) + \dim(R_2)$.

Definition: Let R_1, R_2 be representations, with representation spaces V_1, V_2. Define their *tensor product* to be the representation $R_1 \otimes R_2$ acting on $V_1 \otimes V_2$ by:

$$(R_1 \otimes R_2)(X) = R_1(X) \otimes I_2 + I_1 \otimes R_2(X),$$

where I_1 and I_2 are the identity maps on V_1 and V_2 respectively.

Theorem: This indeed constitutes a representation.

Proof: Only non-trivial property to check is that commutators map to commutators:

$$[R_1 \otimes R_2(X), R_1 \otimes R_2(Y)]$$
$$= [R_1(X) \otimes I_2 + I_1 \otimes R_2(X), R_1(Y) \otimes I_2 + I_1 \otimes R_2(Y)]$$

$$= R_1(X)R_1(Y) \otimes I_2 + R_1(X) \otimes R_2(Y) + R_1(Y) \otimes R_2(X)$$
$$+I_1 \otimes R_2(X)R_2(Y) - R_1(Y)R_1(X) \otimes I_2 - R_1(Y) \otimes R_2(X)$$
$$-R_1(X) \otimes R_2(Y) - I_1 \otimes R_2(Y)R_2(X)$$

$$= R_1 \otimes R_2([X, Y]),$$

since $R_i([X,Y]) = R_i(X)R_i(Y) - R_i(Y)R_i(X)$. □

Definition: A *fully reducible* representation can be written as the direct sum of irreps. For a fully reducible representation, there exists a basis in which the matrix of the representation is:

$$R(X) = \text{diag}(R_1(X), R_2(X), ..., R_n(X)),$$

i.e. it has block-diagonal form, where the R_i are irreps.

Theorem: If R_i are irreps of a simple Lie algebra $\mathfrak{g}$, then the tensor product is fully reducible. That is,

$$R_1 \otimes R_2 \otimes ... \otimes R_m = \tilde{R}_1 \oplus \tilde{R}_2 \oplus ... \oplus \tilde{R}_{\tilde{m}},$$

for some irreps $\tilde{R}_j$, some $\tilde{m}$.

Proof: Too difficult for this course. □

4.9 The Clebsch-Gordan formula

Let R_M and R_N be irreps of $\mathcal{L}_{\mathbb{C}}(SU(2))$ as we found earlier, with highest weights M and N, and representation spaces V_M and V_N.

Theorem (Clebsch-Gordan): We have:

$$R_M \otimes R_N = R_{|N-M|} \oplus R_{|N-M|+2} \oplus ... \oplus R_{N+M}.$$

Proof: We know that $R_M \otimes R_N$ is fully reducible, so is the direct sum of irreps of $\mathcal{L}_{\mathbb{C}}(SU(2))$. Hence:

$$R_M \otimes R_N = \bigoplus_{\Lambda \in \mathbb{Z}_{\geq 0}} \mathcal{L}^{\Lambda}_{M,N} R_{\Lambda}.$$

The coefficients $\mathcal{L}^{\Lambda}_{M,N}$ denote how many times the summand R_Λ is repeated on the RHS. They are called *Littlewood-Richardson coefficients*.

To determine these coefficients, we just recall: **the weights of the LHS and RHS must be the same.** On the LHS, for eigenvectors v_λ, $v'_{\lambda'}$ of $R_M(H)$ and $R_N(H)$ respectively, we have:

$$(R_M \otimes R_N)(H)(v_\lambda \otimes v'_{\lambda'})$$
$$= (R_M(H)v_\lambda) \otimes v'_{\lambda'} + v_\lambda \otimes (R_N(H)v'_{\lambda'}) = (\lambda + \lambda')(v_\lambda \otimes v'_{\lambda'}).$$

So the weight set of $R_M \otimes R_N$ is:

$$S_{M,N} = \{\lambda + \lambda' : \lambda \in S_M, \lambda' \in S_N\}.$$

WLOG, let $M > N$. Let $\lambda = N - 2n$ and $\lambda = M - 2m$ be general weights in S_M, S_N so that:

$$\lambda + \lambda' = N + M - 2(n + m).$$

Then the multiplicities in this weight set are clearly:

- $N + M$ has multiplicity 1 ($n = m = 0$), $N + M - 2$ has multiplicity 2 ($n = 1, m = 0$ or $n = 0, m = 1$), $N + M - 4$ has multiplicity 3, ... etc
- $N + M - 2N$ has multiplicity $N + 1$ ($n = N, m = 0$, or ..., or $n = 0, m = N$). Now n is at its maximum though, so if we want to reduce further, we'll have to increase m.
- $N + M - 2(N + 1)$ has multiplicity $N + 1$ again, ($n = N, m = 1$, or ..., or $n = 0, m = N + 1$), etc.

- $N+M-2M$ has multiplicity $N+1$ ($n=N, m=M-N$, or ..., or $n=0, m=M$). Afterwards, the multiplicities decrease again, since we have reached the maximum m value.
- $N+M-2(M+1)$ has multiplicity N, $N+M-2(M+2)$ has multiplicity $N-1$, etc. $N-M-2(M+N)$ has multiplicity 1.

The highest weight on the LHS is $N+M$, of multiplicity 1. So we must need exactly one summand R_{N+M} on the RHS (any higher, would get different highest weight, any lower, wouldn't get highest weight).

Subtracting off the weight set of R_{N+M} from the list in $S_{M,N}$ leaves us with highest weight $N+M-2N$, now of multiplicity 1 since we've used it in R_{N+M}. So get a single summand of $R_{N+M-2N} = R_{M-N}$ on the RHS. Continuing in this fashion, we get Clebsch-Gordan. □

Example: Let $N = M = 1$. Recall $S_1 = \{1, -1\}$. So

$$S_{1,1} = \{-2, 0, 0, 2\} = S_0 + S_2.$$

Hence $R_1 \otimes R_1 = R_2 \oplus R_0$.

---

5 The Killing form

5.1 Definitions

Definition: An *inner product* on a vector space is a symmetric, bilinear map

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

An inner product is *non-degenerate* if for all $v \in V$, there is a $w \in V$ such that $i(v, w) \neq 0$.

Definition: The *Killing form* is the map $\kappa : \mathfrak{g} \times \mathfrak{g} \to \mathbb{F}$:

$$\kappa(X, Y) = \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y).$$

Theorem: κ is an inner product.

Proof: κ is symmetric, by cyclicity of trace. κ is bilinear; suffices to show linear in first argument by symmetry. Linearity follows from ad_X linear in X, and tr linear. □

5.2 Matrix representations

Theorem: The matrix representation of κ is

$$\kappa^{ab} = f^{ad}{}_c f^{bc}{}_d.$$

Proof: Want the trace of $[X, [Y, \cdot]]$. Consider first $[X, [Y, Z]]$. Let $X = X_a T^a$, $Y = Y_a T^a$, $Z = Z_a T^a$. Then:

$$[X, [Y, Z]] = X_a Y_b Z_c f^{ad}{}_e f^{bc}{}_d T^e = \underbrace{M(X, Y)^c{}_e Z_c}_{\text{components}} T^e.$$

Here, we've expressed this in its final form, since $(M_{ij} x_j)\mathbf{x}_i = (M\mathbf{x})_i \mathbf{x}_i$, where x_j are components and $\mathbf{x}_i$ are the basis vectors. Thus the matrix of $[X, [Y, \cdot]]$ is:

$$M(X, Y)^c{}_e = X_a Y_b f^{ad}{}_e f^{bc}{}_d.$$

Then

$$\kappa(X, Y) = \mathrm{tr}(M(X, Y)) = f^{ad}{}_c f^{bc}{}_d X_a Y_b,$$

which gives the matrix as $\kappa^{ab} = f^{ad}{}_c f^{bc}{}_d$. □

5.3 Properties of the Killing form

Theorem (Invariance): The Killing form is *invariant* under the adjoint action:

$$\kappa([Z, X], Y) + \kappa(X, [Z, Y]) = 0.$$

Proof: By definition of adjoint map, we have:

$$\mathrm{ad}_{[Z,X]} = \mathrm{ad}_X \circ \mathrm{ad}_X - \mathrm{ad}_X \circ \mathrm{ad}_Z.$$

Hence:

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

Sum, use cyclicity, then done. □

Theorem: If $\mathfrak{g}$ is simple, κ is the unique invariant inner product on $\mathfrak{g}$ up to an overall scalar multiple.

Proof: Too difficult for this course. □

Definition: A Lie algebra is of *compact type* if there exists a basis $\{T^a\}$ in which the Killing form has matrix $\kappa^{ab} = -\kappa\delta^{ab}$, for some constant $\kappa > 0$.

Then by Sylvester's Law of Inertia, κ has signature $(-, -, ..., -)$ in all bases.

5.4 Complexification and real forms

Definition: Let $\mathfrak{g}$ be a real Lie algebra with basis $\{T^a\}$ in which the structure constants are real. Then

$$\mathfrak{g} = \mathrm{span}_{\mathbb{R}}\{T^a\}.$$

Define the *complexification* of $\mathfrak{g}$ by:

$$\mathfrak{g}_{\mathbb{C}} = \mathrm{span}_{\mathbb{C}}\{T^a\}.$$

Definition: Let $\mathfrak{g}_{\mathbb{C}}$ be a complex Lie algebra. Then $\mathfrak{g}$ is a *real form* of $\mathfrak{g}_{\mathbb{C}}$ if $\mathfrak{g}$ has complexification $\mathfrak{g}_{\mathbb{C}}$.

Real forms of a complex Lie algebra are not unique:

Example: The complexified Lie algebra of $\mathcal{L}(SU(2))$ is $\mathcal{L}_{\mathbb{C}}(SU(2))$, the traceless 2×2 complex matrices, as we saw earlier. But $\mathcal{L}(SL(2,\mathbb{R}))$ consists of only traceless 2×2 matrices, and so its complexification is also $\mathcal{L}_{\mathbb{C}}(SU(2))$.

To show these aren't the same, we evaluate their Killing forms. The structure constants of $\mathcal{L}(SU(2))$ are ${f^{ab}}_c = \epsilon_{abc}$, and hence

$$\kappa^{ab} = \epsilon_{adc}\epsilon_{bcd} = -\epsilon_{adc}\epsilon_{bdc} = -\delta_{ab}.$$

So $\mathcal{L}(SU(2))$ is of compact type.

We can find the Killing form of $\mathcal{L}(SL(2,\mathbb{R}))$ similarly and see it is not of compact type; so these are not equivalent real forms (i.e. not isomorphic as Lie algebras).

5.5 Cartan's theorem

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

Theorem: Any semi-simple Lie algebra has a real form of compact type.

Proof: Too hard for this course. □

Theorem: A finite dimensional semi-simple Lie algebra can be written as the direct sum of a finite number of simple Lie algebras (by this, we mean the underlying vector space is $\mathfrak{g} \oplus \mathfrak{f} \oplus ...$, with $[\mathfrak{g}, \mathfrak{f}] = 0$ for all summands).

Proof: Let $\mathfrak{g}$ be of compact type. Let $\mathfrak{i}$ be an ideal of $\mathfrak{g}$. Let $\mathfrak{i}_\perp$ be its orthogonal complement with respect to the Killing form. Let $X \in \mathfrak{i}$, $Y \in \mathfrak{g}$, $Z \in \mathfrak{i}_\perp$. Then by invariance:

$$\kappa(X, [Y, Z]) = -\kappa([Y, X], Z) = 0$$

since $[Y, X] \in \mathfrak{i}$, as $\mathfrak{i}$ an ideal. So $[Y, Z] \in \mathfrak{i}_\perp$, and $\mathfrak{i}_\perp$ is thus an ideal. Since κ is an inner product, standard linear algebra then says $\mathfrak{g} = \mathfrak{i} \oplus \mathfrak{i}_\perp$.

The summands commute because of the following:

$$\kappa([X, Z], Y) = -\kappa(Z, [X, Y]) = 0,$$

since $[X, Y] \in \mathfrak{i}$. Since of compact type, Killing form is non-degenerate, so $[X, Z] = 0$ (since holds for all $Y \in \mathfrak{g}$).

Finally, note that $\mathfrak{i}, \mathfrak{i}_\perp$ are closed under taking the bracket so are Lie algebras, and are compact because they inherit a restricted Killing form from $\mathfrak{g}$.

Iterating we see any Lie algebra of compact type is the direct sum of simple Lie algebras of compact type.

Now need to prove the result for a complex semi-simple Lie algebra $\mathfrak{g}_{\mathbb{C}}$. We know from above this has a real form of compact type, say $\mathfrak{g}$. Then from above:

$$\mathfrak{g} = \sum_i \mathfrak{g}_i,$$

for $\mathfrak{g}_i$ simple, compact Lie algebras. Then just take complexification of both sides (obvious that sum of complexifications is complexification of sum). □

Cartan's Theorem: The Killing form κ is non-degenerate if and only if $\mathfrak{g}$ is semi-simple.

Proof: Suppose $\mathfrak{g}$ is not semi-simple. Then $\mathfrak{g}$ has an Abelian ideal $\mathfrak{j}$. Let $\dim(\mathfrak{g}) = D$, and $\dim(\mathfrak{j}) = d$. Choose a basis

$$B = \{T^a\} = \{T^i : i = 1, ..., d\} \cup \{T^\alpha : \alpha = d+1, ..., D\}.$$

where $\{T^i\}$ span $\mathfrak{j}$. As $\mathfrak{j}$ is Abelian, we have $[T^i, T^j] = 0$. As $\mathfrak{j}$ is an ideal, we also have $[T^\alpha, T^j] = {f^{\alpha j}}_k T^k \in \mathfrak{j}$, thus:

$${f^{ij}}_a = 0 \quad -(1), \qquad {f^{\alpha j}}_\beta = 0 \quad -(2).$$

(Early Latin indices for $a, b, c = 1, ..., D$, late Latin indices for $i, j = 1, ..., d$, and Greek indices for $\alpha = d+1, ..., D$.) For $X = X_a T^a$ and $Y = Y_i T^i$, we have:

$$\kappa(X, Y) = \kappa^{ai} X_a Y_i.$$

Here, $\kappa^{ai} = {f^{ab}}_c {f^{ic}}_b$. If $c \leq d+1$, this vanishes by (1). So $\kappa^{ai} = {f^{ab}}_\alpha {f^{i\alpha}}_b$. By (2) can replace b by j: $\kappa^{ai} = {f^{aj}}_\alpha {f^{i\alpha}}_j$.

But ${f^{aj}}_\alpha = 0$ (for $a = 1, ..., d$, get 0 by (2), and $a = d+1, ..., D$ get 0 by (1)). So the Killing form is degenerate: $\kappa(X, Y) = 0$ for all $X \in \mathfrak{g}$ and $Y \in \mathfrak{j}$.

The converse is non-examinable. □

6 The Cartan classification

6.1 Cartan subalgebras

Definition: Let $\mathfrak{g}$ be a Lie algebra. We say $X \in \mathfrak{g}$ is *ad-diagonalisable* if ad_X is diagonalisable.

Definition: A *Cartan subalgebra* $\mathfrak{h} \trianglelefteq \mathfrak{g}$ is a maximal Abelian subalgebra containing only ad-diagonalisable elements. That is:

(i) $H \in \mathfrak{h}$ implies H is ad-diagonalisable;

(ii) $H, H' \in \mathfrak{h}$ implies $[H, H'] = 0$;

(iii) If $X \in \mathfrak{g}$ is ad-diagonalisable and $[X, H] = 0$ for all $H \in \mathfrak{h}$, then $X \in \mathfrak{h}$ (this is called *maximality*).

Theorem: All possible choices of Cartan subalgebra have the same dimension.

Proof: Not required. □

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

Example: $\mathfrak{g} = \mathcal{L}_{\mathbb{C}}(SU(2))$ has Cartan subalgebra $\mathfrak{h} = \mathrm{span}_{\mathbb{C}}\{H\}$, where H is the Cartan element as previously defined. This is because

$$[H, X_H H + X_+ E_+ + X_- E_-] = 2X_+ E_+ - 2X_- E_- = 0$$

if and only if $X_+ = X_- = 0$. So nothing outside $\mathfrak{h}$ commutes with H. Also recall H is ad-diagonalisable, and clearly $\mathfrak{h}$ is Abelian. Hence rank of $\mathcal{L}_{\mathbb{C}}(SU(2))$ is 1.

Example: The Lie algebra $\mathfrak{g} = \mathcal{L}_{\mathbb{C}}(SU(n)) = \{n \times n \text{ traceless matrices}\}$ has a natural basis for its Cartan subalgebra:

$$(H^i)_{\alpha\beta} = \delta_{\alpha i}\delta_{\beta i} - \delta_{\alpha i+1}\delta_{\beta i+1},$$

i.e. $+1$ at ith position on diagonal, and -1 at $(i+1)$th position on diagonal. This can be shown to be the Cartan subalgebra, hence the rank is $n-1$.

6.2 Step operators

Consider a general Lie algebra $\mathfrak{g}$ and Cartan subalgebra $\mathfrak{h}$. Introduce a basis H^i, $i = 1, ..., r$ for the Cartan subalgebra.

Theorem: The r linear maps ad_{H^i} are simultaneously diagonalisable.

Proof: $[H^i, H^j] = 0$ for all i, j, so since ad is a representation:

$$0 = \mathrm{ad}_{[H^i, H^j]} = [\mathrm{ad}_{H^i}, \mathrm{ad}_{H^j}].$$

So all commute, so all simultaneously diagonalisable. □

Definition: The simultaneous eigenvectors of $\{\mathrm{ad}_{H^i}\}$ can be split into two types:

- Let V be an eigenvector with $\mathrm{ad}_{H^i}(V) = 0$ for all i. Then V was in the Cartan subalgebra after all, since commutes with all elements in subalgebra.
- Otherwise, there exists a H^i for which ad_{H^i} gives a non-zero eigenvalue when acting on the eigenvector. Write the eigenvector as E^α. Then

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

  where α^i is the eigenvalue of E^α when ad_{H^i} acts. We know that not all α^i are zero.

We call E^α a *step operator* of the Lie algebra, and the vector $\alpha = (\alpha^1, ..., \alpha^r)$ a *root* of the Lie algebra.

Theorem: The roots are elements of the dual of the Cartan subalgebra, $\mathfrak{h}^*$.

Proof: Consider $H = \rho_i H^i \in \mathfrak{h}$. Then $\mathrm{ad}_H(E^\alpha) = \rho_i \alpha^i E^\alpha$. So α defines a linear map $\alpha(H) = \rho_i \alpha^i$, from $\mathfrak{h}$ to $\mathbb{C}$, i.e. is a dual vector. □

Theorem: The roots are *non-degenerate*. That is, the set of roots Φ consists of $d - r$ distinct element of $\mathfrak{h}^*$.

Proof: Too hard for this course. □

Since all roots are non-degenerate, the E^α are linearly independent. So:

Definition: The *Cartan-Weyl basis* of $\mathfrak{g}$ is:

$$B = \{H^i : i = 1, ..., r\} \cup \{E^\alpha : \alpha \in \Phi\}.$$

Example: Consider $\mathcal{L}_{\mathbb{C}}(SO(2n))$. This is the set of $2n \times 2n$ antisymmetric matrices, which has a natural basis:

$$(T_{ij})_{\alpha\beta} = \delta_{\alpha i}\delta_{j\beta} - \delta_{i\beta}\delta_{\alpha j},$$

where $j > i$ to avoid redundancy. Then a basis for the CSA is: $H_I = T_{(2I-1)(2I)}$, for $I = 1, ..., n$. Let's prove it.

Being Abelian is obvious. The best way to show maximality is to commute with a general antisymmetric matrix $X = X^{ij}T_{ij}$, and the best way to show ad-diagonalisability is to exhibit the step operators, which in this case are:

$$F^{\pm}_{IJ} = T_{(2I-1)(2J-1)} - T_{(2I)(2J)} \pm i\left(T_{(2I-1)(2J)} + T_{(2I)(2J-1)}\right),$$
$$G^{\pm}_{IJ} = T_{(2I-1)(2J)} - T_{(2I)(2J-1)} \pm i\left(T_{(2I-1)(2J-1)} + T_{(2I)(2J)}\right),$$

for $I, J = 1, 2...n$ and $I < J$. For $\mathcal{L}_{\mathbb{C}}(SO(2n+1))$, everything's the same, *except* there's extra step operators:

$$E^{\pm}_I = T_{(2I-1)(2n+1)} \pm iT_{(2I)(2n+1)},$$

for $I = 1, ..., n$. Everything's given explicitly for $n = 2$ later in these notes.

6.3 Killing form in the Cartan-Weyl basis

Assume now we are working with a simple Lie algebra $\mathfrak{g}$. Then by Cartan's theorem, the Killing form is non-degenerate.

Theorem: The Killing form for a simple Lie algebra $\mathfrak{g}$ with the Cartan-Weyl basis obeys:

(i) $\kappa(H, E^\alpha) = 0$ for all $H \in \mathfrak{h}$;

(ii) $\kappa(E^\alpha, E^\beta) = 0$ for $\alpha + \beta \neq 0$;

(iii) For all $H \in \mathfrak{h}$, there exists $H' \in \mathfrak{h}$ with $\kappa(H, H') \neq 0$;

(iv) If α is a root, then $-\alpha$ is a root, and $\kappa(E^\alpha, E^{-\alpha}) \neq 0$.

Proof: (i) Let $H' \in \mathfrak{h}$. Then:

$$\begin{aligned}\alpha(H')\kappa(H, E^\alpha) &= \kappa(H, [H', E^\alpha]) \\ &= -\kappa([H', H], E^\alpha) = -\kappa(0, E^\alpha) = 0.\end{aligned}$$

But $\alpha(H') \not\equiv 0$, so there exists H' such that $\alpha(H') \neq 0$. Thus $\kappa(H, E^\alpha) = 0$.

(ii) Let $H' \in \mathfrak{h}$. Then

$$\begin{aligned}&(\alpha(H') + \beta(H'))\kappa(E^\alpha, E^\beta) \\ &= \kappa([H', E^\alpha], E^\beta) + \kappa(E^\alpha, [H', E^\beta]) = 0,\end{aligned}$$

by invariance. But assumed $\alpha + \beta \neq 0$, so $\kappa(E^\alpha, E^\beta) = 0$.

(iii) Assume $\kappa(H, H') = 0$ for all $H' \in \mathfrak{h}$. From (i), $\kappa(H, E^\alpha) = 0$ for all $\alpha \in \Phi$. Since $\{H^i, E^\alpha\}$ form a basis, get $\kappa(H, X) = 0$ for all $X \in \mathfrak{g}$, which is a contradiction since κ is non-degenerate. So there exists H' with $\kappa(H, H') \neq 0$.

(iv) From (i), $\kappa(E^\alpha, H) = 0$ for all $H \in \mathfrak{h}$, and from (ii) $\kappa(E^\alpha, E^\beta) = 0$ for all $\alpha, \beta \in \Phi$ with $\alpha \neq -\beta$. If $-\alpha$ is not a root then, E^α is orthogonal to all other basis vectors - contradiction, as κ non-degenerate. If $\kappa(E^\alpha, E^{-\alpha}) = 0$, we similarly get κ degenerate, contradiction. $\square$

Upshot of above theorem is:

Theorem: For $\mathfrak{g}$ simple, the Killing form is a non-degenerate inner product on the Cartan subalgebra.

Proof: Follows from (iii) above. $\square$

In particular, the $r \times r$ matrix κ^{ij} of the restricted Killing form defined by $\kappa(H, H') = \kappa^{ij}\rho_i\rho'_j$ is invertible. We use the inverse to define an inner product on $\mathfrak{h}^*$:

Definition: The inner product of roots $\alpha, \beta \in \Phi$ is defined by:

$$(\alpha, \beta) = (\kappa^{-1})_{ij}\alpha^i\beta^j.$$

6.4 Algebra in the Cartan-Weyl basis

So far, we have: $[H^i, H^j] = 0$, and $[H^i, E^\alpha] = \alpha^i E^\alpha$. Remains to evaluate $[E^\alpha, E^\beta]$.

Theorem: Let H^α solve $\kappa(H^\alpha, H) = \alpha(H)$ (exists by (iii) above). Then for undetermined constants $N_{\alpha,\beta}$:

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

Proof: By the Jacobi identity,

$$\begin{aligned}[H^i, [E^\alpha, E^\beta]] &= -[E^\alpha, [E^\beta, H^i]] - [E^\beta, [H^i, E^\alpha]] \\ &= (\alpha^i + \beta^i)[E^\alpha, E^\beta].\end{aligned}$$

Hence for $\alpha+\beta \neq 0$, $[E^\alpha, E^\beta] \propto E^{\alpha+\beta}$ (by non-degeneracy of roots). If $\alpha + \beta \notin \Phi$, we need the constant of proportionality to be 0, else $E^{\alpha+\beta}$ would be an eigenvector.

If $\alpha + \beta = 0$, note that

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

Now (iv) above says $\kappa(E^\alpha, E^{-\alpha}) \neq 0$. So have:

$$\alpha(H) = \kappa\left(\frac{[E^\alpha, E^{-\alpha}]}{\kappa(E^\alpha, E^{-\alpha})}, H\right).$$

Define $H^\alpha = [E^\alpha, E^{-\alpha}]/\kappa(E^\alpha, E^{-\alpha})$ to get result in Theorem.

We can also construct H^α explicitly using the matrix representation. Write $H^\alpha = \rho_i^\alpha H^i$ and $H = \rho_j H^j$. Then need $\kappa^{ij}\rho_i^\alpha\rho_j = \alpha^j\rho_j$. Inverting:

$$\rho_i^\alpha = (\kappa^{-1})_{ij}\alpha^j \qquad \Rightarrow \qquad H^\alpha = (\kappa^{-1})_{ij}\alpha^j H^i. \quad \square$$

We summarise this result as follows:

Algebra in the Cartan-Weyl basis:

$$[H^i, H^j] = 0$$
$$[H^i, E^\alpha] = \alpha^i E^\alpha$$
$$[E^\alpha, E^\beta] = \begin{cases} N_{\alpha,\beta}E^{\alpha+\beta}, \text{ if } \alpha + \beta \in \Phi, \\ \kappa(E^\alpha, E^{-\alpha})H^\alpha, \text{ if } \alpha + \beta = 0, \\ 0 \text{ otherwise,} \end{cases}$$

where the $N_{\alpha,\beta}$ are undetermined constants and H^α solves $\kappa(H^\alpha, H) = \alpha(H)$, i.e.

$$H^\alpha = (\kappa^{-1})_{ij}\alpha^j H^i.$$

We can now normalise the basis to make the algebra more simple. Choose to work with H^α instead. Then want:

Theorem: $[H^\alpha, E^\beta] = (\alpha, \beta)E^\beta$.

Proof: Note:

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

then use symmetry of Killing form: $(\alpha, \beta) = (\beta, \alpha)$. □

We now normalise the basis:

$$h^\alpha = \frac{2}{(\alpha, \alpha)}H^\alpha, \quad e^\alpha = \frac{\sqrt{2}}{((\alpha, \alpha)\kappa(E^\alpha, E^{-\alpha}))^{1/2}}E^\alpha.$$

Here, we have assumed $(\alpha, \alpha) \neq 0$ (this is in fact true). Thus the algebra becomes:

Algebra in the normalised Cartan-Weyl basis:

$$[h^\alpha, h^\beta] = 0$$
$$[h^\alpha, e^\beta] = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}e^\beta$$
$$[e^\alpha, e^\beta] = \begin{cases} n_{\alpha,\beta}e^{\alpha+\beta}, & \text{if } \alpha + \beta \in \Phi, \\ h^\alpha, & \text{if } \alpha + \beta = 0, \\ 0 & \text{otherwise,} \end{cases}$$

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

This representation makes clear the relation of the *rank* and the *dimension* of a Lie algebra: $\dim(\mathfrak{g}) = |\Phi| + \mathrm{rank}(\mathfrak{g})$, where $|\Phi|$ is the root set.

6.5 $\mathcal{L}_{\mathbb{C}}(SU(2))$ subalgebras

Recall for all $\alpha \in \Phi$, we have $-\alpha \in \Phi$. Hence:

Theorem: For each pair $\pm\alpha \in \Phi$, there exists a $\mathcal{L}_{\mathbb{C}}(SU(2))$ subalgebra of $\mathfrak{g}$ with basis $\{h^\alpha, e^\alpha, e^{-\alpha}\}$, with commutators: $[h^\alpha, e^{\pm\alpha}] = \pm 2e^{\pm\alpha}, \quad [e^\alpha, e^{-\alpha}] = h^\alpha$.

Proof: Clear from above. □

Definition: The $\mathcal{L}_{\mathbb{C}}(SU(2))$ subalgebra corresponding to the roots $\pm\alpha$ is called $\mathfrak{sl}(2)_\alpha$.

Definition: Let $\alpha, \beta \in \Phi$. We define the α *root string passing through* β (for $\alpha \neq \beta$) to be the set of roots:

$$S_{\alpha,\beta} = \{\beta + \rho\alpha \in \Phi : \rho \in \mathbb{Z}\}.$$

Define also the vector subspace:

$$V_{\alpha,\beta} = \mathrm{span}_{\mathbb{C}}\{e^{\beta+\rho\alpha} : \beta + \rho\alpha \in S_{\alpha,\beta}\}.$$

Theorem (Quantisation Condition): Let $S_{\alpha,\beta}$ be a root string with parameter ρ. Then for $\rho = n_-$ to $\rho = n_+$, all values of $\beta + \rho\alpha$ are roots (the string is 'unbroken'), where n_+ and n_- must satisfy:

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = -(n_+ + n_-) \in \mathbb{Z}.$$

Proof: Consider the adjoint representation of $\mathfrak{sl}(2)_\alpha$ on $V_{\alpha,\beta}$. We have:

$$[h^\alpha, e^{\beta+\rho\alpha}] = \frac{2(\alpha, \beta + \rho\alpha)}{(\alpha, \alpha)}e^{\beta+\rho\alpha} = \left(\frac{2(\alpha, \beta)}{(\alpha, \alpha)} + 2\rho\right)e^{\beta+\rho\alpha} \in V_{\alpha,\beta},$$

$$[e^{\pm\alpha}, e^{\beta+\rho\alpha}] \propto \begin{cases} e^{\beta+(\rho\pm1)\alpha} \in V_{\alpha,\beta}, & \text{if } \beta + (\rho \pm 1)\alpha \in \Phi, \\ 0 \in V_{\alpha,\beta}, & \text{otherwise.} \end{cases}$$

So $V_{\alpha,\beta}$ is a valid representation space. We have also calculated the weights above. They are:

$$S_R = \left\{\frac{2(\alpha, \beta)}{(\alpha, \alpha)} + 2\rho : \beta + \rho\alpha \in \Phi\right\}.$$

Each of these weights has multiplicity one, and the representation space is finite, so the representation R is irreducible and finite-dimensional.

But we've classified these: $R = R_\Lambda$ for some highest weight $\Lambda \in \mathbb{Z}_{\geq 0}$. It follows the weight set is also equal to

$$S'_R = \{-\Lambda, -\Lambda + 2, ..., \Lambda\} = S_R.$$

So for some $\rho = n_-$ and $\rho = n_+$ we have:

$$-\Lambda = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} + 2n_-, \quad \Lambda = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} + 2n_+.$$

Since all weights in the original S'_R must occur, and all differ by 2, and all possible weights in the original S_R differ by 2, all possible roots $\beta + \rho\alpha$ *must* occur between $n_- \leq \rho \leq n_+$ for the weight sets to be equal.

Finally, adding the two conditions on Λ, $n_\pm$, we get the quantisation condition. □

Two slogans:

1. For any two roots, α, β, we have *quantisation*:

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

2. The α root string passing through β, $S_{\alpha,\beta}$ is equal to (for some n_-, n_+ as yet undetermined):

$$S_{\alpha,\beta} = \{\beta + n\alpha : n \in \mathbb{Z}, n_- \leq n \leq n_+\}.$$

All of its members are roots of the Lie algebra.

6.6 Real geometry of the roots

In this section, we aim to show that the inner product of roots is a real Euclidean inner product.

Theorem: The matrix of the Killing form is given by:

$$\kappa^{ij} = \sum_{\delta \in \Phi} \delta^i \delta^j.$$

Proof: Work in the un-normalised Cartan-Weyl basis. Here, we have $[H^i, E^\delta] = \delta^i E^\delta$ for any $\delta \in \Phi$. Hence:

$$\kappa^{ij} = \kappa(H^i, H^j) = \mathrm{tr}\left(\mathrm{ad}_{H^i} \circ \mathrm{ad}_{H^j}\right).$$

Diagonalise both ad_{H^i} and ad_{H^j}, multiply their matrices together, then take the trace to see the result. □

Definition: We define a root α_i with lowered indices by: $\alpha_i = (\kappa^{-1})_{ij}\alpha^j$; that is, we use κ as a metric for raising and lowering indices.

Theorem: For any two roots, $(\alpha, \beta) \in \mathbb{R}$.

Proof: We have:

$$\begin{aligned}(\alpha, \beta) = \alpha^i \beta^j (\kappa^{-1})_{ij} = \alpha_i \beta_j \kappa^{ij} = \sum_{\delta \in \Phi} \alpha_i \delta^i \delta^j \beta_j \\ = \sum_{\delta \in \Phi} (\alpha, \delta)(\beta, \delta).\end{aligned}$$

Divide by $(\alpha, \alpha)(\beta, \beta)$, and multiply through by 4 to get:

$$\frac{2}{(\beta, \beta)} \cdot \underbrace{\frac{2(\alpha, \beta)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} = \underbrace{\sum_{\delta \in \Phi} \frac{2(\alpha, \delta)}{(\alpha, \alpha)} \frac{2(\beta, \delta)}{(\beta, \beta)}}_{\in \mathbb{Z}}.$$

If $(\alpha, \beta) = 0$, we're done. Otherwise, we can deduce $(\beta, \beta) \in \mathbb{R}\backslash\{0\}$. Then by $2(\alpha, \beta)/(\beta, \beta) \in \mathbb{Z}$, we have $(\alpha, \beta) \in \mathbb{R}$. □

Theorem: The roots Φ span $\mathfrak{h}^*$, the dual of the Cartan subalgebra.

Proof: Suppose not. Then there exists $\lambda \in \mathfrak{h}^*$ with no component in the direction of any of the roots α:

$$(\lambda, \alpha) = (\kappa^{-1})_{ij} \lambda^i \alpha^j = \kappa^{ij} \lambda_i \alpha_j = 0.$$

Then for $H_\lambda = \lambda_i H^i \in \mathfrak{h}$, obeying $[H_\lambda, H] = 0$ and $[H_\lambda, E^\alpha] = (\lambda, \alpha) E^\alpha = 0$. Hence $[H_\lambda, X] = 0$ for all X, so $\mathrm{span}_{\mathbb{C}}\{H_\lambda\}$ is a non-trivial ideal. Contradiction as $\mathfrak{g}$ is simple. □

Hence choose r roots $\{\alpha_{(i)}\}$ which provide a basis for $\mathfrak{h}^*$.

Definition: Define the *real subspace* of the Cartan subalgebra to be $\mathfrak{h}^*_{\mathbb{C}} = \mathrm{span}_{\mathbb{R}}\{\alpha_{(i)}\}$.

Theorem: $\mathfrak{h}^* = \mathfrak{h}^*_{\mathbb{R}}$.

Proof: Clearly, $\mathfrak{h}^*_{\mathbb{R}} \subseteq \mathfrak{h}^*$. Conversely, let $\beta \in \mathfrak{h}^*$. Then for some β^i:

$$\beta = \sum_{i=1}^r \beta^i \alpha_{(i)} \quad \Rightarrow \quad (\beta, \alpha_{(j)}) = \sum_{i=1}^r \beta^i (\alpha_{(i)}, \alpha_{(j)}).$$

Since $(\alpha_{(i)}, \alpha_{(j)})$ essentially defines a matrix, which is invertible by non-degeneracy of the inner product, we can invert to get $\beta^i \in \mathbb{R}$ (since all inner products real). □

Theorem: For all roots λ, we have $(\lambda, \lambda) \geq 0$ with equality if and only if $\lambda = 0$.

Proof: By the expression for the Killing form above, we have:

$$(\lambda, \lambda) = \sum_{\delta \in \Phi} \lambda_i \delta^i \lambda_j \delta^j = \sum_{\delta \in \Phi} (\lambda, \delta)^2 \geq 0,$$

since $(\lambda, \delta) \in \mathbb{R}$. There is equality iff $(\lambda, \delta) = 0$ for all $\delta \in \Phi$, hence $\lambda = 0$ since inner product non-degenerate. □

Real geometry: The roots $\alpha \in \Phi$ are in a real vector space $\mathfrak{h}^*_{\mathbb{R}} \cong \mathbb{R}^r$, of dimension r, equipped with a Euclidean inner product $(\cdot, \cdot)$ obeying:

(a) $(\alpha, \beta) \in \mathbb{R}$;

(b) $(\alpha, \alpha) \geq 0$, with equality iff $\alpha = 0$.

Definition: Since $(\alpha, \alpha) > 0$ for all roots, can define the *length* of a root $|\alpha| = (\alpha, \alpha)^{1/2}$. We can also define the *angle* ϕ between the roots α and β by $(\alpha, \beta) = |\alpha||\beta| \cos(\phi)$.

Theorem: The angle between the roots is one of

$$\left\{0, \frac{\pi}{2}, \pi, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{2\pi}{3}, \frac{3\pi}{4}, \frac{5\pi}{6}\right\}.$$

Proof: The angle is constrained by:

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 2\frac{|\beta|}{|\alpha|}\cos(\phi) \in \mathbb{Z}, \qquad \frac{2(\beta, \alpha)}{(\beta, \beta)} = 2\frac{|\alpha|}{|\beta|}\cos(\phi) \in \mathbb{Z}.$$

So multiplying these, need $4\cos^2(\phi) \in \mathbb{Z}$, so $4\cos^2(\phi) = 0, 1, 2, 3$ or 4. Hence $\cos(\phi) = \pm\sqrt{n}/2$, for $n \in \{0, 1, 2, 3, 4\}$. This gives the possibilities above. □

6.7 Simple roots

The roots live in $\mathfrak{h}^*_{\mathbb{R}} \cong \mathbb{R}^r$, and there are only finitely many of them. So we can pick a hyperplane splitting the roots in half (since there are only finitely many banned directions from the roots, and a continuous degree of freedom associated with the normal to the hyperplane).

Definition: For a choice of hyperplane splitting the roots Φ into Φ^+ and Φ^-, we call the roots in Φ^+ the *positive roots* and the roots in Φ^- the *negative roots*.

Theorem: (i) If $\alpha \in \Phi^+$, then $-\alpha \in \Phi^-$; (ii) If α, β are positive (negative) roots, and $\alpha + \beta$ is a root, then $\alpha + \beta$ is a positive (negative) root.

Proof: Obvious from hyperplane definition. □

Definition: A *simple root* is a positive root which *cannot* be written as a sum of two positive roots. We denote the set of simple roots by Φ_S.

Theorem (Simple root properties): Let α,β be simple roots throughout. The following hold.

(i) $\alpha - \beta$ is not a root.

(ii) The α string passing through β has length:
$$l_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)}.$$

(iii) $(\alpha,\beta) \leq 0$.

(iv) Any positive root $\gamma \in \Phi_+$ can be written as a linear combination of simple roots with positive integer coefficients:
$$\gamma = \sum_i c_i \alpha_{(i)}.$$

(v) Simple roots are linearly independent.

(vi) There are exactly r simple roots, $|\Phi_S| = r$, where r is the rank of $\mathfrak{g}$.

(vii) The simple roots are a basis for $\mathfrak{h}^*_{\mathbb{R}}$.

Proof: (i) If $\alpha - \beta$ is a root, then either $\alpha - \beta$ is a positive root, or $\beta - \alpha$ is a positive root. In the first case, $\alpha = \alpha - \beta + \beta$, so we have a contradiction. Similarly for latter case.

(ii) Recall the string is
$$S_{\alpha,\beta} = \{\beta + n\alpha : n \in \mathbb{Z}, n_- \leq n \leq n_+\},$$
with
$$(n_+ + n_-) = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}.$$
Since α, β are simple, $\beta - \alpha \notin \Phi$ by (i), so $n_- = 0$. Hence $n_+ = -2(\alpha,\beta)/(\alpha,\alpha)$, and the length of the string is $l_{\alpha,\beta} = n_+ + 1 = 1 - 2(\alpha,\beta)/(\alpha,\alpha)$.

(iii) In (ii), note $n_+ \geq n_- = 0$, so $n_+ \geq 0$. Hence $-2(\alpha,\beta)/(\alpha,\alpha) \geq 0$. Recall $(\alpha,\alpha) \geq 0$, and result follows.

(iv) Trivial if $\gamma \in \Phi_S$. If not, $\gamma = \beta_1 + \beta_2$, for $\beta_1, \beta_2 \in \Phi$. Iterate until get solely simple roots. Terminates since all coefficients are positive, and there are finitely many positive roots.

(v) Consider all vectors $\lambda \in \mathfrak{h}^*_{\mathbb{R}}$ of the form
$$\lambda = \sum_{i \in J} c_i \alpha_{(i)},$$
where J indexes the simple roots $\alpha_{(i)}$. For linear independence, we need $\lambda \neq 0$ always. Define:
$$\lambda_+ = \sum_{i \in J_+} c_i \alpha_{(i)}, \quad \lambda_- = -\sum_{i \in J_-} c_i \alpha_{(i)},$$
where $J_+ = \{i \in J : c_i > 0\}$ and $J_- = \{i \in J : c_i < 0\}$. Then $\lambda = \lambda_+ - \lambda_-$. So we have:
$$(\lambda,\lambda) = (\lambda_+,\lambda_+) + (\lambda_-,\lambda_-) - 2(\lambda_+,\lambda_-)$$
$$> -2(\lambda_+,\lambda_-) = 2 \sum_{i \in J_+} \sum_{j \in J_-} c_i c_j (\alpha_{(i)}, \alpha_{(j)}) > 0,$$
since $c_i c_j < 0$ in the sum and $(\alpha_{(i)}, \alpha_{(j)}) < 0$ as these are simple roots. Hence $(\lambda,\lambda) > 0$, and thus $\lambda \neq 0$.

(v) and (vi). From (iv) the simple roots span Φ^+. Simply taking negative of any root in Φ^+ generates all roots in Φ^-, hence simple roots span all of Φ. Simple roots are LI by (iv). So simple roots are a basis for $\mathfrak{h}^*_{\mathbb{R}}$ and there are hence exactly r of them (the dimension of $\mathfrak{h}^* = \mathfrak{h}^*_{\mathbb{R}}$). □

Definition: We encode all inner products between the roots in the $r \times r$ *Cartan matrix*:
$$A^{ij} := \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})}.$$

Note all of the entries are integers, but that A^{ij} is *not*, in general, symmetric.

6.8 The Chevalley basis

The Cartan matrix allows us to define the *Chevalley basis* of the Lie algebra.

Definition: Define $h^i = h^{\alpha_{(i)}}$ and $e^i_\pm = e^{\pm\alpha_{(i)}}$, where $\alpha_{(i)}$ are the simple roots, and h^α, e^α are the generators of the normalised Cartan-Weyl basis. The *Chevalley basis* is the basis of the Lie algebra generated by $\{h^i, e^i_\pm\}$.

Theorem: In the Chevalley basis, the commutators of the Lie algebra are:

$$[h^i, h^j] = 0, \quad [h^i, e^j_\pm] = \pm A^{ji} e^j_\pm, \quad [e^i_+, e^j_-] = \delta_{ij} h^i,$$

subject to the *Chevalley-Serre relations*:

$$(\mathrm{ad}_{e^i_\pm})^{1-A^{ji}} e^j_\pm = 0.$$

Proof: The first commutator is obvious, since h^i, h^j are in the Cartan subalgebra. The second is:

$$[h^i, e^j_\pm] = [h^{\alpha_{(i)}}, e^{\pm\alpha_{(j)}}] = \pm \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(i)}, \alpha_{(i)})} e^{\pm\alpha_{(j)}} = \pm A^{ji} e^j_\pm.$$

The third is: $[e^i_+, e^j_-]$

$$= [e^{\alpha_{(i)}}, e^{-\alpha_{(j)}}] = \begin{cases} n_{\alpha_{(i)}, \alpha_{(j)}} e^{\alpha_{(i)} - \alpha_{(j)}}, \text{ if } \alpha_{(i)} - \alpha_{(j)} \in \Phi, \\ h^{\alpha_{(i)}}, \text{ if } \alpha_{(i)} = \alpha_{(j)}, \\ 0 \text{ otherwise.} \end{cases}$$

But $\alpha_{(i)} - \alpha_{(j)}$ is not a root. So first case cannot occur. So find: $[e^i_+, e^j_-] = \delta_{ij} h^i$ as required.

Finally, need the Chevalley-Serre relations. Consider for example:

$$[e^i_+, e^j_+] \propto \begin{cases} e^{\alpha_{(i)} + \alpha_{(j)}}, \text{ if } \alpha_{(i)} + \alpha_{(j)} \in \Phi, \\ 0 \text{ otherwise.} \end{cases}$$

Similarly:

$$(\mathrm{ad}_{e^i_+})^n e^j_+ \propto \begin{cases} e^{n\alpha_{(i)} + \alpha_{(j)}}, \text{ if } n\alpha_{(i)} + \alpha_{(j)} \in \Phi, \\ 0 \text{ otherwise,} \end{cases}$$

and so the question comes down to lengths of root strings. In particular, we know that the $\alpha_{(i)}$ root string passing through $\alpha_{(j)}$ is

$$\alpha_{(j)}, \quad \alpha_{(j)} + \alpha_{(i)}, \quad \alpha_{(j)} + 2\alpha_{(i)}, \ldots$$

and we know its length is $1 - A^{ji}$. Hence the Chevalley-Serre relations follow. □

6.9 The Cartan classification

Theorem (Cartan Classification Version I):
A finite-dimensional, simple, complex Lie algebra $\mathfrak{g}$ is uniquely determined by its Cartan matrix.

Proof: From the above commutation relations. It can be shown that these generate the full Lie algebra by taking repeated brackets and using the Chevalley-Serre relations. □

6.10 Constraints on the Cartan matrix

Lemma: If α is a root, then $k\alpha$ is a root iff $k = \pm 1$.

Proof: Consider $S_{\alpha,\alpha} = \text{span}_{\mathbb{C}}\{\alpha + l\alpha : l \in \mathbb{C}\}$, the *generalised* α string passing through α. Let $V_{\alpha,\alpha} = \text{span}_{\mathbb{C}}(\{h^\alpha\} \cup \{e^{\alpha + l\alpha} : \alpha + l\alpha \in \Phi, l \in \mathbb{C}\})$.

Clearly $V_{\alpha,\alpha}$ is a rep space for $\mathfrak{sl}(2)_\alpha$. Therefore, we know the weights of $\mathfrak{sl}(2)_\alpha$ acting on $V_{\alpha,\alpha}$. Note that

$$[h^\alpha, e^{\alpha + l\alpha}] = (2 + 2l) e^{\alpha + l\alpha}.$$

Thus for all l, $2 + 2l$ is a weight of a rep of $\mathfrak{sl}(2)_\alpha$. But there are finitely many of these, and they are all integers, so $l \in \mathbb{Z}$.

We know that $\pm\alpha$ are roots, so $V_{\alpha,\alpha}$ already contains $e^\alpha, h^\alpha, e^{-\alpha}$, so already contains an irrep of $\mathfrak{sl}(2)_\alpha$, namely the adjoint irrep.

Now suppose $k\alpha$ is a root for $k > 1$. Then the lowering operator implies that $e^{k\alpha}$ is in the same irrep of $\mathfrak{sl}(2)_\alpha$ as $e^{k\alpha - 1}$, $e^{k\alpha - 2}$, etc. down to e^α. So e^α appears in two irreps.

But irreps are linearly independent spaces of $V_{\alpha,\alpha}$. Contradiction. Similarly for $k\alpha$, where $k < -1$. □

Theorem: The Cartan matrix A^{ij} obeys:

(i) $A^{ii} = 2$;

(ii) $A^{ij} = 0$ if and only if $A^{ji} = 0$;

(iii) $A^{ij} \in \mathbb{Z}_{\leq 0}$ for $i \neq j$.

(iv) $\det(A) > 0$;

(v) For simple Lie algebras, the Cartan matrix is irreducible.

(vi) $A^{ij} A^{ji} \in \{0, 1, 2, 3\}$.

Proof: (i) Obvious by definition; (ii) by symmetry of inner product; (iii) for simple roots, $(\alpha_{(i)}, \alpha_{(j)}) \leq 0$.

For (iv), notice that $(\kappa^{-1})_{ij} = (\alpha_{(i)}, \alpha_{(j)})$ is a symmetric matrix, so can be diagonalised. So if v_ρ is the eigenvector with eigenvalue ρ, we have $0 < (v_\rho, v_\rho) = (\kappa^{-1})_{ij} v^i_\rho v^j_\rho = \rho \sum_i (v^i_\rho)^2$. Thus $\rho > 0$, and it follows $\det(\kappa^{-1}) > 0$.

The Cartan matrix is $A^{ij} = (\kappa^{-1})^{ik} D_k{}^j$, where $D_k{}^j = 2\delta_k{}^j / (\alpha_{(j)}, \alpha_{(j)})$. Then $\det(A) = \det(\kappa^{-1}) \det(D) > 0$.

(v) is genuinely hard.

(vi) Use Cauchy-Schwarz and the Lemma above. □

6.11 Dynkin diagrams

Definition: The *Dynkin diagram* of a Lie algebra is a graph constructed from the Cartan matrix as follows:

1. Draw a node for each simple root $\alpha_{(i)} \in \Phi_S$.
2. Join the nodes corresponding to the simple roots $\alpha_{(i)}$, $\alpha_{(j)}$ by some lines. The number of lines we use is:

$$\max\{|A^{ij}|, |A^{ji}|\} \in \{0, 1, 2, 3\}.$$

3. If the roots have different lengths, draw an arrow pointing from the node corresponding to the longer root towards the shorter root.

All possible Dynkin diagrams were classified by Cartan:

Theorem (Cartan Classification Version II):
A finite-dimensional, simple, complex Lie algebra is uniquely determined by its Dynkin diagram, which must be one of the following types:

1. $A_n \cong \mathcal{L}_{\mathbb{C}}(SU(n+1))$.
2. $B_n \cong \mathcal{L}_{\mathbb{C}}(SO(2n+1))$.
3. $C_n \cong \mathcal{L}_{\mathbb{C}}(SP(2n))$.
4. $D_n \cong \mathcal{L}_{\mathbb{C}}(SO(2n))$.
5. E_6.
6. E_7.
7. E_8.
8. F_4.
9. G_2.

6.12 Example Dynkin diagram constraints

Theorem: Dynkin diagrams:

(i) cannot contain any closed loops;

(ii) each vertex meets at at most 3 lines.

Proof: WLOG work with $\alpha_{(i)}$, the *normalised* simple roots, i.e. $|\alpha_{(i)}| = 1$.

(i) Let $1, 2, ..., k$ be a loop of nodes containing no subloops. Let the corresponding roots be $\alpha_{(i)}$. Since there are no subloops, the only non-vanishing inner products between the $\alpha_{(i)}$'s are: $(\alpha_{(i)}, \alpha_{(i+1)})$ (with $\alpha_{(k+1)} \equiv \alpha_{(1)}$). Also:

$$\underbrace{|2(\alpha_{(i)}, \alpha_{(i+1)})|}_{A^{i,i+1}} \geq 1 \quad \Rightarrow \quad (\alpha_{(i)}, \alpha_{(i+1)}) \leq -\frac{1}{2}.$$

This is because $A^{ij} = 0$ iff $A^{ji} = 0$, and since we have at least one edge between i and $i+1$, need $|A^{i,i+1}| \geq 1$. Also the inner product of simple roots is negative.

Define $\chi = \alpha_{(1)} + ... + \alpha_{(k)}$. Then

$$(\chi, \chi) = \sum_{i=1}^{k} (\alpha_{(i)}, \alpha_{(i)}) + 2\sum_{i=1}^{k} (\alpha_{(i)}, \alpha_{(i+1)} \leq k - \frac{2}{2}k = 0.$$

Hence $\chi = 0$. Contradiction since simple roots are linearly independent.

(ii) Let α be a normalised simple root connected to the root $\alpha_{(i)}$ by n_i edges. If $i \neq j$, $(\alpha_{(i)}, \alpha_{(j)}) = 0$ as there are no closed loops. Set

$$\chi = \alpha - \sum_i (\alpha, \alpha_{(i)})\alpha_{(i)}.$$

Then $(\chi, \alpha_{(i)}) = 0$, and so:

$$1 = (\alpha, \alpha) = (\chi, \chi) + \sum_i (\alpha, \alpha_{(i)})^2 = (\chi, \chi) + \frac{1}{4}\sum_i n_i.$$

But χ cannot vanish by linear independence, so $\sum_i n_i < 4$. So α can have at most 3 edges emanating from it. □

In general, the trick to proving these things is to come up with a clever χ, somehow consider (χ, χ) (perhaps indirectly as in (ii)), and use linear independence of the roots (equivalent to $(\chi, \chi) \neq 0$).

7 Reconstructing Lie algebras

7.1 Example: Root system of A_2

Consider the Cartan matrix:

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

This tells us $\mathfrak{g}$ has two simple roots α, β, with

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = \frac{2(\alpha,\beta)}{(\beta,\beta)} = -1.$$

Thus $|\alpha| = |\beta|$ and the angle between the roots is $2\pi/3$. To find the remaining roots, we use root strings:

Theorem: All the roots of the Lie algebra are generated by the following algorithm:

1. Let $S = \Phi_S$, the set of simple roots.
2. For each $\alpha \in \Phi_S$, compute the α string through each $s \in S$, and add all resulting elements to S.
3. Continue until we get no new elements of S from an application of Step 2.
4. S then contains all positive roots. Full root set is $S \cup (-S)$.

So in our case start with $S = \{\alpha, \beta\}$. Start with α strings. We have $l_{\alpha,\beta} = 1 - 2(\alpha,\beta)/(\alpha,\alpha) = 2$. So β, $\alpha+\beta$ are roots; add them to our set: $S = \{\alpha, \beta, \alpha+\beta\}$.

Next, $l_{\alpha,\alpha+\beta} = 1 - 2(\alpha,\alpha+\beta)/(\alpha,\alpha) = 0$, so no new roots here.

Now compute the β strings. We have $l_{\beta,\alpha} = 2$, so α, $\alpha+\beta$ are roots. Finally, $l_{\beta,\alpha+\beta} = 0$, so we're done. We've found all positive roots: $\alpha, \beta, \alpha+\beta$.

So the root system of A_2 is $\{\pm\alpha, \pm\beta, \pm(\alpha+\beta)\}$.

7.2 Example: recovering Lie brackets

We can now recover all Lie brackets of A_2. In the Cartan-Weyl basis, we have:

$$\begin{aligned}
&[h^{\alpha_1}, h^{\alpha_2}] = 0, \\
&[h^{\alpha_1}, e^{\alpha_2}] = \frac{2(\alpha_1,\alpha_2)}{(\alpha_1,\alpha_1)} e^{\alpha_2}, \\
&[e^{\alpha_1}, e^{\alpha_2}] = \begin{cases} n_{\alpha_1,\alpha_2} e^{\alpha_1+\alpha_2}, \text{ if } \alpha_1+\alpha_2 \in \Phi, \\ h^{\alpha_1}, \text{ if } \alpha_2 = -\alpha_1, \\ 0 \text{ otherwise.} \end{cases}
\end{aligned}$$

How do we figure out the normalisation constants n_{α_1,α_2}, which up to now have been undetermined?

Idea: Let $\theta = \alpha + \beta$. Fix the normalisation constants by considering $[e^\theta, e^{-\theta}]$, which is given by:

$$[e^\theta, e^{-\theta}] = h^\theta = h^\alpha + h^\beta.$$

Expand this in three ways:

1. Expand first argument of bracket:
$$\begin{aligned}
[e^\theta, e^{-\theta}] &= \frac{1}{n_{\alpha,\beta}} [[e^\alpha, e^\beta], e^{-\theta}] \\
&= -\frac{1}{n_{\alpha,\beta}} \left([[e^\beta, e^{-\theta}], e^\alpha] + [[e^{-\theta}, e^\alpha], e^\beta] \right) \\
&= \frac{n_{\beta,-\theta}}{n_{\alpha,\beta}} h^\alpha + \frac{n_{-\theta,\alpha}}{n_{\alpha,\beta}} h^\beta.
\end{aligned}$$
So $n_{\beta,-\theta} = n_{\alpha,\beta} = n_{-\theta,\alpha}$.
2. Expanding the second argument, we get $n_{-\beta,\theta} = n_{\theta,-\alpha} = n_{-\alpha,-\beta}$ similarly.
3. Finally, expand both arguments at the same time:
$$[e^\theta, e^{-\theta}] = \frac{1}{n_{\alpha,\beta} n_{-\alpha,-\beta}} [[e^\alpha, e^\beta], [e^{-\alpha}, e^{-\beta}]]$$
Define $X = [e^{-\alpha}, e^{-\beta}]$. Then using Jacobi, we have:
$$[e^\theta, e^{-\theta}] = -\frac{1}{n_{\alpha,\beta} n_{-\alpha,-\beta}} \left([[X, e^\alpha], e^\beta] + [[e^\beta, X], e^\alpha] \right).$$
We have: $[X, e^\alpha] = -[h^\alpha, e^{-\beta}]$ (using Jacobi and fact $\alpha - \beta$ is not a root), and $[e^\beta, X] = [e^\alpha, h^\beta]$. We're free to set all roots lengths to 1; then using our knowledge of the roots system above (in particular $(\alpha,\beta) = \cos(2\pi/3) = -1/2$, we're left with:
$$[e^\theta, e^{-\theta}] = -\frac{1}{n_{\alpha,\beta} n_{-\alpha,-\beta}} (h^\alpha + h^\beta).$$
It follows $n_{\alpha,\beta} = -n_{-\alpha,-\beta}$.

Hence we've fixed all brackets in terms of a single normalisation constant $n_{\alpha,\beta}$. Rescaling e^θ, we can fix this to a desired value.

---

8 More representations

8.1 Definitions

We can generalise our treatment of the reps of $\mathcal{L}_{\mathbb{C}}(SU(2))$ to general Lie algebras $\mathfrak{g}$.

Let R be an N-dimensional rep of $\mathfrak{g}$. Let H^i be the Cartan generators. Assume $R(H^i)$ is diagonalisable. Then

Theorem: $\{R(H^i)\}$ are simultaneously diagonalisable.

Proof: R is a rep, so $[R(H^i), R(H^j)] = R([H^i, H^j]) = 0$. □

Hence the representation space $V \cong \mathbb{C}^N$ is spanned by simultaneous eigenvectors of $\{R(H^i)\}$:

Definition: Let V_λ be the eigenspace defined by

$$V_\lambda = \{v : R(H^i)v = \lambda^i v, \lambda = (\lambda^1, ..., \lambda^N) \in \mathbb{C}^N\}.$$

We call $\lambda \in \mathfrak{h}^*$ the *weights* of the representation. We call the set of weights $S_R = \{\lambda\}$ the *weight set*. By our previous work, we can decompose the representation space V as:

$$V = \bigoplus_{\lambda \in S_R} V_\lambda.$$

The *multiplicity* of the weight λ is the dimension of V_λ.

Example: The roots of the Lie algebra are the weights of the adjoint representation $R(X) = \mathrm{ad}_X$.

Theorem: Let $v \in V_\lambda$. The reps of the step operators $R(E^\alpha)$ obey:

$$R(E^\alpha)v \begin{cases} \in V_{\lambda+\alpha} \text{ if } \lambda + \alpha \in S_R; \\ = 0 \text{ otherwise.} \end{cases}$$

Proof: We have

$$\begin{aligned} R(H^i)R(E^\alpha)v &= R(E^\alpha)R(H^i)v + [R(H^i), R(E^\alpha)]v \\ &= \lambda^i R(E^\alpha)v + R([H^i, E^\alpha])v \\ &= (\lambda^i + \alpha^i)R(E^\alpha)v. \quad \square \end{aligned}$$

Since both roots and weights are in $\mathfrak{h}^*$, we can consider their inner product. We find we can generalise the quantisation condition as:

Theorem (Quantisation Condition II): For $\lambda \in S_R$, and $\alpha \in \Phi$:

$$\frac{2(\alpha, \lambda)}{(\alpha, \alpha)} \in \mathbb{Z}.$$

Proof: Consider the action of $\mathfrak{sl}(2)_\alpha$ generators: $\{R(h^\alpha), R(e^\alpha), R(e^{-\alpha})\}$ on V. Each generators defines a linear map $V \to V$, so V is a valid representation space. For any $v \in V_\lambda$, we have (by normalisation of Cartan-Weyl basis):

$$\begin{aligned} R(h^\alpha)v = \frac{2}{(\alpha, \alpha)}(\kappa^{-1})_{ij}\alpha^i R(H^j)v &= \frac{2}{(\alpha, \alpha)}(\kappa^{-1})_{ij}\alpha^i \lambda^j v \\ &= \frac{2(\alpha, \lambda)}{(\alpha, \alpha)}v. \end{aligned}$$

The coefficient is a weight of R, and all weights of $\mathcal{L}_\mathbb{C}(SU(2))$ representations are integers, so have result. $\square$

8.2 Root and weight lattices

Definition: The *root lattice* of a Lie algebra is

$$\mathcal{L}[\mathfrak{g}] = \left\{ \sum_{i=1}^r m^i \alpha_{(i)} : m^i \in \mathbb{Z} \right\} = \mathrm{span}_\mathbb{Z} \left\{ \alpha_{(i)} : i = 1, ..., r \right\},$$

where the $\alpha_{(i)}$ are the simple roots.

Note that all roots in Φ lie in the root lattice (from Properties of Simple Roots Theorem), but not all points in the root lattice correspond to roots.

Definition: Define the *simple co-roots* by:

$$\hat{\alpha}_{(i)} = \frac{2\alpha_{(i)}}{(\alpha_{(i)}, \alpha_{(i)})},$$

and the *co-root lattice* by

$$\hat{\mathcal{L}}[\mathfrak{g}] = \mathrm{span}_\mathbb{Z} \left\{ \hat{\alpha}_{(i)} : i = 1, 2, ..., r \right\}.$$

Definition: The *weight lattice* $\mathcal{L}_W[\mathfrak{g}]$ is the lattice dual to the co-root lattice:

$$\mathcal{L}_W[\mathfrak{g}] = \{\lambda \in \mathfrak{h}^*_\mathbb{R} : (\lambda, \mu) \in \mathbb{Z}, \mu \in \hat{\mathcal{L}}[\mathfrak{g}]\}.$$

Writing $\mu = \hat{\alpha}_{(i)} n_i$ for integers n_i, the condition for λ to be in $\mathcal{L}_W[\mathfrak{g}]$ becomes:

$$\frac{2(\alpha_{(i)}, \lambda)}{(\alpha_{(i)}, \alpha_{(i)})} \in \mathbb{Z}.$$

Note: Quantisation Condition II then exactly says: for any representation R, all of its weights λ lie in the weight lattice $\mathcal{L}_W[\mathfrak{g}]$. Note the converse is not true.

Definition: Given a basis for $\hat{\mathcal{L}}[\mathfrak{g}]$, $B = \{\hat{\alpha}_{(i)}\}$, define the *dual basis* for $\mathcal{L}_W[\mathfrak{g}]$ by $B^* = \{\omega_{(i)}\}$, where $\omega_{(i)}$ are called the *fundamental weights* and are determined by: $(\hat{\alpha}_{(i)}, \omega_{(j)}) = \delta_{ij}$, i.e.

$$\frac{2(\alpha_{(i)}, \omega_{(j)})}{(\alpha_{(i)}, \alpha_{(i)})} = \delta_{ij}.$$

Since the simple roots span $\mathfrak{h}^*_\mathbb{R}$, we may write

$$\omega_{(j)} = \sum_{k=1}^r B_{jk}\alpha_{(k)}.$$

Substituting into the quantisation condition, we have $B_{jk}A^{ki} = \delta_j{}^i$. It follows that B is the inverse of the Cartan matrix. So we may write:

$$\alpha_{(i)} = \sum_{j=1}^r A^{ij}\omega_{(j)}.$$

Example: Consider A_2, with Cartan matrix:

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

We have:

$$\alpha_{(1)} = 2\omega_{(1)} - \omega_{(2)}, \quad \alpha_{(2)} = -\omega_{(1)} + 2\omega_{(2)}.$$

Hence

$$\omega_{(1)} = \frac{1}{3}\left(2\alpha_{(1)} + \alpha_{(2)}\right), \quad \omega_{(2)} = \frac{1}{3}\left(\alpha_{(1)} + 2\alpha_{(2)}\right).$$

Definition: For any weight λ in $S_R \subseteq \mathcal{L}_W[\mathfrak{g}]$, we can write

$$\lambda = \sum_{i=1}^{r} \lambda^i \omega_{(i)}.$$

We call the integers $\{\lambda^i\}$ the *Dynkin labels* of λ.

8.3 Highest weight representations

Definition: The *highest weight* of a finite-dimensional representation R of $\mathfrak{g}$ is a weight

$$\Lambda = \sum_{i=1}^{r} \Lambda^i \omega_{(i)}$$

with $\Lambda^i \in \mathbb{Z}_{\geq 0}$ such that its corresponding eigenvector v_Λ obeys

$$R(H^i)v_\Lambda = \Lambda^i v_\Lambda, \quad R(E^\alpha)v_\Lambda = 0.$$

We call the integers Λ^i the *Dynkin labels* of the representation.

We can rephrase the step operator condition using the result about $R(E^\alpha)v$. We have that $R(E^\alpha)v_\Lambda = 0$ for all $\alpha \in \Phi^+$ iff $\Lambda + \alpha$ is not a weight for all $\alpha \in \Phi^+$.

All weights can be generated from the highest weight using lowering operators. This is encoded in the result:

Theorem: If

$$\lambda = \sum_{i=1}^{r} \lambda^i \omega_{(i)}$$

is a weight, then

$$\lambda - m_{(i)}\alpha_{(i)}$$

is a weight for all $m_{(i)} \in \mathbb{Z}$ with $0 \leq m_{(i)} \leq \lambda^i$, $i = 1, 2, ..., r$.

Proof: Encodes fact we get weights from lowering operators. $\square$

8.4 Irreps of $A_2 \cong \mathcal{L}_{\mathbb{C}}(SU(3))$

Notation: Write $R_{(\Lambda^1,\Lambda^2)}$ for the irrep of A_2 with highest weight $\Lambda = (\Lambda^1, \Lambda^2)$.

Example: Consider $A_2 \cong \mathcal{L}_{\mathbb{C}}(SU(3)) = \{3 \times 3$ traceless matrices$\}$ in the fundamental representation. Recall that the Cartan subalgebra is generated by

$$H^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

and therefore the simultaneous evectors are $v_1 = (1,0,0)^T$ (weight $(1,0)^T$), $v_2 = (0,1,0)^T$ (weight $(-1,1)^T$) and $v_3 = (0,0,1)^T$ (weight $(0,-1)^T$).

The step operators are just matrices with a single one on an off-diagonal. Thus we can very quickly calculate the roots of the Lie algebra to be

$$\{(2,-1)^T, (1,1)^T, (-1,2)^T, (-2,1)^T, (-1,-1)^T, (1,-2)^T, \mathbf{0}\}.$$

($\mathbf{0}$ is included by convention.) Splitting these into positive and negative roots, we choose

$$\Phi^+ = \{(2,-1)^T, (1,1)^T, (1,-2)^T\}.$$

Then the simple roots are $\alpha_{(1)} = (1,1)^T$ and $\alpha_{(2)} = (1,-2)^T$. Using our calculation above, it follows that $\omega_{(1)} = (1,0)^T$ and $\omega_{(2)} = (1,-1)^T$.

Examining the weights we calculated above, we see the only weight which when added to any of the Φ^+ elements does not become a weight is $(1,0)^T = \omega_{(1)}$. Hence the Dynkin label of the fundamental representation is $(\Lambda^1, \Lambda^2) = (1,0)$.

Example: If we'd have known the Dynkin labels before, we could have constructed all the weights of A_2's fundamental representation imemdiately:

$$(1,0) \underset{-\alpha_{(1)}}{\rightarrow} (-1,1) \underset{-\alpha_{(2)}}{\rightarrow} (0,-1).$$

So the weights are $\omega_{(1)}, -\omega_{(1)} + \omega_{(2)}$ and $-\omega_{(2)}$. In particular, $\dim(R_{(1,0)}) = 3$, as expected.

Example: The adjoint representation of A_2 has weights equal to the roots. Adding all elements of Φ^+ to all roots that we calculated above, we see that the highest weight is $(2,-1)^T = \omega_{(1)} + \omega_{(2)}$. Hence its Dynkin label is $(1,1)$.

Again, if we'd known this beforehand, we'd have calculated the roots to be:

$$S = \{0, \omega_{(1)} + \omega_{(2)}, -\omega_{(1)} + 2\omega_{(2)}, \omega_{(1)} - 2\omega_{(2)}, -\omega_{(1)} - \omega_{(2)}, 2\omega_{(1)} - \omega_{(2)}, -2\omega_{(1)} + \omega_{(2)}\}.$$

Note there are seven weights, so the dimension of $R_{(1,1)}$ is at least 7. Could there be degenerate weights?

In this case, there obviously are. The Cartan subalgebra contains 2 linearly independent elements, H^1 and H^2, so $R(H^i)H^1 = 0$ and $R(H^i)H^2 = 0$. So the root 0 has multiplicity 2. Thus the dimension of $R_{(1,1)}$ is 8.

In general, we can check dimensions using:

Theorem: The dimension of $R_{(\Lambda^1,\Lambda^2)}$ is:

$$\dim(R_{(\Lambda^1,\Lambda^2)}) = \frac{1}{2}(\Lambda^1+1)(\Lambda^2+1)(\Lambda^1+\Lambda^2+2).$$

Proof: Requires an understanding of character formulae, so beyond scope of course. □

We note that this is symmetric in Λ^1 and Λ^2. This is due to the existence of the conjugate representation: $R_{(\Lambda^2,\Lambda^1)} = \overline{R}_{(\Lambda^1,\Lambda^2)}$.

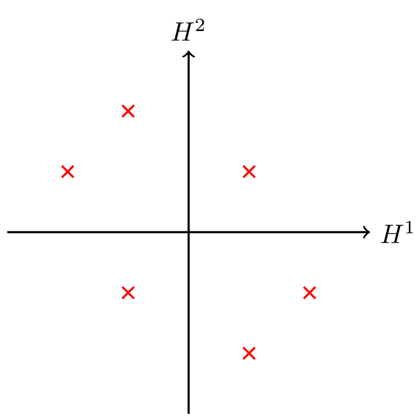
Definition: Physicists use the following notation for A_2 irreps. They denote the irrep by $\underline{n}$, where n is the dimension of the irrep. If the Dynkin label of the irrep is (Λ^1,Λ^2), and satisfies $\Lambda^2 > \Lambda^1$, they put an overline, $\overline{\underline{n}}$, to be suggestive of the conjugate irrep.

Example: The fundamental representation is written $\underline{3}$. The anti-fundamental representation is written $\overline{\underline{3}}$.

8.5 Graphical techniques for A_2

In the above, we calculated the roots α of A_2 with respect to the Cartan subalgebra $\mathrm{span}\{H^1,H^2\}$. It is important to realise that expressions like $\alpha = (2,-1)^T$ are in this basis; they say that α acts as $\alpha(H^1) = 2$ and $\alpha(H^2) = -1$.

In this basis, the root diagram for A_2 looks like:



This is not very symmetric. Perhaps a better choice of Cartan subalgebra basis would help?

Theorem: With the basis

$$H'^1 = \frac{1}{2}\begin{pmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}, \quad H'^2 = \frac{1}{2\sqrt{3}}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{pmatrix},$$

for the Cartan subalgebra, the root system for the Lie algebra A_2 becomes

$$\pm\left\{(1,0)^T, \left(-\frac{1}{2},\frac{\sqrt{3}}{2}\right)^T, \left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)^T\right\}.$$

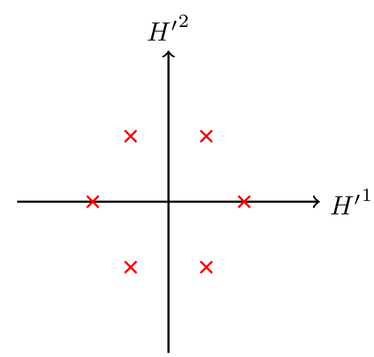
In particular, $\alpha_{(1)} = (\frac{1}{2},\frac{\sqrt{3}}{2})$ and $\alpha_{(2)} = (\frac{1}{2},-\frac{\sqrt{3}}{2})$.

Proof: We notice that

$$\alpha(H'^1) = \frac{1}{2}\alpha(H^1), \quad \alpha(H'^2) = \frac{\sqrt{3}}{6}(\alpha(H^1)+2\alpha(H^2)).$$

So the rule for conversion is: the new first component is half the hold one, and then new second component is the sum of the old first one and twice the old second one, multiplied by $\sqrt{3}/6$. □

This is much prettier, with root diagram now a nice neat hexagon:



Why did this work? Well, in our original basis, the inner product (α_1,α_2) involved a non-diagonal Killing form, since the H^i were not orthogonal with respect to the Killing form. Therefore, (α_1,α_2) seemed like a 'stretched version' of our Euclidean inner product (in particular, the 'dot product' between two roots doesn't give a cosine - we need a matrix in the middle).

The new basis is actually orthonormal with respect to the Killing form, so it looks directly comparable to the standard Euclidean inner product.

The new basis has provided a lot more symmetry for our root diagram. This will be particularly helpful in the following example.

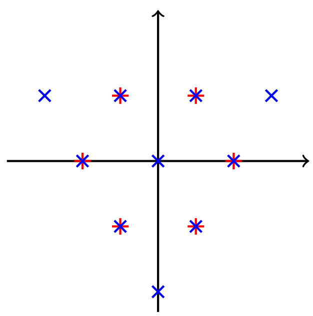
Example: Consider the rep $R_{(3,0)}$ of A_2. In physicist's notation, this is $\underline{10}$, since it is 10-dimensional by the formula above. The weights of the representation are given by the standard algorithm applied to the Dynkin label $(3, 0)$:

$$\{0, 3\omega_{(1)}, \omega_{(1)} + \omega_{(2)}, 2\omega_{(1)} - \omega_{(2)}, -2\omega_{(1)} + \omega_{(2)},\\ -\omega_{(1)} - \omega_{(2)}, -\omega_{(1)} + 2\omega_{(2)}, \omega_{(1)} - 2\omega_{(2)},\\ -3\omega_{(1)} + 3\omega_{(2)}, -3\omega_{(2)}\}$$

In the new symmetric basis, we have

$$\omega_{(1)} = \frac{1}{3}(2\alpha_{(1)} + \alpha_{(2)}) = \left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right)$$

and $\omega_{(2)} = \left(\frac{1}{2}, -\frac{\sqrt{3}}{6}\right)$. Hence the *weight* diagram in this case takes the form:



The roots are drawn as red pluses for comparison.

One question that we've come to ask very often about representations is whether we can decompose them into irreps. In this case, we can decompose everything into $\mathfrak{sl}(2)_\alpha$ irreps *graphically* as follows (here, α is any root - it doesn't matter which we pick by their inherent symmetry, manifest in the diagram and the fact $\mathfrak{sl}(2)_\alpha$ depends on the *pair of roots* $\pm\alpha$).

Write R_Λ for the $\mathfrak{sl}(2)_\alpha$ irrep of dimension $\Lambda + 1$. Now the only thing we need do is recall that $\mathfrak{sl}(2)_\alpha$ has a representation on $V_{\alpha,\beta} = \mathrm{span}_{\mathbb{C}}\{e^{\beta+\rho\alpha} : \beta + \rho\alpha \in \Phi, \rho \in \mathbb{Z}\}$, for each β.

It follows from this fact that we get different representations of $\mathfrak{sl}(2)_\alpha$ depending on our choice of β. Picking any α as our starting point, and looking at lines parallel to α through the roots, we see that there are 4 roots in one of the $V_{\alpha,\beta}$, 3 in another, 2 in another and 1 in the last.

So we see immediately that the decomposition must be of the form:

$$\underline{10} = R_3 \oplus R_2 \oplus R_1 \oplus R_0.$$

We can use similar graphical techniques to evalute *tensor products* of A_2 representations. We need the following Theorem:

Theorem: Let R and $\tilde{R}$ be irreps of a Lie algebra $\mathfrak{g}$ with rep spaces V, $\tilde{V}$ and weight space decompositions:

$$V = \bigoplus_{\lambda\in S} V_\lambda, \quad \tilde{V} = \bigoplus_{\lambda'\in\tilde{S}} \tilde{V}_\lambda,$$

where S and $\tilde{S}$ are their respective weight sets. Then the weight set of $R \otimes \tilde{R}$ is

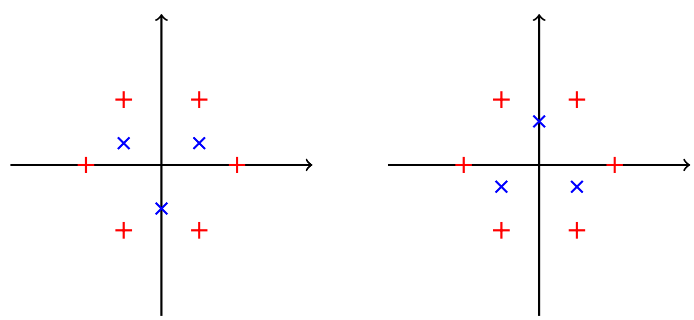
$$S_{R\otimes\tilde{R}} = \{\lambda + \lambda' : \lambda \in S, \lambda' \in \tilde{S}\}.$$

Proof: We have v_λ, $\tilde{v}_{\lambda'}$ such that $R(H^i)v_\lambda = \lambda^i v_\lambda$ and $\tilde{R}(H^i)\tilde{v}_{\lambda'} = \lambda'^i \tilde{v}_{\lambda'}$. Hence $(R \otimes \tilde{R})(H^i)(v_\lambda \otimes \tilde{v}_{\lambda'}) =$
$R(H^i)v_\lambda \otimes \tilde{v}_{\lambda'} + v_\lambda \otimes \tilde{R}(H^i)\tilde{v}_{\lambda'} = (\lambda + \lambda')^i(v_\lambda \otimes \tilde{v}_{\lambda'})$. $\square$

Example: Consider $\underline{3} \otimes \underline{\overline{3}}$. By the standard algorithm, we have that $\underline{3}$ and $\underline{\overline{3}}$ have respective weight sets:

$$S_{(1,0)} = \{\omega_{(1)}, -\omega_{(1)} + \omega_{(2)}, -\omega_{(2)}\},\\ S_{(0,1)} = \{\omega_{(2)}, \omega_{(1)} - \omega_{(2)}, -\omega_{(1)}\}.$$

Inserting the numerical values of $\omega_{(1)}$ and $\omega_{(2)}$ in our numerical basis, we get the weight diagrams:



On the left is the fundamental rep, on the right is the anti-fundamental rep. To work out the weight set of the tensor product, recall that we *add* the weights of the factors in all possible ways.

Doing so, we produce the weight diagram for $\underline{8}$ together with the weight diagram for $\underline{1}$. We deduce that $\underline{3}\otimes\underline{\overline{3}} = \underline{8}\oplus\underline{1}$.

Example: It's easy, but not essential, to use graphical methods. Standard algebraic methods work just as well. For example, to find $\underline{3} \otimes \underline{3} \otimes \underline{3}$, we first work out $\underline{3} \otimes \underline{3}$.

Summing the elements of the weight sets of $\underline{3}$ and $\underline{3}$ in all possible ways, we find: $S_{\underline{3}\otimes\underline{3}} = \{(2, 0), (0, 1), (1, -1), (0, 2), (-2, 2), (-1, 0), (1, -1), (-1, 0), (0, -2)\}$.

Note the highest weight here is $(2, 0)$; computing the weight set of $R_{(2,0)} = \underline{6}$ and subtracting, we see that $\underline{3} \otimes \underline{3} = \underline{6} \oplus \underline{1}$. Iterating, we find $\underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10}$.

8.6 Irreps of $B_2 \cong \mathcal{L}_{\mathbb{C}}(SO(5))$

Example: The Cartan matrix of B_2 is

$$\begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}.$$

It's possible to show that the Dynkin labels of the fundamental representation are again $(1,0)$. Thus using:

$$\alpha_{(1)} = 2\omega_{(1)} - 2\omega_{(2)}, \quad \alpha_{(2)} = -\omega_{(1)} + 2\omega_{(2)},$$

we have that the weights of the fundamental rep are:

$$\{\omega_{(1)}, -\omega_{(1)} + 2\omega_{(2)}, 0, \omega_{(1)} - 2\omega_{(2)}, -\omega_{(1)}\}.$$

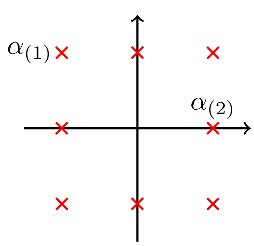
Note the dimensions are correct for the fundamental rep.

Example: Let's consider the adjoint rep of B_2. Again, the weights are just the roots of the Lie algebra. If we set $|\alpha_{(2)}| = 1$, then the Cartan matrix implies $|\alpha_{(1)}| = \sqrt{2}$ and the angle between them is $3\pi/4$.

By considering root strings, we can work out all the roots to be:

$$\pm\{\alpha_{(1)}, \alpha_{(1)} + \alpha_{(2)}, \alpha_{(1)} + 2\alpha_{(2)}, \alpha_{(2)}\}.$$

We can therefore sketch the root system as:



It's now easy, by considering adding roots together, to identify that the highest weight is $2\omega_{(2)} = \alpha_{(1)} + 2\alpha_{(2)}$. Therefore the Dynkin label is $(0,2)$ for the adjoint representation of B_2.

Applying our standard algorithm now, we have that the weights are:

$$\pm\{0, 2\omega_{(2)}, \omega_{(1)}, -2\omega_{(2)} + 2\omega_{(1)}, 2\omega_{(2)} - \omega_{(1)}\}.$$

Again, 0 has multiplicity 2 because the Cartan subalgebra has dimension 2 here, since it's generated by the matrices:

$$H^1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad H^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(Recall the general form of the CSA for $\mathcal{L}_{\mathbb{C}}(SO(2n+1))$, from earlier in these notes.)

ASIDE: While we're discussing $\mathcal{L}_{\mathbb{C}}(SO(5))$, it's worth mentioning the form of its step operators, since these are not too obvious. There are step operators:

$$F^{\pm} = \begin{pmatrix} 0 & 0 & 1 & \pm i & 0 \\ 0 & 0 & \pm i & -1 & 0 \\ -1 & \mp i & 0 & 0 & 0 \\ \mp i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad G^{\pm} = \begin{pmatrix} 0 & 0 & \pm i & 1 & 0 \\ 0 & 0 & -1 & \pm i & 0 \\ \mp i & 1 & 0 & 0 & 0 \\ -1 & \mp i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which are generic to both $\mathcal{L}_{\mathbb{C}}(SO(2n))$ and $\mathcal{L}_{\mathbb{C}}(SO(2n+1))$, as we saw earlier. However there are additional ones for the odd matrices, as in $\mathcal{L}_{\mathbb{C}}(SO(5))$:

$$E_1^{\pm} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \pm i \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & \mp i & 0 & 0 & 0 \end{pmatrix}, \quad E_2^{\pm} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \pm i \\ 0 & 0 & -1 & \mp i & 0 \end{pmatrix}.$$

Example: Consider the tensor product of B_2 representations $R_{(0,1)} \otimes R_{(0,1)}$. The weight set of $R_{(0,1)}$ is

$$S_{(0,1)} = \{\omega_{(2)}, \omega_{(1)} - \omega_{(2)}, -\omega_{(1)} + \omega_{(2)}, -\omega_{(2)}\}.$$

Adding this to itself in all possible ways, and comparing to the weight sets we've calculated above, we see that

$$R_{(0,1)} \otimes R_{(0,1)} = R_{(0,2)} \oplus R_{(1,0)} \oplus R_{(0,0)}.$$

---

9 Symmetries in quantum theory

9.1 Symmetries and conserved quantities

Definition: A *symmetry transformation* of a quantum system is a transformation $|\psi\rangle \mapsto U|\psi\rangle$ for U a unitary operator such that $UHU^\dagger = H$, where H is the Hamiltonian.

Theorem: Symmetry transformations (i) preserve the inner product; (ii) leave the energy levels of the system invariant.

Proof: (i) $\langle\psi_1|\psi_2\rangle \mapsto \langle\psi_1|U^\dagger U|\psi_2\rangle = \langle\psi_1|\psi_2\rangle$ since unitary. (ii) Suppose $H|\psi\rangle = E|\psi\rangle$. Then since $UHU^\dagger = H \Rightarrow [H,U] = 0$, we have $HU|\psi\rangle = UH|\psi\rangle = EU|\psi\rangle$. $\square$

Definition: A *conserved quantity* is an observable θ, i.e. a Hermitian operator, obeying $[\theta, H] = 0$.

Since $U = e^{is\theta}$ is unitary for all s (since $i\theta$ is anti-Hermitian), we have $[H,U] = 0$ also, and thus $|\psi\rangle \to U|\psi\rangle$ is a symmetry transformation. This is the quantum form of Noether's Theorem.

Theorem: Let $\{\theta^a\}$ be a maximal set of conserved quantities. Then

$$\mathfrak{g}_{\mathbb{R}} = \mathrm{span}_{\mathbb{R}}\{i\theta^a\}$$

equipped with the operator commutator is a *real* Lie algebra.

Proof: Obvious. □

If $X \in \mathfrak{g}_{\mathbb{R}}$, we can exponentiate X to get $U = \exp(X)$, the symmetry transformation. But the exponential of a Lie algebra is a Lie group, so such symmetry transformations form a *compact Lie group* G.

Theorem: Let $\mathcal{H}_n$ be the eigenspace of the Hamiltonian with eigenvalue E_n. Then $\mathcal{H}_n$ is a representation space for both $\mathfrak{g}_{\mathbb{R}}$ and G, for some $\dim(\mathcal{H}_n)$ dimensional reps.

Proof: Let d be our candidate representation of $\mathfrak{g}_{\mathbb{R}}$, and set $D(U) = \exp(d(X))$ where $U = e^X$ as our representation of G. Let $|\psi\rangle \in \mathcal{H}_n$.

Now $D(U)$ is a matrix acting on a basis of states $\{|\psi\rangle_1, ..., |\psi\rangle_n\}$ spanning $\mathcal{H}_n$. Thus it lives in a completely different space to the Hamiltonian H. It follows that $[H, D(U)] = 0$ and $D(U)$ thus preserves the energy of the state. Similarly for $d(X)$. □

From the condition $UHU^\dagger = H$, taking D of this, we have $D(U)D(U^\dagger)H = H \Rightarrow D(U)^{-1} = D(U)^\dagger$. So the representation must be unitary. This implies that $d(X)^\dagger = -d(X)$, i.e. the representation of the Lie algebra is anti-Hermitian (though this is often confusingly referred to as *unitary* in the context of a Lie algebra).

---

10 Non-Abelian gauge theory

10.1 Abelian theory

In relativistic electrodynamics, we use a 4-*vector potential*, a_μ, which transforms under a *gauge transformation* as:

$$a_\mu \mapsto a_\mu + \partial_\mu \alpha(x).$$

Definition: The *field-strength tensor* of the theory is:

$$f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu.$$

This is obviously gauge-invariant.

Definition: The *Lagrangian* of the theory is

$$\mathcal{L} = -\frac{1}{4g^2} f^{\mu\nu} f_{\mu\nu}.$$

In physics, we want to couple this to matter fields. Consider coupling to the complex scalar field ϕ with Lagrangian

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - W(\phi^*\phi),$$

where W is some *self-interaction term*.

Theorem: The scalar theory is invariant under the $U(1)$ global internal symmetry $\phi \mapsto g\phi$, $\phi^* \mapsto g^{-1}\phi^*$, where $g \in U(1)$.

Proof: Obvious. □

It's also sometimes useful to consider this transformation infinitesimally. If $g = \exp(\epsilon X)$ for $X \in \mathcal{L}(U(1))$, we have $g \approx 1 + \epsilon X$, and thus

$$\phi \mapsto \phi + \epsilon X \phi, \quad \phi^* \mapsto \phi^* - \epsilon X \phi^*.$$

We write the changes as $\delta_X \phi = \epsilon X \phi$ and $\delta_X \phi^* = -\epsilon X \phi^*$.

How can we retain this symmetry in a gauge-invariant way when we move to consider interactions with electromagnetism? The answer is to *gauge* the theory.

Definition: When we allow g to have an x dependence:

$$\phi \mapsto g(x)\phi, \quad \phi^* \mapsto g^{-1}(x)\phi^*,$$

we say we are *gauging* the theory.

Does this mean that the matter Lagrangian is now gauge invariant? No:

Theorem: Under a gauge transformation, $\delta_X(\partial_\mu \phi) = \epsilon(\partial_\mu X)\phi + \epsilon X(\partial_\mu \phi)$.

Proof: We have $\delta_X(\partial_\mu \phi) = \partial_\mu(\delta_X \phi)$, whence

$$\partial_\mu(\delta_X \phi) = \partial_\mu(\epsilon X \phi). \quad \square$$

From this Theorem, it's evident that $\partial_\mu \phi \partial^\mu \phi$ will pick up many extra unwanted terms. To restore the gauge invariance then, we introduce:

Definition: Write $A_\mu = -ia_\mu$. The *covariant derivative* is defined by

$$D_\mu = \partial_\mu + A_\mu.$$

Since $a_\mu \in \mathbb{R}$, we have $A_\mu \in i\mathbb{R} = \mathcal{L}(U(1))$, so it's common in this course for us to prefer working with A_μ (as we like Lie algebras!).

Recall the transformation law for A_μ from above; under a gauge transformation we have:

$$A_\mu \mapsto A_\mu - \partial_\mu i\alpha(x).$$

We claim that by picking $\alpha = -i\epsilon X$, i.e. a transformation law

$$A_\mu \mapsto A_\mu - \epsilon\partial_\mu X = A_\mu + \delta_X A_\mu.$$

the covariant derivative gives rise to a gauge-invariant kinetic term for the matter fields.

Theorem: $D_\mu\phi D^\mu\phi^*$ is gauge invariant.

Proof: Work infinitesimally. We have

$$\delta_X(D_\mu\phi) = \delta_X(\partial_\mu\phi + A_\mu\phi) = \partial_\mu(\delta_X\phi) + (\delta_X A_\mu)\phi + A_\mu(\delta_X\phi).$$

Substituting all variations in, we find $\delta_X(D_\mu\phi) = \epsilon X D_\mu\phi$. So $D_\mu\phi$ transforms in the same way as ϕ. Thus

$$D_\mu\phi^* = (D_\mu\phi)^*$$

transforms in the same way as ϕ^*, and the result follows (all order $O(\epsilon)$ terms cancel). □

Therefore, the gauge invariant Lagrangian we seek is:

$$\mathcal{L} = \frac{1}{4g^2}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - W(\phi^*\phi),$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = -if_{\mu\nu}$.

10.2 Non-abelian generalisation

Suppose a Lagrangian has gauge symmetry generated by the Lie group G. Let D be a representation of G with representation space $V \cong \mathbb{C}^N$, equipped with the standard inner product $(u, v) = u^\dagger \cdot v$, and suppose that the matter fields ϕ lie in V.

The Lagrangian of the matter fields is now:

$$\mathcal{L} = (\partial_\mu\phi, \partial^\mu\phi) - W[(\phi, \phi)].$$

Theorem: Provided D is unitary, the Lagrangian is invariant under the global internal symmetry $\phi \mapsto D(g)\phi$.

Proof: Simple, using fact $(D(g)\phi, D(g)\phi) = (D(g)^\dagger D(g)\phi, \phi) = (\phi, \phi)$, since $D(g)^\dagger D(g) = 1$ for unitary transformations. □

If $g = \exp(\epsilon X)$ for $X \in \mathcal{L}(G)$, define the representation R by $D(g) = \exp(\epsilon R(X))$. We then have the infinitesimal equivalent of the symmetry transformation:

$$\phi \mapsto \phi + \epsilon R(X)\phi = \phi + \delta_X\phi.$$

Gauge the theory by giving X a spacetime dependence: $X \equiv X(x)$. Again, introduce the covariant derivative to maintain gauge-invariance:

Definition: The *covariant derivative* is

$$D_\mu\phi = \partial_\mu\phi + R(A_\mu)\phi,$$

where A_μ is the *gauge field* for G.

The gauge field transformation law is now

$$\delta_X A_\mu = -\epsilon\partial_\mu X + \epsilon[X, A_\mu].$$

Theorem: We have $\delta_X(D_\mu\phi) = \epsilon R(X)D_\mu\phi$. Since D is unitary, it follows $R(X)^\dagger = -R(X)$, and thus $(D_\mu\phi, D^\mu\phi)$ is gauge invariant.

Proof: We have

$$\delta_X(D_\mu\phi) = \partial_\mu(\delta_X\phi) + R(A_\mu)\delta_X\phi + R(\delta_X A_\mu)\phi.$$

Inserting the variations, and performing some commutation moves (also recall $R([X, A_\mu]) = [R(X), R(A_\mu)]$) we have the result. □

Therefore, the gauge invariant form of the matter field Lagrangian is

$$\mathcal{L} = (D_\mu\phi, D^\mu\phi) - W[(\phi, \phi)].$$

Can we still use $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ for the kinetic terms of the gauge field? No, as the transformation law for A_μ is different now. Instead we must use:

Definition: The *field-strength tensor* is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu].$$

Theorem: $\delta_X(F_{\mu\nu}) = \epsilon[X, F_{\mu\nu}]$.

Proof: Another simple calculation. The Jacobi identity is necessary at one point. □

We can now use the *Killing form* of the Lie algebra to define the kinetic term of the Lagrangian:

$$\mathcal{L} = \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}).$$

This is gauge invariant since

$$\delta_X\mathcal{L} = \frac{1}{g^2}\kappa(\delta_X F_{\mu\nu}, F^{\mu\nu}) + \frac{1}{g^2}\kappa(F_{\mu\nu}, \delta_X F^{\mu\nu}) = 0,$$

by invariance of the Killing form.

Note, however, that this is a sensible kinetic term iff $\mathcal{L}(G)$ is of compact type. That is, there is a basis $\{T^a\}$ in which $\kappa^{ab} = -\kappa\delta^{ab}$, $\kappa > 0$, so that

$$\mathcal{L} = -\frac{\kappa}{g^2}\sum_a F_{\mu\nu,a}F^{\mu\nu,a}.$$

We need this to avoid getting negative energy states.

Summary of infinitesimal gauge theory:

The gauge invariant Lagrangian for a theory with gauge group G is $\mathcal{L} =$

$$\frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) + \sum_{\Lambda \in S} \left((D_\mu \phi_\Lambda, D^\mu \phi_\Lambda) - W[(\phi_\Lambda, \phi_\Lambda)]\right),$$

where ϕ_Λ is in the rep space V_Λ of the irrep R_Λ of $\mathcal{L}(G)$, where Λ is the Dynkin label of the irrep. S is some set of Dynkin labels.

The field-strength tensor is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$$

and the covariant derivative is

$$D_\mu \phi_\Lambda = \partial_\mu \phi_\Lambda + R_\Lambda(A_\mu)\phi_\Lambda$$

The fields transform under gauge transformations as:

$$\delta_X A_\mu = -\epsilon \partial_\mu X + \epsilon[X, A_\mu],$$
$$\delta_X \phi_\Lambda = \epsilon R_\Lambda(X)\phi_\Lambda.$$

10.3 Finite analogues

In the above, we've worked primarily with infinitesimal transformations of the fields. It's possible to work with finite transformations too.

Under a *finite gauge transformation* $\phi \mapsto D(g)\phi$, we choose for the gauge field A_μ to transform as

$$A_\mu \mapsto g A_\mu g^{-1} - (\partial_\mu g) g^{-1}.$$

Theorem: $g A_\mu g^{-1} - (\partial_\mu g) g^{-1}$ is in the Lie algebra.

Proof: A_μ is in the Lie algebra by Definition. The tangent to the curve $g e^{tA_\mu} g^{-1} \in G$ at the identity is:

$$\frac{d}{dt}\left(g e^{tA_\mu} g^{-1}\right)\Big|_{t=0} = g A_\mu g^{-1} \in \mathcal{L}(G).$$

Let $g(t) \in G$ be an arbitrary curve with $g(0) = g_0$. Then $g(t) g_0^{-1}$ passes through the identity at $t = 0$, with tangent

$$\frac{d}{dt}\left(g(t) g_0^{-1}\right)\Big|_{t=0} = \dot{g}(0) g_0^{-1} \in \mathcal{L}(G).$$

Strictly, $g(t) = g(x(t))$, since g depends on spacetime points through $g = e^{X(x)}$, say. Thus $\dot{g}(0) = \dot{x}^\mu(0)\partial_\mu g(0)$. We can of course parametrise such that $\dot{x}^\mu(0) = 1$ is the only non-zero component of $\dot{x}(0)$, and thus $(\partial_\mu g) g^{-1} \in \mathcal{L}(G)$ (since $g(0)$ was arbitrary). $\square$

This result allows us to verify that the covariant derivative transforms correctly under finite transformations.

Theorem: Under a finite gauge transformation $\phi \mapsto D(g)\phi$, we have $D_\mu \phi \mapsto D(g) D_\mu \phi$.

Proof: The only main difficulty is deciding what to do with:

$$R(g A_\mu g^{-1} - (\partial_\mu g) g^{-1}) = R(g A_\mu g^{-1}) - R((\partial_\mu g) g^{-1}).$$

First, note that $R(g A_\mu g^{-1}) =$

$$\frac{d}{dt}\left(D\left(g e^{tA_\mu} g^{-1}\right)\right)\Big|_{t=0} = D(g)\frac{d}{dt}\left(D\left(e^{tA_\mu}\right)\right)\Big|_{t=0} D(g)^{-1}$$

by definition of R and using homomorphism property of D. The remaining derivative is just $R(A_\mu)$ by definition. Hence

$$R(g A_\mu g^{-1}) = D(g) R(A_\mu) D(g)^{-1}.$$

Similarly, using the curve we used in the previous proof, we have $R((\partial_\mu g) g^{-1}) = (\partial_\mu D(g)) D(g)^{-1}$. The rest of the proof is just a simple calculation. $\square$

Thus, it is also possible to verify gauge-invariance using finite methods.

We can also examine the transformation of the field-strength tensor in a finite regime.

Theorem: Under a gauge transformation, we have that $F_{\mu\nu} \mapsto g F_{\mu\nu} g^{-1}$.

Proof: Recalling the transformation law for A_μ and the definition of $F_{\mu\nu}$, this is straightforward but tedious. Need to use $\partial_\mu(g g^{-1}) = 0$. $\square$

A quicker way of getting the transformation law is:

Theorem: For any representation R, we have $R(F_{\mu\nu}) = [D_\mu, D_\nu]$.

Proof: First of all, we need to ask if this makes sense: is $F_{\mu\nu}$ in the Lie algebra? Indeed, it is because $[A_\mu, A_\nu]$ is in the Lie algebra, and $e^{A_\mu(x(t))}$ is a curve in the Lie group, which WLOG we can set up so that $\dot{x}^\nu(0) = 1$ is the only non-zero component of $\dot{x}(0)$. and $A_\mu = 0$ when $t = 0$. This then has tangent $\partial_\nu A_\mu$, and so the derivatives are in the Lie algebra. Now, we have $[D_\mu, D_\nu]\phi =$

$$(\partial_\mu + R(A_\mu))(\partial_\nu \phi + R(A_\nu)\phi) - (\partial_\nu + R(A_\nu))(\partial_\mu \phi + R(A_\mu)\phi)$$

Expanding and simplifying we find this is equal to

$$(\partial_\mu R(A_\nu) - \partial_\nu R(A_\mu) + R([A_\mu, A_\nu]))\phi.$$

Expand $A_\nu = A_\nu^a T^a$ in the basis of the Lie algebra. Then $\partial_\mu R(A_\nu) = (\partial_\mu A_\nu^a(x)) R(T^a)$ since T^a has no x dependence, and thus $\partial_\mu R(A_\nu) = R(\partial_\mu A_\nu)$. So done. $\square$

Now simply pick R to be the fundamental rep. Then under a gauge transformation, we have

$$F_{\mu\nu}\phi = [D_\mu, D_\nu]\phi \mapsto [D_\mu, D_\nu]g\phi = \tilde{F}_{\mu\nu}g\phi,$$

where $\tilde{F}_{\mu\nu}$ is the gauge-transformed field strength tensor. Then

$$[D_\mu, D_\nu]g\phi = g[D_\mu, D_\nu]\phi$$

since $D_\mu D(g)\phi = D(g)D_\mu\phi$, for any rep D. Thus it follows that

$$[D_\mu, D_\nu]\phi = g^{-1}\tilde{F}_{\mu\nu}g\phi,$$

and so $\tilde{F}_{\mu\nu} = gF_{\mu\nu}g^{-1}$, as we found before!

10.4 Some rep-specific examples

Example: Consider the *fundamental rep* $D(g)\phi = g\phi$. The infinitesimal rep corresponding to the fundamental rep can be deduced from the gauge transformation

$$\phi \mapsto D(g)\phi = g\phi = \phi + X\phi.$$

That is, we should use $R(X) = X$, the fundamental rep of the Lie algebra (hence both are aptly named!).

The gauge covariant derivative is

$$D_\mu\phi = \partial_\mu\phi + A_\mu\phi,$$

and so under a finite gauge transformation we have $D_\mu(g\phi) =$

$$\partial_\mu(g\phi) + (gA_\mu g^{-1} - (\partial_\mu g)g^{-1})g\phi = g(\partial_\mu + A_\mu)\phi = gD_\mu\phi,$$

as expected by the general theory.

Example: Consider the *adjoint rep* $D(g)\phi = g\phi g^{-1}$. The infinitesimal rep corresponding to the adjoint rep can be deduced from the gauge transformation

$$\phi \mapsto g\phi g^{-1} = (1+X)\phi(1-X) = \phi + [X, \phi].$$

Hence we should use $R(X) = [X, \cdot] = \mathrm{ad}_X$, the adjoint rep of the Lie algebra (again, both are aptly named!).

Again, we can check that $D_\mu\phi \mapsto g(D_\mu\phi)g^{-1}$ under a gauge transformation, as predicted by the general theory.

Example: In both the above examples, we get gauge-invariant kinetic terms for the matter fields. In the fundamental rep, we have $(D_\mu\phi, D^\mu\phi) \mapsto$

$$(gD_\mu\phi, gD^\mu\phi) = (g^\dagger g D_\mu\phi, D^\mu\phi) = (D_\mu\phi, D^\mu\phi),$$

since g is unitary. In the adjoint rep, we have $(D_\mu\phi, D^\mu\phi) \mapsto$

$$(gD_\mu\phi g^{-1}, gD^\mu\phi g^{-1}) = (g^\dagger g D_\mu\phi, D^\mu\phi(g^{-1})(g^{-1})^\dagger).$$

Again, this is gauge invariant since g is unitary (and so $g^{-1} = g^\dagger$ is also unitary).

10.5 The Killing form in gauge theory

From the transformation law for $F_{\mu\nu}$ above (i.e. $F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1}$), we notice that

$$\frac{1}{g^2}\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})$$

is gauge invariant. So why are we bothering with the Killing form at all?

In fact, this kinetic term is *proportional* to the Killing form. Indeed, any such gauge invariant inner product must be, and so we've actually treated things in their greatest generality above. The result is stated below.

Theorem: The Killing form, $\kappa(X, Y) = \mathrm{Tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y)$ is, up to scalar multiplication, the unique invariant inner product on a simple Lie algebra $\mathcal{L}(G)$. Recall, a bilinear form $B(X, Y)$ is *invariant* if

$$B([Z, X], Y) + B(X, [Z, Y]) = 0.$$

Proof: Not required in this course (though it can be proved by considering $B \circ \kappa^{-1}$ and using Schur's Lemma). □

The 'invariance' property is actually equivalent to gauge invariance here. Recall that $F_{\mu\nu}$ transforms infinitesimally as

$$F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1} = (1+X)F_{\mu\nu}(1-X) = F_{\mu\nu} + [X, F_{\mu\nu}].$$

Therefore, if $B(F_{\mu\nu}, F^{\mu\nu}$ is some kinetic inner product term, for gauge invariance is equivalent to

$$B(F_{\mu\nu}, F^{\mu\nu}) = B(F_{\mu\nu}, F^{\mu\nu}) + B([X, F_{\mu\nu}], F^{\mu\nu}) + B(F_{\mu\nu}, [X, F^{\mu\nu}]).$$

That is, gauge invariant is equivalent to gauge invariance!

Therefore, we've seen that *whatever* inner product term we write down, if it's gauge invariant it's proportional to the Killing form.

Sometimes we'd like to know the constant of proportionality. Clearly this is representation and group-dependent, since the Killing form depends on both of these. Let's see an example of how we'd calculate this:

Example: Consider the gauge invariant Lagrangian when the gauge group is $SU(N)$:

$$\mathcal{L} = \frac{1}{4g^2}\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu}).$$

Since it is gauge-invariant, it is proportional to the Killing form:

$$\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu}) = \lambda\kappa(F^{\mu\nu}, F_{\mu\nu}) = \lambda\mathrm{Tr}\left(\mathrm{ad}_{F^{\mu\nu}} \circ \mathrm{ad}_{F_{\mu\nu}}\right).$$

To calculate the constant of proportionality, begin by working in slightly greater generality. Let's work out $\kappa(X, Y)$ for $X, Y \in \mathcal{L}(GL(N))$.

A convenient basis for $\mathcal{L}(GL(N))$ is $(T_{ij})^{\alpha}{}_{\beta} = \delta^{\alpha}{}_{i}\delta_{\beta j}$. It's a quick exercise to work out the structure constants in this basis:

$$[T_{ij}, T_{kl}] = \delta_{jk}T_{il} - \delta_{li}T_{kj}.$$

Now writing $X = X^{ij}T_{ij}, Y = Y^{ij}T_{ij}$, we may compute

$$\mathrm{ad}_X(T_{ij}) = X^{li}T_{lj} - X^{jl}T_{il},$$

and so $\mathrm{ad}_Y \mathrm{ad}_X(T_{ij}) =$

$$(X^{rk}Y^{ki}\delta_{sj} + X^{ks}Y^{jk}\delta_{ri} - X^{ri}Y^{jk}\delta_{sk} - X^{js}Y^{ki}\delta_{rk})T_{rs}.$$

Thus calculating the trace directly, we have $\mathrm{Tr}(\mathrm{ad}_Y \circ \mathrm{ad}_X) = 2N\mathrm{Tr}(XY) - 2\mathrm{Tr}(X)\mathrm{Tr}(Y)$. To restrict to $\mathcal{L}(SU(N))$, just impose the fact the matrices must be traceless. Thus

$$\mathrm{Tr}\left(\mathrm{ad}_{F^{\mu\nu}} \circ \mathrm{ad}_{F_{\mu\nu}}\right) = 2N\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu}),$$

and it follows $\lambda = 1/2N$. Thus

$$\mathcal{L} = \frac{1}{4g^2}\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu}) = \frac{1}{8Ng^2}\kappa(F^{\mu\nu}, F_{\mu\nu}).$$

---

11 Applications to hadronic physics

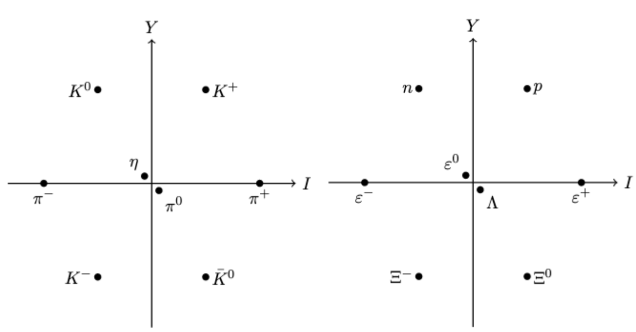
11.1 Isospin and hypercharge

Many years ago, people discovered lots and lots of types of particles. Each particle carried a quantum number called *hypercharge* and another called *isospin* (it doesn't matter what these are in this course). It was found that the sum of the hypercharges before and after any interaction were equal, and the same for isospins. Thus both correspond to *conserved charges*.

Definition: In hadronic physics, *hypercharge* and *isospin* are conserved charges, denoted Y and I respectively.

Since we know how quantum mechanics works (see Symmetries in Quantum Theory chapter), we know that the conserved quantities of a system form a *Lie algebra*, which arises as the symmetry of a Lie group. What algebra and group work for isospin and hypercharge?

Plotting the isospin and hypercharge of the 8 lightest *mesons* and the 8 lightest *baryons*, physicists found the diagrams:



These diagrams looked suspiciously like the weight diagrams of $\underline{8}$, the 8-dimensional irrep of $A_2 = \mathcal{L}(SU(3))$.

Theorem: The weight diagram of $\underline{8}$ reproduces the figures on the left.

Proof: Recall that with the basis of the CSA:

$$H^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

the root system of A_2 is given by

$$\{(2,-1)^T, (1,1)^T, (-1,2)^T, (-2,1)^T, (-1,-1)^T, (1,-2)^T, \mathbf{0}\}.$$

Recall also from earlier in the course that *adjoint* rep was 8-dimensional. So just plot the roots of A_2.

This gives something similar to the diagram, but a bit wonky. It suggests we change basis for the Cartan subalgebra:

$$I = \frac{1}{2}H^1, \quad Y = \frac{1}{3}(H^1 + 2H^2).$$

Indeed, plotting the weight diagram in this new basis, we get the result. □

This suggests that the eight lightest baryons and mesons live in the representation space of the adjoint rep of A_2.

However, physicists kept finding more particles! Continuing to add and add to the weight diagrams, we actually found that the right representation for all of the baryons was:

$$\underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10}$$

and the right representation for the all of the mesons was:

$$\underline{1} \oplus \underline{8}.$$

There is an immediate question to ask here: why not other reps of $\mathcal{L}(SU(3))$? Why not, say $\underline{6}$? Nature seems to have been very picky...

The Quark Model: The *quark model* uses the results we derived earlier in the course:

$$\underline{1} \oplus \underline{8} = \underline{3} \otimes \underline{\bar{3}},$$
$$\underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10} = \underline{3} \otimes \underline{3} \otimes \underline{3}.$$

This suggests that the quantum state of a meson lies in the rep space of $\underline{3} \otimes \underline{\bar{3}}$ and the quantum state of a baryon lives in the rep space of $\underline{3} \otimes \underline{3} \otimes \underline{3}$.

If we consider a meson to be composed of an *anti-quark* and a *quark*, and a baryon to be composed of three quarks, where *quarks* are things living in the rep space of $\underline{3}$, we immediately reproduce all of the hadronic physics we've seen above!