Symmetries, Fields and Particles

based on a course given by   written up by

Nick Dorey   Josh Kirklin

Please send errors and suggestions to jjvk2@cam.ac.uk.

Unlectured. I may sometimes include sections that were not lectured; these will appear like so.

Contents

1 Introduction

1.1 Groups

Definition. A *symmetry* is a transformation of a set of dynamical variables, under which the forms of all physical laws are invariant.

Example. Consider rotations in 3D space:

$$\mathbf{x} \in \mathbb{R}^3 \mapsto \mathbf{x}' = M \cdot \mathbf{x} \in \mathbb{R}^3$$

Here, M is a 3×3 real special orthogonal matrix:

$$\det M = 1 \qquad\qquad M^{\intercal}M = \mathbb{1}$$

Newton's second law is invariant under these transformations, so rotations are a symmetry of Newtonian mechanics.

$$\mathbf{F} = m\mathbf{a} \longrightarrow \mathbf{F}' = m\mathbf{a}'$$

Group theory provides us with a useful mathematical formulation of symmetries.

Definition. A *group* $(G, \cdot)$ is a set G equipped with a multiplication operator $\cdot$ with the following properties:

    (i) *Closure*: $g_1, g_2 \in G \implies g_1 \cdot g_2 \in G$

    (ii) *Identity*: $\exists\, e \in G \,\mathrm{s.\,t.}\, e \cdot g = g = g \cdot e \,\forall\, g \in G$

    (iii) *Inverses*: $\forall\, g \in G \,\exists\, g^{-1} \in G \,\mathrm{s.\,t.}\, g \cdot g^{-1} = g^{-1} \cdot g = e$

    (iv) *Associativity*: $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3 \,\forall\, g_1, g_2, g_3 \in G$

The set in question may be either finite or infinite. We say that $(G, \cdot)$ is *abelian* or *commutative* if $g_1 \cdot g_2 = g_2 \cdot g_1 \,\forall\, g_1, g_2 \in G$.
Often we will omit $\cdot$ and just write $G = (G, \cdot)$ and $g_1 g_2 = g_1 \cdot g_2$.

Example. Rotations in 3D space form a non-abelian group under composition[1]. We call the group of 3×3 real special orthogonal matrices $SO(3)$. The elements of $SO(3)$ depend on three parameters: $\hat{\mathbf{n}} \in S^2, \theta \in [0, \pi]$.

1.2 Manifolds

Definition. An *n-manifold* is a space which locally looks like $\mathbb{R}^n$.

Example. The 2-sphere S^2 is a 2-manifold.

Let $\mathcal{M}$ be a smooth manifold of dimension D, and let p be a point in $\mathcal{M}$. Since $\mathcal{M}$ is locally like $\mathbb{R}^D$, we can introduce a set of coordinates $\{x^i\}, i = 1, \ldots, D$ into an open subset of the manifold, with origin at p.

Definition. The tangent space to $\mathcal{M}$ at p, denoted $\mathcal{T}_p(\mathcal{M})$, is the D-dimensional vector space spanned by the differential operators $\{\frac{\partial}{\partial x^j}\}, j = 1, \ldots, D$.
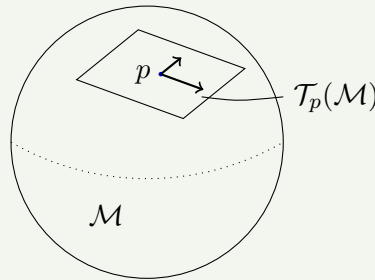
Suppose $f : \mathcal{M} \to \mathbb{R}$ is a function on $\mathcal{M}$, and let $V = V^i \frac{\partial}{\partial x^i} \in \mathcal{T}_p(\mathcal{M})$. The action of V on f is defined as follows:

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

Consider a smooth curve on $\mathcal{M}$:

$$C \; : \; \begin{array}{ccc} \mathbb{R} & \to & \mathcal{M} \\ t & \mapsto & x^i(t) \end{array}$$

Suppose this curve goes through p at $t = 0$. We can associate a tangent vector at p with C in the following way:

$$V_C = \dot{x}^i(0) \frac{\partial}{\partial x^i} \in \mathcal{T}_p(\mathcal{M}) \text{ where } \dot{x}^i = \frac{\mathrm{d}x^i}{\mathrm{d}t}$$
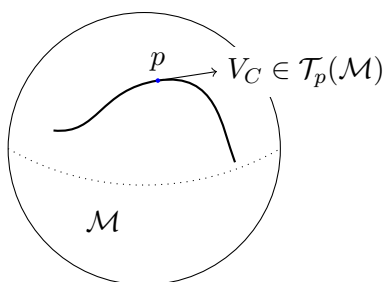
If we let V_C act on a function f we can see that it is simply the total derivative of f along the curve at p:

$$V_C(f) = \dot{x}^i \frac{\partial}{\partial x^i} f(x)\bigg|_{x=0} = \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\partial f}{\partial x^i}\bigg|_{x=0} = \frac{\mathrm{d}f}{\mathrm{d}t}\bigg|_p$$

In physics, often C is the trajectory of a particle or other dynamical object, and V_C represents its velocity vector.

1.3 Lie groups

> **Definition.** A *Lie group G* is a group which is also a smooth manifold. The group and manifold structures must be compatible, so for example the product $\cdot : G \times G \to G$ and inverse operations must be smooth maps.

We will see that G is (almost) completely determined by its behaviour "near" the identity e.

We denote the manifold of G by $\mathcal{M}(G)$. The dimension of G, denoted $\dim G$, is equivalent to the dimension of $\mathcal{M}(G)$.

Let's introduce a set of coordinates $\{\theta_i\}, i = 1, \ldots, D = \dim G$ in some local neighbourhood of the identity $e \in G$, with e at the origin. The group elements in this neighbourhood depend continuously on the $\{\theta_i\}$, and we can write $g = g(\theta)$.

Multiplication of two of these elements corresponds to a smooth map $G \times G \to G$:

$$g(\theta)g(\theta') = g(\phi) \in G$$

1.4 Matrix groups

Let $\mathrm{Mat}_n(F)$ denote the set of $n \times n$ matrices over a field F. In these notes we will henceforth only consider $F = \mathbb{R}$ or $\mathbb{C}$. Matrix multiplication is closed, associative, and there exists an identity, but $\mathrm{Mat}_n(F)$ is *not* a group under multiplication because not all elements have inverses. The "largest" multiplicative group contained in $\mathrm{Mat}_n(F)$ is the *general linear group*, containing all invertible $n \times n$ matrices:

$$GL(n, F) = \{M \in \mathrm{Mat}_n(F) \text{ s. t. } \det M \neq 0\}$$

If we require the determinant to be equal to 1, we obtain the *special linear group*:

$$SL(n, F) = \{M \in \mathrm{Mat}_n(F) \text{ s. t. } \det M = 1\}$$

Note that the closure of and existence of inverses in $SL(n, F)$ follows from the fact that multiplication commutes with taking the determinant, i.e. $\det(M_1 M_2) = \det M_1 \det M_2$.

We can see that these are groups; slightly less obviously they are also Lie groups.

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

> **Definition.** A *subgroup H* of a group G is a subset of G which is also a group. If H is also a smooth submanifold, then we say that H is a *Lie subgroup*.

Of interest are the *orthogonal groups*, which are subgroups of $GL(n, \mathbb{R})$:

$$O(n) = \{M \in GL(n, \mathbb{R}) \text{ s. t. } M^\mathsf{T}M = \mathbb{1}\}$$

Orthogonal transformations preserve the lengths of vectors.

$$|\mathbf{v}'|^2 = \mathbf{v}'^\mathsf{T}\mathbf{v}' = \mathbf{v}^\mathsf{T}M^\mathsf{T}M\mathbf{v} = \mathbf{v}^\mathsf{T}\mathbf{v} = |\mathbf{v}|^2$$

We also have $(\det M)^2 = \det(M^\mathsf{T}M) = \det \mathbb{1} = 1$, so $\det M = \pm 1$. The two cases correspond to the two connected components of $O(n)$. Elements in the positive determinant component correspond to rotations in n-dimensional space, and elements in the negative determinat component correspond to a reflection composed with rotations. The fact that the identity does not contain a reflection allows us to define the *special orthogonal groups*:

$$SO(n) = \{M \in O(n) \text{ s. t. } \det M = +1\}$$

It can be checked that $\dim O(n) = \dim SO(n) = \frac{1}{2}n(n + 1)$.
It is easy to show that if λ is an eigenvalue of an orthogonal matrix M, then we have the following:

(i) λ^* is also an eigenvalue of M.

(ii) $|\lambda|^2 = 1$

> **Example.** Let $M \in SO(2)$. From the above, we can deduce that the eigenvalues of M are $\lambda = e^{\pm i\theta}$ for some $\theta \in \mathbb{R}$, and we can write:
>
> $$M(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
>
> These matrices are uniquely specified by θ if we restrict $0 \le \theta < 2\pi$. Since $M(\theta) = M(\theta + 2\pi)$, we can identify the manifold of $SO(2)$ as the circle:
>
> $$\mathcal{M}(SO(2)) \simeq S^1$$

> **Example.** Let $M \in SO(3)$. From the above properties, we can see that the eigenvalues of M must be $\lambda = 1, e^{\pm i\theta}$ for some $\theta \in \mathbb{R}$. Let $\hat{\mathbf{n}}$ be a normalised eigenvector for $\lambda = 1$. We have $M\hat{\mathbf{n}} = \hat{\mathbf{n}}$; $\hat{\mathbf{n}}$ is parallel to the axis of rotation of M. It can be shown that a general group element of $SO(3)$ can be written in the following form:
>
> $$M(\hat{\mathbf{n}}, \theta)_{ij} = \cos\theta\delta_{ij} + (1 - \cos\theta)n_i n_j - \sin\theta\epsilon_{ijk}n_k$$
>
> Note that $M(\hat{\mathbf{n}}, 2\pi - \theta) = M(-\hat{\mathbf{n}}, \theta)$. We can still specify all group elements if we restrict $0 \le \theta \le \pi$. Note also that $M(\hat{\mathbf{n}}, 0) = \mathbb{1} \, \forall \, \hat{\mathbf{n}}$.

Consider $\mathbf{w} = \theta\hat{\mathbf{n}}$. Values of $\mathbf{w}$ in the 3-ball of radius π specify the elements of $SO(3)$.

$$B_3 = \{\mathbf{w} \in \mathbb{R}^3 \,\text{s.\,t.}\, |\mathbf{w}| \leq \pi\} \subset \mathbb{R}^3 \qquad\qquad \partial B_3 = \{\mathbf{w} \in \mathbb{R}^3 \,\text{s.\,t.}\, |\mathbf{w}| = \pi\} \simeq S^2$$

If we identify $\theta = \pi$ with $\theta = -\pi$, then we uniquely specify points in $SO(3)$. Hence, $\mathcal{M}(SO(3))$ is defined from B_3 by indentifying antipodal points on the boundary ∂B_3.

$$\mathcal{M}(SO(3)) = B_3/\sim \ \text{ where } \mathbf{w} \sim \mathbf{w}' \iff \mathbf{w} = -\mathbf{w}' \text{ and } |\mathbf{w}| = \pi$$

This manifold has the following properties:

- It has no boundary.

- It is connected (a path exists between any two points) but not simply connected (not all closed loops can contract to a point).

- Its fundamental group is equivalent to the integers mod 2: $\Pi_1(SO(3)) = \mathbb{Z}_2$.

- It is compact (closed and bounded).

Note that the fact that orthogonal transformations preserve the lengths of vectors is equivalent to saying that they preserve the Euclidean metric on $\mathbb{R}^n$, i.e. $g = \mathbb{1}$. If we substitute the Euclidean metric for a more general metric of signature (p, q):

$$\eta = \begin{pmatrix} \mathbb{1}_p & 0 \\ 0 & -\mathbb{1}_q \end{pmatrix}$$

then we obtain the group $O(p, q)$:

$$O(p, q) = \{M \in GL(n, \mathbb{R}) \,\text{s.\,t.}\, M^{\mathsf{T}}\eta M = \eta\}$$

Example. The Lorentz group is $O(3, 1)$.

These groups are non-compact.

If we replace $\mathbb{R}$ with $\mathbb{C}$ and transposition with Hermitian conjugation in the above, we can obtain the *unitary groups*:

$$U(n) = \{U \in GL(n, \mathbb{C}) \,\text{s.\,t.}\, U^{\dagger}U = \mathbb{1}\}$$

Note that now, since the U are complex matrices, we have $\det U) = e^{i\delta}$ where $\delta \in \mathbb{R}$. If we enforce that $\det U = 1$, we obtain the *special unitary groups*:

$$SU(n) = \{U \in U(n) \,\text{s.\,t.}\, \det U = 1\}$$

Definition. Two Lie groups G and G' are *isomorphic*, denoted $G \simeq G'$, if there exists a bijective smooth map J from one group to the other, and this map preserves the group structure, i.e.:

$$J(g_1 g_2) = J(g_1)J(g_2)$$

Example. Let's examine $U(1) = \{z \in \mathbb{C} \,\mathrm{s.t.}\, |z|^2 = 1\}$. From this definition we see that elements in $U(1)$ take the form $z = e^{i\theta}, \theta \in [0, 2\pi)$, and we can immediately observe that $U(1) \simeq SO(2)$, with isomorphism function given by:

$$e^{i\theta} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Example. It can be shown that a general form for all $U \in SU(2)$ is:

$$U = a_0 \mathbb{1} + i\mathbf{a} \cdot \sigma$$

where the $\sigma_i, i = 1, 2, 3$ are the Pauli matrices. Importantly, we must have $a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1$, which shows that $\mathcal{M}(SU(2)) = S^3$.

1.5 Lie algebras

Definition. A *Lie algebra* $\mathfrak{g}$ is a vector space over a field F, equipped with a *bracket* $[\,,\,]$: $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ with the following properties:

(i) *Anti-symmetry*: $[X, Y] = -[Y, X]$

(ii) *Linearity*: $[\alpha X + \beta Y, Z] = \alpha[X, Z] + \beta[Y, Z]$

(iii) *Jacobi identity*: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

Example. If a vector space V has an associative linear product $*$, then we can get a Lie algebra by setting:

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

This bracket is known as the *commutator*. The commutator can provide lots of examples of Lie algebras. For example, we can let V be a vector space of matrices, and let $*$ be matrix multiplication.

The dimension of a Lie algebra $\mathfrak{g}$, denoted $\dim \mathfrak{g}$, is given by the dimension of its vector space. If we choose a basis $\{T^a\}$, where $a = 1, \ldots, n = \dim \mathfrak{g}$, we can expand any $X \in \mathfrak{g}$ in terms of its components:

$$X = X_a T^a$$

Definition. If we apply the bracket to any two basis vectors, we obtain the *structure constants* f_c^{ab}:

$$[T^a, T^b] = f_c^{ab} T^c$$

Note that the structure constants are dependent on the basis chosen. From the properties of the bracket, the structure constants obey the following:

$$f_c^{ab} = -f_c^{ba}$$
$$f_c^{[ab} f_e^{d]c} = 0$$

Definition. Two Lie algebras $\mathfrak{g}$ and $\mathfrak{g}'$ are *isomorphic* if there exists a bijective linear map $f : \mathfrak{g} \to \mathfrak{g}'$ that preserves the bracket:

$$[f(X), f(Y)] = f([X, Y])$$

We are concerned with classifying Lie algebras up to isomorphism.

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

Definition. An *ideal* $\mathfrak{h}$ of $\mathfrak{g}$ is a subalgebra such that $[X, Y] \in \mathfrak{h} \, \forall \, X \in \mathfrak{g}, Y \in \mathfrak{h}$.

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

$$\mathfrak{h} = \{0\}, \mathfrak{h} = \mathfrak{g}$$

Example. The *derived algebra*:

$$\mathfrak{i}(\mathfrak{g}) = \{[X, Y] \, \text{s.t.} \, X, Y \in \mathfrak{g}\}$$

is an ideal.

Example. The *center*:

$$J(\mathfrak{g}) = \{X \in \mathfrak{g} \, \text{s.t.} \, [X, Y] = 0 \, \forall \, Y \in \mathfrak{g}\}$$

is an ideal.

Definition. A Lie algebra is said to be *abelian* if its bracket is always equal to 0.

If $\mathfrak{g}$ is abelian, then $\mathfrak{i}(\mathfrak{g}) = \{0\}$ and $J(\mathfrak{g}) = \mathfrak{g}$.

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

If $\mathfrak{g}$ is simple, then $\mathfrak{i}(\mathfrak{g}) = \mathfrak{g}$ and $J(\mathfrak{g}) = \{0\}$.

2 Lie algebras from Lie groups

2.1 $\mathcal{L}(G)$

Let G be a Lie group of dimension D and introduce coordinates $\{\theta^i\}, i = 1, \ldots, D$ in a region containing the identity e at $\theta = 0$. We have that $\mathcal{T}_e(G)$ is a vector space of dimension D. We will define a bracket $[\,,\,] : \mathcal{T}_e(G) \times \mathcal{T}_e(G) \to \mathcal{T}_e(G)$ which will show that $\mathcal{L}(G) = (\mathcal{T}_e(G), [\,,\,])$ is a Lie algebra.

This is easiest do for matrix Lie groups $G \subset \mathrm{Mat}_n(F)$. We can map tangent vectors to matrices:

$$\mathcal{T}_e(G) \ni v^i \frac{\partial}{\partial \theta^i} \mapsto v^i \frac{\partial g(\theta)}{\partial \theta^i}\bigg|_{\theta=0} \in \mathrm{Mat}_n(F)$$

With this mapping we identify $\mathcal{T}_e(G)$ with the subspace of $\mathrm{Mat}_n(F)$ spanned by $\{\frac{\partial g(\theta)}{\partial \theta^i}\}, i = 1, \ldots, D$. Now we can define the obvious bracket, the *matrix commutator*:

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

This obeys most of the properties of a Lie bracket. The only one that is slightly non-trivial is that the Lie algebra is closed under the bracket, i.e. $[X, Y] \in \mathrm{span}\{\frac{\partial g(\theta)}{\partial \theta^i}\}$. We will show this now.

To do so we use the correspondence between curves and tangent vectors shown earlier. Suppose C is a curve with that passes through the identity e with tangent vector $X \in \mathcal{T}_e(G) = \mathcal{L}(G)$:

$$C : t \mapsto g(t) \in G \qquad\qquad g(0) = e = \mathbb{1}$$

Since $\frac{\mathrm{d}g(t)}{\mathrm{d}t}\big|_{t=0}$, we can expand $g(t)$ as a Taylor series near $t = 0$ in the following way:

$$g(t) = \mathbb{1} + tX + O(t^2)$$

Now suppose we have two elements $X_1, X_2 \in \mathcal{L}(G)$ and corresponding curves C_1, C_2:

$$C_1 : t \mapsto g_1(t) \in G \qquad\qquad g_1(0) = g_2(0) = \mathbb{1}$$
$$C_2 : t \mapsto g_2(t) \in G \qquad\qquad \dot{g}_1(0) = X_1, \dot{g}_2(0) = X_2$$

Now expanding to second order, we can write:

$$g_1(t) = \mathbb{1} + X_1 t + W_1 t^2 + O(t^3)$$
$$g_2(t) = \mathbb{1} + X_2 t + W_2 t^2 + O(t^3)$$

for some $W_1, W_2 \in \mathrm{Mat}_n(F)$. In order to show that $\mathcal{L}(G)$ is closed, we need to find a new curve with tangent vector $[X_1, X_2]$. To that end, let us define $h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t)$. Obviously $h(0) = \mathbb{1}$, so we can set $h(t) = \mathbb{1} + h_1 t + h_2 t^2 + O(t^3)$. If we evaluate $g_2(t)g_1(t)h(t) = g_1(t)g_2(t)$ term by term, we find that W_1, W_2 cancel from our equations and we obtain:

$$h_1 = 0, h_2 = [X_1, X_2]$$

So define a new curve C_3:

$$C_3 : s \mapsto g_3(s) = h(\sqrt{s}) \in G$$

We have:

$$g_3(s) = \mathbb{1} + s[X_1, X_2] + O(s^{3/2})$$

Thus the tangent vector to this curve at the identity is $[X_1, X_2]$. So we have shown that $[X_1, X_2] \in \mathcal{L}(G)$ and thus that $\mathcal{L}(G)$ is closed. Therefore, $\mathcal{L}(G)$ is a real[2] Lie algebra of dimension D.

---

[2]That is, it has real structure constants

Example. Consider $G = SO(2)$. A curve on G can be written:

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

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

where we have assumed $g(0) = \mathbb{1}$. Therefore we can deduce the Lie algebra associated with G:

$$\mathcal{L}(SO(2)) = \left\{ \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix}, c \in \mathbb{R} \right\}$$

Example. More generally, consider $G = SO(n)$ or $O(n)$, and let $g(t) = R(t)$ be a curve in G that goes through the identity at $t = 0$. We have:

$$R^{\mathsf{T}}(t)R(t) = \mathbb{1}$$

Differentiating both sides with respect to t gives:

$$\dot{R}^{\mathsf{T}}(t)R(t) + R^{\mathsf{T}}(t)\dot{R}(t) = 0$$

If we now set $t = 0$, and use $R(0) = \mathbb{1}$, we obtain:

$$X^{\mathsf{T}} + X = 0$$

where $\dot{R}(0) = X \in \mathcal{L}(G)$. In other words, X is antisymmetric. Hence:

$$\mathcal{L}(SO(n)) = \mathcal{L}(O(n)) = \{X \in \mathrm{Mat}_n(\mathbb{R}) \,\text{s.\,t.}\, X^{\mathsf{T}} = -X\}$$

By examining the number of independent components in an antisymmetric matrix, we can deduce that $\dim G = \frac{1}{2}n(n-1)$.

Example. Consider $G = SU(n)$. As before let $g(t) = U(t) \in SU(n)$ with $U(0) = \mathbb{1}$.

$$U^{\dagger}(t)U(t) = \mathbb{1} \implies Z^{\dagger} + Z = 0$$

where $Z = \dot{U}(0)$, i.e. Z is anti-Hermitian. But we have another constraint that will be relevant:

$$\det U(t) = 1$$

Let's Taylor expand near $t = 0$:

$$U(t) = \mathbb{1} + tZ + O(t^2)$$
$$\implies \det U(t) = 1 + \mathrm{tr}\, Z t + O(t^2)$$

Thus we must have $\operatorname{tr} Z = 0$. Therefore:

$$\mathcal{L}(SU(n)) = \{Z \in \operatorname{Mat}_n(\mathbb{C}) \text{ s. t. } Z^\dagger = -Z, \operatorname{tr} Z = 0\}$$

There are $2n^2 - n^2 - 1 = n^2 - 1$ independent components of a traceless anti-Hermitian complex matrix, so we have $\dim G = n^2 - 1$.

We did not consider the $\det R = 1$ case for $\mathcal{L}(SO(n))$, but if we had we would have again obtained that $\operatorname{tr} R = 0$. Note however that R being antisymmetric implies that R is traceless, so we do not need to additionally specify this.

Example. From above, we have $\mathcal{L}(SU(2)) = \{2 \times 2 \text{ traceless anti-Hermitian matrices}\}$. Consider the Pauli matrices:

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

These are traceless and Hermitian. Thus we can choose the following basis for for $\mathcal{L}(G)$:

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

where the factor of $-\frac{1}{2}$ is for future convenience. The Pauli matrices obey the following convenient identity:

$$\sigma_a \sigma_b = \delta_{ab} \mathbb{1} + i \epsilon_{abc} \sigma_c$$

With this we can calculate the bracket of any two basis elements:

$$[T^a, T^b] = -\frac{1}{2} [\sigma_a, \sigma_b] = -\frac{1}{2} i \epsilon_{abc} \sigma_c = \epsilon_{abc} T^c$$

So the structure constants are $f_c^{ab} = \epsilon_{abc}$.

Example. We have $\mathcal{L}(SO(3)) = \{3 \times 3 \text{ real antisymmetric matrices}\}$. It is convenient to choose the following basis:

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

The reason this basis is convenient is that we can write $(T^a)_{bc} = -\epsilon_{abc}$. It is not difficult to then show that $[T^a, T^b] = \epsilon_{abc} T^c$, so the structure constants are $f_c^{ab} = \epsilon_{abc}$.

Since we can choose bases in which $SO(3)$ has the same structure constants as $SU(2)$, we have $\mathcal{L}(SO(3)) \simeq \mathcal{L}(SU(2))$, despite the fact that $SO(3) \not\simeq SU(2)$.

2.2 Natural maps on Lie groups

Definition. Let G be a Lie group. For each $h \in G$ we have two smooth maps L_h, R_h, called *left translation* and *right translation* respectively:

$$
\begin{aligned}
L_h \;&:\; G \;\rightarrow\; G & R_h \;&:\; G \;\rightarrow\; G \\
&\quad\; g \;\mapsto\; hg & &\quad\; g \;\mapsto\; gh
\end{aligned}
$$

In what follows we will focus on L_h, but similar results apply for R_h.

Lemma 1. *L_h is a diffeomorphism[3].*

Proof. We have $L_h(h^{-1}g) = g$ for all $g \in G$, so L_h is surjective. Suppose $g, g' \in G$ and $L_h(g) = L_h(g')$. Then $hg = hg' \implies g = g'$, so L_h is injective. Thus L_h is a bijection, and since it is defined as a product of two Lie group elements, it is a smooth one. Furthermore, we have that the inverse map $(L_h)^{-1} = L_{h^{-1}}$ is also smooth. $\qquad\square$

Here we will sketch the consequences of this proposition. Introduce a set of coordinates $\{\theta^i\}, i = 1, \ldots, D$, with e at $\theta = 0$. Let $g = g(\theta) \in G$ and $g' = g(\theta') = L_h(g) = hg(\theta)$. From this we see that L_h is specified by $\theta'^i = \theta'^i(\theta)$. Since L_h is a diffeomorphism, the Jacobian matrix exists:

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

and it must be invertible, i.e. $\det J \neq 0$.

Left translation L_h induces a map L_h^* from tangent vectors at $g \in G$ to tangent vectors at $L_h(g) = hg \in G$.
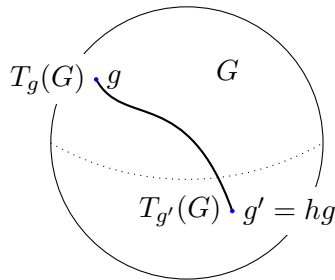
We define this map in the following way:

$$
\begin{aligned}
L_h^* \;&:\; \mathcal{T}_g(G) \;&\rightarrow\;& \mathcal{T}_{hg}(G) \\
& \quad V = V^i \tfrac{\partial}{\partial \theta^i} \;&\mapsto\;& V' = V'^i \tfrac{\partial}{\partial \theta'^i}
\end{aligned}
\quad \text{where } V'^i = J^i_j(\theta) V^j
$$

Definition. A *vector field* V assigns a tangent vector to each point $g \in G$:

$$
V(g) \in \mathcal{T}_g(G)
$$

---

[3]i.e. a smooth bijection with a smooth inverse.

Starting with a non-zero tangent vector at the identity $W \in \mathcal{T}_e(G)$, we can define a vector field by:

$$V(g) = L_g^*(W)$$

This vector field is smooth, and since J is invertible it is non-vanishing. So if we start with a basis $\{W_a\}, a = 1, \ldots, D$ of the tangent space at the identity, then we obtain D independent non-vansishing vector fields $V_a(g) = L_g^*(W_a)$. Although it may not seem so at first, the presence of these vector fields is a very strong constraint.

> **Example.** The so-called "Hairy ball theorem" says that a smooth vector field defined on S^2 must have at least one zero. Thus we cannot have a Lie group whose underlying manifold is the 2-sphere. So suppose G is a compact Lie group of dimension 2. Then we must have $G = T^2 = S^1 \times S^1$, and $G \simeq U(1) \times U(1)$. This is the only Lie group of dimension 2.

Now consider matrix Lie groups.

𝔏emma 2. *Suppose we have a matrix Lie group $G \subset \mathrm{Mat}_n(F)$, and $h \in G, X \in \mathcal{L}(G)$. Then $L_h^*(X) = hX \in \mathcal{T}_h(G)$.*

Proof. Let C be a curve with tangent vector X at the identity:

$$
\begin{array}{rccc}
C & : & \mathbb{R} & \to & G \\
 & & t & \mapsto & g(t)
\end{array}
\qquad\qquad g(0) = e \qquad\qquad \dot{g}(0) = X
$$

Near $t = 0$ we can write:
$$g(t) = \mathbb{1} + tX + O(t^2)$$

Now define a new curve:

$$
\begin{array}{rccc}
C_1 & : & \mathbb{R} & \to & G \\
 & & t & \mapsto & h(t) = hg(t)
\end{array}
$$

We have $h(0) = h$ and $\dot{h}(0) = h\dot{g}(0) = hX$, and hence near $t = 0$:

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

Thus $hX \in \mathcal{T}_h(G)$.   □

𝔠orollary. *Given some smooth curve in G:*

$$
\begin{array}{rccc}
C & : & \mathbb{R} & \to & G \\
 & & t & \mapsto & g(t)
\end{array}
\qquad\qquad \dot{g}(t) \in \mathcal{T}_{g(t)}(G)
$$

we can deduce that:
$$g^{-1}(t)\dot{g}(t) = L_{g^{-1}(t)}^*(\dot{g}(t)) \in \mathcal{T}_e(G) = \mathcal{L}(G)$$

Conversely, given $X \in \mathcal{L}(G)$, the following yields a curve in G:

$$g^{-1}(t)\dot{g}(t) = X \qquad\qquad\qquad g(0) = \mathbb{1} \qquad\qquad\qquad (*)$$

Definition. The *exponential* of a matrix $M \in \text{Mat}_n(F)$ is given by the following:

$$\text{Exp}\,M = \sum_{l=0}^{\infty} \frac{1}{l!} M^l \in \text{Mat}_n(F)$$

We can solve $(*)$ by setting $g(t) = \text{Exp}(tX)$.

Proof. We immediately have $g(0) = \text{Exp}(0) = \mathbb{1}$. Also:

$$g(t) = \text{Exp}(tX) = \sum_{l=0}^{\infty} \frac{1}{l!} t^l X^l$$

$$\implies \dot{g}(t) = X \left(\sum_{l=0}^{\infty} \frac{1}{l!} t^l X^l \right) = \text{Exp}(tX)X = g(t)X$$

$$\implies g^{-1}(t)\dot{g}(t) = X$$

$\square$

A useful identity that holds for all $X \in \text{Mat}_n(F)$ is:

$$\det(\text{Exp}\,X) = \exp(\text{tr}(X))$$

Lemma 3. *With a suitable choice of range J the set $S_{X,J} = \{g(t) = \text{Exp}(tX) \,\forall\, t \in J \subset \mathbb{R}\}$ is an abelian Lie subgroup of G with dimension 1, and $S_{X,J} \simeq U(1)$ or $(\mathbb{R}, +)$.*

Proof. (sketch) We have:

$$g(t_1)(t_2) = g(t_1 + t_2) = g(t_2)g(t_1)$$

so $S_{X,J}$ is abelian and closed. Also we have an identity $g(0) = \mathbb{1}$ and inverses $(g(t))^{-1} = g(-t)$. Associativity is inherited. Now there are two cases to consider:

1. Compact case: there is some value of $t_1 \neq 0$ such that $g(t_1) = \mathbb{1}$. Choose the least such t_1 and set $J = [0, t_1)$.
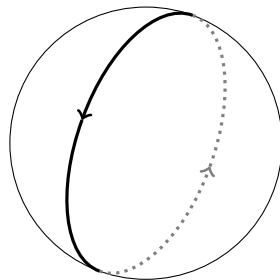
   Then we see that $S_{X,J} \simeq U(1)$ with isomorphism $g(t) \mapsto \exp(2\pi i t/t_1)$.

2. Non-compact case: for all $t \neq 0$, $g(t) \neq \mathbb{1}$. Then set $J = \mathbb{R}$, and we see that $S_{X,J} \simeq (\mathbb{R}, +)$ with isomorphism $g(t) \mapsto t$.

$\square$
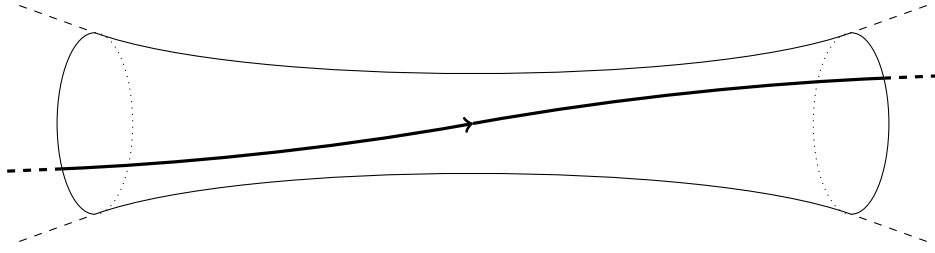
2.3 Reconstructing G from $\mathcal{L}(G)$

> **Definition.** Setting $t = 1$ defines the *exponential map*:
>
> $$\text{Exp} \ : \ \begin{array}{ccc} \mathcal{L}(G) & \to & G \\ X & \mapsto & \text{Exp}\,X \end{array}$$

Although we will not prove it here, in some neighbourhood of the identity, Exp is a bijective map. So, given some $X, Y \in \mathcal{L}(G)$, we have $g_X = \text{Exp}(X), g_Y = \text{Exp}(Y) \in G$, and we expect $g_X g_Y = g_Z = \text{Exp}(Z)$ for some $Z \in \mathcal{L}(G)$. Z is determined using the *Baker-Campbell-Haussdorf formula*:

$$\text{Exp}(X)\,\text{Exp}(Y) = \text{Exp}(Z) \implies Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}\left([X,[X,Y]] - [Y,[X,Y]]\right) + \cdots \quad \text{(BCH)}$$

So $\mathcal{L}$ completely determines G in some neighbourhood of the identity.
Note that Exp is not globally bijective; in particular:

- It is not surjective when G is disconnected.

- It is not injective when G has a $U(1)$ subgroup.

> **Example.** Consider $G = O(n)$. We have shown above that:
>
> $$\mathcal{L}(O(n)) = \{X \in \text{Mat}_n(\mathbb{R}) \,\text{s.\,t.}\, X^\mathsf{T} = -X\}$$
>
> X being antisymmetric implies that it is traceless, and hence that $\det(\text{Exp}\,X) = \exp(\text{tr}\,X) = 1$. Therefore $\text{Exp}(X) \in SO(n) \subsetneq O(n)$, so Exp is not surjective.

More generally, the image of $\mathcal{L}(G)$ under Exp is the connected component of the identity in G.

> **Example.** Consider $G = U(1)$. We have $\mathcal{L}(G) = \mathbb{I}$, the imaginary numbers. But then $\text{Exp}\,x = \text{Exp}(x + 2\pi i)$, so Exp is not injective.

2.4 $SU(2)$ **vs** $SO(3)$

We have already seen that $\mathcal{L}(SU(2)) \simeq \mathcal{L}(SO(3))$. It is possible to construct a double-covering[4] d from $SU(2)$ to $SO(3)$ in the following way:

$$
d \;:\; \begin{matrix} SU(2) & \to & SO(3) \\ A & \mapsto & d(A) \end{matrix} \quad \text{where } d(A)_{ij} = \frac{1}{2}\operatorname{tr}\!\left(\sigma_i A \sigma_j A^\dagger\right)
$$

We have $d(A) = d(-A)$. This map provides an isomorphism $SO(3) \simeq SU(2)/\mathbb{Z}_2$. Note that $\mathbb{Z}_2 = \{\mathbb{1}, -\mathbb{1}\}$ is the centre[5] $Z(SU(2))$ of $SU(2)$.

3 Representations

Definition. For any group G (not necessarily Lie), a *representation* is a set of non-singular matrices
$$
\{D(g) \in GL(n, F), g \in G\}
$$
such that
$$
D(g_1)D(g_2) = D(g_1 g_2)\,\forall\, g_1, g_2 \in G
$$

Definition. For any Lie algebra $\mathfrak{g}$, a *representation* is a set of matrices
$$
\{d(X) \in \operatorname{Mat}_n(F), X \in \mathfrak{g}\}
$$
such that

   (i) $[d(X_1), d(X_2)] = d([X_1, X_2])\,\forall\, X_1, X_2 \in \mathfrak{g}$

   (ii) $d(\alpha X_1 + \beta X_2) = \alpha d(X_1) + \beta d(X_2)\,\forall\, X_1, X_2 \in \mathfrak{g},\ \alpha, \beta \in F$

In both cases:

Definition. The *dimension* of a representation is the dimension of the corresponding matrices.

Definition. Matrices act on a vector space $V = F^n$ known as the *representation space*.

There is a direct relation between the representations of a Lie group G and the representations of the corresponding Lie algebra $\mathcal{L}(G)$. Suppose D is a representation of dimension n of a matrix Lie group $G \subset \operatorname{Mat}_m(F)$ (note that $\dim G \neq \dim D$ in general). For each $X \in \mathcal{L}(G)$, construct a curve $C : t \mapsto g(t)$ with $g(0) = \mathbb{1}$, $\dot{g}(0) = X$, and define $d(X) = \frac{\mathrm{d}}{\mathrm{d}t}D(g(t))\big|_{t=0} \in \operatorname{Mat}_n(F)$.

Lemma 4. *d is a representation of $\mathcal{L}(G)$.*

---

[4]i.e. a globally 2:1 map.
[5]i.e. the normal subgroup that commutes with everything.

Proof. Let $X_1, X_2 \in \mathcal{L}(G)$ and construct the curves C_1, C_2 as described. Define $h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t) \in G$. We established before that this has the following Taylor expansion:

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

Now, since D is a representation of G, we have $D(h) = D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2)$. Let's Taylor expand both sides.

$$D(h) = D(\mathbb{1} + t^2[X_1, X_2] + \cdots)$$
$$= D(\mathbb{1}) + t^2\left(\frac{\mathrm{d}}{\mathrm{d}\,(t^2)}D(h(t))\right)_{t=0} + \cdots$$
$$= \mathbb{1} + t^2 d([X_1, X_2]) + \cdots$$
$$D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2) = \mathbb{1} + t^2[d(X_1), d(X_2)] + \ldots$$

So, by comparing both the coefficient of t^2 on both sides, we have $d([X_1, X_2]) = [d(X_1), d(X_2)]$. Also, linearity is automatic. $\qquad\square$

Conversely, given a representation d of $\mathcal{L}(G)$, define $D(g = \mathrm{Exp}\,X) = \mathrm{Exp}(d(X))$.

Lemma 5. *D is a representation of* $\mathrm{Im}_{\mathrm{Exp}}(\mathcal{L}(G))$.

Unlectured. *Proof.* First note that $D(g)$ is nonsingular for $g \in G$. Suppose $g_1 = \mathrm{Exp}(X_1), g_2 = \mathrm{Exp}(X_2) \in \mathrm{Im}_{\mathrm{Exp}}(\mathcal{L}(G))$. By applying (BCH) twice we have:

$$D(g_1 g_2) = \mathrm{Exp}\left(d\left(X_1 + X_2 + \frac{1}{2}[X_1, X_2] + \cdots\right)\right)$$
$$= \mathrm{Exp}\left(d(X_1) + d(X_2) + \frac{1}{2}[d(X_1), d(X_2)] + \cdots\right)$$
$$= \mathrm{Exp}(d(X_1))\,\mathrm{Exp}(d(X_2)) = D(g_1)D(g_2) \qquad\square$$

Definition. The *trivial representation* d_0 of a Lie algebra $\mathfrak{g}$ is defined by $d_0(X) = 0 \,\forall\, X \in \mathfrak{g}$.

Definition. The *fundamental representation* d_f of a matrix Lie algebra $\mathfrak{g}$ is defined by $d_f(X) = X \,\forall\, X \in \mathfrak{g}$.

Definition. Every Lie algebra $\mathfrak{g}$ has an *adjoint representation* d_{Adj} defined as follows: the representation space is $\mathfrak{g}$ itself, and $d_{\mathrm{Adj}}(X) = \mathrm{ad}_X \,\forall\, X \in \mathfrak{g}$, where ad_X is defined by $\mathrm{ad}_X(Y) = [X, Y] \,\forall\, Y \in \mathfrak{g}$.

If we choose a basis $B = \{T^a\}, a = 1, .., D$ for $\mathfrak{g}$, then we can write:

$$[\mathrm{ad}_X(Y)]_c = [X, Y]_c = X_a Y_b f_c^{ab} = (R_X)_c^b Y_b$$

where $[d_{\mathrm{Adj}}(X)]_c^b = (R_X)_c^b = X_a f_c^{ab}$ is a $D \times D$ matrix.

We should check that d_{Adj} is a valid representation:

$$
\begin{aligned}
[d_{\mathrm{Adj}}(X), d_{\mathrm{Adj}}(Y)](Z) &= (\mathrm{ad}_X \, \mathrm{ad}_Y - \mathrm{ad}_Y \, \mathrm{ad}_X)(Z) \\
&= [X, [Y, Z]] - [Y, [X, Z]] \\
&= [[X, Y], Z] \qquad\qquad\qquad\qquad \text{(Jacobi identity)} \\
&= \mathrm{ad}_{[X,Y]}(Z) = d_{\mathrm{Adj}}([X, Y])(Z)
\end{aligned}
$$

Definition. Two representations R_1, R_2 of a Lie algebra $\mathfrak{g}$ are called *equivalent* or *isomorphic* if there exists a non-singular matrix S such that:

$$
R_2(X) = S R_1(X) S^{-1} \, \forall \, X \in \mathfrak{g}
$$

S represents a change of basis in representation space.

Definition. A representation R of a Lie algebra $\mathfrak{g}$ with representation space V has an *invariant subspace* $U \subset V$ if:

$$
R(X)u \in U \, \forall \, X \in \mathfrak{g}, u \in U
$$

Any representation has two trivial invariant subspaces $\{0\}$ and V.

Definition. An *irreducible representation* or *irrep* of a Lie algebra $\mathfrak{g}$ is a representation of $\mathfrak{g}$ with no non-trivial invariant subspaces.

3.1 Representation theory of $\mathcal{L}(SU(2))$

Before we noted that $\{T^a = -\frac{1}{2}\sigma_a\}$ is a real basis of $\mathcal{L}(SU(2))$. In order to explore the representations of $\mathcal{L}(SU(2))$, we will need to use a new basis:

Definition. The *Cartan-Weyl basis* of $\mathcal{L}(SU(2))$ is a complex basis with the following elements:

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

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

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

We have:

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

$E_\pm$ are known as *step operators*. Note $\mathrm{ad}_H(H) = 0, \mathrm{ad}_H(E_\pm) = \pm 2E_\pm$, so the Cartan-Weyl generators are eigenvectors of ad_H: H, E_+, E_- have eigenvalues $0, 2, -2$ respectively, and these eigenvalues are known as the *roots* of $\mathcal{L}(SU(2))$.

Consider a finite dimensional representation R of $\mathcal{L}(SU(2))$ with representation space V, and assume that $R(H)$ is diagonalisable. Then V is spanned by the eigenvectors of $R(H)$:

$$R(H)v_\lambda = \lambda v_\lambda, \ \lambda \in \mathbb{C}$$

The eigenvalues $\{\lambda\}$ of $R(H)$ are known as the *weights* of R. We have:

$$R(H)R(E_\pm)v_\lambda = (R(E_\pm)R(H) + \underbrace{[R(H), R(E_\pm)]}_{=\pm 2R(E_\pm)})v_\lambda$$

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

Hence $R(E_\pm)v_\lambda$ is either another eigenvector of $R(H)$, or it is equal to 0.

Since R is finite dimensional, it must have a *highest weight* $\Lambda \in \mathbb{C}$, by which we mean one with $R(E_+)v_\Lambda = 0$ (as otherwise we could generate an infinite number of independent vectors in the representation space). If R is irreducible, then we expect to be able to find all of the remaining basis vectors of V by acting on v_Λ with $R(E_\pm)$. To that end, define:

$$v_{\Lambda-2l} = (R(E_-))^l \, v_\Lambda, l \in \mathbb{N}$$

We have $R(H)v_{\Lambda-2l} = (\Lambda - 2l)v_{\Lambda-2l}$, $R(E_-)v_{\Lambda-2l} = v_{\Lambda-2l-2}$. Also, $R(E_+)v_{\Lambda-2l} = r_l v_{\Lambda-2l+2}$ for some $r_l \in \mathbb{C}$. We can find a recurrence relation for r_l:

$$R(E_+)v_{\Lambda-2l} = R(E_+)R(E_-)v_{\Lambda-2l+2}$$

$$= (R(E_-)R(E_+) + \underbrace{[R(E_+), R(E_-)]}_{=R(H)})v_{\Lambda-2l+2}$$

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

$$\implies r_l = r_{l-1} + (\Lambda - 2l + 2)$$

Using $r_0 = 0$, we can solve this to obtain $r_l = (\Lambda + 1 - l)l$.

The fact that R is a finite dimensional irrep implies that there must be a least weight and it must be of the form $\Lambda - 2N$ for some $N \in \mathbb{N}$. That is, $v_{\Lambda-2N} \neq 0, R(E_-)v_{\Lambda-2N} = 0$. This implies that $v_{\Lambda-2N-2} = 0$, which in turn implies that $r_{N+1} = (\Lambda - N)(N + 1) = 0$. Hence, $\Lambda = N$.

So we can conclude that the finite dimensional irreps of $\mathcal{L}(SU(2))$ are R_Λ where $\Lambda \in \mathbb{Z}, \Lambda \geq 0$, with the following properties:

- R_Λ has weights $\{-\Lambda, -\Lambda + 2, \ldots, \Lambda\} \subset \mathbb{Z}$.

- Each weight has multiplicity one.

- $\dim R_\Lambda = \Lambda + 1$ if $\Lambda > 0, = 0$ if $\Lambda = 0$.

> **Example.** We have already seen the following representations:
>
> - $R_0 = d_0$, the trivial representation.
>
> - $R_1 = d_f$, the fundamental representation.
>
> - $R_2 = d_{\mathrm{Adj}}$, the adjoint representation.

This is equivalent to the theory of angular momentum in quantum mechanics. We have the hermitian operators $\mathbf{J} = (J_1, J_2, J_3)$. Eigenstates are labelled by $j \in \mathbb{Z}/2, j \geq 0$ and $m \in \mathbb{Z}/2$ or $\mathbb{Z}, -j \geq m \geq j$, with $J^2 |j, m\rangle = j(j+1) |j, m\rangle$, $J_3 |j, m\rangle = m |j, m\rangle$.

$$J_3 = \frac{1}{2}R(H) \qquad\qquad J_\pm = J_1 \pm iJ_2 = R(E_\pm)$$
$$\Lambda = 2j \qquad\qquad \lambda = 2m$$
$$v_\Lambda \sim |j, j\rangle \qquad\qquad v_{\Lambda - 2l} \sim |j, j - l\rangle$$

Locally, we can parametrise $A \in SU(2)$ as $A = \mathrm{Exp}(X)$, $X \in \mathcal{L}(SU(2))$. Define $D_\Lambda(A) = \mathrm{Exp}(R_\Lambda(X)$. It is easy to check that this gives a valid representation of $SU(2)$, but it does not always give a valid representation of $SO(3) \simeq SU(2)/\mathbb{Z}_2$. For D_Λ to be a representation of $SO(3)$, we require $D_\Lambda(-\mathbb{1}) = D_\Lambda(\mathbb{1}) \implies D_\Lambda(-A) = D_\Lambda(A) \,\forall\, A \in SU(2)$. Note that $-\mathbb{1} = \mathrm{Exp}(i\pi H)$, so $D_\Lambda(-\mathbb{1}) = \mathrm{Exp}(i\pi R_\Lambda(H))$. $R_\Lambda(H)$ has eigenvalues in $S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, \Lambda\}$, so $D_\Lambda(-\mathbb{1})$ has eigenvalues $\exp(i\pi\lambda), \lambda \in S_\Lambda$. We then have two cases:

- $\Lambda \in 2\mathbb{Z}$. This implies that all eigenvalues of $D_\Lambda(-\mathbb{1})$ are equal to 1, so $D_\Lambda(-\mathbb{1}) = \mathbb{1} = D_\Lambda(-\mathbb{1})$. Thus D_Λ is a representation of both $SU(2)$ and $SO(3)$.

- $\Lambda \in 2\mathbb{Z} + 1$. This implies that some eigenvalues of $D_\Lambda(-\mathbb{1})$ are equal to -1, so $D_\Lambda(-\mathbb{1}) \neq D_\Lambda(-\mathbb{1})$. Thus D_Λ is a representation of $SU(2)$ but not $SO(3)$.

Rotational symmetry in 3D is naively $SO(3)$, but we see that in quantum mechanics this is not necessarily the case. For example, electrons have spin $\frac{1}{2}$, and so their angular momentum states live in R_1, which is not a representation of $SO(3)$. In some sense, we can say that if $\Lambda \in 2\mathbb{Z} + 1$, then R_Λ is a "spinor" representation of $SO(3)$ (but this is an abuse of language).

3.2 New representations from old

> **Definition.** If $\mathfrak{g}$ is a real Lie algebra, and R is a representation of $\mathfrak{g}$, then the *conjugate representation* $\bar{R}$ is defined by $\bar{R}(X) = R(X)^* \,\forall\, X \in \mathfrak{g}$.

Sometimes, but not always, $\bar{R} \simeq R$.

> **Definition.** Suppose R_1 and R_2 are representations of a Lie algebra $\mathfrak{g}$ with representation spaces V_1 and V_2, and dimensions d_1 and d_2 respectively.
>
> - The *direct sum* $R_1 \oplus R_2$ is a $(d_1 + d_2)$-dimensional representation of $\mathfrak{g}$ acting on $V_1 \oplus V_2 = \{v_1 \oplus v_2 \,\text{s.t.}\, v_1 \in V_1, v_2 \in V_2\}$ in the following way:
>
> $$(R_1 \oplus R_2)(X)(v_1 \oplus v_2) = (R_1(X)v_1) \oplus (R_2(X)v_2) \,\forall\, X \in \mathfrak{g}$$
>
> - The *tensor product* $R_1 \otimes R_2$ is a $d_1 d_2$-dimensional representation of $\mathfrak{g}$ acting on $V_1 \otimes V_2 = \{v_1 \otimes v_2 \,\text{s.t.}\, v_1 \in V_1, v_2 \in V_2\}$ in the following way:
>
> $$(R_1 \otimes R_2)(X)(v_1 \otimes v_2) = (R_1(X)v_1) \otimes v_2 + v_1 \otimes (R_2(X)v_2) \,\forall\, X \in \mathfrak{g}$$

More explicitly, we can write $(R_1 \oplus R_2)(X)$ as a matrix:

$$(R_1 \oplus R_2)(X) = \left(\begin{array}{c|c} R_1(X) & 0 \\ \hline 0 & R_2(X) \end{array} \right) \begin{array}{l} \updownarrow d_1 \\ \updownarrow d_2 \end{array}$$

If we choose bases $B_1 = \{v_1^j\}, j = 1, \ldots, d_1$ and $B_2 = \{v_2^\alpha\}, \alpha = 1, \ldots, d_2$ for V_1 and V_2 respectively, then we can define a basis for $V_1 \otimes V_2$:

$$B_{1\otimes 2} = B_1 \otimes B_2 = \{v_1^j \otimes v_2^\alpha \,\text{s.\,t.}\, j = 1, \ldots, d_1, \; \alpha = 1, \ldots, d_2\}$$

So the dimension of $R_1 \otimes R_2$ is indeed $d_1 d_2$. Let $w \in V_1 \otimes V_2$ have components $w_{j\alpha}$ in $B_{1\otimes 2}$. Then we can write $(R_1 \otimes R_2)(X)$ in terms of its components:

$$(R_1 \otimes R_2)(X)_{i\alpha, j\beta} = R_1(X)_{ij} \mathbb{1}_{\alpha\beta} + \mathbb{1}_{ij} R_2(X)_{\alpha\beta}$$

Unlectured. We should show that $R_1 \otimes R_2$ is a valid representation. We have:

$$\begin{aligned}
(R_1 \otimes R_2)([X, Y]) &= R_1([X, Y]) \otimes \mathbb{1} + \mathbb{1} \otimes R_2([X, Y]) \\
&= [R_1(X), R_1(Y)] \otimes \mathbb{1} + \mathbb{1} \otimes [R_2(X), R_2(Y)] \\
&= [R_1(X) \otimes \mathbb{1}, R_1(Y) \otimes \mathbb{1}] + [R_1(X) \otimes \mathbb{1}, \mathbb{1} \otimes R_2(Y)] \\
&\quad + [\mathbb{1} \otimes R_2(X), R_1(Y) \otimes \mathbb{1}] + [\mathbb{1} \otimes R_2(X), \mathbb{1} \otimes R_2(Y)] \\
&= [R_1(X) \otimes \mathbb{1} + \mathbb{1} \otimes R_2(X), R_1(Y) \otimes \mathbb{1} + \mathbb{1} \otimes R_2(Y)] \\
&= [(R_1 \otimes R_2)(X), (R_1 \otimes R_2)(Y)]
\end{aligned}$$

where we have used the fact that $A \otimes \mathbb{1}$ commutes with $\mathbb{1} \otimes B$ for any A, B. Also, $R_1 \otimes R_2$ is inherently linear.

Definition. A representation is *fully reducible* if it can be written as a direct sum of non-trivial irreps.

If a representation R is reducible, then it has an invariant subspace, and we can find a basis where:

$$R(X) = \left(\begin{array}{c|c} A(X) & B(X) \\ \hline 0 & C(X) \end{array} \right) \; \forall X \in \mathfrak{g}$$

If R is fully reducible, then we can furthermore find a basis such that $B(X) = 0 \,\forall X \in \mathfrak{g}$. More generally, if R is fully reducible, then we have a basis where $R(X)$ is block diagonal:

$$R(X) = \begin{pmatrix} \boxed{R_1(X)} & & & \\ & \boxed{R_2(X)} & & \\ & & \ddots & \\ & & & \boxed{R_n(X)} \end{pmatrix}$$

and each R_i is an irrep.

Lemma 6. *If $R_i, i = 1, \ldots, m$ are finite dimensional irreps of a complex simple Lie algebra $\mathfrak{g}$, then tensor product of the R_i is fully reducible:*

$$\bigotimes_{i=1}^{m} R_i = \bigoplus_{i=1}^{\tilde{m}} \tilde{R}_i \; \text{some } \tilde{m}, \tilde{R}_i$$

3.3 Tensor products of $\mathcal{L}(SU(2))$ representations

Let R_Λ and $R_{\Lambda'}$ be irreps of $\mathcal{L}(SU(2))$ with highest weights $\Lambda, \Lambda' > 0$ respectively. We have $\dim R_\Lambda = \Lambda + 1, \dim R_{\Lambda'} = \Lambda' + 1$. The tensor product $R_\Lambda \otimes R_{\Lambda'}$ is a fully reducible representation with $\dim(R_\Lambda \otimes R_{\Lambda'}) = (\Lambda + 1)(\Lambda' + 1)$ (since $SU(2)$ is simple):

$$R_\Lambda \otimes R_{\Lambda'} = \bigoplus_{\substack{\Lambda'' \in \mathbb{Z} \\ \Lambda'' > 0}} \mathcal{L}_{\Lambda,\Lambda'}^{\Lambda''} R_{\Lambda''}$$

where $\mathcal{L}_{\Lambda,\Lambda'}^{\Lambda''} \in \mathbb{Z}$. We have bases of $R(H)$ eigenvectors for V_Λ and $V_{\Lambda'}$:

$$\{v_\lambda\}, \lambda \in S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, \Lambda\} \qquad\qquad R_\Lambda(H)v_\lambda = \lambda v_\lambda$$
$$\{v'_{\lambda'}\}, \lambda' \in S_{\Lambda'} = \{-\Lambda', -\Lambda' + 2, \ldots, \Lambda'\} \qquad\qquad R_{\Lambda'}(H)v'_{\lambda'} = \lambda' v'_{\lambda'}$$

So construct a basis for $V_\lambda \otimes V_{\lambda'}$:

$$B = \{v_\lambda \otimes v'_{\lambda'} \text{ s. t. } \lambda \in S_\Lambda, \lambda' \in S_{\Lambda'}\}$$

These basis vectors are eigenvectors of $(R_\Lambda \otimes R_{\Lambda'})(H)$:

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

So the weights of $R_\Lambda \otimes R_{\Lambda'}$ are $S_{\Lambda,\Lambda'} = \{\lambda + \lambda' \text{ s. t. } \lambda \in S_\Lambda, \lambda' \in S_{\Lambda'}\}$. The highest weight is $\Lambda + \Lambda'$, and it has multiplicity one, so we must have:

$$R_\Lambda \otimes R_{\Lambda'} = R_{\Lambda+\Lambda'} \oplus \tilde{R}_{\Lambda,\Lambda'}$$

for some remainder $\tilde{R}_{\Lambda,\Lambda'}$. The highest weight of $\tilde{R}_{\Lambda,\Lambda'}$ is $\Lambda + \Lambda' - 2$, again with multiplicity one, so:

$$R_\Lambda \otimes R_{\Lambda'} = R_{\Lambda+\Lambda'} \oplus R_{\Lambda+\Lambda'-2} \oplus \tilde{\tilde{R}}_{\Lambda,\Lambda'}$$

We can continue in this way to obtain:

$$R_\Lambda \otimes R_{\Lambda'} = R_{\Lambda+\Lambda'} \oplus R_{\Lambda+\Lambda'-2} \oplus \cdots \oplus R_{|\Lambda-\Lambda'|+2} \oplus R_{|\Lambda-\Lambda'|}$$

where we have stopped at $R_{|\Lambda-\Lambda'|}$ in order to satify $\dim(R_\Lambda \otimes R_{\Lambda'}) = (\Lambda + 1)(\Lambda' + 1)$.

Example.
$$
\begin{array}{ccccccc}
R_1 & \otimes & R_1 & = & R_2 & \oplus & R_0 \\
\text{in Q.M.:} \quad \text{spin } \tfrac{1}{2} & \otimes & \text{spin } \tfrac{1}{2} & = & \underbrace{\text{spin } 1}_{\text{triplet}} & \oplus & \underbrace{\text{spin } 0}_{\text{singlet}}
\end{array}
$$

---

4 Cartan Classification

4.1 The Killing Form

Definition. Given a vector space V over a field F, an *inner product* is a symmetric bilinear map $i : V \times V \to F$. i is said to be *non-degenerate* if for each $v \in V$ there is a $w \in V$ such that $i(v, w) \neq 0$.

Given a Lie algebra $\mathfrak{g}$ we have a "natural" inner product:

Definition. The *Killing form* of a Lie algebra $\mathfrak{g}$ over a field F is the following function:

$$
\begin{array}{rccc}
\kappa & : & \mathfrak{g} \times \mathfrak{g} & \to & F \\
& & (X, Y) & \mapsto & \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y)
\end{array}
$$

Lemma 7. *The Killing form is an inner product.*

Proof. $K(X \circ Y)$ is the trace of $\mathrm{ad}_X \circ \mathrm{ad}_Y$, so we should analyse this map. Choose a basis $\{T^a\}$ for $\mathfrak{g}$. We have:

$$
\begin{aligned}
(\mathrm{ad}_X \circ \mathrm{ad}_Y)(Z) &= [X, [Y, Z]] \\
&= X_a Y_b Z_c [T^a, [T^b, T^c]] \\
&= X_a Y_b Z_c f_d^{bc} [T^a, T^d] \\
&= X_a Y_b Z_c f_d^{bc} f_e^{ad} T^e \\
&= M(X, Y)_e^c Z_c T^e
\end{aligned}
$$

where we have defined $M(X, Y)_e^c = X_a Y_b f_d^{bc} f_e^{ad}$. Hence:

$$
\kappa(X, Y) = \mathrm{tr}(M(X, Y)) = X_a Y_b \underbrace{f_c^{ad} f_d^{bc}}_{\kappa^{ab}}
$$

From this we can see that κ is both bilinear and symmetric. $\qquad\square$

It is worth saying what we mean by "natural":

Lemma 8. κ *has the following property (we say it is* invariant*):*

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

Proof. First note that $\mathrm{ad}_{[X,Y]} = [\mathrm{ad}_X, \mathrm{ad}_Y]$. We have:

$$
\begin{aligned}
\kappa([Z, X], Y) &= \mathrm{tr}\left(\mathrm{ad}_{[Z,X]} \circ Y\right) \\
&= \mathrm{tr}(\mathrm{ad}_Z \circ \mathrm{ad}_X \circ \mathrm{ad}_Y - \mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y) \\
&= \mathrm{tr}(\mathrm{ad}_X \circ \mathrm{ad}_Y \circ \mathrm{ad}_Z - \mathrm{ad}_X \circ \mathrm{ad}_Z \circ \mathrm{ad}_Y) \\
&= \mathrm{tr}\left(\mathrm{ad}_X \circ \mathrm{ad}_{[Y,Z]}\right) \\
&= \kappa(X, [Y, Z]) \\
&= -\kappa(X, [Z, Y])
\end{aligned}
$$

where we used in the third line the fact that the trace of a product is cylic in its factors. $\qquad\square$

If $\mathfrak{g}$ is simple then κ is the unique invariant inner product (up to a constant multiple).

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

Lemma 9. *If $\mathfrak{g}$ is a semi-simple Lie algebra, then we can write it as a direct sum of simple Lie algebras:*

$$\mathfrak{g} = \underbrace{\mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \cdots \oplus \mathfrak{g}_n}_{simple}$$

It will be helpful to know when the Killing form is non-degenerate. It turns out:

Theorem 1 (Cartan). *The Killing form is non-degenerate if and only if $\mathfrak{g}$ is semi-simple.*

Proof in one direction. Suppose $\mathfrak{g}$ is not semi-simple. Then $\mathfrak{g}$ has an abelian ideal $\mathfrak{j}$. Let $D = \dim \mathfrak{g}, d = \dim \mathfrak{j}$. Choose a basis:

$$B = \{T^a\} = \{T^i; i = 1, \ldots, d\} \cup \{T^\alpha; \alpha = 1, \ldots, D - d\}$$

where $\{T^i\}$ spans $\mathfrak{j}$. Since $\mathfrak{j}$ is abelian, we have $[T^i, T^j] = 0$ so $f_a^{ij} = 0$. Since $\mathfrak{j}$ is an ideal, we have $[T^i, T^\alpha] \in \mathfrak{j}$ so $f_\beta^{\alpha j} = 0$. Now let $X = X_a T^a \in \mathfrak{g}$ and $Y = Y_i T^i \in \mathfrak{j}$. We have:

$$\begin{aligned}
\kappa^{ai} &= f_c^{ad} f_d^{ic} \\
&= f_\alpha^{ad} f_d^{i\alpha} \quad (f_d^{ic} = 0 \text{ for } c \neq \alpha) \\
&= \underbrace{f_\alpha^{aj}}_{=0} f_j^{i\alpha} \quad (f_d^{i\alpha} = 0 \text{ for } d \neq j) \\
&= 0
\end{aligned}$$

Thus $\kappa(X, Y) = 0$ for all $X \in \mathfrak{g}$, so κ is degenerate. $\qquad\square$

4.2 Complexification

Definition. Given a real Lie algebra $\mathfrak{g}$, we can find a basis $\{T^a\}$ such that $\mathfrak{g} = \text{span}_\mathbb{R}\{T^a\}$. The *complexification* of $\mathfrak{g}$ is the complex Lie algebra $\mathfrak{g}_\mathbb{C} = \text{span}_\mathbb{C}\{T^a\}$ with the same bracket.

$\mathfrak{g}$ is said to be a *real form* of $\mathfrak{g}_\mathbb{C}$. A complex Lie algebra can have more than one real form. $\mathfrak{g}_\mathbb{C}$ is also a real Lie algebra with generators $\{T^a\} \cup i\{T^a\}$. The "real" dimension of $\mathfrak{g}_\mathbb{C}$ is $\dim_\mathbb{R} \mathfrak{g}_\mathbb{C} = 2 \dim \mathfrak{g}$.

Example. We have:

$$\begin{aligned}
\mathcal{L}(SU(2)) &= \text{span}_\mathbb{R}\{T^a = -\frac{i}{2}\sigma_a; a = 1, 2, 3\} \\
&= \{2 \times 2 \text{ traceless anti-Hermitian matrices}\}
\end{aligned}$$

The corresponding complexification is:

$$\begin{aligned}
\mathcal{L}_\mathbb{C}(SU(2)) &= \text{span}_\mathbb{C}\{T^a = -\frac{i}{2}\sigma_a; a = 1, 2, 3\} \\
&= \{2 \times 2 \text{ traceless complex matrices}\}
\end{aligned}$$

In fact we also have:

$$\mathcal{L}(SL(2, \mathbb{C})) = \{2 \times 2 \text{ traceless complex matrices}\}$$

So $\mathcal{L}_\mathbb{C}(SU(2)) = \mathcal{L}(SL(2, \mathbb{C}))$.

Definition. A real Lie algebra of *compact type* has a basis in which $\kappa^{ab} = -\kappa\delta^{ab}$.

𝕿𝖍𝖊𝖔𝖗𝖊𝖒 **2.** *Every complex, semi-simple, finite-dimensional Lie algebra has a real form of compact type.*

4.3 Cartan-Weyl basis

Definition. We say that $X \in \mathfrak{g}$ is ad-diagonalisable if ad_X is diagonalisable.

Definition. A *Cartan subalgebra* (CSA) $\mathfrak{h}$ of $\mathfrak{g}$ is a maximal abelian subalgebra containing only ad-diagonalisable elements:

(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 $[X, H] = 0 \,\forall\, H \in \mathfrak{h}$, then $X \in \mathfrak{h}$.

It is a fact that all possible choices of CSA have the same dimension.

Definition. The dimension of a CSA in $\mathfrak{g}$ is called the *rank r* of $\mathfrak{g}$.

Example. Suppose $\mathfrak{g} = \mathcal{L}_{\mathbb{C}}(SU(2))$. H is ad-diagonalisable, but $E_{\pm}$ are not, so we can choose $\mathfrak{h} = \mathrm{span}_{\mathbb{C}}\{H\}$ to be a CSA. Hence the rank of $\mathfrak{g}$ is 1.

Given a CSA $\mathfrak{h}$ we can choose a basis $\{H^i\}$.

Example. Suppose $\mathfrak{g} = \mathcal{L}_{\mathbb{C}}(SU(n)) = \{\text{traceless complex } n \times n \text{ matrices}\}$. The diagonal elements of $\mathfrak{g}$ are a natural choice of CSA. We can thus choose the following basis:

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

The rank of $\mathfrak{g}$ is $n - 1$.

Note that since $[H^i, H^j] = 0$, we have $[\mathrm{ad}_{H^i}, \mathrm{ad}_{H^j}]$, so the H^i are simultaneously ad-diagonalisable. We can find distinguish the eigenvectors ($\in \mathfrak{g}$) into two types:

- The eigenvectors with eigenvalue zero are the basis vectors of $\mathfrak{h}$ i.e. the H^i.

- Since the CSA is maximal, all remaining vectors must have non-zero eigenvalues.

Definition. Eigenvectors of the CSA with non-zero eigenvalues are called *step operators*, and we denote them as E^α, where α is a collection of complex numbers $\alpha^i, i = 1, \ldots, r$ such that $[H^i, E^\alpha] = \alpha^i E^\alpha$. α is called a *root* of the Lie algebra. The set of all roots of a Lie algebra is known as its *root set*, denoted Φ.

Example. Suppose $\mathfrak{g} = \mathcal{L}_\mathbb{C}(SU(n))$, then as before $\mathfrak{h} = \{\text{traceless diagonal } n \times n \text{ complex matrices}\}$. So we can write $H \in \mathfrak{h}$ as $H = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Suppose $\text{ad}_H X = [H, X] = \mu X$. Then we have:

$$(\lambda_l - \lambda_m) X_{lm} = \mu X_{lm} \quad \text{(no summation)}$$

The non-zero solutions of this equation gives us the step operators of $\mathcal{L}_\mathbb{C}(SU(n))$:

$$X = E^{(r,s)} \text{ where } E^{(r,s)}_{lm} = \delta_{lr}\delta_{ms}, \ r \neq s$$

Suppose $H \in \mathfrak{h}$ is a general element of the CSA. We can write $H = e_i H^i$, where the e_i are complex numbers. Then we have:

$$[H, E^\alpha] = e_i[H^i, E^\alpha] = \alpha(H)E^\alpha$$

where $\alpha(H) = e_i \alpha^i$. So each root α defines a linear map $\alpha : \mathfrak{h} \to \mathbb{C}$. That is, the roots are elements of $\mathfrak{h}^*$, the dual of $\mathfrak{h}$.

Henceforth we shall make the assumption that the roots are non-degenerate. This implies that the root set Φ consists of $d - r$ independent elements of $\mathfrak{h}^*$. Then we have:

Definition. The *Cartan-Weyl basis* for a Lie algebra is given by:

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

Recall that $\mathfrak{g}$ being simple implies that the Killing form is non-degenerate. We will now prove a series of results about the Killing form in the Cartan-Weyl basis.

Lemma 10.   *(i)* $\kappa(H, E^\alpha) = 0 \, \forall \, H \in \mathfrak{h}, \alpha \in \Phi$.

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

$$\alpha(H')\kappa(H, E^\alpha) = \kappa(H, [H', E^\alpha]) = -\kappa(\underbrace{[H, H']}_{=0}, E^\alpha) = 0 \qquad \square$$

   (ii) $\kappa(E^\alpha, E^\beta) = 0 \, \forall \, \alpha, \beta \in \Phi \, \text{s.t.} \, \alpha \neq -\beta$

     Proof. Let $\alpha, \beta \in \Phi$. For all $H' \in \mathfrak{h}$ we have:

$$(\alpha(H') + \beta(H'))\kappa(E^\alpha, E^\beta) = \kappa([H', E^\alpha], E^\beta) + \kappa(E^\alpha, [H', E^\beta]) = 0$$

     Thus $\alpha + \beta \neq 0$ implies $\kappa(E^\alpha, E^\beta) = 0$. $\square$

   (iii) $\forall \, H \in \mathfrak{h}, \, \exists \, H' \in \mathfrak{h} \, \text{s.t.} \, \kappa(H, H') \neq 0$.

Proof. Suppose that for some $H \in \mathfrak{h}$, we have $\kappa(H, H') = 0 \,\forall\, H' \in \mathfrak{h}$. Then from (i) we have $\kappa(H, E^\alpha) = 0 \,\forall\, \alpha \in \Phi$. But then that would mean that $\kappa(H, X) = 0 \,\forall\, X \in \mathfrak{g}$, i.e. that κ is degenerate, which is a contradiction. $\hfill\square$

(iv) $\alpha \in \Phi \implies -\alpha \in \Phi$ and $\kappa(E^\alpha, E^{-\alpha}) \neq 0$.

Proof. (i) and (ii) imply together that unless $-\alpha$ is a root and $\kappa(E^\alpha, E^{-\alpha}) \neq 0$ then we have $\kappa(E^\alpha, X) = 0 \,\forall\, X \in \mathfrak{g}$, implying that κ is degenerate, again a contradiction. $\hfill\square$

κ started life as a non-degenerate product. (iii) implies that κ also gives a non-degenerate inner product on $\mathfrak{h}$. Let $H = e_i H^i, H' = e'_i H^i$. Then we can calculate $\kappa(H, H') = \kappa^{ij} e_i e'_j$, where κ^{ij} are the components of an invertible $r \times r$ matrix. We can find the inverse $(\kappa^{-1})_{ij}$, and from this we now get a non-degenerate product on the span of the roots in $\mathfrak{h}^*$:

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

Let's fully evaluate the structure of the Lie algebra in the Cartan-Weyl basis. We already have:

$$[H^i, H^j] = 0 \;\; \forall\, i, j = 1, \dots, r$$
$$[H^i, E^\alpha] = \alpha^i E^\alpha \;\; \forall\, \alpha \in \Phi$$

Now let's find $[E^\alpha, E^\beta]$. We have:

$$[H^i, [E^\alpha, E^\beta]] = -[E^\alpha, [E^\beta, H^i]] - [E^\beta, [H^i, E^\alpha]]$$
$$= (\alpha^i + \beta^i)[E^\alpha, E^\beta]$$
$$\implies [E^\alpha, E^\beta] \begin{cases} \propto E^{\alpha+\beta} & \text{if } \alpha + \beta \in \Phi \\ = 0 & \text{otherwise (if } \alpha + \beta \neq 0) \end{cases}$$

If $\alpha + \beta = 0$, then consider $[E^\alpha, E^{-\alpha}]$. We have:

$$\kappa([E^\alpha, E^{-\alpha}], H) = \kappa(E^\alpha, [E^{-\alpha}, H])$$
$$= \alpha(H) \underbrace{\kappa(E^\alpha, E^{-\alpha})}_{\neq 0}$$

We can define $H^\alpha = [E^\alpha, E^{-\alpha}]/\kappa(E^\alpha, E^{-\alpha}) \in \mathfrak{h}$. H^α solves $\kappa(H^\alpha, H) = \alpha(H) \,\forall\, H \in \mathfrak{h}$ and moreover since κ is non-degenerate it does so uniquely. Let $H^\alpha = \rho_i^\alpha H^i, H = \rho_i H^i$. We have:

$$\kappa^{ij} \rho_i^\alpha \rho_j = \alpha^j \rho_j \,\forall\, \rho_j$$
$$\implies \kappa^{ij} \rho_i^\alpha = \alpha^j$$
$$\implies \rho_i^\alpha = (\kappa^{-1})_{ij} \alpha^j$$
$$\implies H^\alpha = (\kappa^{-1})_{ij} \alpha^j H^i$$

We thus have, for all $\alpha, \beta \in \Phi$, $[H^\alpha, E^\beta] = (\kappa^{-1})_{ij} \alpha^i \beta^j E^\beta = (\alpha, \beta) E^\beta$.

We have been assuming the roots are non-degenerate. Let us also assume $(\alpha, \alpha) \neq 0$. These assumptions are relaxable, but the proofs become much longer without being very instructive. Let's normalise our basis:

Definition. $\forall\, \alpha \in \Phi$:

$$e^\alpha = \sqrt{\frac{2}{(\alpha,\alpha)\kappa(E^\alpha, E^{-\alpha}}}E^\alpha$$

$$h^\alpha = \frac{2}{(\alpha,\alpha)}H^\alpha$$

Now we have, $\forall\, \alpha, \beta \in \Phi$:

$$[h^\alpha, h^\beta] = 0$$

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

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

where $n_{\alpha,\beta} \in \mathbb{C}$.

Note that we have for each $\alpha \in \Phi$ a $\mathcal{L}_\mathbb{C}(SU(2))$ subalgebra generated by $\{h^\alpha, e^\alpha, e^{-\alpha}\}$.

$$[h^\alpha, e^{\pm\alpha}] = \pm 2e^{\pm\alpha}$$

$$[e^\alpha, e^{-\alpha}] = h^\alpha$$

We call this subalgebra $sl(2)_\alpha$.

4.4 Root strings

Definition. Given $\alpha, \beta \in \Phi$ with $\alpha \neq \beta$, the *α-string passing through β* is the following set of roots:

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

For a given root string, we have a corresponding vector space $V_{\alpha,\beta} = \text{span}_\mathbb{C}\{e^{\beta+\rho\alpha};\ \beta + \rho\alpha \in S_{\alpha,\beta}\}$.
We also have:

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

and

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

$V_{\alpha,\beta}$ is invariant under $sl(2)_\alpha$, and can be viewed as a representation space for a representation R of $sl(2)_\alpha$. From the above, R has the following set of weights:

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

Each of these weights has multiplicity one. Comparing this with what we have discovered about the representations of $sl(2) = \mathcal{L}_\mathbb{C}(SU(2))$, we can see that R must be an irrep of $sl(2)_\alpha$. Furthermore,

it is finite dimensional, so we have $R = R_\Lambda$ for some highest weight $0 < \Lambda \in \mathbb{Z}$. Hence $S_R = \{-\Lambda, -\Lambda + 2, \ldots, +\Lambda\}$ and so we have $\rho \in \mathbb{Z}$ and $n_- \le n = \rho \le n_+$ for integers $n_+ \ge 0, n_- \le 0$ such that:

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

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

Adding these together we find an important constraint:

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

So we have:

$$[h^\alpha, e^\beta] = \underbrace{\frac{2(\alpha, \beta)}{(\alpha, \alpha)}}_{\in \mathbb{Z}} e^\beta = R_{\alpha, \beta} e^\beta$$

Where the last equality is a notational definition. In the Cartan-Weyl basis, we have $[H^i, E^\delta] = \delta^i E^\delta$.

$$\kappa^{ij} = \kappa(H^i, H^j)$$
$$= \operatorname{tr}(\operatorname{ad}_{H^i} \circ \operatorname{ad}_{H^j})$$
$$= \sum_{\delta \in \Phi} \delta^i \delta^j$$

Now multiply both sides by $(\kappa^{-1})_{jk}(\kappa^{-1})_{il}\alpha^k \beta^l$:

$$\kappa^{ij}(\kappa^{-1})_{jk}(\kappa^{-1})_{il}\alpha^k \beta^l = \sum_{\delta \in \Phi} \delta^i \delta^j (\kappa^{-1})_{jk}(\kappa^{-1})_{il}\alpha^k \beta^l$$

$$\implies (\alpha, \beta) = \sum_{\delta \in \Phi} (\alpha, \delta)(\beta, \delta)$$

$$\implies \frac{2}{(\beta, \beta)} = \frac{1}{R_{\alpha, \beta}} \sum_{\delta \in \Phi} R_{\alpha, \delta} R_{\beta, \delta} \in \mathbb{R}$$

Hence $(\alpha, \beta) = R_{\alpha, \beta}(\alpha, \alpha)/2 \in \mathbb{R}$ for all $\alpha, \beta \in \Phi$.

4.5 Real geometry of roots

𝔏𝔢𝔪𝔪𝔞 11. $\mathfrak{h}^* = \operatorname{span}_{\mathbb{C}}\{\alpha \in \Phi\}$.

Proof. Suppose $\alpha \in \Phi$ do not span $\mathfrak{h}^*$. Then there exists a $\lambda \in \mathfrak{h}^*$ with $(\lambda, \alpha) = (\kappa^{-1})_{ij}\lambda^i \alpha^j = \kappa^{ij}\lambda_i \alpha_j = 0$ for all $\alpha \in \Phi$. Define $H_\lambda = \lambda_i H^i$; then we have $[H_\lambda, H] = 0 \, \forall \, H \in \mathfrak{h}$ and $[H_\lambda, E^\alpha] = (\lambda, \alpha)E^\alpha = 0 \, \forall \, \alpha \in \Phi$, so $[H_\lambda, X] = 0 \, \forall \, X \in \mathfrak{g}$. But then $\mathfrak{j} = \operatorname{span}_{\mathbb{C}}\{H_\lambda\}$ is a non-trivial ideal, which contradicts simplicity of $\mathfrak{g}$. $\qquad \square$

We can find $r = \operatorname{rank}[\mathfrak{g}]$ roots to act as a basis for $\mathfrak{h}^*$:

$$\{\alpha_{(i)} \in \Phi; \ i = 1, \ldots, r\}$$

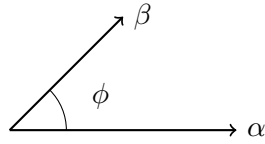Define a real subspace $\mathfrak{h}^*_{\mathbb{R}} \subset \mathfrak{h}^*$ as follows:

$$\mathfrak{h}^*_{\mathbb{R}} = \operatorname{span}_{\mathbb{R}}\{\alpha_{(i)}\}$$

Since the $\alpha_{(i)}$ span $\mathfrak{h}^*$, any root $\beta \in \Phi$ can be written as $\beta = \sum_{i=1}^{r} \beta^i \alpha_{(i)}$, where $\beta^i \in \mathbb{C}$. If we take the inner product on both sides with $\alpha_{(j)}$ we obtain:

$$\underbrace{(\beta, \alpha_{(j)})}_{\in \mathbb{R}} = \sum_{i=1}^{r} \beta^i \underbrace{(\alpha_{(i)}, \alpha_{(j)})}_{\in \mathbb{R}}$$

The inner product is non-degenerate so we can invert it to find $\beta^i \in \mathbb{R}$. Thus $\beta \in \mathfrak{h}^*_{\mathbb{R}}$ for all roots β. Let $\lambda, \mu \in \mathfrak{h}^*_{\mathbb{R}}$, and $\lambda = \lambda^i \alpha_{(i)}, \mu = \mu^i \alpha_{(i)}$. Then $(\lambda, \mu) = \sum_{i,j} \lambda^i \mu^j (\alpha_{(i)}, \alpha_{(j)}) \in \mathbb{R}$. Also, $(\lambda, \lambda) = \sum_{\delta \in \Phi} \lambda_i \delta^i \delta^j \lambda_j = \sum_{\delta \in \Phi} (\lambda, \delta)(\lambda, \delta) \geq 0$ with equality if and only if $(\lambda, \delta) = 0 \, \forall \, \delta \in \Phi$, i.e. iff $\lambda = 0$. Therefore our inner product is a Euclidean one, and $\mathfrak{h}^*_{\mathbb{R}}$ is isomorphic to n-dimensional Euclidean space. This is very nice, as it allows us to apply all of the results familiar to us from Euclidean geometry to the study of Lie algebras.

In particular, since $(\alpha, \alpha) > 0 \, \forall \, \alpha \in \Phi$, we can define the length of a root by $|\alpha| = +(\alpha, \alpha)^{1/2}$, and the angle ϕ between two roots by $(\alpha, \beta) = |\alpha||\beta| \cos \phi$.

We have:

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

$$R_{\beta,\alpha} = \frac{2(\alpha, \beta)}{(\beta, \beta)} = 2 \frac{|\alpha|}{|\beta|} \cos \phi \in \mathbb{Z}$$
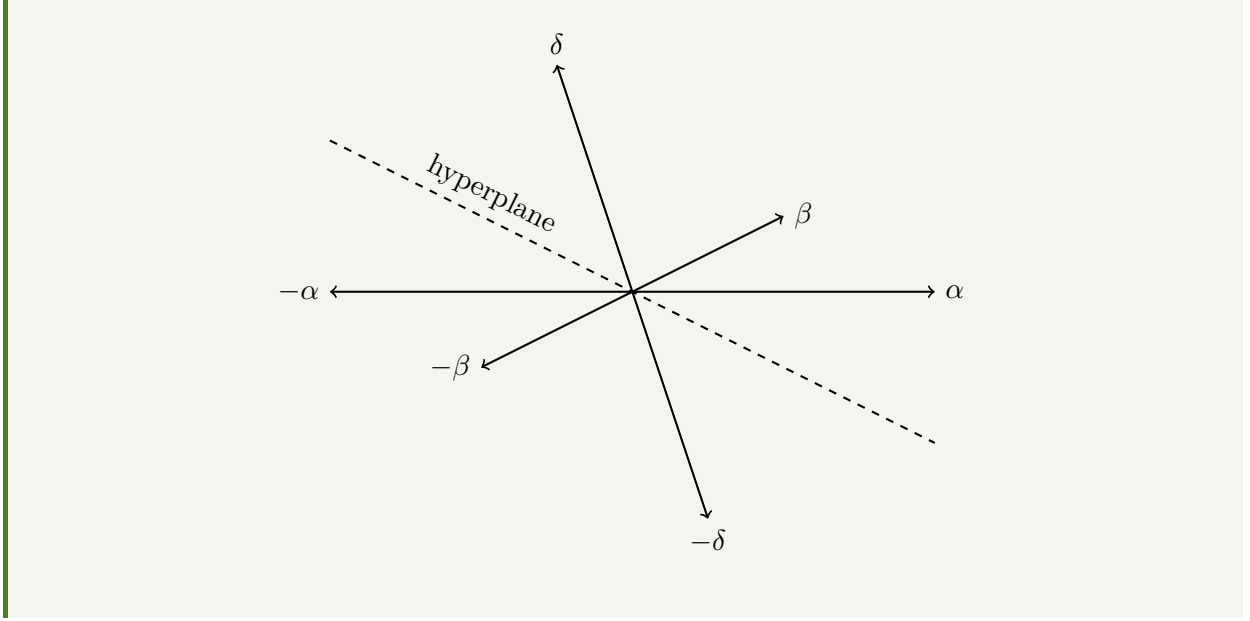
Multiplying these together we obtain:

$$4 \cos^2 \phi \in \mathbb{Z} \implies \cos \phi = \pm \frac{\sqrt{n}}{2} \text{ where } n \in \{0, 1, 2, 3, 4\}$$

This constraint has the following solutions:

$$\phi = \begin{cases} 0 & (\alpha = \beta) \\ \frac{\pi}{2} & ((\alpha, \beta) = 0) \\ \pi & (\alpha = -\beta) \\ \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3} & ((\alpha, \beta) > 0) \\ \frac{5\pi}{6}, \frac{3\pi}{4}, \frac{2\pi}{3} & ((\alpha, \beta) < 0) \end{cases}$$

4.6 Simple roots

Definition. Choose some hyperplane in $\mathfrak{h}^*$. We can divide the root set into *positive roots* in Φ_+ and *negative roots* in Φ_- by imposing that positive roots lie on one side of the hyperplane, and negative roots on the other.

This is equivalent to dividing $\Phi = \Phi_+ \cup \Phi_-$ such that, for all roots α, β:

(i) $\alpha \in \Phi_+ \iff -\alpha \in \Phi_-$.

(ii) $\alpha, \beta \in \Phi_+, \alpha + \beta \in \Phi \implies \alpha + \beta \in \Phi_+$.

> **Definition.** A *simple root* is a positive root which cannot be written as the sum of two positive roots. The set of simple roots is denoted Φ_S.

𝕷emma 12.   *(i) $\alpha, \beta \in \Phi_S \implies \alpha - \beta \notin \Phi$.*

Proof. Suppose $\alpha - \beta \in \Phi$. Then we have two cases:
- $\alpha - \beta \in \Phi_+ \implies \alpha = (\alpha - \beta) + \beta \implies \alpha$ not simple.
- $\alpha - \beta \in \Phi_- \implies \beta - \alpha \in \Phi_+ \implies \beta = (\beta - \alpha) + \alpha \implies \beta$ not simple. □

(ii) $\alpha, \beta \in \Phi_S \implies S_{\alpha,\beta}$, the α-string through β, has length $l_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}$.

Proof. Recall that we can write $S_{\alpha,\beta} = \{\beta + n\alpha; n \in \mathbb{Z}, n_- \le n \le n_+\}$ with $n_+ \ge 0, n_- \le 0$, and $n_+ + n_- = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{Z}$. Now, (i) implies that $n_- = 0$. Hence, $n_+ = -\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ and so $l_{\alpha,\beta} = n_+ - n_- + 1 = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \in \mathbb{N}$. □

(iii) $\alpha, \beta \in \Phi_S, \alpha \ne \beta \implies (\alpha, \beta) \le 0$. (This is a corollary of (ii).)

(iv) Any positive root β can be written as a linear combination of simple roots with positive integer coefficients.

Proof. If $\beta \in \Phi_S$, then we are done. If $\beta \ne \Phi_S$ then we can write $\beta = \beta_1 + \beta_2$ for some positive roots β_1, β_2. Since the rank is finite, we can iterate on this until all of the β_i are simple roots, and then we are done. □

(v) All roots $\alpha \in \Phi$ can be written as $\alpha = \sum_i d_i \alpha_{(i)}$ with $d_i \in \mathbb{Z}, \alpha_{(i)} \in \Phi_S$, so the simple roots span $\mathfrak{h}_\mathbb{R}^*$. (This is a corollary of (iv).)

(vi) Simple roots are linearly independent.

Proof. Consider all vectors $\lambda \in \mathfrak{h}_{\mathbb{R}}^*$. By *(v)* we can write $\lambda = \sum_i c_i \alpha_{(i)}$ where $c_i \in \mathbb{R}, \alpha_{(i)} \in \Phi_S$. Restrict to the case where not all $c_i = 0$. Let $J_\pm = \{i \, \text{s.t.} \, c_i \gtrless 0\}$, and $\lambda_\pm = \pm \sum_{i \in J_\pm} c_i \alpha_{(i)}$; we have $\lambda = \lambda_+ - \lambda_-$. Then

$$
\begin{aligned}
(\lambda, \lambda) &= (\lambda_+, \lambda_+) + (\lambda_-, \lambda_-) - 2(\lambda_+, \lambda_-) \\
&> -2(\lambda_+, \lambda_-) \\
&= +2 \sum_{i \in J_+} \sum_{j \in J_-} \underbrace{c_i c_j}_{<0} \underbrace{(\alpha_{(i)}, \alpha_{(j)})}_{\leq 0 \text{ (by } (iii))} \\
&\geq 0
\end{aligned}
$$

Hence $(\lambda, \lambda) > 0$. So $(\lambda, \lambda) = 0$ if and only if $c_i = 0 \, \forall \, i$. $\square$

(vii) There are exactly $r = \operatorname{rank}[\mathfrak{g}]$ simple roots.

Proof. (v) and *(vi)* tell us that Φ_S is a basis of $\mathfrak{h}_{\mathbb{R}}^*$. Hence $|\Phi_S| = \dim \mathfrak{h}_{\mathbb{R}}^* = r$. $\square$

From now on we will label simple roots as $\alpha_{(i)}, i = 1, \ldots, r$.

Definition. The *Cartan matrix* of a Lie algebra is an $r \times r$ matrix A with components given by:

$$
A^{ij} = \frac{2(\alpha_{(i)}, \alpha_{(j)})}{(\alpha_{(j)}, \alpha_{(j)})} \in \mathbb{Z}
$$

Note that A is not generally symmetric.
We can choose $\{h^i = h^{\alpha_{(i)}}, e_\pm^i = e^{\pm \alpha_{(i)}}\}$ as a basis for the Lie algebra. Then the Lie algebra structure is as follows:

$$
\begin{aligned}
[h^i, e_\pm^i] &= \pm 2 e_\pm^i \\
[h^i, e_\pm^j] &= \pm A^{ji} e_\pm^j \\
[e_+^i, e_-^i] &= h^i
\end{aligned}
$$

with all other brackets $= 0$.

4.7 Classification

The Cartan classification of finite-dimensional simple complex Lie algebras has two steps:

1. Classify all possible Cartan matrices.

2. Show that a Cartan matrix uniquely determines each Lie algebra.

So first let us examine the constraints on the Cartan matrix:

(a) $A^{ii} = 2 \, \forall \, i$.

(b) Symmetry of the linear product gives $A^{ij} = 0 \iff A^{ji} = 0$.

(c) $(\alpha, \beta) \leq 0$ for $\alpha \neq \beta \in \Phi_S \implies A^{ij} \in \mathbb{Z}^{\leq 0}$ for all $i \neq j$.

(d) Euclidean inner product implies $\det A > 0$.

(e) $\mathfrak{g}$ simple implies A is not reducible.

Example. Consider $r = 2$. Then we have:

$$A = \begin{pmatrix} 2 & -m \\ -n & 2 \end{pmatrix}$$

where $m, n > 0$ and $\det A = 4 - mn > 0$. So, up to exchange of m and n, we have $(m, n) \in \{(1,1), (1,2), (1,3)\}$.

It can be shown that $A^{ij}A^{ji} \in \{0, 1, 2, 3\}$ (no summation), and that simple complex Lie algebras can have roots of at most 2 distinct lengths.

To represent the structure of a Lie algebra, we use *Dynkin diagrams*, which we construct in the following way:

- Draw a node $\circ$ for each simple root $\alpha_{(i)}$.

- Join the nodes representing $\alpha_{(i)}$ and $\alpha_{(j)}$ with $\max(|A^{ij}|, |A^{ji}|) \in \{0, 1, 2, 3\}$ lines.

- If more than one line connects two nodes, draw an arrowhead pointing from the longer root to the shorter one.

Example. In rank 2, there are three possible Cartan matrices and associated Dynkin diagrams:

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad \circ\!\!-\!\!\!-\!\!\circ$$

$$\begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad \circ\!\!\Rightarrow\!\!\circ$$

$$\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \quad \circ\!\!\Rrightarrow\!\!\circ$$

Theorem 3 (The Cartan classification). *All finite-dimensional complex simple Lie algebras must have Dynkin diagrams of one of the following forms:*
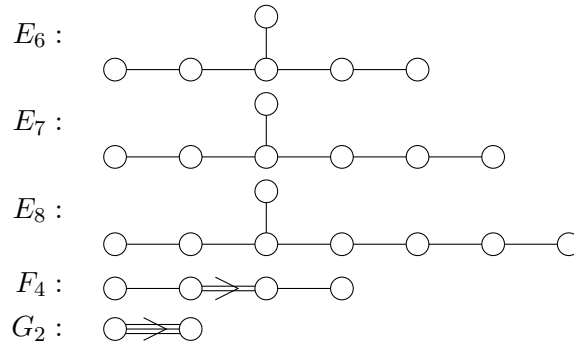

$$A_n: \quad \underset{1}{\circ}\!-\!\underset{2}{\circ}\!-\cdots-\!\underset{n-1}{\circ}\!-\!\underset{n}{\circ} \qquad (\mathcal{L}_{\mathbb{C}}(SU(n+1)))$$

$$B_n: \quad \underset{1}{\circ}\!-\!\underset{2}{\circ}\!-\cdots-\!\underset{n-1}{\circ}\!\Rightarrow\!\underset{n}{\circ} \qquad (\mathcal{L}_{\mathbb{C}}(SO(2n+1)))$$

$$C_n: \quad \underset{1}{\circ}\!-\!\underset{2}{\circ}\!-\cdots-\!\underset{n-1}{\circ}\!\Leftarrow\!\underset{n}{\circ} \qquad (\mathcal{L}_{\mathbb{C}}(SP(2n)))$$

$$D_n: \quad \underset{1}{\circ}\!-\!\underset{2}{\circ}\!-\cdots-\!\underset{n-2}{\circ}\!\begin{smallmatrix} \overset{n-1}{\circ} \\ \diagup \\ \diagdown \\ \underset{n}{\circ} \end{smallmatrix} \qquad (\mathcal{L}_{\mathbb{C}}(SO(2n)))$$

or be one of five exceptional cases:

E_6 :

E_7 :

E_8 :

F_4 :

G_2 :

Example. $A_1 = B_1 = C_1$, and D_1 isn't a thing. So there is only one possible finite-dimensional complex simple Lie algebra of rank 1. We can deduce:

$$\mathcal{L}_\mathbb{C}(SU(2)) \simeq \mathcal{L}_\mathbb{C}(SO(3)) \simeq \mathcal{L}_\mathbb{C}(SP(2))$$

Example. In rank 2, we recover the above Dynkin diagrams. The diagrams for B_2 and C_2 are identical, so we can say $\mathcal{L}_\mathbb{C}(SO(5)) \simeq \mathcal{L}_\mathbb{C}(SP(4))$. Note that D_2's diagram is the same as two A_1s, so we can say $\mathcal{L}_\mathbb{C}(SO(4)) \simeq \mathcal{L}_\mathbb{C}(SU(2)) \oplus \mathcal{L}_\mathbb{C}(SU(2))$.

Now that we know which Dynkin diagrams are avaliable, how do we reconstruct a Lie algebra $\mathfrak{g}$? First we must determine the Cartan matrix A, which can be easily read off from the diagram. A then determines the simple roots in $\mathfrak{h}_\mathbb{R}^*$ (up to one undetermined vector). From the simple roots we can then reconstruct all roots by forming root strings with length $l_{ij} = 1 - A_{ji} \in \mathbb{N}$.

Example. Consider the Dynkin diagram for A_2.

We can see that we must have Cartan matrix given by

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

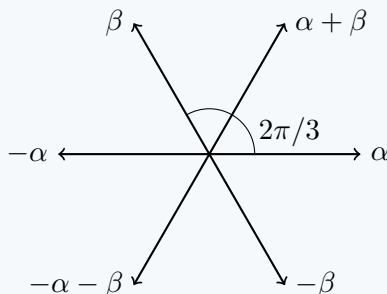Call the simple roots α and β. We have:

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

$$\implies \frac{2|\alpha|}{|\beta|} \cos \phi = \frac{2|\beta|}{|\alpha|} \cos \phi = -1$$

$$\implies |\beta| = |\alpha|, \ \phi = \frac{2\pi}{3}$$

34

The length of the α-string through β is $l_{\alpha,\beta} = 1 - \frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 2$, so $\alpha + \beta$ is a root. No extra roots arise from the other strings. We also have negative roots $-\alpha, -\beta, -\alpha - \beta$. So the full set of roots is $\Phi = \{\alpha, \beta, -\alpha, -\beta, \alpha + \beta, -\alpha - \beta\}$:

5 More representations

5.1 Representation theory of $\mathfrak{g}$

Suppose R is a representation of $\mathfrak{g}$. R takes elements of the Cartan-Weyl basis to $\mathrm{Mat}_N(\mathbb{C})$.

$$H^i \mapsto R(H^i)$$
$$E^\alpha \mapsto R(E^\alpha)$$

Let's assume $R(H^i)$ is diagonalisable. Since $[R(H^i), R(H^j)] = R([H^i, H^j]) = 0$, the $R(H^i)$ are simultaneously diagonalisable. $V = \mathbb{C}^N$ is spanned by the eigenvectors of $\{R(H^i)\}$. We can write $V = \bigoplus_{\lambda \in S_R} V_\lambda$, where if $v \in V_\lambda$ then $R(H^i)v = \lambda^i v$, and $\lambda^i \in \mathbb{C}, i = 1, \ldots, r$. λ is called a *weight* of R; the set of weights S_R is called its *weight set*. Weights in general can have non-trivial multiplicity $m_\lambda = \dim V_\lambda \geq 1$.

Example. The roots are the weights for the adjoint representation.

Consider the action of the representations of the step operators $R(E^\alpha)$ on $v \in V_\lambda$:

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

So $R(E^\alpha)v \in V_{\lambda+\alpha}$ if it is non-zero.

Recall that we have a set of $sl(2)$ generators for each simple root $\{h^i = h^{\alpha_{(i)}}, e^i_\pm = e^{\pm\alpha_{(i)}}\}$, obeying:

$$
\begin{aligned}
[h^i, h^j] &= 0 \\
[h^i, e^j_\pm] &= \pm A^{ji} e^j_\pm \\
[e^i_+, e^i_-] &= h^i
\end{aligned}
$$

with $[e^i_\pm, e^j_\pm]$ generating new basis elements, and all other brackets $= 0$. We have:

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

Since the $\alpha_{(i)}$-string through $\alpha_{(j)}$ has length $1 - A_{ji}$, we can deduce the *Serre relation*:

$$(\mathrm{ad}_{e_\pm^i})^{1 - A^{ji}} e_\pm^j = 0$$

It can be shown that the Serre relation together with the set of bracket relations for $sl(2)_{\alpha_{(i)}}$ uniquely determines $\mathfrak{g}$; this is called *Serre reconstruction.*
Considering the action of the $sl(2)_\alpha$ generators on V, we can see that V is a representation space to some representation R_α of $sl(2)_\alpha$ (in general R_α is reducible). For all $v \in V_\lambda$, we have:

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

So $\frac{2(\alpha,\lambda)}{(\alpha,\alpha)} \in \mathbb{Z}$ for all $\lambda \in S_R, \alpha \in \Phi$.

5.2 Root and weight lattices

Definition. The *root lattice* $\mathcal{L}[\mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$ with simple roots $\alpha_{(i)}$ is defined by:

$$\mathcal{L}[\mathfrak{g}] = \mathrm{span}_\mathbb{Z}\{\alpha_{(i)}\}$$

Definition. The *simple coroots* $\check{\alpha}_{(i)}$ of a Lie algebra $\mathfrak{g}$ with simple roots $\alpha_{(i)}$ are given by:

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

Definition. The *coroot lattice* $\check{\mathcal{L}}[\mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$ with simple coroots $\check{\alpha}_{(i)}$ is defined as:

$$\check{\mathcal{L}}[\mathfrak{g}] = \mathrm{span}_\mathbb{Z}\{\check{\alpha}_{(i)}\}$$

Definition. The *weight lattice* $\mathcal{L}_W[\mathfrak{g}]$ of a Lie algebra $\mathfrak{g}$ is the dual lattice of its coroot lattice:

$$\mathcal{L}_W[\mathfrak{g}] = \check{\mathcal{L}}^*[\mathfrak{g}] = \{\lambda \in \mathfrak{h}_\mathbb{R}^* \,\mathrm{s.\,t.}\,(\lambda, \mu) \in \mathbb{Z}\,\forall\,\mu \in \check{\mathcal{L}}[\mathfrak{g}]\}$$

In particular:

$$\lambda \in \mathcal{L}_W[\mathfrak{g}] \iff (\lambda, \check{\alpha}_{(i)}) = \frac{2(\alpha_{(i)}, \lambda)}{(\alpha_{(i)}, \alpha_{(i)})} \in \mathbb{Z}$$

So the weights of any representation R of $\mathfrak{g}$ all lie in $\mathcal{L}_W[\mathfrak{g}]$. The simple coroots are a basis for the coroot lattice, so we have a dual basis for the weight lattice $\{\omega_{(i)}\}$ where:

$$(\check{\alpha}_{(i)}, \omega_{(j)}) = \frac{2(\alpha_{(i)}, \omega_{(j)})}{(\alpha_{(i)}, \alpha_{(i)})} = \delta_{ij} \tag{$*$}$$

The $\omega_{(i)}$ are called *fundamental weights* of $\mathfrak{g}$. Since the simple roots span $\mathfrak{h}^*_{\mathbb{R}}$, we can write $\omega_{(i)} = \sum_{j=1}^r B_{ij} \alpha_{(j)}$ for some real matrix B. With ($*$) this yields:

$$\sum_{k=1}^r \frac{2(\alpha_{(i)}, \alpha_{(k)})}{(\alpha_{(i)}, \alpha_{(i)})} B_{jk} = A^{ki} B_{jk} = \delta^i_j$$

In other words, B is the inverse of A, so we can write:

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

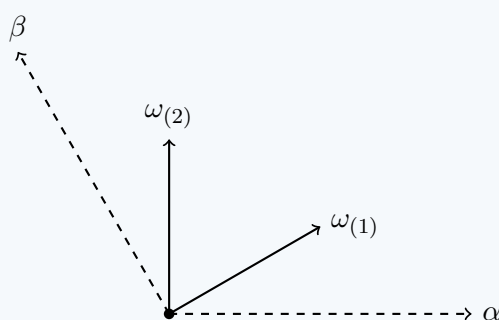**Example.** Consider $\mathfrak{g} = A_2$. Then we have shown:

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

so:

$$\alpha = \alpha_{(1)} = 2\omega_{(1)} - \omega_{(2)}$$
$$\beta = \alpha_{(2)} = -\omega_{(1)} + 2\omega_{(2)}$$

Inverting these, we find:

$$\omega_{(1)} = \frac{1}{3}(2\alpha + \beta)$$
$$\omega_{(2)} = \frac{1}{3}(\alpha + 2\beta)$$

Any weight $\lambda \in S_R \subset \mathcal{L}_W[\mathfrak{g}]$ can be written as a linear combination of the fundamental weights:

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

$\lambda^i \in \mathbb{Z}$ are known as the *Dynkin labels* of λ.

Every finite dimensional representation R of $\mathfrak{g}$ has a highest weight, i.e. a $\Lambda \in S_R$ such that $R(E^\alpha)v = 0$ for all $\alpha \in \Phi_+$ and $v \in V_\Lambda$. If the highest weight is unique and non-degenerate, then R_Λ is a finite dimensional irreducible representation. The rest of the states in the irrep can be found by acting with lowering generators $R(E^{-\alpha}), \alpha \in \Phi_+$ on V_Λ:

$$v_\lambda = R(E^{-\alpha_1})R(E^{-\alpha_2})\dots R(E^{-\alpha_l})v_\Lambda$$

Every weight of the representation, which we will call R_Λ, can be written as:

$$\lambda = \Lambda - \mu \text{ where } \mu = \sum_{i=1}^r \mu^i \alpha_{(i)}, \mu^i \in \mathbb{Z}, \mu^i \geq 0$$

The highest weight is also sometimes called the *dominant weight.*
The following is a useful result:

𝔏emma 13. *For any finite dimensional representation of $\mathfrak{g}$, if $\lambda = \sum_{i=1}^r \lambda^i \omega_{(i)} \in S_R$ then $\lambda - m_{(i)}\alpha_{(i)} \in S_R$ where $m_{(i)} \in \mathbb{Z}, 0 \leq m_{(i)} \leq \lambda^i$.*

Example. Consider $\mathfrak{g} = A_2$, the fundamental representation has highest weight Dynkin labels $(1, 0) = (\Lambda^1, \Lambda^2)$. This implies $\Lambda = \omega_{(1)} \in S_f$, which then by the proposition implies $\Lambda - \alpha_{(1)} = -\omega_{(1)} + \omega_{(2)} \in S_f$. Applying the proposition again, we have $-\omega_{(1)} + \omega_{(2)} - \alpha(2) = -\omega_{(2)} \in S_f$. There are no more weights, so we have:

$$S_f = \{\omega_{(1)}, -\omega_{(1)} + \omega_{(2)}, -\omega_{(2)}\}$$

Example. Consider general irreps of A_2. For each dominant integral weight Λ, we can write $\Lambda = \Lambda^1 \omega_{(1)} + \Lambda^2 \omega_{(2)}$, with $0 \leq \Lambda^1, \Lambda^2 \in \mathbb{Z}$, and we get an irrep $R_{(\Lambda^1, \Lambda^2)}$ of A_2. It can be shown that:

- $\dim R_{(\Lambda^1, \Lambda^2)} = \frac{1}{2}(\Lambda^1 + 1)(\Lambda^2 + 1)(\Lambda^1 + \Lambda^2 + 2)$.

- If $\Lambda^1 \neq \Lambda^2$ then we have $R_{(\Lambda^1, \Lambda^2)} = \bar{R}_{(\Lambda^2, \Lambda^1)}$ and $\lambda \in S_{(\Lambda^1, \Lambda^2)} \implies -\lambda \in S_{(\Lambda^2, \Lambda^1)}$.

---

6 Symmetries in Quantum Mechanics

Consider a generic quantum mechanical system with energy levels $E_0 < E_1 < E_2 < \dots$ for a Hamiltonian $\hat{H}$. The states of the system are elements of a Hilbert space:

$$\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n \text{ where } \hat{H}|\psi\rangle = E_n|\psi\rangle \,\forall\, |\psi\rangle \in \mathcal{H}_n$$

Definition. A *symmetry transformation* of a system is a transformation $|\psi\rangle \mapsto |\psi'\rangle = \hat{U}|\psi\rangle$, where $\hat{U} : \mathcal{H} \to \mathcal{H}$ is a unitary operator such that $\hat{U}\hat{H}\hat{U}^\dagger = \hat{H}$.

Under a symmetry transformation, the inner product is preserved, and the energy is invariant.

> **Definition.** A *conserved quantity* is an observable $\hat{I} = \hat{I}^\dagger$ such that $[\hat{I}, \hat{H}] = 0$.

If $\hat{I}$ is conserved, then $\hat{U} = \exp\left(is\hat{I}\right), s \in \mathbb{R}$ is a symmetry transformation.

Suppose we have a maximal set of linearly independent conserved quantities $\{\hat{I}^a \,\text{s.t.} [\hat{I}^a, \hat{H}] = 0, a = 1, \ldots, d\}$. Then $\mathfrak{g}_\mathbb{R} = \text{span}_\mathbb{R}\{i\hat{I}^a; a = 1, \ldots, d\}$ is a real Lie algebra with bracket equal to the operator commutator.

Consider all symmetry transformations of the form $\hat{U} = \exp\left(\hat{X}\right)$ where $\hat{X} \in \mathfrak{g}_\mathbb{R}$. These form a Lie group G, and moreover G is compact (since G is a subgroup of some product of unitary groups). As $[\hat{X}, \hat{H}] = 0 \,\forall\, \hat{X} \in \mathfrak{g}_\mathbb{R}$, the $\mathcal{H}_n$ are invariant under the action of G or $\mathfrak{g}_\mathbb{R}$.

Each $\mathcal{H}_n$ carries a representation D_n of G, d_n of $\mathfrak{g}_\mathbb{R}$ such that $D_n(\hat{U}) = \exp\left(d_n(\hat{X})\right) \in \text{Mat}_{N_n}(\mathbb{C})$ where $N_n = \dim \mathcal{H}_n$. This representation must be *unitary*, i.e.:

$$D_n^{-1}(\hat{U}) = D_n(\hat{U}) \text{ or equivalently } d_n(\hat{X})^\dagger = -d_n(\hat{X})$$

Note that all finite dimensional representations of a compact G are automatically unitary.

6.1 Hadronic physics

Hadronic physics is the study of the hadrons: mesons (which are bosons) and baryons (which are fermions). The following isn't the most modern approach, but it is one of the first and simplest examples of the application of Lie theory to particle physics. We will refer to *isospin I* and *hypercharge Y*. These are quantum numbers associated with the strong interaction; it is not necessary to know their details. The first few hadrons to be known about are listed below.

|	Q	M	I		Q	M	I
π^+	$+1$	$139\,\text{MeV}$	$+1$	p	$+1$	$938\,\text{MeV}$	$+\frac{1}{2}$
π^0	0	$135\,\text{MeV}$	0	n	0	$940\,\text{MeV}$	$-\frac{1}{2}$
π^-	-1	$139\,\text{MeV}$	$+1$				

(a) Mesons.   (b) Baryons.

Table 1: The first few hadrons discovered.

It was observed that isospin obeys an approximate $SU(2)$ symmetry, with Cartan element $H = 2I$. The proton and the neutron sit in the R_1 representation, while the π mesons sit in the R_2 representation:

$$\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \sim R_1 \qquad\qquad \begin{pmatrix} |\pi^+\rangle \\ |\pi^0\rangle \\ |\pi^-\rangle \end{pmatrix} \sim R_2$$

Soon many more mesons and baryons were discovered, along with the new conserved quantum number hypercharge and many unexplained approximate degeneracies. We can plot the 8 lightest mesons in the *I-Y* plane:
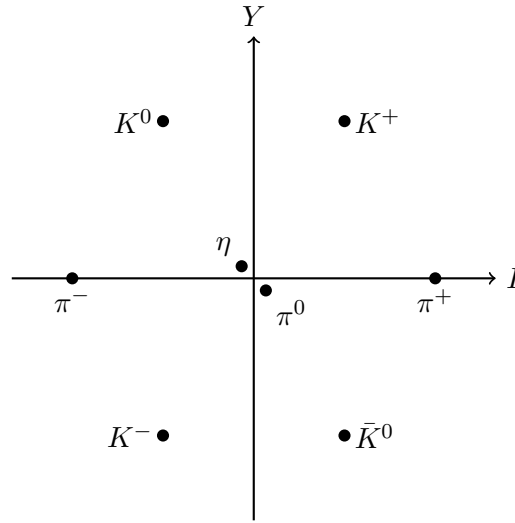
The 8 lightest baryons form a similar shape:
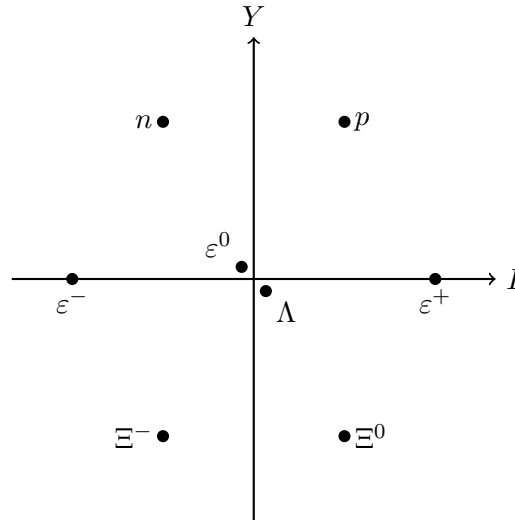
Consider now the Lie algebra $\mathfrak{g}_{\mathbb{R}} = \mathcal{L}(SU(3))$, and in particular its representation $\underline{8} = R_{(1,1)}$. If we choose the following Cartan generators:

$$H^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad\qquad H^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and set $I = \frac{1}{2}H^1, Y = \frac{1}{3}(H^1 + 2H^2)$, then it is easy to calculate that the weights of $R_{(1,1)}$ match the quantum numbers plotted above. If we do a similar investigation for the other known mesons and baryons, we find:

$$\text{Mesons live in: } \underline{1} = R_{(0,0)} \text{ and } \underline{8} = R_{(1,1)},$$
$$\text{Baryons live in: } \underline{8} = R_{(1,1)}, \ \underline{10} = R_{(3,0)} \text{ and } \underline{\overline{10}} = R_{(0,3)}$$

The immediate question to ask is: why only $\underline{1}, \underline{8}$ and $\underline{10}$, and why do the mesons and baryons only live in their corresponding representations? There are more representations of $SU(3)$, e.g. $\underline{3}, \underline{\bar{3}}, \underline{6}, \underline{\bar{6}}, \ldots$

The answer came from Gell-Mann, who proposed that all hadrons are composed of smaller more fundamental fermions called *quarks*. Quarks live in $\underline{3} = R_{(1,0)}$, and antiquarks live in $\underline{\bar{3}} = R_{(0,1)}$. The key insight is to say that mesons are combinations of a quark and an antiquark, and that baryons are combinations of three quarks (or three antiquarks).

Consider a meson in this regime. It has state given by $|\psi_1\rangle \otimes |\psi_2\rangle \in \mathcal{H}_{(1)} \otimes \mathcal{H}_{(2)}$, where $|\psi_1\rangle, |\psi_2\rangle$ are the states of the quark and antiquark respectively. Suppose we have a conserved quantity $\hat{I}_1$ on $\mathcal{H}_{(1)}$ and $\hat{I}_2$ on $\mathcal{H}_{(2)}$. The operator corresponding to their sum for the entire meson is given by:

$$I = \hat{I}_1 \otimes \mathbb{1}_{(2)} + \mathbb{1}_{(1)} \otimes \hat{I}_2$$

We see that this is like the elements of the tensor product of two representation. Thus we can deduce that mesons $(q\bar{q})$ live in $\underline{3} \otimes \underline{\bar{3}} = \underline{1} \oplus \underline{8}$. Similarly we have that baryons (qqq) live in $\underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10}$. So we see exactly why the baryons and mesons have their representations. In order to see exactly why we only have $q\bar{q}$ and qqq, it is necessary to study gauge QCD.

6.2   Gauge Theory

Definition. A *gauge symmetry* is a redundancy in the description of a system.

Example. In classical electromagnetism, we define the non-observable scalar potential $\Phi(\mathbf{x}, t)$ and vector potential $\mathbf{A}(\mathbf{x}, t)$. What we then observe are the electric field $\mathbf{E} = -\boldsymbol{\nabla}\Phi + \frac{\partial \mathbf{A}}{\partial t}$ and magnetic field $\mathbf{V} = \boldsymbol{\nabla} \times \mathbf{A}$. $\mathbf{E}$ and $\mathbf{B}$ are invariant under a gauge transformation given by:

$$\Phi \to \Phi + \frac{\partial \alpha}{\partial t}$$
$$\mathbf{A} \to \mathbf{A} + \boldsymbol{\nabla}\alpha$$

where $\alpha = \alpha(\mathbf{x}, t)$ is some time-varying field. In a relativistic treatment we have a 4-vector potential a_μ given by:

$$a_\mu = \begin{cases} \Phi & \text{if } \mu = 0 \\ A_i & \text{if } \mu = i = 1, 2, 3 \end{cases}$$

and gauge transformations $a_\mu \to a_\mu + \partial_\mu \alpha$. The electric and magnetic fields live in the field-strength tensor $f_{\mu\nu}$:

$$f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu \qquad\qquad E_i = f_{0i}, \; B_i = \frac{1}{2}\epsilon_{ijk}f_{jk}$$

We have a Lagrangian given by:

$$\mathcal{L}_{\text{EM}} = -\frac{1}{4g^2} f_{\mu\nu} f^{\mu\nu}$$

where we use a Minkowski metric with signature $+---$.

Quantisation leads to a spin 1, massles, free particle that we call the *photon*. For later convenience, we will define $A_\mu = -ia_\mu$ and $F_{\mu\nu} = -if_{\mu\nu}$.

Consider a complex scalar field $\phi : \mathbb{R}^{3,1} \to \mathbb{C}$ with Lagrangian given by:

$$\mathcal{L}_\phi = \partial_\mu \phi^* \partial^\mu \phi - \underbrace{W(\phi^* \phi)}_{\text{interactions}}$$

This Lagrangian is invariant under a global $U(1)$ symmetry:

$$\phi \to g\phi, \ \phi^* \to g^{-1}\phi^* \text{ where } g = \exp(i\delta) \in U(1), \ \delta \in [0, 2\pi)$$

Consider infinitesimal such transformations, i.e. those given by $g = \exp(\epsilon X)$ where $\epsilon \ll 1$ and $X \in i\mathbb{R} = \mathcal{L}(U(1))$. We have that $g \approx 1 + \epsilon X$ so these transformations are:

$$\phi \to \phi + \delta_X \phi \text{ where } \delta_X \phi = \epsilon X \phi$$

In general we will write $\delta_X x$ for the difference in a quantity x after the transformation. Working to first order in ϵ, we have $\delta_X(\phi\phi^*) = 0$ and $\delta_X \mathcal{L} = 0$ for all $X \in i\mathbb{R}$.

Now generalise X to depend on spacetime. We have $\delta_X(\partial_\mu \phi) = \partial_\mu(\delta_X \phi) = \partial_\mu(\epsilon X \phi) = \epsilon \partial_\mu X \phi + \epsilon X \partial_\mu \phi$. Now $\mathcal{L}_\phi$ is no longer invariant in general. In order to restore gauge-invariance, we replace all partial derivatives in the Lagrangian with a *covariant derivative* D_μ, defined in the following way:

$$D_\mu = \partial_\mu + A_\mu \text{ where } A_\mu : \mathbb{R}^{3,1} \to \mathcal{L}(U(1)) = i\mathbb{R}$$

A_μ is known as a $U(1)$ *gauge field* and must transform as $\delta_X A_\mu = -\epsilon \partial_\mu X$. Then we have:

$$\begin{aligned}
\delta_X(D_\mu \phi) &= \delta_X(\partial_\mu \phi + A_\mu \phi) \\
&= \partial_\mu(\delta_X \phi) + (\delta_X A_\mu)\phi + A_\mu(\delta_X \phi) \\
&= \partial_\mu(\epsilon X \phi) + A_\mu \epsilon X \phi - \epsilon \partial_\mu X \phi \\
&= \epsilon \partial_\mu X \phi + \epsilon X \partial_\mu \phi + A_\mu \epsilon X \phi - \epsilon \partial_\mu X \phi \\
&= \epsilon X D_\mu \phi
\end{aligned}$$

Thus, to first order, we have that $(D_\mu \phi)^*(D^\mu \phi)$ is gauge invariant. The most general gauge invariant Lagrangian we can have is:

$$\mathcal{L} = \frac{1}{4g^2} F_{\mu\nu} F^{\mu\nu} + (D_\mu \phi)^*(D^\mu \phi) - W(\phi^* \phi)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. It can be shown that if we want our theory to be renormalisable, we must further have $W(Y) = \alpha Y + \beta Y^2$.

Now we will generalise this to any Lie group G. Choose a representation D of G of dimension N with representation space $V \simeq \mathbb{C}^N$. We use the standard inner product $(u, v) = u^\dagger v$. Suppose we have a V-valued scalar field $\phi : \mathbb{R}^{3,1} \to V$, with the following Lagrangian:

$$\mathcal{L}_\phi = (\partial_\mu \phi, \partial^\mu \phi) - W[(\phi, \phi)]$$

If D is unitary (so $D(g)^\dagger = D(g) = \mathbb{1}$), then $\mathcal{L}_\phi$ is invariant under $\phi \to D(g)\phi$ for all $g \in G$. Near the identity, write $g = \exp(\epsilon X)$ where $\epsilon \ll 1$ and $X \in \mathcal{L}(G)$. Then we have $D(g) = \exp(\epsilon R(X))$ where $R : \mathcal{L}(G) \to \text{Mat}_N(\mathbb{C})$ defines a unitary representation of $\mathcal{L}(G)$ (so $R(X) = -R(X)^\dagger$). For $\epsilon \ll 1$, we have $D(g) \approx \mathbb{1} + \epsilon R(X)$. Under an infinitesimal transformation:

$$\phi \to \phi + \delta_X \phi, \ \delta_X \phi = \epsilon R(X)\phi$$

Now let's do what we did before and "gauge" the symmetry i.e. we let X be a $\mathcal{L}(G)$ valued function of spacetime. Define the *covariant derivative* by:

$$D_\mu(\phi) = \partial_\mu \phi + R(A_\mu)\phi \text{ where } A_\mu : \mathbb{R}^{3,1} \to \mathcal{L}(G)$$

A_μ is the gauge field for G, and must transform as:

$$\delta_X A_\mu = -\epsilon \partial_\mu X + \epsilon[X, A_\mu] \in \mathcal{L}(G)$$

To see why, consider the action of the transformation on $D_\mu \phi$:

$$
\begin{aligned}
\delta_X(D_\mu \phi) &= \delta_X(\partial_\mu \phi + R(A_\mu)\phi) \\
&= \partial_\mu(\delta_X \phi) + R(A_\mu)\delta_X \phi + R(\delta_X A_\mu)\phi \\
&= \partial_\mu(\epsilon R(X)\phi) + \epsilon R(A_\mu)R(X)\phi - \epsilon R(\partial_\mu X) + \epsilon R([X, A_\mu]) \\
&= \epsilon(\partial_\mu R(X))\phi + \epsilon R(X)\partial_\mu \phi + \epsilon R(X)R(A_\mu)\phi \\
&\quad + \epsilon[R(A_\mu), R(X)]\phi - \epsilon R(\partial_\mu X)\phi + \epsilon[R(X), R(A_\mu)]\phi \\
&= \epsilon R(X)D_\mu \phi
\end{aligned}
$$

Thus:

$$
\delta_X[(D_\mu \phi, D^\mu \phi)] = \epsilon(R(X)D_\mu \phi, D^\mu \phi) + \epsilon(D_\mu \phi, R(X), D^\mu \phi) = 0
$$

since $R(X)^\dagger = -R(X)$. Hence the Lagrangian is conserved.

Define the field-strength tensor as:

$$
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]
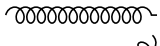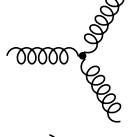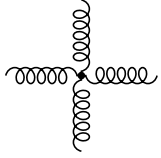$$

Under a gauge transformation, we have:

$$
\begin{aligned}
\delta_X(F_{\mu\nu}) &= \partial_\mu(\delta_X A_\nu) - \partial_\nu(\delta_X A_\mu) + [\delta_X A_\mu, A_\nu] + [A_\mu, \delta_X A_\nu] \\
&= -\epsilon\partial_\mu\partial_\nu X + \epsilon\partial_\mu([X, A_\nu]) + \epsilon\partial_\nu\partial_\mu X - \epsilon\partial_\nu([X, A_\mu]) \\
&\quad - \epsilon[\partial_\mu X, A_\nu] - \epsilon[A_\mu, \partial_\nu X] + \epsilon[[X, A_\mu], A_\nu] + \epsilon[A_\mu, [X, A_\nu]] \\
&= \epsilon[X, \partial_\mu A_\nu] - \epsilon[X, \partial_\nu A_\mu] - \epsilon([A_\nu, [X, A_\mu]] + [A_\mu, [A_\nu, X]]) \\
&= \epsilon[X, \partial_\mu A_\nu - \partial_\nu A_\mu] + \epsilon[X, [A_\mu, A_\nu]] \\
&= \epsilon[X, F_{\mu\nu}]
\end{aligned}
$$

So for a gauge invariant Lagrangian that includes $F_{\mu\nu}$, we can use the Killing form:

$$
\begin{aligned}
\mathcal{L}_A &= \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) \\
\implies \delta_X \mathcal{L}_A &= \frac{1}{g^2}[\kappa(\delta_X F_{\mu\nu}, F^{\mu\nu}) + \kappa(F_{\mu\nu}, \delta_X)] \\
&= \frac{\epsilon}{g^2}[\kappa([X, F_{\mu\nu}], F^{\mu\nu}) + \kappa(F_{\mu\nu}, [X, F^{\mu\nu}])] \\
&= 0
\end{aligned}
$$

This provides a sensible kinetic term if $\mathcal{L}(G)$ is of compact type. Note that $\mathcal{L}_A$ contains both kinetic terms and interaction terms; if we take $A \to gA$ then we have:

$\mathcal{L}_A \sim$	$\partial A \partial A$		kinetic term
	$+g[A, A]\partial A$		3-vertex
	$+g^2[A, A]^2$		4-vertex

This is *Yang-Mills* theory.

There is a large family of consistent gauge theories provided by the Cartan classification. We specify the following:

- The gauge group is specified by the selection of a complex simple Lie algebra $\mathfrak{g}_{\mathbb{C}}$. This has a real form of compact type $\mathfrak{g}_{\mathbb{R}}$ that is equal to $\mathcal{L}(G)$ for a compact Lie group G.

- The matter content is specified by choosing a unitary representation of $\mathfrak{g}_{\mathbb{R}}$ for matter to live in. Then the matter field is $\phi_{\Lambda} : \mathbb{R}^{3,1} \to V_{\Lambda}$, where V_{Λ} is the representation space, and the subscript Λ specifies the dominant integral weights of the representation: $\Lambda \in S = \{\text{dominant integral weights}\}$.

Then the most general possible Lagrangian for our theory is:

$$\mathcal{L}_{\Lambda} = \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}) + \sum_{\Lambda \in S}(D_{\mu}\phi_{\Lambda}, D^{\mu}\phi_{\Lambda}) - W\left(\{(\phi_{\Lambda}, \phi_{\Lambda});\ \Lambda \in S\}\right)$$

6.3 The Standard Model

The gauge group of the Standard Model is:

$$G = U(1) \times SU(2) \times SU(3)$$

This gives rise to the following types of particles:

- Gauge terms give rise to the gauge bosons:

$$A_{\mu} : \mathbb{R}^{3,1} \to \underbrace{\mathcal{L}(U(1))}_{\text{photon}} \oplus \underbrace{\mathcal{L}(SU(2))}_{W^{\pm}, Z} \oplus \underbrace{\mathcal{L}(SU(3))}_{\text{8 gluons}}$$

The part of the Lagrangian which corresponds to the gauge bosons is:

$$\mathcal{L}_g = \frac{1}{4g_{U(1)}^2}F_{\mu\nu}^{(1)}F^{(1)\,\mu\nu} + \frac{1}{g_{SU(2)}^2}\kappa\left(F_{\mu\nu}^{(2)}, F^{(2)\,\mu\nu}\right) + \frac{1}{g_{SU(3)}^2}\kappa\left(F_{\mu\nu}^{(3)}, F^{(3)\,\mu\nu}\right)$$
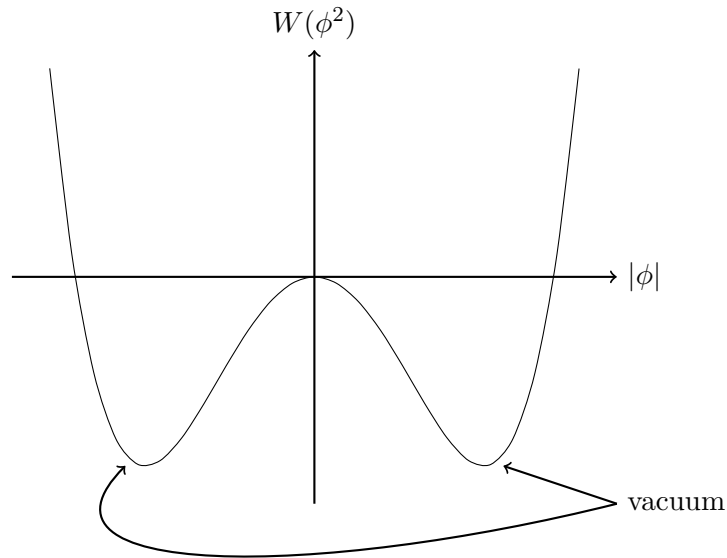
- We have a single scalar particle $\phi : \mathbb{R}^{3,1} \to \mathbb{C}^2$ (the representation space for R_1), the Higgs boson. Higgs boson states live in the following representation:

$$(\underbrace{\mathbf{2}}_{SU(2)}, \underbrace{\mathbf{1}}_{SU(3)})_{+\frac{1}{2}} \leftarrow \text{U(1) charge}$$

The Lagrangian for the Higgs particle is:

$$\mathcal{L}_s = (D_{\mu}\phi, D^{\mu}\phi) - W(\phi^2)$$

where $\phi^2 = (\phi, \phi)$ and $W(\phi^2) = -\mu^2\phi^2 + \lambda\phi^4$. If we plot $W(\phi^2)$ against $|\phi|$ we can see that the vacuum of the Higgs boson is not at $|\phi| = 0$:

$$W(\phi^2)$$

$$|\phi|$$

vacuum

- Finally we have 3 generations of fermions, divided into two types:
    - The quarks in the following representation:

$$\underbrace{(\underline{2},\underline{3})_{\frac{1}{6}}}_{\text{LH spinors}} \oplus \underbrace{(\underline{1},\underline{3})_{\frac{2}{3}} \oplus (\underline{1},\underline{3})_{-\frac{1}{3}}}_{\text{RH spinors}}$$

    - The leptons in the following representation:

$$(\underline{2},\underline{1})_{-\frac{1}{2}} \oplus (\underline{1},\underline{1})_{-1}$$

The terms for the Lagrangian for the fermions are each given by:

$$\mathcal{L}_f = \bar{\psi}^\alpha \gamma^\mu_{\alpha\beta} D_\mu \psi^\alpha + \phi\psi\bar{\psi}$$

The $\phi\psi\bar{\psi}$ terms are known as *Yukawa couplings.*

$$\mathscr{F}\!in$$