Symmetries of Particles and Fields 2019/20

1 Introduction

The laws of physics are invariant under many different symmetries:

- Parity, time reversal, charge conjugation (P, T, & C) - discrete symmetries
- Translations and rotations (Galilean invariance)
- Lorentz transformations (special relativity)
- Internal symmetries (e.g. isospin, flavour $SU(3)$)
- Gauge symmetries (Electrodynamics (Maxwell)), the Standard Model of particle physics: $SU(3) \otimes SU(2) \otimes U(1)$)

We shall study each of these in turn. The overall structure of the course is:

- Introduction to Lie groups and Lie algebras
- Space-time symmetries; constructing actions with symmetries
- Compact groups – representations
- Applications in physics: Noether's theorem; isospin; quark model; spontaneous symmetry breaking; chiral symmetry; gauge theories; QCD; Higgs mechanism; electroweak theory

This is a theoretical physics course, not a mathematics course. We shall introduce the necessary mathematics as we go along, mostly without mathematical rigour and often without proof. A more formal treatment is provided by the School of Mathematics course *Introduction to Lie Groups (MATH11053)*, which is unfortunately not running this year!

2 Lie Groups

Definition 1 *A* group *is a set of elements $G = \{g\}$ with a composition law $G \times G \to G$: The product of any two group elements g_1, g_2 is itself a group element: $\forall\, g_1,\, g_2 \in G,\ g_1\, g_2 = g_3$ for some $g_3 \in G$, with the properties*

1. *Associativity: $(g_1 g_2)\, g_3 = g_1\, (g_2 g_3)$*
2. *Unit element (or identity element) 1: $g1 = 1g = g$*
3. *Unique inverse: $gg^{-1} = g^{-1}g = 1$*

Definition 2 *A* finite group *is a group with a finite number of elements.*

Example.

- $\mathbb{Z}_2$, elements $\{+1, -1\}$. C, P, T (charge conjugation, parity, time reversal) are all $\mathbb{Z}_2$ groups.
- $\mathbb{Z}_N$, elements $\{e^{2\pi i n/N} : n = 1, \ldots, N\}$, the N^{th} roots of unity.
- Dihedral groups D_N: the group of symmetries of regular polygons with N sides, etc. [See the *Symmetries of Quantum Mechanics* course.]

Definition 3 *An* infinite group *is a group with an infinite number of elements, which we will assume to be labelled by a set of continuous parameters $\{\alpha\} = \{\alpha_1, \alpha_2, \ldots\}$. We write $G = \{g(\alpha)\}$. The* dimension *of the group is the number of independent parameters.* [1]

Example. In ordinary three-dimensional (real) space:

- Spatial translations in $\mathbb{R}^3$: $x_i \mapsto x_i + a_i$; $\dim G = 3$.
- Spatial rotations in $\mathbb{R}^3$: $x_i \mapsto R_{ij}\left(\underline{n}, \theta\right) x_j$;
  $R^T R = 1$; $\dim G = 3$ – an angle (ϕ) and an axis $\underline{n}$ (2 parameters).

Definition 4 *A* Lie Group *is a continuous group, for which the group multiplication has an analytic structure: if $g(\alpha_i) = g(\beta_i)\, g(\gamma_i)$ then $\alpha_i = \phi_i(\beta_i, \gamma_i)$ where the ϕ_i are analytic functions of the β_i and γ_i.*

Continuous groups may be compact (the group space is bounded) or non-compact (the group space is not bounded), and connected or not connected.[2]

Example.

- Translations: connected but not compact – the magnitude of a translation is not bounded
- Rotations + reflections: compact but not connected: $\det R = \pm 1$.

  Rotations give $\det R = 1$, reflections give $\det R = -1$, we can't get from $+1$ to -1 continuously, so rotations and reflections are not connected.

---

[1] Recall that the *order* of a group is the total number of *elements* in the group. This is different from the *dimension* of the group, which is the number of independent parameters $\{\alpha\}$.

[2] We'll be a bit more precise when we study specific examples! In previous years, we used topological definitions, which were unfamilar to many students.

2.1 Metric spaces

Metrics will feature in many parts of this course; let's define (or revise) their properties here.

Consider a *real* vector space with a basis $\{x^\mu\}$, $\mu = 1, \ldots, N$. Define a real symmetric matrix called the *metric* $g_{\mu\nu}$, then define $x^2 \equiv g_{\mu\nu} x^\mu x^\nu$ (using the Einstein summation convention). Now *choose* a basis in which:

1. $g_{\mu\nu}$ is diagonal, $g_{\mu\nu} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ where λ_i are the (real) eigenvalues of $g_{\mu\nu}$.
2. The eigenvalues are either zero or ± 1. Under the rescaling $x^\mu \mapsto s x^\mu$, $g_{\mu\nu} \mapsto s^{-2} g_{\mu\nu}$, with $s \in \mathbb{R}$, then x^2 stays fixed. We can rescale each λ_i separately. Since $s^2 > 0$ the *signs* of the eigenvalues cannot change, so in this basis $g_{\mu\nu} = \mathrm{diag}(1, 1, \ldots, 0, 0, \ldots, -1, -1)$. We call this the *canonical* basis.

For example, in three dimensional space $g_{ij} = \delta_{ij}$, with $i, j \in \{1, 2, 3\}$, $x^2 = x_1^2 + x_2^2 + x_3^2$ (Euclidean space).

In four dimensional space-time $g_{\mu\nu} = \mathrm{diag}(1, 1, 1, -1)$, $x^2 = \underbrace{x_1^2 + x_2^2 + x_3^2}_{\text{space}} - \underbrace{x_4^2}_{\text{time}}$. Clearly, x^2 can be positive, negative or zero.

We will generally write $x_4 \equiv x_0$ and use a space-time metric with the *opposite signature*: $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$ whereupon $x^2 = \underbrace{x_0^2}_{\text{time}} - \underbrace{(x_1^2 + x_2^2 + x_3^2)}_{\text{space}}$.

A metric with a mix of $+1$ and -1 elements is called an *indefinite metric*.

Metrics with zero eigenvalues occur in the theory of Lie algebras - see later.

Raising and lowering indices: if $\det g_{\mu\nu} \neq 0$, i.e., the metric has no zero eigenvalues, then $g_{\mu\nu}$ is invertible. We can then use the metric to raise and lower indices: $x_\mu \equiv g_{\mu\nu} x^\nu$ and so forth. Then

$$x^2 = g_{\mu\nu} x^\mu x^\nu = x_\mu x^\mu = x^\mu x_\mu = \delta_\mu{}^\nu x^\mu x_\nu\,.$$

where we used $x_\mu = \delta_\mu{}^\nu x_\nu$ (relabelling of indices), with

$$\delta_\mu{}^\nu = \delta^\nu{}_\mu = \begin{cases} 1 & \mu = \nu \\ 0 & \mu \neq \nu \end{cases}$$

We can also raise indices on the metric: $g^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} g_{\alpha\beta} \Rightarrow g^{\nu\beta} g_{\alpha\beta} = \delta_\alpha{}^\nu$, i.e., $g^{\mu\nu}$ is the *inverse* of $g_{\mu\nu}$.

Similarly for the totally antisymmetric Levi-Civita symbol $\epsilon^{\overbrace{\mu\nu\cdots\alpha\beta}^{N \text{ indices}}}$

$$\epsilon^{\mu\nu\cdots\alpha\beta} = g^{\mu\mu'} g^{\nu\nu'} \cdots g^{\alpha\alpha'} g^{\beta\beta'} \epsilon_{\mu'\nu'\cdots\alpha'\beta'},$$

and derivatives with respect to x^μ

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = g_{\mu\nu} \frac{\partial}{\partial x_\nu} = g_{\mu\nu}\, \partial^\nu \qquad \text{(exercise: check that } \frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha{}_\beta\text{, etc.)}$$

If the metric is *definite* (all the elements have the same sign, with no zeros) we can choose a basis in which $g_{\mu\nu} = \delta_{\mu\nu}$. In this case we can ignore the distinction between upper and lower indices.

2.2 Matrix Groups

A useful way of defining groups is through the linear transformation

$$x_\mu \mapsto D_\mu{}^\nu(\alpha)\, x_\nu \qquad \mu, \nu = 1, \ldots, N.$$

$D_\mu{}^\nu(\alpha)$ is the $(\mu\nu)^{\text{th}}$ element of the invertible $N \times N$ matrix $D(\alpha)$.[3] This gives:

- The *general linear* groups $GL(N, \mathbb{R})$ or $GL(N, \mathbb{C})$ (for real and complex vector spaces, respectively), and the
- The *special linear* groups $SL(N, \mathbb{R})$ or $SL(N, \mathbb{C})$ if $\det D = 1$.

- If $D \in GL(N, \mathbb{R})$ and $D^T D = D D^T = 1$, we get the *orthogonal groups*, $O(N)$.
- If $D \in SL(N, \mathbb{R})$ and $D^T D = D D^T = 1$, we get the *special orthogonal groups*, $SO(N)$.
- If $D \in GL(N, \mathbb{C})$ and $D^\dagger D = D D^\dagger = 1$, we get the *unitary groups*, $U(N)$.
- If $D \in SL(N, \mathbb{C})$ and $D^\dagger D = D D^\dagger = 1$, we get the *special unitary groups*, $SU(N)$.
- We also find the "pseudo" orthogonal groups (e.g., $SO(3, 1)$, the Lorentz group: the metric on the space $\{x_\mu\}$ is indefinite, as it has 3 plus signs and 1 minus sign).
- The "pseudo" unitary groups. These preserve an indefinite sesquilinear metric (involving complex conjugation). [We will not mention them again.]
- The symplectic groups $Sp(2N)$. [See tutorial question (1.5).]
- The exceptional groups $G(2)$, $F(4)$, $E(6)$, $E(7)$, and $E(8)$.

The *orthogonal* and *unitary* groups leave the metric $\delta_{\mu\nu}$ invariant [tutorial]:[4]

$$\delta_{\mu\nu} \overset{G}{\mapsto} D_\mu{}^\alpha D_\nu{}^\beta \delta_{\alpha\beta} = \delta_{\mu\nu}.$$

Similarly the special groups leave the totally antisymmetric Levi-Civita tensor $\epsilon_{\mu\nu\cdots\alpha}$ invariant,

$$\epsilon_{\mu\nu\cdots\alpha} \overset{G}{\mapsto} D_\mu{}^{\mu'} D_\nu{}^{\nu'} \cdots D_\alpha{}^{\alpha'} \epsilon_{\mu'\nu'\ldots\alpha'} = \det D\, \epsilon_{\mu\nu\ldots\alpha} = \epsilon_{\mu\nu\ldots\alpha} \quad (\text{since} \det D = 1)$$

2.3 One Dimensional Groups

Example. Translations on the real line, $x \mapsto x + a$.

Clearly if $g(a)g(b) = g(a + b)$ then the group is Abelian: $g(a)g(b) = g(b)g(a)$, *i.e.* all elements of the group commute with each other. There is a simple unitary representation $g(a) = e^{ia\lambda}$ with λ a fixed parameter (because $g(a)g(b) = e^{ia\lambda} e^{ib\lambda} = g(a + b)$).

[3] Think of x_μ as the μ^{th} component of a N-component column vector x.

[4] We shall see later that $g_{\mu\nu} = \delta_{\mu\nu}$ for compact groups.

Theorem 1 *All one dimensional Lie groups are such that the parameter a can be chosen so that $g(a)g(b) = g(a+b)$ for all a and b.*

Proof. The proof, originally due to Aczel, is sketched below.

Using associativity and analyticity of the composition law, we may write:

$$\begin{array}{rrcll} & g(x)\,(g(y)\,g(z)) & = & (g(x)\,g(y))\,g(z) & \text{(associativity)} \\ \Rightarrow & g(x)\,g(\phi(y,z)) & = & g(\phi(x,y))\,g(z) & \text{(analyticity)} \\ \Rightarrow & \phi(x,\phi(y,z)) & = & \phi(\phi(x,y),z) & \text{(analyticity again)} \end{array}$$

Take the partial derivative of the last expression with respect to z

$$\left.\frac{\partial\phi(x,y)}{\partial y}\right|_{y=\phi(y,z)} \frac{\partial\phi(y,z)}{\partial z} = \frac{\partial\phi(\phi(x,y),z)}{\partial z}$$

where we used the chain rule on the left hand side (LHS).

Set $z = 0$ and choose $g(0) = 1$, then $\phi(y,0) = y$, which gives

$$\frac{\partial\phi(x,y)}{\partial y}\,\psi(y) = \psi(\phi(x,y)) \qquad \text{where we defined} \qquad \psi(y) \equiv \left.\frac{\partial\phi(y,z)}{\partial z}\right|_{z=0}$$

We can solve this differential equation by rewriting it as

$$\frac{1}{\psi(\phi(x,y))}\,\frac{\partial\phi(x,y)}{\partial y} = \frac{1}{\psi(y)}\,, \tag{1}$$

and integrating with respect to y, to obtain

$$\rho(\phi(x,y)) = \rho(y) + c(x) \tag{2}$$

where

$$\rho(x) = \int_0^x \frac{\mathrm{d}t}{\psi(t)}\,, \qquad \rho(0) = 0$$

and $c(x)$ is an arbitrary function of x, an integration 'constant'.

To verify that (2) is indeed the correct solution, differentiate it respect to y, which recovers equation (1) directly.

We can determine $c(x)$ by setting $y = 0$, which gives $\rho(x) = 0 + c(x)$, and hence

$$\rho(\phi(x,y)) = \rho(x) + \rho(y)$$

Finally, if we change the parameterisation of the group element from $g(x)$ to $\bar{g}(\rho(x))$, we have

$$\bar{g}(\rho(x))\,\bar{g}(\rho(y)) = g(x)\,g(y) = g(\phi(x,y)) = \bar{g}(\rho(\phi(x,y))) = \bar{g}(\rho(x)+\rho(y))$$

Thus the group composition law is additive when written in terms of the new (barred) parameters. ■

Corollary 1 *All one dimensional Lie groups are Abelian.*

Example. The rotation group $SO(2)$

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Let $z = x + iy$, then $z \mapsto e^{i\phi}z$. But $e^{i\phi} \in U(1)$. So $U(1)$ and $SO(2)$ are isomorphic.[5]

These groups are distinct from the translation group. The parameter $\phi \in [0, 2\pi)$ so $U(1)$ and $SO(2)$ are compact. But the translation parameter $a \in \mathbb{R}$, *i.e.* $a \in (-\infty, \infty)$ so the translation group is *not* compact.

Theorem 2 *All compact Abelian Lie groups are isomorphic to $U(1) \otimes U(1) \otimes \cdots \otimes U(1)$.*

It follows that in this course most of the groups will be non-Abelian: $g_1 g_2 \neq g_2 g_1$.

Corollary 2 *The (finite) cyclic groups C_n (or $\mathbb{Z}_n$) $= \{e^{2m\pi i/n}, m = 0, \ldots, n-1\}$ are subgroups of $U(1)$.*

2.4 Representations

Any set of matrices ${D_\mu}^\nu(g)$ satisfying the group composition law,

$${D_\mu}^\nu(g_1)\,{D_\nu}^\rho(g_2) = {D_\mu}^\rho(g_1\,g_2)$$

constitutes a *representation* of G.[6]

Representations may be *reducible* (a direct sum of irreducible representations) or *irreducible*, just as for finite groups.

Note that if D is a representation, so is its complex conjugate D^*.

A representation for which $D = D^*$ is a *real* representation.

Any quantity which transforms under G may be classified according to the representation under which it transforms:

- Scalars: $s \mapsto s$ (i.e., it transforms trivially).
- Vectors: $v_\mu \mapsto {D_\mu}^\nu(g)\,v_\nu$ where $D(g)$ is the *defining* or *fundamental* representation.
- Tensors: $t_{\underbrace{\mu\nu\ldots}_{\alpha}} \mapsto {D_\mu}^{\mu'}{D_\nu}^{\nu'}\cdots t_{\underbrace{\mu'\nu'\ldots}_{\beta}}$ or, more succinctly, $t_\alpha \mapsto {D_\alpha}^\beta(g)\,t_\beta$,

  where ${D_\alpha}^\beta$ is some larger (possibly reducible) representation of G.

Invariant Tensors: $t_{\mu\nu\ldots} \mapsto t_{\mu\nu\ldots}$ under G. The components of an *invariant tensor* are *unchanged* by the action of the group.

---

[5] A group isomorphism is a function between two groups that sets up a *one-to-one correspondence between the elements of the groups* in a way that respects the given group operations. If there exists an isomorphism between two groups, then the groups are called isomorphic. From the standpoint of group theory, isomorphic groups have the same properties and need not be distinguished. (Source: Wikipedia!)

[6] We are using the summation convention, pairs of repeated indices are summed over implicitly.

Theorem 3 *If G is compact, then for every representation there is an equivalent unitary representation (just as for finite groups: Maschke's Theorem).*

Proof. We limit ourselves to a sketch of the proof.

Starting from the usual inner product (x, y) on a complex (or real) vector space, we define a new inner product[7]

$$\langle x|y\rangle \equiv \sum_{g\in G} (D(g)\,x, D(g)\,y)$$

where $D(g)\,x$ is a representation $D(g)$ of the group G acting on x, and the 'summation' (integration for continuous groups) is over all elements of G. Let $h \in G$ be an arbitrary group element. Using the rearrangement theorem[8] we can write the new inner product as

$$\begin{aligned}\langle x|y\rangle = \sum_{g\in G}(D(gh)\,x, D(gh)\,y) &= \sum_{g\in G}(D(g)\,(D(h)\,x)\ ,\ D(g)\,(D(h)\,y)) \\ &= \langle D(h)\,x|D(h)\,y\rangle = \langle x|\,D^\dagger(h)D(h)\,|y\rangle\end{aligned}$$

so $D^\dagger D = 1$ with this inner product.

The proof breaks down if G is not compact because the sum/integral over all the group elements is not well defined.

■

The action of a group on a Hilbert space must be through a unitary representation, to conserve probability: $1 = \psi_i^*\psi^i = \psi_i^*\psi_\ell\delta^{i\ell} = (D_i{}^j\psi_j)^* D_\ell{}^k\psi_k\delta^{i\ell} = \psi_j^*\psi_k(D_i{}^{j*}D_\ell{}^k\delta^{i\ell}) = \psi_j^*\psi_k(D^\dagger D)^{jk} \Rightarrow D_i{}^{j*}D^i{}_k = \delta_k^j$, i.e., $D^\dagger D = 1$, (where as usual we have denoted the unit matrix by 1).

3 Lie Algebras

3.1 Generators

Consider a Lie group $G = \{g(\alpha)\}$, and choose α_a, where $a \in [1, \mathrm{Dim}\,G]$, such that $g(0) = 1$. When the α_a are *small*, $g(\alpha)$ is close to 1 (the unit/identity element), and we can use Taylor's theorem to write[9]

$$g(\alpha) = 1 + i\alpha^a T_a + O(\alpha^2) \tag{3}$$

(this uses analyticity), where

$$T_a \equiv \frac{1}{i}\frac{\partial g(\alpha)}{\partial\alpha^a}\bigg|_{\alpha=0}. \tag{4}$$

The matrices $\{T_a\}$ are the *generators* of the group. The factor of i is a useful convention: if $D(g)$ is a *unitary* representation, the corresponding representation of the generators is *hermitian* when α_a are real:

$$1 = D^\dagger D = 1 - i\alpha^a(T_a^\dagger - T_a) + O(\alpha^2) \Rightarrow T_a^\dagger = T_a\,.$$

The number of generators $= \dim G$, *i.e.* the same as the number of parameters.

Example. $SU(N)$: $g^\dagger g = gg^\dagger = 1$, $\det g = 1$, so $T_a^\dagger = T_a$, $\mathrm{tr}\,T_a = 0$ (the latter follows from the constraint $\det g = 1$ and $\det A = \exp\mathrm{tr}\ln A$); the generators are *hermitian* and *traceless*.

For $SU(2)$, we generally *choose* the *Pauli* matrices:

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

For $SU(3)$, the generators are the *Gell-Mann matrices* $\{\lambda_a : a = 1, \ldots, 8\}$ [see later].

Since $U(N)$ doesn't have the constraint $\det D = 1$, the generators of $U(N)$ are $\{1, T_a\}$ where T_a are the $SU(N)$ generators.

Example. $SO(N)$: $g^Tg = gg^T = 1$, $\det g = 1$. Since $SO(N)$ is real, we generally omit the factors of i in the expansion of $g(\alpha)$ in equations (3) and (4), so that

$$g(\alpha) = 1 + \alpha^a T_a + O(\alpha^2) \qquad \text{and} \qquad T_a \equiv \frac{\partial g(\alpha)}{\partial\alpha^a}\bigg|_{\alpha=0}.$$

whereupon $T_a^T = -T_a$, $\mathrm{tr}\,T_a = 0$, i.e., the generators are real antisymmetric. For $SO(2)$

$$T_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0\end{pmatrix}.$$

For $SO(3)$, we may choose

$$T_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix} \qquad T_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0\end{pmatrix} \qquad T_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0\end{pmatrix}.$$

The additional condition $\mathrm{tr}\,T_a = 0$ is *automatic* as the generators are *antisymmetric*, so $O(N)$ and $SO(N)$ share the *same* generators.[10]

3.2 Commutators

Consider $f = ghg^{-1}h^{-1}$ with $f, g, h \in G$. If either $g = 1$ or $h = 1$ then $f = 1$. Take g, h infinitesimal and expand: [11]

$$\begin{aligned} g &= 1 + i\underline{\alpha}\cdot\underline{T} + \tfrac{1}{2}(i\underline{\alpha}\cdot\underline{T})^2 + O(\alpha^3) \\ h &= 1 + i\underline{\beta}\cdot\underline{T} + \tfrac{1}{2}(i\underline{\beta}\cdot\underline{T})^2 + O(\beta^3) \\ \text{then}\quad f &= 1 + i\underline{\gamma}\cdot\underline{T} + \tfrac{1}{2}(i\underline{\gamma}\cdot\underline{T})^2 + O(\gamma^3)\end{aligned}$$

where $\gamma = O(\alpha\beta)$ and $\underline{\alpha}\cdot\underline{T} \equiv \alpha^a T_a$, etc.

[7] $\langle x|y\rangle$ is not the usual inner product in Dirac notation!

[8] Roughly speaking, the rearrangement theorem states that if we multiply all the elements g_i in the group by a fixed element h, we recover all the elements g_i once (and once only), but in a different order, i.e., a "rearrangement".

[9] NB We're using the summation convention here. Take care to contract only upper indices with lower indices.

[10] If we choose *hermitian* generators for $SO(N)$ and $O(N)$, their elements become purely imaginary. Similarly, the generators of $SU(N)$ become *anti*-hermitian $T_a^\dagger = -T_a$ if we omit the factors of i in equations (3) and (4) (exercise). It can be argued that antihermitian generators are more natural as they avoid all factors of i in all these definitions!

[11] Strictly, we take the parameters $|\alpha^a|, |\beta^a| \ll 1$.

How did we get this result?

(a) We've already discussed the linear terms in the expansion of g (and h).

The quadratic terms follow from Taylor-expanding g to second order in α, and using $g(-\alpha) = g(\alpha)^{-1}$ (or $g(\alpha)^2 = g(2\alpha)$) to express the terms with second derivatives in the form shown. This is a useful exercise – do it!

(b) In fact, by substitution (exercise)

$$[\underline{\alpha}\cdot\underline{T}, \underline{\beta}\cdot\underline{T}] = -\frac{i}{2}\underline{\gamma}\cdot\underline{T} + O(\alpha^3, \beta^3)$$

All terms of $O(\alpha)$, $O(\beta)$, $O(\alpha^2)$, and $O(\beta^2)$ cancel (exercise). Since this is true for arbitrary α^a, β^b, we must have a relation $\gamma^c = -2c^c{}_{ab}\alpha^a\beta^b$, for some set of numbers $c^c{}_{ab}$. The factor of -2 is for notational convenience. Hence

$$[T_a, T_b] = iT_c\, c^c{}_{ab} \tag{5}$$

where there is an implicit sum over the index $c \in [1, \mathrm{Dim}\, G]$.

The quantities $c^c{}_{ab}$ are called the *structure constants* of the group: they define the *Lie algebra* L_G of the Lie group G.

(c) We use lower case letters to denote the Lie algebras of $SU(N)$, $SO(N)$, *etc*, by $su(N)$, $so(N)$, *etc.*

(c) If we omit the factors of i in equations (3) and (4), there is no factor of i in equation (5).

3.3 Exponentiation

Consider a one dimensional subgroup $G_1 \subset G$, $\{g(t)\} = \{g(\alpha(t))\}$ labelled by a single parameter t. We can choose $t \in \mathbb{R}$ such that

$$g(s)\, g(t) = g(s+t); \quad g(0) = 1 \qquad \text{(by Theorem 1)}$$

Differentiate the equation above with respect to s:

$$g'(s)\, g(t) = g'(s+t)$$

and set $s = 0$, which gives a differential equation for $g(t)$, which we can solve easily:

$$\frac{d}{dt}g(t) = g'(0)\, g(t) \quad \Rightarrow \quad g(t) = \exp\left(tg'(0)\right).$$

But, we can also write

$$g'(0) = \left.\frac{\partial g}{\partial \alpha^a}\right|_{\alpha=0}\left.\frac{d\alpha^a}{dt}\right|_{t=0} = i\gamma^a\, T_a \qquad \text{where} \quad \gamma^a \equiv \left.\frac{d\alpha^a}{dt}\right|_{t=0},$$

so if $g \in G_1 \subset G$, then

$$g(t) = \exp(i(t\gamma^a)T_a),$$

therefore every element of G that can be connected to the identity has this form. But the set $\{e^{i\beta^a T_a}\}$ has dimension equal to $\dim G$ and coincides with G for β^a infinitesimal. So if we know the generators we know all parts of the group connected to the identity.

Furthermore

$$gh = \exp\{i\underline{\alpha}\cdot\underline{T}\}\,\exp\{i\underline{\beta}\cdot\underline{T}\} = \exp\{i\underline{\gamma}\cdot\underline{T}\}$$

for some $\underline{\gamma}$, because $gh \in G$. But the Baker–Campbell–Hausdorff formula tells us that for any two matrices A and B,

$$\exp(A)\,\exp(B) = \exp(A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{3!}\tfrac{1}{2}([A,[A,B]] + [[A,B],B]) + \cdots)$$

so if we know L_G we can evaluate all the commutators and thus $\underline{\gamma}(\underline{\alpha}, \underline{\beta})$ (at least in principle).

To summarise:

- Lie group $\Rightarrow$ Lie algebra by *linearisation.*
- Lie algebra $\Rightarrow$ *connected part* of a Lie group by *exponentiation.*

Note that $G \Rightarrow L_G$ uniquely, but not vice versa: several groups may have the same Lie algebra. These groups are identical locally (in the neighbourhood of 1), but not globally. Fortunately, for a given L_G there is only one connected group G which is simply connected (no holes): this is called the *universal covering group* (not proven here).

Example. $O(N), SO(N)$ have the same generators, so $o(N) \cong so(N)$ – the Lie algebras are isomorphic.

However, for $O(N)$, since $DD^T = 1$, $(\det D)^2 = 1$, which implies $\det D = \pm 1$, while for $SO(N)$ we have $\det D = +1$ (by definition). We can't get from $\det D = -1$ to $\det = +1$ by a continuous transformation, so $O(N)$ is not connected, whereas $SO(N)$ is connected (but not simply connected – see later). We write $SO(N) \cong O(N)/\mathbb{Z}_2$. For rotations, $N = 2, 3$, think of $\mathbb{Z}_2$ as the *parity.*

For example, in the case of $SO(2)$, choosing the generator to be real antisymmetric,

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad T^2 = -1, \quad \text{and hence} \quad e^{\phi T} = 1\cos\phi + T\sin\phi \quad \text{(exercise).}$$

The matrices $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ are in $O(2)$, but they're not in $SO(2)$ because they're not continuously connected to the identity.

- The simply connected covering group of $SO(N)$ is called $\mathrm{Spin}(N)$.
- The simply connected covering group of $SO(3)$ is $SU(2)$. This is a special case (see tutorial question (1.4)).

3.4 Properties of Structure Constants

Recall the Lie algebra L_G of a Lie group G:

$$[T_a, T_b] = iT_c\, c^c{}_{ab}$$

1. Clearly $c^c{}_{ab} = -c^c{}_{ba}$ by anti-symmetry of the commutator: $[T_a, T_b] \equiv T_aT_b - T_bT_a = -[T_b, T_a]$.

2. Under rescaling $T_a \mapsto \lambda T_a$, $c^c{}_{ab} \mapsto \lambda c^c{}_{ab}$.

3. The $c^c{}_{ab}$ are real.

   Proof. We know already that if $D(g)$ is a representation of G, then $D^*(g)$ is also a representation. So if T_a are a representation of L_G, $-T_a^*$ are also a representation. But $[-T_a^*, -T_b^*] = i(-T_c^*)c^c{}_{ab}{}^*$, so $c^c{}_{ab}{}^* = c^c{}_{ab}$. ■

4. The *Jacobi identity*:

   $$[T_a, [T_b, T_c]] + [T_b, [T_c, T_a]] + [T_c, [T_a, T_b]] \;=\; 0\,.$$

   Note that the Jacobi identity holds for *any* three square matrices T_a, T_b, T_c (exercise).

   Substituting from the Lie algebra, we find (exercise)

   $$c^d{}_{ae}\, c^e{}_{bc} \,+\, c^d{}_{be}\, c^e{}_{ca} \,+\, c^d{}_{ce}\, c^e{}_{ab} \;=\; 0\,.$$

5. The structure constants form a representation of L_G, known as the *adjoint representation*. If we define Dim-G (Dim-G)×(Dim-G) matrices A_a with elements

   $$(A_a)^b{}_c \equiv ic^b{}_{ac}$$

   then the Jacobi identity can be written

   $$(A_a)^d{}_e\, (A_b)^e{}_c - (A_b)^d{}_e\, (A_a)^e{}_c = i(A_e)^d{}_c\, c^e{}_{ab},$$

   where we used the anti-symmetry of the structure constants, i.e. $c^e{}_{ca} = -c^e{}_{ac}$ in the second term on the left-hand side (LHS), and similarly for the first term on the right-hand side (RHS). In matrix notation

   $$[A_a, A_b] = iA_e\, c^e{}_{ab}.$$

3.5 The Adjoint Representation

If a vector v transforms under $G: v \mapsto gv$, then $\forall h \in G$, $(hv) \mapsto g(hv) = (ghg^{-1})(gv)$. So if $v \mapsto gv$, $h \mapsto ghg^{-1}$ gives the action of the group on itself.

Now consider $h' = ghg^{-1}$ with h infinitesimal,[12] then h' is also infinitesimal:

$$\begin{aligned} h &= 1 + i\alpha^a T_a + \cdots \\ h' &= 1 + i\beta^a T_a + \cdots \qquad \text{where } \beta^a = D^a{}_b(g)\,\alpha^b \quad \text{for some } D^a{}_b(g) \text{ (by linearity)} \\ &= 1 + iT_a\, D^a{}_b(g)\,\alpha^b + \cdots \;=\; 1 + iT_b\, D^b{}_a(g)\,\alpha^a + \cdots \end{aligned}$$

[12] *i.e.* h differs from the identity element by the infinitesimal quantity, which in this case is $i\alpha^a T_a$

where we simply relabelled indices ($a \leftrightarrow b$) to get the last expression. It follows from the first and third lines above that

$$gT_ag^{-1} = T_b\, D^b{}_a(g).$$

Pictorially:

$$\begin{array}{ccc} v & \xrightarrow{g} & gv \\ \downarrow_h & \overset{\mathrm{ad}_g}{\Longrightarrow} & \downarrow_{ghg^{-1}} \\ hv & \xrightarrow{g} & ghv \end{array}$$

The matrices $D^b{}_a(g)$ form a representation (the *adjoint representation*) of the group G:

$$D^b{}_a(1) = \delta^b{}_a, \quad T_b\, D^b{}_c(g)\, D^c{}_d(h) = (g\, T_c\, g^{-1})\, D^c{}_d(h) = h\, (g\, T_d\, g^{-1})\, h^{-1} = h\, g\, T_d\, (hg)^{-1} = T_b\, D^b{}_d(hg)$$

so $D(g)D(h) = D(hg)$ as required. (Note that with the definitions used here, D acts to the left if g acts to the right.)

The generators in this representation are just the structure constants described above. To verify this, set $g = 1 + i\alpha^b T_b$, so $g^{-1} = 1 - i\alpha^b T_b$, and write $D^c{}_a = \delta^c{}_a + i\alpha^b (A_b)^c{}_a$. Then

$$\begin{aligned} gT_ag^{-1} &= T_a + i\alpha^b\,(T_b\,T_a - T_a\,T_b) + O(\alpha^2) \\ &= T_a + i\alpha^b\, T_c\, ic^c{}_{ba} + O(\alpha^2) \\ &= T_c\left(\delta^c{}_a + i\alpha^b\, ic^c{}_{ba}\right) + O(\alpha^2)\,. \end{aligned}$$

So $(A_b)^c{}_a = ic^c{}_{ba}$ as claimed.

It follows that the adjoint representation is a 'real' representation. (Recall that the structure constants are real, so the elements of A_b are i times a real number, which we call a 'real' representation). Note that the elements of A_b would actually be real if we used *anti*-hermitian generators!

Example. $SU(2)$ and $SO(3)$.

Take the generators of $SU(2)$ as $\frac{1}{2}\sigma_i$. Then since $\sigma_1\sigma_2 = i\sigma_3$ *etc.*, $[\frac{1}{2}\sigma_i, \frac{1}{2}\sigma_j] = i\epsilon_{ijk}\frac{1}{2}\sigma_k$, so the structure constants are just ϵ_{ijk}. Consider $g = 1 + \frac{i}{2}\underline{\omega}\cdot\underline{\sigma} + O(\omega^2)$. Then

$$\begin{aligned} g\sigma_kg^{-1} &= (1 + \tfrac{i}{2}\underline{\omega}\cdot\underline{\sigma})\sigma_k(1 - \tfrac{i}{2}\underline{\omega}\cdot\underline{\sigma}) \\ &= \sigma_k + \tfrac{i}{2}\omega_i[\sigma_i, \sigma_k] \\ &= \sigma_k + \omega_i\epsilon_{kim}\sigma_m, \\ &= \sigma_k + \left(\underline{\omega}\times\underline{\sigma}\right)_k, \\ \text{i.e.} \qquad g\,\underline{\sigma}\,g^{-1} &= \underline{\sigma} + \underline{\omega}\times\underline{\sigma} + O(\omega^2) \end{aligned}$$

This is clearly an infinitesimal rotation $\in SO(3)$, i.e. an infinitesimal rotation through angle $\omega \equiv |\underline{\omega}|$ about the axis $\underline{\omega}$ in $\mathbb{R}^3$.

It follows that the *adjoint* representation of $su(2)$ is the *defining* representation of $so(3)$, thus that their Lie Algebras are the same: $su(2) \cong so(3)$. In fact $SO(3) \cong SU(2)/\mathbb{Z}_2$: $SU(2)$ is the covering group of $SO(3)$ (see tutorial question (1.4)).

3.6 Invariant Tensors

An *invariant tensor* $d_{a_1 a_2 \ldots a_n}$ is any tensor whose components are invariant under the action of the group, i.e.,

$$d'_{a_1\ldots a_n} \;=\; d_{a'_1\ldots a'_n}\, D^{a'_1}{}_{a_1}(g)\, D^{a'_2}{}_{a_2}(g)\cdots D^{a'_n}{}_{a_n}(g) \;\equiv\; d_{a_1\ldots a_n}$$

The simplest way to construct invariant tensors is to consider traces of products of generators

$$\begin{aligned} d_{a_1\ldots a_n} &\equiv \mathrm{tr}(T_{a_1}T_{a_2}\cdots T_{a_n}) \\ &\overset{G}{\mapsto} d'_{a_1\ldots a_n} \\ &= \mathrm{tr}(T_{a'_1}T_{a'_2}\cdots T_{a'_n})\, D^{a'_1}{}_{a_1}\cdots D^{a'_n}{}_{a_n} \;=\; \mathrm{tr}\left((gT_{a_1}g^{-1})\,(gT_{a_2}g^{-1})\cdots(gT_{a_n}g^{-1})\right) \\ &= \mathrm{tr}\left(T_{a_1}(g^{-1}g)\,T_{a_2}(g^{-1}g))\cdots T_{a_n}(g^{-1}g)\right) \;=\; d_{a_1\cdots a_n} \end{aligned}$$

i.e., $d_{a_1\cdots a_n}$ is an invariant tensor for any representation T_a, including the adjoint representation A_a.

Example. $c^d{}_{bc}$ is invariant (use the Jacobi identity).

3.7 Killing Form

An important invariant tensor is the *Killing form.* This is a *metric* on the group space defined by

$$g_{ab} \equiv \mathrm{tr}\,(A_a\, A_b) \;=\; (A_a)^c{}_d\,(A_b)^d{}_c \;=\; -c^c{}_{ad}\,c^d{}_{bc} \;=\; g_{ba} \qquad \text{(cyclic property of the trace)}$$

where the A_a are the generators in the adjoint representation.

Since the structure constants are real, g_{ab} are the elements of a *real symmetric* $\mathrm{Dim}\,G\times\mathrm{Dim}\,G$ matrix - this is the advantage of using the adjoint representation. Thus we may choose a basis in which it is diagonal

$$g_{ab} = \mu_a\,\delta_{ab} \qquad \text{(no sum on } a.)$$

Since g_{ab} is real symmetric, the eigenvalues μ_a of are real, but may be positive, negative, or zero. Indeed, for Abelian groups $g_{ab} = 0$ because the structure constants are all zero. (Convince yourself of this.)

Under the rescaling $T_a \mapsto \lambda T_a$, with $\lambda \in \mathbb{R}$, then $g_{ab} \mapsto \lambda^2 g_{ab}$ (with $\lambda^2 > 0$). By rescaling each T_a separately we may reduce g_{ab} to the canonical form we introduced earlier

$$\mathrm{diag}(+1,\ldots,+1,0,\ldots,0,-1,\ldots,-1), \qquad \text{i.e., } \mu_a = 0,\ \pm1,\ \forall a.$$

Example. $SU(2)$:

$$g_{ij} = -\epsilon_{ikl}\,\epsilon_{jlk} = 2\delta_{ij}.$$

so there's no difference between upstairs and downstairs indices in the case of $SU(2)$. If the generators were rescaled to $\frac{-1}{\sqrt{2}}\epsilon_{ijk}$ we would have $g_{ij} = \delta_{ij}$,

---

We can use the Killing form to lower indices: for example, we may define structure constants with all lower indices:

$$c_{abc} \equiv g_{ad}\,c^d{}_{bc}.$$

This is also an invariant tensor because we can use the Lie algebra to write it in terms of the trace of products of generators:

$$\mathrm{tr}(A_a[A_b, A_c]) \;=\; i\,\mathrm{tr}(A_a A_d)\,c^d{}_{bc} \;=\; ig_{ad}\,c^d{}_{bc} \;=\; ic_{abc}.$$

Since[13]

$$\mathrm{tr}(A_a[A_b, A_c]) \;=\; \mathrm{tr}([A_a, A_b]A_c) \;=\; \mathrm{tr}([A_c, A_a]A_b)\,.$$

it follows that c_{abc} is *totally antisymmetric on all three indices.*

A second invariant tensor with three indices is $d_{abc} = \mathrm{tr}(A_a\{A_b, A_c\})$ which is totally symmetric.[14]

We need consider only *one* tensor $d_{a_1\cdots a_n}$ for each $n > 3$, the one which is *totally symmetric* because antisymmetric parts can always be expressed in terms of c_{abc} and tensors of lower degree, by using the commutation relations.

Definition 5 *The number of independent invariant tensors (excluding c_{abc}) is called the* rank *of the group.*

4 Simplicity

We may consider a Lie algebra $L = \{T_a : a = 1,\ldots,l\}$ as a vector space spanned by the generators. Then

Definition 6 *The subspace $S \subset L$ is a* subalgebra *of L if the commutator $[s, s'] \in S$ for all $s, s' \in S$.*

In words, the commutator of any two elements in the subspace also lies in the subspace.

S will be spanned by a new set of generators $\{S_i : i = 1,\ldots,l'\}$. S is a *proper* subalgebra provided $0 < l' < l$.

Definition 7 *The subalgebra $S \subset L$ is an* invariant subalgebra *(or* ideal*) if $[s,t] \in S$ for all $s \in S, t \in L$, i.e., if S "absorbs" L or "S knocks L into S".*

This is analogous to the definition of a normal subgroup: $ghg^{-1} \in H, \quad \forall h \in H, g \in G, H \subset G$.

Definition 8 *The Lie algebra L is* simple *if it has no proper invariant subalgebras, whereas L is* semi-simple *if it has no proper Abelian invariant subalgebras.*

Simple and semi-simple Lie groups are those with simple or semi-simple algebras. Clearly all simple groups/algebras are also semi-simple.

[13] $\mathrm{tr}\,X[Y,Z] = \mathrm{tr}(XYZ{-}XZY) = \mathrm{tr}\,XYZ{-}\mathrm{tr}\,XZY = \mathrm{tr}\,XYZ{-}\mathrm{tr}\,YXZ = \mathrm{tr}(XYZ{-}YXZ) = \mathrm{tr}[X,Y]Z$, etc.

[14] The *anti-commutator* $\{A,B\} \equiv AB + BA$.

4.0.1 Non semi-simple algebras

Consider an algebra which is *not* semi-simple. Then it has a proper Abelian invariant subalgebra $S \subset L$.

Let's take a basis $\{S_i\}$ for S, and extend it to a basis $\{S_i, P_p\} = \{T_a\}$ for L; where the indices $a, b, c \in T$, $i, j, k \in S$, and $p, q, r \in P$, i.e. the set $\{P_p\}$ contains all the generators that are in L, but not in S. Then

1. Since S is Abelian: $[S_i, S_j] = iT_a\, c^a{}_{ij} = 0$ (by definition), therefore $c^a{}_{ij} = 0$;
2. Since S is invariant: $[S_i, T_a] = iT_b\, c^b{}_{ia} \subset S$ implies $c^p{}_{ia} = 0$ and, by anti-symmetry of the commutator, $c^p{}_{ai} = 0$.

Then, using 2. (twice) and 1. (once),

$$g_{ia} = \mathrm{tr}(A_i A_a) = (A_i)^c{}_d\,(A_a)^d{}_c = -c^c{}_{id}\,c^d{}_{ac} = -c^k{}_{id}\,c^d{}_{ak} = -c^k{}_{ij}\,c^j{}_{ak} = 0,$$

so g_{ab} has the form $\left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & g_{pq} \end{array}\right)$. It follows that g_{ab} has zero eigenvalues, and $\det g_{ab} = 0$.

4.1 Cartan's Criterion

A Lie algebra is semi-simple *iff*[15] the Killing form is invertible, i.e., $\det g_{ab} \neq 0$.

We denote the inverse by g^{ab}, so that $g^{ac}g_{cb} = \delta^a{}_b$. g_{ab} is now a pseudo-Riemannian metric.

For semi-simple groups we may choose a basis in which g_{ab} and g^{ab} are diagonal and both take the form $\mathrm{diag}(1, \dots, 1, -1 \dots, -1)$, i.e., all eigenvalues are ± 1.

4.2 Casimir Operators

When L is semi-simple, we may define the quadratic Casimir operator

$$C^{(2)} \equiv g^{ab}\, T_a\, T_b$$

and similarly the higher Casimir operators

$$C^{(n)} \equiv d^{a_1 \cdots a_n} T_{a_1} \cdots T_{a_n} \qquad \text{where} \quad d^{a_1 \cdots a_n} = g^{a_1 a_1'} \cdots g^{a_n a_n'}\, d_{a_1' \cdots a_n'}\,,\ n \geq 3.$$

There are as many independent Casimir operators as there are independent symmetric invariant tensors, i.e., the rank of the algebra.

The Casimir operators are useful because they commute with the generators: $[C^{(n)}, T_a] = 0$. To show this, note that

$$\begin{aligned} gC^{(n)}g^{-1} &= d^{a_1 \cdots a_n}\,(gT_{a_1}g^{-1}) \cdots (gT_{a_n}g^{-1}) \\ &= d^{a_1 \cdots a_n}\, T_{a_1'}\, D^{a_1'}{}_{a_1} \cdots T_{a_n'}\, D^{a_n'}{}_{a_n} \\ &= d^{a_1 \cdots a_n} T_{a_1} \cdots T_{a_n} = C^{(n)} \end{aligned}$$

[15] *iff* means 'if and only if'.

where the first expression in the last line follows because $d_{a_1 \cdots a_n}$ is an invariant tensor. Taking $g = 1 + i\alpha^a T_a$ with infinitesimal α^a gives the result:

$$(1 + i\alpha^a T_a)\, C^{(n)}\, (1 - i\alpha^a T_a) = C^{(n)} \qquad \Rightarrow \qquad C^{(n)} = C^{(n)} + i\alpha^a \left[T_a, C^{(n)}\right]$$

Since Casimir operators are constructed from products of generators, it follows that they also commute with each other: in fact they constitute a maximally commuting set. They are thus useful for labelling representations. This is analogous to using a complete set of mutually commuting operators to label states in quantum mechanics.

For semi-simple Lie algebras it is possible to show using Schur's lemma that

$$\tilde{g}_{ab} \equiv \mathrm{tr}\, T_a T_b = C(R)\, g_{ab}$$

for some real number $C(R)$ which depends only on the representation R of the generators T_a. Taking the trace of the expression for the definition of the quadratic Casimir operator above, and using this result gives

$$\mathrm{tr}\, C^{(2)} = g^{ab}\, \mathrm{tr}\,(T_a\, T_b) = g^{ab}\, C(R)\, g_{ab} = \delta^a{}_a\, C(R) = (\mathrm{Dim}\, G)\, C(R)$$

So $C(R) = \mathrm{tr}\, C^{(2)} / \dim G$.

For the adjoint (Ad) representation, $C(\mathrm{Ad}) = 1$ with our normalisation. One often sees instead the normalisation $C(\mathrm{Fun}) = 1$ or 2 or $\frac{1}{2}$ for the fundamental (Fun) representation.

4.3 Cartan Decomposition

Consider a semi-simple Lie algebra L that is not simple: it has a proper invariant subalgebra $S \subset L$.

Use the metric g_{ab} to define an othogonal subspace P, again using indices $i, j, k \in S$, and $p, q, r \in P$:

$$g_{ip} = \mathrm{tr}\,(A_i\, A_p) = 0 \qquad \text{for all } A_i \in S,\ A_p \in P.$$

Now $\mathrm{tr}\,(A_i\,[A_j, A_p]) = \mathrm{tr}\,([A_i, A_j]\, A_p) = 0$ because $[A_i, A_j] \in S$

Thus $[A_j, A_p] \in P$, *i.e.* $[S, P] \subset P$.

Since S is invariant, we already have $[S, P] \subset S$.

But $S \cap P = 0$, so $[S, P] = 0$ (all elements of S commute with all elements of P). We say that L decomposes into a direct sum $L = S \oplus P$.

It is easy to show that P is also an invariant subalgebra of L:

$$\mathrm{tr}\,([A_p, A_q]\, A_i) = \mathrm{tr}\,(A_p\,[A_q, A_i]) = 0 \qquad ([A_q, A_i] = 0 \text{ because } L = S \oplus P)$$

so $[P, P] \subset P$ as required.

It is now easy to prove the following theorem:

Theorem 4 *A semi-simple Lie algebra L is a direct sum of mutually commuting simple Lie algebras:*

$$L = L_1 \oplus L_2 \oplus \cdots \oplus L_n,$$

where the L_i are simple.

Proof. If L is not simple, use the above result to decompose it as

$$L = S \oplus P.$$

Both S and P are semi-simple (otherwise L would not be). If either is not simple, repeat the procedure. Since L is finite dimensional, this procedure obviously stops. ■

Example.

- $so(4) \cong su(2) \oplus su(2)$ (see tutorial question (3.2))
- $sl(2, \mathbb{C}) \cong su(2) \oplus su(2)$ (see later).

The decomposition of the algebra induces a similar decomposition of a semi-simple group

$$G \;\cong\; (G_1 \otimes G_2 \otimes \cdots \otimes G_n) \,/ Z\;,$$

where the G_i are simple, and Z is the *centre* of G, namely the set of all group elements g' that commute with themselves *and* with all other group elements: $Z = \{g' : gg' = g'g, g \in G\}$. Clearly Z is discrete, as otherwise G would not be semi-simple.

Example: $SO(4) \cong SU(2) \otimes SU(2) \,/\, \mathbb{Z}_2$.

5 Compact Groups I

If the Lie group G is compact, the generators of L_G may be taken to be hermitian (because there are unitary representations). This has far-reaching consequences.

Taking the adjoint representation hermitian, $A_a = A_a^\dagger$, so $(A_a)^b{}_c = (A_a^*)^c{}_b$, and hence

$$(A_a)^b{}_c \;=\; ic^b{}_{ac} \;=\; -ic^c{}_{ab}\,.$$

Therefore

$$g_{ab} \;=\; \mathrm{tr}\,(A_a A_b) \;=\; (A_a)^d{}_c\,(A_b)^c{}_d \;=\; (-i\,c^c{}_{ad})\,ic^c{}_{bd} \;=\; c^c{}_{ad}\,c^c{}_{bd}\,.$$

Theorem 5 *If G is compact L_G may be decomposed as*

$$L_G \;\cong\; L_0 \oplus (L_1 \oplus \cdots \oplus L_n)$$

where L_0 is Abelian (so $L_0 \cong u(1) \oplus \cdots \oplus u(1)$) and the L_i are simple.

Proof. If L_G is not semisimple, it has a proper Abelian invariant subalgebra $S \subset L$, so $g_{ia} = 0$ for $S_i \in S$, $T_a \in L$ (derived on page 15).

But we have just shown that for hermitian A_a, $g_{ia} = c^b{}_{ic}\,c^b{}_{ac}$, so $g_{ii} = (c^b{}_{ic})^2 = 0$ (not summed over i); therefore $c^b{}_{ic} = 0 \;\; \forall i, b, c$, whence $[S, L] = 0$. It follows that $L_G = L_0 \oplus L'$ where L_0 is Abelian and L' is semisimple, and thus can be decomposed into simple algebras as previously shown. ■

So we need consider only semisimple (or indeed simple) compact groups. NB, this decomposition does *not* apply to non-compact groups because it relies on the generators being hermitian.

Theorem 6 *If G is compact, the Killing form g_{ab} is positive definite (has non-negative eigenvalues). If L_G is also semisimple, we may choose a basis in which $g_{ab} = \delta_{ab}$, and distinctions between upper and lower indices can thus be ignored.*

Proof. For any real vector x^a

$$g_{ab}x^a x^b \;=\; c^c{}_{ad}\,c^c{}_{bd}\,x^a x^b \;=\; \left(x^a c^c{}_{ad}\right)^2 \;\geq\; 0$$

Therefore all eigenvalues of $g_{ab} \geq 0$. If L_G is semisimple, $\det g_{ab} \neq 0$, so $g_{ab}x^a x^b > 0$ and all eigenvalues > 0. Finally, rescale the generators so that $g_{ab} = \mathrm{diag}(1, \ldots, 1)$. ■

Theorem 7 *If G is compact and semisimple,* $\mathrm{tr}\, T_a T_b = C(R)\delta_{ab}$ *for all irreducible representations. The positive real number $C(R)$ is "the Casimir" of the representation R.*

Proof. Omitted (uses Schur's lemma). ■

5.1 The $su(2)$ algebra

This is the "angular momentum algebra" familiar from quantum mechanics. For $su(2) \cong so(3)$ $[T_i, T_j] = i\epsilon_{ijk}T_k$, $g_{ij} = 2\delta_{ij}$, so it is compact and semisimple. Define the "step operators" $T_\pm = T_1 \pm iT_2$, then

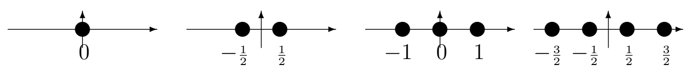
$$\begin{aligned} [T_3, T_\pm] &= \pm T_\pm \\ [T_+, T_-] &= 2T_3 \\ [T_+, T_+] &= [T_-, T_-] = 0. \end{aligned}$$

The quadratic Casimir is $T^2 = T_1^2 + T_2^2 + T_3^2 = T_-T_+ + T_3^2 + T_3$, and $[T^2, T_i] = 0$. There are no other Casimir operators (so $su(2)$ has rank 1): T^2 and T_3 form a complete commuting set.

- T^2 has eigenvalues $j(j+1)$, $2j \in \mathbb{N}$: use to label representations.
- T_3 has eigenvalues $m = j, j-1, \ldots, -j$, $2m \in \mathbb{Z}$: use to label states within representation.

If $T_3\,|jm\rangle \;=\; m\,|jm\rangle$, $T_3T_\pm\,|jm\rangle \;=\; T_\pm(T_3 \pm 1)\,|jm\rangle \;=\; (m \pm 1)T_\pm\,|jm\rangle$, while $T^2\,|jj\rangle \;=\; (T_-T_+ + T_3^2 + T_3)\,|jj\rangle \;=\; j(j+1)\,|jj\rangle$ because $T_+\,|jj\rangle \;=\; 0$. $|jj\rangle$ is the "state of greatest weight". Since we are interested in finite dimensional representations there is also a "state of least weight" $T_-\,|j\bar{m}\rangle = 0$, for which $T^2\,|j\bar{m}\rangle = (T_+T_- + T_3^2 - T_3)\,|j\bar{m}\rangle = \bar{m}(\bar{m}-1)\,|j\bar{m}\rangle$. Since T^2 is a multiple of the unit matrix we must have $\bar{m}(\bar{m}-1) = j(j+1)$, so $\bar{m} + j = \bar{m}^2 - j^2 = (\bar{m}+j)(\bar{m}-j)$, and $\bar{m} = -j$ since $\bar{m} \leq j$.

Representations:



1 2 (fundamental) 3 (adjoint) 4

Combining representations: "start with the state of largest m and work down": $2 \otimes 2 = 3 \oplus 1$, $3 \otimes 3 = 5 \oplus 3 \oplus 1$, etc. We will show later that all compact semisimple algebras contain $su(2)$ subalgebras which allow for similar constructions of their representations, and composition of representations.

6 Space-time Symmetries

In this section we shall study the symmetry groups of the space-time of special relativity, namely the Lorentz and Poincaré groups.

6.1 Kinematics

Let us choose a set of space-time coordinates $x_\mu = (x_0, x_i)$, with $x_0 = ct$ and $i = 1, 2, 3$.

We have a metric tensor[16] $\eta_{\mu\nu}$ which is real, symmetric, invertible, $\det \eta_{\mu\nu} \neq 0$, with inverse $\eta^{\mu\nu}$, so that $\eta^{\mu\nu}\,\eta_{\nu\sigma} = \delta^\mu{}_\sigma$.

We can choose an orthonormal frame in which $\eta_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$. Space-time is non-compact.

Define $x^\mu \equiv \eta^{\mu\nu}\,x_\nu$, so $x_\mu = \eta_{\mu\nu} x^\nu$ etc. $x^\mu = (x_0, -x_i)$.

We use the metric to define "length" of x_μ: $x^2 \equiv \eta_{\mu\nu} x^\mu x^\nu = \eta^{\mu\nu} x_\mu x_\nu = x^\mu x_\mu = x_\mu x^\mu = x_0^2 - |\underline{x}|^2$.

6.2 Lorentz Group

Under a change of frame $x_\mu \mapsto x'_\mu = \Lambda_\mu{}^\nu\, x_\nu$ (so $x^\mu \mapsto x'^\mu = \Lambda^\mu{}_\nu\, x^\nu$).

The postulates of Special Relativity are

- The speed of light c is the same in all frames.
- Space-time is isotropic – the laws of physics are the same in all reference frames.

These lead to:

$$x'^2 = \eta_{\mu\nu}\, x'^\mu\, x'^\nu = \eta_{\mu\nu}\, \Lambda^\mu{}_\rho\, \Lambda^\nu{}_\sigma\, x^\rho x^\sigma = \eta_{\rho\sigma}\, x^\rho\, x^\sigma = x^2 \equiv x_0^2 - |\underline{x}|^2$$

Since this expression is true for all x^μ then

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

thus the transformation preserves the metric.

Since $\Lambda^\mu{}_\rho = (\Lambda^T)^\rho{}_\mu$, we may write this in matrix notation as $\Lambda^T \eta\, \Lambda = \eta$.

Taking the determinant gives: $\det \eta = \det \Lambda^T \det \eta \det \Lambda \Rightarrow \det \Lambda = \pm 1$.

In an orthonormal frame, choosing $\rho = \sigma = 0$ gives

$$(\Lambda^0{}_0)^2 - (\Lambda^i{}_0)^2 = 1$$

Since $(\Lambda^i{}_0)^2 \geq 0$, we must have $(\Lambda^0{}_0)^2 \geq 1$, so either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$. Thus the Lorentz group $O(3,1)$ has 4 disconnected pieces, connected by *time reversal* $T : x_0 \mapsto -x_0$ and *parity* $P : x_i \mapsto -x_i$.

[16] NB the space-time metric tensor $\eta_{\mu\nu}$ is the *not* the Killing form of the Lorentz group – see later.

If we take $\Lambda \in SO_+(3,1)$ (i.e. $\det \Lambda = +1, \Lambda^0{}_0 \geq 1$), we have a *proper orthochronous Lorentz transformation.*[17] We get the other three pieces of the group by taking the products: $P\Lambda$, $T\Lambda$, $PT\Lambda$.

The proper orthochronous transformations are:

- Rotations: $\Lambda = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & R \end{array}\right)$, where R is 3×3 and satisfies $R^T R = 1$, $\det R = 1$, so $R \in SO(3)$.
- Boosts: e.g., $\Lambda = \left(\begin{array}{cc|c} \cosh\eta & -\sinh\eta & \\ -\sinh\eta & \cosh\eta & 0 \\ \hline & 0 & 1 \end{array}\right) \in SO_+(1,1)$.

We get the usual form of the Lorentz boost along the x_1 axis (in units where the speed of light $c = 1$) if we take $\cosh\eta = (1 - v^2)^{-\frac{1}{2}}$, $\sinh\eta = v\,(1 - v^2)^{-\frac{1}{2}}$. The most general orthochronous Lorentz transformation is specified by 3 rotations about orthogonal axes, and 3 Lorentz boosts in orthogonal directions. Therefore $SO(3,1)$ has dimension 6.

6.3 Lorentz algebra

Consider an infinitesimal Lorentz transformation on the (four-)vector [18] V^μ, *i.e.* in the *defining* or *fundamental* or *vector* representation of the Lorentz group:

$$V'^\mu = \Lambda^\mu{}_\nu V^\nu = (\delta^\mu{}_\nu + \omega^\mu{}_\nu)\, V^\nu, \tag{6}$$

Now write $\delta V^\mu = \omega^\mu{}_\nu\, V^\nu = \omega^{\mu\nu}\, V_\nu$ where $\omega^{\mu\nu} = -\omega^{\nu\mu}$ is real and anti-symmetric, so it has 6 real parameters:

- $\omega^{ij} = -\omega^{ji}$, $\forall\, i, j = 1, 2, 3$, give rotations ($\omega^{ij} = \epsilon^{ijk}\,\theta\, n_k$)
- $\omega^{0i} \equiv \nu^i$ give boosts.

In a general representation, D, of the Lorentz group, we write an infinitesimal transformation as

$$D(1 + \omega) = 1 + \frac{i}{2}\omega^{\alpha\beta} M_{\alpha\beta} \tag{7}$$

where $M_{\alpha\beta} = -M_{\beta\alpha}$ with $\alpha, \beta = 0, 1, 2, 3$, are a set of six hermitian matrices, the generators of the Lorentz algebra. (It may help to think of $a = (\alpha, \beta)$ as a single index $a = 1, 2, \ldots 6$, which labels the six generators $M_{\alpha\beta} = M_a$ of the Lorentz group, i.e. similar to T_a in our previous discussions.)

In the *defining* (fundamental, vector) representation, from equations (6) and (7), we find

$$(M_{\alpha\beta})^\mu{}_\nu = -i\left(\delta^\mu_\alpha\, \eta_{\beta\nu} - \delta^\mu_\beta\, \eta_{\alpha\nu}\right) \qquad \text{(exercise)} \tag{8}$$

In words, this is the $({}^\mu{}_\nu)$ element of the $(\alpha, \beta)^{\text{th}}$ (matrix) generator $M_{\alpha\beta}$ in the defining representation of the Lorentz group.

[17] An orthochronous Lorentz transformation preserves the direction of time.

[18] For example $V^\mu = x^\mu$.

To find the commutation relations $[M_{\mu\nu},\, M_{\alpha\beta}]$, we note that $(\Lambda^{-1})^{\mu\nu} = (\Lambda^T)^{\mu\nu}$ for an infinitesimal Lorentz transformation, and use the group properties in an arbitrary representation $D(\Lambda)$ to write

$$\begin{aligned} D(\Lambda)\, D(1+\omega)\, D(\Lambda)^{-1} &= D(1+\Lambda\omega\Lambda^T) \\ \Rightarrow \qquad D(\Lambda)\left(1+\frac{i}{2}\omega^{\alpha\beta}\, M_{\alpha\beta}\right) D(\Lambda)^{-1} &= 1+\frac{i}{2}\left(\Lambda^\alpha{}_\mu\, \omega^{\mu\nu}\, \Lambda^\beta{}_\nu\right) M_{\alpha\beta} \\ \Rightarrow \qquad D(\Lambda)\, \omega^{\alpha\beta}\, M_{\alpha\beta}\, D(\Lambda)^{-1} &= \omega^{\alpha\beta}\Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\, M_{\mu\nu} \end{aligned}$$

In the last two lines, we raised and lowered some pairs of indices and relabelled dummy indices on the RHS ($\alpha \leftrightarrow \mu$ and $\beta \leftrightarrow \nu$) to give $\omega^{\alpha\beta}$ on both sides.

This holds for all infinitesimal parameters $\omega^{\alpha\beta}$, so

$$D(\Lambda)\, M_{\alpha\beta}\, D(\Lambda)^{-1} = \Lambda^\mu{}_\alpha\, \Lambda^\nu{}_\beta\, M_{\mu\nu}.$$

Now take $\Lambda = 1+\epsilon$, *i.e.* $\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu$ and expand through $O(\epsilon)$. Substituting $D(\Lambda) = 1+\frac{i}{2}\epsilon^{\mu\nu} M_{\mu\nu}$ into the above expression, and cancelling the leading $M_{\alpha\beta}$ from both sides, gives

$$\begin{aligned} \frac{i}{2}\epsilon^{\mu\nu}\,[M_{\mu\nu},\, M_{\alpha\beta}] &= \epsilon^\mu{}_\alpha\, \delta^\nu{}_\beta\, M_{\mu\nu} + \delta^\mu{}_\alpha\, \epsilon^\nu{}_\beta\, M_{\mu\nu} \\ &= \epsilon^\mu{}_\alpha\, M_{\mu\beta} - \epsilon_\beta{}^\nu\, M_{\alpha\nu} \\ &= \frac{1}{2}\epsilon^{\mu\nu}\left(M_{\mu\beta}\,\eta_{\nu\alpha} - M_{\nu\beta}\,\eta_{\mu\alpha} - M_{\alpha\nu}\,\eta_{\mu\beta} + M_{\alpha\mu}\,\eta_{\nu\beta}\right) \end{aligned}$$

where the quantity in parentheses is explicitly anti-symmetric under $\mu \leftrightarrow \nu$ and $\alpha \leftrightarrow \beta$, and we used the metric $\eta_{\mu\nu}$ to raise and lower indices. We also used the anti-symmetry $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$, several times. The result holds for all $\epsilon^{\mu\nu}$, which is also anti-symmetric, so

$$[M_{\alpha\beta},\, M_{\mu\nu}] = i\left(M_{\alpha\mu}\,\eta_{\beta\nu} - M_{\alpha\nu}\,\eta_{\beta\mu} - M_{\beta\mu}\,\eta_{\alpha\nu} + M_{\beta\nu}\,\eta_{\alpha\mu}\right) \qquad (9)$$

which defines the algebra, $so(3,1)$, of the Lorentz group $SO(3,1)$

The generators M_{ij}, $i,j = 1,2,3$, generate an $so(3)$ subalgebra: $so(3) \subset so(3,1)$. To see this, we write it in a more familiar form by first defining

$$J_i \equiv -\tfrac{1}{2}\epsilon_{ijk} M_{jk}\,.$$

Then

$$\begin{aligned} [J_i, J_j] &= \frac{1}{4}\,\epsilon_{imn}\,\epsilon_{jpq}\,[M_{mn},\, M_{pq}] \\ &= \frac{i}{4}\,\epsilon_{imn}\,\epsilon_{jpq}\,\left(M_{mp}\,\eta_{nq} - M_{mq}\,\eta_{np} - M_{np}\,\eta_{mq} + M_{nq}\,\eta_{mp}\right) \\ &= -\frac{i}{4}\,\epsilon_{imn}\,\epsilon_{jpq}\,\left(M_{mp}\,\delta_{nq} - M_{mq}\,\delta_{np} - M_{np}\,\delta_{mq} + M_{nq}\,\delta_{mp}\right) \quad (\text{as } \eta_{nq} = -\delta_{nq},\ etc) \\ &= -i\left(\delta_{ij}\delta_{mp} - \delta_{ip}\delta_{mj}\right) M_{mp} \\ &= -i\left(\delta_{ij}\, M_{mm} - M_{ji}\right) = -i\, M_{ij} \\ &= i\,\epsilon_{ijk}\, J_k \end{aligned}$$

where in the last line we inverted the definition of J_i to obtain $M_{ij} = -\,\epsilon_{ijk}\, J_k$. So, finally,

$$[J_i, J_j] = i\epsilon_{ijk} J_k,$$

which is indeed the familiar algebra of $su(2) \cong so(3)$. [See also tutorial question (3.2).] So J_i are just the usual *rotation generators.*

Similarly let $K_i = M_{0i}$, then K_i are the *boost generators.* We find

$$[K_i, K_j] = [M_{0i}, M_{0j}] = iM_{ij} = -i\epsilon_{ijk}\, J_k$$

and

$$[J_i, K_j] = -\tfrac{1}{2}\epsilon_{imn}\,[M_{mn}, M_{0j}] = -\tfrac{1}{2}\epsilon_{imn}\, i\,\left(M_{m0}\,\eta_{nj} - M_{n0}\,\eta_{mj}\right) = i\epsilon_{ijk}\, M_{0k} = i\epsilon_{ijk} K_k.$$

Thus the algebras satisfied by the generators J_i and K_i are *coupled* or *entangled*:

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \qquad [K_i, K_j] = -i\epsilon_{ijk}\, J_k, \qquad [J_i, K_j] = i\epsilon_{ijk} K_k.$$

We can disentangle them by defining new generators N_i, which are linear combinations of J_i and K_i

$$N_i \equiv \tfrac{1}{2}\left(J_i + iK_i\right).$$

Note that these are *not* hermitian, $N_i^\dagger = \frac{1}{2}(J_i - iK_i) \neq N_i$.

We find

$$\begin{aligned} [N_i,\, N_j] &= \tfrac{1}{4}\left([J_i,\, J_j] + i\,[J_i,\, K_j] + i\,[K_i,\, J_j] - [K_i,\, K_j]\right) \\ &= \tfrac{1}{4}\left(i\,\epsilon_{ijk}\, J_k - \epsilon_{ijk}\, K_k + \epsilon_{jik}\, K_k + i\,\epsilon_{ijk}\, J_k\right) \\ &= i\epsilon_{ijk}\, \tfrac{1}{2}\, \left(J_k + iK_k\right) \\ &= i\epsilon_{ijk}\, N_k \end{aligned}$$

Similarly

$$[N_i,\, N_j^\dagger] = \tfrac{1}{4}\left([J_i,\, J_j] - i\,[J_i,\, K_j] + i\,[K_i,\, J_j] + [K_i,\, K_j]\right) = 0.$$

and

$$[N_i^\dagger,\, N_j^\dagger] = i\epsilon_{ijk}\, N_k^\dagger$$

Altogether

$$[N_i,\, N_j] = i\epsilon_{ijk}\, N_k \qquad [N_i^\dagger,\, N_j^\dagger] = i\epsilon_{ijk}\, N_k^\dagger \qquad [N_i,\, N_j^\dagger] = 0$$

i.e., the six generators N_i and $N_i^\dagger$ generate two *separate* $su(2)$ algebras.

These two $su(2)$ algebras are not completely independent, since they are related by hermitian conjugation, or under *parity* ($J_i \mapsto J_i$, $K_i \mapsto -K_i$, so $N_i \mapsto N_i^\dagger$).

Casimirs: we can construct two Casimir operators by taking the product of generators with invariant tensors $\eta_{\mu\nu}$ and $\epsilon_{\mu\nu\rho\sigma}$

$$\tfrac{1}{2}M_{\mu\nu}\, M^{\mu\nu} = J^2 - K^2 \qquad \tfrac{1}{4}\epsilon_{\mu\nu\rho\sigma}M^{\mu\nu}M^{\rho\sigma} = -\underline{J}\cdot\underline{K}.$$

The commutation relations show that $so(3,1)$ has rank 2 and is semi-simple, but it's not compact: the Killing form is $\mathrm{diag}(1,1,1,-1,-1,-1)$: rotations and boosts give opposite signs.

6.3.1 Relation to $sl(2,\mathbb{C})$

The six (Pauli) matrices $\{\frac{1}{2}\sigma_i, -\frac{i}{2}\sigma_j\}$ satisfy the *same* algebra as $\{J_i, K_j\}$, so they may be chosen as a basis of generators that generate rotations and boosts respectively.

These six 2×2 complex traceless (but not all hermitian) matrices are the generators of $sl(2,\mathbb{C})$ in the fundamental representation. Therefore $sl(2,\mathbb{C})$ and $so(3,1)$ are *isomorphic*: $sl(2,\mathbb{C}) \cong so(3,1)$.

In fact $SL(2,\mathbb{C})$ is the covering group of the Lorentz group

$$SO(3,1) \;\cong\; SL(2,\mathbb{C})/\mathbb{Z}_2\,,$$

just as $SU(2)$ is the covering group of $SO(3)$. [See tutorial question (3.3).]

6.4 Representations of $sl(2,\mathbb{C})$

The generators N_3 and $N_3^\dagger$ commute. (Recall that they're not hermitian: $sl(2,\mathbb{C})$ is *not* compact.)

Since the N_i and $N_i^\dagger$ satisfy decoupled $su(2)$ algebras, we can construct two pairs of step operators $N_\pm$ and $N_\pm^\dagger$ as we did for a single $su(2)$ algebra.

There are two Casimir operators:

1. $N^2 \equiv N_i N_i$ with eigenvalues $m(m+1)$
2. ${N^\dagger}^2 \equiv N_i^\dagger N_i^\dagger$ with eigenvalues $n(n+1)$

where $m, n = 0, \frac{1}{2}, 1, \ldots$, just as for a single $su(2)$.

Since we have two $su(2)$ algebras, we label each representation by (m,n), and we label the states within the representation by the eigenvalues of N_3 and $N_3^\dagger$, which run from $-m, -m+1, \ldots, m$, and $-n, \ldots, n$ respectively; i.e., there are $(2m+1)(2n+1)$ states. The spin of the states within the representation (m,n) varies from $|m-n|$ to $m+n$, since $J_i = N_i + N_i^\dagger$. This is just like adding angular momenta in quantum mechanics.

Example.

1. $(0,0)$ spin zero: scalar.
2. $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ spin $\frac{1}{2}$: *spinors* (the fundamental representation of $sl(2,C)$).
3. $(\frac{1}{2},\frac{1}{2})$ spin one: representation of a *vector*, $V_\mu = (V_0, \underline{V})$ (see below).
4. $(1,0)$ and $(0,1)$ spin one antisymmetric tensor $F_{\mu\nu}$ (self-dual and anti-self-dual – see later). The adjoint representation is $(1,0)\oplus(0,1)$ ($F_{\mu\nu}$ has 6 components).
5. $(1,1)$ spin two: a symmetric traceless tensor $T_{\mu\nu}$.

We will explore these in more detail as the course progresses.

6.4.1 Euclidean Group

If we consider "imaginary time", *i.e.* let $t \mapsto it$, we recover "Euclidean space", $-\eta_{\mu\nu} \mapsto \delta_{\mu\nu}$, and the symmetry group is $SO(4) \cong SU(2)\otimes SU(2)/\mathbb{Z}_2$ - see tutorial questions (3.2) and (3.3). Unlike $SO(3,1)$, $SO(4)$ *is* compact: $g_{ab} = \mathrm{diag}(1,1,1,1,1,1)$. The two $su(2)$ subalgebras of $so(4)$ are completely independent, and each state is characterised by two "spins", n, m, both of which are "observables" (correspond to hermitian generators).

6.4.2 $sl(2,\mathbb{C})$ Spinors

The spinor representations $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ of the Lorentz group (or more precisely of $SL(2,\mathbb{C})$) are realised by *two-component complex Weyl spinors*, ψ_L and ψ_R respectively. The choice of labels L, R is a convention.

Spinor indices are usually suppressed (although one sometimes sees $\psi_a \in (\frac{1}{2},0)$, $\psi_{\dot a} \in (0,\frac{1}{2})$, with $a, \dot a = 1,2$). (Note the "dotted index" $\dot a$ in the second case.)

Under an $SL(2,\mathbb{C})$ transformation

$$\begin{aligned} \psi_L &\mapsto \psi_L' = \Lambda_L\,\psi_L \\ \psi_R &\mapsto \psi_R' = \Lambda_R\,\psi_R. \end{aligned}$$

where Λ_L and Λ_R are 2×2 complex matrices.

Under a *rotation* in $SU(2)\subset SL(2,\mathbb{C})$, taking generators $J_i = \frac{1}{2}\sigma_i$,

$$\Lambda_{L,R} = \exp\left\{i\tfrac{1}{2}\underline{\sigma}\cdot\underline{\omega}\right\}\,,$$

where $\omega_i = -\frac{1}{2}\epsilon_{ijk}\omega^{jk}$ are the rotation parameters.

The representation of boosts is not unitary. If we take boost generators $K_i = -\frac{i}{2}\sigma_i$ with boost parameters $\nu^i = \omega^{0i}$, then

$$\begin{aligned} \Lambda_L &= \exp\left\{\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega} - i\underline{\nu})\right\} \\ \Lambda_R &= \exp\left\{\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega} + i\underline{\nu})\right\} \end{aligned}$$

Since the $(\frac{1}{2},0)$ and $(0,\frac{1}{2})$ representations are related by parity, we construct Λ_R from Λ_L by changing the sign of the boost parameters.

So, under a parity transformation, $\Lambda_L \leftrightarrow \Lambda_R$ and $\psi_L \leftrightarrow \psi_R$.

Properties of $\Lambda_{L,R}$:

1. $\Lambda_L^{-1} = \Lambda_R^\dagger$ (NB, in Euclidean space Λ_L and Λ_R are *completely independent.*)
2. Since $\sigma_2\sigma_i\sigma_2 = -\sigma_i^*$, $\sigma_2\,\Lambda_L\,\sigma_2 = \Lambda_R^*$, $\sigma_2\Lambda_L^\dagger\sigma_2 = \Lambda_R^T$.
3. Since $\sigma_2\,\sigma_i\,\sigma_2 = -\sigma_i^T$, $\sigma_2\Lambda_L^T\sigma_2 = \Lambda_L^{-1}$, and hence $\Lambda_L^T\sigma_2\Lambda_L = \sigma_2$.
   Similarly $\Lambda_R^T\sigma_2\Lambda_R = \sigma_2$.

Exercise: check all of these.

Consider a left handed (LH) spinor $\psi_L \in (\frac{1}{2}, 0)$. Under $SL(2,\mathbb{C})$

$$\sigma_2 \psi_L^* \mapsto \sigma_2 \Lambda_L^* \psi_L^* \;=\; \Lambda_R(\sigma_2 \psi_L^*)$$

using result 2. above and $\sigma_2^2 = 1$. So $\sigma_2\,\psi_L^* \in (0, \frac{1}{2})$ is a right handed spinor, which we call the *charge conjugate* of ψ_L. Similarly $\sigma_2\,\psi_R^*$ is a left handed spinor.

Define *charge conjugation* $C:\ \psi_L \mapsto \sigma_2\,\psi_L^*$, which is right handed, and $\psi_R \mapsto \sigma_2\,\psi_R^*$, which is left-handed.

From two left handed spinors ψ_L and χ_L we can construct a *scalar* $(0,0)$ under $SL(2,\mathbb{C})$, since $(\frac{1}{2},0)\otimes(\frac{1}{2},0) = (0,0)\oplus(1,0)$ (just as spin-$\frac{1}{2}$ times spin-$\frac{1}{2}$ gives spin-0 plus spin-1 in quantum mechanics). Explicitly:

$$\chi_L^T\,\sigma_2\,\psi_L \;\mapsto\; \chi_L^T\,\Lambda_L^T\,\sigma_2\,\Lambda_L\,\psi_L \;=\; \chi_L^T\,\sigma_2\,\psi_L$$

where we used result 3. above.

In particular, if we choose χ_L such that $\psi_R = \sigma_2\,\chi_L^*$, and hence $\psi_R^\dagger = \chi_L^T\,\sigma_2$, then $\psi_R^\dagger\,\psi_L$ is a scalar. Similarly, $\psi_L^\dagger\,\psi_R$ is a scalar.

We can also build a 4-vector from two spinors: $(\frac{1}{2},0)\otimes(0,\frac{1}{2}) = (\frac{1}{2},\frac{1}{2})$. Under an infinitesimal Lorentz transformation

$$\begin{aligned}
\psi_L^\dagger\psi_L \;&\mapsto\; \psi_L^\dagger\left(1-\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega}+i\underline{\nu})\right)\left(1+\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega}-i\underline{\nu})\right)\psi_L \\
&=\; \psi_L^\dagger\,\psi_L - \omega_{0i}\,\psi_L^\dagger\,\sigma_i\,\psi_L \qquad (\text{since } \nu_i = \omega_{i0} = -\omega_{0i})
\end{aligned}$$

while

$$\begin{aligned}
\psi_L^\dagger\sigma_i\psi_L \;&\mapsto\; \psi_L^\dagger\left(1-\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega}+i\underline{\nu})\right)\sigma_i\left(1+\frac{i}{2}\underline{\sigma}\cdot(\underline{\omega}-i\underline{\nu})\right)\psi_L \\
&=\; \psi_L^\dagger\left(\sigma_i + \frac{i}{2}\,[\sigma_i,\sigma_j]\,\omega_j + \tfrac{1}{2}\,\{\sigma_i,\sigma_j\}\,\nu_j\right)\psi_L \\
&=\; (\delta_{ij}-\omega_{ij})\psi_L^\dagger\,\sigma_j\,\psi_L + \omega_{i0}\,\psi_L^\dagger\,\psi_L.
\end{aligned}$$

where we recalled that $\omega_i = -\frac{1}{2}\epsilon_{ijk}\,\omega_{jk}$, which we inverted to get $\omega_{ij} = -\epsilon_{ijk}\,\omega_k$.

It follows that if we define $\sigma_\mu \equiv (1,\sigma_i)$, then

$$\psi_L^\dagger\sigma_\mu\psi_L \mapsto (\delta_\mu{}^\nu + \omega_\mu{}^\nu)\,\psi_L^\dagger\,\sigma_\nu\,\psi_L$$

We conclude that $\psi_L^\dagger\,\sigma_\mu\,\psi_L$ is a 4-vector because it transforms in the same way as x_μ under Lorentz transformations.

Similary, if we define $\bar\sigma_\mu \equiv (1,-\sigma_i)$, then $\psi_R^\dagger\,\bar\sigma_\mu\,\psi_R$ is also a 4-vector.

6.4.3 Dirac Spinors

As we saw earlier, the spinors ψ_L and ψ_R transform into each other under parity, $\psi_L \leftrightarrow \psi_R$.

If we want invariance under parity, we need to construct a *Dirac spinor* $\psi \equiv \begin{pmatrix}\psi_L\\ \psi_R\end{pmatrix}$ with *four* components, transforming as $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ under $SL(2,\mathbb{C})$.

$$\psi \;\mapsto\; \begin{pmatrix}\Lambda_L & 0\\ 0 & \Lambda_R\end{pmatrix}\begin{pmatrix}\psi_L\\ \psi_R\end{pmatrix} \;\equiv\; \Lambda\,\psi.$$

(Note: $4 = 2\,+\,2$ here, whereas $4 = 2\times 2$ in our previous construction of a four-vector.)

Under parity, the Dirac spinor transforms as

$$\psi \;\mapsto\; \begin{pmatrix}\psi_R\\ \psi_L\end{pmatrix} \;=\; \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\psi \;\equiv\; \gamma_0\,\psi.$$

where γ_0 is a 4×4 matrix, with each 1 representing a 2×2 unit submatrix

We can project ψ onto L and R spinors using the projection operators $P_{L,R} = \frac{1}{2}(1\pm\gamma^5)$, where

$$\gamma^5 \;\equiv\; \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

NB Conventions for these projection operators with the opposite signs are perhaps more popular!

From Dirac spinors we can construct two scalars under $SL(2,\mathbb{C})$:

$$\begin{array}{lrclcll}
 & \psi_L^\dagger\psi_R + \psi_R^\dagger\psi_L &=& \psi^\dagger\gamma^0\psi &\equiv& \bar\psi\psi & \text{even under parity (scalar)}\\
\text{and} & \psi_L^\dagger\psi_R - \psi_R^\dagger\psi_L &=& -\psi^\dagger\gamma_0\gamma^5\psi &=& -\bar\psi\gamma^5\psi & \text{odd under parity (pseudoscalar).}
\end{array}$$

Note the definition $\bar\psi \equiv \psi^\dagger\gamma_0$.

Similarly there are two 4-vectors:

$$\begin{array}{rcll}
\psi_L^\dagger\sigma_\mu\psi_L + \psi_R^\dagger\bar\sigma_\mu\psi_R &=& \bar\psi\gamma_\mu\psi & \text{odd under parity (vector)}\\
\psi_L^\dagger\sigma_\mu\psi_L - \psi_R^\dagger\bar\sigma_\mu\psi_R &=& \bar\psi\gamma_\mu\gamma^5\psi & \text{even under parity (axial vector)}
\end{array}$$

where $\gamma_\mu = \begin{pmatrix}0 & \bar\sigma_\mu\\ \sigma_\mu & 0\end{pmatrix}$, i.e., $\gamma_i = \begin{pmatrix}0 & -\sigma_i\\ \sigma_i & 0\end{pmatrix}$.

These are the *Dirac matrices* (in the Weyl representation). Remembering that $\{\sigma_i,\sigma_j\} = 2\delta_{ij}$, it's easy to check that they form a *Clifford algebra*

$$\{\gamma_\mu,\gamma_\nu\} \;=\; 2\eta_{\mu\nu}\,,\qquad \gamma^5 \;=\; i\gamma_0\,\gamma_1\,\gamma_2\,\gamma_3\,,\qquad \{\gamma^5,\gamma_\mu\} \;=\; 0\,,\qquad {\gamma_\mu}^\dagger \;=\; \gamma_0\,\gamma_\mu\,\gamma_0\,,\qquad {\gamma^5}^\dagger \;=\; \gamma^5\,.$$

These expressions should be familiar if you're taking the *Quantum Field Theory* course. Note however that we obtained them here using group theory only. We didn't input any quantum field theory or indeed any quantum mechanics.

6.5 The Poincaré Group

Space-time is isotropic and homogeneous: this corresponds to invariance under Lorentz transformations and space-time translations (respectively)

$$x^\mu \mapsto x'^\mu \;=\; \Lambda^\mu{}_\nu\,x^\nu + a^\mu$$

These are called *inhomogeneous Lorentz transformations.*

Translations and Lorentz transformations do not commute, because the translation parameters are 'rotated' by a Lorentz transformation. Under two successive transformations, the second having the tilde

$$x^\mu \mapsto x''^\mu \;=\; \tilde\Lambda^\mu{}_\nu\,x'^\nu + \tilde a^\mu \;=\; (\tilde\Lambda^\mu{}_\nu\,\Lambda^\nu{}_\alpha)\,x^\alpha + (\tilde\Lambda^\mu{}_\nu\,a^\nu + \tilde a^\mu).$$

If we denote a general Poincaré transformation as $U(a, \Lambda)$, then

$$U(a,\Lambda) \;=\; U(a,1)\,U(0,\Lambda) \;\equiv\; U(a)\,U(\Lambda) \qquad \text{(shorthand notation)}$$

and, from the result for two successive transformations above, we find

$$U(\tilde{a},\tilde{\Lambda})\,U(a,\Lambda) \;=\; U(\tilde{\Lambda}a+\tilde{a},\tilde{\Lambda}\Lambda)\,.$$

Let us define the generators $M_{\mu\nu}$ and P_μ of the Poincaré group by

$$\begin{aligned} U(\Lambda) &= 1+\tfrac{i}{2}\omega^{\mu\nu}M_{\mu\nu} && \text{with } \omega^{\mu\nu} \text{ infinitesimal} \\ U(a) &= 1+iP^\mu\,a_\mu && \text{with } a_\mu \text{ infinitesimal} \end{aligned} \tag{10}$$

We can deduce the commutation relations by setting:

1. $\Lambda \,=\, \tilde{\Lambda} \,=\, 1:\; U(\tilde{a})\,U(a) \,=\, U(\tilde{a}+a) \,=\, U(a)\,U(\tilde{a})$, so $[P_\mu,P_\nu] \,=\, 0$.
2. $a \,=\, \tilde{a} \,=\, 0:$ $M_{\mu\nu}$ satisfy the $so(3,1)$ commutation relations (the Lorentz algebra) as before.
3. $\tilde{a} \,=\, 0$, a infinitesimal, $\tilde{\Lambda} \,=\, \Lambda^{-1}$, so $U(\Lambda^{-1})\,U(a)\,U(\Lambda) \,=\, U(\Lambda^{-1}a,1)$. Then
$$U^{-1}(\Lambda)\,P^\mu\,U(\Lambda) \;=\; P^\nu\,\Lambda_\nu{}^\mu,$$
i.e., P_μ transforms as a 4-vector (8). If Λ is infinitesimal, we find (exercise)
$$[M_{\mu\nu},P_\sigma] \;=\; -i(P_\mu\,\eta_{\nu\sigma}-P_\nu\,\eta_{\mu\sigma}).$$
This holds for any 4-vector.

In terms of $J_i = -\frac{1}{2}\epsilon_{ijk}M_{jk}$, and $K_i = M_{0i}$, and writing the four translation generators as $P_\mu \equiv (H,P_i)$, the commutation relations become

$$\begin{aligned} [J_i,P_j] &= i\,\epsilon_{ijk}\,P_k & [J_i,H] &= 0 \\ [K_i,P_j] &= i\,H\,\delta_{ij} & [K_i,H] &= -iP_i\,. \end{aligned}$$

Together with $[P_\mu,P_\nu]=0$, these commutation relations form the algebra of the Poincaré group. The Poincaré group has dimension $6+4=10$, but it's neither compact nor semi-simple, the latter because P_μ form an Abelian invariant subalgebra.

Despite this, we can still find two Casimir operators:

1. $P^2 = P_\mu P^\mu:\; [P_\mu,P^2]=[M_{\mu\nu},P^2]=0$
2. The other Casimir operator is not so obvious, but the square of any 4-vector that commutes with the generators and is linearly independent of P_μ will suffice.

   Define the *Pauli–Lubanski vector*
$$W_\mu \;\equiv\; \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}P^\sigma\,,$$
   where $\epsilon_{\mu\nu\rho\sigma}$ (with $\epsilon_{0123}=+1$) is totally antisymmetric.

   Then $W_\mu P^\mu = 0$ by anti-symmetry, and $[W_\mu,P_\nu]=0$ using the commutation relations between $M^{\rho\sigma}$ and P^σ. W_μ is a 4-vector, so
$$[M_{\mu\nu},W_\sigma] = -i(W_\mu\eta_{\nu\sigma}-W_\nu\eta_{\mu\sigma})\,.$$
   Thus $W^2 \equiv W_\mu W^\mu$ is a scalar, and $[P_\mu,W^2]=[M_{\mu\nu},W^2]=0$, so W^2 is a Casimir operator.

6.6 Classification of Irreducible Representations

Denote the eigenvalues of P_μ by $p_\mu = (p_0,\underline{p})$ (energy and momentum), and of P^2 by p^2. Note that p_μ are *continuous* (not discrete) because we can rescale $P_\mu \mapsto \alpha P_\mu$ for any α without affecting any of the commutation relations. This is always true for generators of invariant Abelian sub-algebras.

Let's consider all possible classes of values of the eigenvalues p_μ:

1. $p^2=0$, $p_\mu=0$, whence $W^2=0$ in all states. This corresponds to the *vacuum* (the lowest energy state) in physics.

2. $p^2>0$ (timelike): set $p^2=m^2$, where m is the *mass*, so $p_0^2=m^2+|\underline{p}^2|$. We then have either $p_0>m$ (positive energy) or $p_0<-m$ (negative energy) states (the sign of p_0 is an invariant for orthochronous Lorentz transformations). If we go to the rest frame $p_\mu=(m,\underline{0})$, then
$$W_\mu = \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}M^{\nu\rho}\,m\,\delta_0^\sigma$$
   so $W_0 \,=\, 0$ and $W_i \,=\, \frac{1}{2}\,m\,\epsilon_{ijk}\,M^{jk} \,=\, -mJ_i$, where J_i satisfy the $su(2)$ algebra. Therefore W^2 (remember this is a Poincaré invariant) has eigenvalues $-m^2s(s+1)$ for $s=0,\frac{1}{2},1,\ldots$.

   We conclude that all representations with $p^2>0$ are labelled by p_i (momentum, continuous), spin s, and $s_3=-s,-s+1,\ldots,s-1,s$, a total of $2s+1$ degrees of freedom). Physically, a state corresponds to a particle of mass m, spin s, 3-momentum p_i, and spin projection s_3. Note: this is pure group theory [Weyl & Wigner (1939)], no quantum mechanics is involved!

3. $p^2=0$: these are massless states, $m=0$, , so there is no "rest frame". Choose instead a frame in which $p_\mu=(p_0,0,0,p_0)$. Only $W^2=0$ seems relevant for physics (otherwise we would have infinite-dimensional representations, which Wigner called "continuous spin representations" – see Ramond's book). Since $W_\mu P^\mu=0$, we have $W_3=W_0$ and $W_1=W_2=0$, so
$$W_\mu = \lambda P_\mu.$$
   Since W_μ and P_μ are four vectors, λ is an invariant, called the *helicity.* Let $\eta_\mu=(1,\underline{0})$, then
$$\lambda \;=\; \frac{W_\mu\eta^\mu}{P_\mu\eta^\mu} \;=\; \frac{\frac{1}{2}\epsilon_{ijk}M^{ij}P^k}{P_0} \;=\; \frac{\underline{J}\cdot\underline{P}}{|\underline{P}|}$$
   which is the usual definition of helicity.

   One can show that: $\lambda \,=\, \pm s$, where s is the (integer or half-integer) spin of the representation, i.e. $\lambda=0$ or $\pm\frac{1}{2}$ or $\pm 1,\ldots$. So massless representations are labelled by λ (left-handed or right-handed states), states within a representation are labelled by p_i.

   For example, the massless spin-1 photon has *only* 2 polarisation states $\lambda=\pm 1$, the massless spin-2 graviton has *only* 2 polarisation states $\lambda=\pm 2$.

4. $p^2<0$: these are tachyonic states and are unphysical. Since they move faster than the speed of light, they are acausal.

6.7 Fields

Motivation: properties of Casimir operators (Casimirs):

- Semi-simple or compact groups: the Casimirs have a discrete spectrum: states correspond to finite dimensional vectors (or spinors or tensors). Our standard example is $SU(2)$ with states $|j,m\rangle$.
- Non-semi-simple groups: some Casimirs have a continuous spectrum: states correspond to continuous functions – fields. Example: Poincaré group.

Because the eigenvalues of P_μ are continuous, we need to introduce fields $\phi_s(x_\mu)$, where s labels the spin components. This is sufficient to construct massive and massless particle states by superposition of states of the form

$$|x,s\rangle \;\equiv\; \phi_s(x_\mu)\,|0\rangle\,,$$

where $|0\rangle$ is the vacuum state.

Under a Poincaré transformation, a state $|\;\rangle \mapsto U^{-1}(a,\Lambda)\,|\;\rangle$, so the field transforms as $\phi(x) \mapsto U^{-1}(a,\Lambda)\,\phi(x)\,U(a,\Lambda)$.

For a *scalar field*

$$U^{-1}\phi\,(x)\,U \;=\; \phi\,(\Lambda x + a)\,. \tag{11}$$

For an infinitesimal transformation, $\phi(x_\mu) \mapsto \phi(x_\mu + \delta x_\mu) = \phi(x_\mu) + \delta\phi(x_\mu)$, therefore $\delta\phi(x_\mu) = \delta x_\mu \frac{\partial}{\partial x_\mu}\phi(x_\mu) \equiv \delta x_\mu\,\partial^\mu\,\phi(x)$ with $\delta x_\mu = a_\mu + \omega_\mu{}^\nu x_\nu$, where we defined the four-vector operators $\partial^\mu = \left(\frac{\partial}{\partial x_0}, \underline{\nabla}\right)$, $\partial_\mu = \left(\frac{\partial}{\partial x_0}, -\underline{\nabla}\right)$.

Note that under a Lorentz transformation

$$\frac{\partial}{\partial x_\mu} \;\mapsto\; \frac{\partial}{\partial x'_\mu} \;=\; \frac{\partial x_\nu}{\partial x'_\mu}\,\frac{\partial}{\partial x_\nu} \;=\; \Lambda^\mu{}_\nu\,\frac{\partial}{\partial x_\nu},$$

so if ϕ is a scalar, $\partial_\mu\phi$ and $\partial^\mu\phi$ are vectors, and ∂_μ is a *vector operator*.

Similarly $\partial_\mu\partial_\nu\phi$ is a tensor, and $\partial^2\phi \equiv \partial_\mu\partial^\mu\phi$ is a scalar. The operator $\partial^2 = \Box = \frac{\partial^2}{\partial x_0^2} - \nabla^2$ is called the d'Alembertian, the relativistic generalisation of the Laplacian.

In terms of generators, from equations (10) & (11),

$$\delta\phi \;=\; -ia_\mu\,[P^\mu,\phi] - \frac{i}{2}\omega_{\mu\nu}\,[M^{\mu\nu},\phi] \;=\; a_\mu\,\partial^\mu\,\phi \;+\; \omega_\mu{}^\nu x_\nu\,\partial^\mu\phi$$

as $\delta x_\mu = a_\mu + \omega_\mu{}^\nu x_\nu$. Since this holds for all (infinitesimal) a_μ and $\omega_{\mu\nu} = -\omega_{\nu\mu}$,

$$\begin{aligned} [P^\mu,\phi] &= i\partial^\mu\phi \\ [M^{\mu\nu},\phi] &= i\,(x^\nu\partial^\mu - x^\mu\partial^\nu)\,\phi \;=\; (x^\nu P^\mu - x^\mu P^\nu)\,\phi. \end{aligned}$$

Writing $P_\mu = (H, P_i)$,

$$\begin{aligned} [H,\phi] &= i\tfrac{\partial}{\partial t}\phi && \text{Energy (Heisenberg equation)} \\ [\underline{P},\phi] &= -i\,\underline{\nabla}\phi && \text{Momentum} \\ [\underline{J},\phi] &= -i\underline{x}\times\underline{\nabla}\phi && \text{Angular momentum} \end{aligned}$$

where $J_i = -\frac{1}{2}\epsilon_{ijk}M_{jk}$.

NB, we are using units where $\hbar = c = 1$. to restore $\hbar$ remember that $\hbar$ has dimensions of angular momentum, so let $P_\mu \mapsto P_\mu/\hbar$, $M_{\mu\nu} \mapsto M_{\mu\nu}/\hbar$.

Since the vacuum has $P_\mu\,|0\rangle = 0$, it follows that

$$H\,|\underline{x},t\rangle \;=\; i\frac{\partial}{\partial t}\,|\underline{x},t\rangle \qquad \underline{P}\,|\underline{x},t\rangle \;=\; -i\underline{\nabla}\,|\underline{x},t\rangle$$

etc. (Schrödinger equation).

A Fourier transformation to momentum space is often useful:

$$\phi(p_\mu) \;\equiv\; \int d^4x\,e^{-ip_\mu x^\mu}\,\phi(x_\mu)$$

etc., then $[P_\mu, \phi(p_\mu)] = p_\mu\,\phi(p_\mu)$ etc.

For fields which are *not scalars* under the Lorentz group, there is an 'extra' factor of $D(\Lambda) = \exp(\frac{i}{2}\omega^{\mu\nu}S_{\mu\nu})$ in the transformation law

$$U^{-1}(\Lambda,a)\,\phi(x)\,U(\Lambda,a) \;=\; D(\Lambda)\,\phi(\Lambda x + a)$$

whence $[P_\mu,\phi] \;=\; i\partial_\mu\phi$ while

$$[M_{\mu\nu},\phi] \;=\; \underbrace{i(x_\nu\partial_\mu - x_\mu\partial_\nu)}_{L_{\mu\nu}}\,\phi + S_{\mu\nu}\phi$$

where $L_{\mu\nu}$ is the 'angular momentum' (generalised to 4 dimensions) we obtained above for the scalar field, while $S_{\mu\nu}$ is the (intrinsic) 'spin' angular momentum (in the appropriate representation of the Lorentz group, $D(\Lambda) = \exp\frac{i}{2}\omega^{\mu\nu}S_{\mu\nu}$).

For example, for a scalar $S_{\mu\nu} = 0$, for a Weyl spinor $\omega^{\mu\nu}S_{\mu\nu} = \frac{i}{2}\omega^{\mu\nu}\sigma_{\mu\nu}$, and for a vector (i.e. the defining or fundamental representation) $(\omega^{\mu\nu}S_{\mu\nu})^\alpha{}_\beta = \omega^\alpha{}_\beta$.

Note that:

1. $L_{\mu\nu}$ satisfy the $so(3,1)$ algebra. (See tutorial question (4.1).)
2. The Pauli–Lubanski vector reduces to $W_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}S^{\nu\rho}P^\sigma$, because $\epsilon_{\mu\nu\rho\sigma}L^{\nu\sigma}P^\sigma = 2\epsilon_{\mu\nu\rho\sigma}x^\nu P^\rho P^\sigma = 0$, as $[P^\rho, P^\sigma] = 0$. Therefore, only the *intrinsic* angular momentum (or "spin") contributes to the eigenvalues of W^2 — single particle states have no angular momentum. (In other words, they can have spin, but not "orbital" angular momentum.)

7 Symmetry in Action

We assume that our dynamics (classical or quantum — here mainly classical) may be derived from an action principle. To construct suitable actions we follow a set of fundamental principles:

1. *Locality*: we assume the action has the form
$$S = \int d^4x\, \mathcal{L}$$
where $d^4x = dx_0\, dx_1\, dx_2\, dx_3$ is the Lorentz invariant integration measure in Minkowski space, and $\mathcal{L}$ is the *Lagrange density* (usually informally called the "Lagrangian"). We assume that $\mathcal{L}$ is a function of local fields Φ (which could be scalars, spinors, vectors, $\ldots$) and their derivatives at one point x^μ only - so there's no "action at a distance"). Locality is necessary to ensure causality in the quantum theory.

2. *Unitarity*: we assume that the action (and thus the Lagrangian) is *real*: this is essential to obtain quantum theories in which probability is conserved (note in classical physics complex potential $\Rightarrow$ absorption).

3. *Causality*: we require that S leads to classical equations of motion that involve no higher than second derivatives: higher order differential equations generally have acausal solutions. This means that $\mathcal{L}$ contains terms with at most two derivatives ∂_μ.

4. *Symmetry*: we assume that S is Poincaré invariant up to surface terms: since d^4x is invariant, this means that $\mathcal{L}$ must be a Lorentz invariant function of the fields $\Phi(x)$ and $\partial_\mu\Phi(x)$ (terms involving x_μ explicitly will not be translation invariant).

   We will also consider further symmetries (internal symmetries, gauge symmetries, etc.) to further restrict the action.

5. *Renormalisability*: terms in the action with coefficients with dimension M^{-n}, $n > 0$, generally imply that the quantum theory loses all predictive power at energies of order M. So we either exclude such terms or assume they are small [technical requirement].

7.1 Examples

Real scalar field:
$$\mathcal{L} \,=\, \tfrac{1}{2}\partial^\mu\phi\,\partial_\mu\phi - V(\phi)$$
where $V(\phi)$ doesn't contain derivatives.

Complex scalar fields:
$$\mathcal{L} \,=\, \partial^\mu\phi^\dagger\,\partial_\mu\phi - V(\phi^\dagger\phi)$$
where $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ (with ϕ_1 and ϕ_2 real).

Weyl spinor:
$$\mathcal{L} \,=\, \psi_L^\dagger\sigma^\mu i\partial_\mu\psi_L \qquad \text{or} \qquad \mathcal{L} \,=\, \psi_r^\dagger\bar{\sigma}^\mu i\partial_\mu\psi_R$$
Terms in the action of the form $m\psi_R^T\sigma_2\psi_R$, where m is the mass, lead to fermion number violation. They can appear in extensions of the Standard Model with right-handed neutrinos.

Dirac spinor:
$$\mathcal{L} \,=\, \bar{\psi}i\partial^\mu\gamma_\mu\psi - m\bar{\psi}\psi$$
Gauge field:
$$\mathcal{L} \,=\, -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} \qquad \text{where} \qquad F_{\mu\nu} \,=\, \partial_\mu A_\nu - \partial_\nu A_\mu$$

7.2 Equations of Motion

Consider the action for some generic field Φ:
$$S[\Phi] = \int d^4x\, \mathcal{L}(\Phi, \partial_\mu\Phi).$$
Under an *arbitrary* change $\Phi \mapsto \Phi + \delta\Phi$
$$\begin{aligned}
\delta S \;&=\; \int d^4x\, \delta\mathcal{L} \;=\; \int d^4x \left(\frac{\partial\mathcal{L}}{\partial\Phi}\,\delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\,\delta(\partial_\mu\Phi)\right)\\
&=\; \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\right)\right]\delta\Phi + \int d^4x\,\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\,\delta\Phi\right)
\end{aligned}$$
Noting that $\delta\left(\partial_\mu\Phi\right) = \partial_\mu\left(\delta\Phi\right)$ by linearity, we split the second term into two parts, one of which is a total derivative. The last term is a surface term: it vanishes if we insist that $\delta\Phi = 0$ on the boundary of space-time. Then Hamilton's principle $\delta S = 0$ for arbitrary $\delta\Phi$ implies
$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\right) - \frac{\partial\mathcal{L}}{\partial\Phi} \,=\, 0\,.$$
These are the Euler-Lagrange equations for the field, which we call the *classical equations of motion* (EoMs), or *classical field equations.*

For a scalar field with $\mathcal{L} \,=\, \frac{1}{2}\partial^\mu\phi\,\partial_\mu\phi - \frac{1}{2}m^2\phi^2$ we obtain
$$\partial^2\phi + m^2\phi \,=\, 0 \qquad \text{the Klein-Gordon equation for a particle of mass } m$$
For a Dirac field with $\mathcal{L} \,=\, \bar{\psi}i\partial^\mu\gamma_\mu\psi - m\bar{\psi}\psi$, varying $\bar{\psi}$ gives
$$\gamma^\mu i\partial_\mu\psi - m\psi = 0 \qquad \text{the Dirac equation}$$
Varying ψ gives an equivalent equation for $\bar{\psi}$.

The classical equations of motion are unchanged if

1. $\mathcal{L} \mapsto \mathcal{L} + \text{constant}$

2. $\mathcal{L} \mapsto \mathcal{L} + \partial_\mu K^\mu$, i.e., a term which is a total derivative. The contribution of such a term to the action can be written using the divergence theorem as an integral over a surface at infinity, which we assume vanishes.

7.3 Noether's Theorem (1918)

Consider $\Phi \mapsto \Phi + \delta\Phi$, with $\delta\Phi \neq 0$ on the boundary of space-time, and $\delta\Phi = \alpha\dfrac{\partial\Phi}{\partial\alpha}$ where α is some infinitesimal parameter.

We call this a *classical symmetry* of the theory if it leaves the classical field equations invariant. [It turns out that this is often (but not always) sufficient to ensure that the quantum theory is invariant too].

If d^4x is invariant under the symmetry, we require that $\mathcal{L}$ must be invariant up to a total derivative, i.e.,

$$\mathcal{L} \mapsto \mathcal{L} + \delta\mathcal{L} = \mathcal{L} + \alpha\,\partial_\mu K^\mu \tag{12}$$

for some K^μ. Of course, for symmetries which leave $\mathcal{L}$ invariant, we have $K^\mu = 0$.

Under variation of the fields

$$\delta\mathcal{L} \;=\; \frac{\partial\mathcal{L}}{\partial\Phi}\delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\delta(\partial_\mu\Phi) \;=\; \alpha\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\,\frac{\partial\Phi}{\partial\alpha}\right) + \alpha\underbrace{\left[\frac{\partial\mathcal{L}}{\partial\Phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\Phi}\right)\right]}_{\text{= 0 by equations of motion}}\frac{\partial\Phi}{\partial\alpha} \tag{13}$$

again using $\delta\left(\partial_\mu\Phi\right) = \partial_\mu\left(\delta\Phi\right)$. It follows from equations (12) and (13) that if we define the current density

$$J^\mu(x) \;=\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\,\frac{\partial\Phi}{\partial\alpha} - K^\mu \qquad \text{the } \textit{Noether current}$$

then this current is conserved *locally*, i.e. at every point x_μ, it satisfies

$$\partial_\mu J^\mu(x) \;=\; 0.$$

If $J_\mu = (J_0, \underline{J})$ then $\dfrac{\partial J^0}{\partial t} + \underline{\nabla}\cdot\underline{J} = 0.$

To see what this means, define the charge

$$Q(t) \equiv \int_V d^3x\; \underbrace{J^0(\underline{x},t)}_{\substack{\text{charge}\\ \text{density}}}.$$

Then

$$\frac{d}{dt}Q(t) \;=\; \int_V d^3x\,\frac{\partial J^0(\underline{x},t)}{\partial t} \;=\; -\int_V d^3x\,\nabla\cdot\underline{J} \;=\; -\int_S d\underline{S}\cdot\underline{J}\,.$$

The last term is (minus) the flux of the current leaving the surface S, so the charge is conserved.

For every classical symmetry we have a (classically) conserved charge! This important result is known as Noether's Theorem.

7.4 Conservation of Energy and Momentum

Under an infinitesimal translation $x_\mu \mapsto x_\mu + a_\mu$, by Taylor's theorem

$$\Phi(x) \;\mapsto\; \Phi(x+a) \;=\; \Phi(x) + a^\nu\partial_\nu\Phi$$

and

$$\mathcal{L}(x) \;\mapsto\; \mathcal{L}(x) + a_\mu\partial^\mu\mathcal{L} \;=\; \mathcal{L} + a^\nu\partial_\mu\left(\delta^\mu{}_\nu\mathcal{L}\right)$$

so $\dfrac{\partial\Phi}{\partial a^\nu} = \partial_\nu\Phi$, and as always $\delta\left(\partial_\mu\Phi\right) = \partial_\mu\left(\delta\Phi\right)$.

By Noether's theorem, defining

$$T^\mu{}_\nu \;\equiv\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\,\partial_\nu\Phi - \delta^\mu{}_\nu\,\mathcal{L} \qquad \text{the } \textit{energy-momentum tensor density}$$

we have *four* conserved currents $T^\mu{}_\nu$, i.e. one for each of $\nu = 0,1,2,3$, which satisfy

$$\partial_\mu T^\mu{}_\nu \;=\; 0\,.$$

The conserved charge associated with time translations is the integral over all space of the local energy density

$$H = \int_V d^3x\; \underbrace{T^0{}_0(\underline{x},t)}_{\substack{\text{Hamiltonian}\\ \text{density}}} \qquad \text{(corresponds to generator of time translations)}$$

where

$$T^0{}_0 \;=\; \frac{\partial\mathcal{L}}{\partial\dot\Phi}\,\dot\Phi - \mathcal{L}$$

is the usual Legendre transform between the Lagrangian and the Hamiltonian, but for the respective "densities".

The conserved quantity associated with spatial translations is the momentum carried by the fields

$$p_i \;=\; \int_V d^3x\; \underbrace{T^0{}_i(\underline{x},t)}_{\substack{\text{momentum}\\ \text{density}}} \qquad \text{where} \qquad T^0{}_i \;=\; \frac{\partial\mathcal{L}}{\partial\dot\Phi}\,\partial_i\Phi\,.$$

NB The momentum density $T^0{}_i(\underline{x},t)$ is distinct from the *canonical momentum* $\pi(\underline{x},t)$ introduced in the QFT course.

Local conservation of energy and momentum is a direct consequence of the homogeneity of space-time.

Similarly, conservation of angular momentum follows from the fact that $\mathcal{L}$ is a Lorentz scalar (isotropy of space-time). [The derivation is more difficult, see tutorial problem (5.4).]

7.5 Real Scalar Field

The most general Lagrangian for one scalar field $\phi(x)$ consistent with fundamental principles (1, 2, 4) is

$$\mathcal{L} = \underbrace{\tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi}_{\substack{\text{kinetic}\\ \text{term}}} - \underbrace{V(\phi)}_{\text{potential}} + \underbrace{\frac{1}{\Lambda_1^2}\partial^2\phi\,\partial^2\phi + \frac{1}{\Lambda_2^2}(\partial_\mu\phi\,\partial^\mu\phi)^2 + \frac{1}{\Lambda_3^2}(\partial_\mu\phi\,\partial_\nu\phi)^2 + \cdots}_{\text{suppress according to principles 3 and 5}}.$$

The coefficient of the kinetic term is a convention. The field has (mass) dimension 1. (If $\hbar = c = 1$ then $[M] = [L^{-1}] = [T^{-1}]$). The classical field equation is

$$\partial^2\phi + V'(\phi) = 0.$$

For renormalisable theories in 3+1 dimensions (condition 5)

$$V(\phi) = \underbrace{\tfrac{1}{2}m^2\phi^2}_{\substack{\text{“mass”}\\ \text{term}}} + \underbrace{\frac{g}{3!}\phi^3 + \frac{\lambda}{4!}\phi^4}_{\substack{\text{“interaction”}\\ \text{terms}\\ \text{(nonlinear)}}}.$$

where g and λ are called coupling constants. (See the QFT course for the technical details on renormalisability.)

For simplicity we assume further a discrete $\mathbb{Z}_2$ symmetry under $\phi \mapsto -\phi$, so $g = 0$.

Now the classical field equation is

$$\left(\partial^2 + m^2\right)\phi(x) = -\frac{\lambda}{3!}\phi^3(x).$$

When $\lambda = 0$

$$\left(\partial^2 + m^2\right)\phi(x) = 0 \qquad \textit{(Klein-Gordon equation)}$$

This can be solved by Fourier transformation: its solutions are superpositions of plane waves $e^{-ip_\mu x^\mu}$, where $p_\mu p^\mu = m^2$, so $\phi(x)$ corresponds to scalar particles of mass m.

Conserved currents:

$$T_{\mu\nu} = \partial_\mu\phi\,\partial_\nu\phi - \eta_{\mu\nu}\,\mathcal{L} = \partial_\mu\phi\,\partial_\nu\phi - \tfrac{1}{2}\eta_{\mu\nu}\,\partial_\rho\phi\,\partial^\rho\phi + \eta_{\mu\nu}\,V(\phi) \qquad \text{the } \textit{stress-energy tensor}$$

Then, by the field equations,

$$\partial^\mu T_{\mu\nu} = \left(\partial^2\phi + V'(\phi)\right)\partial_\nu\phi = 0$$

The energy density is

$$E = T_{00} = \tfrac{1}{2}\dot{\phi}^2 + \tfrac{1}{2}(\nabla\phi)^2 + V(\phi)$$

which is positive definite when $V(\phi) > 0$. In the lowest energy state or *vacuum*, E is minimised, so $\phi =$ constant and the vacuum is given by the minimum of $V(\phi)$.

In the example above, when ($g = 0$, $m^2 > 0$, $\lambda > 0$) the minimum is at $\phi_0 = 0$.

When ($g = 0$, $m^2 < 0$, $\lambda > 0$) we have $\phi_0 \neq 0$. This is *spontaneous symmetry breaking* of the Z_2 symmetry (see later).

The momentum density is

$$P_i = T_{0i} = \dot{\phi}\,\partial_i\phi.$$

It vanishes in the vacuum. [19]

The conserved “angular momenta” are

$$M_{\mu\nu} = \int d^3x\,(x_\mu\,T_{0\nu} - x_\nu\,T_{0\mu})$$

Exercise: check this, noting that $T_{\mu\nu} = T_{\nu\mu}$.

7.6 Complex Scalar Field

If $\phi(x)$ is complex, $\phi = \phi_1 + i\phi_2$, with ϕ_1 and ϕ_2 real, then

$$\mathcal{L} = \tfrac{1}{2}\partial_\mu\phi^*\,\partial^\mu\phi - V(\phi^*\phi) = \tfrac{1}{2}\partial_\mu\phi_1\,\partial^\mu\phi_1 + \tfrac{1}{2}\partial_\mu\phi_2\,\partial^\mu\phi_2 - V(\phi_1^2 + \phi_2^2).$$

We thus have two scalar particles with the same mass, coupled together through interaction terms. For $V(\phi^*\phi) = \frac{1}{2}m^2\phi^*\phi + \frac{1}{2}\lambda(\phi^*\phi)^2$ the field equations are

$$\left(\partial^2 + m^2\right)\phi = -\lambda\phi\,(\phi^*\phi)$$

etc. Now, in addition to Poincaré invariance, we have a further invariance under $\phi \mapsto e^{i\alpha}\phi$. Using Noether's theorem we see that this corresponds to a conserved current

$$j_\mu(x) = i\frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi)}\,\phi - i\frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi^*)}\,\phi^* = \frac{i}{2}\left(\phi\,\partial_\mu\phi^* - \phi^*\,\partial_\mu\phi\right),$$

and thus a charge density $j_0 = \frac{i}{2}\left(\phi\dot{\phi}^* - \phi^*\dot{\phi}\right)$.

This is our first example of an *internal symmetry*, here it's $U(1) \cong SO(2)$. Complex fields often carry such a conserved charge; later we will associate it with *electric charge*.

7.7 Actions for Spinor Fields

For two-component Weyl spinors, we choose to use four-component Dirac spinor notation with

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$$

but we project onto left- and right-handed spinors using the projection operators $P_{L,R} = \frac{1}{2}(1 \pm \gamma^5)$, with $\gamma^5 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. Then

$$\psi_L = P_L\,\psi = \begin{pmatrix} \psi_L \\ 0 \end{pmatrix} \qquad \text{and} \qquad \psi_R = P_R\,\psi = \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}$$

This can be a bit confusing because we have the same symbol ψ_L (say) for 2 and 4-component spinors on the LHS and RHS of the first equation, but it means we can use 4-component spinors for both massless (Weyl) and massive (Dirac) cases.

[19] NB the *canonical* momentum $\pi = \dot{\phi}$ is a different quantity.

For a single left-handed Weyl fermion we have the real 4-vector $\bar{\psi}_L\gamma_\mu\psi_L$ but no real scalar.[20] It follows that (in four-component notation) the only action satisfying principles 1–5 is

$$\mathcal{L}_L = i\bar{\psi}_L\gamma^\mu\partial_\mu\psi_L$$

Similarly for a right-handed Weyl fermion, again in four-component notation,

$$\mathcal{L}_R = i\bar{\psi}_R\gamma^\mu\partial_\mu\psi_R.$$

The field equations may be obtained by varying with respect to $\bar{\psi}_L$ and $\bar{\psi}_R$ respectively:

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi}_L)}\right) - \frac{\partial\mathcal{L}}{\partial\bar{\psi}_L} = 0\,, \qquad \text{and} \qquad \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi}_R)}\right) - \frac{\partial\mathcal{L}}{\partial\bar{\psi}_R} = 0\,,$$

which give

$$\gamma^\mu\partial_\mu\psi_L = 0 \qquad \gamma^\mu\partial_\mu\psi_R = 0 \qquad \textit{Weyl equations}$$

Exercise: verify that you obtain equivalent field equations when ψ_L and ψ_R are varied.

There is no Lorentz-invariant 'mass' term (i.e. no derivatives) that can be included in the action so a single Weyl fermion is automatically massless.

Weyl spinors are complex, and the actions are invariant under $\psi_L \mapsto e^{-i\alpha}\psi_L$, $\psi_R \mapsto e^{-i\alpha}\psi_R$. It follows that we have conserved *chiral currents*

$$j^\mu_L(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi_L)}\frac{\partial\psi_L}{\partial\alpha}\,,$$

and similarly for $j^\mu_R(x)$. These give

$$j^\mu_L = \bar{\psi}_L\gamma^\mu\psi_L \qquad \text{and} \qquad j^\mu_R = \bar{\psi}_R\gamma^\mu\psi_R$$

with conserved *chiral charges* (recalling $\bar{\psi}_L = \psi_L^\dagger\gamma_0$ and $\gamma_0^2 = 1$)

$$Q_L = \int d^3x\,\bar{\psi}_L\gamma_0\psi_L = \int d^3x\,\psi_L^\dagger\psi_L, \qquad \text{and} \qquad Q_R = \int d^3x\,\psi_R^\dagger\psi_R.$$

For Dirac fermions, $\bar{\psi}\psi$ is a scalar, so the most general action is

$$\mathcal{L}_D = \bar{\psi}\left(i\gamma^\mu\partial_\mu - m\right)\psi = \bar{\psi}_L i\gamma^\mu\partial_\mu\psi_L + \bar{\psi}_R i\gamma^\mu\partial_\mu\psi_R - m\left(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R\right) \qquad \text{(exercise)}$$

The field equations are simply

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \qquad \textit{Dirac}\text{ equation}$$

For plane wave solutions, since $(\gamma^\mu p_\mu - m)\,(\gamma^\mu p_\mu + m) = p^2 - m^2 = 0$, Dirac fermions are massive.

In the *chiral limit*, $m \to 0$, ψ_L and ψ_R decouple: $\mathcal{L}_D = \mathcal{L}_L + \mathcal{L}_R$, so we are left with two *independent* Weyl fermions.

[20]None of $\bar{\psi}_L\psi_L$, $\psi_L^\dagger\psi_L$, $\psi_L\psi_L$, nor similar expressions involving γ^5, are scalars.

Invariance under $\psi \mapsto e^{-i\alpha}\psi$ gives the conserved vector current

$$j_\mu = \bar{\psi}\,\gamma_\mu\,\psi$$

with conserved charge

$$Q = \int d^3x\,\bar{\psi}\gamma^0\psi = \int d^3x\,\psi^\dagger\psi = Q_L + Q_R \qquad \text{Fermion number: lepton or baryon number.}$$

However, under

$$\psi \mapsto e^{i\beta\gamma_5}\,\psi \qquad \text{and} \qquad \bar{\psi} \mapsto \bar{\psi}\,e^{i\beta\gamma_5}$$

the mass term in $\mathcal{L}_D$ is *not* invariant. Consequently, the axial current $j^\mu_5 = \bar{\psi}\gamma_5\gamma^\mu\psi$ is *not* conserved, but rather

$$\partial_\mu j^\mu_5 = -2im\bar{\psi}\gamma_5\psi.$$

See tutorial question (5.1).

We say that axial symmetry is *softly broken* by the mass term; it is *restored* as $m \to 0$, whereupon the axial charge $Q_5 \equiv Q_L - Q_R$ is conserved. ($\bar{\psi}\psi$ in the mass term $m\bar{\psi}\psi$ is a dimension 3 operator because each ψ has (mass) dimension 3/2.)

Fermions have no renormalisable self-interactions in $3+1$ dimensions, however they *can* interact with scalars: suitable interactions with a Dirac fermion are

$$g\sigma\bar{\psi}\psi + ig'\pi\bar{\psi}\gamma^5\psi \qquad \text{Yukawa interaction}$$

where $\sigma(x)$ and $\pi(x)$ are scalar fields, and g and g' are coupling constants. If $g = g'$, chiral symmetry is preserved provided that the fields transform as (tutorial question (5.5)):

$$\begin{array}{ll} \delta\psi = i\beta\gamma^5\psi & \beta \text{ infinitesimal} \\ \left.\begin{array}{r} \delta\sigma = 2\beta\pi \\ \delta\pi = -2\beta\sigma \end{array}\right\} & SO(2) \text{ transformation on } \begin{pmatrix}\sigma\\ \pi\end{pmatrix} : \sigma^2 + \pi^2 \text{ is invariant} \end{array}$$

8 Compact Groups II

In the section *Compact Groups I*, we proved two theorems (5) & (6) and stated another (7) without proof:

Theorem 5: If G is compact, L_G may be decomposed as

$$L_G \cong L_0 \oplus (L_1 \oplus \cdots \oplus L_n)$$

where L_0 is Abelian (so $L_0 \cong u(1) \oplus \cdots \oplus u(1)$) and the L_i are simple. So we need consider only semisimple (or indeed simple) compact groups.

NB, this decomposition does *not* apply to non-compact groups because it relies on the generators being hermitian. A good example is the algebra of the Poincaré group: the translation generators P_μ form an Abelian invariant subalgebra, but as we saw they don't commute with the generators $M^{\mu\nu}$ of the Lorentz group, so the Poincaré algebra isn't a direct sum of translation and Lorentz generators.

Theorem 6: If G is compact, the Killing form g_{ab} is positive definite (has non-negative eigenvalues). If L_G is also semisimple, we may choose a basis in which $g_{ab} = \delta_{ab}$ (distinctions between upper and lower indices can thus be ignored).

Theorem 7: If G is compact and semisimple, $\operatorname{tr} T_a T_b = C(R)\, \delta_{ab}$ for all irreducible representations. The positive real number $C(R)$ is "the Casimir" of the representation R.

Then we found the irreducible representations $|jm\rangle$ of the $su(2)$ algebra (see page 18).

This is the "angular momentum algebra" familiar from quantum mechanics. For $su(2) \cong so(3)$, we have $[T_i, T_j] = i\epsilon_{ijk} T_k$ and $g_{ij} = 2\delta_{ij}$, so $su(2)$ is compact and semisimple.

Define the "step operators" $T_\pm = T_1 \pm iT_2$, then

$$\begin{aligned} [T_3, T_\pm] &= \pm T_\pm \\ [T_+, T_-] &= 2T_3 \\ [T_+, T_+] &= [T_-, T_-] = 0. \end{aligned}$$

The quadratic Casimir is $T^2 = T_1^2 + T_2^2 + T_3^2 = T_- T_+ + T_3^2 + T_3$, and $[T^2, T_i] = 0$. There are no other Casimir operators (so $su(2)$ has rank 1): T^2 and T_3 form a complete commuting set.

- T^2 has eigenvalues $j(j+1)$, $2j \in \mathbb{N}$: use to label representations.
- T_3 has eigenvalues $m = j, j-1, \ldots, -j$, $2m \in \mathbb{Z}$: use to label states within representation.

8.1 The Cartan Subalgebra

Consider a compact semisimple algebra L. Choose a hermitian generator $H_1 \in L_G$, then find another hermitian generator H_2 such that $[H_1, H_2] = 0$. Continue enlarging as far as possible.

Definition 9 *The set of generators $H_i \in L$ such that $[H_i, H_j] = 0$ form a* Cartan *subalgebra $H \subset L$.*[21] *The dimension of the Cartan subalgebra (CSA) is the same as the number of independent Casimirs, i.e., the* rank *of L (no proof offered).*

Denote the dimension of L by d, the rank of L by r, and the $d-r$ generators *not* in the CSA by E_m.

1. Choose an orthonormal hermitian basis for L in which $g_{ab} = \delta_{ab}$. Write

$$\begin{array}{llll} H_i \in H & i = 1, \ldots, r & \\ E_m \in E & m = 1, \ldots, d-r & \text{(the space orthogonal to the Cartan subalgebra)} \end{array}$$

Then $[H_i, E_m] = iE_n\, c_{nim}$.[22]

The matrices A_i with $(A_i)_{nm} = ic_{nim}$ are the adjoint representation of the Cartan subalgebra: they're a set of r $(d-r)\times(d-r)$ mutually commuting hermitian matrices.

[21] Note that H can't be an *invariant* subalgebra because L is semi-simple, *i.e.* it has no Abelian invariant subalgebras.

[22] If the indices $i, j \in H$, and $m, n \in E$, then $c_{jim} = 0$ since $c_{mij} = 0$ because $[H_i, H_j] = 0$.

It follows that we can simultaneously diagonalise all the matrices A_i by a unitary transformation, to give real eigenvalues $\lambda_i^\alpha \equiv \alpha_i$ and "eigenvectors" E_α.

In words, $\lambda_i^\alpha \equiv \alpha_i$ is the α^{th} eigenvalue of the i^{th} mutually-commuting $(d-r)\times(d-r)$ generator A_i in the adjoint represention.

In this new basis, the commutation relations are "diagonal"

$$[H_i, E_\alpha] = \alpha_i\, E_\alpha, \qquad (|\alpha| \neq 0 \text{ because } E_\alpha \notin H)$$

As the notation suggests, the α_i (where $i = 1, \ldots, r$) are now regarded as the components of $d-r$ real r-dimensional vectors — called *root vectors* (or just "roots").

Because of the form of the commutation relation in this "diagonal' basis, E_α is the corresponding *step operator* (just like in quantum mechanics).

2. Take the hermitian conjugate of the diagonal commutation relations, noting that the H_i are hermitian:

$$[H_i, E_\alpha]^\dagger = -[H_i, E_\alpha^\dagger] = \alpha_i E_\alpha^\dagger.$$

So if α is a root, then $-\alpha$ is also a root. So roots come in pairs: $\pm\alpha$ with $E_{-\alpha} \equiv E_\alpha^\dagger$, and the E_α are necessarily *not* hermitian. As there are $d-r$ roots, $d-r$ must be *even.*

Example. $su(2)$: rank $r = 1$, T_3 is a Cartan subalgebra, dimension $d = 3$, the roots are just numbers, $\alpha = \pm 1$, and $T_\pm$ are step operators, $T_- = T_+^\dagger$.

3. Next we show that a reasonable basis always occurs. Consider

$$[H_i, E_\alpha E_\beta] = [H_i, E_\alpha]\, E_\beta + E_\alpha\, [H_i, E_\beta] = (\alpha_i + \beta_i)\, E_\alpha E_\beta\,.$$

Take the trace of this equation. The trace of any commutator is zero, so

$$(\alpha_i + \beta_i) \operatorname{tr}(E_\alpha E_\beta) = 0\,.$$

Thus either $\alpha = -\beta$ or $\operatorname{tr}(E_\alpha E_\beta) = 0$. We know that $\operatorname{tr}(E_\alpha E_{-\alpha}) = \operatorname{tr}\left(E_\alpha E_\alpha^\dagger\right) \neq 0$ because $\det g_{ab} \neq 0$.

Choose the following normalisation (Cartan–Weyl)

$$\operatorname{tr}(H_i\, H_j) = \delta_{ij}, \qquad \operatorname{tr}(E_\alpha\, E_{-\alpha}) = 1, \qquad \operatorname{tr}(H_i E_\alpha) = 0\,.$$

The last relation holds because

$$\alpha_j \operatorname{tr}(H_i E_\alpha) = \operatorname{tr}(H_i\, [H_j, E_\alpha]) = \operatorname{tr}([H_i, H_j]\, E_\alpha) = 0$$

where we used the cyclic property of the trace to obtain the expression containing the commutator $[H_i, H_j]$, which is zero. (We can choose j so that $\alpha_j \neq 0$.)

4. Finally, consider

$$\begin{aligned} [H_i, [E_\alpha, E_\beta]] &= -[E_\alpha, [E_\beta, H_i]] - [E_\beta, [H_i, E_\alpha]] \qquad \text{(using the Jacobi identity)} \\ &= (\alpha_i + \beta_i)\, [E_\alpha, E_\beta]\,. \end{aligned}$$

(a) If $\alpha + \beta \neq 0$, *either* $[E_\alpha, E_\beta] = 0$ *or* $E_{\alpha+\beta} \equiv [E_\alpha, E_\beta]$, i.e., $\alpha + \beta$ is a root.

(b) If $\alpha+\beta=0$, then $[E_\alpha, E_{-\alpha}]$ is in the Cartan subalgebra, so it's a linear combination of the H_i:

$$[E_\alpha, E_{-\alpha}] \;=\; x_i H_i \qquad \text{for some } x_i \;\; \text{(implicit sum over } i\text{)}.$$

Multiplying by H_j and taking the trace gives for the RHS

$$x_i \operatorname{tr}(H_i H_j) \;=\; x_i \delta_{ij} \;=\; x_j$$

while for the LHS

$$\operatorname{tr}([E_\alpha, E_{-\alpha}]\, H_j) \;=\; \operatorname{tr}([H_j, E_\alpha]\, E_{-\alpha}) \;=\; \alpha_j\,(\operatorname{tr} E_\alpha E_{-\alpha}) \;=\; \alpha_j,$$

Hence $x_j = \alpha_j$, so

$$[E_\alpha, E_{-\alpha}] \;=\; \alpha_i H_i.$$

In summary:

$$\begin{aligned}
[H_i, E_\alpha] &= \alpha_i\, E_\alpha \\
[E_\alpha, E_{-\alpha}] &= \alpha_i\, H_i \;\equiv\; \underline{\alpha}\cdot\underline{H} \\
\text{while if } \alpha+\beta \neq 0, \qquad [E_\alpha, E_\beta] &= \begin{cases} E_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root} \\ 0 & \text{otherwise} \end{cases} \\
\operatorname{tr}(H_i\, H_j) &= \delta_{ij} \\
\operatorname{tr}(E_\alpha\, E_\beta) &= \delta_{\alpha+\beta,0} \\
\operatorname{tr}(H_i\, E_\alpha) &= 0
\end{aligned}$$

8.2 Chevalley basis

For each root α we have the quantities $\underline{\alpha}\cdot\underline{H}$, E_α and $E_{-\alpha}$.

Define $\hat{\alpha}_i = 2\alpha_i/\alpha^2$, where $\alpha^2 = \alpha_i\,\alpha_i$. These are called the *coroots*.

Define also $\hat{H}^{(\alpha)} = \hat{\alpha}_i H_i = \underline{\hat{\alpha}}\cdot\underline{H}$, $\hat{E}_{\pm\alpha} = \sqrt{\frac{2}{\alpha^2}}E_{\pm\alpha}$, so now $\operatorname{tr}\left(\hat{E}_\alpha \hat{E}_{-\alpha}\right) = 2/\alpha^2$. Then

$$[\hat{H}^{(\alpha)}, \hat{E}_{\pm\alpha}] \;=\; \pm 2\hat{E}_{\pm\alpha}, \qquad [\hat{E}_\alpha, \hat{E}_{-\alpha}] \;=\; \hat{H}^{(\alpha)}.$$

Recall that in $su(2)$

$$[2T_3, T_\pm] \;=\; \pm 2T_\pm, \qquad [T_+, T_-] \;=\; 2T_3$$

So $\hat{H}^{(\alpha)}$ is like $2T_3$, and $\hat{E}_{\pm\alpha}$ is like $T_\pm$ in $su(2)$. This is wonderful, because:

(a) Every compact semisimple Lie algebra splits into these $su(2)$ subalgebras.

(b) We know all about $su(2)$.

In particular, since T_3 has eigenvalues which are integer multiples of $\frac{1}{2}$ in any irreducible representation, then $\hat{H}^{(\alpha)}$ also has integer eigenvalues, called *weights*, in any irreducible representation. This includes the adjoint representation.

An immediate consequence is that for *compact* groups the Casimirs have *discrete* eigenvalues. (These eigenvalues may be continuous for non-compact groups, e.g., the Poincaré group.)

8.3 Angles between roots

Now consider

$$[\hat{H}^{(\alpha)}, \hat{E}_\beta] \;=\; \hat{\alpha}_i [H_i, \hat{E}_\beta] \;=\; \hat{\alpha}_i\,\beta_i \hat{E}_\beta \;=\; \frac{2\underline{\alpha}\cdot\underline{\beta}}{\alpha^2}\,\hat{E}_\beta \qquad \text{i.e.} \quad [\hat{H}^{(\alpha)}, \hat{E}_\beta] \;=\; \frac{2\underline{\alpha}\cdot\underline{\beta}}{\alpha^2}\,\hat{E}_\beta$$

Thus $\hat{E}_\beta$ is a step operator for $H^{(\alpha)}$. Since $H^{(\alpha)}$ has integer eigenvalues, this implies that the "steps" $2\underline{\alpha}\cdot\underline{\beta}/\alpha^2 \in \mathbb{Z}$. By symmetry under $\alpha \leftrightarrow \beta$, also $2\underline{\alpha}\cdot\underline{\beta}/\beta^2 \in \mathbb{Z}$.

But $(\underline{\alpha}\cdot\underline{\beta})^2 \leq \alpha^2\beta^2$ by the Cauchy–Schwartz inequality[23], so

$$\left(2\underline{\alpha}\cdot\underline{\beta}/\alpha^2\right)\left(2\underline{\alpha}\cdot\underline{\beta}/\beta^2\right) \;\leq\; 4\,,$$

where each term in parentheses is also an integer, which must therefore be one of the set $\{\pm 0, \pm 1, \pm 2, \pm 3, \pm 4\}$. Hence there are only a small number of possibilities for the angle between any two roots α and β, namely

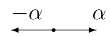
$$\cos\theta \;=\; \frac{\underline{\alpha}\cdot\underline{\beta}}{\sqrt{\alpha^2\beta^2}} \;=\; 0, \pm\tfrac{1}{2}, \pm\tfrac{1}{\sqrt{2}}, \pm\tfrac{\sqrt{3}}{2}, \pm 1, \qquad \text{(exercise)}$$

whence $\theta \in \{\pm\frac{\pi}{2}, \pm\frac{\pi}{3}, \pm\frac{2\pi}{3}, \pm\frac{\pi}{4}, \pm\frac{3\pi}{4}, \pm\frac{\pi}{6}, \pm\frac{5\pi}{6}\}$ (exercise).

Clearly, $\cos\theta = \pm 1$ gives $\alpha = \pm\beta$.

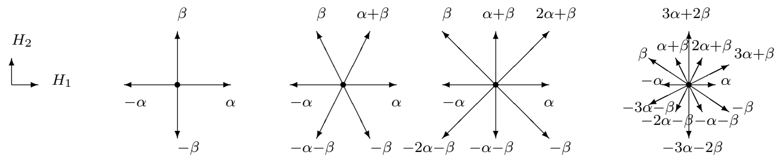
Example.

- **Rank 1** The Cartan subalgebra contains only one element H; there are two step operators E_α and $E_{-\alpha}$, with two roots $\pm\alpha$, which are just numbers (one-dimensional vectors). We represent them graphically by a root diagram:



  Since this is the *only* rank-1 compact semi-simple Lie algebra, then $su(2) \cong so(3) \cong sp(2)$, or $A_1 \cong B_1 \cong C_1$ (classical algebra notation).

- **Rank 2** There are two elements in the Cartan subalgebra, so the roots are two-dimensional vectors. In this case, the number of roots depends on the angles between them. There are 4 rank 2 semi-simple Lie algebras, distinguished by the angles between adjacent roots being $\frac{\pi}{2}$, $\frac{\pi}{3}$, $\frac{\pi}{4}$, $\frac{\pi}{6}$, respectively, which give root diagrams



$D_2 \cong su(2)\oplus su(2) \cong so(4)$ $\quad A_2 \cong su(3)$ $\quad B_2 \cong so(5) \cong C_2 \cong sp(4)$ $\quad G_2$ (exceptional group)

---

[23]To derive this, Minimise $\|\underline{\alpha} - \mu\underline{\beta}\|^2 \geq 0$ with respect to $\mu \in \mathbb{R}$.

One can classify *all* compact semi-simple algebras this way using, for example *Dynkin diagrams*. One finds

The classical algebras:

$$\begin{array}{rclll} A_n & \cong & su(n+1) & & n=1,2,\ldots \\ B_n & \cong & so(2n+1) & & n=2,3,\ldots \\ C_n & \cong & sp(2n) & & n=3,4,5,\ldots \\ D_n & \cong & so(2n) & & n=4,5,\ldots \end{array}$$

In this notation, the subscript on the name of the algebra gives its rank. Many of the classical algebras with low rank n are isomorphic:

$$\begin{array}{ll} B_1 \cong C_1 \cong A_1 & so(3) \cong sp(2) \cong su(2) \\ C_2 \cong B_2 & sp(4) \cong so(5) \\ D_2 \cong A_1 \oplus A_1 & so(4) \cong su2 \oplus su(2) \\ D_3 \cong A_3 & so(6) \cong su(4) \end{array}$$

Finally, there are 5 exceptional algebras that don't fit into the above classical algebra categories:

$$E_6, \quad E_7, \quad E_8, \quad F_4, \quad G_2$$

There are no more...

8.4 The $su(3)$ algebra

In the defining representation the generators are 3×3 traceless hermitian matrices. There are two independent traceless diagonal 3×3 hermitian matrices, which of course commute, so $su(3)$ has rank 2. We choose the *Gell-Mann matrices*:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \text{(c.f., Pauli matrices)}$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_8 = \tfrac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

Then $\mathrm{tr}\,\lambda_i\lambda_j = 2\delta_{ij}$.

Note also the close connection between the pairs λ_4, λ_5 and the Pauli matrices σ_1, σ_2. Similarly for λ_6, λ_7. The diagonal matrices λ_3 and λ_8 span the Cartan subalgebra: $[\lambda_3, \lambda_8] = 0$.

$SU(3)$ is a compact simple group with dimension 8, rank 2.

The Lie algebra is

$$[\tfrac{1}{2}\lambda_i, \tfrac{1}{2}\lambda_j] = i f_{ijk} \tfrac{1}{2}\lambda_k.$$

There are two symmetric invariant tensors: δ_{ij} and $d_{ijk} \equiv \frac{1}{4}\,\mathrm{tr}\,\lambda_i\{\lambda_j, \lambda_k\}$, and thus two Casimir operators: the quadratic and the cubic [see tutorial question (6.3) for an explicit computation of f_{ijk} and d_{ijk}.]

After some algebra:

$$\begin{array}{rcl} f_{123} & = & 1 \\ f_{147} = f_{246} = f_{257} = f_{345} = f_{516} = f_{637} & = & \frac{1}{2} \\ f_{458} = f_{678} & = & \dfrac{\sqrt{3}}{2} \end{array}$$

and

$$\begin{array}{rcl} d_{118} = d_{228} = d_{338} & = & \dfrac{1}{\sqrt{3}} \\ d_{146} = d_{157} = d_{256} = d_{344} = d_{355} & = & \frac{1}{2} \\ d_{247} = d_{366} = d_{377} & = & -\frac{1}{2} \\ d_{448} = d_{558} = d_{668} = d_{778} & = & -\dfrac{1}{2\sqrt{3}} \\ d_{888} & = & -\dfrac{1}{\sqrt{3}}. \end{array}$$

8.5 A more transparent basis

Define

$$\left.\begin{array}{rcl} I_1 & = & \frac{1}{2}\lambda_1 \\ I_2 & = & \frac{1}{2}\lambda_2 \\ I_3 & = & \frac{1}{2}\lambda_3 \end{array}\right| \quad \begin{array}{rcl} U_1 & = & \frac{1}{2}\lambda_6 \\ U_2 & = & \frac{1}{2}\lambda_7 \\ U_3 & = & \frac{1}{2}\left(-\frac{1}{2}\lambda_3 + \frac{\sqrt{3}}{2}\lambda_8\right) \end{array} \quad \left|\begin{array}{rcl} V_1 & = & \frac{1}{2}\lambda_4 \\ V_2 & = & \frac{1}{2}\lambda_5 \\ V_3 & = & \frac{1}{2}\left(\frac{1}{2}\lambda_3 + \frac{\sqrt{3}}{2}\lambda_8\right) \end{array}\right.$$

so that $\lambda_3 = (2/\sqrt{3})(U_3 + V_3)$. Then

$$\left.\begin{array}{rcll} [I_i, I_j] & = & i\epsilon_{ijk} I_k & \text{I-spin} \\ [U_i, U_j] & = & i\epsilon_{ijk} U_k & \text{U-spin} \\ [V_i, V_j] & = & i\epsilon_{ijk} V_k & \text{V-spin} \end{array}\right\} \quad \text{are 3 } su(2) \text{ subalgebras.}$$

It follows that the step operators are

$$\left.\begin{array}{rcl} I_\pm & = & I_1 \pm iI_2 \\ U_\pm & = & U_1 \pm iU_2 \\ V_\pm & = & V_1 \pm iV_2 \end{array}\right\} \begin{array}{rcl} \mathrm{tr}\, I_+ I_- & = & 1 \\ \mathrm{tr}\, U_+ U_- & = & 1 \\ \mathrm{tr}\, V_+ V_- & = & 1 \end{array}$$

with commutation relations

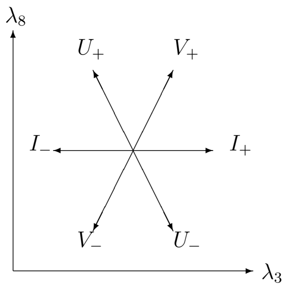
$$[I_+, I_-] = 2I_3 \qquad [U_+, U_-] = 2U_3 \qquad [V_+, V_-] = 2V_3.$$

To get the root vectors, first recall $[\hat{E}_\alpha, \hat{E}_{-\alpha}] = \hat{\alpha}_i H_i$ (implicit sum over $i = 1,2$). Taking the Cartan subalgebra $\{H_i\} = \{\lambda_3, \lambda_8\}$, and using the expressions above for I_3, U_3 and V_3 in terms of $\{\lambda_3, \lambda_8\}$, then the (positive) root vectors in the $\{\lambda_3, \lambda_8\}$ basis are

$$2(1,0) \qquad 2\left(-\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\right) \qquad 2\left(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\right)$$

Note that the third root is the sum of the first two.

Diagrammatically:



The axes refer to the eigenvalues of λ_3 and λ_8. (The angle between adjacent roots is $\pi/3$.)

The root vectors given above are consistent with the commutation relations

$$\begin{array}{rclrcl} [\lambda_3, I_\pm] &=& \pm 2 I_\pm & [\lambda_8, I_\pm] &=& 0 \\ [\lambda_3, U_\pm] &=& \mp U_\pm & [\lambda_8, U_\pm] &=& \pm\sqrt{3} U_\pm \\ [\lambda_3, V_\pm] &=& \pm V_\pm & [\lambda_8, V_\pm] &=& \pm\sqrt{3} V_\pm . \end{array}$$

The remaining commutation relations give consistency between I, U, and V spins:

$$\begin{array}{rclrclrcl} [I_\pm, U_\pm] &=& \pm V_\pm & [I_\pm, V_\pm] &=& 0 & [U_\pm, V_\pm] &=& 0 \\ [I_\pm, U_\mp] &=& 0 & [I_\pm, V_\mp] &=& \mp U_\mp & [U_\pm, V_\mp] &=& \pm I_\mp . \end{array}$$

These are all consistent with the general results for $[E_\alpha, E_\beta]$ in the summary on page 41.

8.6 Representations of su(3)

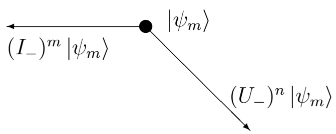
Define a *state of greatest weight* $|\psi_m\rangle$ in a representation $|\psi\rangle$ by the requirement that the state is annihilated by the raising operators

$$I_+ |\psi_m\rangle = U_+ |\psi_m\rangle = V_+ |\psi_m\rangle = 0.$$

The eigenvalues of the operators I_3, U_3 and V_3 take their maximum values in this state.

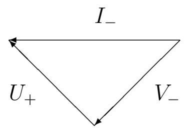
We can find new states by acting on $|\psi_m\rangle$ with the lowering operators I_-, U_-, V_-. The sequence of states

$$\begin{array}{ll} (I_-)^m |\psi_m\rangle & m = 1, 2, \ldots p \\ (U_-)^n |\psi_m\rangle & n = 1, 2, \ldots q \end{array}$$



is non-degenerate, i.e. there is only one state at each point. For example:

$$\begin{array}{rcl} I_- |\psi_m\rangle &=& [U_+, V_-] |\psi_m\rangle = (U_+ V_- - V_- U_+) |\psi_m\rangle \\ &=& U_+ V_- |\psi_m\rangle + 0. \end{array}$$



i.e. I_- and $U_+ V_-$ create the same state when acting on $|\psi_m\rangle$.

Using $[I_-, U_+] = 0$ and $[V_+, I_-] = -U_+$ (repeatedly), one finds (exercise)

$$U_+ (I_-)^m |\psi_m\rangle = V_+ (I_-)^m |\psi_m\rangle = 0$$

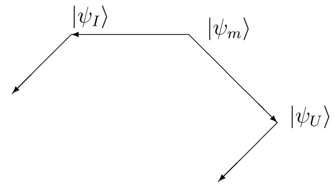
Similarly, using $[I_+, U_-] = 0$ and $[V_+, U_-] = I_+$ (repeatedly), one finds (exercise)

$$I_+ (U_-)^m |\psi_m\rangle = V_+ (U_-)^m |\psi_m\rangle = 0$$

so the new states $(I_-)^n |\psi_m\rangle$ and $(U_-)^n |\psi_m\rangle$ lie on the *boundary* of the representation. There are no states "above or to the right" of these new states.

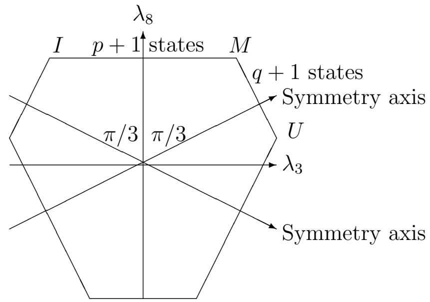
For finite dimensional representations, both sequences will stop after a finite number of steps: i.e. when we have $I_- |\psi_I\rangle = 0$ and $U_- |\psi_U\rangle = 0$.

We can then generate new sequences: $(V_-)^{m'} |\psi_I\rangle$ and $(V_-)^{n'} |\psi_U\rangle$ by operating repeatedly with (V_-) on the states we've just generated, which generates further states on the boundary of the representation, as illustrated below.



If we keep going, the boundary will eventually close (for finite dimensional representations), and we end up with a representation of the general hexagonal form shown below.

There are $p+1$ states with different values of m along the top line (MI), and $q+1$ states with different values of u along the top-right diagonal line (MU). [24]



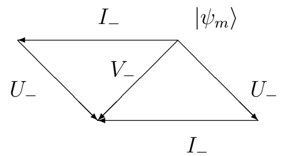
So each representation may be characterised by a pair of non-negative integers (p, q). Either may be zero, in which case the representation is *triangular*.

There are 3 symmetry axes because the boundaries MI and MU form complete I-spin and U-spin representations (respectively), as must the two parallel boundaries.

[24] The symmetry axes don't look like symmetry axes because the figure isn't drawn accurately!

We now fill in the states in the middle using e.g., $(V_-)^r |\psi_m\rangle$ etc. The next layer is doubly degenerate. For example, we can reach $V_- |\psi_m\rangle$ in *three* ways (as shown):

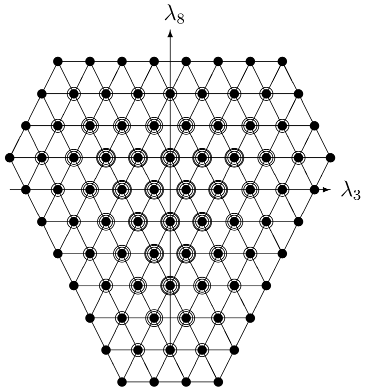
$$V_- |\psi_m\rangle\,, \qquad U_- I_- |\psi_m\rangle\,, \qquad I_- U_- |\psi_m\rangle\,.$$



But $V_- |\psi_m\rangle = U_- I_- |\psi_m\rangle - I_- U_- |\psi_m\rangle$ so only *two* of these states are linearly independent.

Continue by defining a state $|\psi'_m\rangle$ *orthogonal* to $V_- |\psi_m\rangle$, and use this to start the next (third) layer. It can be shown similarly that all states in the third layer are triply degenerate. This layer-by-layer unit increase in degeneracy continues until we reach the centre – for the case $p = q$, *or* we reach a triangular subrepresentation, at which point the degeneracy ceases to increase as we move inwards.

For example:



$(p, q) = (7, 3)$ in this (rather complex!) example. The states on the boundary are non-degenerate; the degeneracies of the states inside the boundary are indicated by concentric circles.

8.7 Examples of irreducible representations of $su(3)$

In the diagrams above, we labelled states in the representations by their eigenvalues of λ_3 and λ_8. Let's now label these states by the eigenvalues of rescaled generators. Define $I_3 = \lambda_3/2 = \mathrm{diag}(1/2,\ \text{-}1/2,\ 0)$ (just in $su(2)$) and $Y = \lambda_8/\sqrt{3} = \mathrm{diag}(1/3,\ 1/3,\ \text{-}2/3)$, both in the fundamental representation.

Note also that $Y = (2/3)(U_3 + V_3)$.

Commutators become

$$\begin{aligned}
[I_3, I_\pm] &= \pm I_\pm & [Y, I_\pm] &= 0 \\
[I_3, U_\pm] &= \mp\tfrac{1}{2} U_\pm & [Y, U_\pm] &= \pm U_\pm \\
[I_3, V_\pm] &= \pm\tfrac{1}{2} V_\pm & [Y, V_\pm] &= \pm V_\pm .
\end{aligned}$$

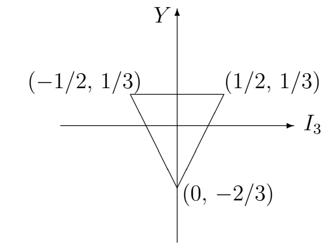
Note that the commutators of I_3 with $U_\pm$ and $V_\pm$ result in a "half step".

The smallest irreducible representations are:

1. $\underset{\sim}{1}$ Singlet $(p, q) = (0, 0)$ $\lambda_3 = \lambda_8 = 0$, and hence $I_3 = Y = 0$.

2. $\underset{\sim}{3}$ Triplet $(p, q) = (1, 0)$. This is the defining representation, so the states are labelled by the eigenvalues of the (Gell Mann) matrices $\lambda_3/2$ and $\lambda_8/\sqrt{3}$.

   There are 3 states, with $(I_3,\, Y) = (1/2,\, 1/3),\ (-1/2,\, 1/3),\ (0,\, -2/3)$, i.e. an I-spin doublet with $Y = 1/3$, and an I-spin singlet with $Y = -2/3$.

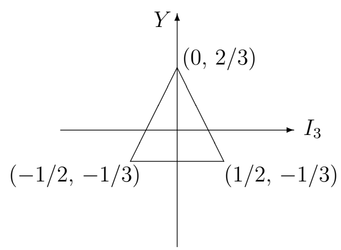
   Equivalently, we may regard the 3 states as a U-spin doublet with $U_3 = \pm 1/2$ and a singlet with $U_3 = 0$, or as a V-spin doublet with $V_3 = \pm 1/2$ and a singlet with $V_3 = 0$.

   Note that any 2 of the 4 mutually commuting operators I_3, U_3, V_3 and Y are linearly independent. Clearly, since $Y = \lambda_3/\sqrt{3} = (2/3)(U_3 + V_3)$, then if we know U_3 and V_3, we know Y! This will be useful below.

3. $\underset{\sim}{3}^*$ Anti-triplet $(p, q) = (0, 1)$

   Note under complex conjugation $\lambda_3 \mapsto -\lambda_3, \lambda_8 \mapsto -\lambda_8$, because infinitesimal group elements $1 + i\underset{\sim}{\alpha} \cdot \underset{\sim}{\lambda}/2 \mapsto 1 - i\underset{\sim}{\alpha} \cdot \underset{\sim}{\lambda}/2$.

   So the 3 states have $(I_3,\, Y) = (1/2,\, -1/3),\ (-1/2,\, -1/3),\ (0,\, 2/3)$, i.e. an I-spin doublet with $Y = -1/3$ and an I-spin singlet with $Y = 2/3$.

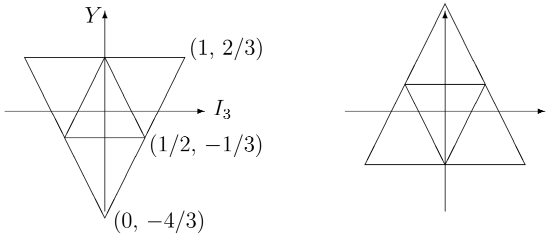
4. $\underline{6}$ sextet $(p,q)=(2,0)$ and anti-sextet $\underline{6}^*$ $(p,q)=(0,2)$.

Since $q=0$ for the $\underline{6}$, and $p=0$ for the $\underline{6}^*$, they are triangular representations.

The $\underline{6}$ contains an I-spin triplet ($I_3 = 1,\ 0,\ -1$) with $Y = 2/3$ (top row of diagram on the left), an I-spin doublet $I_3 = (1/2,\ -1/2)$ with $Y = -1/3$ (middle row), and an I-spin singlet $I_3 = 0$ with $Y = -4/3$ (apex at bottom).

The Y values are perhaps most easily obtained by first deducing the U_3 and V_3 values, and then using $Y = (2/3)(U_3 + V_3)$.

Exercise: complete the labelling of the $\underline{6}$ with the (I_3, Y) labels of the other 3 states. Similarly, label the states in the $\underline{6}^*$.

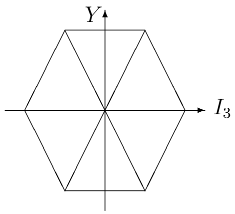


5. $\underline{8}$ The *octet* representation. $(p,q)=(1,1)$. This is the adjoint representation.

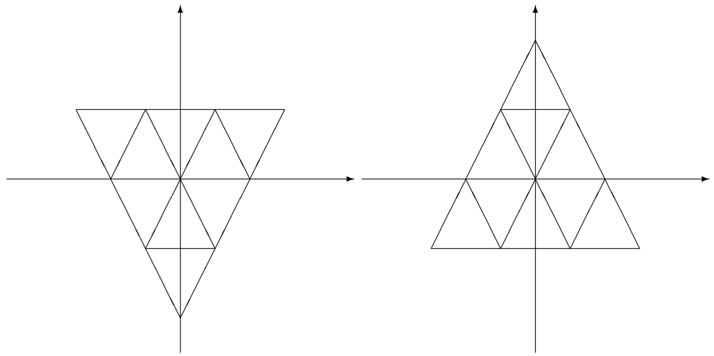
Note that there is only one diagram because the adjoint representation is real, $\underline{8} = \underline{8}^*$. The diagram is invariant under $\lambda_3 \mapsto -\lambda_3$ *and* $\lambda_8 \mapsto -\lambda_8$.

There are two I-spin doublets $I_3 = (1/2,\ -1/2)$ with $Y = 1$ and $Y = -1$ (top and bottom rows), one I-spin triplet $I_3 = (1,\ 0,\ -1)$ with $Y = 0$ (middle row), and one I-spin singlet state $(I_3,\ Y) = (0,\ 0)$ at the origin, i.e. there are two states at the origin (with the same I_3 and Y). Note that the I-spin singlet is not an $SU(3)$ singlet.

Exercise: deduce the values of Y and label all the states in the $\underline{8}$.



6. $\underline{10}$ The decuplet representation $(p,q)=(3,0)$, $\underline{10}^*$, $(p,q)=(0,3)$.



Exercise: label all the states in the $\underline{10}$.

(Hints: starting from the top row, argue that the rows have $Y = 1, 0, -1, -2$, respectively. The states in any given row have the usual I_3 assignments.)

For later use, we note the relation $Y = 2(I-1)$ for states in the decuplet.

8.8 Combining representations

Examples:

$$\left.\begin{aligned} \underline{3}\otimes\underline{3}^* &= \underline{8}\oplus\underline{1} \\ \underline{3}\otimes\underline{3} &= \underline{6}\oplus\underline{3}^* \\ \underline{3}\otimes\underline{3}\otimes\underline{3} &= \underline{10}\oplus\underline{8}_S\oplus\underline{8}_A\oplus\underline{1} \end{aligned}\right\} \quad \textit{Young tableaux} \text{ for complicated cases.}$$

See tutorial questions (6.4) and (6.5).

For combining more complicated representations, we often use *Young Tableaux*, which unfortunately we don't have time to do here.

9 Internal Symmetry

An *internal symmetry* is a symmetry G whose elements $g \in G$ commute with the elements $U(a,\Lambda)$ of the Poincaré group, i.e.

$$g\,U(a,\Lambda) \;=\; U(a,\Lambda)\,g,$$

or in terms of the generators T_a of L_G, and the translation and Lorentz generators P_μ and $M_{\mu\nu}$,

$$[T_a, P_\mu] \;=\; [T_a, M_{\mu\nu}] \;=\; 0.$$

Theorem 8 (Coleman–Mandula (1967)) *Under a general set of assumptions, the symmetry group of any nontrivial theory is the direct product of the Poincaré group and a* compact *internal symmetry group G (no proof offered).*

In other words, due to compactness the internal quantum numbers must be discrete.

Consider a *multiplet* of states $\{g\,|\Phi\rangle : g \in G\}$. If $P_\mu\,|\Phi\rangle = p_\mu\,|\Phi\rangle$, then[25]

$$P_\mu g\,|\Phi\rangle \;=\; gP_\mu\,|\Phi\rangle \;=\; p_\mu g\,|\Phi\rangle \qquad \text{etc.}$$

It follows that the space-time properties of $|\Phi\rangle$ and $g\,|\Phi\rangle$ are the same: in particular all the states $g\,|\Phi\rangle$ have the same Casimirs, i.e., they have the same mass and spin.

If the states $|r\rangle$ are created by a field Φ_r, i.e., $|r\rangle = \Phi_r\,|0\rangle$, then if

$$\Phi_r \mapsto g\Phi_r g^{-1} \;=\; D_{rs}(g)\Phi_s$$

is a symmetry of the Lagrangian of the theory, and *if the vacuum is invariant under G*, i.e., $g\,|0\rangle = |0\rangle$, then under G

$$|r\rangle \mapsto g\,|r\rangle \;=\; g\Phi_r g^{-1} g\,|0\rangle \;=\; D_{rs}\,(g)\,\Phi_s\,|0\rangle \;=\; D_{rs}(g)\,|s\rangle\,,$$

i.e., the states sit in the same representation as the fields.

Since we generally consider unitary representations D_{rs}, the normalisation $\langle r|s\rangle \propto \delta_{rs}$ is unchanged by the group transformation. [For real fields orthogonal (real) representations will do.]

9.1 Noether's Theorem for internal symmetries

For an *internal* symmetry the Lagrangian $\mathcal{L}(\Phi_r, \partial_\mu\Phi_r)$ is conserved under the infinitesimal transformation

$$\Phi_r \;\mapsto\; \Phi_r + \delta\Phi_r, = \;\Phi_r + i\alpha_a\,T^a_{rs}\,\Phi_s, \qquad (\text{with } T^{a\dagger} \;=\; T^a)$$

and thus the Noether currents

$$j^a_\mu \;=\; \frac{\partial\mathcal{L}}{\partial(\partial^\mu\Phi_r)}\,(iT^a_{rs})\,\Phi_s$$

[25] Recall that in our notation the generators P_μ have eigenvalues p_μ.

are conserved: $\partial^\mu j^a_\mu = 0$. The corresponding conserved charges are

$$Q^a \;=\; \int d^3x\, j^a_0(\underline{x}, t).$$

In the quantum theory, the charges may be used as a basis for the Lie algebra $\mathcal{L}_G$, i.e. $[T_a, T_b] = ic_{abc}\,T_c$, so that

$$[Q_a, Q_b] \;=\; ic_{abc}\,Q_c$$

The proof (omitted here, but see tutorial question (7.1)) uses the equal-time canonical commutation relations of the quantised fields.

For infinitesimal α, where $g \;=\; 1 + i\alpha_a Q_a$, whence

$$[Q^a, \Phi_r] \;=\; T^a_{rs}\Phi_s.$$

Again, the proof uses canonical commutation relations (tutorial question (7.1))

It follows that *if the vacuum is invariant* under G, i.e. $g\,|0\rangle \;=\; (1 + i\alpha_a Q_a)\,|0\rangle \;=\; |0\rangle$, and since this holds $\forall\,\alpha_a$, then

$$Q^a\,|0\rangle \;=\; 0 \qquad \text{i.e., the vacuum has zero charge.}$$

Furthermore, since $Q^a\,|0\rangle$ and hence $\langle 0|\,Q^a = 0$, then

$$0 \;=\; \langle 0|\,[Q^a, \Phi_r]\,|0\rangle \;=\; T^a_{rs}\,\langle 0|\,\Phi_s\,|0\rangle \quad \forall a,$$

whence $\langle 0|\,\Phi_s\,|0\rangle = 0$, i.e., the field has zero *vacuum expectation value* (vev).

Because the vacuum is isotropic, this will always be true for fields which transform non-trivially under the Lorentz group, *i.e.* fields with intrinsic spin such as spinor or vector fields, but it may not be true for *scalar* fields (see later).

9.2 Isospin (Heisenberg/Kemmer, 1938)

Strongly interacting particles (hadrons) evidently come in $SU(2)$ multiplets: with the same spin, J, parity, P, and *approximately* the same mass. For example

$$\text{Mesons} \qquad \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix} \qquad \begin{pmatrix} \rho^+ \\ \rho^0 \\ \rho^- \end{pmatrix}$$

$$I = 1,\; J^P = 0^- \qquad I = 1,\; J^P = 1^-$$

$$\approx 140 \text{ MeV} \qquad \approx 770 \text{ MeV}$$

$$\text{Baryons} \qquad \begin{pmatrix} p \\ n \end{pmatrix} \qquad \begin{pmatrix} \Delta^{++} \\ \Delta^+ \\ \Delta^0 \\ \Delta^- \end{pmatrix}$$

$$I = \tfrac{1}{2},\; J^P = \tfrac{1}{2}^+ \qquad I = \tfrac{3}{2},\; J^P = \tfrac{3}{2}^+$$

$$\approx 940 \text{ MeV} \qquad \approx 1230 \text{ MeV}$$

For these multiplets, there is a relation between the electric charge Q and the *baryon number* (defining $B = 1$ for baryons and $B = 0$ for mesons): $Q = I_3 + \frac{1}{2}B$.

Experimentally, Q, B, I, and I_3, are conserved separately in strong decays and strong scattering processes.

States in the same multiplet are related by $SU(2)$ shift operators:

$$I_\pm \,|I, I_3\rangle = \sqrt{I(I+1) - I_3(I_3 \pm 1)}\,|I, I_3 \pm 1\rangle\,.$$

For example

$$I_-\,|p\rangle \;=\; |n\rangle\,, \quad I_+\,|n\rangle \;=\; |p\rangle\,, \quad \text{while} \quad I_-\,\left|\pi^+\right\rangle \;=\; \sqrt{2}\,\left|\pi^0\right\rangle, \quad I_-\,\left|\pi^0\right\rangle \;=\; \sqrt{2}\,\left|\pi^-\right\rangle \quad \text{etc.}$$

We can construct combined (multiparticle) states (e.g., πN states) by combining isospins just like angular momenta. Later, we will construct a Lagrangian for mesons and baryons.

9.3 The Eightfold Way (Gell-Mann, Ne'eman, 1961)

Experimentally, there are more isospin multiplets. These are assigned a new empirical quantum number *strangeness*, which is observed experimentally to be *conserved in strong interactions*.

Mesons $\begin{pmatrix} K^+ \\ K^0 \end{pmatrix}$ $\quad$ $\begin{pmatrix} \bar{K}^0 \\ K^- \end{pmatrix}$

$I = \frac{1}{2}$, $J^P = 0^-$, $S = +1$ $\quad$ $S = -1$

≈ 495 MeV

Baryons $\begin{pmatrix} \Sigma^+ \\ \Sigma^0 \\ \Sigma^- \end{pmatrix}$ $\quad$ $\begin{pmatrix} \Xi^0 \\ \Xi^- \end{pmatrix}$

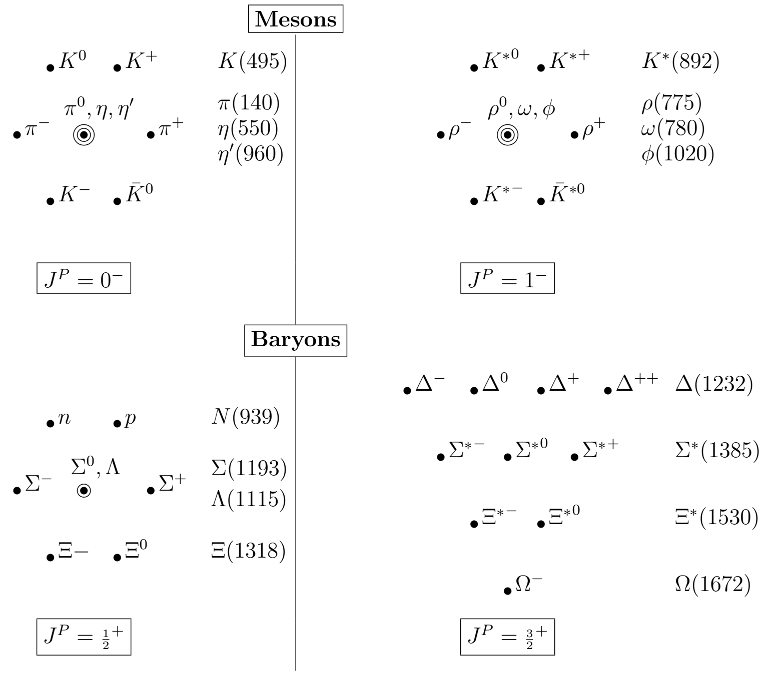
$I = 1$, $J^P = \frac{1}{2}^+$, $S = -1$ $\quad$ $I = \frac{1}{2}$, $S = -2$

≈ 1190 MeV $\quad$ ≈ 1320 MeV

We can assemble these into representations of $SU(3)$:

Mesons:	8	$J^P = 0^-$	Pseudoscalar octet: π, K, η, η'	(η' is an $su(3)$ singlet)
	8	$J^P = 1^-$	Vector Octet: K^*, ρ, ω, ϕ	(ϕ is an $su(3)$ singlet)
Baryons:	8	$J^P = \frac{1}{2}^+$	Baryon octet: p, n, Σ, Ξ, Λ	(no $su(3)$ singlet)
	10	$J^P = \frac{3}{2}^+$	Decuplet: Δ, Σ^*, Ξ^*, Ω	

The η' and ϕ states are $SU(3)$ singlets, which for convenience we place at the centre of the meson octets alongside the two $SU(3)$ octet states that also have $I_3 = 0$. There is no $SU(3)$ singlet baryon, which can be understood in the quark model.



Clearly, the masses of the states in any given $SU(3)$ multiplet are similar, but they're not equal! We refer to this as *softly broken* $SU(3)$ symmetry because as we shall see, the symmetry is broken by disparate mass terms in the Lagrangian.

The $SU(3)$ singlet mesons have significantly higher masses than those in the octets.

Electric charge: All particles in the same U-spin multiplet have the same charge Q, so the charge generator Q must satisfy $[Q, U_\pm] = [Q, U_3] = 0$. Writing $Q = aI_3 + bY$, and using the commutators at the top of page 48, we find $b = a/2$, so $Q = a(I_3 + \frac{1}{2}Y)$. Comparing with the $su(3)$ multiplets, we find $a = 1$, so

$$Q = I_3 + \tfrac{1}{2}Y \qquad \text{with } Y \equiv S + B \qquad \text{[Gell-Mann (1956)–Nishijima (1953)].}$$

Since we *chose* $B = 1$ for baryons, this gives S uniquely.

In terms of Gell-Mann matrices $I_3 = \frac{1}{2}\lambda_3$, $Y = \frac{1}{\sqrt{3}}\lambda_8$, so the charge generator in the fundamental representation is (exercise)

$$Q = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix}$$

9.4 The Quark Model (Gell-Mann, Zweig (1964))

$SU(3)$ states can be combined just as for $SU(2)$.

We can build up all representations of *flavour* $SU(3)$ from the fundamental representation. We postulate a multiplet of *quark* states

$$|q_r\rangle \;=\; \begin{pmatrix} |u\rangle \\ |d\rangle \\ |s\rangle \end{pmatrix}$$

transforming as the $\underline{3}$ (fundamental representation) of $SU(3)$. Then the *antiquark* state

$$|\bar{q}_r\rangle \;=\; \begin{pmatrix} |\bar{u}\rangle \\ |\bar{d}\rangle \\ |\bar{s}\rangle \end{pmatrix}$$

transforms as $\underline{3}^*$.

If quarks have spin $\frac{1}{2}$ and non-zero mass (so correspond to Dirac spinors) then mesons are $|q\bar{q}\rangle$ states with spin 0 and 1: $\underline{3}\otimes\underline{3}^* = \underline{8}\oplus\underline{1}$

Baryons are $|qqq\rangle$ states with spin $\frac{1}{2}$ and $\frac{3}{2}$: $\underline{3}\otimes\underline{3}\otimes\underline{3} = \underline{10}\oplus\underline{8}\oplus\underline{8}\oplus\underline{1}$. (See tutorial question (6.5).)

Since quarks transform as the $\underline{3}$, the quark model *predicts* that quarks have *fractional* electric charges: $Q_u = \frac{2}{3}$, $Q_d = -\frac{1}{3}$, $Q_s = -\frac{1}{3}$ (in units of the charge on the proton).

We can write down a Lagrangian (excluding interactions) for the quark fields q_r, $\bar{q}_r$:

$$\mathcal{L} \;=\; \bar{q}_r\,\delta_{rs}\,\gamma^\mu\, i\partial_\mu\, q_s - \bar{q}_r\, M_{rs}\, q_s$$

where r and s are *flavour* indices $r = 1,2,3$ (or $r = u,d,s$), and M_{rs} are the $3\times 3 = 9$ elements of a 3×3 hermitian mass matrix.

If $SU(3)$ symmetry is *exact*, then $M_{rs} = m\,\delta_{rs}$, i.e., all quarks have the *same mass*. Thus $\mathcal{L}$ is invariant under $q\mapsto Uq$, $\bar{q}\mapsto\bar{q}\,U^\dagger$, where $U\in SU(3)$ (i.e. U is a 3×3 unitary matrix: $U^\dagger U = 1$, $\det U = 1$).

If the strange (s) quark (mass $m' = m+\delta m$) is heavier than the up (u) and down (d) quarks (mass m), we may parameterise

$$M_{rs} \;=\; \left(m + \tfrac{1}{3}\delta m\right)\delta_{rs} - \delta m\tfrac{1}{\sqrt{3}}\,\lambda^8_{rs} \;=\; \mathrm{diag}(m,m,m+\delta m)$$

We say the $SU(3)$ *flavour* symmetry is *softly broken* by a term in the Lagrangian, which is proportional to the 8th component of the octet of operators $\bar{\psi}\lambda^a\psi$. This breaks the symmetry *explicitly*.

Flavour symmetry is *restored* if the *symmetry-breaking term* proportional to $\delta m\bar{\psi}\lambda^8\psi$ goes to zero.

9.5 The Gell-Mann–Okubo mass formula

Let us assume that the mass splittings in a given $SU(3)$ multiplet are due to the 8th component of an octet: in first order perturbation theory the mass splitting is given by the matrix elements of

$$\frac{1}{\sqrt{3}}\,\delta m_1\,\lambda_8 \;+\; \frac{1}{4}\,\delta m_2\, d_{8ab}\,\lambda_a\,\lambda_b$$

since λ_8 and $d_{8ab}\lambda_a\lambda_b$ are the only two independent octet 8th-component operators (because $f_{8bc}\lambda_b\lambda_c \propto \lambda_8$), and δm_1 and δm_2 are free parameters. Using d_{8ab} from tutorial (6.3),

$$d_{8ab}\lambda_a\lambda_b \;=\; -\frac{1}{2\sqrt{3}}\underbrace{\sum_{a=1}^{8}\lambda_a\lambda_a}_{\substack{\text{Casimir for}\\ \text{a given}\\ \text{multiplet:}\\ \text{absorb into}\\ m_0}} + \frac{3}{2\sqrt{3}}\underbrace{\sum_{i=1}^{3}\lambda_i\lambda_i}_{=4\hat{I}^2} - \frac{1}{2\sqrt{3}}\underbrace{\lambda_8^2}_{=3Y^2} \qquad \text{(exercise)}\,,$$

where $d_{811} = d_{822} = d_{833} = \frac{1}{\sqrt{3}}$, $d_{844} = d_{855} = d_{866} = d_{877} = -\frac{1}{2\sqrt{3}}$, $d_{888} = -\frac{1}{\sqrt{3}}$, and m_0 is the unperturbed mass. It follows that

$$m \;=\; m_0 + \delta m_1 Y + \delta m_2\left(I(I+1) - \tfrac{1}{4}Y^2\right) \qquad \text{Gell-Mann–Okubo mass formula}$$

where we absorbed a factor of $3/(2\sqrt{3})$ into δm_2. Applying this to the baryon octet we have 4 masses and 3 unknowns, so we get 1 prediction (or more accurately a *post*-diction):

$$\tfrac{1}{2}(m_N + m_\Xi) \;=\; \tfrac{3}{4}m_\Lambda + \tfrac{1}{4}m_\Sigma \qquad \text{(exercise)},$$

which is accurate to a few MeV.

Applying the Gell-Mann–Okubo formula to the decuplet, and noting that $(Y = 2(I-1))$, the quadratic terms in I and Y cancel, so the δm_2 term is proportional to $Y+$constant, and we end up with an expression of the form $m = m_0 + \delta m\,Y$. This is known as the *equal spacing rule*. From the masses of the Δ, Σ and Ξ, we get $\delta m \approx 145$MeV. This gave a *prediction* for the Ω^- with a mass of about 1675MeV. The Ω^- was discovered shortly afterwards with the expected mass (to within a percent) - a triumph for softly broken $SU(3)$!

9.6 Colour

Consider the Δ^{++} state in the decuplet: this is a $|uuu\rangle$ state consisting of three identical fermions. It has spin $\frac{3}{2}^+$ so all spins are up (or down), which is a symmetric state. This means that the overall wavefunction is totally symmetric under exchange of quarks, which contradicts the Pauli principle since quarks are fermions.

To circumvent this problem we introduce a new exact internal symmetry, *colour* $SU(3)$.[26] We can then construct meson and baryon fields from quark fields with flavour $(r = 1,2,3)$

[26] That the $SU(N)$ groups for colour and flavour both have $N = 3$ is coincidental. If we add a fourth *charm* quark, the flavour symmetry becomes $SU(4)$, etc. This isn't a very good symmetry because the charm mass is *much* heavier than the up, down and strange masses. We shall restrict our discussion to the latter three.

and colour indices ($a = 1, 2, 3$):

$$\left.\begin{array}{ll} \text{Mesons} & \bar{q}^a_r\, q^b_s\, \delta_{ab} \\ \text{Baryons} & q^a_r\, q^b_s\, q^c_t\, \epsilon_{abc} \end{array}\right\} \text{ Recall that } \delta_{ab} \text{ and } \epsilon_{abc} \text{ are } SU(3) \text{ invariant tensors}$$

where (r, s, t) are *flavour* indices (u, d, s), and a, b, c are *colour* indices. Since ϵ_{abc} is totally antisymmetric, the Δ^{++} baryon is antisymmetric overall. This allows us to introduce the theory of strong interactions between quarks: physical states are *singlets* under colour $SU(3)$. No physical states have non-trivial colour or fractional electric charge – in agreement with experiment.

9.7 Spontaneous Symmetry Breaking

In theories with *scalar* fields, symmetries can break *spontaneously*; the field acquires a non-zero *vacuum expectation value* (vev).

9.7.1 Abelian

$SO(2)$: Introduce two real scalar fields in a doublet $\underline{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$ and a Lagrangian

$$\mathcal{L} \;=\; \partial_\mu \underline{\phi} \cdot \partial^\mu \underline{\phi} - \frac{\lambda}{4!}\left(\underline{\phi}\cdot\underline{\phi} - c^2\right)^2,$$

so $m^2 = -\frac{\lambda}{6}c^2 < 0$ if $\lambda > 0$. $\mathcal{L}$ is invariant under $\underline{\phi} \mapsto D\underline{\phi}$ where $D^T D = 1$, so $D \in SO(2)$.

The minimum of $V(\underline{\phi})$ occurs on a *circle*: $\underline{\phi}\cdot\underline{\phi} \;=\; \phi_1^2 + \phi_2^2 \;=\; c^2$. The possible vacua $\underline{\phi} \;=\; c\begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$ are related by $SO(2)$ transformations, so all the vacua are equivalent.

Without loss of generality, we choose to expand about $\underline{\phi} = \begin{pmatrix} c \\ 0 \end{pmatrix}$ (for simplicity). Now write $\underline{\phi} = \begin{pmatrix} c+\chi \\ \pi \end{pmatrix}$, so the fields χ and π both have zero vev. Then

$$\mathcal{L} \;=\; \tfrac12\partial_\mu\chi\partial^\mu\chi + \tfrac12\partial_\mu\pi\partial^\mu\pi - \underbrace{\frac{\lambda}{6}c^2\chi^2}_{\substack{\text{mass}\\\text{term}}} - \underbrace{\frac{\lambda}{6}\chi(\chi^2+\pi^2)}_{\substack{\text{cubic}\\\text{interaction}}} - \underbrace{\frac{\lambda}{4!}\left(\pi^2+\chi^2\right)^2}_{\substack{\text{quartic}\\\text{interaction}}}.$$

We thus have (within perturbation theory) states with masses $m_\chi^2 = \frac{\lambda}{3}c^2$ and $m_\pi^2 = 0$. The π field is massless since there is no term of the form *constant*$\times\pi^2$ in $\mathcal{L}$. It's *massless* because the potential has a *flat direction* due to the equivalent vacua, which lie on the circle. In terms of the fields χ and π, the interaction terms in $\mathcal{L}$ no longer respect the $SO(2)$ symmetry: indeed the Noether charge is no longer conserved, and $Q\,|0\rangle \neq 0$ because $\langle 0|\,\underline{\phi}\,|0\rangle \neq 0$ implies $Q\,|0\rangle \neq 0$. (See previous section on Noether's Theorem.)

9.7.2 Non-Abelian

Consider a theory with n real scalar fields and Lagrangian

$$\mathcal{L} = \tfrac12\partial_\mu\underline{\phi}\cdot\partial^\mu\underline{\phi} - V(\underline{\phi}).$$

Suppose $\mathcal{L}$ is invariant under a group

$$G:\; V(\underline{\phi}) \;=\; V(g\,\underline{\phi}\,g^{-1})$$

where $g\,\phi_r\,g^{-1} = D(g)_{rs}\,\phi_s$, with $D^T D = 1$. Both $\frac12\partial_\mu\underline{\phi}\cdot\partial^\mu\underline{\phi}$ and V are separately invariant.

Assume that $V(\underline{\phi})$ is minimised on a set $\mathcal{V} = \{\underline{\phi}_0 : V(\underline{\phi}_0) = 0\}$ (i.e., vacuum energy zero), and assume further that $\mathcal{V}$ is nontrivial ($\mathcal{V} \neq \{\underline{0}\}$), then for every $\underline{\phi}_0$, $\underline{\phi}'_0 \in \mathcal{V}$ we can find a g such that $\underline{\phi}'_0 = D(g)\,\underline{\phi}_0$ (transitivity).

Define the *little group* or *stabiliser* $H \subset G$ as the group of transformations leaving $\underline{\phi}_0$ invariant

$$H \;=\; \{h \in G : \underline{\phi}_0 \;=\; D(h)\,\underline{\phi}_0,\;\; \underline{\phi}_0 \in \mathcal{V}\}.$$

This definition is independent of the choice of $\underline{\phi}_0$: if $\underline{\phi}_0 \;=\; D(h)\,\underline{\phi}_0$ and $\underline{\phi}'_0 \;=\; D(g)\,\underline{\phi}_0$ then $\underline{\phi}'_0 \;=\; D(g)\,D(h)\,D(g^{-1})\,\underline{\phi}'_0 \;=\; D(ghg^{-1})\,\underline{\phi}'_0$, so H and H' are isomorphic (conjugate) subgroups of G. [27]

Now write the fields $\underline{\phi}$ in terms of new fields $\underline{\chi}$ with zero vev, i.e.

$$\underline{\phi} \;=\; \underline{\phi}_0 + \underline{\chi}, \quad \text{so} \quad \underline{\phi}_0 \;=\; \langle 0|\,\underline{\phi}\,|0\rangle \;\neq\; 0 \quad \text{and} \quad \underline{\chi}_0 \;=\; \langle 0|\,\underline{\chi}\,|0\rangle \;=\; 0\,.$$

The Lagrangian is then

$$\mathcal{L} \;=\; \tfrac12\partial_\mu\,\underline{\chi}\cdot\partial^\mu\,\underline{\chi} - U(\underline{\chi}) \qquad \text{where} \quad U\left(\underline{\chi}\right) \;=\; V(\underline{\phi}_0 + \underline{\chi})\,.$$

Under $h \in H$, and using $D(h)\,\underline{\phi}_0 = \underline{\phi}_0$, we have

$$U\left(D(h)\,\underline{\chi}\right) \;=\; V\left(\underline{\phi}_0 + D(h)\,\underline{\chi}\right) \;=\; V\left(D(h)\,(\underline{\phi}_0 + \underline{\chi})\right) \;=\; U(\underline{\chi}),$$

so H is still an explicit unbroken symmetry of the theory. We say G is *spontaneously broken* to H.

Theorem 9 (Goldstone (1961)) *When G is spontaneously broken to H there are* $\dim G - \dim H$ *massless states called* Goldstone bosons [28]

Proof. [Classical]. Write the potential as

$$U(\underline{\chi}) \;=\; \tfrac12 M_{rs}\chi_r\chi_s \;+\; \text{higher order polynomials in } \chi$$

where the *mass matrix* is $\qquad M_{rs} \;=\; \left.\dfrac{\partial^2 V}{\partial\phi_r\,\partial\phi_s}\right|_{\underline{\phi}=\underline{\phi}_0}$

Now $\qquad V(D_{rs}\,\phi_s) \;=\; V(\phi_r)$

[27] In fact $\mathcal{V}$ is the coset space G/H, e.g., $S^2 = SO(3)/SO(2)$.

[28] There could also be states that are massless for other reasons.

Writing $D = 1 + i\alpha^a T^a$ and using the invariance of V under G gives

$$V(\phi_r + i\alpha^a \, T^a_{rs} \, \phi_s) \; = \; V(\phi_r)$$

Taylor expanding the LHS about ϕ_r implies

$$\frac{\partial V}{\partial \phi_r} \, T^a_{rt} \, \phi_t \; = \; 0 \quad \forall \, \underline{\phi}$$

Differentiate this equation with respect to ϕ_s:

$$\frac{\partial^2 V}{\partial \phi_r \, \partial \phi_s} \, T^a_{rt} \, \phi_t + \frac{\partial V}{\partial \phi_r} \, T^a_{rs} \; = \; 0.$$

The minimum of V is at $\underline{\phi} = \underline{\phi}_0$ where

$$\left. \frac{\partial V}{\partial \phi} \right|_{\underline{\phi} = \underline{\phi}_0} = 0, \quad \text{so from the equation above:} \quad M_{rs} \, T^a_{rt} \, \phi^0_t \; = \; 0$$

Thus $T^a_{rt} \, \phi^0_t$ is a zero eigenvector of M_{rs}. But not all of these eigenvectors are linearly independent. If $h \in H$ then

$$D(h) \, \underline{\phi}_0 \; = \; \underline{\phi}_0 \; = \; \underbrace{(1 + i\epsilon^a \, T^a)}_{D(h)} \, \underline{\phi}_0, \quad \text{so} \quad \epsilon^a \, T^a_{rs} \, \phi^0_s \; = \; 0 \quad \Rightarrow \quad T^a_{rs} \, \phi^0_s \; = \; 0 \quad (\text{as } \epsilon^a \text{ is arbitrary})$$

There are clearly exactly $d_H = \dim H$ such relations, so there are $d_G - d_H$ zero eigenvalues of M_{rs}. ■

Note: Goldstone's theorem also holds in the full quantum theory (no proof offered).

Example. $G = SO(3)$, $\underline{\phi}_0 = (v, 0, 0)$, $H = SO(2)$: there are two Goldstone bosons, since $3 - 1 = 2$.
