UNIVERSITY OF CAMBRIDGE

PART III OF THE MATHEMATICAL TRIPOS

Symmetries, Fields and Particles

Michaelmas 2014, Prof. N. Manton

Notes by:
William I. JAY

Edited and updated by:
Nicholas S. MANTON

Diagrams by:
Ben NACHMAN

Last updated on September 23, 2014

Preface

William Jay typeset these notes from the Cambridge Mathematics Part III course Symmetries, Fields and Particles in Spring 2013. Some material amplifies or rephrases the lectures.
N. Manton edited and updated these notes in Autumn 2013, with further minor changes in Autumn 2014. If you find errors, please contact N.S.Manton@damtp.cam.ac.uk.

Thanks to Ben Nachman for producing all of the diagrams in these notes.

Contents

Chapter 1

Introduction to Particles

The Standard Model incorporates all the fundamental particles including the recently discovered Higgs particle. However, the Standard Model is elaborate and involves many parameters. Possible next steps to make sense of these include:

(a) Beyond-the-Standard-Model physics (including more particles, SUSY, or dark matter), and

(b) Simplification and unification (string theory or competitors).

Experimentally one finds many types of particle in nature, including:

- electrons
- photons
- protons / neutrons
- pions
- quarks
- gluons
- neutrinos
- gauge particles ($W^{\pm}, Z$)
- Higgs bosons

The Standard Model makes detailed sense of these but is not fully understood. Experimentally, the most important properties of the observed particles are *mass* and *spin*. These are related to the *geometry* of Minkowski space. Only massless particles move at the speed of light.

The simplest theory of particles is *perturbative quantum field theory* (pQFT). In pQFT, there is one particle per field (one particle spin state per field component). The theory is approximately linear, but this can fail when interactions between the fields are strong. Then *nonlinearity* between fields becomes crucial. Particles associated with a field may appear as *composites* or *not at all.* Solitons are *particle-like nonlinear* field structures.

1.1 Standard Model Fields

1.1.1 Fermions: Spin 1/2 ("matter")

The fermions occur in three families which are similar apart from their masses

$$\begin{matrix} \begin{pmatrix} e \\ \nu_e \end{pmatrix} & \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} & \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix} & \text{Leptons} \\ \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} & \text{Quarks} \end{matrix}$$

Note that all fermions have *anti-particles*. This was predicted by Dirac.

1.1.2 Bosons: Spin 0 or 1

$$\underbrace{g \text{ (gluon)}, \gamma \text{ (photon)}, W^{\pm}, Z}_{\text{spin 1}}, \underbrace{H \text{ (Higgs)}}_{\text{Spin 0}}$$

Quarks interact through gluons; leptons do not.

1.2 Observed Particles (of "long life")

- Leptons: e, ν_e (stable)
- Mesons: $q\bar{q}$, for example $\pi^+ = u\bar{d}$, $\pi^- = \bar{u}d$
- Baryons: qqq, for example $p = uud$ (stable)
- Gauge particles: γ (stable), $W^{\pm}$, Z, H, g (not seen as tracks, even in glueballs)

Remark. The strongly interacting particles are called *hadrons*: $\{mesons\} \bigcup \{baryons\} = \{hadrons\}$

1.3 Further Remarks on Particles

The pairs $\binom{e}{\nu_e}$ and $\binom{u}{d}$ lead to an $SU(2)$ structure. $SU(2)$ is a three-dimensional Lie group of 2×2 matrices which helps explain the $W^{\pm}, Z$ particles. The qqq baryons lead to an $SU(3)$ structure. $SU(3)$ is an eight-dimensional Lie group of 3×3 matrices, which explains the eight species of gluons. The Standard Model has the gauge group $U(1) \times SU(2) \times SU(3)$ which extends the $U(1)$ gauge symmetry of electromagnetism with its one photon (gauge boson).

1.3.1 Mass of gauge bosons

Naively, one expects the gauge bosons in QFT to be massless. This is evaded by:

(a) *Confinement* for gluons

(b) *Higgs mechanism* for $W^{\pm}, Z$

The Higgs mechanism breaks the $SU(2)$ symmetry. The $U(1) \times SU(3)$ symmetry remains unbroken.

1.3.2 The Poincaré Symmetry

The *Poincaré symmetry* combines translations and Lorentz transformations. The Poincaré group is a ten-dimensional Lie group (think geometrically: 3 rotations, 3 boosts, and 4 translations). The Poincaré symmetry explains the mass, spin, and particle-antiparticle dichotomy of particles. When Poincaré symmetry is broken, particles lose definite values for mass and spin. Gravity bends spacetime, changing the Minkowski metric. Thus we expect breaking of the Poincaré symmetry when gravity becomes significant.

1.3.3 Approximate Symmetries

Approximate symmetries simplify particle classification and properties. The most important example is that $\binom{u}{d}$ have similar masses. Thus $p = uud$ and $n = udd$ have similar masses and interactions ($m_p = 938$ MeV, $m_n = 940$ MeV). This gives rise to an approximate $SU(2)$ symmetry called *isospin*. There is also a less accurate $SU(3)$ flavour symmetry involving the u, d and s quarks.

1.4 Particle Models

(a) Perturbative QFT: quantize linear waves

(b) Point particles: naive quark model, non-relativistic

(c) Composites: baryons (qqq), nuclei (p, n), atoms (nuclei and e's)

(d) Exact field theory: classical localized field structures become solitons / particles after quantization

(e) String theory models of particles

We remark that multi-particle processes are hard to calculate in all models. At the LHC, $pp \longrightarrow$ hundreds of particles, mostly hadrons. Sometimes, one observes a few outgoing *jets*. These are the experimental signatures of quarks and gluons.

1.5 Forces and Processes

1.5.1 Strong Nuclear Force (quarks, gluons, $SU(3)$ gauge fields)

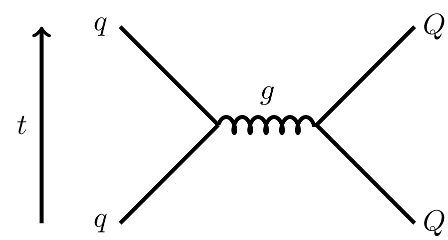


Figure 1.1: Quark Scattering, the process at the heart of hadron scattering

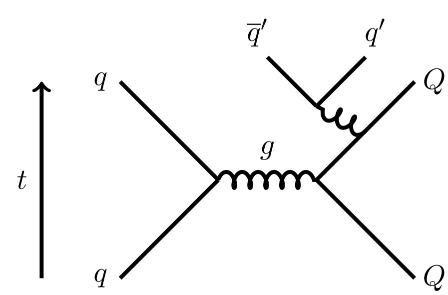


Figure 1.2: Particle Production (perhaps $pn \to pn\pi^0$)

Strong forces are the same for all quarks ("flavour blind"). However, the quark masses differ: $m_u \sim 2-5$ MeV, $m_t \sim 175$ GeV. Note: 1 GeV $= 10^3$ MeV $\sim$ proton mass. 1 TeV $= 10^3$ GeV $\sim$ LHC energies.

Strong interactions *do not change quark flavour*. For each quark flavour, the net number of quarks, i.e. (# quarks − # antiquarks) is conserved. Thus $N_u, N_d, N_c, N_s, N_t, N_b$ are all independently conserved in strong interactions. The conserved total number of quarks $N_q = N_u + N_d + N_c + N_s + N_t + N_b$ is always a multiple of three in any physical state. We write $N_q = 3B$, where B is the *baryon number*, which is evidently also conserved.

1.5.2 Electroweak Forces

Electroweak forces involve the photon γ and the vector bosons $W^{\pm}, Z$ and may or may not change quark flavour. However, the net number of quarks N_q, and hence baryon number B, remains conserved. The lepton number L is also conserved (all leptons have $L = 1$, antileptons have $L = -1$).

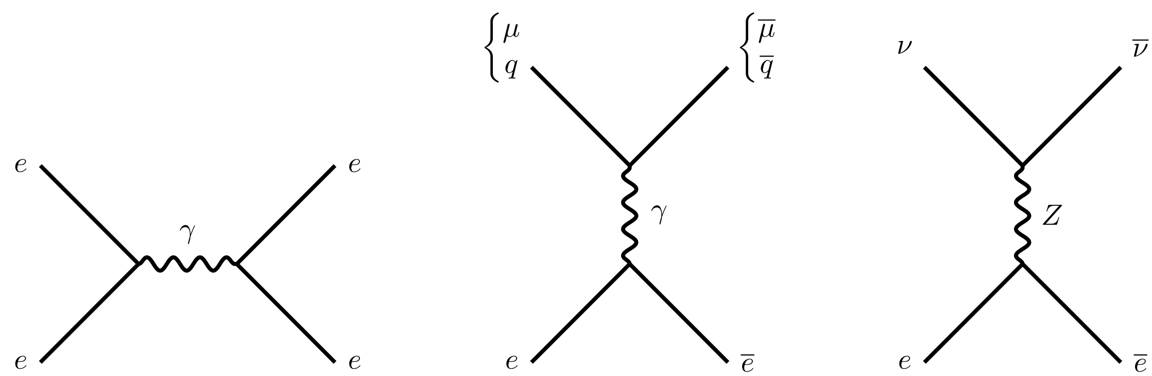


Figure 1.3: Some electroweak interactions with their Feynman diagrams

The photon only couples to electrically charged particles. Z couples to neutrinos too.

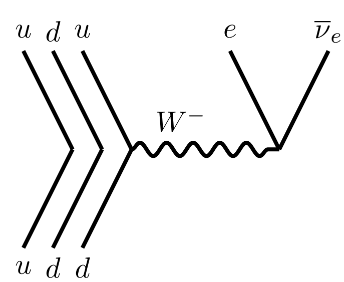


Figure 1.4: Neutron decay

Neutron decay is mediated by quark decay: $n \longrightarrow pe\bar{\nu}_e$. This is a *flavour-changing electroweak process*, involving a W-boson.

Heavy meson decay often involves a *transition between families* (whose strength is controlled by the CKM matrix).

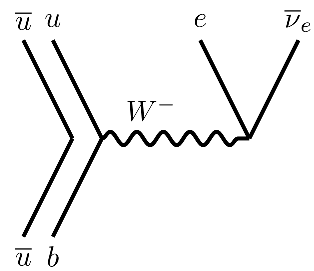


Figure 1.5: Heavy meson decay

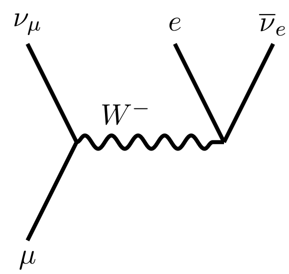


Figure 1.6: Muon decay

Heavy quarks are produced in $Q\bar{Q}$ pairs in strong interactions. They separate and decay weakly, leaving short tracks. The background Higgs field couplings determine the masses of all other particles, but do not appear in Feynman scattering diagrams. The Higgs particle H shows up in diagrams like:

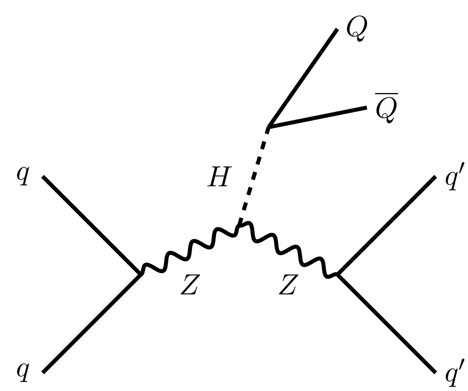


Figure 1.7: An interaction involving the Higgs

The strength of the Higgs particle couplings are also proportional to the masses of other particles, i.e., Z and Q in the case above. H couples preferentially to heavy particles.

Weak interactions are "weak" and hence slow only if the energy available is $\ll M_W, M_Z \sim 80/90$ GeV, as in neutron or muon decay. Strong interactions are "fast," occurring on time scales $\sim 10^{-24}s$ (the time for light to cross a proton).

Chapter 2

Symmetry

Definition 1. A *group* is a set $G = \{g_1 = I, g_2, g_3, \dots\}$ with

(i) a composition rule (binary operation) $g \star g' \in G$ which we usually denote gg',

(ii) a unique identity $I \in G$ such that $Ig = gI = g$ for all $g \in G$,

(iii) associativity: $(gg')g'' = g(g'g'') = gg'g''$ for all $g, g', g'' \in G$, and

(iv) unique inverse: $\forall g \in G, \exists! g^{-1}$ such that $gg^{-1} = g^{-1}g = I$.

If the binary operation is commutative, we say G is *abelian*.

2.1 Symmetry

Many physical and mathematical objects or physical theories possess symmetry. A symmetry is a transformation that leaves the thing unchanged. *The set of all possible symmetries forms a group*:

(i) Symmetries can be composed. We usually interpret gg' as "act with g' first and then act with g".

(ii) Doing nothing is a symmetry, the identity I.

(iii) $gg'g''$ does not need brackets because it means: act with g'' then g' then g.

(iv) A symmetry transformation g can be reversed, which gives the inverse g^{-1}. The inverse is itself a symmetry.

The conclusion is that group theory is the mathematical framework of symmetry. Every group is the symmetry of something, at least itself. A natural question is then, "Why does symmetry occur in nature?" As with most of the "big" questions, there are no easy answers. However, some partial answers are given by the following arguments:

(1) Solutions of variational problems generally exhibit a high degree of symmetry. For example, circles maximize area. As another example, Minkowski space is a stable solution of the Einstein equations (Einstein-Hilbert action) and has a high symmetry compared with a random spacetime. The symmetries of Minkowski space - known as the Poincaré group - are important and discussed in some detail later on.

(2) Physics often uses more mathematical variables than are really present in nature, leading to different descriptions of the same phenomenon. Transformations between them are known as *gauge symmetries*. Gauge symmetries are *exact*. For example:

(a) Coordinate transformations.

(b) Gauge transformations in electrodynamics, where the fields $\vec{B}, \vec{E}$ are physical while the potential $A^\mu = (\varphi, \vec{A})$ is partly non-physical. The potential can be freely gauge transformed without altering the physics.
Remark: Gauge transformations were named by Weyl, who thought physics could not depend on a "ruler." Although this idea ultimately turned out to be wrong, as there are fundamental length scales, the name remains.

(c) Non-physical changes to the phase of a wavefunction in quantum mechanics.

(3) Approximate symmetries arise by ignoring part of the physics or by making simplifying assumptions. For example, one ignores the difference between the masses of the u and d quarks, and also ignores the effects of electric charge, to get the $SU(2)$ isospin symmetry of hadron physics.

Symmetry simplifies analysis; hence it's popular with theorists. Symmetry leads to conservation laws (e.g. energy, angular momentum, electric charge), by Noether's theorem.

Chapter 3

Lie Groups and Lie Algebras

Lie groups have infinitely many elements. The elements depend *continuously* on a number of (real) parameters, called the *dimension* of the group, which will usually be finite here. The group operations (products and inverses) depend continuously (and smoothly) on the parameters.

Definition 2. A *Lie group* G is a smooth manifold which is also a group with smooth group operations.

Definition 3. The *dimension* of G, denoted $\dim G$, is the dimension of the underlying manifold.

The coordinates of gg' depend smoothly on the coordinates of g and g'. The inverse g^{-1} also depends smoothly on the coordinates of g.

Examples:

(i) $(\mathbb{R}^n, +)$. $\mathbb{R}^n$ is a manifold of dimension n. $\vec{x}'' = \vec{x} + \vec{x}'$ is a smooth function of $\vec{x}$ and $\vec{x}'$. The inverse $\vec{x}^{-1} = -\vec{x}$ is also smooth.

(ii) $S^1 = \{\theta : 0 \leq \theta \leq 2\pi\}$ with $\theta = 0$, $\theta = 2\pi$ identified (to skirt the issue that some manifolds require more than one chart). Here the group operation is addition $\bmod 2\pi$. Equivalently, $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, with multiplication in $\mathbb{C}$ giving the group operation. The equivalence between these two portrayals can be seen via $z = \exp i\theta$. S^1 has dimension 1.

3.1 Subgroups of G

Definition 4. A *subgroup* $H \subset G$ is a subset of G closed under the group operation inherited from G. We sometimes write $H \leq G$.

Note that a subgroup can be discrete, e.g., $\{z = 1, -1\} \subset S^1$, but discrete subgroups are *not* Lie subgroups. If H is a continuous subgroup and a *smooth submanifold* of G, then H is called a *Lie subgroup*. A Lie subgroup usually has a smaller dimension than the parent group.

3.2 Matrix Lie Groups

Lie groups of square matrices abound. These are called *linear Lie groups*, as the matrices act linearly on vectors in a vector space. The group operation in matrix Lie groups is always multiplication $M_1 \star M_1 = M_1 M_2$. Although addition is a sensible matrix operation, it is not the group operation. The identity in a matrix Lie group is always the unit matrix, and the inverse of an element M is the inverse matrix M^{-1}. Matrix multiplication is automatically associative, provided the matrix elements multiply associatively (for example, when they belong to a field). We will restrict our study to matrices over $\mathbb{R}$ or $\mathbb{C}$. The principal example of a matrix Lie group is given by the following definition:

Definition 5. The *General Linear group* is:

$$GL(n) = \{n \times n \text{ invertible matrices}\}.$$

$GL(n,\mathbb{R})$ with entries over $\mathbb{R}$ has real dimension n^2. $GL(n,\mathbb{C})$ with entries over $\mathbb{C}$ has real dimension $2n^2$ and complex dimension n^2. The identity in $GL(n)$ is the unit matrix $I = I_n$.

The condition of invertibility is equivalent to the condition $\det M \neq 0$ (for $\mathbb{R}$ and $\mathbb{C}$). This is an "open condition," so $\dim GL(n)$ is not reduced from the dimension of the space of all $n \times n$ matrices. (Put another way, the matrices with $\det = 0$ are a subset of measure zero within the space of all matrices). $GL(n,\mathbb{R})$ has a subgroup $GL^+(n,\mathbb{R}) = \{M \text{ real}, \det M > 0\}$.

3.2.1 Important Subgroups of $GL(n)$

(1) $SL(n) = \{M : \det M = 1\}$, the special linear group. Note that the group composition closes, since determinants of matrices multiply: $\det M_1 M_2 = \det M_1 \det M_2$.

$$\dim SL(n,\mathbb{R}) = n^2 - 1 \text{ (real dimension)}$$
$$\dim SL(n,\mathbb{C}) = 2n^2 - 2 \text{ (real dimension), or } n^2 - 1 \text{ (complex dimension)}$$

Note that the complex dimension is reduced by 1 since we've imposed one (complex) algebraic constraint.

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

(i) $O(n) = \{M : M^T M = I\}$, the orthogonal group. Closure here is again easy to see, since for M_1 and M_2 in $O(n)$

$$(M_1 M_2)^T M_1 M_2 = M_2^T M_1^T M_1 M_2 = I$$

$O(n)$ has inverses by construction. We note that transformations in $O(n)$ preserve length in the sense that if $\vec{v}' = M\vec{v}$ with $M \in O(n)$, then

$$\vec{v}' \cdot \vec{v}' = M\vec{v} \cdot M\vec{v} = (M\vec{v})^T M\vec{v} = \vec{v}^T M^T M\vec{v} = \vec{v} \cdot \vec{v}.$$

If $M \in O(n)$, then $\det M = \pm 1$. This follows using $M^T M = I$ and the determinant property above.

(ii) $SO(n) = \{R \in O(n) : \det R = +1\}$, the special orthogonal group. Geometrically this corresponds to the group of rotations in $\mathbb{R}^n$. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a frame in $\mathbb{R}^n$, then $\{R\vec{v}_1, \ldots, R\vec{v}_n\}$ is a frame *with the same orientation*. Volume elements are preserved by $R \in SO(n)$ (in $\mathbb{R}^3$, $\vec{v}_1 \wedge \vec{v}_2 \cdot \vec{v}_3 = R\vec{v}_1 \wedge R\vec{v}_2 \cdot R\vec{v}_3$). We note that $O(n)$ additionally contains orientation reversing elements, i.e., reflections, which are excluded from $SO(n)$. We note also that $\dim O(n) = \dim SO(n) = \frac{1}{2}n(n-1)$ The argument has to do with the fact that the columns of matrices in $O(n)$ are mutually orthonormal (See Example Sheet 1, Problem 3).

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

(i) $U(n) = \{U \in GL(n, \mathbb{C}) : U^\dagger U = I\}$, the unitary group. Note that $(U^\dagger)_{ij} = U^*_{ji}$. $U(n)$ preserves the norm of complex vectors, and the proof is essentially the same as for the length-preserving property of the orthogonal group. Note that $U^\dagger U = I \implies |\det U|^2 = 1$ (unit magnitude).

(ii) $SU(n) = \{U \in U(n) : \det U = 1\}$, the special unitary group.

$$\dim U(n) = n^2 \text{ (real dimension)}$$
$$\dim SU(n) = n^2 - 1$$

Note that $O(n) \subset U(n)$ and $SO(n) \subset SU(n)$ are the *real* subgroups.

Example: $U(1) \simeq SO(2)$. These have underlying manifold S^1, the circle.

(a) $U(1) = \{\exp i\theta : 0 \le \theta \le 2\pi\}$, with $\theta = 0$, $\theta = 2\pi$ identified. The product is $\exp i\theta \exp i\phi = \exp i(\theta + \phi)$.

(b) $SO(2) = \left\{ R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : 0 \le \theta \le 2\pi \right\}$. Using trigonometric formulae, one finds $R(\theta)R(\phi) = R(\theta + \phi)$. $R(\theta)$ is an counter-clockwise rotation by the angle θ in a plane.

3.2.2 A Remark on Subgroups Defined Algebraically

$GL(n)$ is obviously a smooth manifold with smooth group operations. The coordinates are the matrix elements ($\det M \neq 0$ defines an open subset of $\mathbb{R}^{n^2}$ or $\mathbb{C}^{n^2}$). Subgroups defined by algebraic equations involving matrix entries (e.g. $\det U = 1$ or $M^T M = I$) are "algebraic varieties." The natural question is then, "Are they manifolds?" In general, algebraic varieties can have singularities (non-manifold points). For example, consider the Cassini ovals, defined by $(x^2 + y^2)^2 - 2(x^2 - y^2) + 1 = b$.

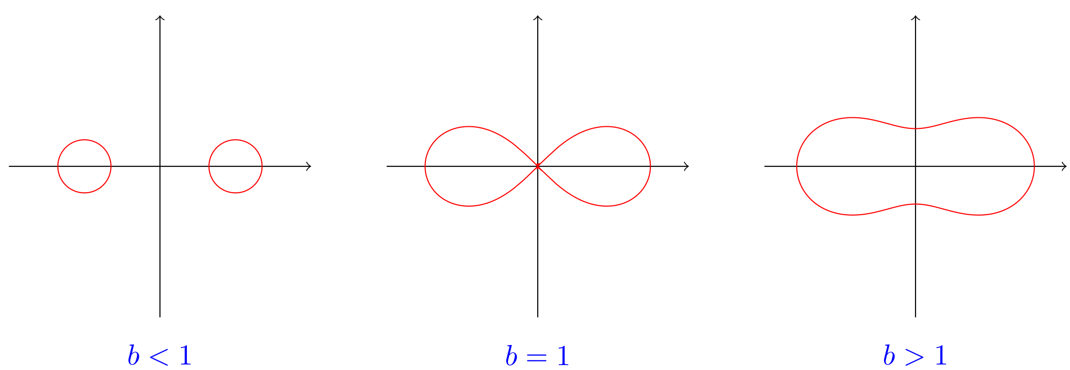
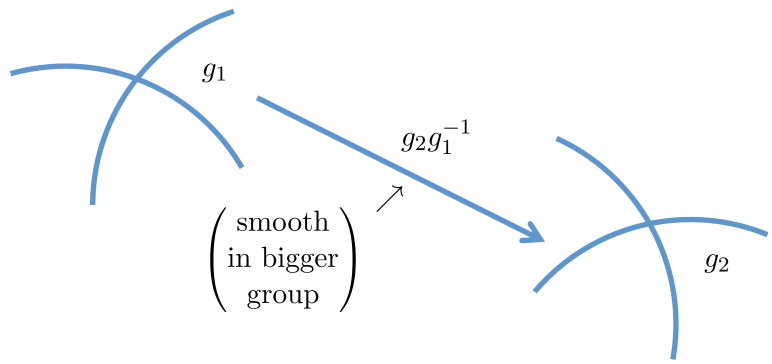


Figure 3.1: The Cassini ovals

The case $b = 1$ has a singularity. At least naively, this algebraic variety isn't a manifold. Fortunately, the *group* structure of our manifolds prevents singularities. The argument is the following:

Assume there is a singularity at $g_1 \in G$. Then there is a singularity at $g_2 \in G$, because the action of $g_2 g_1^{-1}$ by matrix multiplication is smooth (see figure below). Since g_2 was arbitrary, singularities occur everywhere in G. This is a contradiction, since a variety cannot be singular everywhere. Therefore G is a smooth manifold. We conclude that algebraically defined subgroups of $GL(n)$ are Lie groups.



3.3 Lie Algebras

The Lie algebra $L(G)$ of a Lie group G is the tangent space to G at the identity $I \in G$. We study the tangent space by differentiating curves in G. $L(G)$ is a vector space of dimension $\dim G$, with an algebraic structure called the *Lie bracket*. The algebraic structure of $L(G)$ *almost* uniquely determines G. Group geometry thus reduces to algebraic calculations. This fact was useful to Lie and continues to be for physicists (and mathematicians, too). For two matrices X and Y, the Lie bracket is $[X,Y] = XY - YX$, i.e., the commutator. We sometimes denote the Lie algebra of a Lie group G using the lowercase Fraktur script $\mathfrak{g}$.

3.3.1 Lie Algebra of $SO(2)$

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

with $f(0) = 0$ is a curve in $SO(2)$ through the identity. Differentiating with respect to t and evaluating at the origin gives:

$$\left.\frac{dg}{dt}\right|_{t=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \left.\frac{df}{dt}\right|_{t=0}.$$

For any f, this is a multiple of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Thus

$$\mathfrak{so}(2) = \left\{ \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix} : c \in \mathbb{R} \right\}.$$

Note that these matrices are not in $SO(2)$; they're in $L(SO(2))$, the tangent space at the identity.

3.3.2 Lie Algebra of $SO(n)$

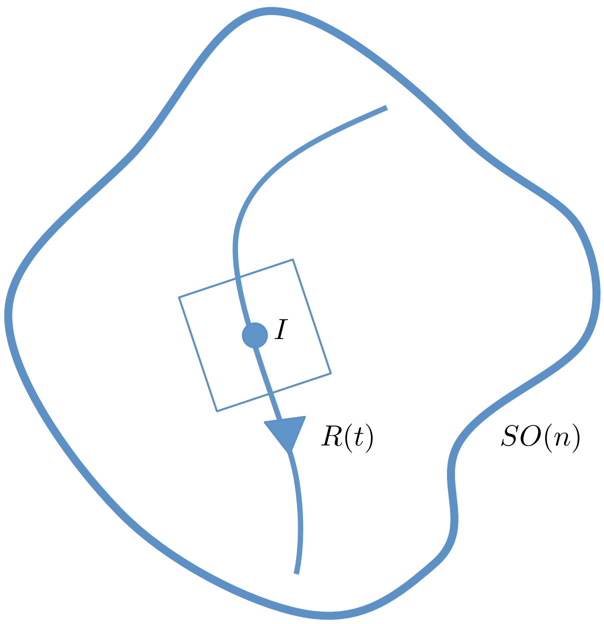
Consider a curve $R(t) \in SO(n)$ with $R(0) = I$. We require $R(t)^T R(t) = I$. Differentiating with respect to t, we find:

$$\frac{d}{dt}(R(t)^T R(t)) = R^T \dot{R} + \dot{R}^T R = \frac{d}{dt} I = 0, \forall t.$$

At $t = 0$, we have $R = I$, which gives the condition $\dot{R} + \dot{R}^T = 0$. In other words, $\dot{R}$ is antisymmetric. Therefore we find that

$$\begin{aligned} L(SO(n)) &= \{X : X + X^T = 0\} \\ &= \{\text{vector space of real antisymmetric } n \times n \text{ matrices}\}. \\ \dim L(SO(n)) &= \underbrace{\frac{1}{2}}_{\text{symmetry}} \underbrace{(n^2 - n)}_{\text{diag} = 0} = \frac{1}{2} n(n-1) \end{aligned}$$

Note that $L(O(n)) = L(SO(n))$ because an $O(n)$ matrix R near the identity has $\det R = 1$. In other words, $SO(n)$ is the part of $O(n)$ that is connected to the identity.



3.3.3 Lie algebra of $SU(n)$ and $U(n)$

Let $U(t)$ be a curve in $SU(n)$ with $U(0) = I$. So $U(t)^\dagger U(t) = I$ and $\det U(t) = 1$. For small t, we assume that $U(t)$ can be expanded as a power series: $U(t) = I + tZ + \ldots$. Z must be anti-hermitian so that $U^\dagger U = I$ to first order in t. Then

$$U(t) = \begin{pmatrix} 1 + tZ_{11} & tZ_{12} & \cdots \\ tZ_{21} & 1 + tZ_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

$$\begin{aligned} \implies \det U &= (1 + tZ_{11})(1 + tZ_{22}) \cdots + \mathcal{O}(t^2) \\ &= 1 + t(Z_{11} + Z_{22} + \ldots) + \mathcal{O}(t^2) \end{aligned}$$

Thus we see that the condition $\det U = 1, \forall t \implies \operatorname{Tr} Z = 0$. Therefore

$$\begin{aligned} L(SU(n)) &= \{Z : Z + Z^\dagger = 0 \text{ and } \operatorname{Tr} Z = 0\} \\ &= \{n \times n \text{ traceless anti-hermitian matrices}\} \end{aligned}$$

For $U(n)$ there is no constraint on the phase of the determinant, so $L(U(n)) = \{Z : Z + Z^\dagger = 0\} = \{n \times n \text{ anti-hermitian matrices}\}$.

3.3.4 General Structure of $L(G)$ for a matrix group G

(1) Vector Space Structure:
Suppose $X_1, X_2 \in L(G)$. Then $X_1 = \dot{g_1}(t)|_{t=0}, X_2 = \dot{g_2}(t)|_{t=0}$ for curves $g_1(t), g_2(t) \in G$ with $g_1(0) = g_2(0) = I$. Let $g(t) = g_1(\lambda t)g_2(\mu t)$, with λ, μ real. Then

$$\begin{aligned}\frac{d}{dt}(g_1(\lambda t)g_2(\mu t))|_{t=0} &= (\lambda \dot{g_1} g_2 + \mu g_1 \dot{g_2})|_{t=0} \\ &= \lambda \dot{g_1} + \mu \dot{g_2} \\ &= \lambda X_1 + \mu X_2 \\ \therefore \lambda X_1 + \mu X_2 \in L(G) &\implies L(G) \text{ is a vector space.}\end{aligned}$$

Note that since the definition of the Lie algebra is the tangent space at the identity, the computation above really just shows that our definition is consistent with group composition, since the tangent space (at any point on any manifold) is a vector space.

(2) Bracket on $L(G)$:
We'll now use more of the group structure. Let $g_1(t)\,, g_2(t)$ be curves in G passing through I at $t = 0$

$$\begin{aligned} g_1(t) &= I + tX_1 + t^2 W_1 + \dots \\ g_2(t) &= I + tX_2 + t^2 W_2 + \dots \end{aligned}$$

Computing products, we find:

$$\begin{aligned} g_1(t)g_2(t) &= I + t(X_1 + X_2) + t^2(X_1X_2 + W_1 + W_2) + \mathcal{O}(t^3) \\ g_2(t)g_1(t) &= I + t(X_2 + X_1) + t^2(X_2X_1 + W_2 + W_1) + \mathcal{O}(t^3) \end{aligned}$$

Define $h(t) = g_1(t)^{-1}g_2(t)^{-1}g_1(t)g_2(t)$ or equivalently $g_1(t)g_2(t) = g_2(t)g_1(t)h(t)$. $h(t)$ is a curve in G. We see that

$$h(t) = I + t^2[X_1, X_2] + \dots, \tag{3.1}$$

where $[X_1, X_2] = X_1X_2 - X_2X_1$. If we reparametrize according to $t^2 = s$, we see that $h(s)$ is a curve (for $s \geq 0$) such that $h(0) = I$ with tangent vector $[X_1, X_2] \in L(G)$. Thus $L(G)$ is closed under the bracket operation. ($h(t)$ has higher order terms at order t^3 etc, and hence at order $s^{3/2}$, but we only need $h(s)$ to have a first derivative, so that's no problem.)

Comment: Non-zero brackets are a measure of the non-commutativity of G. If G is abelian, $h(t) = I\ \forall t$, so $L(G)$ has trivial brackets.

Lemma 3.3.1. *If G is 1-dimensional, $L(G)$ has trivial brackets.*

Proof. $L(G) = \{cX : c \in \mathbb{R}\}$ for some fixed matrix X. $[cX, c'X] = cc'[X, X] = 0$. □

The only *connected* 1-dimensional Lie groups are S^1 and $\mathbb{R}$.

(3) Antisymmetry and Jacobi Identity:
The matrix bracket has the following general properties

(a) Antisymmetry $[X, Y] = -[Y, X]$

(b) Jacobi identity: $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (prove this by expanding out).

Remark: An abstract Lie algebra L is a vector space with a bracket $[\,,\,] : L \times L \to L$ linear in both entries, satisfying antisymmetry and the Jacobi identity (cf. Humphreys, Intro. Lie Alg. and Rep. Thy.)

(4) Basis and Structure Constants:
Let $\{T_i\}$ be a basis for $L(G)$. Define structure constants by $[T_i, T_j] \equiv c_{ijk} T_k$. Antisymmetry says $c_{ijk} = -c_{jik} \Leftrightarrow c_{(ij)k} = 0$. Now computing the nested brackets gives:

$$\begin{aligned}
[[T_i, T_j], T_k] &= c_{ijl}[T_l, T_k] = c_{ijl} c_{lkm} T_m \\
[[T_j, T_k], T_i] &= c_{jkl} c_{lim} T_m \\
[[T_k, T_i], T_j] &= c_{kil} c_{ljm} T_m \\
\text{Jacobi identity} \implies & \; c_{ijl} c_{lkm} + c_{jkl} c_{lim} + c_{kil} c_{ljm} = 0
\end{aligned}$$

3.3.5 $SU(2)$ and $SO(3)$: The Basic Non-abelian Lie Groups

We begin by comparing the Lie algebras of $SU(2)$ and $SO(3)$, $\mathfrak{su}(2)$ and $\mathfrak{so}(3)$.

$\mathfrak{su}(2) = \{2 \times 2$ traceless, anti-hermitian matrices$\}$: A basis is given in terms of the Pauli matrices:

$$T_a = -\frac{1}{2} i \sigma_a .$$

Here the σ_a are the (hermitian) Pauli matrices:

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

Recall the property of the Pauli matrices: $\sigma_a \sigma_b = \delta_{ab} I + i \epsilon_{abc} \sigma_c$. Using this, one finds:

$$\begin{aligned}
[T_a, T_b] &= -\frac{1}{4}(\sigma_a \sigma_b - \sigma_b \sigma_a) \\
&= -\frac{1}{4}(i \epsilon_{abc} \sigma_c - i \epsilon_{bac} \sigma_c) \\
&= -\frac{i}{2} \epsilon_{abc} \sigma_c = \epsilon_{abc} T_c \\
\implies [T_a, T_b] &= \epsilon_{abc} T_c
\end{aligned}$$

$\mathfrak{so}(3) = \{3 \times 3$ antisymmetric real matrices$\}$: A basis for $\mathfrak{so}(3)$ is given by:

$$\tilde{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; \tilde{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; \tilde{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

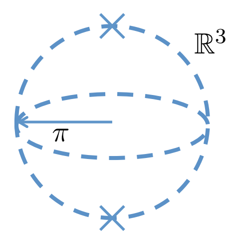
In other words, $(\tilde{T}_a)_{bc} = -\epsilon_{abc}$. Then $[\tilde{T}_a, \tilde{T}_b] = \epsilon_{abc} \tilde{T}_c$, as for $\mathfrak{su}(2)$. So $\mathfrak{su}(2) \simeq \mathfrak{so}(3)$. Therefore we expect the groups $SU(2)$ and $SO(3)$ to be similar.

$SU(2)$: $U^\dagger U = I$ and $\det U = 1$ imply that U has the form (cf. Example Sheet 1, Problem 6) $U = a_0 I + i \vec{a} \cdot \vec{\sigma}$ with $(a_0, \vec{a})$ real and $a_0^2 + \vec{a} \cdot \vec{a} = 1$. So the manifold of $SU(2)$ is

$$SU(2) = S^3 = \text{ unit sphere in } \mathbb{R}^4$$

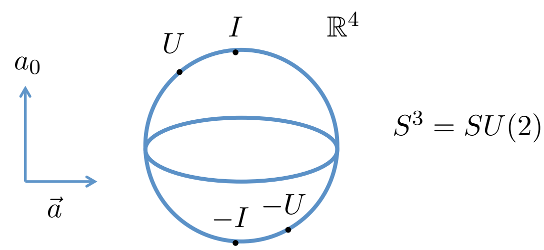
$SO(3)$: A rotation is specified by an axis of rotation $\hat{n}$ (i.e. a unit vector $\in \mathbb{R}^3$) and by an angle of rotation $\psi \in [0, \pi]$. (A rotation by a larger angle is thought of as a rotation about $-\hat{n}$.) We combine these into a 3-vector $\psi\hat{n} \in \{\text{ball of radius } \pi\} \subset \mathbb{R}^3$. Note that a rotation by π about $\hat{n}$ is equivalent to a rotation by π about $-\hat{n}$. Thus opposite points on the boundary are identified, and consequently the manifold $SO(3)$ does not actually have a boundary.

$$SO(3) = \{\text{ball in } \mathbb{R}^3 \text{ of radius } \pi \text{ with opposite points on boundary identified}\}$$

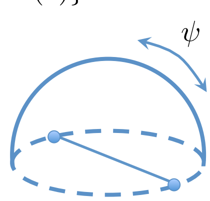


3.3.6 The Isomorphism $SO(3) \simeq SU(2)/\mathbb{Z}_2$

$SU(2)$ has a center $Z(SU(2)) = \mathbb{Z}_2 = \{I, -I\}$. If $U = a_0 I + i\vec{a}\cdot\vec{\sigma}$, then $(-I)U = -U = -a_0 I - i\vec{a}\cdot\vec{\sigma}$. Thus $SU(2)/\mathbb{Z}_2 = S^3$ with antipodal points $\{U, -U\}$ identified.



$$\begin{aligned} \therefore\ SU(2)/\mathbb{Z}_2 &= \{\text{upper half of } S^3\ (a_0 \geq 0) \text{ with opposite points of equator } S^2 \text{ identified}\} \\ &= \{\text{curved version of } SO(3)\}. \end{aligned}$$

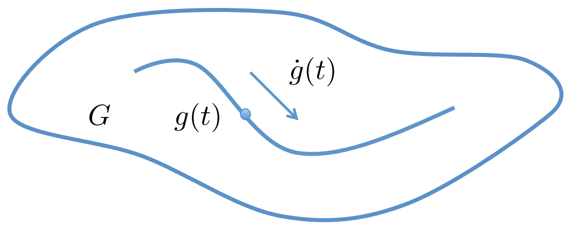


Note that the "curvature" is immaterial to the group or manifold structure. There is an explicit correspondence $U \in SU(2) \mapsto R(U) \in SO(3)$ with $R(-U) = R(U)$, where $U = \cos\frac{\alpha}{2}\, I + i\sin\frac{\alpha}{2}\, \hat{n}\cdot\vec{\sigma} \mapsto R(U) = $ rotation by α about $\hat{n}$.

3.4 Lie Group – Lie Algebra Relation

3.4.1 Tangent space to G at general element g

Let G be a matrix Lie group and $g(t) \in G$ a curve. $\frac{dg}{dt} \equiv \dot{g}$ is the tangent (matrix) at $g(t)$ and $g(t+\epsilon) = g(t) + \epsilon\dot{g}(t) + \mathcal{O}(\epsilon^2)$, where ϵ is infinitesimal. We can also write $g(t+\epsilon)$ as a product in G: $g(t+\epsilon) = g(t)h(\epsilon)$, where $h(\epsilon) = I + \epsilon X + \mathcal{O}(\epsilon^2)$ for some $X(t) \in L(G)$. $h(\epsilon)$ is the group element that generates the translation $t \mapsto t + \epsilon$.



Then $I + \epsilon X(t) = h(\epsilon) = g(t)^{-1}g(t+\epsilon) = g(t)^{-1}(g(t) + \epsilon\dot{g}(t)) = I + \epsilon g(t)^{-1}\dot{g}(t)$ to $\mathcal{O}(\epsilon)$. So $X(t) = g(t)^{-1}\dot{g}(t)$. Thus:

$$g(t)^{-1}\dot{g}(t) \in L(G), \; \forall t. \tag{3.2}$$

Similarly, by putting (a different) $h(\epsilon)$ on the left of $g(t)$ we have $\dot{g}(t)g(t)^{-1} \in L(G)$ (in general $\neq g^{-1}\dot{g}$). We see that the tangent space to G at g is not the Lie algebra $L(G)$, but is mapped to $L(G)$ by either left or right multiplication by g^{-1}.
Important remark: Suppose $g(x_1, x_2, \ldots, x_k)$ is a G-valued function on $\mathbb{R}^k$. Then $\frac{\partial}{\partial x_i} g \equiv \partial_i g$ is in the tangent space to G at g, so $(\partial_i g)g^{-1} \in L(G)$ and $g^{-1}(\partial_i g) \in L(G)$. These formulae appear in gauge theory.
We now consider the converse. Suppose $X(t) \in L(G)$ is a given curve in the Lie algebra. We can write down the equation:

$$\begin{aligned} &g(t)^{-1}\dot{g}(t) = X(t), \\ &g(0) = I \text{ (could be the more general element } g_0). \end{aligned}$$

This equation with an initial condition has a unique solution in G. The equation makes intuitive sense because the "velocity" $\dot{g}$ is always tangent to G. As a special case, consider $X(t) = X =$ const. Then $\dot{g} = gX$ with $g(0) = I$. This equation has the solution $g(t) = \exp tX$.

Proof. By definition

$$\exp tX = \sum_{n=0}^{\infty} \frac{1}{n!}(tX)^n \tag{3.3}$$

and this series converges for all t. Then

$$\begin{aligned} \frac{d}{dt}(\exp tX) &= X + tX^2 + \frac{1}{2}t^2X^3 \\ &= \exp(tX)X \end{aligned}$$

Thus the equation is satisfied and $g(0) = I$. Note also that $g(t) = \exp(tX)$ commutes with X, so also solves the equation $\dot{g} = Xg$. □

Claim: The curve $\{g_X(t) = \exp(tX) : -\infty < t < \infty\}$ is an abelian subgroup of G, generated by X.

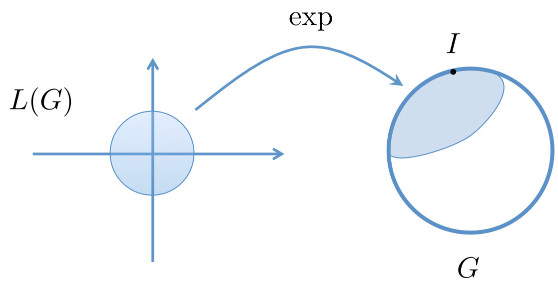
Proof. Using the series definition of $\exp(tX)$ one can verify that

$$\begin{aligned} g(0) &= I, \\ g(s)g(t) &= g(t+s) = g(s+t) \text{ (by combining all terms at a given order in } X), \\ g(t)^{-1} &= g(-t) \end{aligned}$$

□

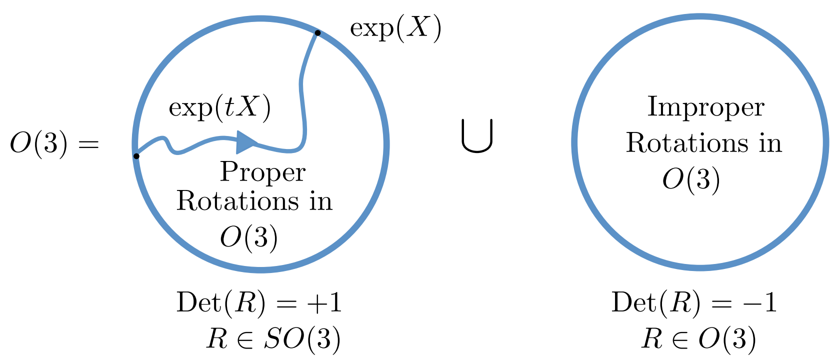
$g_X(t)$ is isomorphic either to $(\mathbb{R},+)$ if $g_X(t) = I$ only for $t = 0$ or to S^1 if $g_X(t_0) = I$ for some $t_0 \neq 0$ and not for all t.

We can gain a general insight from the above considerations. Set $t = 1$, and consider all $X \in L(G)$. We have then found a map $L(G) \to G$, given by $X \mapsto \exp X$. This map is locally bijective (proof omitted), as all elements $g \in G$ close to I can be expressed uniquely as $\exp X$ for some small X.



Note that this map is *not* globally simple and in most cases not even one-to-one. For example, the map $\mathbb{R} \mapsto S^1$ given by $\theta \mapsto \exp i\theta \in \{z \in \mathbb{C} : |z| = 1\}$ is onto but not one-to-one, as $\exp i2\pi n = 1, \forall n \in \mathbb{Z}$.
In general, the image of exp is not the whole group G, but rather the *component connected to* I in G.

Example: $O(3)$ is disconnected



So $\exp \mathfrak{so}(3) = \exp \mathfrak{o}(3) = SO(3)$. Evidently improper rotations R cannot be expressed as $\exp X$ with X antisymmetric and real.

3.4.2 The Baker-Campbell-Hausdorff Formula

If group elements are expressed as $\exp X$, $X \in L(G)$, can we calculate products? Remarkably, there is a universal formula:

$$\exp X \exp Y = \exp Z, \text{ where } Z = X + Y + \frac{1}{2}[X,Y] + \text{higher, nested brackets.}$$

This can be checked up to the order given by expanding out both sides. Finding the structure of the BCH formula to all orders, and proving its validity, is not trivial.
We deduce that the Lie algebra, with its bracket structure, determines the group structure of G near I.

Chapter 4

Lie Group Actions: Orbits

A Lie group G can act in many ways on other objects.

Definition 6. An *action* of G on a manifold $\mathcal{M}$ is a set of smooth maps $g : \mathcal{M} \to \mathcal{M}$ for all $g \in G$, consistent with the group composition $g_1(g_2(m)) = (g_1 g_2)(m)$ for all $g_1, g_2 \in G$ and for all $m \in \mathcal{M}$.

Note: This is equivalent to a map $G \times \mathcal{M} \to \mathcal{M}$, assumed here to be smooth in both arguments.

Definition 7. The *orbit* of a point $m \in \mathcal{M}$ is the set $G(m) = \{g(m) : g \in G\}$

Proposition 1. *If $m' \in G(m)$, $G(m) = G(m')$*

Proof. $m' = g(m) \implies G(m') = G(g(m)) = (Gg)(m) = G(m)$ $\square$

Theorem 4.0.1. *$\mathcal{M}$ is a disjoint union of orbits of G*

Proof. (Omitted.) Two orbits are either the same, as sets, or have no point in common. $\square$

Example. $SO(n)$ acts on $\mathbb{R}^n$. The orbits are the spheres S^{n-1}, labelled by the radius r. The origin $\vec{0} \in \mathbb{R}^n$ is special, as its orbit is this single point.

4.1 Examples of Group Actions

Definition 8. The *left action* of G on G is defined by $g : G \to G$ with $g(g') = gg'$

Definition 9. The *right action* of G on G is defined by $g : G \to G$ with $g(g') = g'g^{-1}$

Note that the right action uses the inverse of g; this is necessary so that the action satisfies the group composition law.

Definition 10. An action $G \times \mathcal{M} \to \mathcal{M}$ is said to be *transitive* if $\mathcal{M}$ consists of one orbit.

Transitivity of the left and right actions on G. Let $g', g'' \in G$. g' and g'' are in the same left orbit, since $g(g') = g''$ when $g = g''g'^{-1}$. A similar argument applies for right orbits. $\square$

Definition 11. *Conjugation* by G on G is the action defined by $g(g') = gg'g^{-1}, \forall g, g' \in G$

The orbit structure under conjugation is more interesting. In particular, one can show (cf. Example Sheet 2, Problem 4) that for matrices "conjugation preserves eigenvalues." Matrices with different eigenvalues must be in distinct orbits. The idea is that conjugation amounts to a change of basis in a matrix Lie group. Note also that one orbit is the identity alone, since $g(I) = gIg^{-1} = I, \forall g$.

Definition 12. The *combined action* of $G \times G$ on G is defined by $(g_1, g_2)(g') = g_1 g g_2^{-1}$.

Note that $G_1 \times G_2 = \{(g_1, g_2) : g_1 \in G_1, g_2 \in G_2\}$ with the product given by $(g_1, g_2) \star (g_1', g_2') = (g_1 g_1', g_2 g_2')$ and identity (I_1, I_2).

Example. The actions of $SU(2)$ on $SU(2)$.
Recall that we can express $g \in SU(2)$ as $g = a_0 I + i\vec{a} \cdot \vec{\sigma}$ with the constraint $a_0^2 + \vec{a} \cdot \vec{a} = 1$. (One way to see this parametrization is to think of $SU(2)$ as the subgroup of the quaternions $\mathbb{H}$ with unit length.) This parametrization leads us to understand $SU(2)$ to be S^3 as a manifold. If $g = a_0 I + i\vec{a} \cdot \vec{\sigma}$ and $g' = b_0 I + i\vec{b} \cdot \vec{\sigma}$, we can look at the left action given by:

$$\begin{aligned} g(g') \equiv gg' &= (a_0 b_0 - \vec{a} \cdot \vec{b})I + i(a_0 \vec{b} + \vec{a} b_0 - \vec{a} \times \vec{b}) \cdot \vec{\sigma} \\ &\equiv c_0 I + i\vec{c} \cdot \vec{\sigma} \end{aligned}$$

(cf. Example Sheet 1, Problem 6 and Sheet 2, Problem 5 for details.) We see that $(c_0, \vec{c})$ depends linearly on $(b_0, \vec{b})$ and $c_0^2 + \vec{c} \cdot \vec{c} = b_0^2 + \vec{b} \cdot \vec{b} = 1$. Thus we see that the left action of g defines an element of $O(4)$ (depending on a_0 and $\vec{a}$).
The identity I acts trivially, and $SU(2)$ is connected, so the left action of g must be a proper rotation (a proper rotation is an element of $SO(4)$ with det $= +1$). We deduce that $SU(2)_L \leq SO(4)$, where the notation $\leq$ denotes a subgroup. Similarly, the right action gives a different subgroup $SU(2)_R \leq SO(4)$. In fact, the combined action of $SU(2)_L \times SU(2)_R$ gives every element of $SO(4)$.

Theorem 4.1.1. $SO(4) \simeq (SU(2)_L \times SU(2)_R)/\mathbb{Z}_2$

Proof. Omitted (cf. Example Sheet 2) □

The subgroup $\mathbb{Z}_2$ consists of (I, I) and $(-I, -I)$, since (g_1, g_2) and $(-g_1, -g_2)$ act identically as $SO(4)$ transformations. Because of the group structure given in the theorem, it follows that the Lie algebra is given by $\mathfrak{so}(4) = \mathfrak{su}(2)_L \oplus \mathfrak{su}(2)_R$.

4.2 The General Nature of an Orbit of G

Definition 13. Let G act on $\mathcal{M}$ transitively (single orbit). Let $m \in \mathcal{M}$. The *isotropy subgroup* (or *stabilizer*) at m is the subset H of G that leaves m fixed:

$$H = \{h \in G : h(m) = m\}$$

We often write $H(m) = m$. H is a subgroup of G, because if $h_1(m) = m$, $h_2(m) = m$ for $h_1, h_2 \in H$, then $(h_1 h_2)(m) = h_1(h_2(m)) = h_1(m) = m$. The inverses and identity are also in H.
Using m as a base point for $\mathcal{M}$, we can identify another point m' with a *coset* of H. If $m' = g(m)$, then $m' = gH(m)$. The element m' is identified not just with one element g that sends m to m' but with the whole (left) coset gH. Thus we have

$$\mathcal{M} \simeq \{\text{space of all cosets of } H \text{ in } G\} \equiv G/H$$

However, there is nothing special about the base point m. The isotropy group at m' is $H' = gHg^{-1}$, and this is structurally the same as H. We can check this:

$$H'(m') = gHg^{-1}(m') = gHg^{-1}g(m) = gH(m) = g(m) = m'$$

Thus we regard G/H and G/H' as the same. This motivates the following definition:

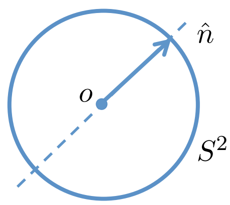
Definition 14. Let G be a group acting smoothly and transitively on the manifold $\mathcal{M}$. $\mathcal{M}$ is said to be a *homogeneous space.*

The critical part of the definition above is the transitivity: since there is just one orbit, all points on $\mathcal{M}$ are "similar."

Proposition 2. *If H is a Lie subgroup of G and $\mathcal{M}$ is as above, then* $\dim \mathcal{M} = \dim G - \dim H$.

We can check this statement near m. The tangent space to $\mathcal{M}$ is $L(G)/L(H)$ (as vector spaces), as H acts trivially. If we find a vector space decomposition $L(G) = L(H) \oplus V_m$, with V_m a vector space complement to $L(H)$ (could be orthogonal complement), then V_m can be identified with the tangent space to $\mathcal{M}$ at m.

Example. $SO(3)$ acts transitively on S^2, the unit sphere, since any point can be rotated into any other point. The isotropy group at $\hat{n}$ is the $SO(2)$ subgroup of $SO(3)$ rotations about the axis through $\vec{0}$ and $\hat{n}$. Thus $S^2 = SO(3)/SO(2)$. Note that we usually choose the base point $\hat{n}$ so that the $SO(2)$ rotations are about the x_3-axis.



Chapter 5

Representations of Lie Groups

Definition 15. A *representation* $D(G)$ of G is a linear group action $v \mapsto D(g)v$ of G on a vector space V, by invertible transformations. Let $\dim V = N$. Then N is called the *dimension* of D and $D(g) \in GL(N), \forall g \in G$.

Linearity says that

$$D(g)(\alpha v_1 + \beta v_2) = \alpha D(g)v_1 + \beta D(g)v_2$$

for $v_1, v_2 \in V$. To be a group action, D must satisfy

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

We see that identity and inverses are particularly well-behaved: $D(I) = I_N$ and $D(g^{-1}) = D(g)^{-1}$. Note that in order to get explicit matrices we must choose a basis for V.

Definition 16. A representation is *faithful* if $D(g) = I_N$ only for $g = I$.

In a faithful representation distinct group elements are represented by distinct matrices. Note that a slightly more sophisticated definition would say that the homomorphism $\phi : G \to GL(N)$ is injective, i.e., that $\ker(\phi)$ is trivial.

Example. Representations of the additive group $\mathbb{R}$. We require $D(\alpha + \beta) = D(\alpha)D(\beta)$.

(a) $D(\alpha) = \exp(k\alpha), k \in \mathbb{R}$. This is faithful if $k \neq 0$.

(b) $D(\alpha) = \exp(ik\alpha), k \in \mathbb{R}$. This is not faithful as $D(\alpha) = 1$ for $\alpha = 2\pi n/k$.

(c) $D(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. This representation has $N = 2$, and is not faithful.

(d) Let $V = \{\text{space of functions of } x\}$, an infinite-dimensional space. D is defined by $(D(\alpha)f)(x) = f(x - \alpha)$. This is a simple example of an induced representation (see Chapter 12).

5.1 Types of Representation

Definition 17. Let G be a matrix Lie group with matrices of dimension N. Then the representation $D(g) = g$ is called the *fundamental representation* of the group, and is N-dimensional.

Definition 18. The representation of G that sets $D(g) = I_N, \forall g \in G$ (and for any N) is called the *trivial representation*.

Definition 19. Let G be a matrix Lie group. Let $V = L(G)$, the Lie algebra of G. The *adjoint representation* of G, denoted Ad, is the natural representation of G on $L(G)$:

$$D(g)X \equiv (\operatorname{Ad} g)X \equiv gXg^{-1}, g \in G, X \in L(G)$$

Note that the adjoint representation Ad is the linearized version of the action of G on itself by conjugation. We do a couple of checks to make sure that we have a well-defined representation:

- Closure: $gXg^{-1} \in L(G)$
  There exists some curve $g(t) = I+tX+\dots$ in G with tangent X at $t = 0$. Then $\tilde{g}(t) = gg(t)g^{-1}$ is another curve in G and $\tilde{g}(t) = I + tgXg^{-1} + \dots$ with tangent gXg^{-1} at $t = 0$, thus $gXg^{-1} \in L(G)$.

- Ad is a representation:

$$\begin{aligned}(\operatorname{Ad} g_1 g_2)X &= g_1 g_2 X (g_1 g_2)^{-1} \\ &= g_1 g_2 X g_2^{-1} g_1^{-1} \\ &= (\operatorname{Ad} g_1)(\operatorname{Ad} g_2)X\end{aligned}$$

Note that we are thinking of $L(G)$ as a real vector space, so $\operatorname{Ad} g \in GL(\dim G, \mathbb{R})$. In fact, for $U(n)$ and $O(n)$, $\operatorname{Ad} g \in SO(\dim G)$.

Definition 20. An N-dimensional representation D of G is said to be *unitary* if $D(g) \in U(N), \forall g \in G$. If D is also real, then $D(g) \in O(N)$, and the representation is said to be *orthogonal*.

Remark: Unitary representations are important in quantum mechanics and its various generalizations because a symmetry group should preserve the *norm* of all wave functions.

Definition 21. Let D be a representation of G acting on the vector space V. Let A be a fixed invertible transformation on V. Then we say that $\widetilde{D}(g) = AD(g)A^{-1}$ is an *equivalent representation* of G.

Note that equivalent representations are related by a change of basis of the vector space V.

Definition 22. Let D be a representation of G acting on V. D is *reducible* if there exists a proper, invariant subspace $W \subset V$, i.e., there exists a subspace W such that $D(G)W \subseteq W$. If no such subspace exists, then D is an *irreducible representation*, which we will sometimes call an "irrep".

Note that $D(G)W = W$, since $I \in G$. (Also, for any $g \in G$, $D(g)W = W$, because $D(g)$ is invertible.)

Definition 23. A representation D is *totally reducible* if it can be decomposed into irreducible pieces, i.e., if there exists a (possibly infinite) direct sum decomposition $V = W_1 \oplus W_2 \oplus \dots W_k$ such that $D(G)W_i = W_i$ and D restricted to W_i is an irrep.

In matrix language, for a totally reducible representation there exists a basis for V such that, simultaneously for all g, $D(g)$ is block diagonal, taking the form:

$$D(g) = \begin{pmatrix} D_1(g) & 0 & \cdots & 0 \\ 0 & D_2(g) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & D_k(g) \end{pmatrix},$$

where the D_i are each irreps. W_i consists of the column vectors which are zero in the subcolumns acted on by $D_j\,, j \neq i$, and non-zero (generally) in the subcolumn acted on by D_i.
Note that it is usually not easy, nor even desirable, to find the basis that makes D block diagonal. One then needs other techniques to determine the decomposition of D into irreps.

Theorem 5.1.1. *A finite-dimensional unitary representation is totally reducible.*

Proof (sketch). For each invariant subspace W, the orthogonal complement $W_\perp$ is also invariant, so $V = W \oplus W_\perp$. Now reduce W and $W_\perp$ until the process ends. Note that the process must end for a finite-dimensional representation.
(cf. Example Sheet 2, Problem 9 for more details) $\square$

The notions of irreducible representation, and total reducibility, are important, because vectors within W_i are actually related by G. The block diagonal form shows that vectors in W_i are not related by G to vectors in W_j (for $j \neq i$). If G is a symmetry acting on the Hilbert space V of all physical states, then only physical states (particles!) within an *irreducible* subspace W have similar properties.

Example. Let $V = \{\text{space of functions } f \text{ of } x \text{ with period } 2\pi\}$, i.e. $f(x + 2\pi) = f(x)$. The circle group S^1 acts on V by the representation $(D(\alpha)f)(x) = f(x - \alpha)$, where $0 \leq \alpha \leq 2\pi$. This representation is infinite-dimensional.

The space $W_n = \{f(x) = c_n e^{inx}\}$ is a 1-dimensional invariant subspace for each $n \in \mathbb{Z}$, because $f(x - \alpha) = c_n e^{in(x-\alpha)} = e^{-in\alpha} c_n e^{inx}$. The 1-dimensional representation that occurs here is $d^{(n)}(\alpha) = e^{-in\alpha}$.

The Fourier series decomposition of a general function f

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

exhibits the complete reducibility of V as

$$V = \cdots \oplus W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots$$

Each distinct 1-dimensional irrep of S^1 occurs once in V.

There is a generalisation of this analysis for the action of $SO(3)$ on functions defined on S^2 (decomposition into spherical harmonics), and more generally for the action of any Lie group G on the functions on any of its coset spaces G/H.

Chapter 6

Representations of Lie Algebras

By restricting a representation D of G to elements close to the identity I, we obtain the notion of a representation of the Lie algebra $L(G)$.

Definition 24. A *representation* d *of* $L(G)$ acting on a vector space V is a linear action $v \mapsto d(X)v$, with $X \in L(G)$ and $v \in V$, satisfying

$$d([X,Y]) = d(X)d(Y) - d(Y)d(X) = [d(X), d(Y)]$$

As with the case of groups, the *dimension* of the representation is $\dim d = N$, where N is the dimension of the vector space V.

Example: For a matrix Lie algebra, the fundamental representation is $d(X) = X$. There also exists the trivial representation in which $d(X) = 0, \forall X \in L(G)$.

A representation d of $L(G)$ is called (anti)hermitian if $d(X)$ is (anti)hermitian for all $X \in L(G)$.

6.1 Representation of $L(G)$ from a Representation of G

Let $g(t) = I + tX + \cdots \in G$. Write $D(g(t)) = I_N + td(X) + \dots$, which defines $d(X)$. Then d is the representation of $L(G)$ associated to D. We check now that the Lie bracket is preserved:
As D is a representation of G

$$D(g_1^{-1}g_2^{-1}g_1g_2) = D(g_1)^{-1}D(g_2)^{-1}D(g_1)D(g_2).$$

Now set

$$\begin{aligned} g_1(t) &= I + tX_1 + t^2W_1 + \dots \\ g_2(t) &= I + tX_2 + t^2W_2 + \dots \\ g_1^{-1}(t) &= I - tX_1 - t^2(W_1 - X_1^2) + \dots \\ g_2^{-1}(t) &= I - tX_2 - t^2(W_2 - X_2^2) + \dots \end{aligned}$$

After a brief computation very similar to the one leading to (3.1) we see that:

$$\begin{aligned} LHS &= D(I + t^2[X_1, X_2] + \dots) = I_N + t^2 d([X_1, X_2]) + \dots \\ RHS &= I_N + t^2[d(X_1), d(X_2)] + \dots \end{aligned}$$

and so we see the bracket is preserved.

Note: If D is a unitary representation of G, then $d(X)$ is antihermitian for all $X \in L(G)$. Why is this true? If $D(g)$ is unitary, then $D(g) = I + td(X)$ and $D(g)^\dagger = I + td(X)^\dagger$, so $D(g)D(g)^\dagger = I$ implies that $d(X) + d(X)^\dagger = 0$.

6.1.1 The adjoint representation of $L(G)$

The representation ad of $L(G)$ is associated to Ad of G (both of these representations act on $L(G)$, though one is a group representation and the other is a Lie algebra representation).
Recall the adjoint representation $(\mathrm{Ad}\, g)Y = gYg^{-1}, Y \in L(G)$. Set $g(t) = I + tX$, so $g^{-1}(t) = I - tX$. Then

$$\begin{aligned}(\mathrm{Ad}\, g)Y &= (I + tX)Y(I - tX) + \mathcal{O}(t^2)\\ &= Y - tYX + tXY + \mathcal{O}(t^2)\\ &= Y + t[X, Y] + \mathcal{O}(t^2)\\ &\equiv (I + t(\mathrm{ad}\, X))Y + \mathcal{O}(t^2)\end{aligned}$$

Thus we see that

$$(\mathrm{ad}\, X)Y = [X, Y]$$

This is the adjoint representation of $L(G)$ acting on itself. We can check easily (about three lines of computations) that $\mathrm{ad}[X, Y] = [\mathrm{ad}\, X, \mathrm{ad}\, Y]$ using the Jacobi identity. (Act with both sides on $Z \in L(G)$.)

6.2 Representation of G from a Representation of $L(G)$

Given $g \in G$, express g as $\exp X$ for $X \in L(G)$, then use the formula: $D(\exp X) = \exp(d(X))$. At least locally (i.e., in a neighborhood of the identity), this defines a representation D of the group G. How do we check this? We need to show that D "commutes appropriately" with exp:

$$D(\exp X)D(\exp Y) = D(\exp X \exp Y)$$

(See Example Sheet 2, Problem 11). Summarizing, we have

(a) Always true: Rep. of $G \longrightarrow$ Rep. of $L(G)$

(b) Mostly true: Rep. of $L(G) \longrightarrow$ Rep. of G (works locally, but can encounter problems globally)

In practice (in physics) it is often easier use representations of $L(G)$ rather than the corresponding representations of G.
Important questions now are: How can we classify and construct the irreps of $L(G)$ and G? What are their dimensions? We start to answer this by looking at $\mathfrak{su}(2)(= L(SU(2)))$. The results are very helpful as most Lie algebras have several inequivalent $\mathfrak{su}(2)$ subalgebras.

6.3 $\mathfrak{su}(2)$: The Mathematics of Quantum Angular Momentum

We know that $\mathfrak{su}(2)$ has the standard basis:

$$\left\{T_a = -\frac{1}{2} i\sigma_a : a = 1, 2, 3\right\}$$

It is convenient for our current purposes to construct a new basis using (non-real) linear combinations:

$$\begin{aligned}
h &= \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} = \frac{1}{2}\sigma_3 = iT_3 \\
e_+ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 + i\sigma_2) = iT_1 - T_2 \\
e_- &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_1 - i\sigma_2) = iT_1 + T_2
\end{aligned}$$

Note: This notation with h and $e_\pm$ appears in the text by Humphreys. h, e_+, e_- can be thought of as analogous to the familiar operators J_3, J_+, J_- in angular momentum theory.
We have the brackets:

$$\begin{aligned}
[h, e_+] &= e_+ \\
[h, e_-] &= -e_- \\
[e_+, e_-] &= 2h
\end{aligned}$$

and one may also express these bracket relations in terms of the adjoint representation ad of $\mathfrak{su}(2)$:

$$\begin{aligned}
(\operatorname{ad} h)e_+ &= e_+ \\
(\operatorname{ad} h)e_- &= -e_- \\
(\operatorname{ad} h)h &= 0
\end{aligned}$$

We see that we have diagonalized the operator $\operatorname{ad} h$, and e_+, e_-, h are eigenvectors. A maximal commuting (i.e., abelian) subalgebra in $\mathfrak{su}(2)$ is generated by h. The non-zero eigenvalues of $\operatorname{ad} h$ are called *roots*.

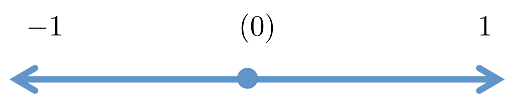


Figure 6.1: The root diagram for $\mathfrak{su}(2)$

6.3.1 Irreducible Representations of $\mathfrak{su}(2)$

The irreducible representations $d^{(j)}$ of $\mathfrak{su}(2)$, with the label j (spin), are as follows:

j	0	$\frac{1}{2}$	1	$\frac{3}{2}$	...
dim	1	2	3	4	...

The representation $d^{(j)}$ acts on $V^{(j)}$, a vector space of dimension $2j+1$. We introduce a basis $\{|j,m\rangle\}$ for $V^{(j)}$, with $m = j, j-1, \ldots, -j+1, -j$ ($2j+1$ values). Then:

$$\begin{aligned}
d^{(j)}(h)\,|j,m\rangle &= m\,|j,m\rangle \\
d^{(j)}(e_-)\,|j,m\rangle &= \sqrt{(j-m+1)(j+m)}\,|j,m-1\rangle \\
d^{(j)}(e_+)\,|j,m-1\rangle &= \sqrt{(j-m+1)(j+m)}\,|j,m\rangle
\end{aligned}$$

One can show that $[d^{(j)}(e_+), d^{(j)}(e_-)] = 2d^{(j)}(h)$ etc. This irreducible representation is antihermitian, since $d^{(j)}(h)^\dagger = d^{(j)}(h)$ and $d^{(j)}(e_+)^\dagger = d^{(j)}(e_-)$, which tells us that $d^{(j)}(T_a)$ is an antihermitian

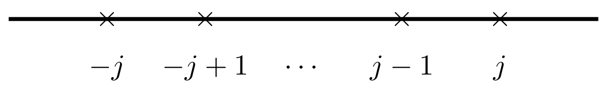


Figure 6.2: The weight diagram of $d^{(j)}$

matrix. (Note: the basis $\{h, e_+, e_-\}$ lives in the complexification $\mathfrak{su}(2)^{\mathbb{C}} = \mathfrak{su}(2) \otimes \mathbb{C}$). The eigenvalues of $d^{(j)}(h)$ are called *weights* and are real. For $d^{(j)}$, the weights are given by the labels m.

Some specific representations are: $d^{(0)}$ is the 1-dimensional trivial representation, $d^{(\frac{1}{2})}$ is the fundamental representation, and $d^{(1)}$ is the adjoint representation.

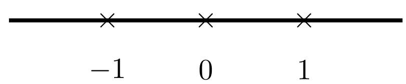


Figure 6.3: The weight diagram of $d^{(1)}$, which is the same as the root diagram including zero

One can show that $d^{(j)}$ exponentiates to an irreducible representation $D^{(j)}$ of $SU(2)$ for all j. As $L(SU(2))$ and $L(SO(3))$ are the same, they have the same irreps $d^{(j)}$. However, only for integer j does the irrep $d^{(j)}$ exponentiate to an irrep of $SO(3)$. One gets an $SO(3)$ irrep because within $SU(2)$, for integer j, $D^{(j)}(-I) = I_{2j+1}$. (If j is half-integer, $D^{(j)}(-I) = -I_{2j+1}$, which forbids the $\mathbb{Z}_2$ quotient).

6.4 Tensor Products of Representations

Tensor products are one of the most useful constructions in physics. Using tensor products one can combine representations of groups to produce a wide variety of physically interesting and useful further representations. Let $D^{(1)}(g)_{\alpha\beta}$ and $D^{(2)}(g)_{ab}$ be representations of G acting on vectors $\phi^{(1)}_\beta \in V^{(1)}, \phi^{(2)}_b \in V^{(2)}$. We define the tensor product $D^{(1)} \otimes D^{(2)}$ acting on $V^{(1)} \otimes V^{(2)}$ by:

$$(D^{(1)} \otimes D^{(2)})(g)_{\alpha a, \beta b} \equiv D^{(1)}(g)_{\alpha\beta} D^{(2)}(g)_{ab}$$

This acts on $\Phi \in V^{(1)} \otimes V^{(2)}$ by:

$$\Phi_{\alpha a} \mapsto \Phi'_{\alpha a} = D^{(1)}(g)_{\alpha\beta} D^{(2)}(g)_{ab} \Phi_{\beta b}$$

A special form of the tensor $\Phi_{\alpha a}$ is the factorized form $\phi^{(1)}_\alpha \phi^{(2)}_a$, but this is not necessary (in other words, not all tensors are the direct product of vectors). The dimension of the tensor product representation is

$$\dim(D^{(1)} \otimes D^{(2)}) = (\dim D^{(1)})(\dim D^{(2)})$$

6.4.1 The Representation of $L(G)$ associated to $D^{(1)} \otimes D^{(2)}$

Let $g \in G$. Set $g = I + tX$. Then, up to order t,

$$\begin{aligned} D^{(1)}(g) \otimes D^{(2)}(g) &= (I + t d^{(1)}(X)) \otimes (I + t d^{(2)}(X)) \\ &= I \otimes I + t(d^{(1)}(X) \otimes I + I \otimes d^{(2)}(X)) \end{aligned}$$

Thus the associated representation of $L(G)$ is:

$$d^{(1 \otimes 2)} = d^{(1)} \otimes I + I \otimes d^{(2)}$$

We see that in the tensor product representation $d^{(1\otimes 2)}$, the eigenvalues of $d^{(1)}$ and $d^{(2)}$ add. Therefore the weights of $d^{(1\otimes 2)}$ are the sums of the weights of $d^{(1)}$ and $d^{(2)}$.

6.4.2 Tensor Products of $\mathfrak{su}(2)$ Irreducible Representations

Let j denote the spin j irreducible representation $d^{(j)}$ of $\mathfrak{su}(2)$. The tensor product $j\otimes j'$ decomposes as

$$j \otimes j' = (j + j') \oplus (j + j' - 1) \oplus \cdots \oplus |j - j'| \tag{6.1}$$

into a direct sum of irreducible representations. This formula, known as the *Clebsch-Gordon series*, is used for combining states of particles with spins j and j'. One can verify the formula above by comparing weights on both sides (cf. Example Sheet 3, Problem 1).

Example:
$j = 1$ has weights $\{-1, 0, 1\}$. $j = \frac{1}{2}$ has weights $\{-\frac{1}{2}, \frac{1}{2}\}$. Thus $1\otimes\frac{1}{2}$ has weights $\{-\frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}\} = \{-\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}\} \cup \{-\frac{1}{2}, \frac{1}{2}\}$. Thus $1 \otimes \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2}$.

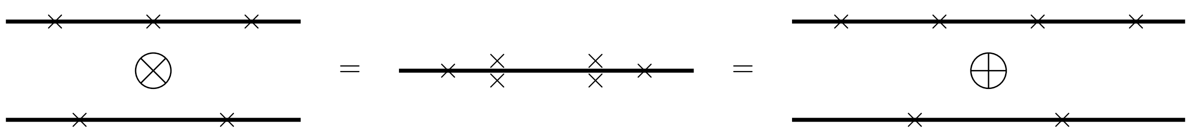


Figure 6.4: A diagrammatic illustration of weights adding.

In Eq.(6.1) each state on the right $|J, M\rangle$ with $|j - j'| \leq J \leq j + j'$ can be expressed explicitly as a linear combination of states on the left:

$$|J, M\rangle = \sum_m c_m\, |j, m\rangle \otimes \left|j', M - m\right\rangle$$

One writes more formally the coefficients c_m as the Clebsch-Gordon coefficients $c^{Jjj'}_{Mmm'}$, which (after a simple rescaling) are also known as the Wigner $3j$ symbols. They are non-vanishing only for $M = m + m'$, which is just the statement that weights add. For more information, see Landau and Lifschitz, Quantum Mechanics, §106.

6.5 Roots and Weights for general $L(G)$

The following applies to the large class of Lie algebras called "semisimple", which includes $L(SU(n))$ and $L(SO(n))$. The definition is clarified in Section 8.2.

Definition 25. Let $L(G)$ be a semisimple Lie algebra. A *Cartan subalgebra* is a subalgebra H of $L(G)$ such that:

(a) H is a maximal Abelian subalgebra of $L(G)$

(b) The adjoint representation ad of H is completely reducible.

Different Cartan subalgebras are related by conjugation by elements of G, so the Cartan subalgebra is essentially unique.

Definition 26. Let H be the Cartan subalgebra of $L(G)$. The dimension of H is called the *rank* of G, and denoted by k.

Definition 27. Let $h_i : i = 1, 2, \ldots, k$ be a basis of H. If $L(G)$ has dimension n, one can find a standard basis $\{h_1, h_2, \ldots, h_k; e_1, e_2, \ldots, e_{n-k}\}$, where each e_p is a simultaneous eigenvector for all $\operatorname{ad} h_i$ (since these mutually commute). The simultaneous eigenvalues of $\operatorname{ad} h_i$ (collected into a non-zero k-component vector) are called *roots*.

The root diagram exhibits the roots in Cartesian k-space (simplest for $k = 1$ or $k = 2$). There are $n - k$ roots.
Now consider a representation d of $L(G)$ acting on the N-dimensional vector space V. With h_i as above, the matrices $d(h_i)$ can be simultaneously diagonalized (since they again mutually commute). One can therefore find a basis for V consisting of simultaneous eigenvectors of $d(h_i)$. Let this basis be $\{\psi_1, \psi_2, \ldots, \psi_N\}$. For each basis element ψ we have

$$d(h_i)\psi = m_i\psi \text{ , where } i = 1, ..., k$$

so ψ is the simultaneous eigenvector of $d(h_i)$ with the vector of eigenvalues $\vec{m} = (m_1, m_2, \ldots, m_k)$. $\vec{m}$ is called the weight vector, or simply *weight*, of ψ. The set of all weights, for $\psi = \psi_1, \psi_2, \ldots, \psi_N$, defines the set of weights for the representation d. The weights are k-vectors, and there are N of them. Some of them can be identical (i.e., there can be “degeneracy”) and some can be zero vectors. (In the case of $\mathfrak{su}(2)$, the weights were single numbers because the Cartan subalgebra was spanned by a single element h.)
The weight diagram of d exhibits the weights in Cartesian k-space (and the roots are sometimes shown at the same time). The most important weight diagrams are those of the irreps of $L(G)$. The weights of a reducible representation d are just the union of the weights of all the irreps in the decomposition of d. Weight diagrams can therefore be used to determine the decomposition of a representation into irreps. (This is much easier than trying to find the block diagonal form, and also easier than using characters.) There is also a (grand) weight diagram, which shows all possible weights of all possible representations of $L(G)$. The grand weight diagram is a lattice in k-space, called the *weight lattice*.
The *roots* of $L(G)$ are the non-zero weights of the adjoint representation ad of $L(G)$. Thus

$$\{\text{roots of } L(G)\} \subset \{\text{weight lattice of } L(G)\}$$

We saw above that in a tensor product representation $d^{(1\otimes 2)}$ the eigenvalues of $d^{(1)}$ and $d^{(2)}$ add. In particular, each eigenvalue of $d^{(1\otimes 2)}(h_i)$ is the sum of an eigenvalue of $d^{(1)}(h_i)$ and an eigenvalue of $d^{(2)}(h_i)$. The weights of $d^{(1\otimes 2)}$ are therefore obtained by adding a weight of $d^{(1)}$ and a weight of $d^{(2)}$ in all possible ways. This will be illustrated in Chapter 9.

Chapter 7

Gauge Theories

Classical and quantum field theory is constructed from a Lagrangian density $\mathcal{L}$ which is gauge invariant and Lorentz invariant. The action is:

$$S = \int_{\mathbb{R}^4} \mathcal{L}\, d^4x$$

A gauge theory has a Lie group G acting as a local symmetry, i.e., acting independently at each spacetime point. Fields differing by a gauge transformation are physically the same.

7.1 Scalar Electrodynamics

Here the gauge group G is $U(1)$, which is abelian. This theory describes the electromagnetic field interacting with other electrically charged fields. The $U(1)$ global symmetry leads to charge conservation via Noether's Theorem. Gauge symmetry is local and leads, additionally, to massless photons (see below).

We begin with the ungauged theory with a complex scalar field $\phi(x)$:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\overline{\phi}\partial^\mu\phi - U(\overline{\phi}\phi)$$

where U is a function (usually a polynomial) of $|\phi|^2 = \overline{\phi}\phi$. We use the convention $\eta_{\mu\nu} = \mathrm{diag}(+1, -1, -1, -1)$ for the Minkowski metric. So $\partial_\mu\overline{\phi}\partial^\mu\phi = \partial_0\overline{\phi}\partial_0\phi - \boldsymbol{\nabla}\overline{\phi}\cdot\boldsymbol{\nabla}\phi$. This convention gives us a positive kinetic term (involving time derivatives) in $\mathcal{L}$.

$\mathcal{L}$ is Lorentz invariant and invariant under the global $U(1)$ symmetry $\phi \mapsto e^{i\alpha}\phi$, $\overline{\phi} \mapsto e^{-i\alpha}\overline{\phi}$. To obtain a $U(1)$ gauge theory, $\mathcal{L}$ must also be invariant under

$$\phi(x) \mapsto e^{i\alpha(x)}\phi(x) \tag{7.1}$$

where $\alpha(x)$ is an arbitrary, smooth real function. The $U(\overline{\phi}\phi)$ term is gauge invariant as it stands, but the derivative terms cause a problem. The way to solve this is to introduce a new field, the *real gauge potential* $a_\mu(x)$, and the *gauge covariant derivative* of ϕ:

$$D_\mu\phi = \partial_\mu\phi - ia_\mu\phi$$

(In this definition of D_μ we have set the coupling to unity.) We *postulate* that under the gauge transformation (7.1) the gauge potential transforms according to:

$$a_\mu(x) \mapsto a_\mu(x) + \partial_\mu\alpha(x)$$

With these definitions, we see that $D_\mu\phi$ transforms in the same way as ϕ ("covariantly" with ϕ):

$$\begin{aligned} D_\mu\phi &\mapsto \partial_\mu(e^{i\alpha}\phi) - i(a_\mu + \partial_\mu\alpha)(e^{i\alpha}\phi) \\ &= i(\partial_\mu\alpha)e^{i\alpha}\phi + e^{i\alpha}\partial_\mu\phi - ia_\mu e^{i\alpha}\phi - i(\partial_\mu\alpha)e^{i\alpha}\phi \\ &= e^{i\alpha}(\partial_\mu\phi - ia_\mu\phi) \\ &= e^{i\alpha}D_\mu\phi \end{aligned}$$

Similarly, the covariant derivative of $\overline{\phi}$ is $D_\mu\overline{\phi} = \overline{D_\mu\phi} = \partial_\mu\overline{\phi} + ia_\mu\overline{\phi}$, and so

$$\overline{D_\mu\phi} \mapsto e^{-i\alpha}\overline{D_\mu\phi}$$

under the gauge transformation. With these modifications, $\overline{D_\mu\phi}D^\mu\phi$ is gauge invariant and Lorentz invariant.
The field a_μ is also dynamical, and so we need its derivatives in $\mathcal{L}$. The electromagnetic field tensor $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$ appears in $\mathcal{L}$. We can easily verify its gauge invariance:

$$\begin{aligned} f_{\mu\nu} &\mapsto \partial_\mu(a_\nu + \partial_\nu\alpha) - \partial_\nu(a_\mu + \partial_\mu\alpha) \\ &= \partial_\mu a_\nu - \partial_\nu a_\mu \text{ , by symmetry of partial derivatives} \\ &= f_{\mu\nu} \end{aligned}$$

We now combine these ingredients to construct a Lagrangian density for scalar electrodynamics:

$$\mathcal{L} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu} + \frac{1}{2}\overline{D_\mu\phi}D^\mu\phi - U(\overline{\phi}\phi).$$

In order to understand the signs, we separate this equation into time and space parts. Recall: $e_i = f_{0i}$ so $\vec{e} = \partial_0\vec{a} - \boldsymbol{\nabla}a_0$, and $b_k = \frac{1}{2}\epsilon_{ijk}f_{ij}$ so $\vec{b} = \boldsymbol{\nabla}\times\vec{a}$, where $\vec{e}$ and $\vec{b}$ are the electric and magnetic fields. Using this notation, we find

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4}(-2\vec{e}\cdot\vec{e} + 2\vec{b}\cdot\vec{b}) + \frac{1}{2}\overline{D_0\phi}D_0\phi - \frac{1}{2}\overline{\vec{D}\phi}\cdot\vec{D}\phi - U(\overline{\phi}\phi) \\ &= \underbrace{\frac{1}{2}\vec{e}\cdot\vec{e} + \frac{1}{2}\overline{D_0\phi}D_0\phi}_{\text{“Kinetic Terms”}} \underbrace{-\frac{1}{2}\vec{b}\cdot\vec{b} - \frac{1}{2}\overline{\vec{D}\phi}\cdot\vec{D}\phi - U(\overline{\phi}\phi)}_{\text{“Potential Terms”}} \end{aligned}$$

This is a "natural Lagrangian density" in the sense that it takes the form $\mathcal{L} = T - V$, where T is quadratic in time derivatives. However, a_0 is *not* dynamical. For completeness, we give the Euler-Lagrange equations, though their derivation is not a part of this course:

$$\begin{aligned} D_\mu D^\mu\phi &= -2U'(\overline{\phi}\phi)\phi \\ \partial_\mu f^{\mu\nu} &= -\frac{i}{2}(\overline{\phi}D^\nu\phi - \phi\overline{D^\nu\phi}) \end{aligned}$$

where U' is the derivative of U with respect to its argument $\overline{\phi}\phi$.

Note: These are second-order evolution equations for $\phi, \vec{a}$. The $\nu = 0$ component is rather different (Gauss' law...).

7.1.1 Field Tensor from Covariant Derivatives

$$\begin{aligned}
[D_\mu, D_\nu]\phi &= (\partial_\mu - ia_\mu)(\partial_\nu - ia_\nu)\phi - (\mu \leftrightarrow \nu) \\
&= (\partial_\mu - ia_\mu)(\partial_\nu\phi - ia_\nu\phi) - (\mu \leftrightarrow \nu) \\
&= \cancel{\partial_\mu\partial_\nu\phi} - i(\partial_\mu a_\nu)\phi - \cancel{ia_\nu\partial_\mu\phi} - \cancel{ia_\mu\partial_\nu\phi} - \cancel{a_\mu a_\nu\phi} - (\mu \leftrightarrow \nu) \\
&= -i(\partial_\mu a_\nu - \partial_\nu a_\mu)\phi \\
&= -if_{\mu\nu}\phi
\end{aligned}$$

So we see that the field tensor arises as a commutator of covariant derivatives.

7.2 A Non-Abelian Gauge Theory: Scalar Yang-Mills Theory

We now extend this formalism to a general gauge group G. We fix $G = U(n)$ or some Lie subgroup of $U(n)$ (e.g., $SU(n), SO(n), \dots$). We introduce the fundamental scalar field

$$\Phi(x) = \begin{pmatrix} \Phi_1(x) \\ \vdots \\ \Phi_n(x) \end{pmatrix}$$

which has n complex components. As before, we can consider the global action of G

$$\Phi(x) \mapsto g\Phi(x)$$

where $g \in G$ is independent of x. However, we will require our theory to be invariant under gauge transformations

$$\Phi(x) \mapsto g(x)\Phi(x)$$

where $g \in G$ *depends on spacetime location.* As previously, this inspires us to introduce the covariant derivative

$$D_\mu\Phi = (\partial_\mu + A_\mu)\Phi,$$

with gauge potential $A_\mu(x) \in L(G)$. In particular, A_μ (as a member of the Lie algebra of $U(n)$) is an $n \times n$ antihermitian matrix. Although there are certain *geometric motivations*, we *postulate* that A_μ gauge transforms as:

$$A_\mu \mapsto gA_\mu g^{-1} - (\partial_\mu g)g^{-1}$$

Remark. Earlier (see the discussion leading to (3.2)) we discussed why expressions of the form $g^{-1}(\partial_\mu g)$ and $(\partial_\mu g)g^{-1}$ are elements of the Lie algebra $L(G)$. In our discussion of the adjoint representation of G, we also found that $(\mathrm{Ad}\, g)X = gXg^{-1} \in L(G)$ where $g \in G$ and $X \in L(G)$. So A_μ remains in $L(G)$ after a gauge transformation.

Remark. This definition for the transformation of the gauge potential A_μ is consistent with our previous postulate for the transformation law in the abelian case of scalar electrodynamics. There, $-ia_\mu \mapsto g(-ia_\mu)g^{-1} - (\partial_\mu g)g^{-1} = -ia_\mu - i\partial_\mu\alpha$ when $g(x) = e^{i\alpha(x)}$, so $a_\mu \mapsto a_\mu + \partial_\mu\alpha$.

We see that $D_\mu\Phi$ transforms covariantly with Φ:

$$\begin{aligned}
D_\mu\Phi = (\partial_\mu + A_\mu)\Phi &\mapsto (\partial_\mu + gA_\mu g^{-1} - (\partial_\mu g)g^{-1})g\Phi \\
&= \cancel{(\partial_\mu g)\Phi} + g(\partial_\mu\Phi) + gA_\mu\Phi - \cancel{(\partial_\mu g)\Phi} \\
&= g(\partial_\mu\Phi + A_\mu\Phi) \\
&= gD_\mu\Phi
\end{aligned}$$

We now look for the *Yang-Mills field tensor* by examining the commutator of covariant derivatives:

$$\begin{aligned}
[D_\mu, D_\nu]\Phi &= (\partial_\mu + A_\mu)(\partial_\nu + A_\nu)\Phi - (\mu \leftrightarrow \nu) \\
&= \partial_\mu\partial_\nu\Phi + A_\mu\partial_\nu\Phi + \partial_\mu(A_\nu\Phi) + A_\mu A_\nu\Phi - (\mu \leftrightarrow \nu) \\
&= \cancel{\partial_\mu\partial_\nu\Phi} + \cancel{A_\mu\partial_\nu\Phi} + \partial_\mu A_\nu\Phi + \cancel{A_\nu\partial_\mu\Phi} + A_\mu A_\nu\Phi - (\mu \leftrightarrow \nu) \\
&= (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])\Phi \\
&= F_{\mu\nu}\Phi
\end{aligned}$$

so we define $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ as the field tensor in the non-abelian case.

Note: $F_{\mu\nu} \in L(G)$ and $F_{\mu\nu} = -F_{\nu\mu}$. The final term in $F_{\mu\nu}$ can be interpreted abstractly as the Lie bracket and not simply as the matrix commutator.

Proposition 3. *Under a gauge transformation $g(x)$, $F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1}$.*

Proof. $[D_\mu, D_\nu]\Phi \mapsto g[D_\mu, D_\nu]\Phi = gF_{\mu\nu}\Phi = (gF_{\mu\nu}g^{-1})(g\Phi)$. Thus we see that $F_{\mu\nu} \mapsto gF_{\mu\nu}g^{-1}$. □

Remark. One can also compute this directly by gauge transforming A_μ, but the calculation is longer.

7.2.1 Lagrangian Density

For the Yang-Mills gauge potential A_μ coupled to the scalar Φ, the Lagrangian density is given by:

$$\mathcal{L} = \frac{1}{4}\operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) + \frac{1}{2}(D_\mu\Phi)^\dagger D^\mu\Phi - U(\Phi^\dagger\Phi).$$

Lorentz invariance of $\mathcal{L}$ is clear, since all the Lorentz indices are contracted using the Minkowski metric. However we still need to check gauge invariance:

$$\begin{aligned}
\Phi &\mapsto g\Phi \\
\Phi^\dagger &\mapsto \Phi^\dagger g^\dagger = \Phi^\dagger g^{-1}\text{, since we're in } U(n) \\
\therefore\ \Phi^\dagger\Phi &\mapsto \Phi^\dagger g^{-1} g\Phi = \Phi^\dagger\Phi,
\end{aligned}$$

so $U(\Phi^\dagger\Phi)$ is invariant. Since $D_\mu\Phi$ transforms as Φ, the same argument applies to the second term in $\mathcal{L}$. For the first term we see:

$$\begin{aligned}
\operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) &\mapsto \operatorname{Tr}(gF_{\mu\nu}g^{-1}gF^{\mu\nu}g^{-1}) \\
&= \operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}g^{-1}g)\text{ , since Tr is invariant under cyclic permutations} \\
&= \operatorname{Tr}(F_{\mu\nu}F^{\mu\nu})
\end{aligned}$$

Thus, the Lagrangian density is gauge invariant, and so is the action.

From the covariant derivative $D_\mu\Phi = (\partial_\mu + A_\mu)\Phi$ and the field tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$ we see that the kinetic terms in the Lagrangian density are:

$$-\frac{1}{2}\operatorname{Tr}(F_{0i}F_{0i}) + \frac{1}{2}(D_0\Phi)^\dagger D_0\Phi$$

As F_{0i} is an antihermitian matrix, $-\frac{1}{2}\operatorname{Tr}(F_{0i}F_{0i})$ is positive. Why is this true? If X is antihermitian,

$$\operatorname{Tr} X^2 = X_{\alpha\beta}X_{\beta\alpha} = -X_{\alpha\beta}X^*_{\alpha\beta} = -\sum_{\alpha,\beta}|X_{\alpha\beta}|^2$$

Hence all kinetic terms are positive. This is a consequence of the gauge group G being unitary (so elements of its Lie algebra are antihermitian).

7.2.2 Adjoint Covariant Derivative

Not all fields transform as $\Phi \mapsto g\Phi$ under gauge transformations. We could also have an adjoint scalar field $\Psi \in L(G)$ transforming as $\Psi \mapsto g\Psi g^{-1}$ (This looks like the transformation of $F_{\mu\nu}$.) In this case, the covariant derivative is:

$$D_\mu \Psi = \partial_\mu \Psi + [A_\mu, \Psi]$$

To check this, one substitutes the transformation laws $\Psi \mapsto g\Psi g^{-1}$ and $A_\mu \mapsto gA_\mu g^{-1} - (\partial_\mu g)g^{-1}$ for the scalar field and gauge potential, respectively, into the proposed formula. One then sees after a few lines of computation that the covariant derivative transforms as Ψ, i.e., $D_\mu \Psi \mapsto g(D_\mu \Psi)g^{-1}$.

7.2.3 General Covariant Derivative

Abstractly, we write $D_\mu = \partial_\mu + A_\mu$, where $A_\mu \in L(G)$ and $[D_\mu, D_\nu] = F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$. In order to act concretely, we need a field $\Phi \in V$ (V is a vector space) and a representation $\mathcal{D}$ of G acting on V. (Note that since D in this chapter denotes a covariant derivative we use $\mathcal{D}$ for a representation.) Then under gauge transformations the field transforms as $\Phi \mapsto \mathcal{D}(g)\Phi$ and $D_\mu \Phi = \partial_\mu \Phi + d(A_\mu)\Phi$, where d is the representation of $L(G)$ associated to $\mathcal{D}$.

Claim: $D_\mu \Phi$ transforms like Φ under gauge transformations.

Proof. We know $\Phi \mapsto \mathcal{D}(g)\Phi$ and $d(A_\mu) \mapsto \mathcal{D}(g)d(A_\mu)\mathcal{D}(g^{-1}) - (\partial_\mu \mathcal{D}(g))\mathcal{D}(g^{-1})$ Thus:

$$\begin{aligned} D_\mu \Phi = \partial_\mu \Phi + d(A_\mu)\Phi &\mapsto \partial_\mu(\mathcal{D}(g)\Phi) + (\mathcal{D}(g)d(A_\mu)\mathcal{D}(g^{-1}) - (\partial_\mu \mathcal{D}(g))\mathcal{D}(g^{-1}))\mathcal{D}(g)\Phi \\ &= \cancel{(\partial_\mu \mathcal{D}(g))\Phi} + \mathcal{D}(g)(\partial_\mu \Phi) + \mathcal{D}(g)d(A_\mu)\Phi - \cancel{(\partial_\mu \mathcal{D}(g))\Phi} \\ &= \mathcal{D}(g)(\partial_\mu \Phi + d(A_\mu)\Phi) \\ &= \mathcal{D}(g)D_\mu \Phi \end{aligned}$$

Thus $D_\mu \Phi$ transforms covariantly with Φ under gauge transformations. □

7.2.4 The Field Equation of Pure Yang-Mills Theory

The field equation for pure Yang-Mills theory is:

$$\partial_\mu F^{\mu\nu} + [A_\mu, F^{\mu\nu}] = 0$$

Remark: This is an equation for A_μ involving second-order derivatives. In Yang-Mills theory, $F^{\mu\nu}$ is a derived quantity, while A_μ is considered fundamental (contrast with electromagnetism). We may rewrite the field equation more compactly using the adjoint covariant derivative:

$$D_\mu F^{\mu\nu} = 0$$

7.2.5 Classical Vacuum

Naively, the classical vacuum is $A_\mu = 0$. We can gauge transform this to $A_\mu = -(\partial_\mu g)g^{-1}$. This is a "pure gauge," which is physically the same as the vacuum. The Yang-Mills field tensor vanishes in both the classical vacuum and the pure gauge transformed vacuum: $F_{\mu\nu} = 0$ for both $A_\mu = 0$ and $A_\mu = -(\partial_\mu g)g^{-1}$.

Proof.

$$\begin{aligned}
F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \\
&= \partial_\mu(-(\partial_\nu g)g^{-1}) - \partial_\nu(-(\partial_\mu g)g^{-1}) + [-(\partial_\mu g)g^{-1}, -(\partial_\nu g)g^{-1}] \\
&= \cancel{-(\partial_\mu\partial_\nu g)g^{-1}} - (\partial_\nu g)(\partial_\mu g^{-1}) + \cancel{(\partial_\nu\partial_\mu g)g^{-1}} + (\partial_\mu g)(\partial_\nu g^{-1}) + (\partial_\mu g)g^{-1}(\partial_\nu g)g^{-1} - (\partial_\nu g)g^{-1}(\partial_\mu g)g^{-1} \\
&= (\partial_\mu g)[(\partial_\nu g^{-1}) + g^{-1}(\partial_\nu g)g^{-1}] - (\partial_\nu g)[(\partial_\mu g^{-1}) + g^{-1}(\partial_\mu g)g^{-1}] \\
&= (\partial_\mu g)[\partial_\nu(g^{-1}g)]g^{-1} - (\partial_\nu g)[\partial_\mu(g^{-1}g)]g^{-1}
\end{aligned}$$

But $g^{-1}g = I = \text{const}$, so $\partial_\nu(g^{-1}g) = 0$. Thus $F_{\mu\nu} = 0$ in the pure gauge $A_\mu = -(\partial_\mu g)g^{-1}$. □

If there is a fundamental scalar field Φ as well, then in the vacuum $\Phi = 0$ (provided $\Phi = 0$ minimizes the potential $U(\Phi^\dagger\Phi)$), and $D_\mu\Phi = 0$.

7.3 A Very Brief Introduction to Mass and the Higgs Mechanism

Recall the Klein-Gordon theory for a scalar field ϕ is governed by the Lagrangian density:

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}m^2\phi^2$$

As shown in quantum field theory, the equation of motion is the Klein-Gordon equation:

$$\partial_\mu\partial^\mu\phi + m^2\phi = 0$$

This equation has wavelike solutions of the form $\phi = \exp[i(\omega t - \vec{k}\cdot\vec{x})]$ provided $\omega^2 = |\vec{k}|^2 + m^2$. Here m is the "mass parameter" of the field. Particles arise when the field is quantized. Particle energy is given by $E = \hbar\omega$; particle momentum is given by $\vec{p} = \hbar\vec{k}$. Thus $E^2 = |\vec{p}|^2 + (\hbar m)^2$. So in units where $c = 1$, the particle has mass $\hbar m$. In units where also $\hbar = 1$ the mass is m.

7.3.1 Electrodynamics

We have the Lagrangian density $\mathcal{L} = -\frac{1}{4}f_{\mu\nu}f^{\mu\nu}$ with $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$. The Maxwell equation says:

$$\partial_\mu f^{\mu\nu} = 0 \implies \partial_\mu(\partial^\mu a^\nu) - \partial^\nu(\partial_\mu a^\mu) = 0$$

If we work in the Coulomb gauge, $\partial_i a_i = 0, i = 1,2,3$, i.e., $\boldsymbol{\nabla}\cdot\vec{a} = 0$, then $\partial_\mu a^\mu = \partial_0 a_0$. Let us take a closer look at the $\nu = 0$ equation $\partial_\mu\partial^\mu a^0 - \partial^0(\partial_0 a_0) = (\partial_0\partial_0 - \partial_i\partial_i)a_0 - \partial_0\partial_0 a_0 = 0 \implies \boldsymbol{\nabla}^2 a_0 = 0$. Without any special boundary conditions or sources, the solution is evidently $a_0 = 0$. The remaining equations are $(\partial_\mu\partial^\mu)a_i = 0$. These are *three massless wave equations.*

We see that polarization is transverse to momentum: If $a_i = \epsilon_i\exp[i(\omega t - \vec{k}\cdot\vec{x})]$, then $\omega^2 = |\vec{k}|^2$ (massless!), and the gauge condition $\partial_i a_i = 0$ tells us $\vec{k}\cdot\vec{\epsilon} = 0$. After quantization, the *photon* has the following properties:

- zero mass, so energy satisfies $E = |\vec{p}|$
- 2 (transverse) polarization states: $\vec{\epsilon} \perp \vec{p}$
- non-vanishing momentum: $\vec{p} \neq 0$

The last condition is because the $\omega = \vec{k} = 0$ wave, $a_i = \epsilon_i$, is pure gauge and has vanishing field tensor. So it is the vacuum.

The quantized particles of pure Yang-Mills theory are also massless, but confined. They are called gluons in the $SU(3)$ theory of quantum chromodynamics. This phenomenon is only partially understood.

7.3.2 Perturbative Effect of Interaction of EM Field with a Charged Scalar Field

The scalar particle mass m is renormalized, thereby changing its value. The photon, however, remains massless. This result is complicated to prove. However, it may be understood as arising from:

(a) No term of the form $\frac{1}{2}M^2 a_\mu a^\mu$ is allowed, because it wouldn't be gauge invariant.

(b) A massive vector particle can have $\vec{p} = 0$, and it then has three polarization states (i.e., $\vec{\epsilon} \cdot \vec{p} = 0$ doesn't constrain $\vec{\epsilon}$ if $\vec{p} = 0$). However a continuous perturbation cannot change the number of polarization states.

7.3.3 Higgs Mechanism

The *Higgs mechanism* evades these two previous arguments. In scalar electrodynamics, the "photon" can become massive. The abelian Higgs model has the Lagrangian density:

$$\mathcal{L} = -\frac{1}{4} f_{\mu\nu} f^{\mu\nu} + \frac{1}{2}\overline{D_\mu \phi} D^\mu \phi + \frac{1}{2}\mu^2 \overline{\phi}\phi - \frac{1}{4}\lambda(\overline{\phi}\phi)^2 \tag{7.2}$$

Evidently, the potential is $U = -\frac{1}{2}\mu^2\overline{\phi}\phi + \frac{1}{4}\lambda(\overline{\phi}\phi)^2 = -\frac{1}{2}\mu^2|\phi|^2 + \frac{1}{4}\lambda|\phi|^4$.

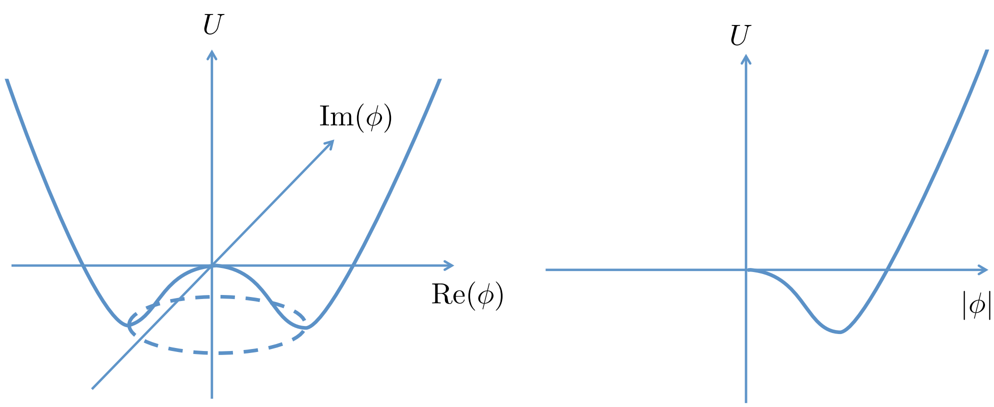


Figure 7.1: The so-called "Mexican hat potential" (before shifting it up)

Often it is convenient to shift the potential up and write $U = \frac{1}{4}\lambda(|\phi|^2 - v^2)^2$, where $\mu^2 = \lambda v^2$. This shift is by a constant value of $\frac{1}{4}\lambda v^4$, which has no effect on the field equations. The vacuum is based around the classical field that minimizes the energy, and in particular minimizes U. It is *not* at $\phi = 0$, but where $|\phi| = v$. Further, the vacuum is *degenerate* (related by gauge transformations); it is a *non-trivial orbit* of the gauge group $U(1)$. We call v the *vacuum expectation value (vev)* of ϕ (or more accurately of $|\phi|$). The simplest vacuum is $\phi = v, a_\mu = 0$. We can gauge transform this vacuum to $\phi = v \exp i\alpha(x), a_\mu = \partial_\mu \alpha$. A natural question is then how this affects the two arguments above that kept the photon massless.

(a') There is now a mass term for a_μ:

Part of $\frac{1}{2}\overline{D_\mu \phi} D^\mu \phi$ is $\frac{1}{2}(\overline{-ia_\mu \phi})(-ia^\mu \phi) = \frac{1}{2}a_\mu a^\mu |\phi|^2$. Close to the vacuum, this is $\frac{1}{2}v^2 a_\mu a^\mu +$ higher order terms. Thus the "photon" acquires mass v and gets the new name *massive vector boson.*

Remark: The term we've singled out would, by itself, violate gauge invariance. However, the other terms present in the Lagrangian combine to preserve the gauge invariance.

(b') The massive vector boson has *three* polarization states. We can impose the gauge condition

$$\text{Im } \phi = 0, \ \phi(x) = v + \eta(x)$$

where $\eta(x)$ is real. We cannot simultaneously impose the Coulomb gauge. Thus the physical particles are:

- a massive vector boson with three states, and
- a real scalar particle H from the field η with one state.

Alternatively, in the Coulomb gauge, we need to allow $\phi(x) = \exp[i\beta(x)](v + \eta(x))$ with $\beta(x), \eta(x)$ real. The third (longitudinal) polarization state of the vector boson now comes from β. For the *mass of the Higgs particle H*, look at the quadratic terms in η (with $a_\mu = 0$).

$$\mathcal{L}_\eta = \frac{1}{2}\partial_\mu \eta \partial^\mu \eta - \frac{1}{4}\lambda((v+\eta)^2 - v^2)^2 = \frac{1}{2}\partial_\mu \eta \partial^\mu \eta - \lambda v^2 \eta^2 + \dots$$

which we find from substituting $\phi(x) = v + \eta(x)$ into $\mathcal{L}$. Thus we see that $m_H = \sqrt{2\lambda} v = \sqrt{2}\mu$.

7.3.4 Higgs Mechanism in the Non-abelian Case

The Standard Model has electroweak gauge group $G = U(2)$ and a scalar field Φ, which is a doublet of complex scalar fields. $U(2)$ acts by the fundamental representation. The Higgs mechanism occurs if $\Phi \neq 0$ in the classical vacuum, i.e., if the minimum of the potential $U(\Phi^\dagger \Phi)$ is on a *non-trivial orbit* of the gauge group G. We expect the orbit to have the form G/K for some subgroup $K \leq G$. The gauge symmetry is still present, but the choice of a vev for Φ (which is just a choice of gauge) apparently breaks the G-symmetry. One says that the G-symmetry is spontaneously broken, while the K-symmetry (which leaves the vev fixed) remains unbroken. Massless gauge bosons remain for K. In the Standard Model, $G = U(2), K = U(1)$, and as $U(2)$ is a 4-dimensional Lie group, we get:

- 3 massive vector bosons W^+, W^-, Z
- 1 massless vector boson, the photon
- 1 real (electrically neutral) Higgs boson

The masses of the vector bosons and Higgs particle are unrelated, so the Higgs mass had to be determined experimentally.

Chapter 8

Quadratic Forms on Lie Algebras and the Geometry of Lie Groups

8.1 Invariant Quadratic Forms

For a matrix Lie algebra $L(G)$, we have the quadratic form $L(G) \times L(G) \to \mathbb{C}$ (symmetric inner product) given by:

$$(X, Y) = \mathrm{Tr}(XY)$$

where $X, Y \in L(G)$. This quadratic form is G-invariant in the following sense:

$$(X, Y) \mapsto (gXg^{-1}, gYg^{-1}) = \mathrm{Tr}(gXg^{-1}gYg^{-1}) = \mathrm{Tr}(XY) = (X, Y)$$

where $g \in G$. One can also use a different representation d of $L(G)$ and define $(X, Y)_d = \mathrm{Tr}(d(X)\, d(Y))$, which is similarly invariant: $(gXg^{-1}, gYg^{-1})_d = (X, Y)_d$. This is a result of the following lemma:

Lemma 8.1.1. *If d is a representation of $L(G)$ associated to the representation D of G, then $d(gXg^{-1}) = D(g)d(X)D(g)^{-1}$.*

Proof. For $g' \in G$

$$D(gg'g^{-1}) = D(g)D(g')D(g)^{-1}$$

Now set $g' = I + tX$ with t small, and ignore terms of order t^2. Then

$$\begin{aligned} D(gg'g^{-1}) &= D(g(I + tX)g^{-1}) \\ &= D(I + tgXg^{-1}) \\ &= I + td(gXg^{-1}) \end{aligned}$$

whereas

$$\begin{aligned} D(g)D(g')D(g)^{-1} &= D(g)(I + td(X))D(g)^{-1} \\ &= I + tD(g)d(X)D(g)^{-1} \end{aligned}$$

Comparing the terms of order t gives the result. $\square$

As a special case, we consider the adjoint representation of $L(G)$, which gives rise to the so-called *Killing form*:

$$\kappa(X,Y) = \mathrm{Tr}(\mathrm{ad}\,X\ \mathrm{ad}\,Y)$$

The Killing form is real and is the most basic inner product, because every Lie algebra has an adjoint representation. In practice, these inner products are usually related. They differ by a constant multiple if $L(G)$ is "simple" (to be made precise soon).

The infinitesimal version of the invariance above is the *associativity property*:

Claim: $(X,[Y,Z])_d = ([X,Y],Z)_d$

Proof. (For the initial inner product in this section)
$\mathrm{Tr}(X,[Y,Z]) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(XZY) = \mathrm{Tr}(XYZ) - \mathrm{Tr}(YXZ) = \mathrm{Tr}([X,Y]Z)$ □

What does this have to do with an infinitesimal version of the invariance under the conjugation $X \mapsto gXg^{-1}$? If we let $g = I + tZ \in G$ with t infinitesimal, then we have to order t:

$$\begin{aligned}
\mathrm{Tr}(XY) &= \mathrm{Tr}((1+tZ)X(1-tZ)(1+tZ)Y(1-tZ)) \\
&= \mathrm{Tr}((X - t[X,Z])(Y + t[Z,Y])) \\
&= \mathrm{Tr}(XY) - t\,\mathrm{Tr}([X,Z]Y) + t\,\mathrm{Tr}(X[Z,Y]) \\
\implies& \mathrm{Tr}([X,Z]Y) = \mathrm{Tr}(X[Z,Y])
\end{aligned}$$

which is just the associativity property.

8.2 Non-Degeneracy of the Killing Form

Let $\{T_i\}$ be a basis for $L(G)$. The Killing form becomes a symmetric matrix with components

$$\kappa_{ij} = \kappa(T_i,T_j) = \mathrm{Tr}(\mathrm{ad}\,T_i\ \mathrm{ad}\,T_j)$$

Now $\mathrm{ad}\,T_i$ is defined by $(\mathrm{ad}\,T_i)X = [T_i, X]$, where $X \in L(G)$. Let $X = X_l T_l$ (basis expansion). Then

$$(\mathrm{ad}\,T_i)X = [T_i, X_l T_l] = X_l[T_i,T_l] = c_{ilk}X_l T_k \implies (\mathrm{ad}\,T_i)_{kl} = c_{ilk}$$

Although the indices look a bit strange in the final equality of the last line, they are in fact correct. Therefore we have $\mathrm{Tr}(\mathrm{ad}\,T_i\ \mathrm{ad}\,T_j) = (\mathrm{ad}\,T_i)_{kl}(\mathrm{ad}\,T_j)_{lk} = c_{ilk}c_{jkl}$. So the $(ij)^{\text{th}}$ component of the Killing form is given by:

$$\kappa_{ij} = c_{ilk}c_{jkl}$$

Example. For the standard basis of $\mathfrak{su}(2)$, $\kappa_{ab} = \epsilon_{adc}\epsilon_{bcd} = -2\delta_{ab}$. This is a non-degenerate Killing form.

Definition 28. The Killing form κ_{ij} is *non-degenerate* if all of its eigenvalues are non-zero. Equivalently, $\det \kappa_{ij} \neq 0$ and κ_{ij} is invertible, since $\det \kappa_{ij} = \prod \lambda_i$, where λ_i are the eigenvalues.

Definition 29. A Lie algebra $L(G)$ is said to be *semi-simple* if its Killing form is non-degenerate.

Theorem 8.2.1. *A semi-simple Lie algebra $L(G)$ has a decomposition into mutually commuting simple factors $L(G_i)$ such that*

$$L(G) = L(G_1) \oplus L(G_2) \oplus \cdots \oplus L(G_k)$$

and $L(G)$ cannot be reduced further.

Proof. Omitted (but see Example Sheet 4, Problem 6 for some related details). The general idea is that a non-degenerate Killing form gives a way to build "orthogonal complements," which leads to a direct sum decomposition of the Lie algebra. □

Fact: In the case of $L(G)$ semi-simple, the associated group G has the structure

$$G = G_1 \times G_2 \times \cdots \times G_k/\{\text{discrete group}\}$$

where G_i are simple Lie groups. (Note: A simple Lie group is a connected, non-abelian Lie group with no (proper) normal Lie subgroups.)

8.3 Compactness

If κ_{ij} is *negative definite*, $L(G)$ is said to be of *compact type*. G is then a *compact group* (as a topological manifold). For example, $SU(n)$ is simple and compact. In the case of $L(G)$ of compact type, we can find an *adapted basis* $\{T_i\}$ such that $\kappa_{ij} = -\mu\delta_{ij}$, for some constant $\mu > 0$.

Then $\kappa(T_i, [T_j, T_k]) = \kappa(T_i, c_{jkl}T_l) = c_{jkl}\kappa_{il} = -\mu c_{jki}$, and $\kappa([T_i, T_j], T_k) = \kappa(c_{ijl}T_l, T_k) = c_{ijl}\kappa_{lk} = -\mu c_{ijk}$. Thus by the associativity property we have:

$$c_{ijk} = c_{jki} = c_{kij} = -c_{jik} = -c_{kji} = -c_{ikj}$$

In other words, *the structure constants are totally antisymmetric in the adapted basis.*

We can now clarify why gauge groups are chosen to be compact. If $L(G)$ is of compact type then the Yang-Mills Lagrangian density

$$\mathcal{L}_{\text{YM}} = \kappa(F_{\mu\nu}F^{\mu\nu})$$

is gauge invariant because of the G-invariance of κ, and has positive kinetic terms (remember the extra minus sign from lowering Lorentz indices). For a simple matrix group G, this density is a positive multiple of $\text{Tr}(F_{\mu\nu}F^{\mu\nu})$. (Non-compact groups have Killing forms κ of mixed signature, which would lead to unphysical negative kinetic energy in some cases; as κ is not positive definite, we cannot just flip the sign.)

8.4 Universal Enveloping Algebra

In a matrix Lie algebra $L(G)$, we can deal with a product XY (which isn't necessarily in $L(G)$) as well as the commutator $[X, Y]$ (which *is* in $L(G)$). In an abstract Lie algebra L, we can define the *universal enveloping algebra* (UEA) to be the formal span of $\{1, L, L \otimes L, L \otimes L \otimes L, \dots\}$, allowing sums and products of elements of L. The UEA is subject to the one rule: $XY - YX = [X, Y]$, as defined in L.

(a) In the UEA we see that there are new identities, e.g.

$$\begin{aligned}[X, YZ] &= XYZ - YZX = XYZ - YXZ + YXZ - YZX \\ &= [X, Y]Z + Y[X, Z]\end{aligned}$$

(b) We have

$$\exp X = 1 + X + \frac{1}{2}X^2 + \cdots \in \text{UEA}$$

Thus the connected component of G is in the UEA of $L(G)$. (Note: This is a bit of a white lie, as the UEA technically only contains *finite* sums, while the definition of exp involves infinite sums. However, we assume that the idea may be extended in a more careful treatment to include a "formal completion of the UEA".)

(c) There is an analogy between a UEA and the Dirac matrix algebra, which is the span of $\{I, \gamma^\mu, \gamma^\mu\gamma^\nu, \gamma^\mu\gamma^\nu\gamma^\sigma, ...\}$ subject to the constraint $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu} I$.

One can extend an N-dimensional representation d of L to the UEA in the obvious way (replace X by $d(X)$ and 1 by I_N). All elements are now $N \times N$ matrices. Usually there are additional identities satisfied by the matrices $d(X)$ which are not universal. It is important not to be deceived by these. E.g. $(\sigma_1)^2 = I_2$ for the Pauli matrix σ_1 does not imply that $(T_1)^2$ is a multiple of 1 in the UEA of $L(SU(2))$.

8.5 Casimir Elements

Let $L(G)$ be a Lie algebra of simple (irreducible in the sense that it isn't a direct sum of mutually commuting subalgebras) compact (κ_{ij} negative definite) type with adapted basis $\{T_i\}$ such that $\kappa_{ij} = -\delta_{ij}$. Here the basis elements are normalized so that $\mu = 1$.

Definition 30. The universal (quadratic) *Casimir element* in the UEA is $C = \sum_i T_i T_i$.

Lemma 8.5.1. $[X, C] = 0$ *for all* $X \in L(G)$*, i.e., the Casimir commutes with every element in the Lie algebra:*

Proof. It is sufficient to set X equal to the basis element T_j.

$$\begin{aligned}
[T_j, C] &= \sum_i [T_j, T_i T_i] \\
&= \sum_i (T_i [T_j, T_i] + [T_j, T_i] T_i) \\
&= \sum_i (T_i c_{jik} T_k + c_{jik} T_k T_i) \\
&= \sum_i c_{jik} (T_i T_k + T_k T_i) = 0
\end{aligned}$$

where the last line follows because the antisymmetric structure constants (adapted basis) are contracted against a symmetric quantity. □

It follows that the Casimir C is a central element (commutes with everything) in the UEA. However, this does not mean that it is a fixed multiple of 1. (The UEA has as further central elements all polynomials in C. It may have yet more independent central elements, higher Casimirs, but these depend on the detailed form of $L(G)$ and do not have a universal form.)

Consider now C in a representation d of $L(G)$ of dimension N:

$$C_d = \sum_i d(T_i) d(T_i),$$

where C_d is now an $N \times N$ matrix. By the same proof as above, $[d(X), C_d] = 0, \forall X \in L(G)$.

Claim: If d is irreducible, then by Schur's lemma $C_d = c_d I_N$.

Proof. Example Sheet 3, Exercise 7 □

Here the Casimir eigenvalue c_d is a useful way to characterize the irreducible representation d.

Example. Consider the fundamental ($j = \frac{1}{2}$) representation of $\mathfrak{su}(2)$. The adapted basis is $T_a = -\frac{i}{2\sqrt{2}}\sigma_a$ (notice the factor of $\sqrt{2}$). Then:

$$C_{(j=\frac{1}{2})} = \sum_a (T_a)^2 = -\frac{1}{8}\sum_a \sigma_a^2 = -\frac{3}{8}I_2.$$

Thus $c_{(j=\frac{1}{2})} = -\frac{3}{8}$. More generally, consider the well-known angular momentum operators J_1, J_2, J_3 from physics. They have the Casimir operator:

$$C = -\frac{1}{2}\vec{J}^2 = -\frac{1}{2}(J_1^2 + J_2^2 + J_3^2)$$

Quoting a familiar result, the eigenvalue in the spin j irreducible representation is $c_{(j)} = -\frac{1}{2}j(j+1)$.

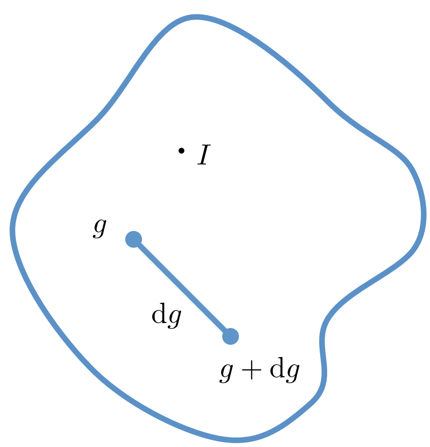
8.6 Metric on G

(Note: In what follows the assumption that G is a simple Lie group is non-essential, but with this assumption the metric we find is unique up to a constant multiple. The assumption of compactness gives a Riemannian metric. Otherwise we can wind up with a Lorentzian metric, for example.)

Suppose G is a simple matrix Lie group with Lie algebra of compact type. Let $X, Y \in L(G)$ be tangent vectors at I. Define $(X, Y) = -\operatorname{Tr}(XY)$. Up to a constant multiple, this is the unique positive definite quadratic form on $L(G)$ which is also invariant under the action of conjugation on X, Y by G. We can use this quadratic form to define a Riemannian metric on G:

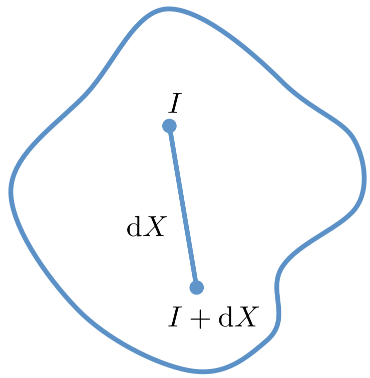
$$ds^2 = -\operatorname{Tr}\left((dg)g^{-1}(dg)g^{-1}\right) \tag{8.1}$$

where dg is an infinitesimal tangent element to G at g. By the argument used previously leading to (3.2), we see that $(dg)g^{-1} \in L(G)$. ds^2 is interpreted as the squared length of dg.



Note: We must use g^{-1} to map dg back to the identity, where we can use the Lie algebra inner product (,).

If we're near the identity, $g = I$ and $g + dg = I + dX$, where dX is an infinitesimal element of $L(G)$. Then $(dg)g^{-1} = dX$ and $ds^2 = -\operatorname{Tr}(dXdX)$.



A natural question to ask is how this metric behaves under the left and right actions of G. If

$$g \mapsto g_1 g g_2^{-1}, \qquad dg \mapsto g_1 dg g_2^{-1}$$

with g_1, g_2 constant, then $dgg^{-1} \mapsto g_1 dgg^{-1} g_1^{-1}$ and

$$ds^2 \mapsto -\operatorname{Tr}(g_1 dgg^{-1} g_1^{-1} g_1 dgg^{-1} g_1^{-1}) = -\operatorname{Tr}(dgg^{-1}dgg^{-1}) = ds^2$$

(dg transforms like g because it is the difference of two infinitesimally separated group elements.) Thus we've found that our metric is invariant under the action of $G \times G$ on G; such a metric is sometimes called a *bi-invariant metric*. This metric is highly symmetric and unique up to a scalar multiple.

Note: The high degree of symmetry is clear from the above discussion, but we haven't shown uniqueness.

8.7 Kinetic Energy and Geodesic Motion

For a particle moving on G along the trajectory $g(t)$, we can define the kinetic energy:

$$T = -\operatorname{Tr}\left(\dot{g}g^{-1}\dot{g}g^{-1}\right), \text{ with } \dot{g} \equiv dg/dt$$

and the action:

$$S = -\int_{t_0}^{t_1} \operatorname{Tr}\left(\dot{g}g^{-1}\dot{g}g^{-1}\right) dt$$

The stationarity condition $\delta S \overset{!}{=} 0$ gives $\frac{d}{dt}(g^{-1}\dot{g}) = 0$ as the equation of motion (Example Sheet 3, Problem 9). As the action is defined using the metric on G, and is purely kinetic, solutions correspond to motion along geodesics at constant speed.

If we integrate the equation of motion once with respect to time, we see:

$$g^{-1}\dot{g} = X_0 = \text{const.} \in L(G)$$

The function $g(t) = \exp(tX_0)$ is a solution with $g(0) = I$. It is easy to check this if one recalls that $\exp(\pm tX_0)$ and X_0 commute. The general solution is a left translation of this special solution given by $g(t) = g_0 \exp(tX_0)$, where $X_0 \in L(G)\,, g_0 \in G$ are both constant. We can check that this does in fact solve the differential equation:

$$\begin{aligned}
\dot{g}(t) &= g_0 X_0 \exp(tX_0), \text{ so} \\
g^{-1}\dot{g} &= (g_0 \exp(tX_0))^{-1} g_0 X_0 \exp(tX_0) \\
&= \exp(-tX_0) g_0^{-1} g_0 X_0 \exp(tX_0) \\
&= X_0
\end{aligned}$$

Remark: Geodesic motion possesses enough symmetry in this case that it is completely integrable.

8.8 $SU(2)$ Metric and Volume Form

Recall (from example sheets) the following parametrizations of $SU(2)$ as a 3-sphere:

$$\begin{aligned}
U &= \begin{pmatrix} a_0 + ia_3 & ia_1 + a_2 \\ ia_1 - a_2 & a_0 - ia_3 \end{pmatrix}, \text{ with } a_0^2 + \vec{a}\cdot\vec{a} = 1 \\
U &= \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \text{ with } |\alpha|^2 + |\beta|^2 = 1
\end{aligned}$$

(These parametrizations are really the same and in fact give a quick way to recall explicit forms for the Pauli matrices.) We know that an $SU(2)\times SU(2)$ invariant metric on $SU(2)$ is $SO(4)$ invariant. This fact follows because we showed that $\mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2)$. The bi-invariant metric on $SU(2)$ is therefore the round metric on S^3:

$$\begin{aligned}
ds^2 &= da_0^2 + d\vec{a}\cdot d\vec{a}, \text{ restricted to } S^3 \\
ds^2 &= d\alpha d\alpha^* + d\beta d\beta^*, \text{ with the restriction } |\alpha|^2 + |\beta|^2 = 1
\end{aligned}$$

Exercise: (Example Sheet 4, Problem 2) Show that $ds^2 = -\frac{1}{2}\operatorname{Tr}\left((dU)U^{-1}(dU)U^{-1}\right)$.

8.8.1 Euler Angle Parametrization

Let $\alpha = \cos\frac{\theta}{2}\exp\left(\frac{i}{2}(\phi+\psi)\right)$, $\beta = \sin\frac{\theta}{2}\exp\left(\frac{i}{2}(-\phi+\psi)\right)$, with $\theta\in[0,\pi]$, $\phi\in[0,2\pi]$, $\psi\in[0,4\pi]$. If one computes the differentials $d\alpha$ and $d\beta$, one finds after a few lines of algebra that:

$$ds^2 = d\alpha d\alpha^* + d\beta d\beta^* = \frac{1}{4}(d\theta^2 + d\phi^2 + d\psi^2 + 2\cos\theta\, d\phi d\psi)$$

This is the round metric on S^3 in terms of Euler angles (cf. Landau and Lifschitz Vol. I for more discussion). The metric tensor is therefore:

$$g_{\mu\nu} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \cos\theta \\ 0 & \cos\theta & 1 \end{pmatrix}$$

The volume form is $dV = \sqrt{\det g_{\mu\nu}}\, d\theta d\phi d\psi = \frac{1}{8}\sin\theta\, d\theta d\phi d\psi$. Using it, we can compute the total volume of $SU(2)$:

$$V = \frac{1}{8}\int \sin\theta\, d\theta d\phi d\psi = 2\pi^2$$

(Note: The volume form is useful in gauge theory quantization. Like the metric, the volume form is bi-invariant. In general a volume form on a differentiable manifold $\mathcal{M}$ is a nowhere-vanishing differential form of top degree. On a manifold $\mathcal{M}$ of $\dim n$, a volume form is an n-form, a section of the line bundle $\Omega^n(\mathcal{M}) = \Lambda^n(T^*\mathcal{M})$.)

Chapter 9

$SU(3)$ and its Representations

9.1 Roots

Recall the brackets of the $SU(2)$ Lie algebra $\mathfrak{su}(2)$:

$$\begin{aligned}[h, e_+] &= e_+ \\ [h, e_-] &= -e_- \\ [e_+, e_-] &= 2h\end{aligned}$$

where $\{e_\pm, h\}$ is a basis for (complexified) $\mathfrak{su}(2)$. Recall that $\mathfrak{su}(3)$ consists of traceless, anti-hermitian 3×3 matrices. A convenient basis (and one with similarities to the basis for $\mathfrak{su}(2)$ above) for the complexified $\mathfrak{su}(3)$ (i.e., for $\mathfrak{su}(3) \otimes \mathbb{C}$) is:

$$e_\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; e_{-\alpha} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; e_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$e_{-\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; e_\gamma = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; e_{-\gamma} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

and two in the span of

$$h_\alpha = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; h_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}, \; h_\gamma = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1/2 \end{pmatrix}$$

With these definitions, we have the following commutation relations (among others):

$$\begin{aligned}[h_\alpha, e_\alpha] &= e_\alpha \\ [h_\alpha, e_{-\alpha}] &= -e_{-\alpha} \\ [e_\alpha, e_{-\alpha}] &= 2h_\alpha\end{aligned}$$

These are precisely the commutation relations fulfilled by $\mathfrak{su}(2)$. Thus $\{e_\alpha, e_{-\alpha}, h_\alpha\}$ is the basis of an $\mathfrak{su}(2)$ subalgebra, and similarly for β, γ. However, $h_\alpha + h_\beta = h_\gamma$, so these subalgebras are not linearly independent. We can choose a basis for the diagonal matrices:

$$h_1 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; h_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

They are orthonormal in the sense: $\mathrm{Tr}({h_1}^2) = \mathrm{Tr}({h_2}^2) = 1/2$ and $\mathrm{Tr}(h_1 h_2) = 0$. We write $\vec{h} = (h_1, h_2)$ as a vector and define:

$$\begin{aligned}
\vec{\alpha} &= (1,0)\\
\vec{\beta} &= (-1/2, \sqrt{3}/2)\\
\vec{\gamma} &= (1/2, \sqrt{3}/2)
\end{aligned}$$

Using this notation we see:

$$\begin{aligned}
\vec{\alpha}\cdot\vec{h} &= (1,0)\cdot(h_1,h_2) = h_1 = h_\alpha\\
\vec{\beta}\cdot\vec{h} &= (-1/2,\sqrt{3}/2)\cdot(h_1,h_2) = -1/2\,h_1 + \sqrt{3}/2\,h_2\\
&= \begin{pmatrix}-1/4 & 0 & 0\\ 0 & 1/4 & 0\\ 0 & 0 & 0\end{pmatrix} + \begin{pmatrix}1/4 & 0 & 0\\ 0 & 1/4 & 0\\ 0 & 0 & -1/2\end{pmatrix} = h_\beta\\
\vec{\gamma}\cdot\vec{h} &= (1/2,\sqrt{3}/2)\cdot(h_1,h_2) = 1/2\,h_1 + \sqrt{3}/2\,h_2\\
&= \begin{pmatrix}1/4 & 0 & 0\\ 0 & -1/4 & 0\\ 0 & 0 & 0\end{pmatrix} + \begin{pmatrix}1/4 & 0 & 0\\ 0 & 1/4 & 0\\ 0 & 0 & -1/2\end{pmatrix} = h_\gamma.
\end{aligned}$$

We can verify, and write compactly:

$$\begin{aligned}
[\vec{h}, e_\alpha] &= (\mathrm{ad}\,\vec{h})e_\alpha = (1,0)e_\alpha = \vec{\alpha}e_\alpha\\
[\vec{h}, e_\beta] &= (\mathrm{ad}\,\vec{h})e_\beta = \left(\frac{-1}{2}, \frac{\sqrt{3}}{2}\right)e_\beta = \vec{\beta}e_\beta\\
[\vec{h}, e_\gamma] &= (\mathrm{ad}\,\vec{h})e_\gamma = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)e_\gamma = \vec{\gamma}e_\gamma
\end{aligned}$$

Note that $[h_\alpha, e_\alpha] = [\vec{\alpha}\cdot\vec{h}, e_\alpha] = (\vec{\alpha}\cdot\vec{\alpha})e_\alpha = e_\alpha$ as $|\vec{\alpha}| = 1$. In other words, all of this notation is consistent. This notation leads to the *root diagram* of $\mathfrak{su}(3)$:

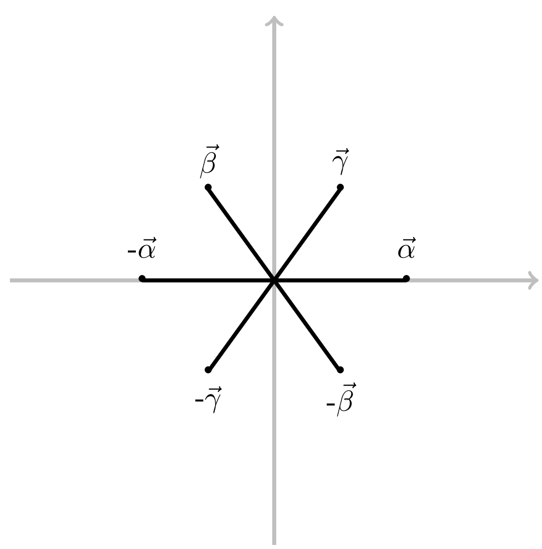


Figure 9.1: The root diagram of $L(SU(3))$

$\pm\vec{\alpha}, \pm\vec{\beta}, \pm\vec{\gamma}$ are the *roots* of $\mathfrak{su}(3)$. (In this case they are of unit length.) The negative roots are included because:

$$[e_{-\alpha}, e_\alpha] = -2h_\alpha = 2h_{-\alpha} = 2(-\vec{\alpha}) \cdot \vec{h}$$

and so on. $\mathfrak{su}(3)$ is said to have *rank* 2. The rank is the dimension of the maximal commuting subalgebra (the Cartan subalgebra, or CSA), which is here spanned by h_1 and h_2. The root diagram is two-dimensional. $\mathfrak{su}(3)$ has three (special) $\mathfrak{su}(2)$ subalgebras, one for each pair $\pm\alpha, \pm\beta, \pm\gamma$. There are also further brackets that we have not yet given, e.g.,

$$\begin{aligned}
&[e_\alpha, e_\alpha] = 0 \\
&[e_\alpha, e_\beta] = e_\gamma \\
&[e_\alpha, e_\gamma] = 0 \\
&\text{and so on} \ldots
\end{aligned}$$

Remark. We show below that these roots are consistent with our previous definition of roots as non-zero eigenvalues of the Cartan subalgebra in the adjoint representation.

9.2 Representations and Weights

Suppose d is a representation of $\mathfrak{su}(3)$ acting on V. Let h_1 and h_2 be as before. Then $d(h_1)$ and $d(h_2)$ commute (as h_1 and h_2 commute) and can be simultaneously diagonalized. Thus V decomposes into subspaces. Let

$$V_{\lambda_1,\lambda_2} = \{v \in V : d(h_i)v = \lambda_i v\}$$

(i.e., V_{λ_1,λ_2} is the span of the simultaneous eigenvectors belonging to eigenvalues λ_1 and λ_2 of $d(h_1)$ and $d(h_2)$). We can simplify notation by writing $\vec{h} = (h_1, h_2)$, $\vec{\lambda} = (\lambda_1, \lambda_2)$. Then

$$V_{\vec{\lambda}} = \{v \in V : d(\vec{h})v = \vec{\lambda}v\}$$

If $V_{\vec{\lambda}}$ is a non-zero subspace, we say that $\vec{\lambda}$ *is a weight* and $V_{\vec{\lambda}}$ *is a weight space for d.*

Example: The fundamental representation of $\mathfrak{su}(3)$.
The fundamental representation of $\mathfrak{su}(3)$ is 3-dimensional, and

$$h_1 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; h_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

are diagonal. Thus we can immediately read off the eigenvalues and hence the fundamental weights:

$$\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \; \left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \; \left(0, -\frac{1}{\sqrt{3}}\right)$$

The weight spaces are clearly one-dimensional and given by:

$$\begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix}, \; \begin{pmatrix} 0 \\ c \\ 0 \end{pmatrix}, \; \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}$$

We can show the weights on the root diagram; they form an equilateral triangle:
Note: We have three weights since the fundamental representation is 3-dimensional. Each weight is a 2-vector, since the Cartan subalgebra is 2-dimensional.

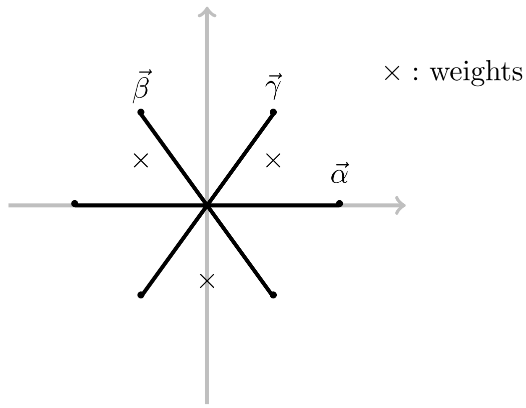


Figure 9.2: Sketch of root diagram of $L(SU(3))$ along with the fundamental weights

9.2.1 General Constraint on Weights

We saw previously that $\{e_\alpha, e_{-\alpha}, h_\alpha\}$ span an $\mathfrak{su}(2)$ subalgebra of $\mathfrak{su}(3)$. By restriction, we see that the representation d of $\mathfrak{su}(3)$ gives a representation of $\mathfrak{su}(2)$ acting on V. From our experience with $SU(2)$ theory (angular momentum), we know that $d(h_\alpha)$ has half-integer or integer eigenvalues. If $v \in V_{\vec{\lambda}}$, then

$$\begin{aligned} d(h_\alpha)v &= d(\vec{\alpha} \cdot \vec{h})v \\ &= \vec{\alpha} \cdot d(\vec{h})v, \text{ by linearity} \\ &= \vec{\alpha} \cdot \vec{\lambda} v, \text{ since } v \in V_{\vec{\lambda}} \end{aligned}$$

Our experience with $SU(2)$ theory therefore tells us $2\vec{\alpha} \cdot \vec{\lambda} \in \mathbb{Z}$.
The same argument and result holds for any root $\vec{\delta}$ (i.e., $\pm\vec{\alpha}, \pm\vec{\beta}, \pm\vec{\gamma}$) of $\mathfrak{su}(3)$. Thus the weights $\vec{\lambda}$, for *any* representation d of $\mathfrak{su}(3)$ are constrained by

$$2\vec{\delta} \cdot \vec{\lambda} \in \mathbb{Z} \tag{9.1}$$

where $\vec{\delta}$ runs over all roots.
The possible weights therefore form a lattice. A pair of basis lattice vectors are $\vec{w}_1 = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)$ and $\vec{w}_2 = \left(0, \frac{1}{\sqrt{3}}\right)$, which satisfy:

$$\begin{aligned} 2\vec{\alpha} \cdot \vec{w}_1 &= 1 \\ 2\vec{\alpha} \cdot \vec{w}_2 &= 0 \\ 2\vec{\beta} \cdot \vec{w}_1 &= 0 \\ 2\vec{\beta} \cdot \vec{w}_2 &= 1 \end{aligned}$$

(Check using the definitions of $\vec{\alpha}, \vec{\beta}$.) A general weight is $\vec{\lambda} = n_1\vec{w}_1 + n_2\vec{w}_2$, where $n_1, n_2 \in \mathbb{Z}$. This satisfies the constraint equation (9.1) for $\vec{\delta} = \vec{\alpha}$ and $\vec{\delta} = \vec{\beta}$, and for $\vec{\delta} = \vec{\gamma} = \vec{\alpha} + \vec{\beta}$.
The collection of all possible weights is the *weight lattice*. The weights of $d = \mathrm{ad}$ (the adjoint representation) are the roots together with $\vec{0}$ (with multiplicity 2). They are shown by the circles in the lattice.

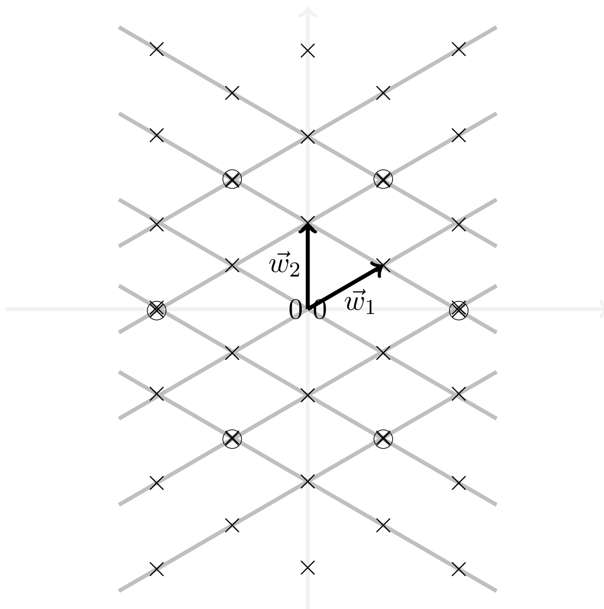


Figure 9.3: The weight lattice

Roots are weights, since ad is of course a representation. In the case of $\mathfrak{su}(3)$,

$$
\begin{aligned}
&(\operatorname{ad} h_1)e_{\pm\alpha} = [h_1, e_{\pm\alpha}] = \pm e_{\pm\alpha} \\
&(\operatorname{ad} h_2)e_{\pm\alpha} = [h_2, e_{\pm\alpha}] = 0 \\
&\implies (\operatorname{ad} \vec{h})e_{\pm\alpha} = (\pm 1, 0)e_{\pm\alpha}, \text{ which just says that } \pm\vec{\alpha} \text{ is a weight pair.}
\end{aligned}
$$

Similarly, we see the other weights are determined according to:

$$
\begin{aligned}
(\operatorname{ad} \vec{h})e_{\pm\beta} &= (\mp 1/2, \pm\sqrt{3}/2)e_{\pm\beta} = \pm\vec{\beta}e_{\pm\beta} \\
(\operatorname{ad} \vec{h})e_{\pm\gamma} &= (\pm 1/2, \pm\sqrt{3}/2)e_{\pm\gamma} = \pm\vec{\gamma}e_{\pm\gamma} \\
(\operatorname{ad} \vec{h})h_1 &= 0 \\
(\operatorname{ad} \vec{h})h_2 &= 0
\end{aligned}
$$

Thus ad has eight weights $\{\pm\vec{\alpha}, \pm\vec{\beta}, \pm\vec{\gamma}, \vec{0}, \vec{0}\}$.

9.2.2 Weights of some Irreps of $\mathfrak{su}(3)$

For a given representation, the weights belong to the weight lattice. They must also form *complete strings* of weights for each $\mathfrak{su}(2)$ subalgebra. (The $\mathfrak{su}(2)$ weights are read off along the lines parallel to the directions of $\vec{\alpha}$, $\vec{\beta}$ or $\vec{\gamma}$.) Weights of some irreps of $\mathfrak{su}(3)$ are given in Fig. 9.4. The label of each irrep indicates its dimension.

9.2.3 Conjugate Representations

The fundamental representation of $SU(3)$ is $D(U) = U$. The conjugate representation is $D(U) = U^*$. This defines a representation because $U_1U_2 = U_3 \mapsto U_1^*U_2^* = U_3^*$. We may also consider the associated representations of the Lie algebra $\mathfrak{su}(3)$:

Fundamental:	$D(I + X) = I + X$, so $d(X) = X$
Conjugate:	$D(I + X) = I + X^*$, so $d(X) = X^*$

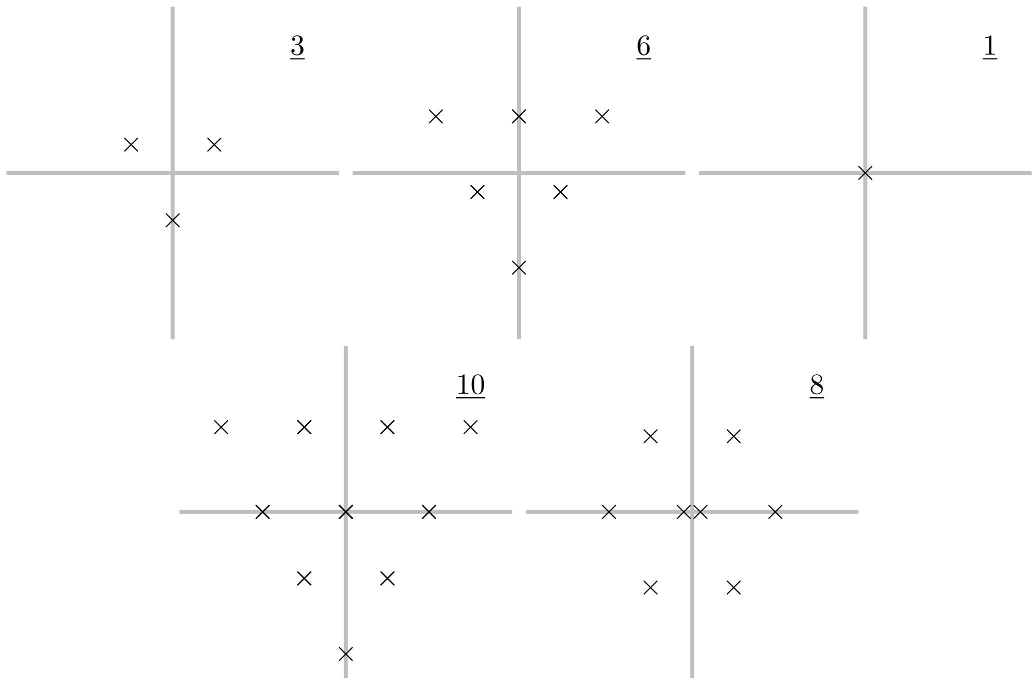


Figure 9.4: The weight diagrams for some irreducible representations of $L(SU(3))$

Since $X \in \mathfrak{su}(3)$, it is antihermitian and thus has imaginary eigenvalues: $Xv = i\mu v, \mu \in \mathbb{R}$. Then for the conjugate: $X^* v^* = -i\mu v^*$, and so we see that the eigenvalues change sign. Now we move to the complexification $\mathfrak{su}(3) \otimes \mathbb{C}$, i.e., we now allow for complex linear combination such as e_α, h_α, and so on. The weights still change sign. The weight diagrams of the fundamental irrep $\underline{3}$ and its conjugate $\overline{\underline{3}}$ are:

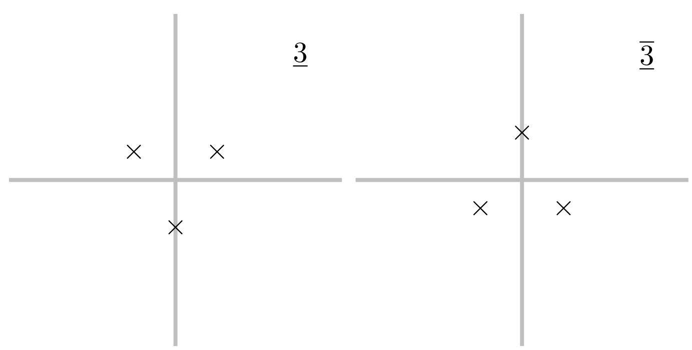


Figure 9.5: The weight diagrams of the fundamental and conjugate representations of $L(SU(3))$

Note that these are different, so $\underline{3}$ and $\overline{\underline{3}}$ are *not equivalent*. (Check that equivalent reps have the same weights.) Compare this to the fundamental representation of $\mathfrak{su}(2)$, which is self-conjugate. Similarly, the conjugates of $\underline{6}$ and $\underline{10}$ are distinct, $\overline{\underline{6}}$ and $\overline{\underline{10}}$, respectively. However, $\underline{8} = \overline{\underline{8}}$ is self-conjugate, as the representation ad involves the structure constants, which are real.

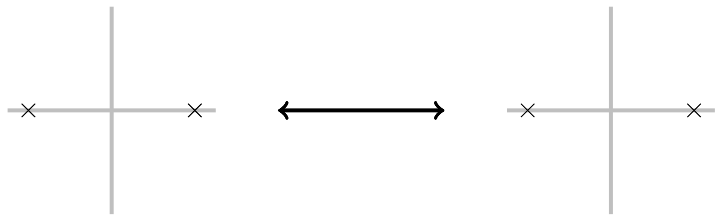
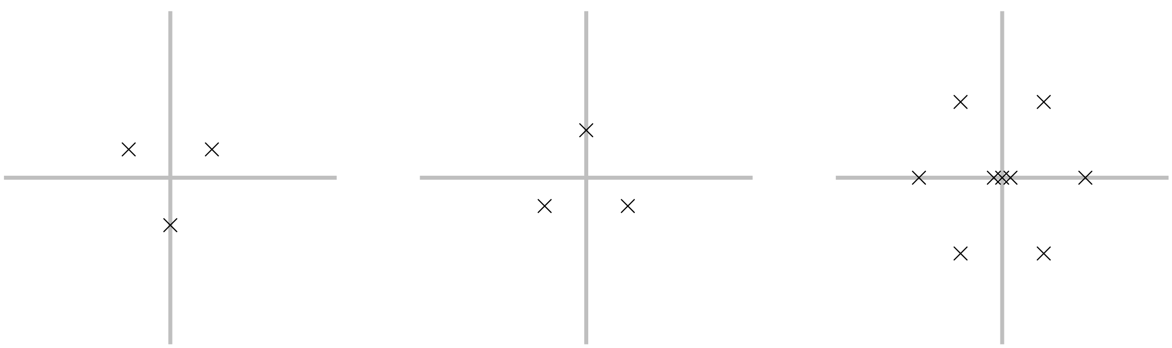


Figure 9.6: The weight diagrams for the fundamental and conjugate representations of $L(SU(2))$.

9.2.4 Tensor Products for $\mathfrak{su}(3)$

We can construct all irreducible representations of $\mathfrak{su}(3)$ from tensor products of $\underline{3}$ (the fundamental representation) and $\underline{\overline{3}}$ (its conjugate) and then reducing the tensor product to irreducible representations. One *adds weights* to identify the representations present in the tensor product.

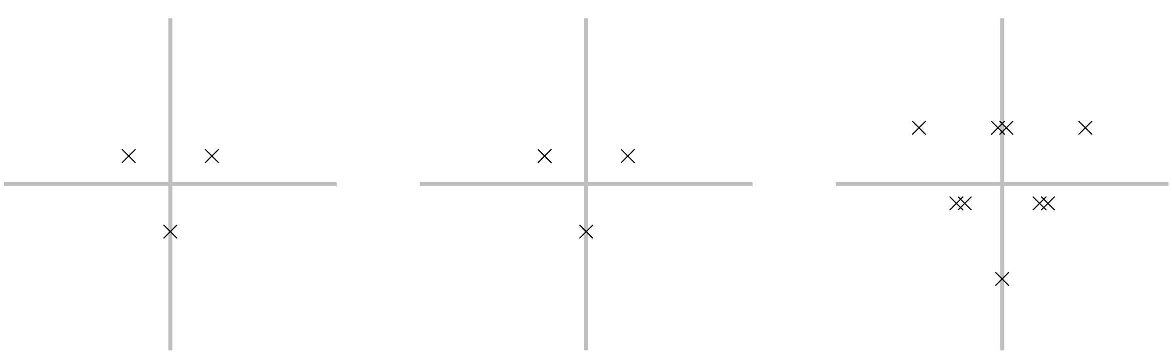
Example 1: $\underline{3} \otimes \underline{\overline{3}} = \underline{8} \oplus \underline{1}$



$$\underline{3} \qquad \otimes \qquad \underline{\overline{3}} \qquad = \qquad \underline{8} \oplus \underline{1}$$

The singlet is invariant under the action of $SU(3)$. If a tensor has the form $v^*_\beta u_\alpha$, then the singlet is $\sum_{\alpha=1}^{3} v^*_\alpha u_\alpha = v^\dagger u$.

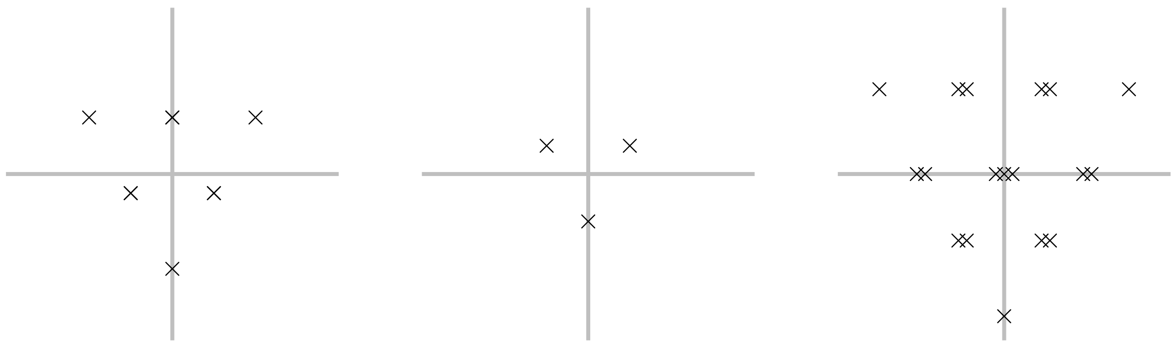
Example 2: $\underline{3} \otimes \underline{3} = \underline{6} \oplus \underline{\overline{3}}$



$$\underline{3} \qquad \otimes \qquad \underline{3} \qquad = \qquad \underline{6} \oplus \underline{\overline{3}}$$

$\underline{6}$ is the symmetric 3×3 tensor, $S_{\alpha\beta}$, and $\underline{\overline{3}}$ is the antisymmetric 3×3 tensor, $A_{\alpha\beta}$. Since these are acted upon by $SU(3)$, they are complex.

Example 3: $\underline{3} \otimes \underline{3} \otimes \underline{3} = (\underline{6} \oplus \underline{\overline{3}}) \otimes \underline{3} = (\underline{6} \otimes \underline{3}) \oplus \underline{8} \oplus \underline{1} = \underline{10} \oplus \underline{8} \oplus \underline{8} \oplus \underline{1}$



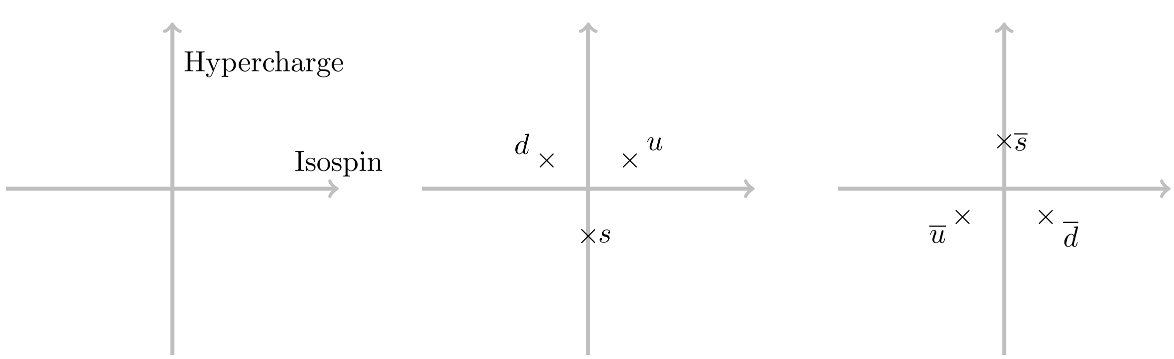
$\underline{6} \qquad \otimes \qquad \underline{3} \qquad = \qquad \underline{10} \oplus \underline{8}$

$\underline{10}$ is a totally symmetric 3-index tensor. The singlet $\underline{1}$ is a totally antisymmetric tensor. In terms of three 3-vectors, the singlet is $\epsilon_{\alpha\beta\gamma} u^\alpha v^\beta w^\gamma$. The singlet is important because it is $SU(3)$-invariant.

9.3 Quarks

Heisenberg noted the close degeneracy (in the physical sense) of the proton and neutron masses $(p, n) = (938 \text{ MeV}, 940 \text{ MeV})$ and observed that (with the exception of electric charge) the two particles have similar properties in nuclei. This led him to propose *isospin symmetry*, an $SU(2)$ symmetry with (p, n) as a doublet. Isospin was confirmed by the discovery of three pions (π^+, π^0, π^-) of similar mass, and interactions that were approximately $SU(2)$ invariant. The discovery of more particles, like (K^+, K^0), and the new *strangeness* conservation led Gell-Mann to suggest a larger $SU(3)_{\text{flavor}}$ approximate symmetry with $SU(2)_{\text{isospin}}$ as a subgroup. The particles transforming under the *fundamental representation* ($\underline{3}$) of flavour $SU(3)$ are now understood to be the quarks (u, d, s), i.e., $(1, 0, 0)$ is identified with a u quark state and so on. All other hadrons are multi-quark states in $SU(3)_{\text{flavor}}$ multiplets. Mathematically, we use tensor products to produce these multiplets. Particles in irreducible representations have approximately the *same mass*, and related strong interactions. (Different irreducible representations are *not* related by $SU(3)_{\text{flavor}}$ symmetry.)

In the multiplets, quarks q are in $\underline{3}$, while antiquarks $\bar{q}$ are in $\underline{\overline{3}}$



The masses of the up and down quarks $\begin{pmatrix} u \\ d \end{pmatrix}$ are very similar (and small, on the order of a few MeV). The strange quark (s) has a larger mass. (The Particle Data Group says 95 ± 5 MeV for the mass of the strange quark, still small compared to the proton or even pion masses.)

9.3.1 Meson Octet

The $q\bar{q}$ meson octet comes from the tensor product $\underline{3} \otimes \underline{\bar{3}}$.

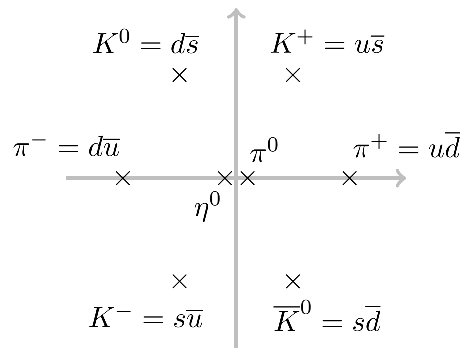


Figure 9.7: The meson octet

This agrees with what is seen: 4 kaons (mass ∼500 MeV), 3 pions (mass ∼140 MeV), 1 eta (mass ∼500 MeV). There is also a further *singlet* η' (also neutral, but heavier than η^0). π^0, η^0 and η' are linear combinations of $u\bar{u}$, $d\bar{d}$ and $s\bar{s}$ states.

9.3.2 Baryon Octet and Decuplet

We obtain qqq baryons in representations $\underline{8}$ and $\underline{10}$ from $\underline{3} \otimes \underline{3} \otimes \underline{3}$.

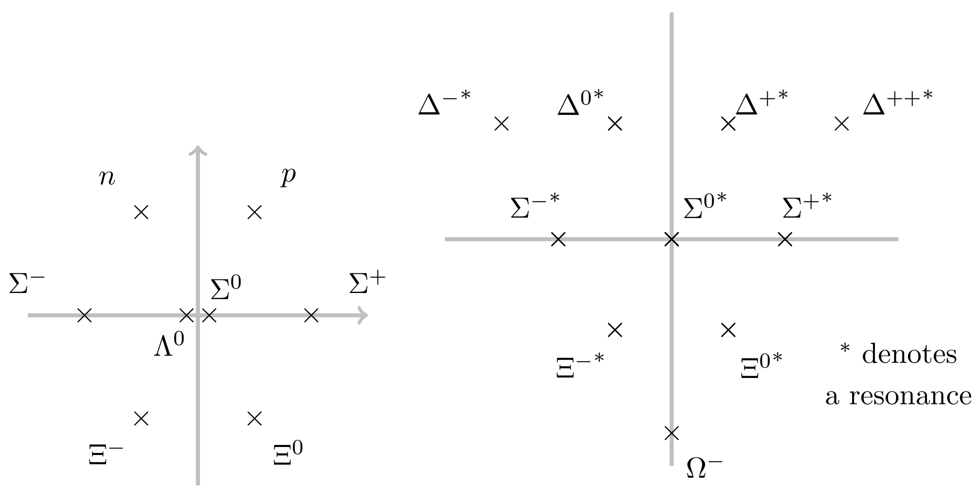
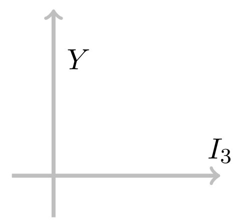


Figure 9.8: The baryon octet and decuplet

The diagrams above are consistent with what is seen in hadron collisions and decays. The baryon octet contains the long-lived baryons, which have spin 1/2 and decay weakly (only p is stable). The baryon decuplet consists mostly of resonances, which are produced in, for example, pion-proton and kaon-proton collisions. They are short-lived and have spin 3/2. The sss baryon Ω^- was a successful prediction of this scheme and is the only long-lived particle in this group.

There are three conserved quantities in the *strong interactions* of these particles:

- Y (related to strangeness S), hypercharge
- I_3 (related to electric charge Q), 3rd component of isospin
- B, baryon number



I_3 and Y are the eigenvalues of $d(h_1)$ and $\frac{2}{\sqrt{3}}\,d(h_2)$, respectively. The net quark numbers

$$\begin{aligned} N_u &= \#u - \#\bar{u} \\ N_d &= \#d - \#\bar{d} \\ N_s &= \#s - \#\bar{s} \end{aligned}$$

are all conserved under strong interactions. We also have the following relationships:

$$\begin{aligned} B &= \frac{1}{3}(N_u + N_d + N_s) \\ I_3 &= \frac{1}{2}N_u - \frac{1}{2}N_d \\ Y &= \frac{1}{3}N_u + \frac{1}{3}N_d - \frac{2}{3}N_s \\ Q &= \frac{2}{3}N_u - \frac{1}{3}N_d - \frac{1}{3}N_s \\ S &= -N_s \end{aligned}$$

Of these five related quantum numbers, only three are independent. One relation is $Q = I_3 + \frac{1}{2}Y$. For the quarks, the electric charges are $Q(u) = \frac{2}{3}$, $Q(d) = -\frac{1}{3}$, $Q(s) = -\frac{1}{3}$. Originally thought to be algebraic curiosities, these values have found verification from deep inelastic scattering experiments.

9.3.3 The Pauli Principle and Color

Consider $\Delta^{++} = u^\uparrow u^\uparrow u^\uparrow$, consisting of three spin-up quarks. This is a $|3/2, 3/2\rangle$ spin state and also $|3/2, 3/2\rangle$ isospin state. Through models, we believe that this state should have a spatially symmetric wavefunction. Since it also has a symmetric spin and flavor state, this appears to violate the Pauli principle for spin-1/2 quarks. This leads one to propose a further "color" label for quarks, $q_\nu, \nu = 1, 2, 3$. States such as qqq must be totally antisymmetric in color: $\epsilon_{\mu\nu\sigma} q_\mu q_\nu q_\sigma$. Color algebra motivated the $SU(3)$ gauge theory, QCD, where color symmetry rules become a consequence of gauge invariance. Gauge invariance of physical states also explains why only color singlet $\underline{3} \otimes \overline{\underline{3}}$ $(q\bar{q})$, $\underline{3} \otimes \underline{3} \otimes \underline{3}$ (qqq), and $\overline{\underline{3}} \otimes \overline{\underline{3}} \otimes \overline{\underline{3}}$ $(\bar{q}\bar{q}\bar{q})$ states are seen physically. Note that free quarks q are not gauge invariant. Rather surprisingly, there is no evidence for $qq\bar{q}\bar{q}$ or $qqqqqq$ states, even though color singlets exist here, and there are seemingly no glueballs. The dynamics of color confinement remain mysterious.

Chapter 10

Complexification of $L(G)$, Representations

10.1 $L(G)^{\mathbb{C}}$

Consider a real Lie algebra $L(G)$ with basis $\{T_k\}$ and brackets $[T_i, T_j] = c_{ijk}T_k$, where $c_{ijk} \in \mathbb{R}$. Here the T's themselves need *not* be real matrices, but the structure constants *are* real. The full Lie algebra then consists of the real linear span of the basis matrices. The *complexification* of $L(G)$, denoted $L(G)^{\mathbb{C}}$, has a general element $\sum_k \lambda_k T_k$, with $\lambda_k \in \mathbb{C}$. The bracket is unchanged:

$$[\lambda_i T_i, \mu_j T_j] = \lambda_i \mu_j c_{ijk} T_k$$

Here *the real dimension doubles*, while the *complex dimension is the same* as the real dimension of the original Lie algebra. As we saw earlier, for the complexified algebra it is sometimes convenient to use new basis elements that are complex linear combinations of the $\{T_k\}$.
A representation d of $L(G)$ becomes a rep d of $L(G)^{\mathbb{C}}$:

$$d(\lambda_k T_k) = \lambda_k d(T_k)$$

10.2 $L(G)^{\mathbb{C}}$ as a real Lie algebra $\Re\{L(G)^{\mathbb{C}}\}$

The real algebra $\Re\{L(G)^{\mathbb{C}}\}$ has dimension $2\dim(L(G))$. A basis for $\Re\{L(G)^{\mathbb{C}}\}$ is given by $\{X_k = T_k, Y_k = iT_k\}$. (We assume here that these are all independent, i.e., that multiplication by i does not reproduce a basis element. Abstractly, they *are* independent.) We have brackets:

$$\begin{aligned}
[X_i, X_j] &= c_{ijk} X_k \\
[X_i, Y_j] &= c_{ijk} Y_k \\
[Y_i, Y_j] &= -c_{ijk} X_k
\end{aligned}$$

We can produce two representations of $\Re\{L(G)^{\mathbb{C}}\}$ using a representation d of $L(G)$:

(i) $d(X_k) = d(T_k)$
$d(Y_k) = i\, d(T_k)$.
This is effectively the representation of $L(G)^{\mathbb{C}}$ we considered above.

(ii) $d(X_k) = d(T_k)$
$d(Y_k) = -i\, d(T_k)$.
This representation is conjugate to (i).

Note that these representations preserve the brackets, as they must:

(i)

$$\begin{aligned}
[d(X_i), d(Y_j)] &= [d(T_i), i\,d(T_j)] \\
&= d([T_i, iT_j]) = d(ic_{ijk}T_k) \\
&= c_{ijk}d(Y_k) \\
[d(Y_i), d(Y_j)] &= [i\,d(T_i), i\,d(T_j)] \\
&= d([iT_i, iT_j]) = d(-c_{ijk}T_k) \\
&= -c_{ijk}d(X_k)
\end{aligned}$$

(ii)

$$\begin{aligned}
[d(X_i), d(Y_j)] &= [d(T_i), -i\,d(T_j)] \\
&= d([T_i, -iT_j]) = d(-ic_{ijk}T_k) \\
&= c_{ijk}d(Y_k) \\
[d(Y_i), d(Y_j)] &= [-i\,d(T_i), -i\,d(T_j)] \\
&= d([-iT_i, -iT_j]) = d(-c_{ijk}T_k) \\
&= -c_{ijk}d(X_k)
\end{aligned}$$

We can combine these two types of representation via a tensor product to get a new representation. Start with $d^{(1)}, d^{(2)}$ as irreps of $L(G)$. The representation of $\Re\{L(G)^{\mathbb{C}}\} = \text{span}\{X_k, Y_k\}$ defined by

$$\begin{aligned}
d(X_k) &= d^{(1)}(T_k) \otimes I + I \otimes d^{(2)}(T_k) \\
d(Y_k) &= i\,d^{(1)}(T_k) \otimes I - i\,I \otimes d^{(2)}(T_k)
\end{aligned}$$

is an irreducible representation d of $\Re\{L(G)^{\mathbb{C}}\}$.

This is exactly what is needed to understand the Lorentz group and its representations, as we will see.

10.3 Another Point of View

From the algebra $\Re\{L(G)^{\mathbb{C}}\}$ we can construct two commuting copies of the $L(G)$ algebra:

$$\begin{aligned}
Z_k &= \frac{1}{2}(X_k - iY_k) \\
\tilde{Z}_k &= \frac{1}{2}(X_k + iY_k)
\end{aligned}$$

(Note, this uses the complexification of $\Re\{L(G)^{\mathbb{C}}\}$, which is a bit complicated.) One easily computes the brackets and finds:

$$\begin{aligned}
[Z_i, Z_j] &= \frac{1}{4}[X_i - iY_i, X_j - iY_j] \\
&= \frac{1}{4}([X_i, X_j] - i[X_i, Y_j] + i[X_j, Y_i] - [Y_i, Y_j]) \\
&= \frac{1}{4}(c_{ijk}X_k - ic_{ijk}Y_k + ic_{jik}Y_k + c_{ijk}X_k) \\
&= \frac{1}{2}c_{ijk}(X_k - iY_k) \\
&= c_{ijk}Z_k
\end{aligned}$$

By similar calculations we see $[\tilde{Z}_i, \tilde{Z}_j] = c_{ijk}\tilde{Z}_k$ and $[Z_i, \tilde{Z}_j] = 0$. We represent $\{Z_k\}$ using $d^{(1)}$ and $\{\tilde{Z}_k\}$ using $d^{(2)}$ to get the tensor products:

$$\begin{aligned} d(Z_k) &= d^{(1)}(T_k) \otimes I \\ d(\tilde{Z}_k) &= I \otimes d^{(2)}(T_k) \end{aligned}$$

Adding and subtracting then gives the same formulae for $d(X_k)$ and $d(Y_k)$ as above. (One copy of $L(G)$ and hence G acts via $d^{(1)}$ on the first index of a tensor, the other copy acts via $d^{(2)}$ on the second index. This point of view makes it clearer that d is irreducible.)

Example:

$$\begin{aligned} &\mathfrak{su}(n) = \{\text{traceless antihermitian matrices}\} \\ &\mathfrak{su}(n)^{\mathbb{C}} = \{\text{traceless complex matrices}\} \\ &\Re\{\mathfrak{su}(n)^{\mathbb{C}}\} = \text{span}\{\text{traceless antihermitian matrices, traceless hermitian matrices}\} \end{aligned}$$

In the last case one also obtains all traceless complex matrices, but a basis has twice as many matrices as a basis of $\mathfrak{su}(n)^{\mathbb{C}}$, and the coefficients are real.

Remark: Previously in our discussion of $\mathfrak{su}(3)$ we had said that the 3×3 matrices $\{e_{\pm\alpha}, e_{\pm\beta}, e_{\pm\gamma}, h_\alpha, h_\beta\}$ were a basis for $\mathfrak{su}(3)^{\mathbb{C}}$. This is consistent with the second example. The coefficients multiplying these basis elements are complex.

Chapter 11

Lorentz Group and Lie Algebra, Representations

Lorentz transformations $L^\mu{}_\nu$ act on 4-vectors $x^\mu \to L^\mu{}_\nu x^\nu, \mu, \nu = 0, 1, 2, 3$. Lorentz transformations by definition preserve the square of the relativistic interval $s^2 \equiv \eta_{\mu\nu} x^\mu x^\nu$, where $\eta_{\mu\nu} = \eta^{\mu\nu} = diag(1, -1, -1, -1)$ is the Minkowski metric. The metric signature is $(1, 3)$. Preservation of the relativistic interval is the condition:

$$L^\mu{}_\sigma L^\nu{}_\tau \eta_{\mu\nu} = \eta_{\sigma\tau} \tag{11.1}$$

which is the defining equation for the Lorentz group $O(1, 3)$. This group is a variant of $O(4)$. Near the group identity, $L^\mu{}_\sigma = \delta^\mu{}_\sigma + \epsilon\, l^\mu{}_\sigma$, where ϵ is infinitesimal. Substituting this into (11.1) one finds, collecting the terms of order ϵ:

$$(l^\mu{}_\sigma \delta^\nu{}_\tau + \delta^\mu{}_\sigma l^\nu{}_\tau)\eta_{\mu\nu} = l^\mu{}_\sigma \eta_{\mu\tau} + \eta_{\sigma\nu} l^\nu{}_\tau = 0$$

Lowering indices, we see that $l_{\tau\sigma} + l_{\sigma\tau} = 0$, i.e., l with lowered indices is antisymmetric. Thus we may write

$$l^\mu{}_\sigma = \begin{pmatrix} 0 & a & b & c \\ a & 0 & -f & e \\ b & f & 0 & -d \\ c & -e & d & 0 \end{pmatrix}$$

where $a, b, c, d, e, f \in \mathbb{R}$. The first column and the first row are symmetrically related, which comes from raising the first index with the Minkowski metric (which flips the sign of the bottom three rows). These matrices form the 6-dimensional Lie algebra $L(O(1, 3))$. A basis is given by:

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

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

The K's are generators for the boosts, while the J's are generators for the rotations. They are sometimes combined into $M^{\mu\nu} = -M^{\nu\mu}$ with $M^{0k} = K_k$, $M^{ij} = \epsilon_{ijk} J_k$. We have the following brackets of the Lorentz Lie algebra:

$$\begin{aligned}
[J_i, J_j] &= \epsilon_{ijk} J_k \\
[J_i, K_j] &= \epsilon_{ijk} K_k \\
[K_i, K_j] &= -\epsilon_{ijk} J_k
\end{aligned} \tag{11.2}$$

The first set are just the brackets of $\mathfrak{su}(2)$. Note that the minus sign in the third set is critical. It means that this is precisely the algebra of $\Re(\mathfrak{su}(2)^{\mathbb{C}})$ that we met previously. Note that $\mathfrak{su}(2)^{\mathbb{C}}$ is the same as $\mathfrak{sl}(2;\mathbb{C})$. To construct an irreducible representation, use the spin j representation of $\mathfrak{su}(2)$. Let $\{T_i, i = 1, 2, 3\}$ be the standard basis of $\mathfrak{su}(2)$, and set:

$$\begin{aligned}
d^{(j)}(J_i) &= d^{(j)}(T_i) \\
d^{(j)}(K_i) &= \pm i\, d^{(j)}(T_i)
\end{aligned}$$

With our conventions, $d^{(j)}(T_i)$ is antihermitian and has imaginary eigenvalues. In the second line the plus sign corresponds to a type (i) representation of $\Re\{L(G)^{\mathbb{C}}\}$ while the minus sign corresponds to a type (ii) representation.
The general irrep of the Lorentz Lie algebra is a tensor product with spin labels (j_1, j_2):

$$\begin{aligned}
d^{(j_1,j_2)}(J_i) &= d^{(j_1)}(T_i) \otimes I + I \otimes d^{(j_2)}(T_i) \\
d^{(j_1,j_2)}(K_i) &= i\, d^{(j_1)}(T_i) \otimes I - i\, I \otimes d^{(j_2)}(T_i)
\end{aligned}$$

The $\mathfrak{su}(2)$ subalgebra $\{J_i, i = 1, 2, 3\}$ is represented by the standard tensor product $j_1 \otimes j_2$.
Remark: One can also consider $Z_i = \frac{1}{2}(J_i - iK_i)$ and $\tilde{Z}_i = \frac{1}{2}(J_i + iK_i)$ in the complexification of $\Re(\mathfrak{su}(2)^{\mathbb{C}})$, and then $d^{(j_1,j_2)}(Z_i) = d^{(j_1)}(T_i) \otimes I$ and $d^{(j_1,j_2)}(\tilde{Z}_i) = I \otimes d^{(j_2)}(T_i)$.

Global Aspects
$O(1, 3)$ has four disconnected components. They are labelled by whether $\det L^\mu{}_\nu = \pm 1$ and whether $L^0{}_0 \geq 1$ or ≤ -1. The part of $O(1, 3)$ connected to the identity has $\det L = 1$ and $L^0{}_0 \geq 1$. This subgroup is called $SO(1, 3)^\uparrow$. The rotation subgroup generated by $\{J_i\}$ is a copy of $SO(3)$. True representations of $SO(1, 3)^\uparrow$ have integer spin for the $SO(3)$ subgroup. This requires $j_1 + j_2$ to be an integer (the weights of $j_1 \otimes j_2$ are then integral), e.g. $(j_1, j_2) = (0, 0)$ or $(\frac{1}{2}, \frac{1}{2})$. The Lorentz group has a double cover with spinor representations, e.g., $(\frac{1}{2}, 0)$.

Examples of Representations

(a) $(\frac{1}{2}$,0) spinor (spin $\frac{1}{2}$)
This corresponds to the fundamental representation of $\mathfrak{sl}(2, \mathbb{C})$.

(b) (0,$\frac{1}{2}$) spinor (spin $\frac{1}{2}$)
This is conjugate to the fundamental representation of $\mathfrak{sl}(2, \mathbb{C})$.

(c) $(\frac{1}{2},\frac{1}{2})$ is the 4-vector representation of the Lorentz group. Note that *under* $SO(3)$ *rotations*, $(\frac{1}{2},\frac{1}{2})$ is reducible: $\underline{\frac{1}{2}} \otimes \underline{\frac{1}{2}} = \underline{1} \oplus \underline{0}$. This corresponds to the 4-vector decomposition $(x^0, \vec{x})$. Note of course that $(\frac{1}{2},\frac{1}{2})$ is not reducible under general Lorentz transformations.

Note: (a) and (b) are left- and right-handed Weyl spinors. A Dirac spinor is $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$.

Chapter 12

Poincaré Group and Particle States

12.1 Lie algebra and Casimirs

The Poincaré group ($Poinc$) combines Lorentz transformations and spacetime translations and acts transitively on Minkowski space $\mathcal{M}_4$. It is 10-dimensional. The isotropy group at one point (e.g., the origin) is the Lorentz group, so $\mathcal{M}_4 = Poinc/Lorentz$. More explicitly, elements of the Poincaré group are $(L^\mu{}_\nu, a^\mu)$ and the action is $x^\mu \to L^\mu{}_\nu x^\nu + a^\mu$. There is a convenient 5×5 matrix realization of this:

$$\begin{pmatrix} x \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} L & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} Lx + a \\ 1 \end{pmatrix}$$

where x and a are column 4-vectors and the above matrices are in block form. Using this realization, group multiplication is given by:

$$\begin{pmatrix} L & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} L' & a' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} LL' & La' + a \\ 0 & 1 \end{pmatrix}$$

Both the Lorentz group and the translation group are subgroups of the Poincaré group, but they do not commute. The Lie algebra $L(Poinc)$ has the basis:

$$\tilde{M}^{\rho\sigma} = \begin{pmatrix} M^{\rho\sigma} & 0 \\ 0 & 0 \end{pmatrix}, \; \tilde{P}^{\tau} = \begin{pmatrix} 0 & P^{\tau} \\ 0 & 0 \end{pmatrix}$$

where $(M^{\rho\sigma})^\alpha{}_\beta = \eta^{\rho\alpha}\delta^\sigma{}_\beta - \eta^{\sigma\alpha}\delta^\rho{}_\beta$ and $(P^\tau)^\beta = \eta^{\tau\beta}$. (Note, the indices ρ, σ, τ label the 10 basis matrices – the matrix indices are α, β. The matrices $M^{\rho\sigma}$ encode what we previously presented as the matrices J_k and K_k.)

The brackets for the Poincaré Lie algebra are the following:

$$\begin{aligned}
[\tilde{M}^{\rho\sigma}, \tilde{M}^{\tau\mu}] &= \eta^{\sigma\tau}\tilde{M}^{\rho\mu} - \eta^{\rho\tau}\tilde{M}^{\sigma\mu} + \eta^{\rho\mu}\tilde{M}^{\sigma\tau} - \eta^{\sigma\mu}\tilde{M}^{\rho\tau} \\
[\tilde{M}^{\rho\sigma}, \tilde{P}^{\tau}] &= \eta^{\sigma\tau}\tilde{P}^{\rho} - \eta^{\rho\tau}\tilde{P}^{\sigma} \\
[\tilde{P}^{\tau}, \tilde{P}^{\mu}] &= 0
\end{aligned}$$

The first equation is equivalent to the brackets given earlier for the J's and K's. The third is obvious: translations in flat space commute. The second equation isn't immediately obvious, but may be quickly verified using the matrix form for the basis elements. From now on we omit the ˜ and think of the basis elements and their brackets abstractly rather than using 5×5 matrices.

Since the Poincaré Lie algebra is not semi-simple, the Killing form is not useful (it is degenerate). However, we do have the following Casimirs in the universal enveloping algebra (i.e. commuting with both the M and the P generators):

$$P^2 = P_\mu P^\mu \quad \text{(a quadratic Casimir)}$$
$$W^2 = W_\mu W^\mu \quad \text{(a quartic Casimir)}$$

where $W_\mu = \frac{1}{2}\epsilon_{\mu\nu\rho\tau}M^{\nu\rho}P^\tau$. Here, the symbol $\epsilon_{\mu\nu\rho\tau}$ is Lorentz invariant (for the subgroup with $\det L = 1$), totally antisymmetric, and its normalisation is fixed by $\epsilon_{0123} = 1$. W_μ satisfies $W_\mu P^\mu = \epsilon_{\mu\nu\rho\tau}M^{\nu\rho}P^\tau P^\mu = 0$, as $P^\tau P^\mu$ is symmetric in the indices τ, μ while ϵ is totally anti-symmetric.

12.2 Representations

The Poincaré group has finite-dimensional representations (e.g., the five-dimensional representation above), but these are not unitary! In quantum mechanics and quantum field theory, particle states transform under symmetries by unitary operators. Thus we need unitary irreps of the Poincaré group, which will necessarily be infinite-dimensional. We'll find these as induced representations of *Poinc* acting on spaces of functions over the homogeneous space $\mathcal{M}_4 = Poinc/Lorentz$.

12.2.1 General Idea of Induced Representation

Consider the space of scalar functions $\Phi(m)$ defined on an orbit (coset space) $\mathcal{M} = G/H$ of G, where H is the isotropy group at some basepoint $m_0 \in \mathcal{M}$. We can define an infinite-dimensional representation $\mathcal{D}$ of G using the definition

$$(\mathcal{D}(g)\Phi)(m) = \Phi(g^{-1}(m))$$

(i.e. the transformed function at m has the value of the original function at the point mapped to m).

$\mathcal{D}$ is a representation as it is linear in Φ and the composition rule is satisfied:

$$\begin{aligned}(\mathcal{D}(g_1 g_2)\Phi)(m) &= \Phi(g_2^{-1}g_1^{-1}(m)) \\ &= (\mathcal{D}(g_2)\Phi)(g_1^{-1}(m)) \\ &= (\mathcal{D}(g_1)\mathcal{D}(g_2)\Phi)(m)\end{aligned}$$

Note that H acts trivially at m_0:

$$(\mathcal{D}(h)\Phi)(m_0) = \Phi(h^{-1}(m_0)) = \Phi(m_0)$$

which is why we regard Φ as a scalar function.

We can generalise this construction to a space of vector functions $\Psi(m)$ taking values in a vector space V. It is necessary that there is a finite-dimensional representation D of H acting on V. Now H acts non-trivially at m_0:

$$(\mathcal{D}(h)\Psi)(m_0) = D(h)\Psi(m_0)$$

The full action of G is defined in the following way:

$$(\mathcal{D}(g)\Psi)(m) = D(h(g, m))\Psi(g^{-1}(m))$$

The interesting new ingredient is the H-valued function $h(g, m)$. What is this? There is some freedom in fixing it. If one checks the composition rule one finds that $h(g, m)$ has to satisfy the *cocycle condition*

$$h(g_1 g_2, m) = h(g_1, m)h(g_2, g_1^{-1}(m))$$

If $h(g, m)$ doesn't depend on the second argument (m) then this simplifies to the condition for a homomorphism from G to H. But one usually cannot impose this simplification. Generally, the cocycle condition is solved as follows. One chooses, in a way depending smoothly on m, an element $g_0(m)$ of G that maps m to m_0. This exists if the action of G is transitive, but is not unique because one can follow $g_0(m)$ by any element of H. A further condition one imposes is that $g_0(m_0) = I$.

Now notice that for any g and m there are two routes from $g^{-1}(m)$ to m_0. One route is by $g_0(g^{-1}(m))$, and the other is by g followed by $g_0(m)$. These differ by some element of H, which is the function $h(g, m)$ we want. Thus we define

$$h(g, m) = g_0(m)\, g\, g_0(g^{-1}(m))^{-1}$$

One can check that this formula satisfies the cocycle condition. It also simplifies in the way one wants when $g \in H$ and $m = m_0$.

The result is that $\mathcal{D}$ is again a representation of G. It is called the representation of G *induced* from the representation D of H. Usually one supposes that D is irreducible.

12.2.2 Application to Representations of the Poincaré Group

The functions spaces above (scalar Φ or vector Ψ on $\mathcal{M}_4$) lead to reducible representations of $Poinc$. To obtain irreducible representations one needs to impose an additional linear, Poincaré-invariant condition. For a scalar function $\Phi(x)$ one may assume it obeys the Klein-Gordon equation

$$\partial_\mu \partial^\mu \Phi + M^2 \Phi = 0$$

The space of solutions is built up from simple exponentials $\exp(ikx) = \exp(ik_\mu x^\mu)$ with $k^2 = M^2$. Under the translation by a, $\exp(ikx) \mapsto \exp(ik(x - a)) = \exp(-ika)\exp(ikx)$. Thus we get a 1-dimensional irreducible representation of the translation group labelled by k_μ, i.e., momentum. Under a Lorentz transformation $x \mapsto Lx$ we have $\exp(ikx) \mapsto \exp(ikL^{-1}x) = \exp(i(kL^{-1})x)$. In indices, $k_\mu x^\mu \mapsto k_\mu (L^{-1})^\mu{}_\nu x^\nu = (k_\mu (L^{-1})^\mu{}_\nu) x^\nu$. In other words, k undergoes a Lorentz transformation. However, as a Lorentz scalar, $k^2 = k_\mu k^\mu$ is conserved. Thus in an *irreducible representation* we can fix $k^2 = M^2$. (Note that M^2 is the eigenvalue of the Casimir P^2.) The following space of functions transforms irreducibly under the Poincaré group:

$$\Phi(x) = \int f(k)\delta^+(k^2 - M^2)\exp(ikx)\, d^4k$$

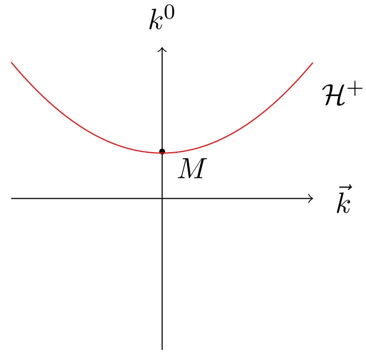
where $f(k)$ is an arbitrary (well-behaved) function. (As we've seen in quantum field theory, the δ-function may be integrated out in integrals like these, leaving an integral over 3-momentum.) f is really a function on $\mathcal{H}^+$, the hyperboloid $k^2 = M^2$ with $k_0 > 0$ in k-space. Thus we shift our focus from $\mathcal{M}_4$ to $\mathcal{H}^+$. The above function space can be thought of as a space of classical functions, but it is also a space of relativistic wavefunctions, that is, positive-energy 1-particle states for a scalar particle of mass M.

12.2.3 Irreducible Representations with Spin

We can extend the above discussion to allow f to be valued in some vector space V. The key idea is the following: $\mathcal{H}^+$ is a coset space for the Lorentz group:

$$\mathcal{H}^+ = SO(1,3)^{\uparrow}/SO(3)$$

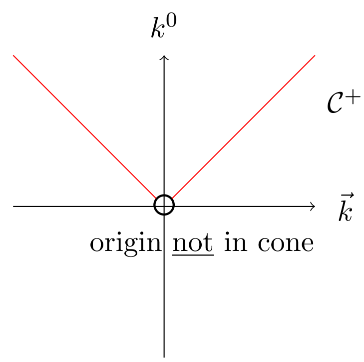
where $SO(3)$ is the isotropy group of $k^\mu = (M, \vec{0})$. We now apply the idea of an induced representation to this coset space. For an irreducible representation of the Poincaré group, we need $SO(3)$ to act irreducibly on the vector-valued function $\vec{f}(M,\vec{0})$. Thus $\vec{f}$ is a state with *definite spin* j (integer or half-integer). The basis states $\vec{f}$ at $\vec{k} = \vec{0}$ are labelled by j and $j_3 \in \{-j, -j+1, ..., j-1, j\}$. ($j_3$ is an eigenvalue of iJ_3 in the chosen irreducible representation.)



The result is a function space which can be identified with 1-particle states for a particle of mass M and spin j. M and j are related to the eigenvalues of the Casimirs P^2 and W^2. (Note that W^2 is proportional to $\vec{J}\cdot\vec{J}$ when acting on a state with momentum $k^\mu = (M, \vec{0})$.)

12.2.4 Massless Case

If $M = 0$, the orbit of $SO(1,3)^{\uparrow}$ in k-space is $\mathcal{C}^+$, where $\mathcal{C}^+$ is the positive lightcone $k^2 = 0\,,\ k^0 > 0$ shown in the picture below. We choose as base point of $\mathcal{C}^+$, $k^\mu = (1,0,0,1)$. Then we see that $\mathcal{C}^+ = SO(1,3)^{\uparrow}/E$, where E is the non-compact 3-dimensional isotropy subgroup of $(1,0,0,1)$. E has the three generators $K_1 - J_2, K_2 + J_1$ and J_3, as one can check using the 4×4 matrices for these.



The generators of E have the following brackets:

$$\begin{aligned}
[J_3, K_1 - J_2] &= K_2 + J_1 \\
[J_3, K_2 + J_1] &= -(K_1 - J_2) \\
[K_1 - J_2, K_2 + J_1] &= 0
\end{aligned}$$

which follow from (11.2). (Note in particular the final bracket, which differs from the corresponding bracket in $L(SO(3))$). These are the brackets of the Lie algebra of the (special) Euclidean group in two dimensions, which is usually called $SE(2)$ or $ISO(2)$. This group consists of translations in a plane, which commute, along with rotations about an axis perpendicular to the plane which form a compact $SO(2)$ subgroup. Finite-dimensional irreducible representations of E are labelled by the *helicity* eigenvalue j_3 of the (hermitian) $SO(2)$ generator iJ_3, which is either integer or half-integer. The translation generators have to act trivially, otherwise one would need to work with functions in the plane, lying in an infinite-dimensional space. The irreducible representation of E labelled by j_3 is simply 1-dimensional:

$$\begin{aligned} d(K_1 - J_2) &= 0 \\ d(K_2 + J_1) &= 0 \\ d(J_3) &= -ij_3 \end{aligned}$$

Thus a Poincaré irreducible representation is constructed from single-component complex functions $\vec{f}(k)$ on $\mathcal{C}^+$ with helicity j_3. iJ_3 is the projection of the physical spin along the direction of the spatial momentum $\vec{k} = (0, 0, 1)$, and j_3, the helicity, is the eigenvalue of this projected spin. For a massless particle, helicity is a Poincaré invariant, so it is the projection of spin along the momentum direction, whatever that direction is and whatever the particle energy is. There is no meaning for spin in other directions. Examples:

- Neutrinos: helicity $\pm\frac{1}{2}$ in the Standard Model.
- Photons: helicity ± 1, i.e., *not* 0. These correspond to photons circularly polarized in opposite directions.