Lecture Notes on

Group Theory

Kevin Zhou

kzhou7@gmail.com

These notes cover group theory as used in particle physics, ranging from the elementary applications of isospin to grand unified theories. The main focus is on practical computations; many core statements are not proven, nor are many algorithms proven to work. The primary sources were:

- Nick Dorey's Symmetries, Fields, and Particles lectures as transcribed by Josh Kirklin. Introduces Lie algebras and Lie groups and outlines the Cartan classification.
- Andre Lucas' Groups and Representations lectures. Covers more ground than the previous course, but rather telegraphically.
- Zee, *Group Theory in a Nutshell for Physicists*. A very readable and easygoing book developing group theory by example, spending significant time on finite groups and applications in quantum mechanics. This is a good first book to get the idea of how group theory is used in physics.
- Georgi, *Lie Algebras in Particle Physics*. The standard text for particle physicists. This thin book crams lots of content into its 300 pages, squeezing the Cartan classification into just 5, and gives many practical algorithms, at the cost of omitting proofs. The emphasis is on grand unified theories; finite groups receive little coverage.
- Wu-Ki Tung, *Group Theory in Physics*. A methodical group theory textbook that clearly covers the material that no introductory book teaches, but every advanced book assumes you already know, such as Wigner's classification, the Wigner–Eckart theorem, and Young tableaux. More rigorous and formal than most group theory books for physicists.
- Sternberg, *Group Theory and Physics*. A more formal book that focuses on applications in quantum mechanics. Uses differential geometry and bundles freely throughout.
- Ramond, *Group Theory*. A short, clean book which also briefly covers topics of mathematical interest, such as Kac–Moody algebras. Beware: the reference tables in the appendix contain a number of mistakes.
- Kirillov, *An Introduction to Lie Groups and Lie Algebras*. A standard introductory graduate math textbook for Lie theory. Written in the pure math style, but still informal.

- Fuchs and Schweigert, *Symmetries, Lie Algebras, and Representations*. Covers the standard material rigorously and goes far beyond; a useful reference for theoretical work.
- Yamatsu, *Finite-Dimensional Lie Algebras and Their Representations for Unified Model Building*. A massive reference containing every group theory constant, branching, and multiplication rule a practicing physicist could ever want; expands on Slansky's classic reference. A shorter reference focusing on generators of $\mathfrak{su}(n)$ is here.

Sections covered in the 2017 Part III course *Symmetries, Fields, and Particles* are marked with $\star$. Sections covered in the 2018 MMathPhys course *Groups and Representations* are marked with $\circ$. Groups are introduced in the lecture notes on Undergraduate Math. The most recent version is here; please report any errors found to kzhou7@gmail.com.

Contents

1 Introduction

In this course, we will investigate symmetries in physical laws and their mathematical representations. First, we consider how symmetries arise in nature.

- In general, we will consider a symmetry to be a mapping of the physical states of a system which leaves the dynamics invariant.
- Solutions of variational problems are often symmetrical. For example, soap bubbles are spherical because they minimize their area, giving $SO(3)$ symmetry. Similarly, paths $\mathbf{x}(t)$ that minimize the action for a rotationally symmetric Lagrangian remain valid paths when rotated.
- Exact symmetries can arise from redundancies in our description, e.g. gauge symmetries.
  - Coordinate transformations contain rotations $SO(3)$, or more generally Lorentz transformations $SO(3,1)$, or even more generally Poincare transformations. In general relativity we will also include arbitrary diffeomorphisms. These symmetries are collectively called spacetime symmetries; all others are called internal symmetries.
  - We also have gauge symmetries from the $U(1)$, $SU(2)$, and $SU(3)$ gauge groups. These are not truly 'symmetries' in the sense we have defined since they map a physical state to a different description of the exact same state.
  - The Coleman–Mandula theorem says that spacetime and internal symmetries cannot be combined nontrivially in relativistic quantum field theory. The famous exception is supersymmetry.
- Approximate symmetries can arise from neglecting the difference between two things. For example, isospin symmetry holds for up and down quarks if we neglect their charges and the difference in their masses, giving an $SU(2)$ symmetry. Adding the strange quark gives $SU(3)$, realized in the Eightfold Way.

As shown above, many of the groups relevant in physics are Lie groups.

- A Lie group is a group G which is also a smooth manifold, where the group operation is compatible with the smooth structure.
- The above definition is strongly constraining. A Lie group is nearly determined by its tangent space at the identity $T_e(G)$, which defines a Lie algebra $\mathfrak{g} = \mathcal{L}(G)$ when equipped with a bracket operation. In this course we will work with matrix Lie groups ($G \subset GL(n, \mathbb{F})$) for which the bracket is the matrix commutator.
- The Cartan classification states that all finite-dimensional semi-simple Lie algebras over $\mathbb{C}$ belong to four infinite families, A_n, B_n, C_n, and D_n (for positive integer n) and the five exceptional cases E_6, E_7, E_8, G_2, and F_4.
- In classical mechanics, Lie groups and Lie algebras correspond to finite and infinitesimal symmetry transformations, which yield conserved quantities by Noether's theorem.
  - In Lagrangian mechanics, symmetries preserve the action, and the conserved quantity can be read off from the Euler-Lagrange equations.

  - In Hamiltonian mechanics, symmetries are phase space flows which preserve the Hamiltonian H, and they are generated by quantities $G(q, p)$ which are conserved. Mathematically, we require the Poisson bracket of G and H to vanish.

- Quantum mechanics is similar to Hamiltonian mechanics; groups and algebras correspond to unitary and Hermitian operators which commute with the Hamiltonian, and the latter are exactly the conserved physical quantities.

- The crucial feature is the antisymmetry of the commutator/Poisson bracket. This tells us, e.g. that rotating a configuration don't change the energy if and only if time evolution doesn't change the angular momentum. This is how symmetry and conservation are generically related.

- There is one more piece to this statement: in classical mechanics, the symmetry generators are the same as the real-valued observables that we measure. In quantum mechanics, this is not true, because they must be anti-Hermitian and Hermitian respectively. We instead relate them by a factor of i, one reason that quantum mechanics requires complex numbers.

- Mathematically, the commutator and Poisson bracket make the set of infinitesimal transformations into a Lie algebra. Using the Jacobi identity shows that the set of infinitesimal symmetries is closed under the bracket and is hence a Lie subalgebra.

The link between symmetry and conservation can be somewhat subtle.

- There can be conserved quantities not associated with a continuous symmetry. For example, if a space is disconnected, the connected component a particle is in is conserved, but there is no associated continuous symmetry because the components are topologically invariant. Conserved quantities of this type are called topological charges.

- There can be algebras of symmetries that do not exponentiate to groups of symmetries; these appear in the study of supersymmetry and quantum groups. However, we still find conserved quantities, as Noether's theorem only requires infinitesimal symmetries, not finite ones.

- In both classical and quantum mechanics, continuous symmetries lead to locally conserved currents by Noether's theorem, which integrate to conserved charges. In addition, in quantum mechanics, discrete symmetries also lead to conserved charges; they partition the Hilbert space into subspaces which evolve independently.

- In simple cases, these conserved quantities can be described in terms of "quantum numbers". For example, a continuous $U(1)$ symmetry yields a conserved quantity that adds between particles. For example, translations, rotations about a single axis, and phase rotations correspond to momentum, angular momentum, and charge.

- For discrete symmetries, a unitary symmetry operator that is its own inverse is also Hermitian, and so corresponds to a conserved quantity that multiplies between particles. Examples include the eigenvalues of parity, charge conjugation, and time reversal. The reason the quantum numbers from continuous symmetries add instead of multiply is because they are defined via the eigenvalues of infinitesimal symmetries, which effectively takes a logarithm.

- However, in general symmetries groups can be non-Abelian, where the associated conservation laws cannot be understood by quantum numbers alone. Instead, particle states can be classified into irreducible representations of the group, and the allowed processes can be found by

decomposing the tensor products of representations. For example, this reasoning gives rise to selection rules associated with angular momentum conservation.

- Another phenomenon occurs with non-Abelian symmetry groups. In general, the Hilbert space is partitioned into irreducible representations of the symmetry group, and states within each irrep have the same energy. Since the irreps of abelian groups are one-dimensional, only non-Abelian symmetries can explain degeneracies. For example, the approximate eight-fold degeneracy in the baryon spectrum can be explained by an approximate $SU(3)$ flavor symmetry.

Symmetries in quantum mechanics are a bit more subtle because quantum states are only defined up to phase factors.

- Wigner's theorem states that symmetries which preserve the norms of inner products must be either unitary linear or unitary anti-linear; the latter appears for time reversal symmetry.
- Even for unitary operators, we can have projective representations $U(g)$, which satisfy

$$U(g_i)U(g_j) = e^{i\gamma(g_i,g_j)}U(g_ig_j).$$

  Note that by phase redefinitions we can always set $U(e) = I$.

- In the case where

$$\gamma(g_i, g_j) = \alpha(g_ig_j) - \alpha(g_i) - \alpha(g_j)$$

  we can remove all phase factors by redefining $U(g)$ as $e^{i\alpha(g)}U(g)$. If the phase factors are nontrivial (in a topological sense) they cannot be removed and give an inherently projective representation.

- Often, but not always, finding the projective representations of G is equivalent to finding the representations of the universal cover $\widetilde{G}$. This leads to the introduction of spinors, representations of $SU(2)$, since $SU(2)$ is the universal cover of $SO(3)$.

We will return to these issues in greater detail later.

2 Finite Groups

2.1 ○ Representations

We now review the basics of group representation theory.

- A representation of a group G is an action of G on a vector space V by linear transformations, where the element g corresponds to the linear operator $D(g)$. Equivalently, a representation of G is a group homomorphism $D : G \to GL(V)$. The representation is faithful if $D(g)$ is distinct for every distinct g. Physicists often think of the $D(g)$ as matrices, so we will write many results below in both notations.
- The dimension of a representation is the dimension of V. In physics, especially in quantum mechanics, the space V itself may be called the representation because it contains the particle states, which are often of more interest than the operators $D(g)$.
- Two representations $D(g)$ and $D'(g)$ are equivalent if they are related by $D'(g) = SD(g)S^{-1}$. In mathematics, S is called an intertwiner.
- Using Dirac notation for vectors, all groups have the regular representation,

  $$D(g)|g'\rangle = |gg'\rangle$$

  which has dimension $|G|$. Each representation matrix has exactly one 1 in each row and column, with all other elements zero. As another example, taking the determinant of every $D(g)$ always gives a one-dimensional representation.
- A representation is real/complex if V is a real/complex vector space. In physics, we're almost always interested in the complex case, partly because of quantum mechanics and partly because we may freely complexify linear problems involving only real quantities. For this section only, we'll explicitly state if representations are complex; everywhere else they are implicitly complex.
- A representation is reducible if there is a subspace $U \subset V$ so that $D(g)$ keeps U inside itself. Equivalently, a representation is reducible if the $D(g)$ can be brought into block upper-triangular form. An irreducible representation is called an irrep.
- A representation is completely reducible if the $D(g)$ can be brought into block-diagonal form, with each block irreducible.

Example. Not all representations are completely reducible. Consider the representation of $\mathbb{Z}$,

$$D(n) = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

It is reducible, but not completely reducible as the $D(n)$ are not diagonalizable for $n \neq 0$.

Lemma (Schur)**.** If $D_1(g)$ and $D_2(g)$ are two irreps, then

$$SD_1(g) = D_2(g)S$$

for all g implies either $S = 0$ or the irreps are equivalent. Also, for a complex irrep $D(g)$, if

$$SD(g) = D(g)S$$

for all g then $S \propto I$.

Proof. For the first part, note that $\mathrm{Im}(S)$ and $\mathrm{Ker}(S)$ are both invariant subspaces. Since D_1 and D_2 are irreps, they must be trivial or the entire space. The only two possibilities are $S = 0$ or S invertible, but in the latter case this means the representations are equivalent.

For the second part, note that S must have an eigenvector since we're working with complex representations. The fact that $D(g)$ and S commute means that $D(g)$ maps this eigenvector to other eigenvectors of S with the same eigenvalue. But since D is irreducible, the $D(g)$ act transitively, so $S \propto I$.

Definition. A unitary representation is one where every matrix is unitary,

$$D(g)^\dagger = D(g)^{-1}.$$

This implies the $D(g)$ preserve the inner product. Equivalently, if we think of the $D(g)$ as linear transformations, there exists a Hermitian form on V preserved by the $D(g)$,

$$\langle D(g)v, D(g)w\rangle = \langle v, w\rangle.$$

This implies all unitary representations are completely reducible, because if a subspace W is preserved, then $W^\perp$ must be preserved as well.

Prop. For a finite group, all complex representations are equivalent to unitary representations.

Proof. We use the "group averaging trick". Define

$$S = \sum_i D(g_i)^\dagger D(g_i).$$

Then we have

$$SD(g)^{-1} = \sum_i D(g_i)^\dagger D(g_i g^{-1}) = \sum_i D(g_i g)^\dagger D(g_i) = D(g)^\dagger \sum_i D(g_i)^\dagger D(g_i) = D(g)^\dagger S$$

where we reindexed in the second step. Now, S is positive definite and thus has a well-defined square root, and the above shows that $D'(g) = S^{1/2} D(g) S^{-1/2}$ give a unitary representation.

Alternatively, if we think of the $D(g)$ as linear transformations, we can phrase this proof in terms of choosing a Hermitian form by averaging. Given any Hermitian form $\langle \cdot, \cdot \rangle_0$, we define

$$\langle v, w\rangle = \sum_i \langle D(g_i)v, D(g_i)w\rangle_0.$$

Then it's straightforward to see by reindexing that the $D(g)$ preserve this form.

We make some remarks about this result, previewing results to come.

- The key intuition for the result comes from the one-dimensional case. Since all elements of a finite group have finite order, they must be represented by pure phases, which are unitary. Similarly, to maintain finite order in higher dimensions, we should only represent by rotations, not scalings.
- The construction works for real representations as well, replacing 'unitary' with 'orthogonal'. By combining these results, we have Maschke's theorem: every real or complex representation of a finite group is completely reducible.

- The same proof can be used for a compact continuous group, with integration over a group measure replacing the sum. Since we'll deal almost entirely with compact or finite groups, all representations we consider below will be assumed to be unitary.
- It also turns out that all irreps of a compact group are finite-dimensional, though this is difficult to prove.
- As a converse, consider a finite-dimensional unitary representation of a non-compact connected semisimple Lie group G. If G is simple, or more generally if G contains no nontrivial normal closed subgroups, then the representation must be trivial.
- This results applies to the connected Lorentz group, but this is acceptable because the representations used in relativistic field theory are non-unitary; quantum fields are not state vectors.
- In the case of the Poincare group, the translations form a nontrivial normal closed subgroup, so we can have nontrivial finite-dimensional unitary representations that only respond to translations. These are uninteresting, and demanding that rotations and boosts act nontrivially forces the representation to be trivial. This is acceptable because the Hilbert spaces in quantum field theory are always infinite-dimensional.

Finally, we cover some ways to construct new representations from old.

- We define the direct sum and tensor product of two representations $D^r(g)$ and $D^s(g)$ by

$$D^{r\oplus s}(g) = D^r(g) \oplus D^s(g), \quad D^{r\otimes s}(g) = D^r(g) \otimes D^s(g).$$

  Generally, the product of two irreps will not be an irrep. Tensor products are most easily managed in the physicist's index notation,

$$(D^{r\otimes s}(g))_{iajb} = D^r(g)_{ij} D^s(g)_{ab}$$

  where we index components product space with one index from each element in the product.

- We define the dual, or contragradient representation by

$$(D'(g)\omega)(D(g)v) = \omega(v), \quad \omega \in V^*.$$

  Converting the linear operator to a matrix by working in the dual basis, we have

$$D'(g) = D(g^{-1})^T.$$

  The dual representation isn't emphasized in physics because for unitary matrix representations, the dual representation is the same as the conjugate representation.

- Any representation of G induces a representation of $H \subset G$ by restriction. Restricting to a subgroup can break apart irreps, by "branching". Similarly, moving to a larger group can combine irreps together, a key idea behind unification.
- Given a representation D of $H \subseteq G$ on V, we may induce a representation of G. The induced representation on G acts on $|G|/|H|$ copies of V, one copy for each coset. The elements of this space are written as v_i where i indexes the coset.

- Taking coset representatives g_i, an arbitrary group element g may be written as $g = g_i^{-1} g_j h$, and it acts as
$$v_i \mapsto D(h) v_j.$$
That is, we act as H would have, and also shift the coset accordingly. If $D(h)$ is irreducible, so is the induced representation. The most important application of induced representations is Wigner's classification, which we'll see below.

- Let $\mathrm{Hom}(V, W)$ be the set of linear maps $V \to W$. Then
$$\mathrm{Hom}(V, W) \cong W \otimes V'.$$
Given representations D_V and D_W of G, we hence have an induced representation
$$D_{\mathrm{Hom}(V,W)} = D_W \otimes D'_V.$$
More explicitly, in mathematical notation a linear operator φ is mapped as
$$D_{\mathrm{Hom}(V,W)}(g)(\varphi) = D_W(g) \circ \varphi \circ D_V^{-1}(g)$$
while in the physicist's notation a matrix M is mapped as
$$M \mapsto D_W(g) M D_V(g)^{-1}.$$

2.2 ○ Characters and Orthogonality

Character theory will help us identify irreps and decompose general representations. We begin by proving a powerful orthogonality theorem.

Theorem (Great Orthogonality)**.** Given a d-dimensional irrep $D(g)$ of a finite group G,
$$\sum_g D(g^{-1})_{ij} D(g)_{k\ell} = \frac{|G|}{d} \delta^i_\ell \delta^k_j$$
Moreover, given two inequivalent irreps $D^r(g)$ and $D^s(g)$,
$$\sum_g D^r(g^{-1})_{ij} D^s(g)_{k\ell} = 0.$$

Proof. To prove the first part, we use the reindexing trick. For an arbitrary X, let
$$A = \sum_g D(g^{-1}) X D(g).$$
By reindexing the sum, we find
$$A = D(g^{-1}) A D(g) \text{ for all } g \in G$$
which implies that $A = \lambda I$ by Schur's lemma. Taking the trace of both sides gives $\lambda = (|G|/d) \operatorname{tr} X$. The theorem follows by setting X to be the matrix with $X_{jk} = 1$ and zeroes everywhere else. The second part of the theorem uses the other half of Schur's lemma; taking X to be an arbitrary matrix with the right dimensions, we have $A = 0$, and setting X as before gives the result.

Note. The intuition for this result is that all 'orientational' information is washed out by averaging over the group. (This is even more explicit when we integrate over, say, the continuous group of rotations.) Thus the only possible results for the sum are the 'invariant tensors' built from δ. In the case of inequivalent representations, we cannot form any δ's at all because the indices ij and $k\ell$ live in 'different spaces' and cannot be contracted, so the result must be zero.

This theorem is especially useful when applied to the characters of a representation.

- Given a representation $D^r(g)$, the character is
$$\chi^r(g) = \operatorname{tr} D^r(g).$$
Since the trace is invariant under conjugation, the character is constant on conjugacy classes.

- Under direct sum, tensor product, and dual we have
$$\chi^{r\oplus s}(g) = \chi^r(g) + \chi^s(g), \quad \chi^{r\otimes s}(g) = \chi^r(g)\chi^s(g), \quad \chi'(g) = \chi(g)^*.$$

- Setting $i = j$ and $k = l$ in the orthogonality theorem gives the orthogonality of characters,
$$\langle \chi_r, \chi_s \rangle \equiv \frac{1}{|G|}\sum_g (\chi^r(g))^* \chi^s(g) = \sum_c n_c (\chi^r(c))^* \chi^s(c) = \delta^{rs}$$
where c is a conjugacy class representative and n_c is the size of the class. Here we used the fact that the representations were unitary, so taking the inverse conjugates the character.

- More generally, suppose a reducible representation contains the irrep r, n_r times. Then the norm of the character is $\langle \chi, \chi \rangle = \sum n_r^2$, so a representation is reducible only if the norm of its character is 1. We compute the numbers n_r by computing $\langle \chi, \chi_r \rangle$.

- Going back to the orthogonality theorem, we can think of $D(g)_{ij}$ for fixed i and j as a vector in a $|G|$-dimensional complex vector space. If we work with unitary representations, then
$$\sum_g D(g)^*_{ji} D(g)_{k\ell} = \frac{|G|}{d}\delta^i_\ell \delta^k_j$$
which tells us that these vectors are all orthogonal. There are $\sum_r d_r^2$ vectors in total, giving the bound $\sum_r d_r^2 \le |G|$.

- Let the group algebra consist of complex linear combinations of the group elements under group multiplication. The regular representation acts on it by $D^{\text{reg}}(g)g' = gg'$ extended by linearity. An alternate physics-inspired notation is $g|g'\rangle = |gg'\rangle$.

- The characters of all classes but the identity vanish. Carrying out the sum shows that the regular representation contains every irrep D^r precisely d_r times, so $\sum_r d_r^2 = |G|$.

- This tells us that the $|G|$ 'vectors' $D^r(g)_{ij}$ form a complete set, so
$$\sum_{r,i,j} d_r\, D^r(g)_{ij} D^{r\dagger}(g')_{ji} = |G|\delta_{gg'}.$$
This is the completeness relation that complements the great orthogonality theorem.

- Suppose there are $N(C)$ conjugacy classes and $N(R)$ irreps. Then character orthogonality implies we must have $N(R) \le N(C)$. The completeness relation above can be used to show that $N(R) = N(C)$.

2.3 Examples and Applications

Example. Consider an abelian group G. There are $|G|$ conjugacy classes, so $|G|$ irreps, each of which must have dimension one. Our results above essentially say the character table is a unitary matrix; every irrep simply assigns a phase to every group element which adds under group multiplication. This generalizes the Fourier modes from characters of $\mathbb{Z}_n$. More generally, the orthogonality of characters generalizes Fourier modes to general groups.

Groups and their irreps appear often in quantum mechanics.

- Let the Hamiltonian H have a symmetry group G, represented on the Hilbert space by a unitary representation $D(g)$ where $[H, D(g)] = 0$ for all g. Then if $|v\rangle$ is an energy eigenstate with energy E, so is $D(g)|v\rangle$. Therefore, the Hilbert space breaks into irreps of G containing degenerate states. We can also see this directly from Schur's lemma, which implies H must be proportional to the identity when restricted to each irrep.
- A standard example of this is rotational invariance, where the Hilbert space decomposes into irreps of $G = SU(2)$. In the case where G is abelian, no degeneracy can result since all irreps are one-dimensional, but it can still be computationally useful because time evolution will keep states within one class of irrep by Schur's lemma. For example, for a $\mathbb{Z}_2$ parity symmetry, we may decompose the Hilbert space into odd and even parts, and an even state remains even for all times. The same applies to T^2 where T is time reversal.
- We may also have 'accidental degeneracies' where distinct irreps may have equal energies. However, this is often due to fine tuning of parameters, or a consequence of a larger symmetry group we haven't taken into account.

Example. For a discrete translational symmetry, the irreps are Bloch wavefunctions, which are periodic up to a fixed phase factor per translation. This does not yield any degeneracy, but if we have both translational symmetry and parity symmetry, there are two-dimensional irreps corresponding to the degeneracy of the $\pm k$ states. This also holds for a free particle, where it explains the degeneracy of $e^{\pm ikx}$.

Example. A set of N masses in d dimensions connected by springs has equation of motion

$$H\mathbf{x} = \omega^2\mathbf{x}$$

where $\mathbf{x}$ is a vector with Nd entries containing all the positions concatenated together. Now suppose the system has a group of symmetries G. This furnishes an Nd-dimensional representation $D(g)$. Since H is rotationally invariant, each irrep is a set of normal modes with the same frequency.

As a simple example, consider two atoms connected in one dimension. The atoms can be exchanged, giving the symmetry group S_2 with character table:

class c	n_c	1	$\overline{1}$
e	1	1	1
(12)	1	1	-1

This tells us the normal modes are either even or odd under reflection, corresponding to the 1 and $\overline{1}$ representation. Less trivially, consider three atoms in an equilateral triangle, where there are six normal modes and the symmetry group is S_3, with character table:

class c	n_c	1	$\bar{1}$	2	phys
e	1	1	1	2	6
(123)	2	1	1	-1	0
(12)	3	1	-1	0	2

Here, the last column denotes the physical representation. We know that $\chi(e) = 6$ since it is six-dimensional, and $\chi((123)) = 0$ since all atoms are moved by the transformation. However, (12) fixes the third atom, giving $\chi((12)) = 2$. Taking the inner product gives phys $= 1 \oplus 1 \oplus 2 \oplus 2$. Thus, there are two two-fold degeneracies.

The normal modes in 1 must be rotationally invariant, while the normal modes in each 2 must transform into each other under rotation. Hence the two translations form a 2 with frequency zero, while uniform rotation and "breathing" are the 1's. The other 2 contains the "scissoring" modes.

Next, we turn to the Wigner–Eckart theorem.

- Consider the tensor product of two irreps, indexed by μ and ν, whose vectors are indexed by i and j. The resulting vector space has a basis $\{|\alpha\lambda\ell\rangle\}$, where $|\alpha\lambda\ell\rangle$ is vector ℓ in occurrence λ of irrep α. The two bases are related by 'Clebsch–Gordan coefficients',

  $$|\alpha\lambda\ell\rangle = |ij\rangle\langle ij|\alpha\lambda\ell\rangle$$

  where the summation convention is used. The coefficients implicitly depend on μ and ν.

- The two bases transform as

  $$U(g)|ij\rangle = D^\mu(g)_{i'i}D^\nu(g)_{j'j}|i'j'\rangle, \quad U(g)|\alpha\lambda\ell\rangle = D^\lambda(g)_{\ell'\ell}|\alpha\lambda\ell'\rangle.$$

  By inserting some copies of the identity, we have

  $$D^\mu(g)_{i'i}D^\nu(g)_{j'j} = \langle i'j'|\alpha\lambda\ell'\rangle D^\lambda(g)_{\ell'\ell}\langle\alpha\lambda\ell|ij\rangle.$$

  Conceptually, this shows how the tensor product representation matrices are decomposed. Applying it in reverse allows us to construct the matrices for larger irreps from smaller ones.

- Next, consider a general representation in a vector space containing two irreps $|\mu i\rangle$ and $|\nu j\rangle$, where the two irreps are not equivalent. Then we claim the vectors are orthogonal. First,

  $$\langle\nu j|\mu i\rangle = \langle\nu j|U^\dagger(g)U(g)|\mu i\rangle = \langle\nu k|\mu\ell\rangle D^{\nu\dagger}(g)_{jk}D^\mu(g)_{\ell i}.$$

  Now we average over the group and apply the orthogonality theorem, giving

  $$\langle\nu j|\mu i\rangle = \frac{1}{|G|}\langle\nu k|\mu\ell\rangle\sum_g D^{\nu\dagger}(g)_{jk}D^\mu(g)_{\ell i} = d_\mu^{-1}\langle\nu k|\mu\ell\rangle\delta^\mu_\nu\delta^j_i\delta^k_\ell = 0$$

  since $\mu \neq \nu$ and d_μ is the dimension of μ. In the case where the irreps are equivalent, but not identically the same, we can apply a change of basis to restore orthogonality.

- This result is a generalization of the fact that eigenvectors of a Hermitian operator with different eigenvalues are orthogonal. In the case of rotational symmetry, we already knew that states with distinct values of ℓ are orthogonal because they have different L^2 eigenvalues. But the fact we've shown above applies much more generally.

- A set of operators O_i^μ is said to be a set of irreducible operators if

$$U(g)O_i^\mu U(g)^{-1} = O_j^\mu D^\mu(g)_{ji}.$$

They are also sometimes called irreducible tensors.

- Note that the vectors $O_i^\mu|\nu j\rangle$ transform under the direct product representation,

$$U(g)O_i^\mu|\nu j\rangle = U(g)O_i^\mu U(g)^{-1}U(g)|\nu j\rangle = D^\mu(g)_{ki}D^\nu(g)_{\ell j}O_k^\mu|\nu\ell\rangle.$$

Therefore, we may decompose the vectors as

$$O_i^\mu|\nu j\rangle = |\alpha\lambda\ell\rangle\langle\alpha\lambda\ell|ij\rangle.$$

Acting with $\langle\lambda\ell|$, using orthogonality, and reindexing, we have the Wigner–Eckart theorem

$$\langle\lambda\ell|O_i^\mu|\nu j\rangle = \sum_\alpha \langle\alpha\lambda\ell|ij\rangle \left(\frac{1}{d_\lambda}\sum_k \langle\lambda k|\alpha\lambda k\rangle\right).$$

The quantity in parentheses is called a reduced matrix element; crucially, it only depends on the on the irrep indices. The remaining factor is purely group-theoretic.

- Note the labeling conventions here: λ indexes irreps in the full space, with different λ values possibly corresponding to equivalent but distinct irreps, while $\alpha\lambda$ indexes irreps in the direct product representation, where different λ values automatically mean inequivalent irreps. These schemes coincide when no irreps appear multiple times, as happens for simple $\mathfrak{su}(2)$ setups.

- In many practical contexts, we instead work infinitesimally, i.e. with representations of $\mathfrak{su}(2)$ instead of $SU(2)$. Much of the reasoning above goes through unchanged, with conjugation replaced with commutation with generators. Computationally, the Wigner–Eckart theorem just says we can move around an irrep by applying raising and lowering operators to both sides.

Example. An explicit $\mathfrak{su}(2)$ example. Letting α and β be irrep indices, let

$$\langle 1/2, 1/2, \alpha|r_3|1/2, 1/2, \beta\rangle = A, \quad \langle 1/2, 1/2, \alpha|r_1|1/2, -1/2, \beta\rangle = B.$$

Here, the states $r_i|1/2, m, \beta\rangle$ form the reducible representation $1/2 \times 1 = 1/2 + 3/2$, and our goal is to relate A and B. To begin, we need to change the r_i basis. Since $[J_3, r_3] = 0$, we know that r_3 carries $m = 0$, so

$$r^0 = r_3.$$

We can then find the $m = 1$ and $m = -1$ operators by raising and lowering,

$$r^1 = -\frac{r_1 + ir_2}{\sqrt{2}}, \quad r^{-1} = \frac{r_1 - ir_2}{\sqrt{2}}.$$

The operators $\{r^{-1}, r^0, r^1\}$ are irreducible tensors, as desired. Thus to complete the problem we just need to know a Clebsch–Gordan coefficient. Alternatively, for this simple case we can explicitly compute them using raising and lowering operators. We start with

$$|3/2, 3/2\rangle = r^1|1/2, 1/2, \beta\rangle$$

and lower both sides to yield

$$|3/2, 1/2\rangle = \sqrt{\frac{2}{3}} r^0 |1/2, 1/2, \beta\rangle + \sqrt{\frac{1}{3}} r^1 |1/2, -1/2, \beta\rangle.$$

Finally, using the fact that $\langle 1/2, 1/2, \alpha | 3/2, 1/2\rangle = 0$, we find

$$0 = \sqrt{\frac{2}{3}} \langle 1/2, 1/2, \alpha | r^0 | 1/2, 1/2, \beta\rangle + \sqrt{\frac{1}{3}} \langle 1/2, 1/2, \alpha | r^1 | 1/2, -1/2, \beta\rangle.$$

We thus conclude that $A = B$.

Note. Note that in the above example, we couldn't have found $|1/2, 1/2\rangle$ in terms of the β states by orthogonality with $|3/2, 1/2\rangle$, because we don't know the norms of the states $r^i|1/2, m, \beta\rangle$. Instead, we can find $|1/2, 1/2\rangle$ by demanding that it be annihilated by J^+. The same goes for putting operators in the irreducible tensor basis: we simply use the algebra, without the 'crutch' of orthogonality.

2.4 Real and Complex Representations

From this point on, all representations will be complex, as promised earlier. Hence it is useful and conventional in physics to use the word "complex" to mean something else.

- Given a representation $D(g)$, its conjugate representation is $D^*(g) = D(g)^*$. A representation is said to be complex if it is not similar to its conjugate; then every representation with a non-real character is complex.

- Now suppose an irrep $D(g)$ is not complex. Then

  $$D(g)^* = SD(g)S^{-1}.$$

  Transposing and using unitarity gives

  $$D(g^{-1}) = S^{-1^T} D(g)^T S^T$$

  and plugging this equation into itself gives

  $$D(g) = (S^{-1}S^T)^{-1} D(g) (S^{-1}S^T).$$

  By Schur's lemma, this implies $S^{-1}S^T = \eta I$, so

  $$S^T = \pm S.$$

  If S is symmetric, we say the representation is real; if S is antisymmetric, we say the representation is pseudoreal, or quaternionic.

- Note that taking the transpose of an antisymmetric $n \times n$ matrix keeps its determinant the same, but also multiplies the determinant by $(-1)^n$. Therefore pseudoreal representations can only exist in even n since S must be invertible.

- The matrix S can always be chosen unitary. Solving for S above gives

$$S = D(g)^T S D(g), \quad S^\dagger = D(g)^\dagger S^\dagger D(g)^*.$$

  Multiplying these gives

$$S^\dagger S = D(g)^\dagger S^\dagger S D(g)$$

  so that $S^\dagger S$ is proportional to the identity, again by Schur's lemma. A bit more work shows that the proportionality constant is real. Hence we can always scale S by a constant so $S^\dagger S = I$.

- The matrices in a real representation are always equivalent to real matrices. First, note that since S is unitary and symmetric, we have $S = e^{iH}$ where H is symmetric. Letting $W = e^{iH/2}$, we have $S = W^2$ where W is also unitary and symmetric. Now

$$W^2 D(g) W^{-2} = D(g)^*$$

  which implies

$$W D(g) W^{-1} = W^{-1} D(g)^* W = (W D(g) W^{-1})^*.$$

  Then the matrices $W D(g) W^{-1}$ are real as desired.

- Equivalently, we may work with Lie algebra representations, letting $D(e^{iX}) = e^{id(X)}$ for a generator X, where $d(X)$ is Hermitian. In this case the conjugate representation of d is the negative conjugate, so for a non-complex representation we have

$$d(X) = -S d(X)^* S^{-1}$$

  where S is the same as for the Lie group representation. For real representations, $d(X)$ can be made pure imaginary, so $D(g)$ is real. For pseudoreal representations, $d(X)$ can be made real.

- Another way of thinking about S is that it is a rank 2 invariant tensor. For example, rearranging the above equation we have

$$d(X) S + S d(X)^T = d(X)^k_i S_{kj} + d(X)^k_j S_{ik} = 0$$

  which is just the infinitesimal transformation of an invariant tensor.

- For a real or pseudoreal representation d, the trivial representation appears exactly once in $d \otimes d$. The tensor S projects out the trivial representation, by contraction with it, so d is real if the trivial representation is in $(d \times d)_s$ and pseudoreal if it is in $(d \times d)_a$.

- Physically, the conjugate representation is important because antiparticles transform in the conjugate representation. Intuitively, any representation that arises from physical rotations of coordinates is real.

Next, we build a 'reality checker' for general irreps.

- If an irrep is real or pseudoreal, then $y^T S x$ is an invariant bilinear, as

$$y^T S x \to y^T D(g)^T S D(g) x = y^T S D(g)^\dagger S^{-1} S D(g) x = y^T S x.$$

  Conversely, the existence of such an invariant bilinear shows the irrep is real or pseudoreal.

- Note that $y^\dagger x$ is always trivially an invariant since the $D(g)$ are unitary; the difference here is that we have y^T rather than $y^\dagger$. In the simplest case where the matrices of $D(g)$ are already real, the invariant bilinear is just $y^T x$, i.e. the $D(g)$ are orthogonal.

- To construct S, we use the averaging trick. Define

$$S = \sum_g D(g)^T X D(g)$$

for arbitrary X. Then $D(g)^T S D(g) = S$, giving an invariant bilinear as desired. Since we can always define S this way, we must have $S = 0$ for a complex representation. If we suppose that $X_{i\ell} = 1$ with all other entries zero, we have

$$\sum_g D(g)^{ij} D(g)^{\ell k} = 0.$$

Finally set $j = \ell$ and $i = k$ to find

$$\sum_g \chi(g^2) = 0.$$

This is our test for a complex representation.

- On the other hand, for a pseudoreal representation, $S^T = \eta S$, which gives

$$\sum_g D(g)^T X^T D(g) = \eta \sum_g D(g)^T X D(g).$$

Performing the same procedure above, we find

$$\sum_g \chi(g^2) = \eta \sum_g \chi(g)^2 = \eta |G|$$

where we used the fact that the character is real. The quantity $\sum_g \chi(g^2)$ is called the Frobenius-Schur indicator.

Example. The fundamental representation of $\mathfrak{su}(2)$. The representations of the generators are the Pauli matrices. They clearly cannot be made pure imaginary by a basis change, since there is only one antisymmetric pure imaginary 2×2 matrix. The representation is instead pseudoreal, with

$$\sigma_a = -\sigma_2 \sigma_a^* \sigma_2$$

where σ_2 is antisymmetric as expected. The fact that we use σ_2 to do the transformation is purely a matter of convention; it is because the usual phase conventions make σ_2 alone non-real. Up to a phase factor, σ_2 is just the invariant tensor ϵ^{ab}, and the fundamental is pseudoreal because the antisymmetric combination is in the trivial representation. This latter reasoning makes it easy to see that all nontrivial representations of $\mathfrak{su}(2)$ are pseudoreal.

2.5 ○ The Group Algebra

We return to the group algebra, to prepare for classifying the representations of S_n. For clarity, we reserve the letter g for group elements, while r and s stand for algebra elements, and write the group algebra as $\tilde{G}$.

- Representations of the group yield representations of the group algebra by linearity, and the group multiplication itself yields the regular representation, which decomposes into irreps as

$$\widetilde{G} = \bigoplus_{\mu} L^\mu, \quad L^\mu = \bigoplus_{a=1}^{n_\mu} L^\mu_a, \quad n_\mu = \dim D^\mu$$

  where μ indexes the distinct irreps.

- An subrepresentation L of D^R is spanned by a basis $|r\rangle$ so that $|sr\rangle \equiv s|r\rangle \in L$ for all $s \in \tilde{G}$. Thinking of the group algebra as a ring, subrepresentations of the regular representation are left ideals, and irreps L^μ_a are minimal left ideals. Thus finding all minimal left ideals will give all inequivalent irreps.

- We define projection operators P^μ_a onto the minimal left ideals L^μ_a by

$$P^\mu_a \widetilde{G} = L^\mu_a, \quad P^\mu_a = \text{identity on } L^\mu_a, \quad P^\mu_a P^\nu_b = \delta^{\mu\nu}\delta_{ab} P^\mu_a.$$

  We also define projection operators P^μ onto the left ideals L^μ by $P^\mu = \sum_a P^\mu_a$.

- Note that the projection operators commute with all of $\widetilde{G}$,

$$P^\mu_a r = r P^\mu_a.$$

  This can be shown by acting on an arbitrary element with decomposition $\sum_{\mu,a} |s^\mu_a\rangle$. Intuitively, it's because group multiplication can't take an element in or out of a left ideal.

Next, we construct the projection operators more explicitly.

- Decompose the identity element e as

$$e = \sum_\mu e_\mu, \quad e_\mu \in L^\mu.$$

  Then we claim the projection operators are given by right-multiplication by e_μ,

$$P^\mu |r\rangle = |r e_\mu\rangle.$$

  Using right-multiplication is useful, because P^μ automatically commutes with any r since left-multiplication and right-multiplication commute.

- To show this, consider an arbitrary element $r = \sum_\mu r_\mu$. Then

$$r = re = \sum_\mu r e_\mu$$

  but the fact that the L^μ are left-ideals means that $r_\mu = re_\mu$ by taking the μ component of both sides. Hence P^μ is indeed a projector.

- Finally, note that

$$e_\nu = e_\nu e = \sum_\mu e_\nu e_\mu$$

  and taking the μ component of both sides gives

$$e_\nu e_\mu = \delta_{\mu\nu} e_\mu.$$

  In particular, $e_\mu e_\mu = e_\mu$. Any algebra element satisfying this relation is called an idempotent and can be used to define a projection operator onto a left ideal.

- Similarly we may define e_μ^a for the L_μ^a. The e_μ^a are primitive idempotents, generating minimal left ideals, while the e_μ are not, as they can be written as a sum of idempotents, $e_\mu = \sum_a e_\mu^a$.
- An idempotent e_i is primitive if and only if $e_i r e_i = \lambda_r e_i$ for all $r \in \widetilde{G}$, where λ_r is a scalar.

  To prove the forward direction, note that e_i generates a minimal left ideal L, and right-multiplication by $e_i r e_i$ is also a projection operator onto L, which commutes with all elements of $\widetilde{G}$. Then by Schur's lemma, $e_i r e_i$ is proportional to the identity on L, giving the result.

  To show the converse, suppose that $ere = \lambda_r e$ and that e decomposes into idempotents as $e = e' + e''$ where $e'e'' = 0$. Then $ee'e = e'e = e'$ and hence $e' = \lambda e$. Then e' generates exactly the same left ideal as e does, so e is primitive.
- Two primitive idempotents e_1 and e_2 generate equivalent irreps if and only if $e_1 r e_2 \neq 0$ for some $r \in \widetilde{G}$.

  Let e_1 and e_2 generate the minimal left ideals L_1 and L_2. To prove the backward direction, note that $|q\rangle \to |qe_1re_2\rangle$ is a nonzero linear transformation from L_1 to L_2 which commutes with (left-multiplication by) any element of $\widetilde{G}$. Then by Schur's lemma, L_1 and L_2 are equivalent. The forward direction is similar.

Example. The reduction of the regular representation of C_3. Let the generator be a. The idempotent for the identity representation is always just the average of the group elements,

$$e_1 = \frac{1}{3}(e + a + a^{-1}).$$

Next, suppose another idempotent is $e_2 = xe + ya + za^{-1}$. We have the constraints

$$e_1 e_2 = 0, \quad e_2 e_2 = e_2$$

which yield the three solutions

$$e' = \frac{1}{3}(2e - a - a^{-1}), \quad e_+ = \frac{1}{3}(e + \omega a + \omega^{-1}a^{-1}), \quad e_- = \frac{1}{3}(e + \omega^{-1}a + \omega a^{-1}), \quad \omega = e^{2\pi i/3}.$$

One can see that e' is not primitive since it is the sum of e_+ and e_-. To check that e_+ and e_- are idempotent, we manually compute $e_\pm a e_\pm$ and $e_\pm a^{-1} e_\pm$ and apply the above theorem; this is simple since $e_\pm$ just pick up phase shifts upon multiplication by any group element. We do a similar check to show that e_+ and e_- generate inequivalent representations, giving the three irreps of C_3.

We can combine the group algebra with character theory to provide alternative proofs of some of the representation theory results above.

- For each conjugacy class c, we define the class average

  $$K(c) = \frac{1}{n_c}\sum_i g_i^{(c)}$$

  in the group algebra. Class averages can be multiplied together using the same operation as the group algebra, and it can be shown they also form an algebra called the class algebra, with

  $$K(c)K(d) = \sum_e \Gamma(c, d; e)K(e)$$

  where the Γ's are positive integers.

- The group and class algebra can be realized in representations, where the addition operation is simply matrix addition. For an irrep $D(g)$, the class average is

$$\mathcal{D}(c) = \frac{1}{n_c} \sum_{g \in c} D(g).$$

  By reindexing, we have $D(g'^{-1})\mathcal{D}(c)D(g') = \mathcal{D}(c)$, so by Schur's lemma

$$\mathcal{D}(c) = \frac{\chi(c)}{\chi(I)} I.$$

- Therefore, the identity

$$\mathcal{D}(c)\mathcal{D}(d) = \sum_e \Gamma(c, d; e)\mathcal{D}(e)$$

  translates to

$$\chi(c)\chi(d) = \chi(I) \sum_e \Gamma(c, d; e)\chi(e).$$

- Now we restore the irrep index r and sum over all irreps, giving

$$\sum_r \chi^r(c)\chi^r(d) = \sum_e \Gamma(c, d; e) \sum_r \chi^r(I)\chi^r(e) = \sum_e \Gamma(c, d; e)\chi^{\text{reg}}(e) = \Gamma(c, d; I)|G|$$

  where we used the decomposition of the regular representation.

- If c is a conjugacy class, then so is $\bar{c}$, the set of inverses of elements of c. Moreover, $\Gamma(c, d; I)$ is only nonzero for $d = \bar{c}$, giving

$$\sum_r \chi^r(c)\chi^r(d) \propto \delta^{d\bar{c}}.$$

  Since the representations are unitary, this implies

$$\sum_r \chi^r(c)^*\chi^r(c') \propto \delta^{cc'}.$$

  Thus, the rows of the character table are orthogonal, proving that $N(R) \leq N(C)$.

- A useful result is that in a representation D, the projector onto irreps isomorphic to D_i is

$$P_i = \frac{\dim D_i}{|G|} \sum_g \chi_i(g)^* D(g).$$

  To see this, note that P_i commutes with all group elements, $hP_ih^{-1} = P_i$, by reindexing. Thus P_i is block diagonal, proportional to the identity on each irrep. In particular, restricting to D_j,

$$\operatorname{tr} P_i|_{D_j} \propto \sum_g \chi_i(g)^*\chi_j(g) \propto (\chi_i, \chi_j) = \delta_{ij}.$$

  Hence P_i vanishes on all irreps not isomorphic to D_i and is the identity otherwise. The projector P_i works whether it acts on the left or the right.

- For example, $\sum_g D(g)$ projects onto the trivial irrep as we've seen above; it yields vectors that are fixed by the group. For the regular representation, the projectors are the e_μ.

2.6 ○ Representations of S_n

We now use our machinery to find the irreps of S_n. We warm up with the one-dimensional irreps.

- All transpositions are conjugate, so we must assign them all the same number. Since transpositions have order 2, the number must be ± 1. Since transpositions generate the entire group, there are only two distinct one-dimensional representations.
- Denote the sign of a permutation p by $(-1)^p$. Then the two one-dimensional representations map p to one, and p to $(-1)^p$. The corresponding primitive idempotents, up to a constant, are

$$s = \sum_p p, \quad a = \sum_p (-1)^p p.$$

- To check this, note that for any q,

$$qs = sq = s, \quad qa = aq = (-1)^q a, \quad sa = as = 0.$$

  Then we have $sqs = ss = n!s$ and $aqa = (-1)^q aa = (-1)^q n!a$, so both are indeed primitives. The corresponding irreps have basis vectors $|qs\rangle$ and $|qa\rangle$ for some arbitrary q. To check the irreps are not equivalent, note that $sqa = sa = 0$ for any q.

For higher-dimensional irreps, it is useful to introduce Young diagrams.

- A partition $\lambda = \{\lambda_1, \ldots, \lambda_r\}$ of n is a sequence of positive integers λ_i satisfying

$$\lambda_i \geq \lambda_{i+1}, \quad \sum_{i=1}^{r} \lambda_i = n.$$

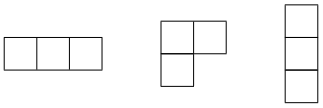
  We say $\lambda > \mu$ if the first nonzero number in the sequence $\lambda_i - \mu_i$ is positive.
- A partition λ is represented by a Young diagram, which consists of n squares arranged in r rows, where row i contains λ_i squares.
- Partitions of n are in one-to-one correspondence with possible cycle structures of permutations, where elements that are left alone are regarded as 1-cycles. Then there is one conjugacy class and hence one irrep for every Young diagram. For example, the identity element is $1+1+\ldots+1$, and a transposition is $2+1+\ldots+1$.
- A Young tableau is obtained by labeling the squares of a Young diagram with the numbers 1 through n, using each number once. The normal Young tableau Θ_λ associated with λ has the numbers in left-to-right, top-to-bottom order. A standard Young tableau is one where the numbers always increase when going down or to the right.
- A permutation acts on a Young tableau by permuting the numbers. Then an arbitrary Young tableau can be written as $p\Theta_\lambda = \Theta_\lambda^p$.
- Given a Young tableau Θ_λ^p, the horizontal permutations h_λ^p are the ones which only permute numbers within rows, and the vertical permutations v_λ^p are the ones which only permute numbers within columns.

- The symmetrizer s^p_λ, antisymmetrizer a^p_λ, and irreducible symmetrizer or Young symmetrizer e^p_λ associated with the Young tableau Θ^p_λ are defined as

$$s^p_\lambda = \sum_h h^p_\lambda, \quad a^p_\lambda = \sum_v (-1)^{v_\lambda} v^p_\lambda, \quad e^p_\lambda = s^p_\lambda a^p_\lambda = \sum_{h,v} (-1)^{v_\lambda} h^p_\lambda v^p_\lambda.$$

  Our main result will be that e^p_λ is a primitive idempotent. To avoid confusion, note that the Young symmetry acts on the right, and multiplication of permutations obeys $f \star g = f \circ g$ when $f, g \in S_n$ are viewed as automorphisms of S_n.

Example. The simple example of S^3. There are three Young diagrams, shown below.



Labeling the corresponding normal Young tableaux Θ_1, Θ_2, and Θ_3, we have

- $s_1 = s$, $a_1 = e$, $e_1 = s$.
- $s_2 = e + (12)$, $a_2 = e - (13)$, $e_2 = s_2 a_2 = e + (12) - (13) - (321)$.
- $s_3 = e$, $a_3 = a$, $e_3 = a$.
- The fourth standard Young tableau, $\Theta^{(23)}_2$, has $s^{(23)}_2 = e + (13)$, $a^{(23)}_2 = e - (12)$, and $e^{(23)}_2 = e + (13) - (12) - (123)$.

We note the following features of the example and state they hold generally without proof.

- The horizontal permutations form a subgroup and s_λ is its symmetrizer. Then by similar logic to before, s_λ is idempotent with $s_\lambda s_\lambda = |\{h_\lambda\}| s_\lambda$. Similarly, a_λ is idempotent.
- Generally, neither s_λ nor a_λ are primitive, but e_λ is. Moreover, e^p_λ generates an irrep distinct from but isomorphic to that of e_λ. All nonisomorphic irreps correspond to exactly one e_λ.
- The four minimal left ideals generated by the idempotents of the standard Young tableaux span the entire group algebra space S_3. Explicitly, $e = (1/6)(e_1 + 2e_2 + 2e^{(23)}_2 + e_3)$.
- Hence in general, we claim that normal Young tableaux correspond to inequivalent irreps, while standard Young tableaux decompose the regular representation.
- Since a d-dimensional irrep occurs d times in the regular representation, the dimension of the irrep corresponding to a Young diagram is the number of standard Young tableaux on it.
- For each box in a Young diagram, a hook is a right angle with vertex in that box, which opens down and to the right; the length of the hook h_i is the number of boxes it contains. For each hook, the box with the vertex is the smallest number with probability $1/h_i$ so we expect that

$$\dim(e_\lambda \text{ irrep}) = \text{standard tableaux} = \frac{n!}{\prod_i h_i}.$$

  This is called the hook length formula.

- The proof above is incorrect because we cannot simply multiply probabilities for nonindependent events. However, it serves as a useful mnemonic, since the real proof is much more difficult.

Example. We give a few more details. First, we explicitly construct the irrep generated by e_2.

- The irrep certainly contains e_2 itself, i.e. $e + (12) - (13) - (321)$. It is convenient to write the vectors in terms of the image of the string "123" under the corresponding permutations, so that $e_2 = |123\rangle + |213\rangle - |321\rangle - |312\rangle$.
- The irrep also contains all the vectors $|pe_2\rangle$, which can be written in our notation by applying a permutation p to the numbers in the states. For example, for $p = (12)$, we get

$$|123\rangle + |213\rangle - |321\rangle - |312\rangle \quad \to \quad |213\rangle + |123\rangle - |312\rangle - |321\rangle.$$

  while for $p = (13)$ we get

$$|123\rangle + |213\rangle - |321\rangle - |312\rangle \quad \to \quad |321\rangle + |231\rangle - |123\rangle - |132\rangle.$$

  Computing the other permutations, we find a two-dimensional irrep, in accordance with the hook length formula.
- The ordering is a bit confusing. Given a permutation p, the Young symmetrizer is applied on the right, giving pe_2. If the $\{pe_2\}$ are an irrep, then gpe_2 is in their span for any g. But to see how pe_2 acts on an element g of the regular representation, we compute pe_2g.

Example. Symmetry classes of tensors. Given an m-dimensional real vector space V_m, the tensor product space V_m^n of rank n tensors is acted on by $GL(m, \mathbb{R})$, and we would like to decompose V_m^n into irreps of $GL(m, \mathbb{R})$. The symmetric group enters because S_n acts on $GL(m, \mathbb{R})$ by permuting the indices, and this commutes with the action of $GL(m, \mathbb{R})$, which does the same thing to each index. Therefore, irreps of $GL(m, \mathbb{R})$ in V_m^n have definite symmetry; specifically they are

$$\{e_\lambda^p|\alpha\rangle \,|\, |\alpha\rangle \in V_m^n\}$$

for each λ and p. These are tensors of the symmetry class Θ_λ^p. Tensors of the symmetry type Θ_λ correspond are the set of all tensors of the symmetry class Θ_λ^p for all p.

As an example, consider rank 3 tensors with $m = 2$, and the same numbering as in the previous example. Then $e_3 = a$ annihilates everything, while $e_1 = s$ gives the irrep

$$\{|+++\rangle, s|++-\rangle, s|+--\rangle, |---\rangle\}$$

containing all totally symmetric tensors. The irrep given by e_2 is spanned by

$$e_2|++-\rangle = 2|++-\rangle - |-++\rangle - |+-+\rangle$$

and

$$e_2|--+\rangle = 2|--+\rangle - |+--\rangle - |-+-\rangle$$

where e_2 acting on any other basis element gives zero or one of these two. Finally, $e_2^{(23)}$ gives the last irrep, accounting for all of the states. Note that in general, some irreps will correspond to non-standard Young tableaux, since nothing here is sensitive to the order of the indices.

Note. One might be tempted to say the e_2 irrep contains tensors symmetric in the first two indices and antisymmetric in the first and third, but this is wrong: the symmetrization partially destroys the result of the antisymmetrization. In fact, it's simply impossible for an index to be simultaneously symmetrized with one and antisymmetrized with another. If a rank 3 tensor is symmetric in its first indices and antisymmetric in the last two,

$$T_{abc} = T_{bac} = -T_{bca} = -T_{cba} = T_{cab} = T_{acb} = -T_{abc}$$

so the tensor is identically zero.

Note. Different sources can differ on the action of the e_λ in several ways.

- While we antisymmetrize first here, one could also symmetrize first; this also yields irreps. However, we generally won't get the same irreps. Our convention above splits the e_2 symmetry class into two irreps both with symmetry in two indices, but we could also split it into two irreps both with antisymmetry in two indices.
- This alternate splitting is legal, for the same reason that we can choose any basis in a degenerate subspace: we have two copies of the same irrep, so any "change of basis" between them will also yield two irreps.
- We have taken the permutations to act on the slots, so that

  $$(23)(12)|ijk\rangle = (23)|jik\rangle = |jki\rangle.$$

  However, we could also label the permutation with the same letters we index the basis of V_m^n, then act on the letters. That is, we could have

  $$(jk)(ij)|ijk\rangle = (jk)|jik\rangle = |kij\rangle$$

  where in the last step we swapped the letters j and k, not the last two slots.
- The latter is more intuitive if we are thinking in terms of tensor components, where we usually act on letters ("antisymmetrize i and j") rather than slots. Swapping between these two conventions is equivalent to swapping the order of operations: if the first convention is $g_1 g_2$, then the second is equivalent to doing $(g_2 g_1 g_2^{-1})g_2 = g_2 g_1$. Indeed, we have

  $$(12)(23)|ijk\rangle = (12)|ikj\rangle = |kij\rangle.$$

- Finally, we may phrase everything in terms of transforming the basis states $|ijk\rangle$ or transforming the tensor components $\langle ijk|\alpha\rangle$. Generally, these transform oppositely. For example, we have

  $$(\alpha')^{ijk} = \langle ijk|(23)(12)|\alpha\rangle = \langle ikj|(12)|\alpha\rangle = \langle kij|\alpha\rangle = \alpha^{kij}.$$

  If we wanted to compute this directly by moving the tensor indices around, we would first swap the second and third indices, then swap the first and second indices. That is, the order of operations is flipped because we're acting on $|ijk\rangle$ from the right.
- In summary, there are three possible convention differences, each of which is equivalent to flipping the order of symmetrization and antisymmetrization. Our convention matches that of most pure math sources. Most physics sources flip all three, ending up with a net order flip.

For concrete computations, we'll generally use the physics convention. The difference between the two generally won't matter because we're mostly concerned with how tensor products of irreps decompose, or how irreps split upon restricting to subgroups.

Example. Consider the symmetry classes of the Riemann tensor. We know that

$$R_{abcd} = -R_{bacd}, \quad R_{abcd} = -R_{abdc}, \quad R_{abcd} = R_{cdab}, \quad R_{[abcd]} = 0.$$

Taking the physics convention, the first two identities indicate that ab and cd have to be in columns, giving the candidates

$$\begin{array}{|c|c|}\hline a & c \\ \hline b & d \\ \hline \end{array} \qquad \begin{array}{|c|}\hline a \\ \hline b \\ \hline c \\ \hline d \\ \hline \end{array}.$$

Both satisfy the third identity, but the fourth rules out the second candidate. Hence the Riemann tensor is a $GL(n, \mathbb{R})$ irrep. Restricting to a subgroup breaks apart irreps, and indeed restricting to the Lorentz transformations $SO(n-1, 1)$, the Riemann tensor decomposes into the Ricci scalar, Ricci tensor, and Weyl tensor.

3 Lie Groups

3.1 ⊛ Matrix Groups

We review some basic examples of matrix Lie groups.

- Let $\mathrm{Mat}(n, \mathbb{F})$ denote the set of $n \times n$ matrices with entries in $\mathbb{F}$, which will be $\mathbb{R}$ or $\mathbb{C}$. It is not a group, because not all matrices are invertible, so we define $GL(n, \mathbb{F})$ to be the subset of invertible matrices.

- We define the special linear group $SL(n, \mathbb{F})$ to be the subset of $GL(n, \mathbb{F})$ with unit determinant. In general, 'special' stands for 'unit determinant'. Considering these groups as real manifolds,

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

  The unit determinant constraint $\det M = 1$ is one constraint over $\mathbb{R}$ and two constraints over $\mathbb{C}$, as it sets $\det M = 1 + 0i$.

- A Lie group is a group that is a smooth manifold, where the group operations are smooth; we'll skip the proof that the groups above are Lie groups. We define a Lie subgroup to be a subgroup of a Lie group that is also a smooth submanifold; one can show that Lie subgroups are Lie groups in themselves.

- The orthogonal group is

  $$O(n) = \{M \in GL(n, \mathbb{R}) \,|\, M^T M = 1\}$$

  which implies $\det M = \pm 1$. The subset $SO(n)$ is called the proper rotations, and

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

  To show this, note that $M^T M$ is symmetric, so subtracting the constraints gives $n^2 - n(n+1)/2$. Alternatively, note that $n(n-1)/2$ is the number of independent planes in n dimensions.

- More generally, define $O(p, q)$ as the subset of $GL(n, \mathbb{R})$, with $n = p + q$, where

  $$M^T \eta M = \eta, \quad \eta = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}.$$

  For example, the Lorentz group is $O(3, 1)$. It is noncompact, and splits into four components since we also have time reversal. The dimension is still $n(n-1)/2$.

- Define the unitary groups as

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

  Unitary matrices preserve length and have determinants with unit norm. We have

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

  To derive this, note that we start with $2n^2$ degrees of freedom, and the matrix $U^\dagger U = 1$ is Hermitian. Then it contains $n(n-1)$ constraints on the off-diagonal, but only n constraints on the diagonal, because the diagonal is automatically real. Switching to the special unitary group gives one constraint since $\det U$ is a phase.

- As a simple example, $U(1) \cong SU(2)$. This is the first of a few 'accidental' Lie group isomorphisms which we'll understand later in terms of the classification of Lie algebras.
- The elements of $SU(2)$ can be parametrized as

$$U = a_0 1 + i\mathbf{a} \cdot \boldsymbol{\sigma}, \quad a_0^2 + a_1^2 + a_2^2 + a_3^2 = 1$$

  where the $\boldsymbol{\sigma}$ are the Pauli matrices. Thus $SU(2) \cong S^3$ as a manifold.
- The centers of these groups only contain elements proportional to the identity, by Schur's lemma and the fact that the fundamental representation is irreducible. For example, the center of $U(n)$ is $\{e^{i\theta} I\} \cong U(1)$, while the center of $SU(n)$ is $\{e^{2\pi i k/n} I\} \cong \mathbb{Z}_n$.
- By varying the metric, we may also define $O(p,q)$, $SU(p,q)$, and $U(p,q)$. Since Lie algebras only depend on infinitesimal structure, changing the signature doesn't change the dimension. The $SU(n)$ and $SO(n)$ groups are often used in particle physics, while $SU(p,q)$ and $SO(p,q)$ appear more often in string dualities.
- Globally, the (special) unitary groups are all connected and simply connected, and compact when the signature is definite. The groups $SO(n)$ are connected and compact, but not simply connected, with double cover $\mathrm{Spin}(n)$.

Note. Generally, when we define a subset of $\mathbb{R}^n$ by constraint equations, the result is not necessarily a manifold; we get an object called an algebraic variety which may have singular points. In this case the group structure forbids this: if there were a singularity at g_1, then there must be a singularity at any other group element g_2, since the action of multiplication by $g_2 g_1^{-1}$ is smooth. But varieties cannot be singular everywhere, so the group must be smooth everywhere.

Example. We consider the possible eigenvalues of $M \in O(n)$. Since M is real, eigenvalues come in complex conjugate pairs. Moreover, since M preserves lengths, they have norm 1. Thus for $n = 2$ we have eigenvalues $e^{\pm i\theta}$ and for $n = 3$ we have eigenvalues 1 and $e^{\pm i\theta}$. The eigenvector with $\lambda = 1$ specifies the axis of rotation.

The general group element of $SO(3)$ can be written as

$$M(\hat{\mathbf{n}}, \theta)_{ij} = \cos\theta\, \delta_{ij} + (1 - \cos\theta) n_i n_j - \sin\theta\, \epsilon_{ijk} n_k.$$

This is redundant, since $M(\hat{\mathbf{n}}, 2\pi - \theta) = M(-\hat{\mathbf{n}}, \theta)$. To remove this we restrict $\theta \in [0, \pi]$ and identify $(\hat{\mathbf{n}}, \pi)$ with $(-\hat{\mathbf{n}}, \pi)$. Then the $SO(3)$ group manifold is obtained by taking the ball B_3 and identifying antipodal points on the boundary.

Example. There is an embedding of $SU(n)$ in $SO(2n)$. This is intuitive because each complex dimension can be thought of as two real dimensions. Formally, the map is

$$U = U_1 + iU_2 \mapsto \begin{pmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{pmatrix}$$

where U_1 and U_2 are real. This is an orthogonal matrix, and it is in $SO(2n)$ since $SU(n)$ is connected.

3.2 ⊛ Lie Algebras

We begin with some fundamental definitions.

- A Lie algebra $\mathfrak{g}$ is a vector space with a bracket operation $[\cdot,\cdot]:\mathfrak{g}\times\mathfrak{g}\to\mathfrak{g}$ satisfying antisymmetry, linearity, and the Jacobi identity,

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

  Its dimension is the dimension of $\mathfrak{g}$ as a vector space. The Lie algebra is real if it is over $\mathbb{R}$ and complex if it is over $\mathbb{C}$.

- For any vector space V, given an associative linear product $\star:V\times V\to V$, the commutator $[X,Y]=X\star Y-Y\star X$ is a Lie bracket. For example, for the vector space of matrices such a product is matrix multiplication.

- For a Lie algebra $\mathfrak{g}$ over $\mathbb{F}$ with basis T^a, we define

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

  where the $f^{ab}{}_c\in\mathbb{F}$ are called the structure constants. They are antisymmetric in a and b by the antisymmetry of the bracket, while the Jacobi identity gives

$$f^{ab}{}_cf^{cd}{}_e+f^{da}{}_cf^{cb}{}_e+f^{bd}{}_cf^{ca}{}_e=0.$$

  Here one lower index is free, another is contracted, and the rest are cyclically permuted.

- A Lie algebra isomorphism is a map $\phi:\mathfrak{g}\to\mathfrak{g}'$ which is an isomorphism of vector spaces and preserves the bracket. A Lie algebra homomorphism is defined similarly.

- A subalgebra $\mathfrak{h}\subset\mathfrak{g}$ is a subspace of $\mathfrak{g}$ which satisfies $[\mathfrak{h},\mathfrak{h}]\subset\mathfrak{h}$. An ideal of $\mathfrak{g}$ is a subalgebra which satisfies $[\mathfrak{h},\mathfrak{g}]\subset\mathfrak{h}$. We now give some examples of ideals.
  - The empty set and all of $\mathfrak{g}$ are trivially ideals; all others are nontrivial ideals.
  - The center $\mathfrak{z}(\mathfrak{g})$ is the set of elements in $\mathfrak{g}$ which have vanishing bracket with all of $\mathfrak{g}$.
  - If $\mathfrak{h}$ and $\mathfrak{k}$ are ideals, so are $\mathfrak{h}+\mathfrak{k}$, $\mathfrak{h}\cap\mathfrak{k}$, and $[\mathfrak{h},\mathfrak{k}]$.
  - The kernel of any Lie algebra homomorphism is an ideal; this is really the motivation behind the definition. They are analogous to normal subgroups of groups.

- We write $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{k}$ if this holds considering the Lie algebras as vector spaces, and $[\mathfrak{h},\mathfrak{k}]=0$.

- An abelian Lie algebra is one where the bracket identically vanishes, i.e. $\mathfrak{z}(\mathfrak{g})=\mathfrak{g}$.

- We say $\mathfrak{g}$ is simple if it is nonabelian and has no nontrivial ideals, and $\mathfrak{g}$ is semi-simple if it has no abelian ideals; then semi-simple Lie algebras are direct sums of simple ones.

Note. For completeness, we'll say a bit more about the general structure of Lie algebras.

- The derived algebra $\mathfrak{i}=\mathrm{span}([\mathfrak{g},\mathfrak{g}])$ is the set of linear combinations of brackets in $\mathfrak{g}$. A Lie algebra is perfect if $\mathfrak{i}=\mathfrak{g}$, so all semi-simple Lie algebras are perfect.

- The derived series is defined by $\mathfrak{i}_n = \text{span}([\mathfrak{i}_{n-1}, \mathfrak{i}_{n-1}])$ with $\mathfrak{i}_1 = \mathfrak{i}$, and a Lie algebra is solvable if $\mathfrak{i}_n$ is empty for some n. Intuitively a solvable Lie algebra is "almost commutative", while a semisimple Lie algebra is the opposite: we could have equivalently defined a semi-simple Lie algebra as one with no solvable ideals.

- As an example, the subalgebra of upper triangular matrices in $\mathfrak{gl}(n, \mathbb{F})$ is solvable, as taking repeated commutators sets more and more diagonals to zero. A theorem due to Lie states that all representations of a solvable Lie algebra $\mathfrak{g}$ are upper-triangular in some basis, and moreover every complex irrep of a solvable Lie algebra is one-dimensional. Hence solvable Lie algebras behave just like abelian ones in terms of representation theory.

- The radical $\text{rad}(\mathfrak{g})$ is the sum of all solvable ideals of $\mathfrak{g}$. Then there is a short exact sequence

$$0 \to \text{rad}(\mathfrak{g}) \to \mathfrak{g} \to \mathfrak{g}/\text{rad}(\mathfrak{g}) \to 0$$

  where $\mathfrak{g}/\text{rad}(\mathfrak{g})$ is semisimple.

- Levi's theorem states that, furthermore, every Lie algebra is the direct sum of a solvable Lie algebra and a semisimple one, which is in turn the direct sum of simple Lie algebras.

- For example, the Poincare algebra decomposes as $\mathbb{R}^4 \oplus \mathfrak{so}(3,1)$. Note that the semisimple part is not an ideal. In all complex irreps, elements in the radical act by scalar multiplication.

We now formally relate Lie groups and Lie algebras, for the case of matrix Lie groups.

- Given an n-dimensional matrix Lie group G, the tangent space $\mathfrak{g} = T_e(G)$ is an n-dimensional vector space. Given coordinates $\boldsymbol{\theta}$ on G, the tangent space is spanned by the vectors

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

  The bracket operation is simply the matrix commutator, which we know satisfies the axioms.

- The nontrivial step is to show closure under the bracket. Consider two paths

$$g_i(t) = 1 + X_i t + W_i t^2 + O(t^3)$$

  which satisfy

$$g_1(t)g_2(t) = 1 + (X_1 + X_2)t + (X_1X_2 + W_1 + W_2)t^2 + O(t^3)$$

  with a similar expression for $g_2(t)g_1(t)$. Then $h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t)$ obeys

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

  and hence we have

$$h(t) = 1 + [X_1, X_2]t^2 + O(t^3).$$

  Then $h(\sqrt{t})$ is a curve with tangent vector $[X_1, X_2]$, as desired. We see the Lie bracket is the infinitesimal version of the commutator of group elements.

We now give some examples of matrix Lie algebras. Note that all these examples are real Lie algebras even though the matrices are complex; the dimension are real dimensions. A complex Lie algebra would instead give us a complex Lie group, which lives on a complex manifold; we will not consider such objects here.

- For $GL(n,\mathbb{F})$, the Lie algebra is $\mathrm{Mat}(n,\mathbb{F})$ since the determinant is continuous. Restricting to $SL(n)$ restricts the Lie algebra to contain matrices with zero trace, because

$$\det \exp X = \exp \operatorname{tr} X.$$

  This provides an easy way to compute $\dim GL(n,\mathbb{F})$, as it is equal to $\dim \mathrm{Mat}(n,\mathbb{F})$.

- For $O(n)$, we have $R^T R = 1$ and setting $R = 1 + X$ gives $X^T = -X$. Note that the Lie algebra of $SO(n)$ is the same, because antisymmetric matrices are automatically traceless.

- Similarly, for $U(n)$, we get $X^\dagger = -X$. We restrict to $SU(n)$ by requiring $\operatorname{tr} X = 0$, which counts as a single constraint since the trace of an anti-Hermitian matrix is imaginary.

- We consider the structure of $\mathfrak{su}(2)$ in detail. One basis is

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

  This differs from the angular momentum by a factor of i, since it is anti-Hermitian. Now

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

  which implies that in our basis, $f^{ab}{}_c = \epsilon_{abc}$.

- Similarly, we may define the following basis for $\mathfrak{so}(3)$,

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

  where has the properties

$$(T^a)_{bc} = -\epsilon_{abc}, \quad [T^a, T^b] = \epsilon_{abc} T^c$$

  which establishes that $\mathfrak{su}(2) \cong \mathfrak{so}(3)$.

Next we define some maps on the Lie group and Lie algebra.

- Left translation L_h is a diffeomorphism of G with corresponds to multiplying by h on the left,

$$L_h g = hg.$$

  Taking the differential gives a map $L_h^* : T_g \to T_{hg}$.

- For matrix Lie groups, L_h^* is implemented by matrix multiplication. That is, for $X \in \mathfrak{g}$,

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

  To prove this explicitly, note that a curve $g(t) = I + tX + O(t^2)$ maps by L_h to the curve $h(t) = h + thX + O(t^2)$ which passes through h with tangent vector hX.

- Given $X \in \mathfrak{g}$, we may define the left-invariant vector field V by

$$V(g) = L_g^*(X).$$

  Conversely, if a manifold has a Lie group structure, this means it has a global field of frames, which is a strong constraint; it means that S^2 cannot be a Lie group.

- An integral curve $g(t)$ of the left-invariant vector field V associated with $X \in \mathfrak{g}$ satisfies

$$\frac{dg(t)}{dt} = V(g(t)) = L_{g(t)}^*(X) = g(t)X, \quad g(t) = \exp(tX)$$

  where the exponential for a matrix Lie group is defined as a series.

- The curve $\exp(tX)$ defines a one-parameter subgroup with

$$\exp(t_1 X)\exp(t_2 X) = \exp(t_2 X)\exp(t_1 X) = \exp((t_1 + t_2)X)$$

  by expanding out the series. Note that the subgroup might be isomorphic to either $\mathbb{R}$ or $U(1)$.

- Setting $t = 1$, we have a map $\exp : \mathfrak{g} \to G$ which is bijective in a neighborhood of the identity. It is not injective if G has a $U(1)$ subgroup, and it is obviously not surjective if G is not connected.

- If G is connected, it turns out that the exponential map is not necessarily surjective if G is not compact, though any group element can still be written as the product of two exponentials. On the other hand, it turns out that the exponential map *is* surjective for the Lorentz group, which isn't compact, but the proof is difficult and relies on particular properties of that group.

- Writing $g_X = \exp(X)$, the Baker-Campbell-Hausdorff theorem states

$$g_X g_Y = g_Z, \quad Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \dots.$$

  Thus the group multiplication law is encoded in the Lie bracket.

Example. The group $SU(2)$ is a double cover of $SO(3)$ by the map

$$d : SU(2) \to SO(3), \quad d(A)_{ij} = \frac{1}{2}\operatorname{tr}(\sigma_i A \sigma_j A^\dagger)$$

where $\ker d = \{\pm I\}$, as can be checked by using the identity

$$\sum_i (\sigma_i)_{\alpha\beta}(\sigma_i)_{\delta\gamma} = 2\delta_{\alpha\gamma}\delta_{\delta\beta} - \delta_{\alpha\beta}\delta_{\delta\gamma}.$$

The inverse of this map is

$$A = \pm\frac{I_2 + \sigma_i R_{ij} \sigma_j}{2\sqrt{1 + \operatorname{tr} R}}.$$

3.3 ⊛ Representations

We now define representations for Lie group and Lie algebras. We will be casual with notation, writing equations that only work for matrix Lie groups.

- A representation $D(g)$ of a Lie group is a homomorphism $D : G \to \mathrm{Mat}_n(\mathbb{F})$, or alternatively an action of G on V by linear transformations. A representation of a Lie algebra is a linear map $d : \mathfrak{g} \to \mathrm{Mat}_n(\mathbb{F})$ which preserves the bracket. Here, n is the dimension of the representation, which may be different from the dimension m of the group and algebra.

- Every representation D of G corresponds to a representation d of $\mathfrak{g}$ by

$$d(X) = \frac{d}{dt} D(g(t))\bigg|_{t=0}$$

  for a curve $g(t)$ passing through the identity at $t = 0$ with tangent vector X. That is,

$$D(1_m + tX + O(t^2)) = 1_n + td(X) + O(t^2).$$

- Linearity is easy to see, but checking the bracket is preserved is a bit trickier. Consider curves $g_1(t)$ and $g_2(t)$ as defined above. Then as before,

$$h(t) = g_1^{-1}(t)g_2^{-1}(t)g_1(t)g_2(t) = 1_m + t^2[X_1, X_2] + O(t^3).$$

  Taking the representation of both sides,

$$D(h(t)) = 1_n + t^2 d([X_1, X_2]) + O(t^3).$$

  On the other hand, we can also write

$$D(h(t)) = D(g_1^{-1}(t))D(g_2^{-1}(t))D(g_1(t))D(g_2(t)) = 1_n + t^2[d(X_1), d(X_2)] + O(t^3)$$

  by the same logic, giving the result.

- Similarly, given a representation $d(X)$, we can define $D(g) = \exp d(X)$ where $g = \exp X$. Since the exponential map is not generally bijective, this may not be a representation of G, but it is at least a representation 'locally'. To verify it is a representation, note that if $g_1 = \exp(X_1)$ and $g_2 = \exp(X_2)$,

$$\begin{aligned} D(g_1 g_2) &= \exp\left(d\left(X_1 + X_2 + \frac{1}{2}[X_1, X_2] + \dots\right)\right) \\ &= \exp\left(d(X_1) + d(X_2) + \frac{1}{2}[d(X_1), d(X_2)] + \dots\right) \\ &= \exp(d(X_1)) \exp(d(X_2)) = D(g_1)D(g_2) \end{aligned}$$

  where we applied the Baker-Campbell-Hausdorff theorem twice.

- The representation space of a representation of a Lie algebra $\mathfrak{g}$ is also called a $\mathfrak{g}$-module. In general, for a ring R, an R-module is simply a vector space where the scalars are elements of R. A Lie algebra is a ring where the operation is the bracket, and 'scalar multiplication' by X in the $\mathfrak{g}$-module is application of $d(X)$.

We now give some basic examples of Lie group and Lie algebra representations.

- Starting with the trivial representation $D(g) = I$, we get the trivial representation $d(X) = 0$.
- Starting from the fundamental representation $D(g) = g$, for an n-dimensional matrix Lie group, we have the fundamental representation $d(X) = X$, which is also n-dimensional.
- The adjoint representation of G has representation space $\mathfrak{g}$, and

  $$D(g)X = (\operatorname{Ad} g)X = gXg^{-1}.$$

  To show the action of $\operatorname{Ad} g$ is closed, note that if X is the tangent vector of a curve $h(t)$, then $(\operatorname{Ad} g)X$ is the tangent vector of a curve $gh(t)g^{-1}$.
- The adjoint representation of G corresponds to the adjoint representation of $\mathfrak{g}$,

  $$d(X) = \operatorname{ad}_X, \quad \operatorname{ad}_X(Y) = [X, Y].$$

  whose dimension is also $\dim \mathfrak{g}$. Concretely, expand in a basis T^a of $\mathfrak{g}$, giving

  $$\operatorname{ad}_X(Y) = [X, Y] = X_a Y_b [T^a, T^b] = X_a Y_b f^{ab}{}_c T^c.$$

  Therefore, we have

  $$[\operatorname{ad}_X(Y)]_c = (X_a f^{ab}{}_c) Y_b, \quad (\operatorname{ad}_X)^b_c = X_a f^{ab}{}_c.$$
- If we didn't know how the adjoint representation was derived, we would have to check that

  $$[\operatorname{ad}_X, \operatorname{ad}_Y] = \operatorname{ad}_{[X,Y]}$$

  where $[\cdot, \cdot]$ means the commutator on the left and the Lie bracket on the right. This follows directly from the Jacobi identity, so one might say the point of the Jacobi identity for an abstract Lie algebra is to ensure that the adjoint representation still exists.
- If $\mathfrak{g}$ is semi-simple, then the adjoint representation is faithful, because its kernel is the center of $\mathfrak{g}$, i.e. an abelian ideal. If $\mathfrak{g}$ is simple, then the adjoint representation is irreducible, because if $\mathfrak{h} \subset \mathfrak{g}$ were an nontrivial invariant subspace then $\mathfrak{h}$ would be a nontrivial ideal.

Note. Below are some powerful general results, which we will not prove.

- For every Lie group G there is a bijection between connected Lie subgroups $H \subseteq G$ and Lie subalgebras $\mathfrak{h} \subset \mathfrak{g}$ given by the tangent map.
- Any finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to the Lie algebra of a Lie group. Specifically, there is a unique connected simply-connected Lie group G with Lie algebra $\mathfrak{g}$, and all others are of the form G/Z for a discrete central subgroup $Z \subseteq G$.
- If G_1 and G_2 are Lie groups and G_1 is connected and simply connected, then $\operatorname{Hom}(G_1, G_2) = \operatorname{Hom}(\mathfrak{g}_1, \mathfrak{g}_2)$. Combining this with the previous point, the categories of finite-dimensional Lie algebras and connected, simply-connected Lie groups are equivalent.
- A crucial ingredient in the proof of this statement is Ado's theorem, which states that every Lie algebra is isomorphic to a matrix Lie algebra. This allows us to focus on matrix Lie algebras without loss of generality.

3.4 Integration

In this section, we define a measure for integration over a Lie group.

- Previously, we showed the orthogonality of characters for a finite group. We would like to do the same for a continuous group, but this requires replacing the sum $\sum_g$ with an integration measure $\int d\mu(g)$.
- In the case of a finite group, the crucial step was being able to 'shift the sum',

  $$A = \sum_g D^\dagger(g) X D(g) \text{ satisfies } D^\dagger(g) A D(g) = A.$$

  For a Lie group, we have

  $$D^\dagger(g) \left(\int d\mu(g')\, D^\dagger(g') X D(g') \right) D(g) = \int d\mu(g')\, D^\dagger(g'g) X D(g'g) = \int d\mu(g'g^{-1})\, D^\dagger(g') X D(g').$$

  Then the analogous requirement is $d\mu(g) = d\mu(g')$ for any two group elements.
- More concretely, suppose we take a small patch of the group manifold. The patch can be moved around by multiplication by a group element, and we demand this leaves the measure of the patch invariant. Alternatively, we are defining a volume form invariant under group multiplication, which shows that all Lie groups are orientable.
- The 'volume' of a group is $\int d\mu(g)$. Unlike the cardinality of a group, it has no canonical normalization; it is finite when the group is compact.
- Expanding in terms of coordinates, we are requiring

  $$d\mu(g) = dx^1 \dots dx^n \rho(x^1, \dots, x^n), \quad dx^1 \dots dx^n \rho(x^1, \dots, x^n) = dx'^1 \dots dx'^n \rho(x'^1, \dots, x'^n)$$

  where the primed quantities are defined by group multiplication.
- As a first example, consider $SO(2)$. If we parametrize by θ, then the group multiplication law $R(\theta)R(\theta') = R(\theta + \theta')$ means that a segment of length $\delta\theta$ is mapped to a segment of length $\delta\theta$. Then we have $\rho(\theta) = \rho(\theta')$ for any two angles, so $d\mu(\theta) = d\theta$. In this case, orthogonality of characters recovers Fourier series.
- Next, consider the restricted Lorentz group $SO(1,1)$. If we parametrize by rapidity φ, we have $L(\varphi)L(\varphi') = L(\varphi + \varphi')$ so $d\mu(\varphi) = d\varphi$. On the other hand, if we parametrize by velocity, we have to calculate. We have

  $$L(u)L(v) = L(v'), \quad v' = \frac{v+u}{1+vu}.$$

  Now consider the segment $[v, v + dv]$. A direct computation gives

  $$dv' = \frac{1-u^2}{(1+uv)^2} dv, \quad \rho(v) = \frac{1-u^2}{(1+uv)^2} \rho\left(\frac{v+u}{1+uv} \right).$$

  Finally, setting $v = 0$ gives $\rho(u) = \rho(0)/(1-u^2)$. We could also have found this by changing variables from rapidity, picking up a Jacobian. In both cases, the group volume is infinite.

- The analogue of the great orthogonality theorem is called the Peter-Weyl theorem, and it applies to all compact Lie groups. It contains as special cases many of the orthogonality and completeness results we've used.

As an application, we consider the extended example of $SO(3)$.

- We recall that elements $SO(3)$ can be parametrized as $R(\mathbf{n}, \psi)$ where ψ is the rotation angle. Hence $SO(3)$ is a three-dimensional ball with opposite points on the boundary identified. The equivalence classes are rotations with the same ψ.

- Now consider computing the character of the spin–j representation of $SO(3)$ on the equivalence class with angle ψ. It is convenient to choose $\mathbf{n} = \mathbf{z}$, so that

$$R(\mathbf{z}, \psi)|jm\rangle = e^{i\psi J_3}|jm\rangle = e^{im\psi}|jm\rangle.$$

  Then the character is

$$\chi(j, \psi) = \sum_{m=-j}^{j} e^{im\psi} = \frac{\sin(j + 1/2)\psi}{\sin(\psi/2)}.$$

- Choosing coordinates (θ, φ, ψ), the measure on $SO(3)$ has the form

$$d\mu(g) = d\Omega d\psi f(\psi), \quad d\Omega = d\theta d\varphi \sin\theta$$

  by rotational invariance. For small ψ, we expect $f(\psi) \propto \psi^2$ since the group is 'locally Euclidean'.

- The trick is to consider rotations next the identity, expanded as

$$R(\delta, \epsilon, \sigma) = I + \begin{pmatrix} 0 & -\delta & \sigma \\ \delta & 0 & -\epsilon \\ -\sigma & \epsilon & 0 \end{pmatrix} = I + A.$$

  By direct multiplication, we have $R(\delta, \epsilon, \sigma)R(\delta', \epsilon', \sigma') \approx R(\delta + \delta', \epsilon + \epsilon', \sigma + \sigma')$ up to second order terms. Then the measure is $d\delta d\epsilon d\sigma$.

- Next, we transport this result across the group by multiplying it with a finite rotation and seeing how it changes (θ, φ, ψ). We have

$$R(\mathbf{n}, \psi') = R(\mathbf{z}, \psi)R(\delta, \epsilon, \sigma), \quad R(\mathbf{z}, \psi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

  A long calculation shows that

$$n_1 = \frac{\epsilon \sin\psi + \sigma(1 + \cos\psi)}{2\sin\psi}, \quad n_2 = (-\sigma\sin\psi + \epsilon(1 + \cos\psi))(2\sin\psi), \quad n_3 = 1, \quad \psi' = \psi + \delta$$

  to first order.

- Switching back to our original parametrization, we have

$$d\Omega d\psi = d\epsilon d\sigma \left(\frac{\partial(n_1, n_2)}{\partial(\epsilon, \sigma)}\right) d\delta$$

where the factor in parentheses is a Jacobian, and we only need to evaluate a 2×2 Jacobian because δ only affects ψ. We thus find

$$d\mu(g) = d\Omega d\psi \, \sin^2(\psi/2)$$

which indeed is proportional to ψ^2 for small ψ. Thus the integral of a class function $F(g)$ is

$$\int d\mu(g)\, F(g) = \int_0^\pi d\psi \, \sin^2(\psi/2) F(\psi).$$

- Using this result, we can check character orthogonality,

$$\int d\mu(g)\, \chi(k,\psi)^* \chi(j,\psi) = \frac{\pi}{2}\delta_{jk}.$$

We can also run this argument in reverse; character orthogonality can be used to find the measure. We can also decompose tensor products of representations, where the characters multiply, reproducing the usual Clebsch–Gordan decomposition.

Note. The general procedure works as follows. For a matrix Lie group, consider some parametrization $A(\xi)$ where the coordinates are ξ and $A \in G$. Take a basis J_α of $\mathfrak{g}$, and let

$$A^{-1}\frac{\partial A}{\partial \xi_i} = J_\alpha \widetilde{A}(\xi)_i^\alpha.$$

Then the weight function for the measure is

$$\rho(\xi) = \det \widetilde{A}(\xi).$$

To understand this, note that the vectors $\partial A/\partial \xi_i$ form a parallelopiped at $A(\xi)$ and left multiplying by A^{-1} moves it to the identity. Here its volume can be compared to the parallelopiped formed by the Lie algebra elements J_α, with the conversion factor being the Jacobian of the transformation between them. We have essentially done this procedure above.

Finally, we turn to the final example of $SU(2)$.

- As we've seen, $SU(2)$ is geometrically the sphere S^3 by the parametrization

$$U = t + i\mathbf{x}\cdot\boldsymbol{\sigma}, \quad t^2 + x^2 = 1.$$

By symmetry, the measure on $SU(2)$ is just the rotationally symmetric measure on the sphere.

- To find the measure on the sphere, we define $t = \cos\zeta$ and parametrize $\mathbf{x}/|\mathbf{x}|$ by the usual spherical coordinates. For a fixed value of t, the $\mathbf{x}$ coordinates trace out a sphere of radius $\sin\zeta$. Then the measure is $\sin^2\zeta \, d\Omega d\zeta$.

- To compare this with our earlier result, note that $\psi = 2\zeta$ is the angle of rotation, so the integral of a class function of $SU(2)$ would be

$$\int_0^{2\pi} d\psi \, \sin^2(\psi/2) F(\psi).$$

This is identical to our result for $SO(3)$, which makes sense since the two are locally isomorphic, but has double the integration range since $SU(2)$ double covers $SO(3)$.

- We can now apply a 'reality check' to the spin-j representation of $SU(2)$, using

$$\eta^{(j)} = \frac{1}{|G|}\sum_g \chi^{(j)}(g^2) = \frac{\int d\mu(g)\,\chi(j,2\psi)}{\int d\mu(g)}.$$

  Carrying out the integral shows that the integer spin representations are real and the half-integer spin representations are pseudoreal.

- Finally, there is a local isomorphism between $SO(4)$ and $SU(2)\times SU(2)$ given by

$$W \to W' = U^\dagger W V$$

  where W, U, and V are in $SU(2)$, and W and W' are regarded geometrically as points in S^3. Then two elements of $SU(2)$ yield a rotation of the sphere, and it can be shown that any small rotation of the sphere can be written this way.

4 Representations

4.1 ⊛ Representations of $\mathfrak{su}(2)$

Now we find the representations of $\mathfrak{su}(2)$. We begin with some remarks about complexification.

- The standard basis of the Lie algebra is

$$T^a = -\frac{1}{2} i\sigma_a, \quad [T^a, T^b] = f^{ab}{}_c T^c, \quad f^{ab}{}_c = \epsilon_{abc}.$$

  It can also be regarded as the fundamental representation. Note that the representation is complex, because it contains complex matrices like T^z, but the Lie algebra $\mathfrak{su}(2)$ is real. Finally, the representation is real in the physical sense since it is similar to its conjugate.

- In general, we will care about complex representations of real Lie algebras $\mathfrak{g}$ since Hilbert spaces are complex vector spaces. However, it is much simpler to find and classify complex representations of their complexifications $\mathfrak{g}_\mathbb{C}$.

- Every complex representation d of $\mathfrak{g}$ extends to a complex representation $d_\mathbb{C}$ of $\mathfrak{g}_\mathbb{C}$ by

$$d_\mathbb{C}(X + iY) = d(X) + id(Y).$$

  Conversely, given a representation $d_\mathbb{C}$ of $\mathfrak{g}_\mathbb{C}$ we may define a representation d of $\mathfrak{g}$ by restriction.

- Given d and $d_\mathbb{C}$ as defined above, it can be shown that $d_\mathbb{C}$ is an irrep of $\mathfrak{g}_\mathbb{C}$ if and only if d is an irrep of $\mathfrak{g}$. Thus to classify irreps of $\mathfrak{g}$ it is completely equivalent to classify irreps of $\mathfrak{g}_\mathbb{C}$.

- Given a complex Lie algebra, there are multiple ways to restrict to a real Lie algebra, i.e. to take a 'real form'. For example, the complexification of $\mathfrak{su}(n)$ is the set of $n \times n$ traceless complex matrices, which is $\mathfrak{sl}(n, \mathbb{C})$. But this has $\mathfrak{sl}(n, \mathbb{R})$ as a real form, so the representation theory of $\mathfrak{sl}(n, \mathbb{R})$ is the same as that of $\mathfrak{su}(n)$.

- Complexification behaves nicely for Lie algebras, but it is much subtler for Lie groups. A complex Lie group G is a complex-analytic manifold with a holomorphic group operation, and we say $H \subseteq G$ is a real form of G if $\mathfrak{h}$ is a real form of $\mathfrak{g}$. However, $SL(2, \mathbb{R})$ is not a real form of any complex Lie group.

We now turn to finding the irreps of $\mathfrak{su}(2)_\mathbb{C}$. From this point on we'll suppress the $\mathbb{C}$ subscript, always implicitly working with a complexified Lie algebra.

- It is convenient to work in the Cartan–Weyl basis

$$H = 2iT^3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad E_+ = iT^1 + T^2 = \begin{pmatrix} & 1 \\ & \end{pmatrix}, \quad E_- = iT^1 - T^2 = \begin{pmatrix} & \\ 1 & \end{pmatrix}.$$

  The matrices $E_\pm$ are called the raising and lowering operators.

- The commutation relations are

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

  which imply that

$$\mathrm{ad}_H(E_\pm) = \pm 2E_\pm, \quad \mathrm{ad}_H(H) = 0.$$

  Therefore, the generators $\{H, E_\pm\}$ are the eigenvectors of $\mathrm{ad}_H : \mathfrak{su}(2) \to \mathfrak{su}(2)$. The eigenvalues $\{0, \pm 2\}$ are known as the roots of $\mathfrak{su}(2)$.

- Now consider a finite-dimensional representation R with representation space V and suppose that $R(H)$ is diagonalizable, with
$$R(H)v_\lambda = \lambda v_\lambda.$$
This will always be the case in quantum mechanics, where $R(H)$ is Hermitian. The eigenvalues v_λ are the weights of the representation R, and they are raised and lowered by 2 by $E_\pm$.

- There exists a highest weight Λ with
$$R(H)v_\Lambda = \Lambda v_\Lambda, \quad R(E_+)v_\Lambda = 0.$$
If the representation is irreducible, then v_Λ must be unique, and we must be able to reach all of V by applying E_- repeatedly. Therefore the other eigenvectors are
$$v_{\Lambda-2n} = (R(E_-))^n v_n.$$

- To see when this chain must terminate, note that
$$R(E_+)v_{\Lambda-2n} = R(E_+)R(E_-)v_{\Lambda-2n+2} = (R(E_-)R(E_+) + (\Lambda - 2n + 2))v_{\Lambda-2n+2}.$$
Therefore we have
$$R(E_+)v_{\Lambda-2n} = r_n v_{\Lambda-2n+2}, \quad r_n = r_{n-1} + \Lambda - 2n + 2, \quad r_0 = 0$$
which, by induction, gives the solution
$$r_n = (\Lambda + 1 - n)n.$$
Now, let $\Lambda - 2n$ be the lowest weight, so $R(E_-)v_{\Lambda-2N} = 0$. Then we must have $r_{N+1} = 0$, which implies $\Lambda = N$.

- We have thus shown that finite dimensional irreps R_Λ of $\mathfrak{su}(2)$ are labeled by their highest weight $\Lambda \in \mathbb{Z}$, with weights $S_\Lambda = \{-\Lambda, -\Lambda + 2, \ldots, \Lambda - 2, \Lambda\}$ and $\dim R_\Lambda = \Lambda + 1$.

- Decomplexifying back down to $\mathfrak{su}(2)_\mathbb{R}$, R_0 is the trivial representation, R_1 is the fundamental representation, and R_2 is the adjoint representation.

- We can also decomplexify to $\mathfrak{sl}(2)_\mathbb{R}$. This is easier, as H, E_+, and E_- are all traceless and real and hence already form a basis for $\mathfrak{sl}(2)_\mathbb{R}$.

We now relate our results to the familiar theory of angular momentum in quantum mechanics.

- To establish notation, in quantum mechanics we have
$$J^2|jm\rangle = j(j+1)|jm\rangle, \quad J_3|jm\rangle = m|jm\rangle, \quad J^2 = J_1^2 + J_2^2 + J_3^2 = J_3^2 + \frac{1}{2}(J_+J_- + J_-J_+).$$
Then $J_3 = R(H)/2$, since its eigenvalues are half-integer rather than integer, and $J_\pm = R(E_\pm)$. The spin value j is $\Lambda/2$. We call J^2 a 'quadratic Casimir'. It is not part of the Lie algebra, but useful for classifying the irreps.

- Our irreps $R_\Lambda(X)$ exponentiate to representations of $SU(2)$ by

$$D_\Lambda(A) = \exp R_\Lambda(X), \quad A = \exp X.$$

All of the irreps give representations of $SU(2)$, which in turn yield irreps of $SO(3)$ if $D_\Lambda(I) = D_\Lambda(-I)$. Now note that $-I = \exp(i\pi H)$ and

$$D_\Lambda(-I) = \exp(i\pi R_\Lambda(H)).$$

Plugging in the eigenvalues of $R_\Lambda(H)$, the right-hand side is I for $\Lambda \in 2\mathbb{Z}$ and $-I$ otherwise. Then the former give representations of $SO(3)$, and the latter yield projective representations.

Note. The Casimir operator is not an element of the Lie algebra. Formally, define the universal enveloping algebra (UEA) of a Lie algebra $\mathfrak{g}$ to be the algebra generated by elements of $\mathfrak{g}$ subject to the relation $T^aT^b - T^bT^a = f^{ab}{}_c T^c$.

Casimir operators are elements of the UEA which commute with all other elements. Representations may be extended from $\mathfrak{g}$ to the UEA, and by Schur's lemma, Casimir operators are represented in irreps by multiples of the identity. The quadratic Casimir is only one of many examples, and we use it to index irreps.

New representations can be built from our irreps.

- Given a representation R of a real Lie algebra $\mathfrak{g}$, the conjugate representation $\overline{X}$ satisfies

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

In this case, there is only one irrep for each dimension, so each irrep is its own conjugate.

- The contragradient representation, or dual representation R^* is

$$R^*(X) = -R(X)^T.$$

This is simply the infinitesimal version of $D^*(g) = D(g^{-1})^T$ for Lie groups.

- The direct sum of two Lie algebra representations R_1 and R_2 is defined as

$$(R_1 \oplus R_2)(X) = R_1(X) \oplus R_2(X)$$

and their tensor product is defined as

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

The latter is simply the infinitesimal version of the definition for Lie groups, because taking the logarithm of a product yields a sum. That is,

$$\exp(tX_1) \otimes \exp(tX_2) = (1 + tX_1) \otimes (1 + tX_2) + O(t^2) = 1 + t(X_1 \otimes I_2 + I_1 \otimes X_2) + O(t^2).$$

For example, when we consider two particles with spin, the Hilbert space is the tensor product space, but the angular momentum observable is $\mathbf{J}_1 + \mathbf{J}_2$ (with implicit identities), not $\mathbf{J}_1\mathbf{J}_2$.

- It can be shown that the tensor product of finite-dimensional irreps of a simple Lie algebra $\mathfrak{g}$ is always fully reducible. As an example, we explicitly decompose $R_\Lambda \otimes R_{\Lambda'}$. It is useful to again work in the Cartan–Weyl basis, where, for example,

$$(R_\Lambda \otimes R_{\Lambda'})(H) = R_\Lambda(H) \otimes I_2 + I_1 \otimes R_{\Lambda'}(H).$$

  Then we know the weight set is

$$S_{\Lambda,\Lambda'} = \{\lambda + \lambda' | \lambda \in S_\Lambda, \lambda' \in S_{\lambda'}\}$$

  from which we conclude

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

- Note that the Cartan–Weyl basis is not unique; as we've seen, it corresponds to picking out J_z, and we can just as well pick J_x. But the decomposition of a representation into irreps is unique; changing the basis just changes the preferred basis within each irrep.

4.2 The Heisenberg Algebra

As a second example, we consider the infinite-dimensional algebra of creation and annihilation operators for a bosonic field.

- The Lagrangian density is

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

  and the equal-time canonical commutators are

$$[\phi(\mathbf{x}), \dot{\phi}(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y}).$$

- For simplicity, suppose the fields live on the unit circle. Then the modes are the Fourier components,

$$\phi(\mathbf{x}) = a_0 + \sum_{n \neq 0} \frac{1}{n} a_n e^{2\pi i n x}$$

  and applying the equal-time commutators gives

$$[a_n, a_m] = n\delta_{n,-m}.$$

  This is known as the Heisenberg algebra. It splits into three subalgebras: a_0, the creation operators $\{a_n, n \geq 1\}$, and the annihilation operators $\{a_n, n \leq -1\}$ where $a_n^\dagger = a_{-n}$.

- It is convenient to rescale the operators as

$$b_n = \frac{1}{\sqrt{n}} a_{-n}.$$

  Then for every $n > 0$, the operators b_n and $b_n^\dagger$ form a copy of the algebra of creation and annihilation operators for a single harmonic oscillator. Using this observation, we can construct the Fock space starting from a vacuum, assumed to be unique.

This example has a number of common features with the previous example.

- In both cases, we construct a maximal set of commuting operators, $\{H\}$ and $\{a_0\}$, called the Cartan subalgebra. This is an analogue of a complete set of commuting operators in quantum mechanics, but we do not allow operators that are not in the algebra, such as L^2 or $a_n^\dagger a_n$.
- The rest of the algebra splits into two subalgebras, consisting of raising and lowering operators, which are related by conjugation. They are also called creation and annihilation operators, or step operators.
- The representations can be understood by starting from a 'highest weight' or 'vacuum' vector, which is annihilated by the annihilation operators, and applying creation operators. Acting with these operators modify the quantum numbers of the state, i.e. the eigenvalues of the operators in the Cartan subalgebra, also called the weights.
- The dimension of the Cartan subalgebra is called the rank of the Lie algebra; it is the number of quantum numbers we have to work with. So far we have only seen rank one, but in the next section we consider an example with rank two.

4.3 Representations of $SO(3)$

Now we explore some of the representations of $SO(3)$, with some results generalizing to $SO(n)$.

- The group $SO(n)$ has the vector representation $V^i \to R^{ij}V^j$. Tensors transform in tensor powers of the vector representation, and we may break them into irreps.
- Note that antisymmetric rank 2 tensors transform to antisymmetric tensors. The same goes from symmetric tensors, so the rank 2 tensor representation decomposes as $3 + 3 = 3 + 6$. However, the trace is also invariant under rotations (using $R^T = R^{-1}$), so we have the further decomposition $6 = 5 + 1$, where the 5 contains symmetric traceless tensors.
- This decomposition can be expressed in index notation. For the rank 2 tensor T^{ij}, we have

  $$A^{ij} = T^{ij} - T^{ji}, \quad S^{ij} = T^{ij} + T^{ji}, \quad \tilde{S}^{ij} = S^{ij} - \delta^{ij}S^{kk}/N.$$

  Then the A^{ij} are the antisymmetric representation, in the sense that the vector space spanned by the A^{ij} (for all 9 values of (i, j)) is the antisymmetric irrep.
- The exact same reasoning holds for the rank 2 tensor in $SO(n)$.
- The fact that the elements of $SO(n)$ are rotations means that δ is an 'invariant symbol', i.e.

  $$\delta^{ij}R^{ik}R^{jl} = \delta^{kl}.$$

  Similarly, we generally have

  $$\epsilon^{ij\ldots n}R^{ip}R^{jq}\cdots R^{ns} = \epsilon^{pq\ldots s}\det R$$

  so that ϵ is an invariant symbol. We call these objects symbols instead of tensors because they are defined by their components.

- In general, contracting tensors with invariant symbols yields other tensors. For example, the 'dual tensor' of A^{ij} is $B^k = \epsilon^{ijk}A^{ij}$. It is a vector, since

$$B^k \to \epsilon^{ijk}R^{ip}R^{jq}A^{pq} = \epsilon^{ijk'}R^{ip}R^{jq}R^{k'r}R^{kr}A^{pq} = \epsilon^{pqr}R^{kr}A^{pq} = R^{kr}B^r.$$

  Moreover, contraction with invariant tensors projects representations to subrepresentations. For example, for T^{ij}, contraction with δ^{ij} gives the 1 and contraction with ϵ^{ijk} gives the 3. Also note that contraction with ϵ always removes symmetric parts.

- Contraction with ϵ yields strong constraints on the irreps of $SO(3)$ that appear in the tensor product representations, because we can always trade two antisymmetric indices for a single index. We claim the only new irrep for rank n is the traceless totally symmetric rank n tensor.

  We have already shown this for $n \leq 2$. Now consider T^{ijk}. It splits into the representations $T^{[ij]k}$ and $T^{\{ij\}k}$, and the former is just the rank 2 antisymmetric tensor. Now note that

$$3T^{\{ij\}k} = (T^{\{ij\}k} + T^{\{jk\}i} + T^{\{ki\}j}) + (T^{\{ij\}k} - T^{\{jk\}i}) + (T^{\{ij\}k} - T^{\{ki\}j})$$

  where the first term is totally symmetric, and the other terms are antisymmetric in ki and kj respectively. Thus the only new representation is totally symmetric. Finally, given a totally symmetric tensor S^{ijk} we can always remove all of its traces,

$$\tilde{S}^{ijk} = S^{ijk} - \frac{1}{N+2}(\delta^{ij}S^{hhk} + \delta^{ik}S^{hhj} + \delta^{jk}S^{hhi})$$

  leaving only the new irrep, as claimed. A similar proof holds for general n.

- We now count the dimension of these representations. A totally symmetric rank n tensor has

$$\binom{n+2}{2}$$

  degrees of freedom. By symmetry, there is only one independent trace to remove, which is a totally symmetric rank $n-2$ tensor, so the dimension is

$$\binom{n+2}{2} - \binom{n}{2} = 2n+1.$$

  These are, in fact, all of the irreps of $SO(3)$.

- A similar argument holds for $SO(2)$. In this case, the dimensions are all $(n+1)-(n-1) = 2$. This is not in contradiction with the fact that the irreps of $U(1) \cong SO(2)$ are all one-dimensional, because we are considering real representations.

- For $SO(4)$, contraction with ϵ does not decrease the rank, while for $SO(5)$ and higher contraction increases the rank. Hence the tricks above fail, and the irreps can have a more complex symmetry structure. They can be enumerated using Young tableaux, as we'll see below.

We now make some extra remarks.

- Consider the representation of totally antisymmetric rank n tensors of $SO(2n)$. This representation is mapped to itself by duality; hence it breaks into two irreps, containing 'self-dual' and 'anti self-dual' tensors.

- Consider an irrep of a group G. If we restrict to a subgroup $H \subset G$, then the irrep will generally not be an irrep in H. For example, the vector representation of $O(4)$ splits into $4 = 3 + 1$ when restricting to $SO(3)$, corresponding to the splitting of space and time. The antisymmetric rank 2 tensor splits into $6 = 3 + 3$, corresponding to the splitting of $F^{\mu\nu}$ into $\mathbf{E}$ and $\mathbf{B}$. The symmetric rank 2 tensor splits into $9 = 5 + 3 + 1$.

- The adjoint representation of a Lie group G has representation space $\mathfrak{g}$, where $g \in G$ acts by conjugation. For $SO(n)$, the adjoint representation is simply the antisymmetric tensor representation, as
$$A^{ij} \to R^{ip} R^{jq} A^{pq} = (RAR^{-1})^{ij}.$$
For $SO(3)$ only, the adjoint representation coincides with the vector representation. That is, angular momentum (as a physical quantity) is a vector only in three dimensions.

- Since we are working with $SO(n)$ rather than $O(n)$, we have neglected the difference between vectors and axial vectors, and so on. Since the metric is Euclidean, we haven't bother with raising or lowering indices.

Example. Decomposing $P^{ijk} = S^{ij}T^k$ where S is symmetric and traceless, in $SO(3)$. First, construct the symmetric tensor
$$U^{ijk} = S^{ij}T^k + S^{jk}T^i + S^{ki}T^j.$$
This is a totally symmetric rank 3 tensor, so it decomposes into a traceless part and a trace; the trace yields the vector representation. The other degrees of freedom are in the antisymmetric part
$$V^{i\ell} = S^{ij}T^k \epsilon^{jk\ell}.$$
The tensor V is neither symmetric nor antisymmetric, but it is traceless. Then it decomposes into a symmetric traceless and antisymmetric part; the latter is just the vector we have already found. Hence we have shown $5 \times 3 = 7 + 5 + 3$.

More generally, suppose we multiply two symmetric traceless tensors, with j and j' indices. Then we can construct a symmetric traceless tensor with $j + j'$ indices, as shown above. The remainder of the degrees of freedom are given by contracting with ϵ^{ijk}, which leaves $j + j' - 1$ indices. We can then take out the symmetric traceless part again and repeat the procedure, so
$$j \times j' = (j + j') + (j + j' - 1) + \ldots + |j - j'|.$$
This is simply the usual Clebsch–Gordan decomposition. Note that we have switched notation so that j is really an irrep of dimension $2j + 1$. However, we'll usually stick to the convention where the name indicates the dimension.

4.4 Representations of $SU(3)$

Next, we extend our tensor methods to $SU(3)$. We warm up with $SU(2)$.

- The fundamental representation transforms as
$$\psi^i \to \psi'^i = U^{ij}\psi^j.$$
By analogy with our earlier work, we consider tensors $\varphi^{i_1 \cdots i_n}$.

- For example, the totally symmetric tensor φ^{ijk} of $SU(3)$ has 10 degrees of freedom (corresponding to the baryon decuplet). At this point one might take out the trace, but the trace does not transform correctly,
$$\delta^{ij}\varphi^{ijk} \to (\delta^{ij}U^{if}U^{jg})U^{kh}\varphi^{fgh}.$$
The quantity in parentheses is U^TU, which was the identity for $SO(n)$, but nothing here; we need an extra complex conjugation.

- To address this, we introduce the conjugate/antifundamental representation $\psi_i = \psi^{i*}$. Introducing upper and lower indices, we have
$$\psi^i \to U^i{}_j\psi^j, \quad \psi_i \to \psi_j(U^\dagger)^j{}_i.$$
We are allowed to contract upper and lower indices together; for instance, $\psi^i\psi_i$ is a scalar.

- Formally, for $SO(n)$, the fundamental representation tensored with itself contains the trivial representation, but in $SU(n)$, we must tensor the fundamental and antifundamental representation. Physically, a color singlet meson is made of a quark and an antiquark.

- We can think of a conjugate representation as living in the dual space of the original representation; our transformation rule for ψ_i is simply the transformation rule for a bra.

- As a result, we can consider tensors with arbitrary mixed rank (r, s). When both r and s are nonzero, we can subtract out traces with δ. Finally, since the elements of $SU(n)$ have determinant one,
$$\epsilon_{i_1\dots i_N}U_1^{i_1}\cdots U_N^{i_N} = \epsilon^{i_1\dots i_N}U^1_{i_1}\cdots U^N_{i_N} = 1$$
yielding two more invariant symbols, which allow raising and lowering of indices.

- As an example, the totally symmetric tensor φ^{ijk} considered earlier is irreducible. However, a tensor T^{ij}_k breaks into four irreps. As usual, the symmetric and antisymmetric parts S^{ij}_k and A^{ij}_k (in the upper two indices) form subrepresentations. However, both of these have traces, which form copies of the fundamental representation. The dimensions are
$$N^3 = \left(\frac{1}{2}N^2(N+1) - N\right) + \left(\frac{1}{2}N^2(N-1) - N\right) + N + N.$$
The naming conventions are somewhat more complex. We will usually stick with naming by dimension, adding a star for representations with mostly lower indices.

- When we work with the Lie algebra, we will have several types of indices. The generators are named T^a, where a has $\dim G$ possible values. If we work in an irrep of dimension d, then each of the generators is a $d \times d$ matrix, so we let p and q index over the irrep, e.g.
$$\delta\varphi^p = i\theta^a(T^a)^p{}_q\varphi^q.$$
The φ^p can also be written as traceless tensors with definite symmetry properties, as we did above; in that case the indices on the ranges i and j range from 1 to N. Note that there is no meaning to the upstairs or downstairs placement of the other types of indices.

- The adjoint representation turns out to be the antisymmetric $(1, 1)$ tensor representation φ^i_j, which indeed has dimension $N^2 - 1$.

Example. All irreps of $SU(2)$. We can use ϵ^{ij} to raise any downstairs indices, so all the irreps have solely upstairs indices. Furthermore, the irreps must be totally symmetric, because any antisymmetric part can be projected out by contraction with ϵ_{ij}, reducing it to a lower rank tensor. Hence the irreps are totally symmetric tensors with n upper indices, which have dimension $n+1$.

We can go further and use this to find the representation matrices $D^s(R)$. The fundamental representation is a two-element spinor, which is multiplied by $D^{1/2}$. Higher irreps are symmetrized tensor powers of that spinor, and a rank n tensor transforms with a factor of $D^{1/2}$ on each index. Hence it is not surprising that the d-functions $d^j(\beta)_{mm'}$ are polynomials in $\cos(\beta/2)$ and $\sin(\beta/2)$.

Note that the antifundamental representation does not appear here; it is similar to the fundamental representation, which turns out to be pseudoreal. Explicitly, we have

$$\sigma_2 \sigma_a^* \sigma_2 = -\sigma_a$$

so the change of basis matrix is σ_2 itself. Since σ_2 is antisymmetric, the representation is pseudoreal.

Note. Consider a tensor product $j \otimes j'$. The resulting $j + j'$ irrep is symmetric, because its highest J_z state is symmetric and symmetry is preserved by lowering. Then the $j + j' - 1$ irrep is antisymmetric, since its highest J_z is determined by orthogonality. This pattern continues, with symmetry alternating between irreps.

Note. Naively, one might say that $U(N) = SU(N) \times U(1)$. However, this is incorrect because in general, N elements of the form $e^{i\theta} I$ are in $SU(N)$. The actual relationship is

$$U(N) = (SU(N)/\mathbb{Z}_N) \times U(1).$$

When one says the Standard Model gauge group is $SU(3) \times SU(2) \times U(1)$, one is really talking about the Lie algebra; there are several possibilities for the Lie group and the correct one is unknown.

Next, we turn to the representations of $SU(3)$.

- In this case, we have the invariant symbols ϵ^{ijk} and ϵ_{ijk}, and the fundamental and antifundamental representations are not equivalent.
- We claim that all irreps are traceless (m, n) tensors that are totally symmetric in both the upper and lower indices. The claim is obvious for rank $r = m + n = 1$. For $r = 2$, the antisymmetric tensors φ^{ij} and φ_{ij} are equivalent to vectors, so they give nothing new, and we may subtract out the trace of φ^i_j.
- For the case $r = 3$, consider the tensor φ^{ijk}. The antisymmetric part $\varphi^{[ij]k}$ can be reduced to rank $r = 2$. The symmetric part $\varphi^{(ij)k}$ can be split into a totally symmetric part and antisymmetric tensors in ki and kj, as shown for $SO(3)$, so the new irrep is totally symmetric.
- Now consider the (m, n) irrep. Without accounting for the traceless condition, the dimension is

$$\binom{n+2}{2}\binom{m+2}{2}.$$

  The trace of a totally symmetric (m, n) tensor is a totally symmetric $(m-1, n-1)$ tensor, so

$$\binom{n+2}{2}\binom{m+2}{2} - \binom{n+1}{2}\binom{m+1}{2} = \frac{1}{2}(m+1)(n+1)(m+n+2).$$

- A few of the low-dimensional irreps are

$$(1,0) = 3, \quad (0,1) = 3^*, \quad (1,1) = 8, \quad (2,0) = 6, \quad (3,0) = 10, \quad (2,1) = 15, \quad (2,2) = 27.$$

  Many of these numbers play a role in the Eightfold Way.

Next, we derive some multiplication rules for irreps of $SU(3)$.

- Note that $(1,0) \times (0,1)$ is a general $(1,1)$ tensor. Subtracting its trace gives

$$(1,0) \times (0,1) = (1,1) + (0,0), \quad 3 \times 3^* = 8 + 1.$$

- Next, $(1,0) \times (1,0)$ is a general $(2,0)$ tensor, so

$$(1,0) \times (1,0) = (2,0) + (0,1), \quad 3 \times 3 = 6 + 3^*$$

  where the two terms on the right are simply the symmetric and antisymmetric parts.

- Next, consider $(1,0) \times (2,0)$, i.e. the tensor $\psi^i \varphi^{jk}$. Consider the antisymmetric and symmetric parts in ij. The antisymmetric part gives a $(1,1)$ tensor which is automatically traceless. The symmetric part is symmetric in all three indices, so

$$(1,0) \times (2,0) = (3,0) + (1,1), \quad 3 \times 6 = 10 + 8.$$

  Therefore $3 \times 3 \times 3 = 10 + 8 + 8 + 1$, which we will see applied in the Eightfold Way.

- Finally, consider $(1,1) \times (1,1)$. The tensor $\psi^i_j \chi^k_\ell$ has two distinct traces,

$$\psi^i_j \chi^j_\ell, \quad \psi^i_j \chi^k_i.$$

  These yield two copies of $(1,1)$ plus a copy of $(0,0)$, since their traces are equal. Now consider the traceless part $T^{ik}_{j\ell}$. It can be shown that $T^{[ik]}_{j\ell}$ is automatically symmetric in j and ℓ, providing a $(0,3)$. Similarly antisymmetrizing $j\ell$ gives a $(3,0)$ and the symmetric remainder is $(2,2)$, for

$$(1,1) \times (1,1) = (2,2) + (3,0) + (0,3) + (1,1) + (1,1) + (0,0).$$

  This computation contains all the ideas necessary for the general case.

Example. Consider the three-dimensional harmonic oscillator. Such a system has an $SU(3)$ symmetry; the generators take the form

$$Q_\alpha = a^\dagger_k [T_\alpha]_{k\ell} a_\ell, \quad T_\alpha = \frac{\lambda_\alpha}{2}.$$

The creation operators transform in the fundamental representation while the lowering operators transform in the antifundamental. Note that the $SU(2)$ rotational symmetry is a subset; it is generated by

$$L_3 = 2Q_2, \quad L_1 = 2Q_7, \quad L_2 = -2Q_5.$$

The $SU(3)$ symmetry explains additional degeneracy that the $SU(2)$ symmetry does not. For example, all six $n = 2$ states are degenerate, even though they split up as $5 + 1$ under $SU(2)$.

Example. Consider two three-dimensional harmonic oscillators. If they are uncoupled, we have an $SU(3) \times SU(3)$ symmetry. Now let

$$H^{\text{int}} = \lambda a_k^\dagger b_k^\dagger a_\ell b_\ell.$$

This interaction is designed so it commutes with

$$Q_\alpha = a_k^\dagger [T_\alpha]_{k\ell} a_\ell - b_k^\dagger [T_\alpha^*]_{k\ell} b_\ell$$

so an $SU(3)$ symmetry remains. Here, the $a^\dagger$'s transform in the 3 and the $b^\dagger$'s transform in the $\overline{3}$. Then we have $[Q_\alpha, a_\ell b_\ell] = 0$ because removing a red quark and red antiquark conserves color. The Hamiltonian commutes with the number operators, so the energy eigenstates have definite number. In particular, the highest weight state of an (n, m) irrep is

$$(a_1^\dagger)^n (b_3^\dagger)^m (\mathbf{a}^\dagger \cdot \mathbf{b}^\dagger)^k |0\rangle$$

for any k, and lowering generates degenerate states.

4.5 The Symplectic Groups

The symplectic groups are the least familiar of the matrix Lie groups.

- The real symplectic group $Sp(2n, \mathbb{R})$ contains real $2n \times 2n$ matrices which satisfy

  $$R^T J R = J, \quad J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

  That is, they are matrices that preserve the quadratic form J, like how the elements of $O(p, q)$ preserve $\text{diag}(I_p, -I_q)$.

- To count parameters, note that the left-hand side is automatically antisymmetric, so the number of constraints is $(2n)(2n-1)/2$. Then

  $$\dim Sp(2n, \mathbb{R}) = 4n^2 - n(2n-1) = n(2n+1).$$

- Directly from the definition, we see $\det R = \pm 1$. However, it turns out that we automatically have $\det R = 1$. For example, for $n = 1$, the symplectic condition is simply $\det R = 1$, so $Sp(2, \mathbb{R}) \cong SL(2, \mathbb{R})$.

- The symplectic groups appear in Hamiltonian mechanics. Combining the canonical positions and momenta into a single vector Z, Hamilton's equations are

  $$\frac{dZ_a}{dt} = J_{ab} \frac{\partial H}{\partial Z_b}$$

  and the form of the equations are preserved under any symplectic transformation $Z \to RZ$.

- Similarly, one may define the complex symplectic groups $Sp(2n, \mathbb{C})$, where the definition still contains a transpose rather than a conjugate transpose, $R^T U R = U$.

- Finally, we define the compact symplectic group

$$USp(2n) = Sp(2n, \mathbb{C}) \cap U(2n).$$

This group is sometimes also called $Sp(2n)$ or $Sp(n)$. To count parameters, we write $U = I + iH$ where H is Hermitian, which gives $H^T J + JH = 0$. Then the general form of H is

$$H = \begin{pmatrix} P & W^* \\ W & -P^T \end{pmatrix}, \quad P^\dagger = P, \quad W^T = W, \quad \dim USp(2n) = n(2n+1).$$

Note that H is generally traceless, so $USp(2n) \subset SU(2n)$. In particular, $USp(2) \cong SU(2)$.

- The compact symplectic groups will appear in the Cartan classification; they also are involved when writing down Lagrangians for Majorana spinors.

It is also useful to write an explicit basis for the algebra $\mathfrak{usp}(2n)$.

- Using tensor product notation, e.g. $J = I \otimes i\sigma_2$, we claim that a basis for $\mathfrak{usp}(2n)$ is given by

$$iA \otimes I, \quad S_1 \otimes \sigma_1, \quad S_2 \otimes \sigma_2, \quad S_3 \otimes \sigma_3$$

where A is real antisymmetric and the S_i are real symmetric. This can be seen by directly comparing with our general form above.

- Using this basis, it is easy to check that the algebra closes. For example,

$$[iA \otimes I, S_a \otimes \sigma_a] = i[A, S_a] \otimes \sigma_a = iS'_a \otimes \sigma_a$$

since the commutator of an antisymmetric and symmetric matrix is symmetric. We also have

$$[S_1 \otimes \sigma_1, S'_1 \otimes \sigma_1] = [S_1, S'_1] \otimes I = i(-iA) \otimes I$$

as well as

$$[S_1 \otimes \sigma_1, S_2 \otimes \sigma_2] = iS_1S_2 \otimes \sigma_3 + iS_2S_1 \otimes \sigma_3 = iS_3 \otimes \sigma_3.$$

Similarly, it is straightforward to check that the symplectic condition $JHJ = H^T$ is satisfied.

- As an example, note that linear combinations of the generators $I \otimes iA$ and $\sigma_3 \otimes S$ yield $H = \mathrm{diag}(P, -P^T)$ for Hermitian p. This is the $U(n)$ subgroup of $USp(2n)$.

4.6 Projective Representations

We now consider how projective representations arise in quantum mechanics.

- Naively, symmetries in quantum mechanics are represented by unitary operators $U(T_i)$. However, since symmetries need only preserve probabilities, the most general possibility allows extra phases. The most general possibility, by Wigner's theorem, is

$$U(T_2)U(T_1)|n\rangle = e^{i\phi_n(T_2,T_1)}U(T_2T_1)|n\rangle$$

where the extra phase $\phi_n(T_2, T_1)$ is allowed to depend on the state $|n\rangle$, and the $U(T_i)$ may be unitary or antiunitary. From here on we only consider the unitary case.

- Now consider a superposition of two states. Then we have

$$e^{i\phi_{nm}}U(T_2T_1)(|n\rangle + |m\rangle) = e^{i\phi_n}U(T_2T_1)|n\rangle + e^{i\phi_m}U(T_2T_1)|m\rangle$$

  where we used linearity. Multiplying by $U(T_2T_1)^{-1}$, we conclude $\phi_{nm} = \phi_n = \phi_m$. Hence

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2T_1).$$

  Hence the $U(T_i)$ form a unitary projective representation.

- If ϕ_n and ϕ_m were not equal, we could not even have a projective representation, which would be an annoying mathematical obstacle. The exception is if the states $|n\rangle$ and $|m\rangle$ cannot be superposed. For example, it is believed that it is impossible to superpose boson and fermion states; it is forbidden by a "superselection rule".

- This need not be viewed as a fundamental restriction, because as we will see, a symmetry group can always be formally enlarged to remove superselection rules. For example, in the Galilean group, we find that the projective phases picked up by multiplying boosts and translations is proportional to the mass M, giving a superselection rule for mass. But if we formally add M to the Galilean algebra, then there is no problem; all the phases ϕ_n become identically zero. Hence the set of superselection rules is a matter of convention.

- Now we consider projective representations in detail. Associativity requires

$$\phi(T_2, T_1) + \phi(T_3, T_2T_1) = \phi(T_3, T_2) + \phi(T_3T_2, T_1).$$

  If the phases have the form

$$\phi(T, \overline{T}) = \alpha(T\overline{T}) - \alpha(T) - \alpha(\overline{T})$$

  then these conditions are clearly satisfied; in this case all phases can be removed by taking $\overline{U}(T) = U(T)e^{i\alpha(T)}$. We call the set of equivalence classes of $\phi(T_2, T_1)$ up to such phase definitions two-cocycles.

- Hence we would like to ask when a representation is intrinsically projective, i.e. when it lies in a nontrivial two-cocycle. Obstructions can come either from the Lie algebra itself, or from the global topology of the Lie group.

We now focus on the Lie algebra.

- Redefining $U(1)$ to be the identity, we must have

$$\phi(T, 1) = \phi(1, \overline{T}) = 0.$$

  Hence when both T and $\overline{T}$ are near the identity, $\phi(T, \overline{T})$ is small. At lowest order,

$$\phi(T(\theta), T(\overline{\theta})) = f_{ab}\theta^a\overline{\theta}^b + \dots$$

  where the θ^a parametrize the Lie algebra.

- This leads to the modified commutation relations

$$[t_b, t_c] = i f^a{}_{bc} t_a + i C_{bc} 1, \quad C_{bc} = f_{cb} - f_{bc}.$$

  The extra terms proportional to the identity are called central charges, and are the signature of a projective representation at the level of the Lie algebra.

- The Jacobi identity still holds, and leads to the constraint

$$f^a{}_{bc} C_{ad} + f^a{}_{cd} C_{ab} + f^a{}_{db} C_{ac} = 0.$$

  This is automatically satisfied if $C_{ab} = f^c{}_{ab} \phi_c$ for real constants ϕ_c. If this is true, we may eliminate the central charges by switching to the generators $\tilde{t}_a = t_a + \phi_a$.

- It turns out that for finite-dimensional representations, it is always possible to remove the central charges. More generally, Bargmann's theorem states that the central charges can always be eliminated if the Lie algebra cohomology group $H^2(\mathfrak{g}, \mathbb{R})$ is trivial. This group is trivial for all semisimple Lie algebras, which include the Lorentz algebra but not the Galilean algebra. It is also trivial for the Poincare algebra; to show this explicitly we can eliminate the central charges by hand.

- The existence of central charges is tied to the existence of central extensions. A central extension of $\mathfrak{g}$ is an exact sequence

$$0 \to \mathfrak{a} \to \mathfrak{e} \to \mathfrak{g} \to 0$$

  so that $\mathfrak{a}$ is in the center of $\mathfrak{e}$. It is nontrivial if the sequence does not split.

Next, we turn to the topology.

- It can be shown that a representation of $\mathfrak{g}$ with no central charges must lift to a proper representation of G if G is simply connected. More generally, any projective representation of G corresponds to a proper representation of the universal cover of G.

- Intuitively, this can be done by simply exponentiating our representation of $\mathfrak{g}$ naively. We know the phases should work out near the origin, and the exponential map defines a "standard path" from the origin to other points in G. We can only run into an inconsistency if there are paths that cannot be deformed into each other, which correspond to nontrivial loops. We can get all projective representations by combining a proper representation of G and a one-dimensional proper representation of $\pi_1(G)$, giving the phases for each nontrivial loop.

- Hence, we can always avoid projective representations by passing to a central extension of the algebra, exponentiating, then passing to the universal cover of the group. (This is really the same thing; a universal cover is a central group extension by $\mathbb{Z}_n$.) This also removes all superselection rules. The of symmetry group and superselection rules is really just a matter of convenience, which depends on which superpositions we know how to prepare, and which groups we think are easy to work with.

- Topologically, the Lorentz group is $\mathbb{R}^3 \times S^3/\mathbb{Z}_2$, and its universal/double cover is $SL(2, \mathbb{C}) = \mathbb{R}^3 \times S^3$. Hence the Lorentz group has projective representations. Since we are working with a double cover, the phases can only be signs, $U(\Lambda)U(\overline{\Lambda}) = \pm U(\Lambda\overline{\Lambda})$. These simply correspond to integer and half-integer spin.

5 Physical Applications

5.1 Isospin

We give a historical account of the development of isospin and some of its successes.

- In 1932, the neutron was discovered and found to have a mass very similar to that of the proton. It was immediately proposed that the neutron and proton form a doublet under an $SU(2)$ symmetry of the strong interaction, called isospin, with the symmetry broken by electromagnetic effects.
- We will refer to both the quantum numbers s and m as the 'spin'. Similarly we refer to both I and I_3 as the isospin.
- In 1935, Yukawa proposed that the nuclear force could be mediated by the exchange of mesons. In 1947, the charged pions $\pi^\pm$ were discovered which participated in the processes

  $$p \to n + \pi^+, \quad n \to p + \pi^-.$$

  Applying isospin addition, the isospin of the charged pions can be either 0 or 1. Since the charged pions have nearly the same mass, we suppose they are part of an isospin triplet, leading to the prediction of a third pion π^0 which was found in 1950.
- The Gell-Mann–Nishijima formula is the empirical result

  $$Q = I_3 + \frac{Y}{2}$$

  where Y is the hypercharge, an operator lying outside of $SU(2)$ conserved by strong interactions. For nucleons, $Y = 1$, while for pions $Y = 0$.
- As an example, the deuteron is a bound state of the proton and neutron, and can be produced in the processes

  $$p + p \to d + \pi^+, \quad p + n \to d + \pi^0.$$

  Then the isospin of the deuteron is either 0 or 1. In the case of isospin 1, applying the isospin raising and lowering operators implies the existence of p-p and n-n bound states, which are not observed. Hence the deuteron has isospin 0. Since the deuteron has zero orbital angular momentum, it must thus have spin 1 to make the full wavefunction antisymmetric.
- Isospin can also make quantitative predictions. The amplitudes for these two processes are proportional to the Clebsch–Gordan coefficients

  $$\left\langle 1,1 \middle| \frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2} \right\rangle = 1, \quad \left\langle 1,0 \middle| \frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2} \right\rangle = \frac{1}{\sqrt{2}}$$

  respectively. The cross section for the first process is twice as large, as confirmed in 1953.
- As another example, it was found that the cross section for

  $$\pi^+ + p \to \pi^+ + p$$

  had a sharp peak, which could be interpreted as the formation of a short-lived particle, or 'resonance', called the N^*. By adding I values, the resonance could have $I = 1/2$ or $I = 3/2$,

but since it has $I_3 = 3/2$ it must have $I = 3/2$. Then isospin predicts three additional resonances, which were shortly found. Again, the Clebsch–Gordan coefficients provide simple relationships between the cross sections.

- Finally, we turn to the couplings between the pions and nucleons. The couplings allowed by isospin are the two shown above, plus $p \to \pi^0 + p$ and $n \to \pi^0 + n$. Using the same Clebsch–Gordan symbols as above we find

$$g_{p,\pi^0 p} = g, \quad g_{p,\pi^+ n} = -\sqrt{2}g, \quad g_{n,\pi^- p} = \sqrt{2}g, \quad g_{n,\pi^0 n} = -g.$$

  Therefore the cross sections for pp and nn scattering are equal, $\sigma \propto g^2$. For pn scattering, we can have either an intermediate π^0 or intermediate charged pion (by the u-channel), giving $\sigma \propto (2-1)^2 g^2 = g^2$. Thus all nucleons are interchangeable under the strong force as expected.

- Note that there aren't two separate contributions for an intermediate π^+ or intermediate π^-. These two possibilities form a single Feynman diagram, where the virtual particle can have either positive or negative energy.

- Finally, the couplings above can also be found by building an isospin-scalar Lagrangian. The proton and neutron are combined into fields N^i and N_i, while the pions form a traceless tensor ϕ^i_j. Hence the only possible term is $\Delta\mathcal{L} = N_i \phi^i_j N^j$, which gives the couplings above.

We can also formalize isospin using creation and annihilation operators.

- For the proton and neutron, we define

$$|p, \alpha\rangle = a^\dagger_{1/2,\alpha}|0\rangle, \quad |n, \alpha\rangle = a^\dagger_{-1/2,\alpha}|0\rangle$$

  where the $\pm 1/2$ stands for the isospin and α stands for everything else about the state. Since nucleons are fermions, these creation operators anticommute.

- We can also write the isospin generators in terms of these operators,

$$T_a = \frac{1}{2} a^\dagger_{m',\alpha} (\sigma_a)_{m'm} a_{m,\alpha}, \quad T_a|m,\alpha\rangle = (J^{1/2}_a)_{m'm}|m',\alpha\rangle = \frac{1}{2}(\sigma_a)_{m'm}|m',\alpha\rangle$$

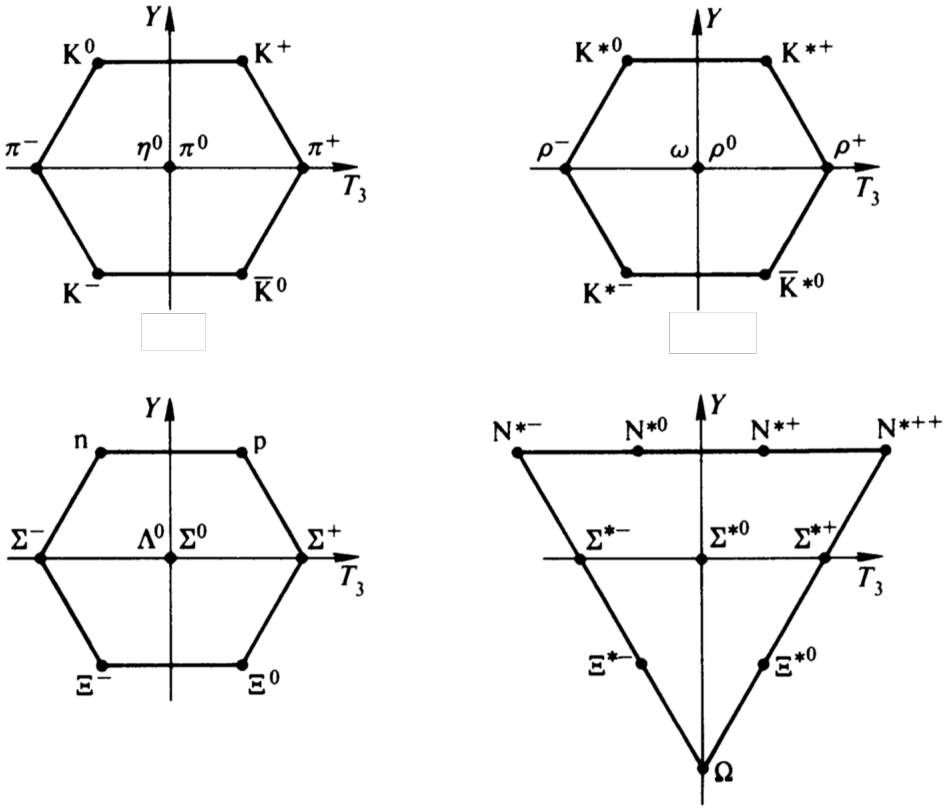
  where there are implicit sums over m, m', and α.

- Note that $|0\rangle$ is an isospin singlet while the $|m,\alpha\rangle$ states form an isospin doublet. Then the operators $a^\dagger_{\pm 1/2,\alpha}$ are an isospin 1/2 tensor operator, e.g.

$$[T_a, a^\dagger_{m,\alpha}] = (J^{1/2}_a)_{m'm} a^\dagger_{m',\alpha}.$$

  For more species of particles, we simply augment $a^\dagger$ with an extra index and sum over it. Note that whether the creation operators commute or anticommute depends on the particle, but the tensor operator relation above always uses commutators.

5.2 ∗ The Eightfold Way



Next, we introduce the Eightfold Way historically.

- In the early 1950s, a number of new particles were discovered, in particular four pseudoscalar K mesons. Since the three pions were known, there was a total of seven pseudoscalar mesons, which were proposed to form an irrep (since all particles in an irrep automatically have the same parity). Gell-Mann proposed that the irrep was the $(1,1)$ of $SU(3)$, thus predicting an eighth particle, the η^0. (There is also another octet containing vector mesons.)

- Part of the confusion was that the $SU(3)$ symmetry was much more badly broken than isospin; the masses of the K mesons were over three times the masses of the pions, though they were still significantly lighter than any baryons.

- In addition, experimentalists found an isospin triplet of Σ baryons, an isospin doublet of Ξ baryons, and an isospin singlet Δ baryon, which Gell-Mann proposed fit with the neutron and proton in another $(1,1)$ of $SU(3)$.

- Finally, a number of short-lived hadron resonances were known, including the four N^* particles above. They were proposed to form the $(3,0)$ of $SU(3)$, predicting a tenth resonance, the Ω.

- The triality of the (m,n) irrep of $SU(3)$ is defined as $(m-n) (\mathrm{mod}\, 3)$. Note that all of the irreps listed above have zero triality. Physically, all of these representations are built from the 3 (containing quarks) and the $\overline{3}$ (containing antiquarks). Thus triality zero ensures that observed mesons and baryons have integer charge and baryon number. Confinement in QCD ensures that all observable free particles have zero triality.

- By restricting to $SU(2) \subset SU(3)$, an $SU(3)$ irrep breaks into isospin irreps, e.g. $3 \to 2 + 1$, so we can recover isospin from the Eightfold Way. The isospin and hypercharge obey

$$I_3 = \frac{1}{2}\operatorname{diag}(1, -1, 0), \quad Y = \frac{1}{3}\operatorname{diag}(1, 1, -2)$$

  so the up and down quark have hypercharge $1/3$, and the strange quark has hypercharge $-2/3$. Then hypercharge generates a $U(1)$ subgroup and we can think of restricting $SU(3)$ to $SU(2) \times U(1)$, so that every isospin irrep is labeled by a hypercharge.

- We write I_{3Y} to indicate an isospin I irrep with hypercharge Y. Then we have

$$3 \to 2_1 + 1_{-2}, \quad 3^* \to 2_{-1} + 1_2.$$

  For example, decomposing both sides of $3 \times 3^* = 8 + 1$ yields

$$8 \to 3_0 + 1_0 + 2_3 + 2_{-3}$$

  which reproduces the structure of the meson and baryon octets. This is experimentally useful, since isospin is a much more accurate symmetry than the $SU(3)$.

- Finally, assigning the quarks charges of $(2/3, -1/3, 2/3)$ and isospin $(1/2, -1/2, 0)$ recovers the Gell-Mann–Nishijima formula. Physically, a formula like this had to work because electromagnetic interactions preserve isospin and hypercharge, so the electromagnetic field has to couple to some combination of I_3 and Y, and we call this combination the charge.

5.3 Roots and Weights for $\mathfrak{su}(3)$

In this section, we investigate the structure of the Lie algebra $\mathfrak{su}(3)$.

- The Gell-Mann matrices are

$$\lambda_i = \begin{pmatrix} \sigma_i & \\ & 0 \end{pmatrix}, \quad \lambda_4 = \begin{pmatrix} & & 1 \\ & & \\ 1 & & \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} & & -i \\ & & \\ i & & \end{pmatrix}$$

  and

$$\lambda_6 = \begin{pmatrix} & & \\ & & 1 \\ & 1 & \end{pmatrix}, \quad \lambda_7 = \begin{pmatrix} & & \\ & & -i \\ & i & \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix}.$$

  The pairs λ_4 and λ_5 are analogous to σ_x and σ_y but act on the first and third column, while the pair λ_6 and λ_7 does the same on the second and third column. The matrices are normalized so that $\operatorname{tr} \lambda_a \lambda_b = 2\delta_{ab}$.

- For concreteness, we work in the fundamental representation of $\mathfrak{su}(3)$ using the universal physics normalization convention for the fundamental,

$$T_a = \lambda_a/2, \quad \operatorname{tr} T_a T_b = \frac{1}{2}\delta_{ab}.$$

  All our results will hold in general, since they depend only on the structure constants of $\mathfrak{su}(3)$.

- The Lie algebra contains three copies of the $\mathfrak{su}(2)$ algebra, defining

$$I_\pm = T_1 \pm iT_2, \quad U_\pm = T_6 \pm iT_7, \quad V_\pm = T_4 \pm iT_5.$$

  These are the raising and lowering operators for

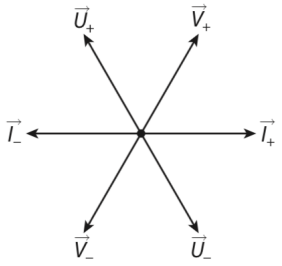
$$[I_+, I_-] = 2I_3, \quad [U_+, U_-] = 2U_3, \quad [V_+, V_-] = 2V_3.$$

  There are also nontrivial commutators between distinct raising and lowering operators.

- We identify (the z-component of) isospin with T_3 and hypercharge Y with $(2/\sqrt{3})T_8$. Then the step operators change the isospin and hypercharge by

$$I_\pm : \pm(1, 0), \quad U_\pm : \pm(-1/2, 1), \quad V_\pm : \pm(1/2, 1).$$

  These vectors are known as the roots of $\mathfrak{su}(3)$.



- The raising and lowering operators have a simple interpretation in the quark model. Here, I_+ replaces a down quark with an up quark, U_+ replaces a strange quark with a down quark, and V_+ replaces a strange quark with an up quark.

- Reverting to the 'mathematical' normalization (i.e. T_8 instead of Y), the root vectors are

$$\pm(1, 0), \quad \pm(-1/2, \sqrt{3}/2), \quad \pm(1/2, \sqrt{3}/2).$$

  Then all of the roots have equal length, and they form a regular hexagon.

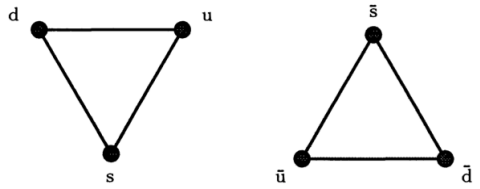
The root diagram gives us some insight into the structure of the Lie algebra and its representations.

- Given a representation, we may simultaneously diagonalize T_3 and T_8. Then the weights are the set of vectors of eigenvalues. Since the raising and lowering operators modify weights by roots, the weights must form part of a hexagonal lattice.

- The roots give us some commutators for free. For example, $[U_+, I_-]$ must be proportional to a step operator with root $U_+ + I_-$, but there is no such root, so the commutator is zero. Similarly, $[U_+, I_+]$ must be proportional to V_+.

- In general, it turns out all irreps of $\mathfrak{su}(3)$ have weight sets that are hexagons built around an equilateral triangle core, with the degeneracy increasing by one every time we go inward by one hexagonal layer. We'll prove this later with tensor methods.

- The simplest examples are the fundamental representation 3 and the antifundamental representation $\overline{3}$, occupied by the quarks and antiquarks.



- Note that the treatment is completely symmetric between I, U, and V. Previously, we used isospin and Clebsch–Gordan coefficients to relate cross sections involving particles in the same isospin irrep. Similarly we can use "U-spin" to relate particles in the same U-spin irrep.

Note. The Schwinger model. We can get intuition for the algebra for $\mathfrak{su}(n)$ in general using something like the quark model. First, consider two uncoupled harmonic oscillators of equal frequency, creation operators $a_i^\dagger$, and number operators N_i. Alternatively, these could be the two components of a two-dimensional harmonic oscillator. Then the operators

$$J_+ = a_1^\dagger a_2, \quad J_- = a_2^\dagger a_1, \quad J_3 = N_1 - N_2$$

form an $\mathfrak{su}(2)$ algebra. The states $|n, m\rangle$ decompose into $\mathfrak{su}(2)$ irreps; acting with J_- on the highest weight vector $|n, 0\rangle$ gives the $n + 1$-dimensional irrep. This is the Schwinger model of angular momentum. We may physically interpret the state $|n, m\rangle$ as containing $n + m$ identical spin $1/2$ particles, symmetrized, with n in the spin up state and m in the spin down state. In the case of isospin, the excitations of the oscillators are up and down quarks, which have isospin $\pm 1/2$. More generally, the Schwinger model works for $\mathfrak{su}(n)$, where the raising and lowering operators create one of n particles and destroy another; the quark model is just the case $n = 3$.

Next, we apply our results to find mass splittings for the four hadron octets/decuplets.

- We begin with the pseudoscalar meson octet, using the wavefunction technique. First, we construct the meson states in terms of the quark states. For example,

  $$|\pi^+\rangle = |u\overline{d}\rangle$$

  which means by isospin lowering that

  $$|\pi^0\rangle = \frac{1}{\sqrt{2}}\left(|u\overline{u}\rangle + |d\overline{d}\rangle\right).$$

  We can then infer $|\eta\rangle$ by orthogonality. None of these manipulations require quarks to actually exist; we can simply think of them at this stage as a useful notational device.

- Continuing, we find the meson wavefunctions fit into a traceless tensor Φ^i_j so that

  $$\Phi = \begin{pmatrix} \pi^0/\sqrt{2} + \eta/\sqrt{6} & \pi^+ & K^+ \\ \pi^- & -\pi^0/\sqrt{2} + \eta/\sqrt{6} & K^0 \\ K^- & \overline{K}^0 & -2\eta/\sqrt{6} \end{pmatrix}.$$

  Under this notation, we have

  $$\Phi^i_j|i\overline{j}\rangle = \pi^0|\pi^0\rangle + \ldots + \eta|\eta\rangle.$$

  If we wanted to work in more detail, the coefficients of the kets here could be other wavefunctions that account for, e.g. position degrees of freedom.

- In principle we could have parametrize the diagonal in a different way. However, mesons with strange quark content are quite different from those without it, because of the strange quark mass, so the π^0 should be split off on its own. Then the η is determined by orthogonality.

- The utility of this notation is that we can write down Lagrangians or Hamiltonians that are $SU(8)$ scalars by just properly contracting all the indices. We are interested in mass terms which are bilinear in Φ, and

$$8 \times 8 = 27 + 10 + 10^* + 8 + 8 + 1.$$

  One has to be careful because matrix and tensor notation conflict. With tensors, the 1 gives

$$\langle \hat{H}_0 \rangle = \overline{\Phi}^i_j \Phi^j_i.$$

  where the bar denotes a bra wavefunction. In matrix notation, taking the bar means taking the adjoint, giving

$$\langle \hat{H}_0 \rangle = {\Phi^\dagger}^i_j \Phi^j_i = {\Phi^*}^j_i \Phi^j_i = \sum_{ij} |\Phi^j_i|^2$$

  where, since we're no longer working with tensors, the indices no longer line up. All the masses are the same, as expected. If we worked only with matrices, we could also have constructed the 1 by constructing a valid scalar expression from $\Phi^\dagger$ and Φ, which here would be $\operatorname{tr} \Phi^\dagger \Phi$.

- Next, we break $SU(3)$ while preserving isospin. The situation is simplified because the mesons are each others' antiparticles; since antiparticles have the same mass as the corresponding particles, the perturbing Hamiltonian $\hat{H}'$ must be symmetric under $\Phi \to \Phi^\dagger$. Thus it must contain one of the symmetric pieces of 8×8, leaving only the 27 and 8.

- We guess the 27 does not contribute. Then the $\hat{H}'$ wavefunction H' is in the 8, so it is again a traceless matrix in flavor space, which must commute with isospin and hypercharge. Thus we must have

$$\langle \hat{H}' \rangle = \operatorname{tr}((\Phi^\dagger \Phi + \Phi \Phi^\dagger) H') = \lambda \operatorname{tr}((\Phi^\dagger \Phi + \Phi \Phi^\dagger) T^8).$$

  Since $\hat{H}_0$ and $\hat{H}'$ are described by two parameters and we have three distinct masses (four isospin triplets, but two related by $\widehat{C}$), we expect to get one nontrivial relation.

- For convenience, we may shift H' by the identity so only H'^3_3 is nonzero. Switching to matrix notation, we have

$$\langle \hat{H}' \rangle \propto \sum_i |\Phi^i_3|^2 + |\Phi^3_i|^2 = |K^-|^2 + |\overline{K}^0|^2 + |K^0|^2 + |K^+|^2 + \frac{4}{3}|\eta|^2$$

  from which we conclude

$$4m_K = 3m_\eta + m_\pi.$$

  This is reasonably accurate, but becomes much more accurate if we square all the masses. This is fair, since both expressions hold to first order in the perturbation, and the squared version can be justified to be more accurate by chiral perturbation theory.

- There's an easier way to come to this conclusion: if the entire Hamiltonian is determined by two parameters, those parameters are essentially the up/down quark mass and the strange quark mass. Then the mass splittings are entirely due to the amount of strange quark content in each meson. This isn't true in general but it's a nice shortcut.

Now we consider the other octets and decuplets, which present more challenges.

- In the vector meson octet, the analogous formula is wrong, because the mesons we observe don't have the naive quark content shown above; mesons with the same quantum numbers can mix, and the isospin singlet ω mixes with the ϕ so that two of the physical states are

$$|s\overline{s}\rangle, \quad \frac{1}{\sqrt{2}}\left(|u\overline{u}\rangle + |d\overline{d}\rangle\right).$$

  This makes sense because the mixing is induced by the $SU(3)$-breaking perturbation.

- Given that mixing can happen, we must explain why the η doesn't mix with the η'. This has to do with instanton effects and the axial anomaly, which conspire to make the η' much heavier.

- Next, consider the baryon octet. The reasoning is similar, with

$$\Phi = \begin{pmatrix} \Sigma^0/\sqrt{2} + \Lambda/\sqrt{6} & \Sigma^+ & p \\ \Sigma^- & -\Sigma^0/\sqrt{2} + \Lambda/\sqrt{6} & n \\ \Xi^- & \overline{\Xi}^0 & -2\Lambda/\sqrt{6} \end{pmatrix}.$$

  The antiparticles sit in a different octet, so we can't use symmetry. Then we have two terms,

$$\langle \hat{H}' \rangle = \lambda_1 \operatorname{tr} \Phi^\dagger \Phi T^8 + \lambda_2 \operatorname{tr} \Phi \Phi^\dagger T^8 \sim \sum_i \lambda_1 |\Phi_3^i|^2 + \lambda_2 |\Phi_i^3|^2.$$

  Since the Hamiltonian has three terms and there are four masses, we again have a relation,

$$2(m_p + m_\Xi) = 3m_\Lambda + m_\Sigma.$$

  This is the Gell-Mann–Okubo formula.

- Finally, we consider the baryon decuplet, also called the 'hadron resonances'. Given the above successes we again assume H' is in an 8. But since $10 \times \overline{10} \times 8$ has only one factor of 1, there is only one term. It must be proportional to the hypercharge, so the splittings are uniform,

$$m_{\Sigma^*} - m_\Delta = m_{\Xi^*} - m_{\Sigma^*} = m_\Omega - m_{\Xi^*}.$$

  This was used by Gell-Mann to predict the mass of the Ω.

- Given the quark model, the part of the Hamiltonian that is bilinear in the quarks must be in $3 \times \overline{3} = 8 + 1$, justifying the assumption that H' is in an 8.

- One final example is the computation of the baryon octet magnetic moments. The magnetic moment operator must be proportional to the charge Q, so it is in 8 by the Gell-Mann–Nishijima formula. Thus there are two allowed terms, so all of the magnetic moments can be written in terms of the proton and neutron magnetic moments.

6 The Cartan Classification

6.1 ⊛ The Cartan–Weyl Basis

We now introduce the Cartan–Weyl basis. In this section, all Lie algebras are implicitly complex.

- We say $X \in \mathfrak{g}$ is ad-diagonalizable if $\mathrm{ad}_X : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable. A Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a maximal abelian subalgebra containing only ad-diagonalizable elements. They are not unique, but their dimension $r = \dim \mathfrak{h}$ is, and is called the rank of $\mathfrak{g}$. Physically, the rank is the number of independent quantum numbers.

- The adjoint maps of the generators H^i of a Cartan subalgebra $\mathfrak{g}$ commute, as

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

  Then they are simultaneously diagonalizable, and the rest of the Lie algebra is spanned by simultaneous eigenvectors E^α, called step operators, which satisfy

$$[H^i, E^\alpha] = \alpha^i E^\alpha$$

  where α is an r-dimensional complex vector called a root. Note that α is nonzero, because elements with $\alpha = 0$ would be in the Cartan subalgebra. A basis consisting of the H^i and E^α is called a Cartan–Weyl basis of $\mathfrak{g}$.

- The set of roots Φ of $\mathfrak{g}$ is called the root space. We can think of each root α as an element of the dual space $\mathfrak{h}^*$, so that $\alpha(H)$ is the eigenvalue of E^α under ad_H,

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

  Here the $e_i \in \mathbb{C}$ are the components of H.

Example. As we've already seen, for $\mathfrak{su}(2)$ we have $r = 1$, where we may take $\mathfrak{g}_0$ to be spanned by $H = 2iT^3 = \mathrm{diag}(1, -1)$. For $\mathfrak{su}(n)$ note that we have the commuting operators

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

To show that this is indeed a Cartan subalgebra, note that the general element in the span of the H^i has the form $H = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with the λ_i summing to zero. Then if $\mathrm{ad}_H X = \mu X$,

$$(\lambda_\ell - \lambda_m) X_{\ell m} = \mu X_{\ell m}$$

with no summation. The solutions are of the form

$$X = E^{(r,s)}, \quad E^{(r,s)}_{\ell m} = \delta_{\ell r}\delta_{ms}, \quad \mu = \lambda_r - \lambda_s$$

with $r \neq s$. Together, the H^i and the $E^{(r,s)}$ span the algebra, and all of the roots of the $E^{(r,s)}$ are nonzero. Hence the H^i are a Cartan subalgebra and $\mathfrak{su}(n)$ has rank $n - 1$.

Note. In the special case of $\mathfrak{su}(n)$, the ad-diagonalizable elements are precisely the diagonal elements. In general, for a matrix Lie algebra, a good guess for a Cartan subalgebra is the elements that are diagonal or almost diagonal, as we'll see in more detail below.

To make further progress, we introduce the Killing form.

- The Killing form is defined as
$$\kappa(X, Y) = \operatorname{tr}(\operatorname{ad}_X \operatorname{ad}_Y).$$
It is symmetric by the cyclic property of the trace, which is a trace for linear operators on $\mathfrak{g}$, and bilinear, so it is an inner product.

- Explicitly, working in a basis $\{T^a\}$, we have
$$[X, [Y, Z]] = X_a Y_b Z_c [T^a, [T^b, T^c]] = X_a Y_b Z_c f^{ad}{}_e f^{bc}{}_d T^e.$$
Finally, taking the trace contracts the indices e and c together, giving
$$\kappa(X, Y) = \kappa^{ab} X_a Y_b, \quad \kappa^{ab} = f^{ad}{}_c f^{bc}{}_d$$
where the κ^{ab} are the components of the Killing form.

- The Killing form is invariant under the adjoint action of $\mathfrak{g}$,
$$\kappa([Z, X], Y) + \kappa(X, [Z, Y]) = 0.$$
To show this, we simply expand the definitions,
$$\kappa([Z, X], Y) = \operatorname{tr}(\operatorname{ad}_{[Z,X]} \operatorname{ad}_Y) = \operatorname{tr}(\operatorname{ad}_Z \operatorname{ad}_X \operatorname{ad}_Y - \operatorname{ad}_X \operatorname{ad}_Z \operatorname{ad}_Y).$$
This cancels with the other term using the cyclic property of the trace. Intuitively, this is just the infinitesimal version of the adjoint action of G on $\mathfrak{g}$, i.e. conjugation by a group element, and the two terms come from the product rule. This indicates that the Killing form is 'the same' everywhere in the group.

- For a simple Lie algebra, it can be shown that the properties of symmetry, linearity, and invariance under the adjoint action determine the inner product up to scalar multiples.

- A real Lie algebra is of compact type if there is a basis where the Killing form is negative definite. It can be shown that every finite-dimensional complex semi-simple Lie algebra has a real form of compact type, and if a Lie group G is compact, its Lie algebra is of compact type.

- Cartan's criterion states the Killing form is nondegenerate if and only if $\mathfrak{g}$ is semi-simple.

  We'll show only the forward direction. Suppose $\mathfrak{g}$ has an abelian ideal $\mathfrak{i}$ so that the ideal has basis $\{T^i\}$ and the rest of the Lie algebra has basis $\{T^a\}$. Then $[T^i, T^j] = 0$ since $\mathfrak{i}$ is abelian, and $[T^i, T^a] \in \mathfrak{i}$ since $\mathfrak{i}$ is an ideal.

  Now consider $\operatorname{ad}_X \operatorname{ad}_Y Z$ where $Y \in \mathfrak{i}$. If $Z \in \mathfrak{i}$, this is automatically zero. If $Z \notin \mathfrak{i}$, then $\operatorname{ad}_Y Z \in \mathfrak{i}$, and hence $\operatorname{ad}_X \operatorname{ad}_Y Z \in \mathfrak{i}$. Hence $\operatorname{tr} \operatorname{ad}_X \operatorname{ad}_Y = 0$, so $\kappa(X, Y) = 0$ for any $Y \in \mathfrak{i}$, so κ is degenerate.

- As a partial converse, we'll show that if $\mathfrak{g}$ is simple, the Killing form is nondegenerate. If the Killing form were degenerate, then the set of elements Y so that $\kappa(X, Y) = 0$ for all $X \in \mathfrak{g}$ forms a nontrivial ideal $\mathfrak{i}$, as
$$\kappa(X, [Z, Y]) = -\kappa([Z, X], Y) = 0$$
for any $Z \in \mathfrak{g}$, so $[Z, Y] \in \mathfrak{i}$.

Example. Consider a simple matrix Lie algebra. Then $\operatorname{tr} XY$ satisfies the properties of the Killing form, so it is proportional to the Killing form; this makes computations much easier. In particular, it's easy to show the Killing form is negative definite for $\mathfrak{u}(n)$, since for $X \in \mathfrak{u}(n)$,

$$\operatorname{tr} X^2 = -\operatorname{tr} XX^\dagger = -\sum_{ij} |X_{ij}|^2 < 0$$

for any nonzero X.

One might wonder why $\operatorname{tr} X \operatorname{tr} Y$ wouldn't work as well. This quantity vanishes identically, since $\operatorname{tr} E^\alpha = 0$ since there are no diagonal elements, while $\operatorname{tr} H^i = 0$ because the eigenvalues of $\mathfrak{su}(2)$ representations sum to zero. Note that the latter statement fails when there are abelian ideals.

Note. The Killing form defines a metric on T_eG, and we may extend this to a metric on G by left and right-translation. Adding an extra minus sign, if $\mathfrak{g}$ is of compact type, then the metric on G is positive definite, and it can be shown that G is compact. Geodesics corresponding to the Levi–Civita connection of this metric are precisely the one-parameter subgroups of G.

Note. The Killing form is used to raise and lower all indices in $\mathfrak{g}$, e.g. we may define $f^{abc} = \kappa^{cd} f^{ab}{}_d$. Now for a Lie algebra of compact type, suppose we choose a basis T^a where $\kappa^{ab} = -\delta^{ab}$. Then

$$0 = \kappa(T^a, [T^b, T^c]) + \kappa([T^b, T^a], T^c) = f^{bca} + f^{bac}$$

so in this basis, f^{abc} is totally antisymmetric. We've often chosen our bases to satisfy this, such as in $\mathfrak{su}(2)$, where we had $f^{abc} = \epsilon^{abc}$. As another example, the quadratic Casimir is generally defined as $\kappa^{ab} T_a T_b$, but we defined it earlier as $\sum_a (T_a)^2$ because we used this special basis.

Next, we find some more properties of the root system of a semi-simple Lie algebra.

- The roots span all of $\mathfrak{h}^*$. This is just because if some direction were not represented, we would have a Cartan subalgebra element that commuted with everything in the group, giving an abelian ideal.
- The roots are nondegenerate, so there is exactly one step operator E^α for each root α. We'll prove this fact below and simply take it as given here.
- Note that $\operatorname{ad}_{E_\alpha}$ raises roots by α, where we regard the Cartan subalgebra itself as associated with the zero root. Then we must have

  $$\kappa(H^i, E^\alpha) = 0, \quad \kappa(E^\alpha, E^\beta) = 0 \text{ if } \alpha + \beta \neq 0$$

  because only $\operatorname{ad}_{E^\alpha} \operatorname{ad}_{E^{-\alpha}}$ and $\operatorname{ad}_{H^i} \operatorname{ad}_{H^j}$ can take roots to themselves.
- To prove these results more formally, we can use the invariance of the Killing form. Since

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

  for any H', we must have $\kappa(H, E^\alpha) = 0$. Similarly

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

  so if $\alpha + \beta \neq 0$, then $\kappa(E^\alpha, E^\beta) = 0$.

- Next, we can get constraints from the nondegeneracy of the Killing form.
  - If α is a root, so is $-\alpha$, with $\kappa(E^\alpha, E^{-\alpha}) \neq 0$, because otherwise $\kappa(E^\alpha, \cdot) = 0$.
  - The Killing form is nondegenerate on $\mathfrak{h}$, as if $\kappa(H^i, H^j) = 0$ for all j, then $\kappa(H^i, \cdot) = 0$.
- Since the Killing form is a nondegenerate inner product on $\mathfrak{h}$, we may use it to correspond elements of $\mathfrak{h}$ and $\mathfrak{h}^*$ and hence define a nondegenerate inner product on $\mathfrak{h}^*$. In components,
$$\kappa(H, H') = \kappa^{ij} e_i e'_j, \quad H = H^i e_i, \quad H' = H^i e'_i$$
and the dual element H^α of α is defined by
$$\kappa(H^\alpha, H) = \alpha(H), \quad H^\alpha = (\kappa^{-1})_{ij} \alpha^j H^i.$$
We define an inner product on the H^α, and hence an inner product on the roots, by
$$(\alpha, \beta) = \kappa(H^\alpha, H^\beta) = (\kappa^{-1})_{ij} \alpha^i \beta^j.$$
The inverse here is just the result of lowering indices, analogous to how the metric and inverse metric are related in differential geometry.

Next, we work out more of the algebra in the Cartan–Weyl basis.

- By the Jacobi identity, we have
$$[H^i, [E^\alpha, E^\beta]] = -[E^\alpha, [E^\beta, H^i]] - [E^\beta, [H^i, E^\alpha]] = (\alpha^i + \beta^i)[E^\alpha, E^\beta].$$
Therefore, $[E^\alpha, E^\beta]$ is proportional to $E^{\alpha+\beta}$ if $\alpha + \beta$ is a root, as anticipated above.
- For the case $\alpha + \beta = 0$, note that
$$\kappa([E^\alpha, E^{-\alpha}], H) = \kappa(E^\alpha, [E^{-\alpha}, H]) = \alpha(H)\kappa(E^\alpha, E^{-\alpha}) = \kappa(H^\alpha, H)\kappa(E^\alpha, E^{-\alpha})$$
where we used the invariance of the Killing form. By the results we found from nondegeneracy of the Killing form, we may conclude
$$H^\alpha = \frac{[E^\alpha, E^{-\alpha}]}{\kappa(E^\alpha, E^{-\alpha})}.$$
- In summary, the algebra for the step operators takes the form
$$[E^\alpha, E^\beta] = \begin{cases} \kappa(E^\alpha, E^{-\alpha}) H^\alpha & \alpha + \beta = 0, \\ N_{\alpha,\beta} E^{\alpha+\beta} & \alpha + \beta \in \Phi, \\ 0 & \text{otherwise.} \end{cases}$$
Here, the $N_{\alpha,\beta}$ are unknown complex numbers. Finally, we have
$$[H^\alpha, E^\beta] = (\kappa^{-1})_{ij} \alpha^i [H^j, E^\beta] = (\kappa^{-1})_{ij} \alpha^i \beta^j E^\beta = (\alpha, \beta) E^\beta.$$
- To simplify these relations, we rescale all of our operators, as
$$e^\alpha = \sqrt{\frac{2}{(\alpha, \alpha)\kappa(E^\alpha, E^{-\alpha})}} E^\alpha, \quad h^\alpha = \frac{2}{(\alpha, \alpha)} H^\alpha.$$
Here we've implicitly assumed that $(\alpha, \alpha) \neq 0$, which we will show below. Our algebra simplifies to the final form
$$[h^\alpha, h^\beta] = 0, \quad [h^\alpha, e^\beta] = \frac{2(\alpha, \beta)}{(\alpha, \alpha)} e^\beta, \quad [e^\alpha, e^\beta] = \begin{cases} h^\alpha & \alpha + \beta = 0, \\ n_{\alpha,\beta} e^{\alpha+\beta} & \alpha + \beta \in \Phi, \\ 0 & \text{otherwise.} \end{cases}$$

6.2 ⊛ The Root Space

Our normalization above makes it easy to identify $\mathfrak{sl}(2)$ subalgebras of $\mathfrak{g}$.

- For any root α, the set $\{h^\alpha, e^\alpha, e^{-\alpha}\}$ obeys
$$[h^\alpha, e^{\pm\alpha}] = \pm 2e^{\pm\alpha}, \quad [e^\alpha, e^{-\alpha}] = h^\alpha$$
which are exactly the commutation relations for $\mathfrak{sl}(2)$. We call this subalgebra $\mathfrak{sl}(2)_\alpha$.

- Define the α-string passing through β as the set
$$S_{\alpha,\beta} = \{\beta + n\alpha \in \Phi \,|\, n \in \mathbb{Z}\}$$
and define a corresponding subspace of $\mathfrak{g}$,
$$V_{\alpha,\beta} = \text{span}(\{e^\delta \,|\, \delta \in S_{\alpha,\beta}\}).$$

- Next, consider the action of $\mathfrak{sl}(2)_\alpha$ on $V_{\alpha,\beta}$. We have
$$[h^\alpha, e^{\beta+n\alpha}] = \frac{2(\alpha, \beta + n\alpha)}{(\alpha,\alpha)} e^{\beta+n\alpha} = \left(\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n\right) e^{\beta+n\alpha}$$
and
$$[e^{\pm\alpha}, e^{\beta+n\alpha}] \propto \begin{cases} e^{\beta+(n\pm1)\alpha} & \text{if } \beta + (n \pm 1)\alpha \text{ is a root,} \\ 0 & \text{otherwise.} \end{cases}$$
Therefore, $V_{\alpha,\beta}$ is a representation of $\mathfrak{sl}(2)_\alpha$.

- Now we can prove the nondegeneracy of the roots. Let $V_{\alpha,\alpha}$ contain all raising and lowering operators with roots $n\alpha$, plus the single operator h^α for the 'zero root'. By similar reasoning to above, it is a representation of $\mathfrak{sl}(2)_\alpha$, where all the weights are even integers.

  Since there is only one element with weight zero, $V_{\alpha,\alpha}$ is an odd-dimensional $\mathfrak{sl}(2)_\alpha$ irrep. But we also know $\mathfrak{sl}(2)_\alpha$ is a subrepresentation of $V_{\alpha,\alpha}$, so $\mathfrak{sl}(2)_\alpha = V_{\alpha,\alpha}$. Then the roots $\pm\alpha$ are nondegenerate, and furthermore no other integer multiples of α can be roots.

- Note that roots of the form $\beta = \pm\alpha/2$ above are allowed by $\mathfrak{sl}(2)_\alpha$ representation theory. But if β is a root, then 2β is a root, a contradiction. Therefore if α is a root, the only nonzero multiple of it that is also a root is $-\alpha$.

- Returning to the general case $V_{\alpha,\beta}$ note that the weights are
$$S = \left\{ \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n \,\middle|\, \beta + n\alpha \in S_{\alpha,\beta} \right\}.$$
Since the weights are nondegenerate and evenly spaced, the representation is a finite-dimensional irrep. Therefore we must have $S = \{\Lambda, -\Lambda + 2, \ldots, \Lambda - 2, \Lambda\}$ for some integer Λ.

- If the range of n is $n_- \leq n \leq n_+$, we have
$$-\Lambda = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_-, \quad \Lambda = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n_+.$$
Adding these equations gives
$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = -(n_+ + n_-) \in \mathbb{Z}.$$

The result above gives a strong constraint as the roots, as we now show.

- First, we return to the original definition of the Killing form. Since $[H^i, E^\alpha] = \alpha^i E^\alpha$,

$$\kappa^{ij} = \kappa(H^i, H^j) = \mathrm{tr}(\mathrm{ad}_{H_i}\,\mathrm{ad}_{H_j}) = \sum_{\delta\in\Phi} \alpha^i\alpha^j.$$

  This is essentially a completeness relation, as

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

- Now, we know the ratio $(\alpha,\beta)/(\alpha,\alpha)$ is real from our work above. But then

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

  which tells us that (β,β) is real, and hence (α,β) is real. Moreover, we have

$$(\alpha,\alpha) = \sum_{\delta\in\Phi}(\alpha,\delta)^2 > 0$$

  by nondegeneracy. Therefore, if we restrict the root space $\mathfrak{h}^*$ to the real span of the roots $\mathfrak{h}^*_{\mathbb{R}}$, we have a Euclidean inner product. This is important, since so far every structure introduced has been complex.

- The real span of the roots has the same dimension, as a real vector space, as $\mathfrak{h}^*$ does as a complex vector space. To see this, choose a basis of roots $\alpha_{(i)} \in \mathfrak{h}^*$. Then for any $\beta \in \Phi$,

$$\beta = \sum_i \beta^i\alpha_{(i)}, \quad (\beta,\alpha_{(j)}) = \sum_i \beta^i(\alpha_{(i)},\alpha_{(j)}).$$

  Since the inner product is nondegenerate, combining these equations for all j gives $\beta^i \in \mathbb{R}$, so $\Phi \subset \mathrm{span}_{\mathbb{R}}(\{\alpha_{(i)}\})$, which is an r-dimensional real vector space.

- We can thus define the length of a root $|\alpha| = \sqrt{(\alpha,\alpha)}$, as well as angles ϕ between roots in the standard way. Now we apply our earlier constraints, for

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

  Multiplying these equations gives

$$4\cos^2\phi \in \mathbb{Z}$$

  which implies the angles between roots must be 0, $\pi/6$, $\pi/4$, $\pi/3$, $\pi/2$, or their supplements.

- These conditions are geometrically intuitive. The constraint $2(\alpha,\beta)/(\alpha,\alpha) \in \mathbb{Z}$ says that the states in the α-string through β have α values that are half-integers or integers, in accordance with $\mathfrak{sl}(2)$ representation theory. The other constraint says the same for the β-string through α. Then the constraint is trivial when α and β are orthogonal, but very restrictive otherwise.

- To summarize, in the Cartan–Weyl basis the Killing form is a block-diagonal matrix, with a block for the Cartan subalgebra, and 2×2 blocks for pairs of roots $\pm\alpha$.

Next, we define simple and positive roots.

- We divide Φ into two halves $\Phi_\pm$, called the positive and negative roots, by drawing an arbitrary hyperplane through the origin that does not intersect any root. Note that the $\Phi_\pm$ are each closed under addition. We call step operators associated with positive roots raising operators, and with negative roots lowering operators.
- A simple root is a positive root which cannot be written as a positive linear combination of positive roots. Geometrically, these are typically the positive roots closest to the hyperplane.
- If α and β are simple roots, then $\alpha - \beta$ is not a root. To see this, note that if $\alpha - \beta$ were a positive root, α could not be simple, while if it were a negative root, β could not be simple.
- If α and β are distinct simple roots, then the α-string through β has

$$n_- = 0, \quad n_+ = -\frac{2(\alpha, \beta)}{(\alpha, \alpha)}$$

  Since we know $n_+ \geq 0$, we have $(\alpha, \beta) \leq 0$.
- The simple roots are linearly independent. Denoting the simple roots by $\alpha_{(i)}$, let

$$\lambda = \sum_{c_i \geq 0} c_i \alpha_{(i)} + \sum_{c_i < 0} c_i \alpha_{(i)} \equiv \lambda_+ - \lambda_-.$$

  Then we have, assuming the c_i are not all zero,

$$(\lambda, \lambda) = (\lambda_+ - \lambda_-, \lambda_+ - \lambda_-) > -2(\lambda_+, \lambda_-) = 2 \sum_{c_i \geq 0} \sum_{c_j < 0} c_i c_j (\alpha_{(i)}, \alpha_{(j)}) > 0$$

  because $(\alpha, \beta) \leq 0$ for simple roots, so $\lambda \neq 0$ as desired.
- Every positive root can be written as a linear combination of the simple roots with nonnegative integer coefficients. This can be shown recursively: if α is positive and simple, we're done. Otherwise $\alpha = \alpha_1 + \alpha_2$ and we may repeat the procedure until we get the desired decomposition.
- As a result, all roots can be written as an integer combination of the simple roots. Then the simple roots are a basis for $\mathfrak{h}^*_\mathbb{R}$, so there are r of them.

6.3 ⊛ The Cartan Matrix

The content of the simple roots can be encoded in the Cartan matrix.

- Define the elements of the Cartan matrix by

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

  We note that $A^{ii} = 2$, and $A^{ij} = 0$ if and only if $A^{ji} = 0$. Since the inner product of roots is negative, $A^{ij} \leq 0$ for $i \neq j$. Intuitively, $-A^{ij}$ is the number of times $\alpha_{(j)}$ can be added to $\alpha_{(i)}$ and yield a root; this interpretation also works for the diagonal elements.

- For each simple root $\alpha_{(i)}$, we have an $\mathfrak{su}(2)$ subalgebra with operators $h^i \equiv h^{\alpha_{(i)}}$, $e^i_\pm \equiv e^{\alpha_{(i)}}_\pm$,

$$[h^i, e^i_\pm] = \pm 2 e^i_\pm, \quad [e^i_+, e^i_-] = h^i.$$

  These $3r$ operators together generate all of $\mathfrak{g}$ by brackets. A Cartan–Weyl basis chosen so that these $3r$ operators are normalized in this way is called a Chevalley basis.

- In a Chevalley basis, the algebra is

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

  We also have the Serre relations

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

  by the intuitive interpretation of A^{ij} above.

- A finite-dimensional, semi-simple complex Lie algebra is uniquely determined by its Cartan matrix. To reconstruct $\mathfrak{g}$, we start with the simple roots and construct all $\alpha^{(i)}$ strings through $\alpha^{(j)}$, with the Cartan matrix telling us the length of all root strings. We then repeat this procedure until no more new roots appear.

- The Cartan matrix satisfies $\det A > 0$. To see this, note that the inner product in the basis of simple roots is

$$(\lambda, \mu) = (\alpha_{(i)}, \alpha_{(j)}) \lambda^i \mu^j, \quad \lambda = \sum_i \lambda^i \alpha_{(i)}, \quad \mu = \sum_i \mu^i \alpha_{(i)}.$$

  Since the inner product is positive definite, the matrix with entries $(\alpha_{(i)}, \alpha_{(j)})$ is positive definite and hence has positive determinant. The Cartan matrix is the product of this matrix with a diagonal matrix with entries $2/(\alpha_{(i)}, \alpha_{(i)})$ which also has positive determinant.

- If $\mathfrak{g}$ is simple, then the Cartan matrix is not reducible, i.e. there is no reordering of the simple roots that makes A block-diagonal. Essentially, if the Cartan matrix were reducible, then the step operators generated by one of the blocks of simple roots (along with the corresponding Cartan subalgebra elements) would form a proper ideal of $\mathfrak{g}$.

The above constraints very strongly restrict the form of the Cartan matrix.

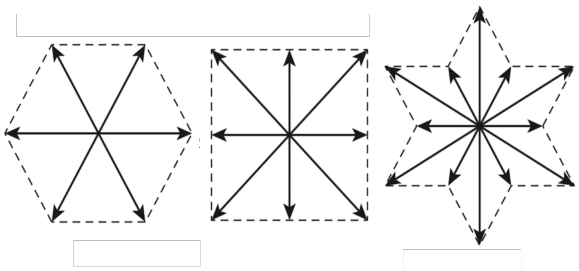
- By our previous identities, we have

$$A^{ij} A^{ji} = 4 \cos^2 \phi \in \mathbb{Z}$$

  which implies that $A^{ij} A^{ji} \in \{0, 1, 2, 3\}$. There are only a few possibilities. Taking the $\alpha^{(i)}$ root to not be shorter without loss of generality, we have:

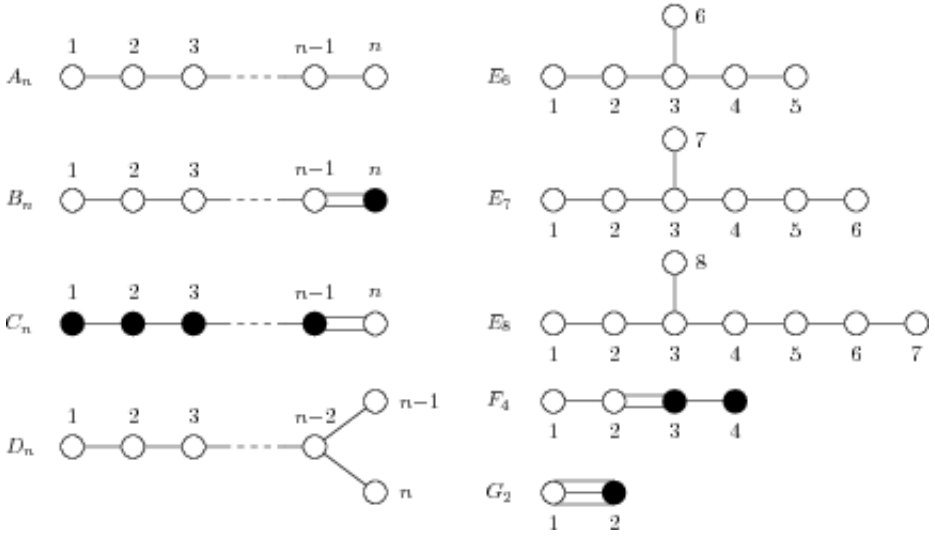
  - $A^{ij} = A^{ji} = 0$. The simple roots are perpendicular with indefinite ratio of lengths; no other roots are attained by adding them.
  - $A^{ij} = A^{ji} = -1$. The simple roots have equal length and angle $120°$, and one additional root is attained by adding them.
  - $A^{ij} = -2$, $A^{ji} = -1$. Then $|\alpha_{(i)}| = \sqrt{2} |\alpha_{(j)}|$ and the angle is $135°$.
  - $A^{ij} = -3$, $A^{ji} = -1$. Then $|\alpha_{(i)}| = \sqrt{3} |\alpha_{(j)}|$ and the angle is $150°$.

The roots generated by these three latter possibilities are shown below.



There are no more possibilities, by the $\det A > 0$ constraint.

- One useful fact is that if α and β are roots, then so is the 'Weyl reflection' of β in the hyperplane normal to α. This can be shown by casework with our results on root strings. The set of Weyl reflections form the Weyl group.
- If $\mathfrak{g}$ is simple, it can be shown that the simple roots can only have two distinct lengths. In fact, one can show there are only two distinct lengths among all the roots. The proof is easy if we ignore the case of perpendicular roots; we can deal with those by Weyl reflection.
- The information in the Cartan matrix can be written in a Dynkin diagram. We draw a node for every simple root $\alpha_{(i)}$, then connect nodes i and j with $\max(|A^{ij}|, |A^{ji}|)$ lines. If the roots have different lengths, we draw an arrow pointing from the longer root to the shorter, or shade the shorter roots black. Note that Dynkin diagrams must be connected, by simplicity.
- The set of all Dynkin diagrams is shown below, classifying all simple complex Lie algebras.



Apart from the five exceptional cases, the four infinite families are known in physics as

$$A_n = \mathfrak{su}(n+1) \cong \mathfrak{sl}(n+1), \quad B_n = \mathfrak{so}(2n+1), \quad C_n = \mathfrak{sp}(2n), \quad D_n = \mathfrak{so}(2n).$$

In all cases the subscript indicates the number of roots. Note that D_1 isn't counted since $\mathfrak{so}(2)$ is not simple. Also, D_2 technically does not belong since it is disconnected.

- Generally, any angle besides 90° provides a strong constraint since it generates new roots, so most angles must be 90°. After that, angles of 120° are nice, as they form regular polyhedra in

higher dimensions. An angle of 150° makes such a strange pattern that it only appears in the rank 2 Lie algebra G_2, whose root system is shown above.

- Looking at the Dynkin diagrams, we can read off some low-dimensional coincidences for the $\mathfrak{so}(n)$. The isomorphic algebras correspond to the spin groups, which double cover $SO(n)$.
  - We have $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ from $B_1 \cong A_1$, and $SU(2) = \text{Spin}(3)$.
  - We have $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ from $D_2 \cong A_1 \oplus A_1$, and $SU(2) \times SU(2) = \text{Spin}(4)$.
  - We have $\mathfrak{so}(5) \cong \mathfrak{sp}(4)$ from $B_2 \cong C_2$, and $Sp(4) = \text{Spin}(5)$.
  - We have $\mathfrak{so}(6) \cong \mathfrak{su}(4)$ from $D_3 \cong A_3$, and $SU(4) = \text{Spin}(6)$.

  Note that E_5 and E_4 aren't listed as they duplicate others.

- It should be kept in mind that we are working at the level of complex Lie algebras; the complexification of $\mathfrak{so}(p,q)$ is equivalent to the complex Lie algebra $\mathfrak{so}(n)_{\mathbb{C}}$, so our results don't depend on the signature. However, the signature does affect the Lie *groups*, e.g. the orthogonal groups of indefinite signature are noncompact. The spin groups of indefinite signature are

  $$\text{Spin}(1,1) \cong GL(2,\mathbb{R}), \quad \text{Spin}(1,2) \cong SL(2,\mathbb{R}), \quad \text{Spin}(1,3) \cong SL(2,\mathbb{C}), \quad \text{Spin}(1,4) \cong Sp(1,1)$$

- Another useful application is that we can 'cut' Dynkin diagrams to find subalgebras. For example, cutting off the long root of B_n shows that $\mathfrak{su}(n) \subset \mathfrak{so}(2n+1)$. Cutting off the short root of $\mathfrak{sp}(2n)$ gives $\mathfrak{su}(n) \subset \mathfrak{sp}(2n)$ as we observed earlier.

Note. A systematic algorithm for reconstructing all roots from the simple roots.

- For a general root $\beta = \sum_i \alpha_{(i)} k_i$, let $k = \sum_i k_i$. Here the Cartan subalgebra has $k = 0$ and the simple roots have $k = 1$.

- For each simple root $\alpha_{(j)}$, we draw a box with entries

  $$c_i = A_{ji} = -(n_+ + n_-).$$

  That is, we just write row j of the Cartan matrix. Hence if $c_i < 0$ then we can add the root $\alpha_{(i)}$, and hence we can construct the $k = 2$ roots.

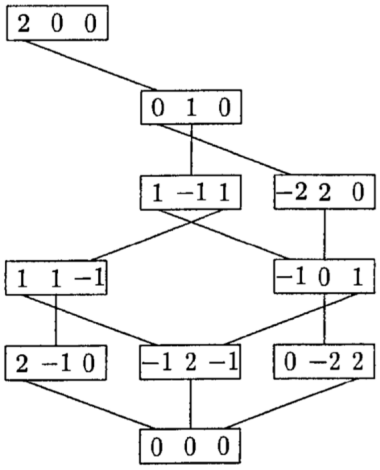
- For each $k = 2$ root, we again draw a box; its entries c_i are the sum of the root it came from and the simple root added to it. We can then continue systematically upward in k, since at every point we know the value of n_-, until the procedure terminates. We then reflect to get all the negative roots.

- A slightly faster method is to start from the simple roots and draw all possible root strings all the way to the end. This will produce some roots with new negative c_i, which sit at the bottom of new root strings. We then repeat this process until termination.

Example. The algebra C_3. The Cartan matrix is

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{pmatrix}$$

and the resulting positive roots are shown below.



Note. Reconstructing the algebra from the roots. For a simple root α, if the α-string through β has length $2s+1$, then the action of the $E^{\pm\alpha}$ operators on this subspace by commutator is just like that of the angular momentum raising and lowering operators for spin s, with the same constants up to phases. This gives the action of E^{α} on everything, and we can find the action of $E^{\alpha+\beta}$ on everything by the Jacobi identity, and so on. The 'up to phases' is because we don't have enough freedom to take Cordan-Shortley phases everywhere.

Note. The only normed division algebras are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$. The classification of Lie algebras is related to this, because $SO(n)$ describes linear transformations that preserve the length of a vector in $\mathbb{R}$, $SU(n)$ does the same for $\mathbb{C}$, and $Sp(n)$ does the same for $\mathbb{H}$. There is no analogue for $\mathbb{O}$ because its multiplication law is not associative, so we don't get a group. But all of the other Lie groups are related to $\mathbb{O}$ in some way.

6.4 ⊛ Representations and Weights

In this section we apply our results to representations of $\mathfrak{g}$.

- We consider an N-dimensional representation R, and assume that the $R(H^i)$ are diagonalizable. Since the $R(H^i)$ all commute, they can be simultaneously diagonalized, with

$$R(H^i)v = \lambda^i v, \quad v \in V_\lambda$$

  where $\lambda \in \mathfrak{h}^*$ is a weight of R and the set of weights is S_R.

- Roots are the weights of the adjoint representation $R(X) = \mathrm{ad}_X$ and also live in $\mathfrak{h}^*$. But unlike roots, weights can be degenerate, with multiplicity $m_\lambda = \dim V_\lambda \geq 1$. It also doesn't make sense to associate a weight with an element of $\mathfrak{g}$. This only made sense for roots because the representation space was the algebra itself.

- As with roots, step operators raise and lower the weights,

$$R(H^i)R(E^\alpha)v = R(E^\alpha)R(H^i)v + [R(H^i), R(E^\alpha)]v = (\lambda^i + \alpha^i)R(E^\alpha)v.$$

  Then if $v \in V_\lambda$, then $R(E^\alpha)v \in V_{\lambda+\alpha}$.

- Next, we consider the action of the $\mathfrak{sl}(2)_\alpha$ generators $\{R(h^\alpha), R(e^\alpha), R(e^{-\alpha})\}$ on V. This makes V into the representation space for a representation R_α of $\mathfrak{sl}(2)$. The $\mathfrak{sl}(2)$ weights are

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

  Since the $\mathfrak{sl}(2)$ weights must be integers, we have

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

  This is just the same condition we found for roots; the only difference is that we don't also have the same constraint with α and λ swapped.

- To understand this constraint geometrically, note that all the roots lie in the root lattice

$$\mathcal{L}[\mathfrak{g}] = \mathrm{span}_{\mathbb{Z}}\{\alpha_{(1)}, \ldots, \alpha_{(r)}\}.$$

  Now define the simple coroots and coroot lattice

$$\alpha^\vee_{(i)} = \frac{2\alpha_{(i)}}{(\alpha_{(i)},\alpha_{(i)})}, \quad \mathcal{L}^\vee[\mathfrak{g}] = \mathrm{span}_{\mathbb{Z}}\{\alpha^\vee_{(1)}, \ldots, \alpha^\vee_{(r)}\}.$$

- Then the weight lattice is the dual of the co-root lattice,

$$\mathcal{L}_W[\mathfrak{g}] = \mathcal{L}^\vee[\mathfrak{g}]^* = \{\lambda \in \mathfrak{h}^* \,|\, (\lambda,\mu) \in \mathbb{Z} \text{ for all } \mu \in \mathcal{L}^\vee[\mathfrak{g}]\}$$

  Consider the basis $\{\alpha^\vee_{(i)}\}$ of $\mathcal{L}^\vee[\mathfrak{g}]$. The weight lattice has the dual basis $\{w_{(i)}\}$ where

$$(\alpha^\vee_{(i)}, w_{(j)}) = \delta_{ij}.$$

  This basis is called the Dynkin basis of the weight space, and its elements are called the fundamental weights of $\mathfrak{g}$.

- Now consider the expansion

$$w_{(i)} = \sum_j B_{ij}\alpha_{(j)}.$$

  Taking the inner product of both sides with $\alpha^\vee_{(k)}$ and relabeling indices,

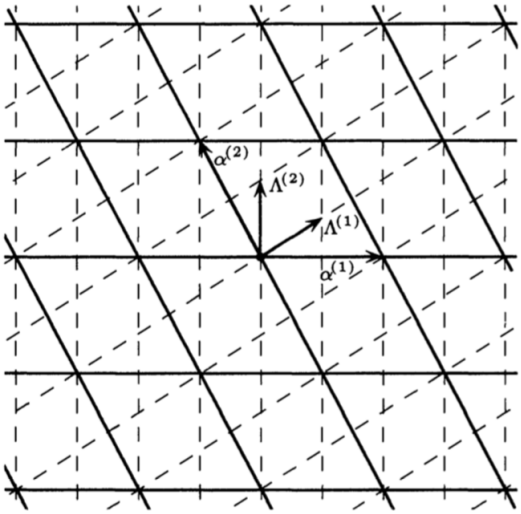
$$\sum_k \frac{2(\alpha_{(i)},\alpha_{(k)})}{(\alpha_{(i)},\alpha_{(i)})}B_{jk} = \delta^i_k$$

  which shows that $B = A^{-1}$ where A is the Cartan matrix, so

$$\alpha_{(i)} = \sum_j A^{ij}w_{(j)}.$$

  Thus starting from the Cartan matrix we can read off the fundamental weights. The matrix B_{ij} is sometimes called the metric tensor.

Example. The root and weight lattices of A_2 are shown below.



The relation between the two can be read off from the Cartan matrix,

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

Note that the root lattice is contained in the weight lattice. This must hold in general because the simple roots are weights of the adjoint representation.

The fundamental weights allow us to extract the weight set from the highest weight alone.

- Any weight can be expanded as
$$\lambda = \lambda^i w_{(i)}$$
and the integers λ^i are called the Dynkin labels of λ. We've already seen these; in our graphical calculation above, the numbers in the boxes were the Dynkin labels of roots/weights.

- Every finite-dimensional representation has a highest weight
$$\Lambda = \Lambda^i w_{(i)}$$
so that all $v \in V_\Lambda$ are annihilated by all raising operators,
$$R(E^\alpha)v_\Lambda = 0, \quad \alpha \in \Phi_+.$$
The Dynkin labels of the highest weight are called the Dynkin labels of the representation.

- Starting from the highest weight, we can find more weights by applying the lowering operators, $R(E^{-\alpha})$ for $\alpha \in \Phi^+$. The reasoning is just like how we used simple roots to construct all the roots, but in reverse; starting with any positive Dynkin label, we can go downward. We can get every weight this way if the representation is an irrep; note that the adjoint representation itself is an irrep when the algebra is simple.

- Note that we never have to go 'upward'. Suppose we had a state of the form
$$E_1 E_2 \dots E_n |\mu\rangle$$

where $|\mu\rangle$ is the highest weight state and the E_i are raising and lowering operators. If any of the E_i are raising operators, we may commute it all the way to the right, picking up extra terms as we go, until it annihilates $|\mu\rangle$. We can repeat this procedure until all operators are lowering operators, which can be taken to be those of simple roots. As a corollary, the highest weight of an irrep is nondegenerate.

- The general principle is that if $\lambda = \sum_i \lambda^i w_{(i)}$ is a weight, then we also have the weights

$$\lambda - m^i \alpha_{(i)}, \quad 0 \leq m_i \leq \lambda^i$$

because representations of $\mathfrak{sl}(2)_{\alpha_{(i)}}$ must have weights symmetric about zero.

- The weight set of a tensor product is the set of sums of weights, which gives a visual method for decomposing tensor products. We get one factor for free, since one of the highest weights is the product of the individual highest weights; for example, $(1,2) \times (5,2)$ contains a $(6,4)$.

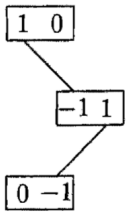
Example. The fundamental representation of $A_2 \cong \mathfrak{su}(3)$ has Dynkin labels $(1,0)$, so $\Lambda = w_{(1)}$, and

$$\Lambda - \alpha_{(1)} = w_{(1)} - (2w_{(1)} - w_{(2)}) = -w_{(1)} + w_{(2)}$$

is also a weight. This yields the new weight

$$(\Lambda - \alpha_{(1)}) - \alpha_{(2)} = -w_{(2)}$$

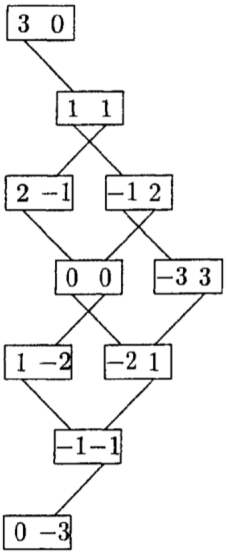
and this weight produces no further weights. Alternatively, we can use a diagram.



Thus the representation has three weights, which form a small upside-down triangle. More generally, the Dynkin labels agree with the labels we assigned earlier based on the ranks of symmetric traceless tensors. Then $R_{(1,1)}$ is still the adjoint representation, $R_{(3,0)}$ is still the 10, and so on.

Note. Finding the degeneracy of weights. Clearly, any weight that can be reached in only one way by lowering from the highest weight is nondegenerate. Also note that a degenerate weight can't create more degeneracy 'further down' unless it sits at the top of a new $\mathfrak{su}(2)$ representation. Another trick is that if a possibly degenerate weight can be related to a nondegenerate weight by Weyl reflection, it is nondegenerate. In general, the degeneracy may be computed by the Freudenthal formula or the Kostant multiplicity formula, though these are rather complicated to use.

Example. The $(3,0)$ representation of $\mathfrak{su}(3)$. We get the diagram shown below.



First, note that there is a 'spine' on the right leading from the highest to the lowest weight. In general for the (n, m) irrep, we can reach it by lowering with α^1 n times, lowering with α^2 $n+m$ times, then lowering with α^1 m times, yielding the lowest weight $(-m, -n)$. There is also the possibility of degeneracy; by the remark above, we only need to check the degeneracy of the $\boxed{0\ 0}$ state; we can't use Weyl reflection since it is at the origin.

Let E_1 and E_2 be the lowering operators for $\alpha_{(1)}$ and $\alpha_{(2)}$. Then

$$E_2E_1E_1 - E_1E_2E_1 = [E_2, E_1]E_1 = E_1[E_2, E_1]$$

where the second step follows because $-2\alpha^1 - \alpha^2$ is not a root. Now act with both sides on the highest weight $|\mu\rangle$. On the right-hand side we have

$$E_1[E_2, E_1]|\mu\rangle = E_1E_2E_1|\mu\rangle$$

because $E_2|\mu\rangle = 0$. This shows that the two possibly degenerate states are proportional.

Note. There is a fully general method for determining whether two states are degenerate. Consider two states of the form

$$|A\rangle = E_{a_1} \dots E_{a_n}|\mu\rangle, \quad |B\rangle = E_{b_1} \dots E_{b_n}|\mu\rangle.$$

Then the states are linearly independent if and only if

$$\langle A|B\rangle\langle B|A\rangle \neq \langle A|A\rangle\langle B|B\rangle$$

where all of these inner products can be computed systematically using the algebra.

Note. Taking the dual/conjugate of a representation just flips the sign on every weight, because the Cartan subalgebra elements $R(H^i)$ become $-R(H^i)^* = -R(H^i)$. For $\mathfrak{su}(3)$, the (n, m) representation has lowest weight $(-m, -n)$, which implies the conjugate representation is (n, m). For example, the antifundamental representation has $\Lambda = w_{(2)}$ and its weights form a triangle. We see that a representation is real if its weights are symmetric about the origin.

Note. In general, we call the representations with exactly one nonzero Dynkin label, which is equal to one, the fundamental representations. For example, $\mathfrak{su}(4)$ has rank three, and hence has three

'fundamental' representations. One is its 'usual' fundamental representation in terms of matrices, one is its conjugate, and the third is something completely different; it is not even four-dimensional. All representations can be found by multiplying fundamental representations, as the representation with Dynkin labels Λ^i is contained in $\otimes_i V_i^{\otimes \Lambda^i}$ where V_i is the i^{th} fundamental.

Example. Consider $B_2 \cong C_2$, where $B_2 \cong \mathfrak{so}(5)$ and $C_2 \cong \mathfrak{sp}(4)$. Then there are two fundamental representations with dimensions 5 and 4; they correspond to the fundamental matrix representations of $\mathfrak{so}(5)$ and $\mathfrak{sp}(4)$. The adjoint representation has Dynkin labels $(0, 2)$.

Note. If we use an explicit representation, rather than just the abstract algebra, everything simplifies. We defined the Killing form and worked to show that it gives a natural inner product on the roots. But we could also work in the adjoint representation and take the inner product on the Hilbert space. The action of X_a in the adjoint representation is

$$X_a|X_b\rangle = |[X_a, X_b]\rangle$$

so it's easy to see how the E^α function as both weights and roots,

$$H^i|E^\alpha\rangle = \alpha^i|E^\alpha\rangle, \quad [H^i, E^\alpha] = \alpha^i E^\alpha.$$

Moreover, we have $H_\alpha^\dagger = H_\alpha$ since its eigenvalues are real, while taking the adjoint of the above commutation relation gives $E^{\alpha\dagger} = E^{-\alpha}$.

6.5 ○ Examples of Roots

The roots are the nonzero weights of the adjoint, and the adjoint is often contained in the product of the fundamental and antifundamental representation. (Here, "fundamental" is used in the physicist's sense; the weights of the physicist's fundamental should not be confused with the fundamental weights.) Therefore, the roots are a subset of the differences of the weights of the fundamental representation. In this section, we'll use this as a shortcut to give examples of roots. Since we deal with only simple matrix Lie algebras, we use the Killing form $\text{tr}(XY)$.

Example. The roots of $\mathfrak{su}(3)$. An orthonormal basis for the Cartan subalgebra in the fundamental representation is

$$H^1 = \text{diag}(1, -1, 0)/\sqrt{2}, \quad H^2 = \text{diag}(1, 1, -2)/\sqrt{6}.$$

These are already diagonal in the standard basis, consisting of the vectors $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$. Their corresponding weights are

$$w^1 = (1, 1/\sqrt{3})/\sqrt{2}, \quad w^2 = (-1, 1/\sqrt{3})/\sqrt{2}, \quad w^3 = (0, -2/\sqrt{3})/\sqrt{2}.$$

These three weights form an equilateral triangle, so the differences are the vertices of a hexagon. As we've seen before, these are precisely the roots of $\mathfrak{su}(3)$. Another common choice is to use a basis like $H^1 = \text{diag}(1, -1, 0)$ and $H^2 = \text{diag}(0, 1, -1)$. While this simplifies the weights, it leads to a non-Euclidean metric in root space, distorting the picture.

Example. The roots of $\mathfrak{so}(4)$. In this case an orthonormal basis for the Cartan subalgebra is, up to an overall scaling factor,

$$H^1 = J^{12} = \begin{pmatrix} & 1 & & \\ -1 & & & \\ & & & \\ & & & \end{pmatrix}, \quad H^2 = J^{34} = \begin{pmatrix} & & & \\ & & & \\ & & & 1 \\ & & -1 & \end{pmatrix}.$$

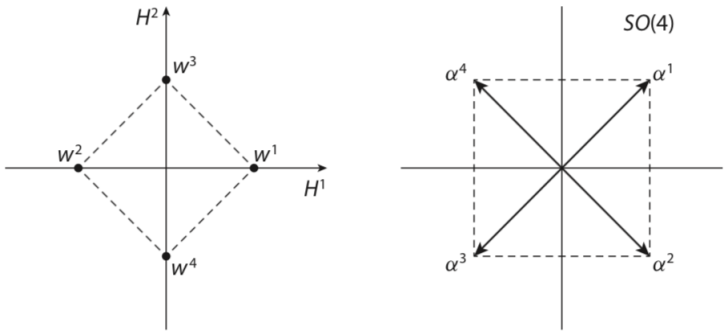
These may be simultaneously diagonalized and scaled to

$$H^1 = \mathrm{diag}(1, -1, 0, 0), \quad H^2 = (0, 0, 1, -1)$$

by switching to the 'polar' basis $(x_1, x_2, x_3, x_4) \to x_1 \pm ix_2, x_3 \pm ix_4$. We can then read off the weights of the fundamental representation,

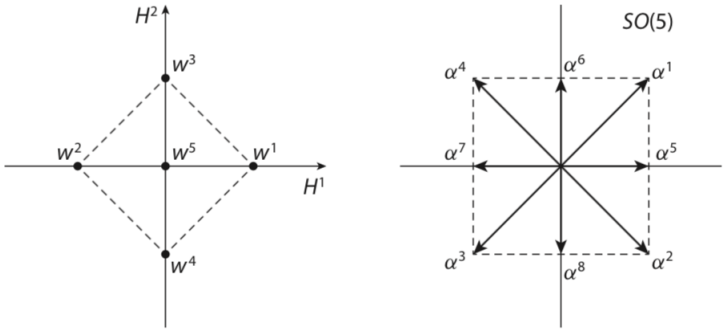
$$w^1 = (1, 0), \quad w^2 = (-1, 0), \quad w^3 = (0, 1), \quad w^4 = (0, -1)$$

which form the vertices of a square. In this case there are some extra differences of weights. As an unjustified ad hoc fix, throwing out $w^1 - w^2$ and its rotations gives the root diagram below.



This root diagram visually shows that $\mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2)$, i.e. the Lie algebra is not semisimple. In this degenerate case, which we only consider because of the importance of the Lorentz group, the adjoint is not an irrep, but rather decomposes as $3 + \overline{3}$.

Example. The roots of $\mathfrak{so}(5)$. There are still two Cartan generators; the only difference is that we now have a zero weight from the vector $(0, 0, 0, 0, 1)$.



In this case the procedure is less unambiguous. If we keep $w^1 - w^5$, then we must toss out $w^1 - w^2$ because it is double $w^1 - w^5$, and likewise for its rotations. Ignoring repeated roots, we recover a familiar root pattern.

Example. The roots of $\mathfrak{so}(6)$. Letting the e_i be unit vectors in root space, and following the pattern of $\mathfrak{so}(4)$, the roots are

$$\{\eta e_i + \eta' e_j \,|\, \eta, \eta \in \{-1, 1\}, i, j \in \{1, 2, 3\}, i \neq j\}.$$

This generalizes directly to $\mathfrak{so}(2n)$. The pattern for $\mathfrak{so}(2n + 1)$ is just slightly more complicated, with the addition of the roots $\pm e_i$. The simple roots are

$$\mathfrak{so}(2n) : e^1 - e^2, \ldots, e^{n-1} - e^n, e^{n-1} + e^n, \quad \mathfrak{so}(2n + 1) : e^1 - e^2, \ldots, e^{n-1} - e^n, e^n$$

from which one can read off the Dynkin diagram.

Example. The roots of $\mathfrak{su}(n)$. Since $\dim \mathfrak{su}(n) = n^2 - 1$, there are $n^2 - n$ roots. Similarly, there are $n(n-1)$ nonzero differences of weights, so all of them must correspond with roots, which is what we saw for $\mathfrak{su}(3)$ above. In the case of $\mathfrak{su}(4)$, we have

$$H^1 = \operatorname{diag}(1, -1, 0, 0)/\sqrt{2}, \quad H^2 = \operatorname{diag}(1, 1, -2, 0)/\sqrt{6}, \quad H^3 = \operatorname{diag}(1, 1, 1, -3)/2\sqrt{3}.$$

The weights are the vertices of a tetrahedron, so the roots are the sides of a tetrahedron. In particular, adjacent sides make an angle of $60°$, while nonadjacent sides make an angle of $90°$, in accordance with the Dynkin diagram. For $\mathfrak{su}(n)$, the pattern continues, with a higher-dimension analogue of the tetrahedron. Unlike for $\mathfrak{so}(n)$, the explicit coordinate expression of the roots is a bit nasty; instead we'll just continue to think of them as differences of fundamental weights.

Example. The roots of $\mathfrak{sp}(2n)$. Earlier, we showed that $\mathfrak{sp}(2n)$ had basis

$$iA \otimes I, \quad S_i \otimes \sigma_i$$

from which we identify a Cartan subalgebra

$$H^i = u^i \otimes \sigma_3, \quad u^i_{jk} = \delta^i_j \delta^i_k.$$

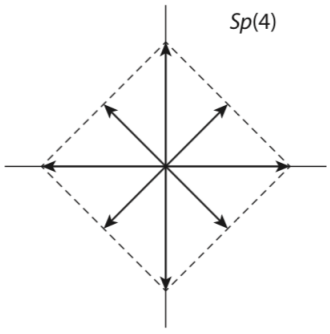
For example, for $\mathfrak{sp}(4)$ we have

$$H^1 = \operatorname{diag}(1, 0, -1, 0), \quad H^2 = \operatorname{diag}(0, 1, 0, -1)$$

from which we read off the weights

$$w^1 = (1, 0), \quad w^2 = (0, 1), \quad w^3 = (-1, 0), \quad w^4 = (0, -1)$$

These are the same weights as for $\mathfrak{so}(4)$, but $\mathfrak{sp}(4)$ has higher dimension. In this case, all of the differences are roots.



Since this is just the root diagram of $\mathfrak{so}(5)$ tilted, $\mathfrak{sp}(4) \cong \mathfrak{so}(5)$.

6.6 ○ Dynkin Diagrams

In this section, we give a quick outline of the Cartan classification.

- First, we can establish the existence of the infinite families A_n, B_n, C_n, and D_n and their Dynkin diagrams by our work in the previous section, reading off angles between the simple roots. We also discovered G_2 earlier by classifying all rank 2 Lie algebras.

- Given simple roots α_i, we define the unit vectors $u_i = \alpha_i/|\alpha_i|$. In this section, we work solely with roots normalized to unit length. Then

$$2u_i \cdot u_j = -\sqrt{\zeta_{ij}}, \quad \zeta_{ij} \in \{0, 1, 2, 3\} = \text{number of lines connecting } i \text{ and } j.$$

  Note that if u_i and u_j are connected at all, then $u_i \cdot u_j \leq -1/2$.

- First, there are no Dynkin diagrams with loops. Summing over the k roots in the loop,

$$\left(\sum_i u_i\right)^2 = k + \sum_{i \neq j} u_i \cdot u_j \leq 0$$

  but the left-hand side must be positive.

- Next, the number of lines coming out of a root cannot be more than three. Let u be directly connected to the roots $w_1, \ldots, w_k$. The roots cannot be directly connected to each other otherwise, since there are loops, so they must be orthogonal. Now note that

$$\sum_i (u \cdot w_i)^2 = \frac{1}{4} \sum_i \zeta_{ui} \geq 1.$$

  But the left-hand side is the length of u when projected down to the subspace spanned by the w_i, so it must be less than one, a contradiction.

- Next, shrinking a chain of roots each connected with a single line gives a valid Dynkin diagram. To see this, let $u_1, \ldots, u_k$ be such a chain. Then we can replace these vectors with the vector $u = \sum_i u_i$, which is properly normalized since

$$u^2 = \sum_{i=1}^{k} u_i^2 + 2 \sum_{i=1}^{k-1} u_i \cdot u_{i+1} = k - 2\frac{k-1}{2} = 1.$$

  Any other root w can be directly connected to the chain only once, or else there will be loops. If w connects to u_i, then $w \cdot u_i = w \cdot u$. Hence everything about the Dynkin diagram stays the same when we collapse the u_i to u.

- Suppose a Dynkin diagram contains a 'fork', i.e. there is a root connected to three other roots. Then all of the connections must be by single lines, and moreover, each of the three roots can only continue in a linear chain made of single lines. We'll return to this case below.

- Our results above show that G_2 is the only Dynkin diagram with a triple line. Next, consider a double line; we can only connect single line chains on both ends. Suppose there are n and m of them, u_i and v_i, with the last ones connected by the double line, so $u_m \cdot v_m = -1/\sqrt{2}$.

  We define $u = \sum_i i u_i$ and $v = \sum_j j v_j$. Then we have

$$u^2 = \frac{n(n+1)}{2}, \quad v^2 = \frac{m(m+1)}{2}, \quad u \cdot v = -\frac{nm}{\sqrt{2}}.$$

  Then the Cauchy–Schwarz inequality gives $(m-1)(n-1) \leq 2$. The cases $m = 1$ or $n = 1$ give the infinite families B_n and C_n. The only other case is $m = n = 2$, which gives F_4.

- Finally, we classify Dynkin diagrams with a fork, with n, m, and p roots in each chain, including the central root. Similar considerations show that

$$\frac{1}{n}+\frac{1}{m}+\frac{1}{p}>1.$$

  Then if $n=m=1$, then p is arbitrary, recovering A_n. If all of the numbers are at least 3, the inequality is violated, so otherwise one of them must be two, say $p=2$. Then

$$\frac{1}{n}+\frac{1}{m}>\frac{1}{2}.$$

  In the case $n=2$, m is arbitrary, recovering D_n. Finally, the only other solutions are $(3,3)$, giving E_6, $(3,4)$, giving E_7, and $(3,5)$, giving E_8.

- Incidentally, the same inequality occurs when classifying the platonic solids; there E_6 corresponds to the tetrahedron, E_7 to the cube, and E_8 to the icosahedron.

- It is possible to have higher rank exceptional algebras such as E_9 if we generalize to Kac-Moody algebras, which are generically infinite-dimensional.

As an application, we consider regular subalgebras.

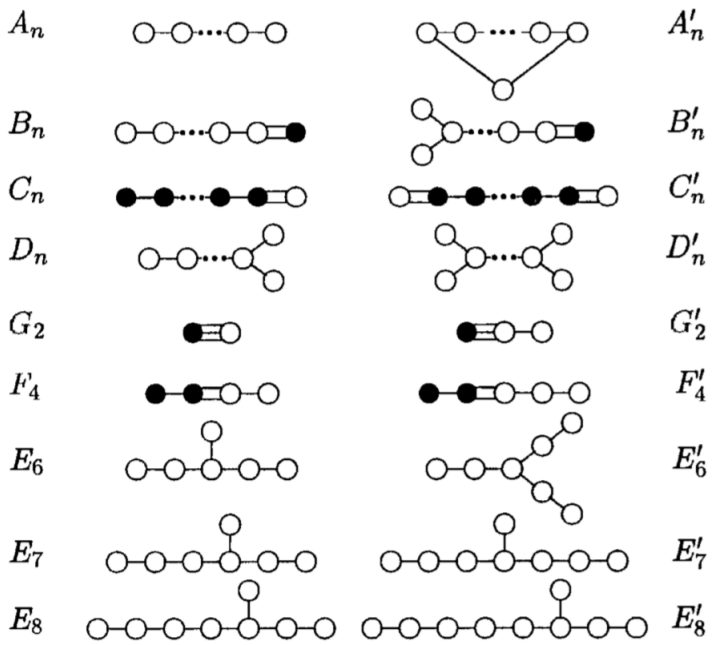
- A regular subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is a subalgebra whose Cartan generators are linear combinations of the Cartan generators of A. A regular maximal subalgebra is not strictly contained in any proper subalgebras of $\mathfrak{g}$.

- In physics, we care about regular subalgebras because if G is a symmetry group, the conserved charges are the Cartan generators. Then the conserved charges in regular subalgebras are built from the original ones, so we may reach them by spontaneous symmetry breaking.

- We can construct regular maximal subalgebras by deleting a root from a Dynkin diagram and replacing it with a $\mathfrak{u}(1)$ factor. For example, applying this to $SU(5)$ gives $SU(3)\times SU(2)\times U(1)\subset SU(5)$, as used in grand unification.

- To construct semisimple regular maximal subalgebras, we use a trick. Given simple roots α^i, we may add the lowest root α^0 to the Dynkin diagram. This root has appropriate angles with the others, because $\alpha^0-\alpha^i$ is not a root for any i. However, the augmented root system is linearly dependent.

- Working case by case, one can construct the extended Dynkin diagrams below.



Semisimple regular maximal subalgebras can then be constructed by removing a root from the extended Dynkin diagram. For example, we find $\mathfrak{so}(2n) \subset \mathfrak{so}(2n+1)$, which we observed earlier, and $\mathfrak{so}(2k) \oplus \mathfrak{so}(2n-2k) \subset \mathfrak{so}(2n)$.

7 More Representations

7.1 Group Theory Constants

In this section, we'll develop powerful calculation tools for representations. First we'll consider the "group theory constants" that often appear in computations. We refer to abstract generators as T^a and generators in a specific representation R as t^a_R with the R subscript suppressed.

- We assume we are working with a simple compact Lie group G, so that $\mathfrak{g}$ is of compact type. We can thus choose a basis T^a where the Killing form is $\kappa_{ij} = -\delta_{ij}$.
- Given this basis, we automatically have

  $$\operatorname{tr} t^a t^b = T(R)\delta^{ab}, \quad T(R) > 0$$

  for any representation R. To see this, define $M^{ab} = \operatorname{tr} t^a t^b$. Then

  $$[t^a_{\text{Ad}}, M]^{bc} = \operatorname{tr}\left((t^a_{\text{Ad}})^{bd} t^d t^c - t^b t^d (t^a_{\text{Ad}})^{dc}\right) = \operatorname{tr}\left(f^{abd} t^d t^c - f^{adc} t^b t^d\right).$$

  Using the definition of the structure constants, we have

  $$[t^a_{\text{Ad}}, M]^{bc} = \operatorname{tr}\left([t^a, t^b]t^c + [t^a, t^c]t^b\right) = 0$$

  which means M must be proportional to the identity by Schur's lemma, since the adjoint representation is irreducible.
- In physics, we usually pick the normalization,

  $$T(\text{fund}) = 1/2$$

  matching the existing convention for the spinor representation of $SU(2)$, $t^a = \sigma^a/2$. However, in mathematics it is more common to pick $T(\text{fund}) = 1$.
- The quadratic Casimir $T^a T^a$ commutes with everything else, so by Schur's lemma,

  $$t^a t^a = C_2(R)$$

  in any irrep R. By contracting indices in the definition of $T(R)$,

  $$(\dim R)\, C_2(R) = (\dim G)\, T(R).$$

  Alternatively, one may also take this equation to be the definition of $T(r)$.
- For $SU(N)$, in the physicist's normalization, we have

  $$C_2(\text{fund}) \equiv C_F = \frac{N^2 - 1}{2N}, \quad C_2(\text{adjoint}) \equiv C_A = T(\text{adj}) = N.$$

  The latter result can be found by a trick. For the adjoint representation, $(t^b)_{ac} = i f^{abc}$ and the structure constants are real, giving

  $$\operatorname{tr} t^c t^c = -\sum_{ab} f_{abc} f_{bac} = \sum_{ab} |f_{abc}|^2 = \sum_{ab} |t^c_{ab}|^2 = N$$

  since t^c is unitary.

- The Dynkin index obeys the rules

$$T(R) = T(\overline{R}), \quad T(R_1 \oplus R_2) = T(R_1) + T(R_2), \quad T(R_1 \otimes R_2) = d(R_1)T(R_2) + d(R_2)T(R_1).$$

  To prove the last identity using matrices, we note that

$$\operatorname{tr} t_R^a = 0$$

  for any representation of a semi-simple Lie algebra. This is because every generator can be written in terms of commutators of generators, and commutators have zero trace.

- The anomaly coefficient obeys the rules

$$A(R) = -A(\overline{R}), \quad A(R_1 \oplus R_2) = A(R_1) + A(R_2), \quad A(R_1 \otimes R_2) = A(R_1)d(R_2) + d(R_1)A(R_2).$$

  The first implies anomaly coefficients vanish unless the representation is complex.

Now we'll state some general facts about representations without proof.

- First, we summarize Lie algebra dimensions and the dimensions of the fundamental, or minimal representation F.

$\mathfrak{g}$	$\dim \mathfrak{g}$	$\dim F$
$\mathfrak{su}(N)$	N^2-1	N
$\mathfrak{so}(N)$	$N(N-1)/2$	N
$\mathfrak{sp}(N)$	$N(2N+1)$	$2N$
E_6	78	27
E_7	133	56
E_8	248	248
F_4	52	6
G_2	14	7

- For Lie algebra representations, characters are the analogue of the character (trace) for Lie group representations. Specifically, for a representation R and Cartan subalgebra element t, $\operatorname{ch}(R)(t) = \operatorname{tr}_R(\exp(t))$. They place similarly powerful constraints on representation theory, via the Weyl character formula.

- A useful special case of the Weyl character formula is the Weyl dimension formula: for a representation R of $\mathfrak{g}$ with highest weight Λ,

$$d(R) = \prod_{\alpha \in \Delta_+} \frac{(\Lambda + g, \alpha)}{(g, \alpha)}, \quad g = \frac{1}{2} \sum_{\alpha \in \Delta_+} \alpha$$

  where Δ_+ is the set of positive roots. In fact, one can show that g is equal to the sum of the fundamental weights, i.e. its Dynkin labels are $(1, 1, \ldots, 1)$. Explicit special cases of this formula are given here.

- In this notation, the quadratic Casimir is $C_2(r) = (\Lambda, \Lambda + 2g)$ in the mathematician's normalization; in the physicist's normalization there is a factor of $1/2$, as for the Dynkin index.

- The only Lie groups with complex representations are

$$SU(n) \text{ for } n \geq 3, \quad SO(4n+2) \text{ for } n \geq 2, \quad E_6.$$

  where we don't include $SO(6)$ because $\mathfrak{so}(6) \cong \mathfrak{su}(4)$. For the latter two, all anomaly coefficients vanish. Hence anomalies may only come from $U(1)$ factors and $SU(n)$ for $n \geq 3$.

- One way to see the SM is anomaly free is to note that its matter fields form a representation of $SO(10)$, whose representations are all anomaly free. The existence of right-handed neutrinos doesn't matter for this argument, since they don't contribute to any anomalies in the SM.

- Since left-handed and right-handed fermions transform in conjugate representations, a real GUT representation would yield them in pairs. This cannot match experiment because the SM is chiral, unless one chirality is much lighter, which is unnatural. Hence only the gauge groups $SU(n)$, $SO(4n+2)$, and E_6 are usable for GUTs. The appearance of E_6 isn't too surprising, because the most common GUT candidates are $SU(5) = E_4$ and $SO(10) = E_5$.

- For the representation R of $\mathfrak{su}(n)$ with Dynkin labels $(a_1, \ldots, a_{n-1})$, the anomaly coefficient is

$$A(R) = d(R) \sum_{i,j,k=0}^{n-1} a_{ijk}(a_i+1)(a_j+1)(a_k+1)$$

  where the tensor a_{ijk} is completely symmetric with

$$a_{ijk} = N_n i(n-2j)(n-k), \quad i \leq j \leq k$$

  where N_n is a conventional normalization factor, which we usually set by $A(\text{fund}) = 1$.

- We've encountered Casimir operators above and used them to classify irreps. In fact, in general Casimir operators give a complete classification; any rank n simple Lie algebra has n fundamental Casimir invariants. For $SU(n)$, their degrees are $2, 3, \ldots, n$.

- Similarly, one can treat the anomaly coefficient and Dynkin index above as the cubic and quadratic special cases of a higher-order Dynkin index. Higher orders are relevant for anomaly cancellation in higher-dimensional theories.

- Dynkin's theorem for the second highest representation is occasionally useful. We know that the product of irreps with highest weights Λ and Λ' contains an irrep with highest weight $\Lambda + \Lambda'$. We may get from Λ to Λ' or vice versa by adding simple roots; suppose $\Lambda' = \Lambda + \alpha_1 + \ldots + \alpha_k$ where the path is chosen so k is minimal. Then the product contains an irrep with highest weight $\Lambda + \Lambda' - \alpha_1 - \ldots - \alpha_k$.

Note. A proof of the formula for the Casimir invariant, $C = \kappa^{ab} T_a T_b = (\kappa^{-1})_{ab} T^a T^b$. We use the basis consisting of E^α and $H^{\alpha_{(i)}}$ for roots α and simple roots $\alpha_{(i)}$. Then

$$C = \sum_{ij} (\kappa^{-1})_{ij} H^{\alpha_{(i)}} H^{\alpha_{(j)}} + \sum_\alpha \frac{E^\alpha E^{-\alpha}}{\kappa(E^\alpha, E^{-\alpha})}.$$

Now act with both sides on a vector v with weight λ. If we restrict the sum in the second term to positive roots, we get anticommutators $\{E^\alpha, E^{-\alpha}\}$ for positive α. This is inconvenient, but if

we choose v to have the highest weight Λ the anticommutators may be replaced with commutators. Hence the second term is

$$\sum_\alpha \frac{E^\alpha E^{-\alpha}}{\kappa(E^\alpha, E^{-\alpha})} v = \sum_{\alpha\in\Delta_+} \frac{[E^\alpha, E^{-\alpha}]}{\kappa(E^\alpha, E^{-\alpha})} = \sum_{\alpha\in\Delta_+} H^\alpha v = (2g, \Lambda)v.$$

The first term is

$$\sum_{ij} (\kappa^{-1})_{ij} H^{\alpha_{(i)}} H^{\alpha_{(j)}} v = \sum_{ij} (\kappa^{-1})_{ij} (\alpha_{(i)}, \Lambda)(\alpha_{(j)}, \Lambda) v.$$

For a general v, this is quite complicated, but it simplifies if v has the highest weight Λ. At this point it is convenient to introduce the "inverse Cartan matrix" or "metric"

$$G_{ij} = (A^{-1})_{ij} \frac{(\alpha_{(j)}, \alpha_{(j)})}{2}$$

which confusingly isn't actually the inverse of the Cartan matrix. If we define

$$\tilde{A}_{ij} = (\alpha_{(i)}, \alpha_{(j)}), \quad D = \mathrm{diag}((\alpha_{(i)}, \alpha_{(i)})/2)$$

then we have the relations

$$A = \tilde{A} D^{-1}, \quad G = D \tilde{A}^{-1} D$$

so we see G is symmetric. It is convenient because

$$(w_{(i)}, w_{(j)}) = G_{ij}.$$

Expanding the highest weight as $\Lambda = a^i w_{(i)}$ and plugging these results in, the first term becomes

$$\sum_{ij} a^i G_{ij} a^j v = (\Lambda, \Lambda) v$$

where we used $\kappa_{ij} = \tilde{A}_{ij}$. Hence we have the desired result.

Example. Calculating the Casimir invariant for $\mathfrak{su}(3)$. It is convenient to expand everything in terms of fundamental weights, so the positive roots are

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

This simple result for g matches the expected general result. The inverse Cartan matrix is

$$G = \frac{1}{3} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Therefore, the Casimir invariant is

$$C = \frac{1}{3} \begin{pmatrix} a^1 & a^2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a^1 + 2 \\ a^2 + 2 \end{pmatrix} = \frac{2}{3}(a_1^2 + a_2^2 + a_1 a_2 + 3a_1 + 3a_2).$$

In this normalization the Casimir invariant of the fundamental (10) is $8/3$, so we must divide by 2 to get the physical normalization. While we're at it, it's not too much more work to get the Weyl dimension formula,

$$d(a_1 a_2) = (1 + a_1)(1 + a_2)\left(1 + \frac{a_1 + a_2}{2}\right).$$

7.2 The Lorentz and Galilean Groups

Inonu–Wigner contraction can be used to generate new Lie algebras from old, and take physically relevant limits. Our main application of this technique will be the nonrelativistic limit, where it explains some puzzling features of Galilean invariance.

- Consider the rotation algebra $\mathfrak{so}(3)$ and let $z = L\zeta$, so $\partial_z = L^{-1}\partial_\zeta$. Then if we take L to infinity while keeping ζ of order one, to lowest order in L we have

$$J_z = -i(x\partial_y - y\partial_x), \quad J_x = -i(y\partial_z - z\partial_y) \to i(L\zeta)\partial_y, \quad J_y \to -i(L\zeta)\partial_x.$$

  Thus, after some rescaling J_x and J_y yield translations P_y and P_x, where $P_i = i\partial_i$.

- Intuitively, imagine the original vector fields as describing symmetries of the sphere. Then taking z to be large corresponds to "reverting to the flat Earth", by zooming in on the flat patch near the North pole. The resulting symmetry group is $ISO(2)$ or $E(2)$, containing symmetries of two-dimensional Euclidean space.

- Similar reasoning works for $\mathfrak{so}(n)$. Letting a Latin index denote a value from 1 to $n-1$, the J_{ij} are unaffected while J_{in} contracts to yield P_i, giving $E(n-1)$.

- The same logic applies for the Lorentz algebra $\mathfrak{so}(3,1)$. We start with the generators

$$J_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu), \quad [J_{\mu\nu}, J_{\rho\sigma}] = -i(\eta_{\mu\rho}J_{\nu\sigma} + \eta_{\nu\sigma}J_{\mu\rho} - (\rho \leftrightarrow \sigma)).$$

  We contract by setting $x_0 = ct$ with $c \to \infty$, $t = O(1)$. Then J_{ij} is unmodified and we have

$$J_{0i} \to ctP_i.$$

  By treating t as a constant, just as we treated ζ, the boosts reduce to translations P_i. We hence recover $E(3)$, up to some possible signs which may have to be scaled away.

- Similarly, one can see how the de Sitter symmetry group $SO(4,1)$ would contract to the Poincare group; the $SO(3,1)$ subgroup gives the Lorentz group, while the remaining four operators give spacetime translations. Here it is important that we contract away a spatial dimension, since contracting away the temporal dimension would instead yield $E(4)$.

Note that we didn't get the full Galilean group above, because we treated t as a constant. By maintaining time-dependence, we can contract the Poincare group to the full Galilean group.

- We begin with the Poincare commutation relations in the $3+1$ split. By defining

$$J_{ij} = \epsilon^{ijk}J_k, \quad J_{i0} = K_i$$

  the Lorentz algebra is

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

  We see that $\mathbf{J}$ and $\mathbf{K}$ are both vectors. The minus sign in the last commutation relation indicates the Lorentz group is noncompact. One way to see that $\mathbf{J}$ must appear there is to note that $\mathbf{J}$ is a pseudovector and $\mathbf{K}$ is a vector, so $\mathbf{J}$ is the only result that obeys parity.

- We split the translations by defining $H = P_0 = i\partial_0$, to give

$$[J_i, P_j] = i\epsilon_{ijk}P_k, \quad [K_i, P_j] = -iH\delta_{ij}, \quad [J_i, H] = [P_i, H] = 0, \quad [K_i, H] = -iP_i.$$

  These new results are intuitive: we see that translations are vectors, and the difference between boosting now and later is a translation.

- As before, we define $x_0 = ct$ and switch variables to t. To remove factors of c from the generators, we rescale $K_i \to K_i/c$ and $H \to cH$. The commutation relations that change are

$$[K_i, K_j] = -\frac{1}{c}\, i\epsilon_{ijk}K_k, \quad [K_i, P_j] = -i\frac{H}{c^2}\delta_{ij}.$$

  The first equation indicates that Galilean boosts commute. However, the second is more subtle, because the energy takes the form $H = Mc^2 + T$. The huge relativistic contribution to the energy allows the right-hand side to survive, giving

$$[K_i, K_j] = 0, \quad [K_i, P_j] = -iM\delta_{ij}, \quad [K_i, T] = -iP_i.$$

  All other commutation relations remain the same.

- Thus, if we don't count the c-number M as a member of the Galilean algebra, then M is a central charge, indicating that we must use projective representations of the Galilean group in nonrelativistic quantum mechanics. There is a superselection rule forbidden superposition of states with different masses, though this may be removed by enlarging the Galilean algebra with the generator M, yielding the Bargmann algebra.

- In terms of differential operators, we could also have computed this explicitly as

$$K_i \to -it\partial_i - Mx_i$$

  which immediately yields $[K_i, P_j] = -iM\delta_{ij}$. Another check is that

$$[K_i, P^2/2M] = -\frac{1}{2M}(2iM\delta_{ij})P_j = -iP_i = [K_i, T]$$

  which confirms $T = P^2/2M$ up to a c-number.

- Also note that there's a pesky sign here: the physical momentum is $P^i = -P_i = -i\partial_i$. The easiest way to see this is in quantum mechanics, where a plane wave $e^{ipx/\hbar}$ has momentum p.

These results provide some insight into nonrelativistic classical and quantum mechanics.

- In Lagrangian mechanics, the Galilean algebra can be represented by differential operators on the configuration space. As usual, lifting this action to quantum mechanics produces central extensions because we allow projective representations. The same result occurs if we go to Hamiltonian mechanics: the results are identical but with Poisson brackets instead of commutators.

- Under an active Galilean boost of velocity u,

$$p' = p + Mu, \quad E' = E + up + \frac{1}{2}Mu^2.$$

On the other hand, the wavenumber and frequency of a plane wave transform as

$$k' = k, \quad \omega' = \omega + uk$$

where the uk term is from the Doppler shift, assuming the phase is Galilean invariant. Naively, these results are not compatible with the de Broglie relations $p = \hbar k$ and $E = \hbar\omega$.

- The resolution is that Galilean invariance only requires that $|\psi'(x', t')|^2 = |\psi(x, t)|^2$, and allows the addition of an arbitrary phase. To find this phase, let the original wavefunction be a plane wave $\psi(x, t) = e^{i(px-Et)/\hbar}$. Then the transformed wavefunction must be

$$\psi'(x, t) = e^{i(p'x-E't)/\hbar} = e^{i(Mux-upt-Mu^2t/2)/\hbar}\psi(x, t) = e^{i(Mux-Mu^2t/2)/\hbar}\psi(x - ut, t).$$

For an infinitesimal boost, we neglect the u^2 term and find

$$\psi'(x, t) = (1 - iuK)\psi(x, t), \quad K = -it\partial_x - Mx.$$

The first term in K is the obvious one. Before relativity, the extra Mx term was mysterious, but in hindsight we understand it as the vestige of the $E = Mc^2$ rest energy after group contraction.

- We may also see this through the Lagrangian. The standard action

$$S = \int \frac{1}{2}mv^2\, dt$$

is not Galilean invariant, but rather changes by a total derivative,

$$L \to L + \frac{d}{dt}\left(mux + \frac{1}{2}mu^2t\right).$$

This is acceptable classically, because it leaves the equations of motion invariant, but seems a little strange because the same caveat does not apply to the relativistic action. And as we've seen above, it becomes an issue in the quantum theory: phase is Lorentz invariant but not Galilean invariant.

- We can understand the issue by considering the nonrelativistic limit. Relativistically,

$$S = -mc^2 \int \sqrt{1 - v^2/c^2}\, dt = \int \left(-mc^2 + \frac{1}{2}mv^2 + \dots\right) dt.$$

We then drop the first term, corresponding to rest energy, to get to the nonrelativistic Lagrangian. But this term has a nontrivial effect even in the limit $c \to \infty$, because the tininess of the time dilation effect is canceled by the size of the rest energy. Specifically, the change in this term is

$$mc^2(t - t') = mc^2\gamma\left(t' + \frac{ux'}{c^2}\right) - mc^2t' = mux' + \frac{1}{2}mu^2t' + \dots$$

which is precisely the extra phase picked up.

- Yet another way to see this is to consider the process by which we move from the Klein–Gordan equation to the Schrodinger equation. In the standard account, we factor out a rapidly oscillating phase e^{-imc^2t} in the wavefunction. This phase is precisely what is missing.

For completeness, we'll briefly overview the representations of the Lorentz group.

- As expected from Dynkin diagrams, we have two independent copies of $\mathfrak{su}(2)$ by defining

$$M_i = \frac{J_i + iK_i}{2}, \quad N_i = \frac{J_i - iK_i}{2}.$$

  The two are swapped by parity. This allows us to easily find the finite-dimensional representations of the Lorentz group from those of $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$, though they won't be unitary.

- Such nonunitary representations are extremely important, since many physical observables transform in them, including positions, momenta, and field values. Nonunitarity is acceptable because these objects do not correspond directly to quantum states. Later, we will show in detail how quantum states in an infinite-dimensional, unitary Poincare irrep correspond to quantum fields in a finite-dimensional, nonunitary Lorentz irrep.

- Restricting to $SO(3)$, each Lorentz irrep breaks into $SO(3)$ irreps attained by adding spins u and v. Hence we can parametrize irreps by the minimum/maximum spin

$$j_0 = |u - v|, \quad j_1 = u + v$$

  and vectors within them by the usual spin and z-component, j and m.

- There are also infinite-dimensional unitary representations. Acting with the J_i does not change the value of j, but acting with the K_i does. Since the K_i form a vector, we have a set of irreducible tensor operators $\{K_-/\sqrt{2}, K_3, -K_+/\sqrt{2}\}$, so by the Wigner–Eckart theorem

$$\langle j'm'|K_3|jm\rangle = A_j^{j'} \langle j'm'|10jm\rangle, \quad \langle j'm'|K_\pm|jm\rangle = \mp\sqrt{2} A_j^{j'} \langle j'm'|1 \pm 1jm\rangle.$$

  The Clebsch–Gordan coefficient is only nonzero when $|j' - j| \leq 1$, so only a few of the $A_j^{j'}$ coefficients matter. This gives a constraint on the matrix elements of K_i in any representation.

- One can then show the general solution is parametrized by a complex number ν, and imposing unitarity gives two classes of irreps,

$$\text{principal series: } \nu = -iw, j_0 = 0, 1/2, 1, \ldots, \quad \text{complementary series: } -1 \leq \nu \leq 1, j_0 = 0$$

  where w is real and j_1 is infinite.

- The finite-dimensional irreps can be labeled by the quantum numbers u and v, where the Casimir operators M^2 and N^2 have values $u(u+1)$ and $v(v+1)$. For example, we have

$$\text{scalar: } (0,0), \quad \text{four-vector: } (1/2,1/2), \quad \text{Weyl spinors: } (1/2,0), (0,1/2).$$

- A rank two tensor transforms in $(1/2,1/2) \otimes (1/2,1/2)$ which decomposes as

$$\text{symmetric: } (1,1), \quad \text{trace: } (0,0), \quad \text{antisymmetric: } (1,0), (0,1).$$

  Note that the antisymmetric electromagnetic field $F_{\mu\nu}$ decomposes under spatial $SO(3)$ into two vectors, which can be arbitrary linear combinations of $\mathbf{E}$ and $\mathbf{B}$. Under the full Lorentz group, the two irreps must be the self-dual and anti self-dual fields $\mathbf{E} \pm i\mathbf{B}$ corresponding to clockwise/counterclockwise circularly polarized waves. In fact, this is precisely the decomposition we applied to $J_{\mu\nu}$ to get the two $\mathfrak{su}(2)$ subalgebras in the first place.

- We can also phrase this by using $\mathfrak{so}(3,1) \cong \mathfrak{sl}(2,\mathbb{C})$. Temporarily switching to mathematics conventions, so that the commutators have no factors of i, the fundamental representation is

$$d(J_i) = -\frac{i}{2}\sigma_i, \quad d(K_i) = -\frac{1}{2}\sigma_i$$

which corresponds to

$$d(M_i) = -\frac{i}{2}\sigma_i, \quad d(N_i) = 0.$$

This is the left-handed Weyl spinor representation $(1/2, 0)$.

- Similarly, the conjugate of the fundamental representation has

$$d(M_i) = 0, \quad d(N_i) = -\frac{i}{2}\sigma_i$$

and is the right-handed Weyl spinor.

- Finally, consider the four-vector representation $(1/2, 1/2)$. Recalling that $(1/2, 1/2)$ stands for a tensor product, we have

$$d(M_i) = -\frac{i}{2}\sigma_i \otimes I_2, \quad d(N_i) = I_2 \otimes \left(-\frac{i}{2}\sigma_i\right).$$

This is the correct result, though not in the familiar four-vector basis.

7.3 Tensor Methods

We'll now introduce tensor methods for finding finite-dimensional irreps of the classical groups, starting with the simplest example of $GL(n,\mathbb{C})$.

- For brevity, let $g \in \mathrm{Mat}(n,\mathbb{C})$ denote the representation of a group element $g \in GL(n,\mathbb{C})$ in the fundamental representation. Then there are four closely related irreps of $GL(n,\mathbb{C})$,

$$\{g\}, \quad \{g^*\}, \quad \{{g^{-1}}^T\}, \quad \{{g^{-1}}^\dagger\}.$$

These are called the fundamental representations of $GL(n,\mathbb{C})$. They are inequivalent, as

$$g^* = SgS^{-1}$$

cannot hold for $g = \alpha I$ for complex α, with similar logic for the others.

- For simplicity, we focus on $\{g\}$ and the dual/contragradient representation $\{{g^{-1}}^T\}$, which act on vector spaces V and $\tilde{V}$ respectively. Vectors in these spaces transform as

$$x^a \to x'^a = g^a{}_b x^b, \quad y_a \to y'_a = y_b ({g^{-1}})^b{}_a.$$

Here, the transpose swaps the order of the indices on g but maintains their vertical positions; we will always denote the fundamental representation with an upper index and the dual representation with a lower index; note that the contraction $x^a y_a$ is invariant.

- For the other representations, we would need more types of indices to keep our notation explicit. For example, we may use dotted indices,

$$z_{\dot{a}} \to z'_{\dot{a}} = (g^*)_{\dot{a}}{}^{\dot{b}} z_{\dot{b}} = z_{\dot{b}} (g^\dagger)^{\dot{b}}{}_{\dot{a}}, \quad w^{\dot{a}} \to w'^{\dot{a}} = (g^{-1\dagger})^{\dot{a}}{}_{\dot{b}} w^{\dot{b}}$$

  for $\{g^*\}$ and $\{g^{-1\dagger}\}$ respectively. The contraction $w^{\dot{a}} z_{\dot{a}}$ is invariant. For simplicity, we'll ignore these representations and return to them later.

- We define the tensor product space of tensors of type (i, j) by

$$T^i_j = V^{\otimes i} \otimes \tilde{V}^{\otimes j}.$$

  We can define a map $T^i_j \to T^{i-1}_{j-1}$ by contracting any pair of upper and lower indices.

- Specializing to T^1_1, the tensor $\delta^a{}_b$ is invariant; its components stay the same under any $GL(n, \mathbb{C})$ transformation. One can show this the only elementary invariant tensor; all other invariant tensors can be built out of it, e.g. by tensor product.

- Using the invariant tensor $\delta^a{}_b$, we can decompose T^1_1 into irreps as

$$t^a{}_b = \hat{t}^a{}_b + \delta^a{}_b \frac{t^c{}_c}{n}$$

  where $\hat{t}^a{}_b$ is the traceless part of $t^a{}_b$. We can think of the second term as the image of $t^a{}_b$ under the projection $\delta^b{}_a \delta^{a'}{}_{b'}/n$.

- Generally, let I be an arbitrary invariant tensor that is also a projection operator; that is, I is a tensor of type (i, i) which acts by contracting half of its indices, and $I^2 = I$. Then for any tensor t we can write

$$t = I(t) + (1 - I)(t)$$

  and we claim that if we repeat this process on each term until it terminates, we will have the maximal decomposition, i.e. t will be decomposed into irreps.

- To see this, let P be a projection operator onto an irrep, so

$$gPt = Pgt.$$

  Since t is general, this is equivalent to $P = gPg^{-1}$, so P is an invariant tensor. So all projection operators will be accounted for in our decomposition.

- Next, we consider T^2_0. Neither of the naive candidates $\delta^{a'}_a \delta^{b'}_b$ and $\delta^{b'}_a \delta^{a'}_b$ are projection operators. We could find projection operators by brute force, but a better way is to recall that the projection operators in the symmetric group S_2 are the symmetrizer and antisymmetrizer, and S_2 acts on T^2_0, commuting with the action of $GL(n, \mathbb{C})$. Thus these projection operators indeed produce $GL(n, \mathbb{C})$ irreps, explicitly

$$t^{ab} = t^{(ab)} + t^{[ab]}.$$

  This is a full decomposition, as there is no trace to remove.

- Now consider the general case T^i_j. We take a Young tableau λ from S_i for the upper indices and a Young tableau σ from S_j for the lower indices. Performing the resulting projection maps $t^{\{a\}}_{\{b\}}$ to $t^{\lambda\{a\}}_{\sigma\{b\}}$, where we use multi-index notation. Now we haven't taken care of the traces yet, because they deal with contractions between upper and lower indices, while we've only treated these sets separately.

- To remove the traces in practice, we work recursively. For T^2_2, we may write

$$t^{ac}_{bd} = r^{ac}_{bd} + \delta^a_b (s_1)^c_d + \delta^a_d (s_2)^c_b + \delta^c_b (s_3)^a_d + \delta^c_d (s_4)^a_b + \delta^a_b \delta^c_d u_1 + \delta^a_d \delta^c_d u_2$$

  where r, the s_i, and the u_i are all traceless. We determine the u_i by contracting both sides with two delta functions, giving

$$n(nu_1 + u_2) = t^{ac}_{ac}, \quad n(u_1 + nu_2) t^{ac}_{ca}.$$

  Next, we contract both sides of the original equation with one delta function, determining the s_i. This procedure ensures we don't 'double count' any traces. It works independently of the Young tableaux, which simply set some of the s_i and u_i equal to each other.

- There are no more available invariant tensors, so we've arrived at an irrep. One can show that the irrep associated with Young tableaux (λ, σ) is equivalent to that of (λ', σ') if and only if λ and λ' have the same Young diagram, and σ and σ' have the same Young diagram.

- Finally, restoring the $\{g^*\}$ and $\{g^{-1\dagger}\}$ representations, the inequivalent irreps of $GL(n, \mathbb{C})$ generated by the four fundamental representations are indexed by four Young diagrams; they are fully traceless amongst the dotted indices and amongst the undotted indices.

Note. In physics, every Lie group is regarded as a subgroup of $GL(n, \mathbb{C})$ for some n, and we have a fundamental, n-dimensional representation by taking matrices, called the defining representation. We focus on this representation and its relatives here, and generally there are 4, 2, or 1. They are a subset of the mathematician's fundamental representations. In most but not all cases we'll consider, the "physicist's fundamentals" are sufficient to build all representations.

The analysis for other classical linear groups is similar, but with more invariant tensors.

Example. The unitary group $U(n)$. In this case, the four fundamental representations collapse to just two, $\{g\}$ and $\{g^*\}$, which are usually called the fundamental and antifundamental. Another way of seeing this is that we have a new invariant tensor $\delta^a_{\dot{b}}$, since

$$g^a{}_b \delta^b_{\dot{d}} (g^\dagger)^{\dot{d}}{}_{\dot{c}} = \delta^a_{\dot{c}}$$

along with $\delta^{\dot{a}}_b$. These invariant tensors relate dotted and undotted indices; hence we can remove all dots before performing the usual procedure. In practical terms, this is like how we can raise indices using the metric tensor in relativity, so we call these new invariant tensors metrics. The results for $GL(n, \mathbb{C})$ go through unchanged, except that we eliminate the dotted indices.

Example. We may also consider indefinite metrics. Suppose that the matrix of $\delta^{\dot{a}}_b$ has n_+ positive eigenvalues and n_- negative eigenvalues. Then demanding it is an invariant tensor yields the unitary group $U(n_+, n_-)$, and the finite-dimensional irreps are just the same as that of $U(n_+ + n_-)$.

However, making the metric indefinite does have important effects. The resulting group is not compact, so there are no finite-dimensional unitary representations. On the other hand, for definite metric, the group is compact, so there are no infinite-dimensional representations.

Example. The special linear group $SL(n, \mathbb{C})$. In this case we pick up another invariant tensor, the Levi–Civita tensor. Specifically, we have four Levi–Civitas with n indices, which are all upper/lower and dotted/undotted. By itself, the Levi–Civita doesn't further decompose the irreps, but it does set irreps equal; for example, the fully antisymmetric T_0^n irrep is now equivalent to the trivial irrep.

In general we can replace m antisymmetric upper/lower indices with $n - m$ antisymmetric lower/upper indices. For a Young diagram λ with column lengths m_i, define $\tilde{\lambda}$ to have column lengths $n - m_i$. Then $(\lambda, \mu; \tau, \kappa)$ is equivalent to $(\lambda', \mu'; \tau', \kappa')$ if the Young diagrams obtained by combining the columns of λ and $\tilde{\mu}$, and of τ and $\tilde{\kappa}$, are the same as their primed counterparts.

Example. In the special case $SU(m_+, m_-)$, the same final statement is true, except that we only have undotted indices.

Example. In the more special case $SL(2, \mathbb{C})$, the Levi–Civita ϵ^{ab} is more useful. Since it relates one upper index to one lower index, the fundamental representations collapse to just two, $\{g\}$ and $\{g^*\}$, and all indices can be taken to be upper indices. Moreover, by the logic above, we can replace any two antisymmetric indices with nothing. Then the irreps are indexed by two integers (i, k) and consist of tensors with i symmetric undotted indices and k symmetric dotted indices. This is as one would expect, since $SL(2, \mathbb{C})$ is the double cover of $SO(3, 1)$.

Example. Next, we consider some even more special cases.

- In the case $SU(2)$, two of the above simplifications combine, and we only have a single type of index. The irreps are indexed by a single integer.
- In the case $SU(3)$, we only have undotted indices, and we can replace any two antisymmetric upper/lower indices with a single lower/upper index. The irreps are indexed by two integers (i, j) and are traceless tensors with i symmetric upper indices and j symmetric lower indices.
- In the case $SU(4)$, there are three fundamental representations: 4, $\overline{4}$, and 6. However, the physicists' fundamentals are sufficient because the 6 is the antisymmetric part of 4×4. However, $\mathfrak{su}(4) \cong \mathfrak{so}(6)$, and in that context the 4 and $\overline{4}$ are spinors that can't be built out of the fundamental 6. In that case, the physicists' fundamentals are not sufficient.

Example. When we restrict to $GL(n, \mathbb{R})$, we lose all dotted indices, retaining only $\{g\}$ and $\{g^{-1^T}\}$. Now in $O(n_+, n_-)$, the new invariant tensor relates upper and lower indices; it is a metric tensor ξ^{ab} in the usual sense, and we can raise all indices, so there is only one fundamental representation. We can also use the metric tensor to remove traces between two upper indices. Thus the irreps are traceless tensors with upper indices and definite symmetry. The invariant tensor δ_b^a is rendered obsolete since there are no lower indices; alternatively, it can be built out of the metric, $\xi^{ab}\xi_{bc} = \delta_c^a$.

The addition of the metric splits up $GL(n, \mathbb{R})$ irreps. For example, t_b^a is a $GL(n, \mathbb{R})$ irrep, but we can raise the index with the metric, then split t^{ab} into symmetric and antisymmetric parts.

Example. The case of $O(2)$. Consider the first two indices in $t^{ab\cdots}$. We can take out the antisymmetric part, and the trace using the metric. But these two irreps are not equivalent, as they differ by how they transform under reflection; they are the scalar and pseudoscalar.

By repeating this procedure, we find rank n symmetric traceless tensors, for all $n > 1$. Imagine starting with a symmetric rank n tensor. Then the trace is a symmetric rank $n - 2$ tensor, so the symmetric traceless rank n tensor has

$$(n + 1) - (n - 1) = 2$$

degrees of freedom. All of these two-dimensional irreps are inequivalent, corresponding to 'angular momentum' eigenvalue n.

Example. The case of $SO(n)$. Since there's only one kind of index, the Levi–Civita can be used to break apart irreps, specifically those from tensors of rank $n/2$, into self-dual and anti self-dual components. We'll consider a few low-dimensional cases.

- For $SO(2)$, the one-dimensional irreps of $O(2)$ become equivalent, since one can multiply by ϵ^a_b where an index is lowered using the metric. The two-dimensional irreps decompose, by 'taking out the trace' using ϵ^a_b.
- For $SO(3)$, we can symmetrize all the indices by the same logic as for $SU(3)$. By similar counting as for $O(2)$, we confirm that the symmetric traceless tensors indeed have the expected integer dimensions. Irreps of $O(3)$ do not break apart, but do become equivalent, identifying the pseudovector and vector, and pseudoscalar and scalar.
- For $SO(3,1)$, the situation is similar to $SO(2)$, in that irreps of $O(3,1)$ do break apart; for example, $F^{\mu\nu}$ decomposes into $\mathbf{E} \pm i\mathbf{B}$. These are self-dual and anti self-dual components. This happens in general for rank n tensors in $SO(p,q)$ with $p+q=2n$.
- For $SO(8)$, the Lie algebra $\mathfrak{so}(8)$ has threefold symmetry, called triality, so there are three fundamental representations of dimension 8. Only one is a physicist's fundamental; the other two are spinors and cannot be built from the first.

7.4 ○ Representations of $SU(n)$

In this section, we give some practical applications of tensor methods, focusing on $SU(n)$.

Note. We can think of tensors by themselves, but in the context of quantum mechanics, they can be interpreted as wavefunctions. For example, consider the state

$$|\psi\rangle = \psi^i |i\rangle$$

where the $|i\rangle$ transform in the fundamental representation of $SU(3)$, say

$$|i\rangle \to U^j_i |j\rangle.$$

Then we have

$$|\psi\rangle \to U^j_i \psi^i |j\rangle = U^i_j \psi^j |i\rangle$$

so we may alternatively transform the wavefunction as

$$\psi^i \to U^i_j \psi^j$$

while keeping the basis vectors $|i\rangle$ fixed. We use upper and lower indices to distinguish the fundamental and antifundamental. Note that the transformations for the states and wavefunction differ by the usual active/passive flip. Now for two particles,

$$|\psi\rangle = \psi^{ij} |ij\rangle$$

the wavefunction ψ^{ij} transforms in a tensor product representation, and we can decompose it using the methods above. The bra transforms in the contragradient representation, which is just the

conjugate representation here; thus upstairs and downstairs indices are switched. A inner product $\langle\phi|\psi\rangle$ can be computed by contracting all the indices of the respective tensors; it is zero when the indices don't match up, reflecting the fact that different irreps are orthogonal. Similarly, tensor operators can be expanded in terms of wavefunctions, e.g. $\hat{A} = A^i\hat{O}_i$.

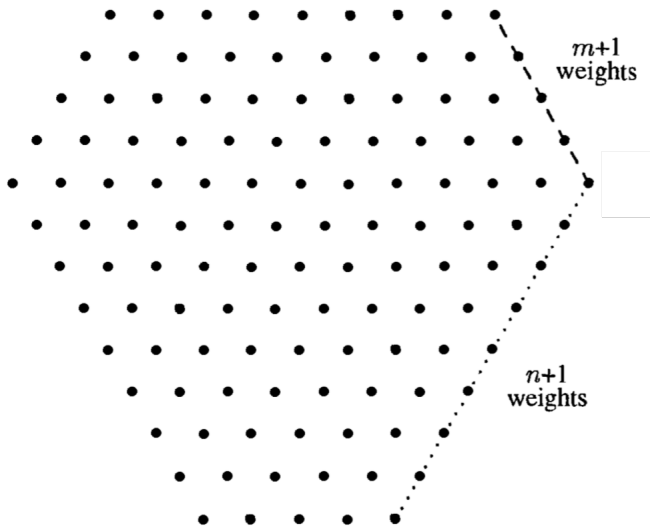
Note. Note that the $(1,0)$ irrep of $SU(3)$ corresponds to a $(1,0)$ tensor, while the $(0,1)$ irrep corresponds to a $(0,1)$ tensor. We also know that the only new irrep that appears in $(1,0)^{\otimes n} \otimes (0,1)^{\otimes m}$ is the (n,m) irrep. But in the tensor picture, the only new irrep is the symmetric traceless tensor with m upper indices and n lower indices. Hence the two labeling schemes of Dynkin labels and tensor ranks coincide, and this reasoning generalizes.

Note. Tensors explain the 'triality' symmetry of the $SU(3)$ representations. For example, in

$$3 \times 8 = 15 + \overline{6} + 3, \quad (1,0) \times (1,1) = (2,1) + (0,2) + (1,0)$$

all of the (n,m) irreps on the right-hand side have the same value of $(n-m)$ mod 3. This holds because all of the invariant tensors in the problem have type (n,m) where $n-m$ is divisible by 3.

Example. The weights of the (n,m) irrep of $SU(3)$ are shown below.



We start at the top-right and initially can move leftward or down-right. Then the states on the upper and upper-right edges are nondegenerate, so all the outer states are nondegenerate by Weyl reflection. In general, the degeneracy increases by one every time we move in a layer, until we reach a triangular layer, at which point it remains constant.

Each of the states shown here can be associated with a tensor, i.e. viewing the tensor as the state's wavefunction. For example, when we move along the path shown, the tensors change as

$$v_{3333}^{11111111} \to v_{2222}^{11111111} \to v_{2222}^{33333333}.$$

Now consider the states one layer inward. These contain tensors with one index in common between the top and bottom. To count the number of states in this layer, we find the number of symmetric tensors of this form (where the symmetry allows us to neglect index order) and subtract the number of traces, arriving at a degeneracy of 2. The same logic holds for all layers. The casework changes when we hit a triangular layer, because from that point on one of m or n is zero.

Note. Suppose we're consider the matrix elements $\langle u|W|v\rangle$, where W, $\langle u|$, and $|v\rangle$ transform in irreps R_W, R_u, and R_v. By the Wigner–Eckart theorem, all such matrix elements are specified by only a few numbers, specifically the number of factors of 1 in $R_W \otimes R_u \otimes R_v$. We've seen examples of this in $SU(2)$, where we found the other matrix elements by raising and lowering, but with tensors we can easily find a general expression by contracting indices.

Example. Suppose everything transforms in the 8 of $SU(3)$. There are two ways to contract the tensors W^i_j, $\overline{u}^i_j$, and v^i_j to yield a 1, either 'clockwise' or 'anticlockwise'. (Note that by convention u stands for the wavefunction of $|u\rangle$, so $\overline{u}$ is the bra wavefunction.) In this case all of these tensors can be written as matrices, giving

$$\langle u|W|v\rangle = \lambda_1 \operatorname{tr} u^\dagger W v + \lambda_2 \operatorname{tr} u^\dagger v W.$$

Here, we replaced $\overline{u}$ with the matrix $u^\dagger$ because taking the bra conjugates and exchanges upper and lower indices, and the latter is just a transposition in matrix notation.

Example. Suppose u and v are 10's and the W is an 8. Then there is only one contraction,

$$\langle u|W|v\rangle = \lambda \overline{u}_{ijk} W^k_\ell v^{ij\ell}.$$

This is more economical in $SU(3)$ than $SU(2)$, because the tensors tend to have lower ranks; the dimension of the irrep grows quadratically in the rank rather than linearly. However, such methods also work for $SU(2)$, where they can supply explicit formulas for the Clebsch–Gordan coefficients.

Next, we apply Young tableaux to tensor products in $SU(n)$. We use the physics conventions detailed in section 2.6 and warm up with $SU(3)$.

- As we've seen, the (n, m) irrep has n upper indices and m lower indices, where everything is traceless and symmetric. To describe the irrep with a single Young tableau, we convert every lower index into two antisymmetric upper indices by raising with ϵ^{ijk}.
- Given a tensor $T^{ijk\cdots}$, we apply a Young tableau to it by labeling the boxes with the indices, symmetrizing over indices in rows, then antisymmetrizing over indices in columns.
- For example, we project out the $(1, 1)$ irrep of $T^{ik\ell}$ corresponding to $S^i_j \epsilon^{jk\ell}$ by applying

  $$\begin{array}{|c|c|}\hline k & i \\ \hline \ell \\ \cline{1-1}\end{array}.$$

  To see this, note there are two constraints: the k and ℓ indices must be antisymmetric, enforced by the columns, and the original tensor must have been traceless, which is equivalent to

  $$\epsilon_{ik\ell} u^{ik\ell} = 0$$

  and enforced by the rows. Similarly, the (n, m) irrep has m columns with two boxes each followed by n columns with 1 box each.
- In general, given a tensor with n upper indices, we can project out an irrep by applying a Young tableau with n boxes. Columns with more than 3 boxes automatically give zero. Columns with 3 boxes mean an ϵ^{ijk} factors out of the tensor, so they yield an isomorphic irrep to a Young tableau without them.

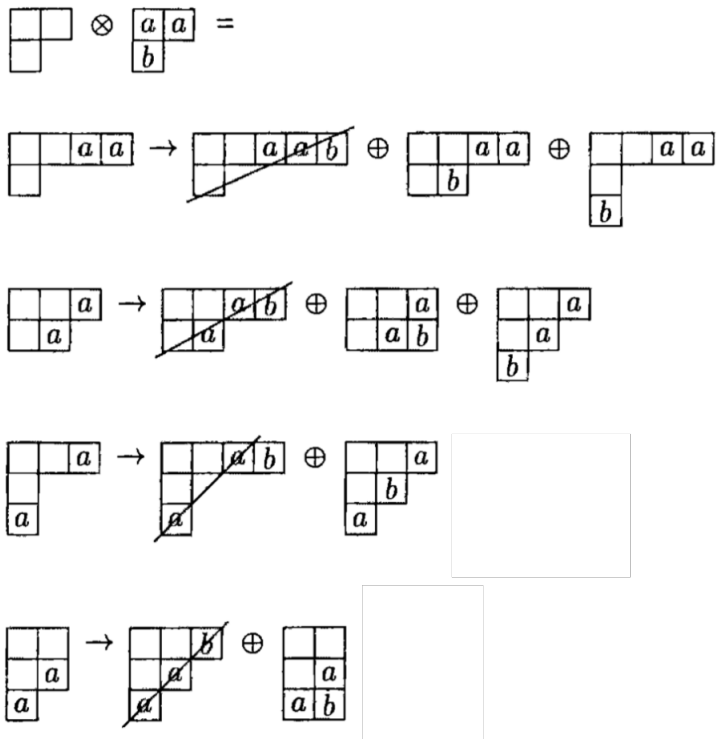
- There is a simple algorithm to compute the product of the irreps α and β corresponding to tableaux A and B.

  1. Write a's in the first row of B and b's in the second row.
  2. Add the a boxes to A anywhere, as long as a valid tableau is formed and no two a's are in the same column.
  3. Then add the b boxes to A similarly. Reading the boxes in Hebrew order and ignore the tableau if there are more total b's than a's at any point. (We could use a different order, but this is more convenient because we're attaching boxes on the right anyway.)

  We won't prove this, though intuitively it just takes all ways to combine the indices, maintaining the symmetry properties found above. In this system, triality is the number of boxes mod 3. The result can also be computed more directly using the Littlewood-Richardson rule.

- In general, the computation is much faster if we choose B to have as few boxes as possible. In particular, conjugating everything and then conjugating the final result can save time.

Example. Below we show that $8 \times 8 = 27 + 10 + \overline{10} + 8 + 8 + 1$.



Here we're ignored combinations that yield columns with more than three boxes, as these are simply zero. After ignoring columns with three boxes, we arrive at the result, recalling that 27 is $(2, 2)$ and 10 is $(3, 0)$.

Next, we move on to $SU(n)$.

- To do representation theory, we only need to relate between the simple roots and fundamental weights. Unfortunately, this is a bit complicated, because inverting the Cartan matrix isn't trivial. Instead, it's easier to instead express *both* of these quantities in terms of the weights of

the defining representation. (Note they are sometimes called the "weights of the fundamental", but one should not confuse them with the fundamental weights!)

- Let the weights of the defining representation be $\{\nu^1, \ldots, \nu^n\}$. As shown earlier, these form the vertices of a tetrahedron in $n-1$-dimensional space.

- The roots are simply the differences of these weights. For convenience, we define positive roots to have the form $\nu^i - \nu^j$ for $i < j$. Then the simple roots are

$$\alpha^i = \nu^i - \nu^{i+1}, \quad i = 1, 2, \ldots, n-1.$$

  We can easily check the angles are in accordance with the Dynkin diagram.

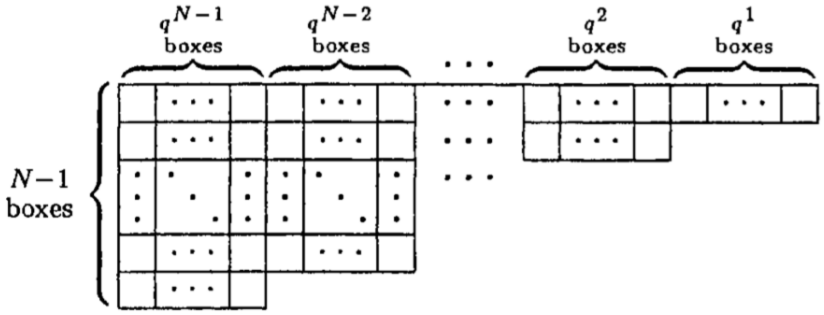
- Taking the dual basis, the fundamental weights are

$$\mu^i = \sum_{j=1}^{i} \nu^j, \quad i = 1, 2, \ldots, n-1.$$

  In particular, the highest weight of the defining representation is $\mu^1 = \nu^1$, as we can then lower with each simple root to get the other ν^i. Then the second-highest weight is ν^2, followed by ν^3.

- Now consider the antisymmetric tensor product of m copies of the defining representation. Then the highest weight is $\nu^1 + \ldots + \nu^m$ because the indices must be distinct, so we have the m^{th} fundamental representation! So just like in $SU(3)$, we can build everything out of tensor products of only the defining representation.

- General irreps of $SU(n)$ can thus be identified by a Young tableau. The tableau



  represents the irrep with Dynkin indices q^k. We specify a Young tableau with the notation $[\ell_1, \ell_2, \ldots]$ where ℓ_i is the length of column i, so the i^{th} fundamental is specified by $[i]$. The adjoint representation has one defining index and one lowered defining index, so it is $[n-1, 1]$.

- Consider the conjugate of the defining representation [1]. Its lowest weight is ν^n, but since all the ν^i sum to zero, this is equal to $-\mu^{n-1}$. Then the conjugate representation is $[n-1]$. Similarly, the conjugate of $[j]$ is $[n-j]$, and hence the conjugate of any Young tableau can be found by rotating it by 180° and interpreting the top edge as a new bottom edge.

- The algorithm for products above also works, with more letters. To avoid overcounting, there should be at least as many a's as b's, at least as many b's as c's, and so on, at every point. Triality generalizes to N-ality, the number of boxes mod N. This in turn generalizes to the "conjugacy class" of a representation of any Lie algebra, which can be useful for constraining tensor products.

- There is a useful formula for the dimension of an $SU(n)$ irrep, which can be derived from the Weyl character formula. We place an n in the top-left box, then place factors in the other boxes, adding one when we move right and subtracting one when we move down. The product of these factors divided by the product of the hook lengths is the dimension.

7.5 ○ Branching Rules

Finally, Young tableaux can be used to decompose representations when restricting to subgroups.

- First, we consider the subgroup

$$SU(n) \times SU(m) \times U(1) \subset SU(n+m).$$

  In terms of indices, we imagine the indices can go from 1 to $n+m$, the $SU(n)$ part acts on the indices from 1 to n only, and the $U(1)$ generator is $\text{diag}(m, \dots, m, -n, \dots, -n)$ to ensure tracelessness.

- The fundamental decomposes as

$$\square = \left(\square\,,\ \cdot\right)_M + \left(\cdot\,,\ \square\right)_N$$

  where the first element in each pair is the $SU(n)$ representation, and the subscript indicates the $U(1)$ charge.

- More generally, to decompose a general irrep C we consider the Young tableau. If it has k boxes, then we can only split into a pair of irreps A and B with n and m boxes so that $n+m=k$, and the $U(1)$ charge is $nM - mN$.

- Now, we need to account for the symmetry of C. Let A' and B' be $SU(n+m)$ irreps with the same tableau as A and B. Then if C doesn't appear in $A' \otimes B'$, then $A \otimes B$ surely can't appear in the decomposition of C. In fact, in general the number of times $A \otimes B$ appears in C is the number of times C appears in $A' \otimes B'$.

- Another important situation is the embedding

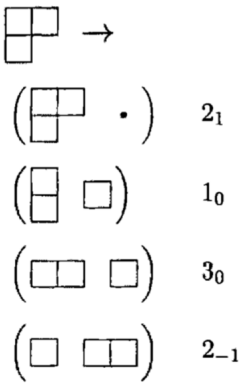
$$SU(n) \times SU(m) \subset SU(nm)$$

  which occurs when we work with tensor product spaces. The best notation here is to have two types of indices: an index i for $SU(n)$ and an index α for $SU(m)$, so that indices in $SU(nm)$ are composites $i\alpha$. Then each factor acts on its index, leaving the other alone.

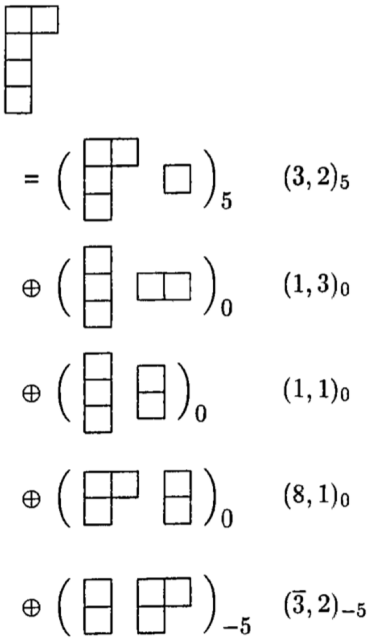
- Consider an irrep D of $SU(nm)$ with K boxes. Then its tensor has K indices of the $SU(n)$ type and K indices of the $SU(m)$ type, so it decomposes into (D_1, D_2) where D_1 and D_2 both have K boxes, up to eliminating full columns.

- The difference from the previous part is that the Young tableau for D describes the S_K symmetry associated with permuting the composite indices $i\alpha$. Thus, to see if we can recover this symmetry, we regard D_1 and D_2 as Young tableaux for S_K irreps, multiply them as S_K irreps, and look for a factor of D. We don't have an algorithm for this, because we only know how to multiply $SU(n)$ irreps diagrammatically.

Example. The case of $SU(3) \times SU(2) \times U(1) \subset SU(5)$. The adjoint decomposes as



Example. The case of $SU(2) \times U(1) \subset SU(3)$, where the '$SU(1)$' factor is trivial. If we're working with flavor $SU(3)$, these components are isospin and hypercharge. The adjoint decomposes as

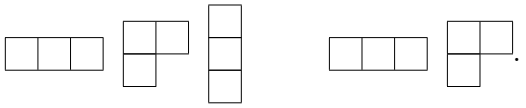


Here the trivial '$SU(1)$' irreps are all one-dimensional, so we don't mark them. However, the fact that $SU(1)$ has only one index value constrains the Young tableau to have a single row.

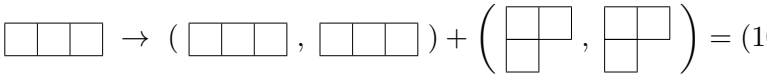
Example. The case of $SU(3) \times SU(2) \subset SU(6)$, where we interpret the factors as quark flavor and quark spin. For the lowest energy baryons, we need total symmetry between the quarks, so we must decompose

$$56 = \square\square\square.$$

There are three irreps of $SU(3)$ with three boxes, and two of $SU(2)$,



Then it turns out the two possibilities are



These are the baryon decuplet, which is totally symmetric in both flavor and spin space, and the baryon octet, which has mixed symmetry in both.

Note. Note that the adjoint A obeys the property

$$R \subset A \otimes R$$

for any nontrivial representation R. This is clear from the Young tableaux method for $SU(n)$. More generally, consider the map

$$A \otimes R \to R, \quad x \otimes v \to x(v).$$

It can be checked this is a map of representations. Since the image is R, R is a quotient of $A \otimes R$.

We can compute branching rules in general using projection matrices.

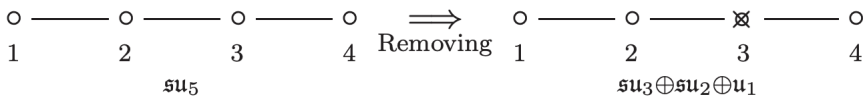
- Suppose $\mathfrak{h} \subset \mathfrak{g}$. We can specify weights in $\mathfrak{g}$ by the coefficients $w_{(i)}$ of the fundamental weights, and similarly in $\mathfrak{h}$ by coefficients $v_{(i)}$. If $\mathfrak{h}$ is not simple, we simply find the coefficients within each component and concatenate the vectors.
- As a simple example, consider the branching associated with $\mathfrak{su}(3) \supset \mathfrak{su}(2) \oplus \mathfrak{u}(1)$. We have

  $$3 \to 2_1 + 1_{-2}$$

  because the weight vectors (10), $(\bar{1}1)$, and $(0\bar{1})$ in $\mathfrak{su}(3)$ map to $(1)(+1)$, $(\bar{1})(+1)$, and $(0)(-2)$ respectively, where a bar stands for a negative sign.
- Hence the weights $v_{(i)}$ and $w_{(i)}$ are related by the projection matrix

  $$P = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}.$$

  Once we have the projection matrix, we can compute branching rules in general by projecting all of the weights of a given representation.
- If the subalgebra $\mathfrak{h}$ is found by a Dynkin diagram, it is simple to write down the projection matrix. For example, consider the subalgebra of $\mathfrak{su}(5)$ below.

  It is clear the projection matrix is

  $$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

  The final row, corresponding to the $\mathfrak{u}(1)$ factor, depends on how the $\mathfrak{u}(1)$ generator is normalized; it may be found by considering an explicit simple representation.

- The projection matrix is not unique. For instance, the usual embedding of $\mathfrak{su}(3) \oplus \mathfrak{su}(2) \oplus \mathfrak{u}(1)$ in $\mathfrak{su}(5)$ used in grand unification instead has

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ -2/3 & 1/3 & -1/3 & 2/3 \end{pmatrix}.$$

  However, the branching rules computed will be the same, up to $\mathfrak{u}(1)$ normalization.

- For multiple layers of subalgebras, $\mathfrak{h} \subset \mathfrak{k} \subset \mathfrak{g}$, we may simply multiply the projection matrices.

7.6 ○ Representations of $SO(n)$

Next, we turn to $SO(n)$ and spinor representations. We begin with $SO(2n+1)$.

- We label the generators as $M_{ab} = -M_{ba}$, where in the defining representation

$$[M_{ab}]_{xy} = -i(\delta_{ax}\delta_{by} - \delta_{bx}\delta_{ay}).$$

  Then a basis for the Cartan subalgebra is $H_j = M_{2j-1,2j}$, for $j = 1, \ldots, n$.

- We've already heuristically found that the roots are

$$e^j \pm e^k, \quad -e^j \pm e^k, \quad \pm e^j.$$

  In terms of the defining representation, they are

$$E_{\eta e^j} = \frac{1}{\sqrt{2}}(M_{2j-1,2n+1} + i\eta M_{2j,2n+1})$$

  and

$$E_{\eta e^j + \eta' e^k} = \frac{1}{2}(M_{2j-1,2k-1} + i\eta M_{2j,2k-1} + i\eta' M_{2j-1,2k} - \eta\eta' M_{2j,2k})$$

  as can be checked by direct commutation, where $\eta, \eta' = \pm 1$.

- Under a suitable definition of positivity, the simple roots are

$$\alpha^j = e^j - e^{j+1} \text{ for } j = 1, \ldots, n-1, \quad \alpha^n = e^n.$$

  Then the fundamental weights are

$$\mu^j = \sum_{k=1}^{j} e^k \text{ for } j = 1, \ldots, n-1, \quad \mu^n = \frac{1}{2}\sum_{k=1}^{n} e^k.$$

  The representation corresponding to μ^n is qualitatively difference since it is associated with the short root; it is the spinor representation.

- Starting with highest weight μ^n, by Weyl reflection in the roots e_j, we get the 2^n weights $(\pm e^1 \pm \ldots \pm e^n)/2$, all of which are nondegenerate since the highest weight was. By some choice of positivity, each of these weights could be the highest weight, and lowering any of them just gives another one, so these are all the weights.

- It is convenient to write the spinor as the tensor product of n two-dimensional spaces,

$$|\pm e^1/2 \pm \ldots \pm e^n/2\rangle = |\pm e^1/2\rangle \otimes \ldots \otimes |\pm e^n/2\rangle$$

and we define σ_a^j to be the Pauli matrix σ_a acting on slot j, so

$$H_j = \frac{1}{2}\sigma_3^j, \quad H_j^2 = \frac{1}{4}.$$

For any $a \neq b$, M_{ab} could have been a Cartan generator since the ordering of axes is arbitrary, so in the spinor representation $M_{ab}^2 = 1/4$ for all $a \neq b$.

- Since any state can only be raised once, $(E_{e^j})^2 = 0$, which implies

$$\{M_{2j-1,2n+1}, M_{2j,2n+1}\} = 0$$

and again since the ordering of axes is arbitrary, we have

$$\{M_{j\ell}, M_{k\ell}\} = 0, \quad j \neq k \neq \ell \neq j.$$

We now construct the generators in the spinor representation.

- First we construct some of the roots. For simplicity we'll set $n = 2$, though the reasoning will be easy to generalize to arbitrary n. First off, we know that

$$E_{e^1}|-e^1/2 \pm e^2/2\rangle \propto |e^1/2 \pm e^2/2\rangle, \quad E_{-e^1}|-e^1/2 \pm e^2/2\rangle = 0.$$

To compute the proportionality constant, take the norm of both sides for

$$\langle -e^1/2 \pm e^2/2|E_{e^1}^\dagger E_{e^1}|-e^1/2 \pm e^2/2\rangle = \langle -e^1/2 \pm e^2/2|\{E_{e^1}^\dagger, E_{e^1}\}|-e^1/2 \pm e^2/2\rangle = \frac{1}{2}$$

using the above results. Hence we have $E_{\pm e^1} = \sigma_\pm^1/2$, where we have implicitly fixed a phase convention. Similarly, we can choose a phase convention so that

$$E_{e^2}|e^1/2 - e^2/2\rangle = \frac{1}{\sqrt{2}}|e^1/2 + e^2/2\rangle.$$

- At this point we have exhausted the phase freedom. Now we compute

$$E_{e^2}|-e^1/2-e^2/2\rangle = \sqrt{2}E_{e^2}E_{-e^1}|e^1/2-e^2/2\rangle = -\sqrt{2}E_{-e^1}E_{e^2}|e^1/2-e^2/2\rangle = -\frac{1}{\sqrt{2}}|-e^1/2+e^2/2\rangle$$

where we used $\{E_{\pm E^1}, E_{\pm e^2}\} = 0$. This extra sign means we need a factor of σ_3^1,

$$E_{\pm e^2} = \frac{1}{2}\sigma_3^1\sigma_\pm^2.$$

- By continuing this reasoning for general n, we have

$$E_{\pm e^j} = \frac{1}{2}\sigma_3^1 \ldots \sigma_3^{j-1}\sigma_\pm^j.$$

Thus, for the generators we have

$$M_{2j-1,2n+1} = \frac{1}{2}\sigma_3^1 \ldots \sigma_3^{j-1}\sigma_1^j, \quad M_{2j,2n+1} = \frac{1}{2}\sigma_3^1 \ldots \sigma_3^{j-1}\sigma_2^j.$$

From these generators we can construct all other generators by anticommutation, as

$$M_{ab} = -i[M_{a,2n+1}, M_{b,2n+1}].$$

Using $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$, each generator is $\pm 1/2$ times a product of Pauli matrices.

- The spinor representation is not complex, since its weights are symmetric. For $n = 1$, we have seen it is equivalent to its conjugate representation by $S = \sigma_2$,

$$\sigma_a = -\sigma_2 \sigma_a^* \sigma_2.$$

Demanding the same for the generators $M_{2j-1,2n+1}$ and $M_{2j,2n+1}$ yields

$$S = \prod_{j \text{ odd}} \sigma_2^j \prod_{k \text{ even}} \sigma_1^k.$$

This implies the spinor representations are real for $SO(8n+1)$ and $SO(8n+7)$, and pseudoreal for $SO(8n+3)$ and $SO(8n+5)$.

Next, we apply similar reasoning for $SO(2n+2)$.

- In this case the simple roots are

$$\alpha^j = e^j - e^{j+1} \text{ for } j = 1, \ldots, n, \quad \alpha^{n+1} = e^n + e^{n+1}$$

where α^n and α^{n+1} are the distinguished roots. The last two fundamental weights are

$$\mu^n = \frac{1}{2}(e^1 + \ldots + e^n - e^{n+1}), \quad \mu^{n+1} = \frac{1}{2}(e^1 + \ldots + e^n + e^{n+1})$$

and correspond to the two spinor representations D^n and D^{n+1}.

- In $SO(2n+2)$, all of the roots have the form $\pm e^j \pm e^k$. Then by similar logic to the case of $SO(2n+1)$, the spinor representations have weights

$$\frac{1}{2}\sum_{j=1}^{n+1} \eta_j e^j, \quad \eta_j = \pm 1$$

but with the additional restriction that the η_j multiply to -1 for D^n and 1 for D^{n+1}. Note that both of these representations have dimension 2^n. We see that the spinor representations are complex exactly when n is even, in which case they are each others' conjugates.

- To construct the representations more explicitly, restrict to $SO(2n+1)$ generated by M_{jk}, with $j, k \leq 2n+1$. Then we lose the last Cartan generator $H_{n+1} = M_{2n+1,2n+1}$, and both of the spinor representations reduce to the one found above.

- Thus, using the same notation, all we have to do is construct H_{n+1}. It is

$$H_{n+1} = \frac{1}{2}\sigma_3^1 \ldots \sigma_3^n \times \begin{cases} -1 & D^n, \\ +1 & D^{n+1}. \end{cases}$$

All the other missing generators can be found by commutation.

- Finally, we determine reality and pseudoreality. We can define R exactly as we did for $SO(2n+1)$, and by the same logic find that the spinors are real for $SO(8n)$ and pseudoreal for $SO(8n+4)$. In summary, the Weyl spinors are:

$SO(8n)$	real
$SO(8n+1)$	real
$SO(8n+2)$	complex
$SO(8n+3)$	pseudoreal
$SO(8n+4)$	pseudoreal
$SO(8n+5)$	pseudoreal
$SO(8n+6)$	complex
$SO(8n+7)$	real

Example. In the case of $\mathfrak{so}(4)$ the algebra is not simple, as it breaks into $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$. Nonetheless the arguments above apply, because at each step we simply ignored one of the two special simple roots. When we do this, we restrict to $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, and the arguments go through as before. Note that the resulting spinors are pseudoreal, but the Weyl spinors of $\mathfrak{so}(3,1)$ are complex; the results depend on the metric signature. We cover the case of Lorentzian signature in the notes on Supersymmetry.

We can also arrive at these results by the alternate route of Clifford algebras.

- A Clifford algebra is a set of N operators satisfying the anticommutation relations

$$\{\Gamma_j, \Gamma_k\} = 2\delta_{jk}.$$

  Given a Clifford algebra, we may define

$$M_{jk} = \frac{1}{4i}[\Gamma_j, \Gamma_k]$$

  which generate a representation of $SO(N)$. We won't show this here, but it is done in detail in the notes on Quantum Field Theory.

- The index on Γ_j may be viewed as a vector index, in the sense that the Γ's transform in the N-dimensional fundamental representation of $SO(N)$,

$$[M_{jk}, \Gamma_\ell] = i(\delta_{j\ell}\Gamma_k - \delta_{k\ell}\Gamma_j) = \Gamma_m (M^{D^1}_{jk})_{m\ell}$$

  where M^{D^1} is in the fundamental representation,

$$[M^{D^1}_{jk}]_{\ell m} = -i(\delta_{j\ell}\delta_{km} - \delta_{jm}\delta_{k\ell}).$$

  That is, the Γ's form a irreducible set of tensor operators.

- For $SO(2n+1)$, the Clifford algebra

$$\Gamma_1 = \sigma_2^1\sigma_3^2\ldots\sigma_3^n, \quad \Gamma_2 = -\sigma_1^1\sigma_3^2\ldots\sigma_3^n, \quad \Gamma_3 = \sigma_2^2\ldots\sigma_3^n, \quad \Gamma_4 = -\sigma_1^2\ldots\sigma_3^n$$

  and so on, up to

$$\Gamma_{2n-1} = \sigma_2^n, \quad \Gamma_{2n} = -\sigma_1^n, \quad \Gamma_{2n+1} = \sigma_3^1\ldots\sigma_3^n$$

  yields a 2^n dimensional representation, precisely the spinor representation we found earlier.

- Using this Clifford algebra, we can simply forget about Γ_{2n+1} to get a Clifford algebra with $2n$ elements, yielding a 2^n dimensional representation of $SO(2n)$. This representation is reducible since Γ_{2n+1} commutes with all the generators, because

$$\Gamma_{2n+1} = (-i)^n \Gamma_1 \Gamma_2 \dots \Gamma_{2n}.$$

  Indeed, we saw above that Γ_{2n+1} is equal to -1 on D^{n-1} and 1 on D^n, so

$$\frac{1}{2}(1 - \Gamma_{2n+1}) \text{ projects onto } D^{n-1}, \quad \frac{1}{2}(1 + \Gamma_{2n+1}) \text{ projects onto } D^n.$$

  This is familiar from the study of Lorentz spinors, where Γ_{2n+1} is called γ^5.

Clifford algebras arise naturally when considering Majorana spinors.

- Suppose we have a set of n independent fermionic creation and annihilation operators. Then their anticommutation relations are

$$\{c_i, c_j^\dagger\} = \delta_{ij}, \quad \{c_i, c_j\} = \{c_i^\dagger, c_j^\dagger\} = 0.$$

  Given these, we may always construct the self-conjugate operators

$$\gamma_i = c_i + c_i^\dagger, \quad \{\gamma_i, \gamma_j\} = 2\delta_{ij}, \quad \gamma_i = \gamma_i^\dagger.$$

  These are called Majorana operators and satisfy the Clifford algebra. However, typically these operators don't create physical excitations, because if the c_i are charged then γ_i does not have definite charge.

- If we do have a set of $2n$ Majorana operators γ_i, we may recover n independent fermionic harmonic oscillators by defining

$$\Psi_k = \frac{1}{2}(\gamma_{2k-1} + i\gamma_{2k}).$$

  Starting from the vacuum, these generate a 2^n dimensional Hilbert space, which is the same reducible representation we saw above. The two pieces correspond to having an even or odd number of particles.

- A set of n harmonic oscillators has an $SU(n)$ symmetry associated with rotating between them. These rotations are generated by

$$T_a = \sum_{jk} \Psi_j^\dagger [T_a]_{jk} \Psi_k$$

  where $[T_a]_{jk}$ is the defining representation of $SU(n)$. This yields a convenient embedding of $SU(n)$ in $SO(2n)$.

- Because the $SU(n)$ rotations preserve particle number, the 2^n dimensional Hilbert space breaks into $n+1$ representations under $SU(n)$, with the m-particle subspace containing $\binom{n}{m}$ states. Since the fermionic oscillators anticommute, these correspond to antisymmetric powers of the fundamental, which we write as $[m]$. For example, for $SU(5) \subseteq SO(10)$ we have

$$D^5 = 16 \to [1] + [3] + [5] = 5 + \overline{10} + 1, \quad D^4 = \overline{16} \to [0] + [2] + [4] = 1 + 10 + \overline{5}.$$

  The latter makes $SO(10)$ suitable for grand unification.

8 Relativistic Fields

8.1 Representations of the Euclidean Group

In this section we introduce a method for finding the representations of the noncompact Euclidean group, as practice for the more complicated Poincare group.

- The Euclidean group E_n is the group of linear transformations of n-dimensional Euclidean space connected to the identity that leave the length of all vectors invariant. It is generated by translations and rotations.

- In the special case of E_2, a general group element $g(\mathbf{b}, \theta)$ is a rotation by θ followed by a translation by $\mathbf{b}$. The multiplication law is

$$g(\mathbf{b}_2, \theta_2) g(\mathbf{b}_1, \theta_1) = g(\theta_1 + \theta_2, R(\theta_2)\mathbf{b}_1 + \mathbf{b}_2)$$

  and inverses are given by

$$g(\mathbf{b}, \theta)^{-1} = g(-R(-\theta)\mathbf{b}, -\theta).$$

- Both translations and rotations can be written as linear transformations on $\mathbb{R}^3$, where the third component of the vector is always unity,

$$g(\mathbf{b}, \theta) = \begin{pmatrix} \cos\theta & -\sin\theta & b^1 \\ \sin\theta & \cos\theta & b^2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x^1 \\ x^2 \\ 1 \end{pmatrix}.$$

  This allows us to deduce commutation relations from the usual matrix algebra.

- The rotations and translations are generated by

$$J = \begin{pmatrix} & -i & \\ i & & \\ & & \end{pmatrix}, \quad P_1 = \begin{pmatrix} & i & \\ & & \\ & & \end{pmatrix}, \quad P_2 = \begin{pmatrix} & & \\ & i & \\ & & \end{pmatrix}, \quad R(\theta) = e^{-i\theta J}, \quad T(\mathbf{b}) = e^{-i\mathbf{b}\cdot\mathbf{P}}.$$

  Note that P_1 and P_2 are not Hermitian. The Lie algebra is

$$[P_1, P_2] = 0, \quad [J, P_k] = i\epsilon^{km} P_m$$

  where $\epsilon^{12} = 1$. Physically, the latter relation states that the translation $\mathbf{b}$ is a vector; exponentiating it gives

$$e^{-i\theta J} P_k e^{i\theta J} = P_m R(\theta)_{mk}, \quad e^{-i\theta J} T(\mathbf{b}) e^{i\theta J} = T(R(\theta)\mathbf{b})$$

  and the latter is just the group multiplication law in another form.

- The translations T_2 form a normal subgroup of E_2, and E_2/T_2 contains pure rotations, so it is isomorphic to $SO(2)$. Since E_2 has an abelian normal subgroup, it is not semi-simple.

- We are interested in unitary representations, but since E_2 is not compact, it has no finite-dimensional faithful unitary representations. The non-faithful finite-dimensional unitary irreps are just those of $SO(2)$, mapping $g(\mathbf{b}, \theta)$ to the phase $e^{im\theta}$.

Next, we construct the faithful unitary irreps of E_2 with the usual method, starting with a reference vector and applying raising and lowering operators.

- We define the raising and lowering operators

$$P_\pm = P_1 \pm iP_2, \quad [P_+, P_-] = 0, \quad [J, P_\pm] = \pm P_\pm.$$

Next, we define the squared momentum operator

$$P^2 = P_1^2 + P_2^2 = P_+P_- = P_-P_+, \quad [P^2, J] = [P^2, P_\pm] = 0.$$

Therefore, P^2 is a Casimir operator and has a single eigenvalue for each irrep.

- In a unitary representation, J, P_1, and P_2 are Hermitian operators; to save space we write the operator associated with a generator with the same symbol. Then $P_+^\dagger = P_-$, so P^2 is positive definite. We thus write its eigenvalue as $p^2 > 0$.

- Since P^2 and J commute, we can simultaneously diagonalize the two, with normalized eigenvectors

$$P^2|pm\rangle = p^2|pm\rangle, \quad J|pm\rangle = m|pm\rangle, \quad m \in \mathbb{Z}$$

where the eigenvalues of J come from our knowledge of the irreps of $SO(2)$. In principle there could be multiple vectors in the same irrep with the same p^2 and m, but we'll see below this isn't the case.

- In the case where $p^2 = 0$, we have $P_\pm|0m\rangle = 0$, so the representation is one-dimensional. This simply reproduces the non-faithful irreps we found above.

- For $p^2 > 0$, we take the phase convention

$$|p, m \pm 1\rangle = (\pm i/p)P_\pm|pm\rangle.$$

Starting with any reference vector $|pm_0\rangle$, repeated application of $P^\pm$ gives all integer values of m precisely once. The matrix elements are

$$\langle pm'|J|pm\rangle = m\delta_{m'm}, \quad \langle pm'|P_\pm|pm\rangle = \mp ip\delta_{m',m\pm1}.$$

- We claim the representation matrices for finite transformations are

$$D^p(\mathbf{b}, \theta)_{m'm} = e^{i(m-m')\phi} J_{m-m'}(pb) e^{-im\theta}$$

where (b, ϕ) are the polar coordinates of $\mathbf{b}$ and J_n is the Bessel function of the first kind. The $e^{-im\theta}$ factor is from the rotation part. The translation can be decomposed as

$$T(b, \phi) = R(\phi)T(b, 0)R(\phi)^{-1}, \quad \langle pm'|T(b, \phi)|pm\rangle = e^{i(m-m')\phi}\langle pm'|T(b, 0)|pm\rangle.$$

Therefore, it suffices to show that

$$\langle pm'|T(b, 0)|pm\rangle = J_{m-m'}(pb).$$

We have $T(b, 0) = e^{-ibP_1} = e^{-ib(P_+ + P_-)/2}$. This can be expanded in a double series, which collapses to a single series by orthogonality, giving the defining series for the Bessel function. This is not unexpected; the Bessel function is fundamentally tied to cylindrical symmetry, as the spherical Bessel function is tied to spherical symmetry.

We call the $|pm\rangle$ basis found above the angular momentum basis. Now we work in the "plane-wave" basis using the method of induced representations.

- This method works for groups with abelian normal subgroups. The idea is to use the generators of this subgroup as a starting point, i.e. working with eigenvectors of $\mathbf{P}$ instead of P^2 and J.
- We take the reference vector $\mathbf{p}_0 = (p, 0)$ and consider a reference ket $|\mathbf{p}_0\rangle$ with $\mathbf{P}|\mathbf{p}_0\rangle = \mathbf{p}_0|\mathbf{p}_0\rangle$. The only operation which yields new kets is the rotation $R(\theta) = e^{-i\theta J}$, and we have

  $$P_k R(\theta)|\mathbf{p}_0\rangle = R(\theta)(R(\theta)^{-1} P_k R(\theta))|\mathbf{p}_0\rangle = p_k R(\theta)|\mathbf{p}_0\rangle, \quad p_k = R(\theta)_{k\ell} p_{0\ell}.$$

  The notation here is overloaded, as $R(\theta)$ is both an operator on the irrep and a 3×3 matrix.
- We see that $R(\theta)|\mathbf{p}_0\rangle$ is also an eigenvector of $\mathbf{P}$ with the rotated momentum $\mathbf{p}$, so we write

  $$|p, \theta\rangle = R(\theta)|\mathbf{p}_0\rangle.$$

  Since $R(2\pi) = 1$, this shows that $|\mathbf{p}_0\rangle$ is the only vector with eigenvalue $\mathbf{p}_0$ in the irrep. Thus the set of $|p, \theta\rangle$ defined above form an irrep of E_2.
- Stepping back, we used the fact that the translation subgroup is abelian to label states with momentum, and we used its normality to show that rotations take states of definite momentum to other states of definite momentum. Then the rest of the procedure is simply generating all the kets we can to get an irrep. In more general situations, there can be group generators that commute with $\mathbf{P}$, yielding multiple kets with the same momentum.
- We choose to normalize the vectors so that

  $$\langle p, \theta' | p, \theta\rangle = 2\pi\delta(\theta' - \theta).$$

  We don't need to normalize over p, since p is fixed in this irrep due to the Casimir operator P^2.
- Next, we find the relationship between the angular momentum and plane wave basis. Dropping the p index, consider the state

  $$|\widetilde{m}\rangle = \int_0^{2\pi} \frac{d\phi}{2\pi} e^{im\phi} |\phi\rangle.$$

  By construction, we have

  $$R(\theta)|\widetilde{m}\rangle = e^{-im\theta}|\widetilde{m}\rangle$$

  by shifting the integration variable, so $|\widetilde{m}\rangle$ is proportional to $|m\rangle$. To fix the phases, note

  $$P_\pm|\widetilde{m}\rangle = |\widetilde{m} \pm 1\rangle, \quad |m\rangle = i^m|\widetilde{m}\rangle, \quad \langle\phi|m\rangle = e^{im(\phi+\pi/2)}.$$

  The inner product allows us to switch between bases both ways, e.g.

  $$|\phi\rangle = \sum_m |m\rangle e^{-im(\phi+\pi/2)}.$$

- As an application, note that under a translation, we have

  $$T(\mathbf{b})|m\rangle = \int \frac{d\phi}{2\pi} e^{im(\phi+\pi/2)} e^{-ipb\cos(\theta-\phi)} |\phi\rangle$$

where $\mathbf{b}$ has polar angle θ. Then we immediately have

$$\langle m'|T(\mathbf{b})|m\rangle = \int \frac{d\phi}{2\pi} e^{i(m-m')(\phi+\pi/2)} e^{-ipb\cos(\theta-\phi)} = e^{i(m-m')\theta} \int \frac{d\psi}{2\pi} e^{i(m-m')\psi - ipb\sin\psi}$$

where $\psi = \pi/2 + \phi - \theta$, which recovers our earlier result using the integral representation of the Bessel function. We can also show that the position-space wavefunctions $\langle \mathbf{r}|m\rangle$ are Bessel functions times $e^{im\theta}$. **(verify)** Further investigation recovers many of the standard properties of Bessel functions, such as their differential equation, recursion formulas, and addition theorems.

Note. As we've seen, E_2 is the group contraction of $SO(3)$ with

$$J_y/R \to P_x, \quad -J_x/R \to P_y, \quad J_z \to J$$

in the limit $R \to \infty$. Then we can recover the representations of E_2 from those of $SO(3)$. This sounds strange since the former has infinite-dimensional representations, but note that for the spin j representation,

$$\langle m'|J_x \pm iJ_y|m\rangle = \delta_{m',m\pm1}\sqrt{j(j+1) - m(m\pm1)}.$$

Then when we replace the J_i with P_i, the inner product is proportional to $1/R$, so the representation is automatically one-dimensional unless j grows with R. Specifically, if we pick $j = pR$, then

$$\langle m'|P_x \pm iP_y|m\rangle = p\delta_{m',m\pm1}$$

which exactly matches the E_2 representation with squared momentum p^2. As an application, it is possible to express Bessel functions in terms of limits of d-functions.

We now apply the same methods to E_3.

- The group E_3 is generated by rotations and translations, and the Lie algebra is

  $$[P_i, P_j] = 0, \quad [J_i, J_j] = i\epsilon_{ijk}J_k, \quad [P_i, J_j] = i\epsilon_{ijk}P_k$$

  which says that both $\mathbf{P}$ and $\mathbf{J}$ are vectors. The translations T_3 form a normal subgroup.

- The general group element can be written with the Euler angle parametrization

  $$g = T(\mathbf{b})R(\alpha,\beta,\gamma) = e^{-i\mathbf{b}\cdot\mathbf{P}} e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3}.$$

  Alternatively, since conjugation by rotations rotates a translation vector, we can write

  $$g = R(\phi,\theta,0)T(\mathbf{b}_0)R(\alpha',\beta',\gamma')$$

  where $\mathbf{b}$ has spherical coordinates (b,ϕ,θ) and $\mathbf{b}_0$ has spherical coordinates $(b,0,0)$.

- We can construct three independent rotational scalars, P^2, $\mathbf{J}\cdot\mathbf{P}$ and J^2, which are candidates for Casimir operators. We have $[J^2, \mathbf{P}] \neq 0$, intuitively because the rotation operators pick out a fixed origin, which is shifted by translations. However,

  $$[\mathbf{J}\cdot\mathbf{P}, P_j] = [J_i, P_j]P_i = -i\epsilon_{ijk}P_kP_i = 0$$

  and clearly $[P^2, \mathbf{P}] = 0$, so the two Casimir operators are P^2 and $\mathbf{J}\cdot\mathbf{P}$.

- To work in the angular momentum basis, we simultaneously diagonalize $\{P^2, \mathbf{J}\cdot\mathbf{P}, J^2, J_z\}$. To work in the plane wave basis, we simultaneously diagonalize $\{P^2, \mathbf{J}\cdot\mathbf{P}, \mathbf{P}\}$.
- In the plane wave basis, the eigenvalues are $\{p^2, \lambda p, \mathbf{p}\}$ and we write a ket with these eigenvalues as $|p, \lambda, \hat{\mathbf{p}}\rangle$ where $\hat{\mathbf{p}} = \mathbf{p}/p$ and λ is called the helicity.
- As before, we start with the standard vector $\hat{\mathbf{p}}_0 = \hat{\mathbf{e}}_z$. Define the little group to be the set of group elements in the quotient group $E_3/T_3 \cong SO(3)$ that leave the standard vector $\hat{\mathbf{p}}_0$ invariant. In this case, the little group is $SO(2)$, generated by $R_3(\phi) = e^{-i\phi J_3}$.
- We can begin by constructing an irrep of the little group, by applying its group operations on the reference ket. Then, as for E_2, we can use this to construct ("induce") an irrep of E_3 using the rotations outside the little group. (The first step was not necessary for E_2, because there the little group was trivial.)
- Here, the irreps of the little group are all one-dimensional, and they are indexed by the eigenvalue of J_3, which must be an integer. Thus the little group irrep is specified by $\lambda \in \mathbb{Z}$. The rest of the E_3 irrep is generated by rotations,

  $$|p, \lambda, \hat{\mathbf{p}}\rangle = R(\phi, \theta, 0)|p, \lambda, \hat{\mathbf{p}}_0\rangle, \quad \hat{\mathbf{p}} = (\theta, \phi).$$

  Suppressing the p and λ indices, translations simply act by

  $$T(\mathbf{b})|\hat{\mathbf{p}}\rangle = e^{-i\mathbf{b}\cdot\mathbf{p}}|\hat{\mathbf{p}}\rangle.$$

- Rotations are a bit more complex. We have

  $$R(\alpha, \beta, \gamma)|\hat{\mathbf{p}}\rangle = R(\alpha, \beta, \gamma)R(\phi, \theta, 0)|\hat{\mathbf{p}}_0\rangle = R(\phi', \theta', \psi)|\hat{\mathbf{p}}_0\rangle$$

  where the second equality defines ϕ', θ', and ψ. Then we have

  $$R(\alpha, \beta, \gamma)|\hat{\mathbf{p}}\rangle = R(\phi', \theta', 0)R(0, 0, \psi)|\hat{\mathbf{p}}_0\rangle = e^{-i\lambda\psi}|\hat{\mathbf{p}}'\rangle, \quad \hat{\mathbf{p}}' = (\theta', \phi').$$

  The point is that $\hat{\mathbf{p}}$ gets rotated in the obvious way, but with a more subtle extra phase $e^{-i\lambda\psi}$, which has already been fixed by the action of the little group on the reference ket.
- We normalize the vectors by

  $$\langle\hat{\mathbf{p}}'|\hat{\mathbf{p}}\rangle = 4\pi\delta(\cos\theta' - \cos\theta)\delta(\phi' - \phi)$$

  since this cancels the invariant measure $\sin\theta\, d\theta d\phi/4\pi$ on $SO(3)$.
- Working in the angular momentum basis, we have vectors $|p, \lambda, j, m\rangle$ corresponding to eigenvalues $\{p^2, \lambda p, j(j+1), m\}$ of $\{P^2, \mathbf{J}\cdot\mathbf{P}, J^2, J_3\}$. We can then construct raising and lowering operators $J_\pm$ which raise and lower m. Since the translations don't commute with J^2, they can change j. It turns out that j takes on all positive integer values, making the representation infinite-dimensional as expected.
- The radial dependence of the angular momentum basis states takes the form of spherical Bessel functions. **(verify)** By similar reasoning to E_2, we can use the group structure of E_3 to recover many of their properties.

8.2 ○ Representations of the Poincare Group

Finally, we turn to the Poincare group, our original goal. We use the $(-+++)$ metric convention.

- The Poincare group is the analogue of the Euclidean group for Minkowski space. Its elements take the form
$$g(b, \Lambda) = T(b)\Lambda, \quad x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + b^\mu$$
where Λ is a Lorentz transformation connected to the identity. By similar reasoning to before, the translations form a normal subgroup and the factor group is the Lorentz group. Again, we will overload the notation so that Λ is both a 4×4 matrix and an operator.

- A general Lorentz transformation can be written as
$$\Lambda = R(\alpha, \beta, 0)L_3(\xi)R(\phi, \theta, \psi)^{-1}$$
where $L_3(\xi)$ is a boost of rapidity ξ along the z-axis. This is the same idea as before: a boost can be conjugated by rotations to change its direction to the z-axis.

- To work with the generators, we have to be careful with index placement. An infinitesimal translation by δb is
$$T(\delta b) = I - i\delta b^\mu P_\mu, \quad T(b) = \exp(-ib^\mu P_\mu).$$
The translation generators P_μ are related to the physical four-momentum by raising an index, so $P^0 = -P_0$ and $P^i = P_i$. Under the multiplication law, P_μ indeed transforms as a covector.

- The Lorentz generators are defined as
$$\Lambda(\delta\omega) = I - \frac{i}{2}\delta\omega^{\mu\nu}J_{\mu\nu}$$
where $\delta\omega$ is an antisymmetric tensor that parametrizes the transformation as
$$\delta\omega^{23} \text{ is angle of rotation about } x, \quad \delta\omega^{10} \text{ is boost along } x.$$
In terms of differential operators, we have
$$J_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu), \quad P_\mu = i\partial_\mu.$$
There may be extra signs here, depending on the convention.

- The commutation relations are
$$[P_\mu, J_{\lambda\sigma}] = i(P_\lambda\eta_{\mu\sigma} - P_\sigma\eta_{\mu\lambda}), \quad [J_{\mu\nu}, J_{\lambda\sigma}] = i(J_{\lambda\nu}\eta_{\mu\sigma} - J_{\sigma\nu}\eta_{\mu\lambda} + J_{\mu\lambda}\eta_{\nu\sigma} - J_{\mu\sigma}\eta_{\nu\lambda}).$$
These can be shown using 5×5 matrices or by using the differential operators. Intuitively, these results just say that P_μ is a vector and $J_{\mu\nu}$ is a tensor.

Next, we apply the method of induced representations to the Poincare group.

- We use the abelian subgroup of translations, and note there is a Casimir operator

$$C_1 = -P_\mu P^\mu = P_0^2 - P^2.$$

We let the value of the operator be c_1. For a momentum p^μ, there are several cases:

$$\text{trivial: } c_1 = p^\mu = 0, \quad \text{timelike: } c_1 > 0, \quad \text{lightlike: } c_1 = 0,\ p^\mu \neq 0, \quad \text{spacelike: } c_1 < 0.$$

In the trivial case, the little group is the Lorentz group, but translations do nothing; this recovers the Lorentz irreps found above. The resulting states represent vacuum states, rather than one-particle states.

- For the timelike case $c_1 = M^2$, we choose the standard momentum $p_t^\mu = (M, \mathbf{0})$. The little group is $SO(3)$, so we write the states with the standard momentum as $|\mathbf{0}\lambda\rangle$ where

$$P^\mu|\mathbf{0}\lambda\rangle = p_t^\mu|\mathbf{0}\lambda\rangle, \quad J^2|\mathbf{0}\lambda\rangle = s(s+1)|\mathbf{0}\lambda\rangle, \quad J_3|\mathbf{0}\lambda\rangle = \lambda|\mathbf{0}\lambda\rangle, \quad \lambda \in \{-s, -s+1, \ldots, s\}.$$

Here the indices M and s are suppressed.

- Next, we construct the rest of the states by applying Lorentz transformations. To do this, we need to pick a conventional Lorentz transformation that maps the reference momentum to any desired momentum $\mathbf{p}$. Noting again that a general Lorentz transformation can be written as

$$\Lambda = R(\alpha, \beta, 0)L_3(\xi)R(\phi, \theta, \psi)^{-1} \equiv H(p)R(\phi, \theta, \psi)^{-1}$$

where the rightmost factor does nothing since it is in the little group, we choose to define

$$|p\hat{\mathbf{z}}\,\lambda\rangle = L_3(\xi)|\mathbf{0}\lambda\rangle, \quad p = M\sinh\xi$$

and therefore

$$|\mathbf{p}\lambda\rangle = R(\alpha, \beta, 0)|p\hat{\mathbf{z}}\,\lambda\rangle = H(p)|\mathbf{0}\lambda\rangle.$$

The $|\mathbf{p}\lambda\rangle$ are now a full irrep of the Poincare group, and the equations above simply define, conventionally, how the λ index behaves for nonzero momenta. We would have gotten a different convention if we had used a different choice of "standard Lorentz transformation" $H(p)$.

- By the same proof as for E_2, the $|\mathbf{p}\lambda\rangle$ states indeed have momentum $\mathbf{p}$, in the sense that

$$T(b)|\mathbf{p}\lambda\rangle = e^{-ib^\mu p_\mu}|\mathbf{p}\lambda\rangle.$$

To see how the Lorentz transformations act, we argue as for E_3. We have

$$\Lambda|\mathbf{p}\lambda\rangle = \Lambda H(p)|\mathbf{0}\lambda\rangle = H(\Lambda p)R(\Lambda, p)|\mathbf{0}\lambda\rangle, \quad R(\Lambda, p) = H^{-1}(\Lambda p)\Lambda H(p).$$

The "Wigner rotation" $R(\Lambda, p)$ maps $p_t \mapsto p \mapsto \Lambda p \mapsto p_t$, so it is in the little group.

- On the other hand, we fixed how the little group acts in the very first step, which gives

$$\Lambda|\mathbf{p}\lambda\rangle = D^s(R(\Lambda, p))_{\lambda'\lambda}|\Lambda\mathbf{p}\,\lambda'\rangle$$

where the D^s are the spin s representation matrices, and $\Lambda\mathbf{p}$ is the spatial part of Λp. This is the generalization of the extra phase picked up for E_3.

- "Common sense" would expect that when Λ is a pure rotation, the Wigner rotation $R(\Lambda, p)$ is precisely that same rotation, for any p. This is indeed true, as can be checked in a tedious, explicit calculation.

- We now give interpretations for all the parameters above. The irrep is labeled by s and M, where s indicates the spin and M indicates the mass, as seen from the relativistic dispersion relation. We've just shown that $\mathbf{p}$ is the spatial momentum, and by the same argument as for E_3, λ may be identified with the helicity, i.e. the eigenvalue of $\mathbf{J} \cdot \mathbf{P}/|\mathbf{p}|$, when $|\mathbf{p}| \neq 0$. The difference is that the helicity is no longer a Casimir invariant, as it can be changed by boosts.

- There is a second Casimir invariant related to the spin s. As we've seen, it can't be J^2, since that doesn't commute with translations. Another guess is $J_{\mu\nu}J^{\mu\nu}$, since it's also quadratic in $J_{\mu\nu}$, but it doesn't commute with translations either, and moreover isn't simply related to s.

- Instead, we define the Pauli–Lubanski vector

$$W^\lambda = \epsilon^{\lambda\mu\nu\sigma} J_{\mu\nu} P_\sigma / 2.$$

  It has the properties

$$W^\lambda P_\lambda = 0, \quad [W^\lambda, P^\mu] = 0, \quad [W^\lambda, J^{\mu\nu}] = i(W^\mu \eta^{\lambda\nu} - W^\nu \eta^{\mu\lambda}), \quad [W^\lambda, W^\sigma] = i\epsilon^{\lambda\sigma\mu\nu} W_\mu P_\nu.$$

  The proofs are straightforward, mostly using the antisymmetry of ϵ. The second and third properties say W^λ is a translationally invariant four-vector. The second Casimir operator is

$$C_2 = W^\lambda W_\lambda.$$

  This is because C_2 is a Lorentz scalar, and it is translationally invariant since W^λ is. Also, note that W^0 is $\mathbf{J} \cdot \mathbf{P}$, so W^λ essentially completes the helicity to a four-vector.

- To understand the meaning of C_2, we return to the timelike case. In the rest frame, i.e. for the states with momentum p_t^μ, we may replace P_σ with $p_{\sigma t}$ to give w^λ, where

$$w^0 = 0, \quad w^i = M J^i$$

  so C_2 reduces to $M^2 J^2$ and hence gives the spin parameter s. Thus we recover the situation with E_2 and E_3 where the irrep is fully specified by the values of the Casimir operators.

- The previous point shows that in the timelike case, the w^μ span the algebra of the little group. This is in fact true in general, and we'll use this to help handle the light-like case below.

- Next, we turn to the light-like case, with standard momentum

$$p_\ell^\mu = (\omega_0, 0, 0, \omega_0).$$

  In this case the generators of the little group are

$$w^0 = w_3 = \omega_0 J_3, \quad w_1 = \omega_0(J_1 + K_2), \quad w_2 = \omega_0(J_2 - K_1)$$

  which have commutation relations

$$[w_1, w_2] = 0, \quad [w_2, J_3] = i w_1, \quad [w_1, J_3] = -i w_2.$$

  This is exactly the algebra of E_2, and the little group is indeed E_2.

- The little group elements here are a little less intuitive. The rotations in E_2 correspond to rotations about the x-axis. The translations correspond to boosting along the y-axis and then rotating about the z-axis (or vice versa), so that the momentum is realigned along $\hat{\mathbf{x}}$.
- The infinite-dimensional irreps of E_2 correspond to particles with infinitely many internal states. These strange "continuous spin particles" do not seem to appear in nature.
- There are also one-dimensional irreps indexed by λ, the eigenvalue of J_3, giving states

$$P^\mu|\mathbf{p}_\ell\lambda\rangle = p_\ell^\mu|\mathbf{p}_\ell\lambda\rangle, \quad J_3|\mathbf{p}_\ell\lambda\rangle = \lambda|\mathbf{p}_\ell\lambda\rangle, \quad W_1|\mathbf{p}_\ell\lambda\rangle = W_2|\mathbf{p}_\ell\lambda\rangle = 0.$$

  That is, the other little group transformations must do nothing to the states. It turns out that for photons ($\lambda = \pm 1$), at the level of the fields, W_1 and W_2 generate gauge transformations.
- We then construct the states $|\mathbf{p}\lambda\rangle$ in the same way as before. One can show that

$$T(b)|\mathbf{p}\lambda\rangle = e^{-ib^\mu p_\mu}|\mathbf{p}\lambda\rangle, \quad \Lambda|\mathbf{p}\lambda\rangle = e^{-i\lambda\theta(\Lambda,p)}|\Lambda\mathbf{p}\,\lambda\rangle$$

  where the phase is

$$e^{-i\lambda\theta(\Lambda,p)} = \langle\mathbf{p}_\ell\lambda|H^{-1}(\Lambda p)\Lambda H(p)|\mathbf{p}_\ell\lambda\rangle.$$

  Unlike the timelike case, the helicity λ is now Poincare invariant.
- CPT symmetry flips λ, so a relativistic theory must have pairs $\pm\lambda$. For instance, the photon has $\lambda = \pm 1$ and the graviton has $\lambda = \pm 2$. Sometimes we describe this as "spin 1" and "spin 2", but this is not technically correct, because spin labels $SO(3)$ irreps, not E_2 irreps.
- For a single helicity, the Wigner rotation only produces a physically irrelevant phase. But for photons and gravitons, whose one-particle states may involve superpositions of helicities, the Wigner rotation may change their relative phase, changing the polarization. For example, for linear photon polarizations, a rotation about the propagation direction rotates the plane of polarization accordingly.
- We can also consider double-valued representations, giving $\lambda = -1/2$ to describe massless neutrinos and $\lambda = 1/2$ to describe massless antineutrinos. Note that we only have double-valued representations, even though the universal cover of $SO(2)$ is $\mathbb{R}$, because it is part of the Lorentz group, which has only a double cover.
- For the spacelike case, taking the standard momentum $p_s^\mu = (0, 0, 0, Q)$, we have

$$w^0 = QJ_3, \quad w_1 = QJ_{20} = QK_2, \quad w_2 = QJ_{01} = -QK_1$$

  and $w_3 = 0$. The little group is thus $SO(2,1)$, which is noncompact.
- Just like E_2, $SO(2,1)$ thus admits no faithful finite-dimensional unitary representations. However, it turns out that there are no nonfaithful finite-dimensional unitary representations either. For example, setting the noncompact boost generators to zero does not work, because we need to have $[K_1, K_2] = -iJ_3$. As in the light-like case, the infinite-dimensional irreps do not seem to appear in nature.
- Finally, if we wish we can normalize the states to be compatible with the Lorentz invariant integration measure, $đ\mathbf{p}/2p^0$. Then we have, in the spacelike and lightlike cases,

$$\langle\mathbf{p}'\lambda'|\mathbf{p}\lambda\rangle = 2p^0\text{đ}(\mathbf{p} - \mathbf{p}')\delta_\lambda^{\lambda'}.$$

  Of course, in any case the little group representation is unitary.

8.3 Relativistic Field Equations

We now connect the transformation properties of fields and particles. We warm up with the case of fields transforming under rotations. Though index placement is not important here, we maintain it since it'll be needed later; for matrices, the upper index is always the first index. For clarity, we always distinguish abstract operators and their representations.

- Under a rotation R, the position states of a spinless particle in $\mathbb{R}^3$ transform as

$$|\mathbf{x}\rangle \mapsto U(R)|\mathbf{x}\rangle = |R\mathbf{x}\rangle.$$

  If the particle had spin 1/2, we would instead have

$$U(R)|\mathbf{x},\sigma\rangle = D^{1/2}(R)^{\lambda}_{\sigma}|R\mathbf{x},\lambda\rangle$$

  where $D^j(R)$ is the representation matrix for spin j.

- For spinless particles, wavefunctions are defined by $\psi(\mathbf{x}) = \langle\mathbf{x}|\psi\rangle$, so $|\psi'\rangle = U(R)|\psi\rangle$ has wavefunction

$$\psi'(\mathbf{x}) = \psi(R^{-1}\mathbf{x}).$$

  The inverse here can be understood as the usual active/passive inverse. Similarly, for particles with spin, the wavefunction picks up a spin index which also transforms,

$$\psi'^{\lambda}(\mathbf{x}) = D^{1/2}(R)^{\lambda}_{\sigma}\psi^{\sigma}(R^{-1}\mathbf{x}).$$

  This generalizes directly to the wavefunction of a spin j particle.

- Next, we consider the transformation of observables. The position operator satisfies $X^i|\mathbf{x}\rangle = x^i|\mathbf{x}\rangle$, and using the transformation of $|\mathbf{x}\rangle$ gives

$$U(R)X^iU(R)^{-1} = (R^{-1})^i_j X^j.$$

  Using the orthogonality of R, we have

$$U(R)X_iU(R)^{-1} = X_jR^j_i.$$

  This is an example of an irreducible tensor operator.

- Finally, we consider second quantized fields, which represent local observables. For a free spin 1/2 field $\Psi^{\sigma}(\mathbf{x})$ and a one-particle state $|\psi\rangle$, the field is constructed so that

$$\langle 0|\Psi^{\sigma}(\mathbf{x})|\psi\rangle = \psi^{\sigma}(\mathbf{x})$$

  where $\psi^{\sigma}(\mathbf{x})$ is the wavefunction. Fiddling around with the above properties gives

$$U(R)\Psi^{\sigma}(\mathbf{x})U(R)^{-1} = D^{1/2}(R^{-1})^{\sigma}_{\lambda}\Psi^{\lambda}(R\mathbf{x}).$$

  The factor of $D^{1/2}$ is the same as for tensor operators, while the position argument is changed since rotating, creating a particle, and then rotating back creates a particle in a different place. A field of general spin behaves similarly.

- We can also try to move the factor of $D^{1/2}$ to the right. Since $U(R)$ and $D^{1/2}(R)$ are unitary,

$$U(R)\Psi^\sigma(\mathbf{x})U(R)^{-1} = \Psi^\lambda(R\mathbf{x})D(R)^{*\lambda}_{\ \ \sigma}.$$

  The complex conjugate and index placement is unappealing, but we can remove it by taking the adjoint of both sides. Note that $D(R)$ is a set of numbers, not an operator on the Hilbert space, so it gets conjugated rather than adjointed, for

$$U(R)\Psi^\dagger{}_\sigma(\mathbf{x})U(R)^{-1} = \Psi^\dagger{}_\lambda(R\mathbf{x})D(R)^\lambda_\sigma.$$

  In some, but not all cases, Ψ is Hermitian, so we get the same result as for tensor operators.

Next, we move to the relativistic case.

- In general, the Hilbert space carries an infinite-dimensional unitary representation $U(\Lambda, \mathbf{b})$ of the Poincare group. We identify individual particles with irreps of the Poincare group; this is sensible, as taking a particle and moving it around in some way should keep it the "same" particle. For concreteness, we consider a particle with mass m and spin s.

- Equivalently, we could define a one-particle state as a state which can be specified by its four-momentum, and a list of additional indices with discrete values. This rules out two-particle states because the relative velocity is continuous, but allows us to count the ground state of hydrogen as a one-particle state, i.e. it does not distinguish between elementary and composite particles.

- The set of one-particle states of a given species is hence $\{|\mathbf{p}\lambda\rangle\}$, while at the level of the free theory, multiple-particle states can be built from this irrep by the Fock construction. We define creation operators, which satisfy

$$|\mathbf{p}\lambda\rangle = a^\dagger(\mathbf{p}\lambda)|0\rangle.$$

  Using our previous results, under Lorentz transformations the creation operators obey

$$U(\Lambda)a^\dagger(\mathbf{p}\lambda)U(\Lambda^{-1}) = a^\dagger(\Lambda\mathbf{p}\,\lambda')D^s(R(\Lambda,p))^{\lambda'}_\lambda.$$

  Taking the adjoint, the annihilation operators transform as

$$U(\Lambda)a(\mathbf{p}\lambda)U(\Lambda^{-1}) = D^s(R(\Lambda,p)^{-1})^\lambda_{\lambda'}a(\Lambda\mathbf{p}\,\lambda').$$

  Here, $R(\Lambda, p)$ is as defined earlier, the D^s are the spin s representation matrices of $SO(3)$, and p is defined to have spatial part $\mathbf{p}$, $p^0 > 0$, and $p^2 = -m^2$.

- As in the nonrelativistic case, we may define the wavefunction of a one-particle state by projecting onto the $|\mathbf{p}\lambda\rangle$ states. Then the wavefunction should transform with a factor of D^s, but this formalism is not manifestly Lorentz invariant.

- Instead, we prefer to let the wavefunction transform under a finite-dimensional representation of the Lorentz group $D(\Lambda)$ as

$$\psi'^\alpha(x) = D(\Lambda)^\alpha_\beta\psi^\beta(\Lambda^{-1}x), \quad \langle x\alpha|\psi\rangle = \psi^\alpha(x)$$

  where the α and λ indices are related by mode functions,

$$\langle x\alpha|\mathbf{p}\lambda\rangle = u^\alpha(\mathbf{p}\lambda)e^{ipx}.$$

  For a fixed value of $\mathbf{p}$, the $u^\alpha(\mathbf{p}\lambda)$ are also called polarizations.

- Switching to the set of states $|\mathbf{x}\alpha\rangle$ severely enlarges the Hilbert space; it contains particles with the wrong mass, negative energy, and the wrong spin, as a generic Lorentz representation contains multiple spins. Hence we'll have to project out an appropriate subset later.
- Once we have defined wavefunctions, we can define relativistic field operators as

$$\langle 0|\Psi^\alpha(\mathbf{x})|\psi\rangle = \psi^\alpha(\mathbf{x})$$

  which by the same proof transform as

$$U(\Lambda)\Psi^\alpha(x)U(\Lambda^{-1}) = D(\Lambda^{-1})^\alpha_\beta \Psi^\beta(\Lambda x).$$

  The nomenclature is a bit confusing, as both the states and the observables can be called fields, since they both depend on position. For example, the wavefunction for a Dirac fermion $\psi^\alpha(x)$ is though of as a field in first quantization. We will distinguish operator fields by capital letters.
- In other words, the role of the mode functions is to intervene between particle states, which carry momenta and little group indices, and fields, which carry momenta and Lorentz indices.

We restrict to the Poincare irrep using the relativistic wave equations that the field satisfies. These constraints will ensure that the field only creates particles in the desired Poincare irrep.

- The general form of such an equation is

$$\Pi(m, -i\partial)^\alpha_\beta \Psi^\beta(x) = 0$$

  which in Fourier space becomes

$$\Pi(m,p)\Phi^\alpha(p) = 0$$

  where $\Phi(p)$ is the Fourier transform of $\Psi(x)$, with identical transformation properties. For example, the Klein–Gordan field has $\Pi(m,p) = p^2 + m^2$.
- We demand the wave equation be relativistically invariant, so

$$D(\Lambda)\Pi(m,p)D(\Lambda^{-1}) = \Pi(m,\Lambda p).$$

  In other words, if $\Pi(m,p)\Phi^\alpha(p) = 0$, then $\Pi(m,\Lambda p)\Phi(\Lambda p) = 0$, so different observers agree on whether the equation of motion is satisfied.
- The equation of motion must impose the mass shell condition

$$(p^2 + m^2)\Phi(p) = 0.$$

  We then parametrize the on-shell degree of freedom by $\Phi(p) = \delta(p^2 + m^2)\widetilde{\Phi}(p)$. This ensures that we only get particles with mass m. The Klein–Gordan equation clearly does this. The Dirac equation does it as well, since the Dirac operator squares to $p^2 + m^2$, and additionally prevents an extraneous doubling of the degrees of freedom, since a Dirac spinor field transforms as $(1/2, 0) + (0, 1/2)$.
- Note that the mass shell condition has two possible energies for each momentum; that is, it doesn't rule out negative energies. This is related to the prediction of antiparticles in quantum field theory, as discussed in the notes on Quantum Field Theory.

- Finally, $\Pi(m,p)$ must act like a projection matrix that selects out states with spin s. This is necessary, for instance, to remove the spin 0 part of a vector field A^μ if we wish to use it to describe only a massive spin 1 particle, as discussed below.
- With this setup, the general quantum field solution has the form

$$\Psi^\alpha(x) = \sum_\lambda \int \widetilde{dp}\, b(\mathbf{p}\lambda) u^\alpha(\mathbf{p}\lambda) e^{ipx} + \text{negative energy term}$$

  where $\widetilde{dp}$ is the invariant measure and the $b(\mathbf{p}\lambda)$ are some unknown operators; plugging in the definition of the quantum field shows they are simply the annihilation operators $a(\mathbf{p}\lambda)$.

Example. Consider a massive vector field V^μ. In this case, the mode functions carry a Lorentz index, and it is conventional to write them as $\epsilon^\mu_\lambda(\mathbf{p})$. The option that gives the spin 0 particle is

$$\epsilon^\mu(\mathbf{p}) \propto p^\mu$$

where there is no little group index, since the index has only a single value. Upon plugging this in, we find that V^μ has precisely the mode expansion of the derivative of a scalar field, $\partial^\mu\phi$. Hence it is much more useful to work with ϕ directly. A "normal" Lagrangian in terms of V^μ would be effectively higher in derivatives than one for ϕ, and wouldn't be able to describe interactions such as ϕ^4. One could describe such interactions if we allowed the Lagrangian to depend on $(1/\partial^\mu)V^\mu$, but then it would be nonlocal.

The option that gives the spin 1 particle is

$$\epsilon_0(\mathbf{0}) \propto \hat{e}_z, \quad \epsilon_{\pm1}(\mathbf{0}) \propto \hat{e}_x \pm i\hat{e}_y$$

where the little group index gives the spin along the z-direction, and the result is extended to nonzero $\mathbf{p}$ by applying the standard Lorentz transformation, as discussed in the previous section. Notice that these polarization are always orthogonal to p^μ, $\epsilon^\mu p_\mu = 0$. At the level of the fields, this corresponds to the constraint $\partial_\mu V^\mu = 0$, which follows from the equations of motion.

Note that there are, confusingly, two uses of "spin" above. The spins of the particles are associated with their transformations under the little group $SO(3)$, embedded in the Poincare group. In this case, the spin 1 particles are created by the part of the field with $\partial_\mu V^\mu = 0$. We can also decompose the field into irreps of spatial rotations, using $SO(3)$ embedded in the Lorentz group. In this case the spin 0 part is simply V^0, while the spin 1 part is V^i. The assignment of spins is in correspondence, but what we mean when we talk about "the spin 1 part of the field" is ambiguous.

Another lesson of this example is that if one is simply given a field-theoretic action, one cannot infer the types of particles present after quantization from the transformation properties of the fields alone. A vector field can end up creating and annihilating either spin 0 or spin 1 particles, depending on the action.

Example. The situation gets more complicated when we take the mass to zero, i.e. when we want to describe a particle with helicity ±1. We need to remove another polarization, but there are no further Lorentz invariant constraints we can impose. If we just start with

$$\epsilon_{\pm1}(p\hat{\mathbf{z}}) \propto \hat{e}_x \pm i\hat{e}_y$$

then under a little group transformation, we will exit the vector space spanned by the $\epsilon_{\pm1}(p\hat{\mathbf{z}})$. The only way to avoid this problem is to impose a gauge symmetry, regarding distinct polarizations

as physically equivalent. This descends to imposing the gauge symmetry $V^\mu \to V^\mu + \partial^\mu \alpha$ on the field, which implies that for the Lagrangian to contain an interaction $V_\mu J^\mu$, the current J^μ must be conserved.

This is the modern way to introduce gauge theories like electromagnetism, but textbooks often present another perspective, which is more historical but somewhat less satisfying. Historically, we figured out Maxwell's equations by experiment, which implies that electromagnetism has a gauge symmetry when written in terms of potentials. The coupling of electromagnetism to matter must respect the gauge symmetry. Thus, textbooks often talk about "introducing" gauge symmetry to a matter sector, with some vague muttering about how this is motivated by "elegance" or "beauty". It isn't! Gauge symmetry is an inconvenience, which is *forced* on us by the mismatch between the mismatch of degrees of freedom of massless particles and fields.

The same logic for helicity ± 2 and a rank 2 tensor field $h_{\mu\nu}$ implies that for direct coupling $h_{\mu\nu} S^{\mu\nu}$, the tensor $S^{\mu\nu}$ must be conserved, so the only option is the stress-energy tensor $T^{\mu\nu}$. For higher helicity, there are no high rank conserved tensors available. One can still write down interactions, but they must involve derivatives, and hence do not produce long-range forces (i.e. forces falling off as $1/r^2$).

Example. A bit more about higher spin fields. Notice that we can also get spin 1 particles from the $(1,0)$ or $(0,1)$ representations of the Lorentz group, which correspond to self-dual and anti self-dual antisymmetric rank 2 tensors. Typically, this is not useful, for the same reason that it isn't useful to use a vector field $\partial_\mu \phi$ to create a scalar particle: the field is just the derivative of a vector field that already does the same job, $F_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$. (In the massless case, using $F_{\mu\nu}$ has the benefit that it is gauge invariant. But we again have the problem that simple interactions in terms of $F_{\mu\nu}$ can't yield, e.g. long range forces, because they are higher order in derivatives, while the usual interactions in terms of A_μ itself would have to be described using nonlocal Lagrangian terms involving inverse derivatives.)

A particle of integer spin s can be embedded in a field transforming in $(s/2, s/2)$, i.e. as a rank s tensor, where the extraneous degrees of freedom can be removed by symmetrizing the tensor and removing all traces. In order to get particles of half-integer spin, it suffices to add a single spinor index. For example, for spin 3/2, we can use a vector and a Dirac spinor index ψ^μ_a, since

$$(1/2,1/2) \times ((1/2,0) + (0,1/2)) = (1/2,1) + (1,1/2) + (1/2,0) + (0,1/2).$$

This is known as a Rarita–Schwinger field. To remove the last two representations, it suffices to impose $\gamma_\mu \psi^\mu = 0$. The first two representations also yield extraneous spin 1/2 particles, which may be removed by imposing $\partial_\mu \psi^\mu = 0$. Finally, just like the Dirac field, the Rarita–Schwinger field obeys a first order equation like the Dirac equation, which avoids double counting the degrees of freedom. To describe massless helicity $\pm 3/2$ particles, one can additionally impose a fermionic gauge symmetry $\psi^\mu_a \to \psi^\mu_a + \partial^\mu \epsilon_a$.

Another approach is to construct everything out of spinor indices from the start. We know that $(j_1/2, 0)$ is the symmetric traceless part of the j_1-fold tensor product of $(1/2, 0)$. Therefore, the representation $(j_1/2, j_2/2)$ can be isolated by taking a field with j_1 and j_2 dotted and undotted Weyl spinor indices, with everything symmetrized and traces removed. We didn't do this in our previous examples because we preferred to use vector indices as much as possible, as they are more intuitive. However, depending on the theory, this approach might ultimately be simpler.

Note. The connection between particles and fields sheds some light on spontaneous symmetry breaking. For example, consider two scalar fields related by an $SO(2)$ symmetry, $\phi_1 = i[Q, \phi_2]$,

where Q is Hermitian and $[Q, H] = 0$. To convert this into a statement about the corresponding particles, we note that at the level of the annihilation operators, we have $a_1 = i[Q, a_2]$. Therefore,

$$|1\rangle = a_1^\dagger|\Omega\rangle = i[Q, a_2^\dagger]|\Omega\rangle = iQ|2\rangle - ia_2^\dagger Q|\Omega\rangle.$$

Therefore, as long as the vacuum state is invariant, $Q|\Omega\rangle = 0$, the one-particle states $|1\rangle$ and $|2\rangle$ are related by the symmetry, and we can conclude, e.g. that they are degenerate. If the symmetry is spontaneously broken, $Q|\Omega\rangle \neq 0$, this logic fails. The symmetry still exists, i.e. Q commutes with the Hamiltonian either way, but it may not yield any useful information about the particles.

It is simple to connect this to the usual way spontaneous symmetry breaking is described. Suppose at least one of the fields gains a vacuum expectation value that is not invariant under the symmetry, $\langle\Omega|[Q, \phi_i]|\Omega\rangle \neq 0$. Since Q is Hermitian, it immediately follows that we must have $Q|\Omega\rangle \neq 0$.

In this case, we typically work with fields $\varphi_i = \phi_i - \langle\Omega|\phi_i|\Omega\rangle$ expanded about the new vacuum. However, the symmetry operator Q acts nonlinearly on these new fields, which means it does not act on the new operators a_i' in a simple way.

In summary, particles live in a unitary Poincare irrep, and the states in these irreps correspond to plane wave solutions $u^\alpha(\mathbf{p}\lambda)e^{ipx}$ for a quantum field, which transforms in a nonunitary Lorentz representation. The field is not strictly necessary, but it allows us to construct theories with locality and causality manifest; the cost is the extra step of converting between particle and field representations with mode functions, as well as the necessity of imposing constraints or gauge symmetries to get the right degrees of freedom in the fields. On the other hand, the on-shell approach to scattering amplitudes proposes to throw away the fields and work with scattering amplitudes directly, precisely to avoid these complications. For much more on these issues, see the notes on Quantum Field Theory.

8.4 ∗ Yang–Mills Theory

In this section, we construct the Yang–Mills Lagrangian. We warm up with a $U(1)$ gauge group.

- The gauge potential is a_μ with gauge transformations $a_\mu \to a_\mu + \partial_\mu\alpha$, and the gauge-invariant field strength is $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$. The Lagrangian is

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

  For convenience, we define $A_\mu = -ia_\mu$ and $F_{\mu\nu} = -if_{\mu\nu}$. This is useful because the Lie algebra $\mathfrak{u}(1)$ may be identified with the imaginary axis $i\mathbb{R}$, so A_μ and $F_{\mu\nu}$ naturally live in it.

- Now consider adding a complex scalar field ϕ with Lagrangian

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

  This is invariant under the $U(1)$ global symmetry

  $$\phi \to g\phi, \quad \phi^* \to g^{-1}\phi^*, \quad g = e^{i\delta} \in U(1).$$

  Now we consider an infinitesimal global symmetry $g = \exp(\epsilon X) \approx 1 + \epsilon X$, where $X \in \mathfrak{u}(1)$ is a pure imaginary number. Then we have

  $$\delta_X\phi = \epsilon X\phi, \quad \delta_X\phi^* = -\epsilon X\phi^*, \quad \delta_X\mathcal{L}_\phi = 0.$$

  To gauge the symmetry, we promote g to $g(x)$.

- The Lagrangian is not invariant under the gauged symmetry because

$$\delta_X(\partial_\mu\phi) = \partial_\mu\delta_X\phi = \epsilon\partial_\mu(X\phi) = \epsilon(\phi\partial_\mu X + X\partial_\mu\phi).$$

  We get an extra term that does not cancel out.

- We restore gauge invariance by promoting the partial derivative to a covariant derivative,

$$D_\mu = \partial_\mu + A_\mu.$$

  We let the gauge field also transform as $\delta_X A_\mu = -\epsilon\partial_\mu X$. Then direct calculation gives

$$\delta_X(D_\mu\phi) = \epsilon X D_\mu\phi.$$

  Then by the same logic as in the global case, $\delta_X\mathcal{L}_\phi = 0$ as desired.

We have constructed the theory of scalar QED above. We now consider a general Lie group G.

- For simplicity, we consider a set of scalar fields ϕ which transform in some representation D of G, with representation space $V \cong \mathbb{C}^N$. We begin with the Lagrangian

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

  where we use the standard inner product on $\mathbb{C}^N$.

- Assuming that D is a unitary representation, $\mathcal{L}_\phi$ is invariant under the global symmetry

$$\phi \to D(g)\phi, \quad D(g)^\dagger = D(g)^{-1}$$

  which is infinitesimally

$$D(g) = \exp(\epsilon R(X)) \approx 1 + \epsilon R(X), \quad R(X)^\dagger = -R(X), \quad \delta_X\phi = \epsilon R(X)\phi$$

  where $R(X)$ is the representation of $\mathfrak{g}$ corresponding to D.

- Next, we gauge the symmetry by allowing X to depend on x and using the covariant derivative

$$D_\mu\phi = \partial_\mu\phi + R(A_\mu)\phi$$

  where A_μ is a $\mathfrak{g}$-valued vector field, which transforms as

$$\delta_X A_\mu = -\epsilon D_\mu X = -\epsilon\partial_\mu X - \epsilon[A_\mu, X]$$

  where the second term is new; it is the Lie bracket, which vanished in the $\mathfrak{u}(1)$ case.

- By direct calculation, we may verify

$$\delta_X(D_\mu\phi) = \epsilon R(X) D_\mu\phi$$

  where we use the fact that R is linear, so it commutes with derivatives, and that R is a representation, so $R([X, A_\mu]) = [R(X), R(A_\mu)]$. Then we have

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

  since $R(X)$ is anti-Hermitian.

- Next, we need a kinetic term for the gauge field. We define the field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu] \in \mathfrak{g}.$$

  Then by direct calculation, we have

$$\delta_X F_{\mu\nu} = \epsilon[X, F_{\mu\nu}]$$

  where we use standard properties and the Jacobi identity.

- We then use the Killing form to define a kinetic term,

$$\mathcal{L}_A = \frac{1}{g^2}\kappa(F_{\mu\nu}, F^{\mu\nu}).$$

  This term is gauge-invariant by the invariance of the Killing form,

$$\delta_X \mathcal{L}_A \propto \kappa([X, F_{\mu\nu}], F^{\mu\nu}) + \kappa(F_{\mu\nu}, [X, F^{\mu\nu}]) = 0.$$

- To get a sensible theory, we need the Killing form to be negative definite, i.e. that $\mathfrak{g}$ is of compact type; this also requires $\mathfrak{g}$ to be simple. That is, we can choose a basis of $\mathfrak{g}$ where

$$\kappa^{ab} = \kappa(T^a, T^b) = -\kappa\delta^{ab}, \quad \kappa > 0$$

  so that in components,

$$\mathcal{L}_A = -\frac{\kappa}{g^2} F_{\mu\nu a} F^{\mu\nu a}, \quad F_{\mu\nu} = F_{\mu\nu a} T^a.$$

- In summary, a gauge theory is specified by a simple compact gauge group G, where $A_\mu \in \mathfrak{g}$ (i.e. the gauge field transforms in the adjoint representation), and a set of matter fields that transform in representations of $\mathfrak{g}$. In nature, the matter fields are always fermions that transform in the fundamental representation.

Note. Some more explicit expressions in the case where G is a matrix Lie group, so that $A_\mu \in \mathfrak{g}$ and $g \in G$ are matrices. The gauge field transforms as

$$A_\mu \to A'_\mu = gA_\mu g^{-1} - (\partial_\mu g)g^{-1}$$

which yields the same $\delta_X A_\mu$ by letting $g = \exp(\epsilon X)$.

A field in the fundamental representation transforms as $\phi \to g\phi$, so its covariant derivative is simply $D^F_\mu \phi = \partial_\mu \phi + A_\mu \phi$. On the other hand, a matrix-valued field in the adjoint representation transforms as $\phi \to g\phi g^{-1}$, so its covariant derivative is $D^A_\mu \psi = \partial_\mu \psi + [A_\mu, \psi]$. These are both special cases of our general expression above.

The field strength tensor is the commutator of covariant derivatives,

$$F_{\mu\nu} = [D^A_\mu, D^A_\nu]$$

as can be checked by acting with both sides on ψ. Directly expanding shows that this matches our earlier definition. Moreover, $F_{\mu\nu}$ itself transforms in the adjoint representation, which implies that $\mathrm{tr}(F_{\mu\nu}F^{\mu\nu})$ is gauge invariant. This is equivalent to what we found earlier, since $\mathrm{tr}(XY)$ and $\kappa(X, Y)$ are proportional for a simple Lie algebra.

Note. For reference, the one-loop running of the Yang–Mills coupling g at one-loop is

$$\frac{d\alpha}{d\log\mu^2} = \frac{b\alpha^2}{4\pi}, \quad \alpha = \frac{g^2}{4\pi}$$

where

$$b = -\frac{11}{3}C_2(G) + \sum_F \frac{4}{3}T(F) + \sum_S \frac{1}{3}T(S)$$

where the terms are due to the gauge bosons, Dirac fermions, and complex scalars. We are using the physics normalization convention, which implies

$$C_2(SU(N)) = N, \quad T(\text{fund}) = \frac{1}{2}.$$

In the case of $G = U(1)$, we have

$$C_2(U(1)) = 0, \quad T(\text{charge } q) = q^2.$$

If Weyl or Majorana fermions are present, they contribute with a coefficient of $2/3$ because they're half a Dirac fermion, while real scalars contribute with a coefficient of $1/6$.

8.5 Grand Unification

We begin with a brief review of the gauge groups and matter content of the Standard Model.

- The Standard Model has gauge group $SU(3) \times SU(2) \times U(1)$. The gauge bosons in each of these factors carry color, weak isospin, and hypercharge Y. The fermions transform in representations of this group; we will indicate representations of $SU(3)$ and $SU(2)$ by their dimensions are usual, and representations of $U(1)$ by half their hypercharge, $Y/2$.
- The structure is the same in each generation, so we focus on the first, containing the quarks u and d, the electron e^-, and the neutrino ν_e. The quarks and electron are Dirac fermions, so they contain left-handed and right-handed Weyl fermions, while the massless neutrino is only left-handed.
  - The left-handed quarks transform in $(3, 2, 1/6)$. Counting quarks of different colors as different particles, there are six; the weak force exchanges the up and down quarks.
  - The right-handed up quarks transform in $(3, 1, 2/3)$.
  - The right-handed down quarks transform in $(3, 1, -1/3)$.
  - The left-handed leptons transform in $(1, 2, -1/2)$.
  - The right-handed electron transforms in $(1, 1, -1)$. It has to transform trivially under $SU(2)$ since there's nothing left; there is no right-handed neutrino.
  - All of the antiparticles transform in the corresponding conjugate representation.
- To build a grand unified theory, we would like to combine these representations together. Since gauge transformations commute with Lorentz transformations, they can only work within one Lorentz irrep, so it is convenient to only work with left-handed Weyl spinors. Since the conjugate of a left-handed Weyl spinor is a right-handed Weyl spinor, we have

$$u \text{ and } d \text{ quarks} : (3, 2, 1/6), \quad u^c \text{ quark} : (\overline{3}, 1, -2/3), \quad d^c \text{ quark} : (\overline{3}, 1, 1/3)$$

and

$$e \text{ and } \nu_e \text{ leptons} : (1, 2, -1/2), \quad e^c \text{ lepton} : (1, 1, 1)$$

where the c superscript denotes charge conjugation, so u^c is an anti-up quark.

- The $SU(2) \times U(1)$ symmetry is spontaneously broken down to the $U(1)$ of electromagnetism; the photon is a mixture of the $U(1)$ generator and one of the $SU(2)$ generators, say the third. The explicit relation is given by a modern version of the Gell-Mann–Nishijima formula,

$$Q = T^3 + \frac{1}{2}Y.$$

  Using this relation, we can reverse engineer the values of Y above, where T^3 is the third generator of $SU(2)_L$. The Y here is totally different from the hypercharge in the Gell-Mann–Nishijima formula, $Q = I^3 + Y/2$ where I^3 is the third component of isospin, and was normalized so the two formulas look the same.

- The naive Dirac mass terms don't work, because left-handed and right-handed quarks and electrons transform differently under $SU(2)$. The resolution is to introduce a new complex scalar field φ transforming in $(1, 2, -1/2)$, called the Higgs field; the hypercharge is chosen so that φ times the Dirac mass term is allowed.

- As a result, the Higgs field contains two uncharged particles, a particle with charge $+1$, and a particle with charge -1. During symmetry breaking, three of these particles combine with the broken $SU(2)$ and $U(1)$ gauge bosons to produce the $W^\pm$ and Z, while the remaining degree of freedom acquires a vev, becoming 'the' Higgs field.

We now turn to grand unification under $SU(5)$, first proposed by Georgi and Glashow in 1972.

- One clue that the Standard Model gauge group should be part of a larger one is that the hypercharges of the 15 Weyl fields above add up to zero. This is what we would expect if hypercharge were one of the generators of a larger gauge group, as it should be traceless.

- The smallest group that contains the Standard Model gauge group is $SU(5)$, generated by traceless anti-Hermitian 5×5 matrices. We identify $SU(3)$ with generators in the upper-left 3×3 block, $SU(2)$ with generators in the bottom-right 2×2 block, and $U(1)$ with generator

$$\frac{1}{2}Y = \text{diag}(-1/3, -1/3, -1/3, 1/2, 1/2).$$

  There are 12 more gauge bosons, which we'll get back to later.

- First, consider the fundamental representation ψ^μ of $SU(5)$. This irrep decomposes as

$$5 \to (3, 1, -1/3) + (1, 2, 1/2)$$

  which implies that

$$\overline{5} \to (\overline{3}, 1, 1/3) + (1, 2, -1/2).$$

  This perfectly accommodates the anti-down quark and the leptons.

- The remaining 10 particles fit into the antisymmetric tensor representation $\psi^{\mu\nu}$. To understand this representation, it's easiest to break the symmetry first; then we want the antisymmetric part of $(3,1,-1/3)+(1,2,1/2)$ times itself. We have

$$(3,1,-1/3)\times(3,1,-1/3)=(\overline{3},1,-2/3)+(6,1,-1/3)$$

and

$$(3,1,-1/3)\times(1,2,1/2)=(3,2,1/6),\quad (1,2,1/2)\times(1,2,1/2)=(1,1,1)+(1,3,1).$$

Taking the antisymmetric part eliminates the $(6,1,-1/3)$ and $(1,3,1)$. It also ensures we only get one copy of $(3,2,1/6)$ though we would naively have two. Then

$$10\to(\overline{3},1,-2/3)+(3,2,1/6)+(1,1,1)$$

which is exactly the anti-up quark, the quarks, and the anti-electron.

- Therefore, the matter content of the Standard Model fits into the $\overline{5}+10$ of $SU(5)$. Taking the reasoning in reverse, this explains many features of the Standard Model; it forces charge to be quantized, and it ensures the proton and electron charge are exactly opposite.

Next, we give a brief taste of dynamics in the $SU(5)$ GUT.

- We write mass terms using the ψ_μ and $\psi^{\mu\nu}$ fields. We let

$$\mu,\nu=1,2,3,4,5,\quad \alpha=1,2,3,\quad i=4,5.$$

There are spinor indices everywhere, which we suppress. In terms of the usual particles, $\psi_\alpha=\overline{d}$, $\psi_i=(\nu,e)$, $\psi^{\alpha\beta}=\overline{u}$, $\psi^{\alpha i}=(d,u)$, and $\psi^{ij}=\overline{e}$.

- In terms of representations, we can write a Dirac mass term for the up quark if $u^c(Cu)$ contains a singlet, where C is charge conjugation. But we run into the same problem as in the Standard Model: it's impossible to do this. Instead we introduce a Higgs field φ^μ transforming in the 5, with symmetry breaking so that φ^4 acquires a vev, and mass term

$$\epsilon_{\mu\nu\rho\sigma\tau}\psi^{\mu\nu}C\psi^{\rho\sigma}\varphi^\tau\to\epsilon^{\mu\nu\rho\sigma 4}\psi^{\mu\nu}C\psi^{\rho\sigma}\varphi^4\sim\psi^{12}C\psi^{35}+\text{permutations}\sim u^cCu.$$

We choose φ^4 to have the vev since it is electrically neutral.

- For the down quark and electron, we introduce the mass term

$$\psi_\mu C\psi^{\mu\nu}\varphi_\nu\to\psi_\mu C\psi^{\mu 4}\sim d^cCd+e^cCe$$

with no mass term for the neutrino since $\psi^{44}=0$, as desired.

- The gauge bosons are in the adjoint representation and hence can be written as components of a traceless tensor A^μ_ν. Then the 12 gauge bosons we haven't accounted for, called the X and Y bosons, mix quarks and leptons because they exist in the same $SU(5)$ irrep. (Specifically, the X and Y mix among themselves under $SU(3)_C$, which is why we distinguish them.)

- For example, we have
$$\psi^{\alpha 4} = d, \quad A^5_\alpha \psi^{\alpha 4} \sim \psi^{54} = e^+.$$
Similarly, we can convert an up quark into an anti-up quark,
$$\psi^{5\alpha} = u, \quad A^5_\alpha \psi^{5\alpha} \sim \psi^{\alpha\beta} = u^c.$$
Therefore, proton decay can occur by emission and reabsorption of an A^5_α boson,
$$p = u + u + d \to u^c + e^+ + d = \pi^0 + e^+.$$
This process occurs very slowly due to high mass of the A^5_α, at the GUT scale.

- Proton decay can also be described below the GUT scale by an effective field theory. In the Standard Model, we have the accidental global symmetries of quark rotation and lepton rotation, yielding conservation of baryon and lepton number B and L. Then to add proton decay, we simply write down terms that don't obey these global symmetries.

- Since we are far below the GUT scale, any terms we add should be scalars under the Standard Model gauge group. Then a qqq term is unacceptable since it has nonzero hypercharge, but a dimension 6 $qqq\ell$ term turns out to be allowed. This term describes proton decay, and we can relate its rate to the rates of other exotic processes just as we did for isospin.

- We notice that in the proton decay process, B and L change, but $B - L$ is conserved. To see why, note that there are only two mass terms but three irreps in play. Therefore we can construct a global symmetry of the Lagrangian, i.e. a conserved quantity X. We have
$$X(10) + X(10) + X(5_\varphi) = X(\bar{5}) + X(10) + X(\bar{5}_\varphi) = 0$$
which implies
$$X(10) = 1, \quad X(5_\varphi) = -2, \quad X(5^*) = -3.$$
However, this symmetry is broken by the Higgs vev, φ_4. Note that φ_4 has hypercharge $Y/2 = -1/2$ and $X = 2$, with some sign flips since the index is lowered. Then neither X nor Y is conserved, but $X + 4(Y/2) = B - L$ is. In more general GUTs, $B - L$ might be violated.

- At the unification scale $M_{\text{GUT}} \approx 2 \times 10^{16}\,\text{GeV}$, we have $\alpha(\mu) \approx 1/25$, so the theory remains perturbative. Note that for comparison, the reduced Planck mass is about $2 \times 10^{18}\,\text{GeV}$, so we can hopefully ignore quantum gravity effects at the GUT scale.

- We expect the most common proton decay mode to be $p \to \pi^0 e^+$, and a rough estimate of the decay rate is
$$\Gamma \sim \frac{g(M_{\text{GUT}})^4}{M_{\text{GUT}}^4} m_p^5$$
where the prefactor is due to the two vertices and the X/Y gauge boson propagator squared, and the m_p^5 factor is the phase space factor. This is raised to the fifth power by dimensional analysis, and m_π and m_e don't enter because they are much smaller than m_p. Numerically,
$$\tau_{p \to \pi^0 e^+} \sim 10^{31}\,\text{years}.$$

Note. Electroweak unification is not the same as grand unification! Grand unification combines three forces into one simple gauge group, where they necessarily have one coupling constant. The electroweak theory instead describes the breaking of $SU(2)_L \times U(1)_Y$ to $U(1)_{\text{EM}}$. Since the unbroken group is a product group, there are still two independent coupling constants above the electroweak scale. The feature they have in common is a Higgs symmetry breaking.

Note. More about picking the Higgs fields irreps for giving mass to matter. Given a left-handed particle field in a representation R and a left-handed antiparticle field in a representation R', we can write a Higgs Yukawa coupling if the Higgs field irrep or its conjugate is in $R \otimes R'$.

Using a notation where, e.g. $[m, n]$ stands for a Young tableau with a column of m and a column of n, for the down quark and lepton we have

$$\overline{5} \times 10 = [4] \times [2] = [1] + [4, 2] = 5 + 45$$

while for the up quark we have

$$10 \times 10 = [2] \times [2] = [4] + [3, 1] + [2, 2] = \overline{5} + \overline{45} + \overline{50}.$$

For simplicity, we'd like to minimize the number of distinct Higgs fields, so we need to take either 5 or 45. We'd like our Higgs to contain the Standard Model Higgs which breaks $SU(2) \times U(1)$ to $U(1)$. Both 5 and 45 work for this purpose, and we took 5 above for simplicity.

The $SU(5)$ GUT comes with a variety of complications.

- Though the strong, weak, and electromagnetic coupling constants do get close at a high energy scale, they don't converge exactly. However, the agreement is significantly improved by adding superpartners. This agreement has had a strong historical impression on the community, with some calling it the best BSM prediction ever made.

- Experiments such as Super-K have placed a stringent upper bound on the proton decay rate of $\tau \gtrsim 2 \times 10^{34}$ years, beyond what would be natural for the $SU(5)$ GUT. These experiments work by taking a large tank of water and looking for the gamma rays produced in the decay.

- We need *two* Higgs fields. The first is in the adjoint and acquires the vev $\text{diag}(2, 2, 2, -3, -3)$, breaking $SU(5)$ to $SU(3) \times SU(2) \times U(1)$ and giving mass to the extra gauge bosons. The fermion masses come from the second Higgs field φ^μ as discussed above.

- Our mass term sets $m_d = m_e$ at the GUT scale. This is modified by RG flow, which gives

  $$\frac{m_b}{m_\tau} \approx \frac{m_s}{m_\mu} \approx \frac{m_d}{m_e} \approx 3.$$

  This is acceptably accurate for the last two generations, but totally wrong for the first generation. One can fix this by adding a third Higgs field, but this starts to make the theory complicated.

- Since the GUT scale is so high, the hierarchy problem remains, due to the distance between the GUT scale and the electroweak scale. There is nothing that protects the SM Higgs from getting mass corrections up to the GUT scale.

- This question is sharpened by looking more closely at the Higgs sector. The SM Higgs is contained in a 5, which also contains a color-triplet Higgs. This color-triplet can mediate proton

decay, so it must be very heavy, near the GUT scale. A light color-triplet Higgs would also ruin gauge coupling unification. So we need a mechanism to make the Higgs doublet very light and simultaneously make the Higgs triplet very heavy. This doesn't hold even at tree level, since the vev of the 24 Higgs field generically makes both parts of the 5 Higgs field heavy, and moreover the separation is not stable under radiative corrections. This is the doublet-triplet splitting problem.

- There is no explanation for the family structure of the SM. More complicated GUT models can account for this, but in all cases the fact that there are three generations seems to be put in "by hand", in the sense that similar, equally elegant theories could have four or five generations. The number of new heavy particles also tends to explode.
- One could quibble about the proton decay measurement. For example, the neutron decays, but becomes stable in atoms because of its nuclear environment; however, this is very unlikely for the proton because its decay is strongly kinematically allowed. Another way out is to note that the matching of leptons to quarks in each generation is in principle arbitrary, so if the e^+ were not in the same generation of u and d quarks, the decay would be suppressed. However, this would make the relations between lepton and quark masses within each generation extremely far off, which would have to be explained.

Note. The basic $SU(5)$ GUT was invented in 1973, and supersymmetric GUTs burst onto the scene in the 1980s, leading to a flurry of thousands of papers. (See this lecture for a fascinating account of the early history.) There was a bewildering array of grand unified theories produced. The simplest next example is $SO(10)$. Here, an entire generation of the Standard Model can be placed in a single irrep 16, which decomposes as

$$16 \to 10 + \overline{5} + 1$$

under $SU(5)$. The singlet 1 may be interpreted as a right-handed sterile neutrino. Somewhat more complicated is E_6, where we can fit a generation into the representation 27. This is the smallest representation, and decomposes as

$$27 \to 16 + 10 + 1$$

under $SO(10)$. The 10 and 1 must be given a high mass by symmetry breaking to match experiment. In the case of E_6, the simplest possible Higgs representation is the adjoint 78. Furthermore, using extended Dynkin diagrams, we can see that $SU(3) \times SU(3) \times SU(3) \subseteq E_6$, so another theory can be constructed with gauge group $SU(3)^3$. In this theory, called "trinification", each Standard Model gauge group comes out of one $SU(3)$ factor. While the gauge group is not simple, it shares many properties with other grand unified theories.

In the early 2000s, as excitement for the LHC built up, model building using group theory reached its apex. In addition to the mundane $SU(5)$, $SO(10)$, and E_6 SUSY GUTs, there were left-right symmetric models, Pati–Salam models (featuring lepton number as the "fourth color"), flipped $SU(5)$, supergravity theories, and family unification models, augmented with a wide variety of Higgs sectors (possibly composite), supersymmetry breaking mechanisms, extra dimensions, orbifolds, and branes. These sharp tools were used to address many subtle problems which are unknown among students today, such as the LEP paradox, the μ problem, the doublet-triplet splitting problem, and even the "gluino sucks" problem.

The products of this golden age had one thing in common: a belief that it was possible to guess *every* particle that existed, reaching the end of physics in a single paper. When the LHC turned on,

all of the models which made sharp predictions were brutally slain by reality. Today, they stand as a testament to the creativity and industry of a bygone era. We look upon their works, and despair.