Chapter 1 Groups

1.1 Definition of a Group

Let us ask what are the general properties which a set of symmetries has.

a) Suppose A and B are operators on the Hilbert space of wavefunctions, and they are symmetries of the Hamiltonian. Then the operation which consists of first applying B and then applying A, called the composition of A and B and written $A \circ B$, is also a symmetry. Thus the product of two elements of the set is also in the set.

b) Clearly if we do nothing, the Hamiltonian is unchanged. Thus the do-nothing, or identity operator, is always in the set.

c) The operations we call symmetries have to preserve the norm of the states, so they are unitary operators on the Hilbert space and therefore have an inverse. If A left the Hamiltonian unchanged, clearly A^{-1}, which brings it back to what it was, must leave it unchanged. Thus the inverse of every element of the set exists and is in the set.

These properties are essentially the definition of a group, except that we need to add one property we have missed, which is that transformations which map any space into itself are associative under composition, A ◦ (B ◦ C) = (A ◦ B)◦ C, because however you place the parentheses, either side means apply C, then B, and then A.

Putting this together, we get the **Definition:** A group G is a set of distinct elements A, B, ... together with a multiplication law A ⊙ B or simply AB, such that

a) If $A \in G$ and $B \in G$, then $AB \in G$. This is called **closure**, the set is closed under group multiplication.

b) For all $A, B, C \in G$, $A(BC) = (AB)C$. This is called **associativity**.

c) $\exists e \in G$ such that $\forall A \in G, eA = Ae = A$. e is called the **identity**, and can also be written as 1 or $\mathbb{1}$. [3]

d) For each $A \in G$, there is a $B \in G$ such that $AB = BA = e$. B is called the inverse of A and is written A^{-1}.

Note that the multiplication law is not necessarily ordinary multiplication.

Some examples of groups are

Set	Multiplication Law	Name
all integers	addition	$\mathbb{Z}$
integers $0, 1, ...p-1$	addition mod p	$\mathbb{Z}_p$
rotations in three dimensions	composition	SO(3)
rotations in two dimensions	composition	$\mathrm{SO}(2) = \mathrm{U}(1)$
symmetries of a square	composition	C_{4v} or D_4

Some examples on non-groups are

non-negative integers	addition	inverses do not exist
rotations through angles $< \pi/2$	composition	not closed

Definition: A **subgroup** H of a group G is a subset of elements $H \subseteq G$, having the same multiplication law as G, which is closed under multiplication and forms a group by itself.

Two groups G and G′ are **isomorphic** if there exists a bijective map $f : G \to G'$ which respects multiplication, $f(a)\, f(b) = f(ab)$ for all a and b in G.

By **bijective map** we mean that f is one-to-one and onto, or in other words no two distinct elements in G are mapped into the same element in G′, and every element in G′ is the image of some element in G.

Without the requirement that f be either one-to-one or onto, but still requiring that $f(a)\, f(b) = f(ab)$, we say f is a **homomorphism** from G into G′

The symbol representing "is isomorphic to" is $\cong$, so $G_1 \cong G_2$ means G_1 is isomorphic to G_2.

Note isomorphism is an equivalence relation (to be defined shortly), so that we usually think of a group as a representative of all groups isomorphic to it.

1.3 Conjugacy

Two elements B and C of a group G are said to be **conjugate** if there exists an element $A \in G$ such that

$$A^{-1}BA = C.$$

Conjugacy is easily shown to be an equivalence relation.

A **relation** on a set S is a boolean valued function on S x S.

That is to say, for each ordered pair it is either true or not.

It is usually stated differently: the relation A ~ B holds for certain pairs of elements A and B in G, and not for some other pairs of elements.

A relation $\cong$ is an **equivalence relation** if

1. it is **reflexive**. That is, $A \cong A$ for every $A \in G$.

2. it is **symmetric**. That is, if $A \cong B$ then $B \cong A$, for any A and B in G.

3. it is **transitive**. That is, if $A \cong B$ and $B \cong C$, then $A \cong C$, for A, B, and C in G.

Any equivalence relation on a set divides the set up into disjoint **equivalence classe**s, all of the elements in one equivalence class being equivalent to each other and not to any of those in other classes.

As conjugacy in a group G is an equivalence relation, it divides G into **conjugacy classes**.

This division plays an important role in the analysis of representations of a finite group.

Any element B which commutes with all other elements of the group G is clearly in a conjugacy class by itself, for then $A^{-1}BA = B$ for any A.

The set of all such elements is called the **center** of the group.

In particular, an Abelian group of order n has n conjugacy classes, each with one element, and then the center is the whole group.

Also, for any group, the identity element is always in a class by itself.

1.4 Cosets

If a group G has a subgroup H, there is another kind of equivalence relation

$$g_1 \cong g_2 \text{ if there is an element } h \in H \ni g_1 h = g_2.$$

H can't just be any subset; the requirements for the relation to be an equivalence relation require H to be a group.

The equivalence class of g is just the set

$$gH := \{gh | h \in H\}$$

which is the set of elements gh for all h ∈ H.

It is called the **left coset** of H in G with respect to g.

We can also define right cosets analogously—they are not necessarily the same.

Note that each coset has the same number of elements as H, and the whole group is the disjoint union of the cosets.

The number of elements in a group is called the **order** of the group, and we have just shown that the order of G is equal to the order of H multiplied by the number of cosets of H in G.

In particular, the order of G is divisible by the order of H.

We define the **index** of H in G to be the number of cosets of H in G, which we see is given, for finite groups, by

$$\text{index of } H \text{ in } G = \text{order}(G)/\text{order}(H).$$

1.5 Cyclic Group

Any nonidentity element g of a finite group generates a sequence {𝕀, g, g^2 , ...} which must eventually repeat an element.

The first one repeated must be 𝕀, so we can take the smallest n > 0 for which g_n = 𝕀, and call it the **order** of the element g in G.

The set H= {𝕀, g, g^2, ...g^{n-1}} is a subgroup of order n.

Such groups, which are constructed of powers of a single generator, are called **cyclic** groups, and may be written as < g >.

Every nonidentity element of order n generates a cyclic subgroup, so there is a subgroup of order n, and thus the order of each element divides the order of the group G.

1.6 Normal Subgroups

H is called a **normal subgroup** of G (H $\triangleleft$ G) if the left and right cosets are the same for every element $x \in G$, or

$$xH = Hx \quad \text{for all } x \in G.$$

This does not imply that each element $h \in H$ commutes with x, but rather that for each h and x there is an h' with $xh' = hx$, or $h' = x^{-1}hx$, with $h' \in H$.

As an example, consider $G = D_4 = C_{4v}$.

The subgroup C_4 generated by C is a normal subgroup, as shown below.

The subgroup H= {𝕀, m_x} is not normal ($H \not\triangleleft G$), because CH= {C, Cm_x}= {C, σ_+}, while HC = {C, m_xC}= {C, σ_-}.

If H is a normal subgroup of G, the cosets of H in G may be considered elements of another group, the **quotient group** G/H, by defining the multiplication law

$$xH \odot yH = xyH.$$

This looks like it might work for any subgroup H, normal or not, but it doesn't, because the coset xH does not uniquely specify an element $x \in G$, and the definition of $\odot$ above only works if xyH is independent of this ambiguity.

For a normal subgroup Hy= yH, so the set of elements xHyH (the set is independent of the ambiguity) is xyHH= xyH.

1.7 Direct Products

Let H and K be two groups of orders η and κ respectively.

Then the set H x K is the set of pairs of elements, one from H and one from K, which we can write as $\{(h, k), h \in H, k \in K\}$.

There is no particular meaning to calling this multiplication — it is just notation for an ordered pair of elements, one from each set. Notice that H x K has $\eta\kappa$ elements.

Now we can turn H x K into a group by defining the group multiplication law

$$(h_1, k_1) \odot (h_2, k_2) := (h_1 \odot h_2, k_1 \odot k_2),$$

where the two $\odot$'s on the right are the ones for H and for K respectively.

It is easy to show that H x K, which is called the **direct product** of H and K, is a group, with identity $(\mathbb{I}, \mathbb{I})$, and with the inverse of (h, k) given by (h^{-1}, k^{-1}).

The two groups act independently.

For example, in elementary particle physics there is invariance under the rotation group SO(3) in space, and also under the color group SU(3).

These act completely independently, so one barely mentions that there is a SO(3) x SU (3) symmetry.

However, in electro-weak theory, there is a symmetry under weak isospin and another U(1) group, SU(2) x U (1), which is often mentioned like this because a subgroup, which is not a direct product of two groups, is the one which survives the spontaneous symmetry breaking.

If G = H x K, it contains normal subgroups $\{(h, \mathbb{I})|h \in H\}$ and $\{(\mathbb{I}, k)|k \in K\}$, which loosely speaking we could just call H and K.

So we can form the quotient group G/H and it is clearly isomorphic to K, $K \cong G/H$, and also $H \cong G/K$.

But this is not true for a general group G with a normal subgroup H.

1.8 Kernel

Let G and K be groups and $f : G \to K$ be a homomorphism with **kernel** H, i.e.

$$H = \{h \in G \ni f(h) = \mathbb{I}_K\}.$$

Then H is a normal subgroup of G, $\ker(f) \triangleleft G$.

It is easy to see that H is closed under multiplication and contains the identity of G, and is therefore not empty.

To show that it is normal, we need only show $g^{-1}hg \in H$ for all $g \in G, h \in H$.

But $f(g^{-1}hg) = f(g^{-1})f(h)f(g) = f(g^{-1})f(g) = \mathbb{I}_K$, because a homomorphism preserves products and inverses and the identity.

Thus $g^{-1}hg$ is in H, and $H \triangleleft G$.

The notion of kernels, and equivalencing by ignoring things in the kernel, is a common notion with many applications in mathematics.

One physics application will come much later, when we consider spontaneous symmetry breaking.

There the full dynamical symmetry group does not leave the vacuum state invariant, but the map from the group into the variations of the vacuum state has a kernel, the symmetries which are not broken, which is a normal subgroup, and the quotient group is generated by the massless Goldstone bosons.

1.9 Permutations

Consider n boxes labelled 1, 2, ..., n and n objects a, b, ... which fit one to each box.

Represent the composite action of moving the object in the i^{th} box to the r_i^{th} box (all simultaneously) by

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & n \\ r_1 & r_2 & \dots & r_i & \dots & r_n \end{pmatrix}.$$

The r_i's are just the integers 1, ..., n but in some different order, with n! possibilities.

These composite movements are called **permutations** of n objects, and, for a given n, form a group called the **symmetric group** Sn.

theLet us specialize to S_3.

Short Name = Full Name	Action	Effect on $+ - *$
$\mathbb{1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$	Identity	$+ - *$
$(12) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$	interchange 1 and 2	$- + *$
$(13) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$	interchange 1 and 3	$* - +$
$(23) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$	interchange 2 and 3	$+ * -$
$(123) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$	$1 \to 2 \to 3 \to 1$	$* + -$
$(132) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$	$1 \to 3 \to 2 \to 1$	$- * +$

The short names we have given are called cycles, and $(n_1, n_2, ..., n_r)$ might have been better written as $(n_1 \to n_2 \to \ldots \to n_r \to n_1)$,

Thus (123) = (231), etc.

All permutations in general can be written as products of disjoint cycles, but not necessarily as a single cycle as we can for S_3, which doesn't have enough elements to have two nontrivial disjoint cycles.

It is also possible to write all permutations in terms of a product of **transpositions**, which are cycles of two elements, or simple interchanges, but these are not disjoint.

For example, (123) = (13)(12).

The sign of a permutation is ±1, depending on whether it can be written as a product of an even or odd number of transpositions.

Although the decomposition of a permutation into transpositions is not unique, the sign is well defined.

Otherwise there could be no fermions!

The sign of the permutation P is often written $(-1)^P$, which has no other intrinsic meaning.

The even permutations, i.e., those with positive sign, obviously form a subgroup A_n, called the **alternating group**.

1.10 Groups of low order

Every group needs to include the identity.

Thus the smallest group is the only group of order one, {𝕀}.

Any other element g of a finite group generates a cyclic subgroup of order the same as the order of g.

If there are no other elements, the group is the cyclic group of order n, Z_n, which can be thought of as the integers {0, 1, ..., n-1} with ⊙ taken to be addition mod n, or as the complex numbers of unit magnitude $z_k = e^{2\pi ik/n}$, with ⊙ ordinary complex multiplication.

Thus for every integer n > 0 there is at least one group of order n.

If n is a prime, the order of any non-identity element must be n because it must divide n.

Then it generates the full group, and there is nothing unresolved in its multiplication table.

Thus there is only one group of order p for each prime p.

Thus there is one group of order two and one of order three.

For order four, in addition to the cyclic group Z_4, we can look for a group with no element of order 4.

As the orders must divide 4, we must have three elements, A, B, C, each of order two, so each has square 1.

What is AB?

It can't be 𝕀 because A ≠ B, and it can't be either of A or B, as the other is not the identity, so it must be C.

The same argument determines that BA is neither 𝕀 nor A nor B, so it must also be C, AB = BA.

As A and B were arbitrary non-identity elements, the group is Abelian and the multiplication fully determined.

This group is called the **Vierergruppe**.

It is the symmetry group of a rectangle, the subgroup of the symmetry group of the square, D4,

For order 6 we already have the cyclic group Z_6 and the permutation group S_3. T

These are all groups of order 6.

Part 2 Representations

In the last part we learned something about the structure of the groups themselves, but we are really interested in how the symmetry groups act on the eigenstates of a physical Hamiltonian.

The set of states corresponding to a given eigenvalue of degeneracy ℓ forms a vector space of dimension ℓ.

The result of applying the symmetry operation to any of these states gives a state with the same energy eigenvalue, and thus a state in the same ℓ-dimensional vector space.

The symmetry acts linearly on the states,

$$S(\alpha\,|\psi_1\rangle + \beta\,|\psi_2\rangle) = \alpha S(|\psi_1\rangle) + \beta S(|\psi_2\rangle),$$

so the operators are linear operators on a vector space, and may be considered ℓ x ℓ matrices acting on this eigenspace.

Let $\{e_i, i = 1, ..., \ell\}$ be an orthonormal basis of the vector space.

If $A \in G$ is one of the symmetry operators, it will act on e_i to give a linear combination of the basis vectors,

$$Ae_i = \sum_j e_j T_{ji}(A).$$

Notice the order of the indices on T .

If $\psi = \sum_i \psi_i e_i$ is an arbitrary vector in the eigenspace, and if $\psi' = \sum_i \psi'_i e_i = A\,\psi$, then

$$\psi' = \sum \psi'_i e_i = A \sum \psi_i e_i = \sum \psi_i A e_i = \sum_{ij} \psi_i T_{ji}(A) e_j,$$

or

$$\psi'_j = \sum_k T_{jk}(A) \psi_k,$$

by the linear independence of the e_j's.

Note the order of the indices now.

Tij can be considered the matrix element $T_{ij}(A) = \langle e_i \mid A \mid e_j \rangle$.

We can find the composition law for these matrices by noting that

$$\begin{aligned} ABe_i &= \sum_k e_k T_{ki}(AB) = A(Be_i) = A\Big(\sum_j e_j T_{ji}(B)\Big) = \sum_j T_{ji}(B) A e_j \\ &= \sum_j T_{ji}(B) \sum_k T_{kj}(A) e_k, \end{aligned}$$

$$\text{so} \qquad T_{ki}(AB) = \sum_j T_{kj}(A) T_{ji}(B),$$

$$\text{or} \qquad T(AB) = T(A)T(B),$$

where the multiplication is just ordinary matrix multiplication.

Thus we have a homomorphism from the group G onto a group of ℓ x ℓ nonsingular matrices.

This is called an ℓ-**dimensional representation** of the group G.

If no two elements of G are mapped into the same matrix, then the homomorphism
$G \rightarrow T = \{T(A) | A \in G\}$ is an isomorphism, $G \cong T$, and we call the representation **faithful**.

An example of an unfaithful representation for any group is the homomorphism $g \in G \longmapsto \mathbb{I}$, in which every element is mapped into the identity matrix.

This trivial representation is called the **identity representation** of the group.

Note that for any representation

$$T(\mathbb{I})T(A) = T(\mathbb{I}A) = T(A), \text{ so } T(\mathbb{I}) = \mathbb{I}_{\ell \times \ell},$$

as T(A) is nonsingular.

Also

$$T(A^{-1})T(A) = T(A^{-1}A) = T(\mathbb{I}) = \mathbb{I}, \text{ so } T(A^{-1}) = [T(A)]^{-1}.$$

If R is an ℓ-dimensional representation of G, and if S is a fixed nonsingular ℓ x ℓ matrix, then

$$T(A) := S^{-1}R(A)S \text{ for all } A \in G,$$

is also clearly an ℓ-dimensional representations of G.

Two representations related by such a fixed similarity transformation are **equivalent representations**.

Consider an n-dimensional representation T of G on a vector space L_n.

Suppose we take a particular vector v and act on it with all the elements T(A) and then consider the vector space Lm spanned by these vectors, i.e. the space of all linear combinations of $\{T(A)v|A \in G\}$.

Surely L_m is contained in L_n, but it might be the whole of L_n or it might be a proper subspace of dimension $m < n$.

A subspace of a vector space is a subset closed under addition and scalar multiplication.

A **proper subspace** is a subspace which is not the whole space and not just the point $\{0\}$.

Clearly L_m is closed under the action of the group, as for any $A \in G$, the action of A on an arbitrary vector $\sum a_B\ T(B)v$ in L_n is $A : \sum a_B\ T(B)v \longmapsto \sum a_B\ T(A)T(B)v = \sum a_B\ T(AB)v$, which is included in the definition of L_m.

If we choose a new basis for L_n such that the first m basis vectors span the subspace L_m, we see that

$$T(A)e_i = \sum_{j=1}^{m} e_j T_{ji}(A) \qquad \text{for } i \leq m$$

or $Tji(A) = 0$ for all $i \ni m, j > m, A \in G$.

So T has the form

$$T(A) = \begin{pmatrix} D^{(1)}(A) & X(A) \\ 0 & D^{(2)}(A) \end{pmatrix} \tag{2.1}$$

with $D^{(1)}$ an m x m matrix, $D^{(2)}$ an (n - m) x (n - m) matrix, and X an m x (n- m) matrix.

If we check how two such matrices multiply we can easily show that D(1)(AB) = D(1)(A)D(1)(B) and D(2)(AB) = D(2)(A)D(2)(B), so D(1) and D(2) are m and n→ m dimensional representations respectively, and our original representation T is said to be a **reducible representation**.

The subspace Lm is called an **invariant subspace**.

If there is no proper subspace of Ln which is closed under action by the group, we say that T is an irreducible representation.

We expect symmetry transformations on our hilbert space to be unitary transformations, preserving the norm of the states.

Then because we asked that e_i be an orthonormal basis, the matrices T(A) will be unitary matrices, and the representation is called a **unitary representation**.

We will now prove that any finite dimensional representation of a finite group is equivalent to a unitary representation.

This will also be true for infinite groups if we can define finite invariant integration measures, but this we will consider later.

Consider a representation T of a finite group G.

Define the matrix

$$H = \sum_{A \in G} T(A) T^{\dagger}(A).$$

H is clearly hermitian so it can be diagonalized by a unitary U , $H_d = U^{-1}HU$, $(H_d)_{ij} = d_i \delta_{ij}$.

So

$$\begin{aligned} d_i &= (H_d)_{ii} = \sum_A \sum_k \left(U^{-1} T(A)\right)_{ik} \left(T^{\dagger}(A) U\right)_{ki} \\ &= \sum_A \sum_k \left|\left(T^{\dagger}(A) U\right)_{ki}\right|^2 > 0, \end{aligned}$$

as $U^{-1} = U^{\dagger}$ because U is unitary, and both T and U are nonsingular.

Thus we can define $(H^{1/2}_d)_{ij} = d^{1/2}_i \delta_{ij}$ and $V = UH^{1/2}_d$ will provide the similarity transformation from T to an equivalent unitary representation Γ - let

$$\Gamma(A) = V^{-1} T(A) V = H_d^{-1/2} U^{-1} T(A) U H_d^{1/2}.$$

then

$$\begin{aligned} \Gamma(A)\Gamma^{\dagger}(A) &= H_d^{-1/2} U^{-1} T(A) \underbrace{U H_d^{1/2} H_d^{1/2} U^{-1}}_{H = \sum_B T(B) T^{\dagger}(B)} T^{\dagger}(A) U H_d^{-1/2} \\ &= H_d^{-1/2} U^{-1} \underbrace{\sum_B T(A) T(B) T^{\dagger}(B) T^{\dagger}(A)}_{H} U H_d^{-1/2} \\ &= H_d^{-1/2} H_d H_d^{-1/2} = \mathbb{1}, \end{aligned}$$

where in the first line we used that H_d is hermitian and U is unitary, and in the second that the underbraced sum is $\sum_B T(AB)T^\dagger(AB) = \sum_B T(B)T^\dagger(B) = H$, where the first equality is due to the left invariance of sums over all the elements in the group.

This is the first example we have seen, but far from the last, of the usefulness of summing over all the elements of a group with equal weight.

For a Lie group, one with elements indexed by a continuous parameter, we will need carefully to give meaning to a sum with equal weights over all elements.

We will need to define an invariant integration measure.

We will then find that for compact Lie groups, for which the total group volume in finite, we can again find a similarity transformation, so that we will have proven:

Theorem: *Any finite dimensional representation of a finite group or of a compact Lie group is equivalent to a unitary representation*.

If T was reducible, so is Γ, as there is still an invariant proper subspace.

The necessary change of basis to get Γ into the form of eq. 2.1 can be chosen unitary, so the new form is also unitary.

But then

$$\Gamma(A) = \left(\Gamma(A^{-1})\right)^{\dagger} = \begin{pmatrix} D^{(1)\dagger}(A^{-1}) & 0 \\ X^{\dagger}(A^{-1}) & D^{(2)\dagger}(A^{-1}) \end{pmatrix}.$$

But we know that the lower left block of Γ is zero, so X = 0, and the representation is block diagonal,

$$\Gamma(A) = \begin{pmatrix} D^{(1)} & 0 \\ 0 & D^{(2)} \end{pmatrix}.$$

A representation which can be successively reduced in this manner to a block diagonal form of two or more irreducible representations is said to by **fully reducible**.

We say that Γ is the direct sum of the irreducible representations.

Now to go ahead we must first prove

2.1 Schur's First Lemma

> *Any matrix which commutes with all the matrices of an irreducible representation of a finite or compact Lie group must be a multiple of the identity.*

We can assume we have made the similarity transformation so that our representatives Γ(A) are unitary.

Whether or not the commuting M is a multiple of the identity will not be affected by this similarity transformation.

Then if [M, Γ(A)] = 0 for all A ∈ G,

$$\Gamma(A)M = M\Gamma(A), \quad \text{and also } M^\dagger\Gamma^\dagger(A) = \Gamma^\dagger(A)M^\dagger.$$

We multiply the last equation by Γ(A) on the left and on the right, using the fact that Γ(A) is unitary, we get

$$\Gamma(A)M^\dagger = M^\dagger\Gamma(A),$$

so $M^\dagger$ also commutes with all Γ(A), and so do the hermitean matrices $H_+ = M + M^\dagger$ and $H_- = i(M^\dagger - M)$.

Now either of these hermitean matrices can be diagonalized by a unitary matrix U , $H = U D U^{-1}$, where $D_{ij} = d_i\delta_{ij}$.

But e can also transform our representation Γ by U, defining a new representation $\Gamma'(A) = U^{-1}\Gamma(A)U$, which is still irreducible and unitary, and $[\Gamma'(A), D] = U^{-1}[\Gamma(A), H]U = 0$.

Thus

$$(\Gamma'(A)D)_{ij} = \Gamma'(A)_{ij}d_j = (D\Gamma'(A))_{ij} = \Gamma'(A)_{ij}d_i \Rightarrow \Gamma'(A)_{ij}(d_j - d_i) = 0.$$

Now either all the d_i are equal, in which case D is a multiple of the identity, or $\Gamma'(A)_{ij} = 0$ for all A, except for i and j in the same block, that is for i and j having the same d.

But that would mean $\Gamma'(A)$ is reducible, and so is Γ(A), which contradicts the hypothesis.

This argument works for both H_+ and H_-, and tells us that both must be multiples of the identity, as must be $M = H_+ + iH_-$.

This lemma has a useful contrapositive: if a matrix M which is not a multiple of the identity matrix commutes with all the representatives of elements in the group, the representation must be reducible.

2.2 Schur's Second Lemma

If Γ^i and Γ^j are irreducible unitary representations of G, of dimensions ℓ_i and ℓ_j respectively, and if

$$\Gamma^i(A)M = M\Gamma^j(A) \quad for\ all \quad A \in G,$$

then either M = 0 or $\Gamma^i \cong \Gamma^j$ and M is nonsingular.

2.3 The Great Orthogonality Theorem

Let Γ^i and Γ^j be irreducible unitary representations of a group G, of finite dimensions ℓ_i and ℓ_j respectively.

Let X be a fixed ℓ_i x ℓ_j matrix, and

$$M = \sum_{A \in G} \Gamma^i(A) X \Gamma^j(A^{-1}).$$

Then for any B,

$$\begin{aligned} \Gamma^i(B)M &= \sum_A \Gamma^i(B)\Gamma^i(A)X\Gamma^j(A^{-1})\Gamma^j(B^{-1})\Gamma^j(B) \\ &= \sum_A \Gamma^i(BA)X\Gamma^j((BA)^{-1})\Gamma^j(B) \\ &= M\Gamma^j(B). \end{aligned}$$

Thus by Schur's second lemma, if Γ^{i} is not equivalent to Γ^{j}, M = 0 for any X.

In particular, take any $b \le \ell_i$ and $d \le \ell_j$, and let $X_{ef} := \delta_{eb}\delta_{fd}$.

Then

$$M_{ac} = \sum_{A\in G} \Gamma^{i}_{ab}(A)\Gamma^{j}_{dc}(A^{-1}) \qquad \text{must vanish}$$

$$\text{so} \quad 0 = \sum_{A\in G} \Gamma^{i}_{ab}(A)\Gamma^{j*}_{cd}(A) \qquad (\text{for } \Gamma^{i} \ncong \Gamma^{j}), \tag{2.2}$$

as Γ^{j} is a unitary representation.

On the other hand, if Γ^{i} and Γ^{j} are the same representation, then Schur's first lemma tells us $M_{ac} = c\delta_{ac}$.

Tracing,

$$c\,\ell_i = \sum_{A\in G}\sum_{a} \Gamma^{i}_{ab}(A)\Gamma^{i}_{da}(A^{-1}) = \sum_{A\in G} \Gamma^{i}_{db}(\mathbb{I}) = g\delta_{bd},$$

where g is the order of G, so $c = g\delta_{bd}/\ell_i$, and

$$\sum_{A\in G} \Gamma^{i}_{ab}(A)\Gamma^{i*}_{cd}(A) = \frac{g}{\ell_i}\delta_{ac}\delta_{bd}. \tag{2.3}$$

If we understand the index i to run over all inequivalent irreducible representations, we can write both results (2.2) and (2.3) as

$$\sum_{A\in G} \Gamma^i_{ab}(A)\Gamma^{j*}_{cd}(A) = \frac{g}{\ell_i}\delta_{ij}\delta_{ac}\delta_{bd}. \tag{2.4}$$

This is known as the **great orthogonality theorem**.

For the continuum version, if $\int \mathrm{d}\mu\ 1 = \mathrm{V}$,

$$\int d\mu \Gamma^i_{ab}(A)\Gamma^{j*}_{cd}(A) = \frac{V}{\ell_i}\delta_{ij}\delta_{ac}\delta_{bd}.$$

In what sense is this orthogonality?

Think of the space of all complex-valued functions of the group elements.

Define a norm on this space of functions by $(f, \mathrm{h}) = \sum_{\mathrm{A}\in\mathrm{G}} f^*(\mathrm{A})\ \mathrm{h}(\mathrm{A})$, or for a compact Lie group, $(f, \mathrm{h}) = \int \mathrm{d}\mu\ f^*(\mathrm{A}(\mu))\ \mathrm{h}(\mathrm{A}(\mu))$.

Then the great orthogonality theorem tells us that

$$\sqrt{\frac{\ell_i}{g}}\Gamma^i_{ab}(A)$$

form an orthonormal set of vectors in the vector space of all functions on the group.

If the group G= $\{A_1, ...A_g\}$ has a finite number of elements, one orthonormal basis for functions on G is the set e_α of functions $e_\alpha(A_\beta) = \delta_{\alpha\beta}$, for α and $\beta = 1, ..., g$.

This set clearly spans the space of all functions, as for any ψ,

$$\psi = \sum_\alpha \psi(A_\alpha) e_\alpha, \quad \text{that is} \quad \psi(B) = \sum_\alpha \psi(A_\alpha) e_\alpha(B)$$

and is linearly independent.

Thus the space of functions L_G on G is g-dimensional.

The orthonormal set we got from the great orthogonality theorem has dimension $\sum_i \ell^2_i$, so it must be true that $\sum_i \ell^2_i \leq g$.

We will soon show it is an equality.

In general, if we have a set of functions $\psi_j(A)$ on the group, we can define the operation of the group on these functions by

$$(B\psi_j)(A) := \psi_j(AB).$$

The order is important, in order to preserve the group multiplication law, (CB)ψ= C(Bψ).

To show this let both functions act on A.

((CB)ψ)(A) = ψ(ACB), (C(Bψ))(A) = (Bψ)(AC) = ψ(ACB) good!.

If you had tried to define $(B\psi)(A) := \psi(BA)$, (wrong!), you would not find $(CB)\psi = C(B\psi)$.

Now the function $B\psi$ is a function of the group elements.

Although it is not the function ψ, it is certainly in the space of all functions on the group.

Thus the space L_G of all functions on the group provides a representation of the group, g dimensional.

Let us use the basis e_α.

Then the function Be_α can be expanded with coefficients we call $\Gamma^{reg}{}_{\beta\alpha}(B)$ of the basis vectors e_β ,
$Be_\alpha = \sum_\beta \Gamma^{reg}{}_{\beta\alpha}(B)\, e\beta$.

Acting on A_γ, we have

$$Be_\alpha(A_\gamma) = \sum_\beta \Gamma^{\rm reg}_{\beta\alpha}(B) \underbrace{e_\beta(A_\gamma)}_{\delta_{\beta\gamma}} = e_\alpha(A_\gamma B) = \delta_{A_\gamma B,\, A_\alpha}\,.$$

So we see

$$\Gamma^{\rm reg}_{\gamma\alpha}(B) = \delta_{A_\gamma B,\, A_\alpha}\,.$$

This representation is called the **regular representation** of the group.

Included in this space of all functions on G is a set of functions

$$\psi_b(A) = \Gamma^i_{ab}(A),$$

where i is any of the irreducible representations of G and a is a fixed index a ∈ [1, ℓ_i].

This set of ℓ_i functions transform by

$$(B\psi_b)(A) = \psi_b(AB) = \Gamma^i_{ab}(AB) = \sum_c \Gamma^i_{ac}(A)\Gamma^i_{cb}(B) = \sum_c \psi_c(A)\Gamma^i_{cb}(B),$$

so $B\psi_b = \sum_c \psi_c \Gamma^i{}_{cb}(B)$, and this set transforms like the i^{th} irreducible representation.

Actually we see that we have ℓ_i such sets of functions, one set for each a, and they are guaranteed linearly independent by the Great Orthogonality Theorem.

Thus we see that each irreducible representation Γ^i appears at least ℓ_i times in the reduction of the regular representation.

This again tells us that $\sum_i \ell^2_i \leq g$.

Very soon, we shall see that it appears exactly ℓ_i times and $\sum_i \ell^2_i = g$.

Any function on the group is a linear combination of the e_α which provide the basis of the regular representation.

In particular, the function $f_B : G \to C$ given by $f_B(A) = \delta_{A,B}$ is just the $(\mathbb{1}, B)$ element in the regular representation Γ^{reg} as we have defined it.

But the regular representation is reducible, so it is equivalent to a direct sum of irreducible representations: $\Gamma^{\rm reg} = S^{-1}(\bigoplus_{i,n_i}\Gamma^{\rm i})\,S$, where $(\bigoplus_{i,n_i}\Gamma^{\rm i})$ is essentially a block diagonal matrix with n_i blocks for each irreducible representation $\Gamma^{\rm i}$, each block an ℓ_i x ℓ_i matrix function on the group.

This direct sum matrix thus has two compound indices, each one a combination of i giving the irreducible representation, n = 1, . . . , n_i telling which one of the n_i copies, and a or b giving which basis vector within that representation.

Thus $S = S_{(i,n,b),\beta}$ and

$$\Gamma^{\rm reg}_{\alpha\beta} = \sum_{i,n,a,b} \left(S^{-1}\right)_{\alpha,(i,n,a)} \Gamma^{i}_{ab} S_{(i,n,b),\beta},$$

and if we define $a^{\rm i}_{\rm ab}(B) = \sum_{\rm n} (S^{-1})_{\mathbf{1},({\rm i,n,a})}\, S_{({\rm i,n,b}),B}$ we have

$$\sum_{iab} a^{(i)}_{ab}(B)\Gamma^{i}_{ab}(A) = \Gamma^{\rm reg}_{\mathbf{1},B}(A) = \delta_{A,B}. \tag{2.5}$$

Multiply this by $\Gamma^{\rm j*}_{\rm cd}(A)$ and sum over $A \in G$,

$$\Gamma^{j*}_{cd}(B) = \sum_{iab} a^{(i)}_{ab}(B) \underbrace{\sum_{A} \Gamma^{i}_{ab}(A)\Gamma^{j*}_{cd}(A)}_{\frac{g}{\ell_i}\delta_{ij}\delta_{ac}\delta_{bd}} = \frac{g}{\ell_j} a^{(j)}_{cd}(B),$$

using the Great Orthogonality Theorem.

When we insert this back into (2.5), we have

$$\delta_{AB} = \sum_{iab} \frac{\ell_i}{g} \Gamma^i_{ab}(A) \Gamma^{i*}_{ab}(B), \tag{2.6}$$

which is sort of orthonormality of the g vectors $\Gamma(A)$ (A is an index here) in the $\sum \ell^2_i$ dimensional space indexed by (i, a, b).

In particular, g vectors in a $\sum_i \ell^2_i$ dimensional space can be independent, which they are, only if $g \leq \sum_i \ell^2_i$.

Perhaps it would be clearer to say this: Using (2.6), any function on G

$$\begin{aligned} f(A) &= \sum_B \delta_{AB} f(B) = \sum_{iab} \left(\frac{\ell_i}{g} \Gamma^{i*}_{ab}(B) \Gamma^i_{ab}(A) \right) f(B) \\ &= \sum_{iab} \left(\frac{\ell_i}{g} \Gamma^{i*}_{ab}(B) f(B) \right) \Gamma^i_{ab}(A), \end{aligned}$$

is a linear combination of the $\sum \ell^2_i$ functions $\Gamma^i_{ab}(A)$, so these are a complete set of functions on G, and that space is g-dimensional.

So again $g \leq \sum_i \ell^2_i$.

But we have already shown $g \geq \sum_i \ell^2_i$, so we have an important statement about the full set of irreducible representations and their dimensions:

$$g = \sum_i \ell_i^2. \tag{2.7}$$

2.4 Characters

We have seen that two representations are considered equivalent if they are related by a similarity transformation $\Gamma(1)(A) = U\Gamma^{(2)}(A)U^{-1}$ for all $A \in G$.

To characterize a representation we are not interested in all the equivalent forms, so we would like something invariant under similarity.

This is done by defining the character of a representation Γ to be a function $\chi : G \to C$ on the group given by

$$\chi(A) = \mathrm{Tr}\,\Gamma(A).$$

Note χ is a complex-valued function on the group, not a matrix-valued function as Γ is.

Also note it is obviously unaffected by similarity transformations.

But this has an important consequence on the function χ — namely, if A and B are conjugate group elements, i.e., $A = C^{-1}BC$ for some $C \in G$, then

$$\begin{aligned}\chi(A) &= \mathrm{Tr}\,\Gamma(A) = \mathrm{Tr}\,\Gamma(C^{-1}BC) = \mathrm{Tr}\left(\Gamma^{-1}(C)\Gamma(B)\Gamma(C)\right) = \mathrm{Tr}\,\Gamma(B) \\ &= \chi(B).\end{aligned}$$

Thus the character is actually only a function on the conjugacy classes of the group.

Now the great orthogonality theorem (2.4)

$$\sum_{A\in G}\Gamma^{i}_{ab}(A)\Gamma^{j*}_{cd}(A)=\frac{g}{\ell_i}\delta_{ij}\delta_{ac}\delta_{bd}$$

can be traced in ab and in cd, to give

$$\sum_{A\in G}\chi^{i}(A)\chi^{j*}(A)=\frac{g}{\ell_i}\delta_{ij}\underbrace{\sum_{ac}\delta_{ac}\delta_{ac}}_{\ell_i}=g\delta_{ij}.$$

Thus two different representations have different characters—in fact they form a set of linearly independent functions on the conjugacy classes of the group.

Thus the number of inequivalent irreducible representations of G must be ≤ the dimension of the space of complex functions on the classes of G, which is just the number of classes of G.

Suppose we have a representation which might be reducible.

Then $\Gamma = \bigoplus_i a_i\Gamma^i$, which means

$$\underbrace{\Gamma^{i_1}\bigoplus\Gamma^{i_1}\bigoplus\ldots\bigoplus\Gamma^{i_1}}_{a_{i_1}\ \text{times}}\bigoplus\Gamma^{i_2}\ldots$$

That is, we know in principle that Γ is reducible to a direct sum of irreducible representations, but we haven't done it.

Now we will see how to extract the a_i's.

The representation Γ has a character χ, so for each possible irreducible representation Γ^j which might be included in the direct sum, we can form

$$\sum_{A\in G}\chi^{j*}(A)\chi(A)=\sum_i a_i\sum_{A\in G}\chi^{j*}(A)\chi^i(A)=\sum_i a_i g\delta_{ij}=ga_j, \tag{2.8}$$

so

$$a_j=\frac{1}{g}\sum_{A\in G}\chi^{j*}(A)\chi(A). \tag{2.9}$$

If Γ is in fact irreducible, one a will come out one and the others zero.

For the one j which gives $a_j = 1$, $\chi_j = \chi$, so

$$\sum_{A\in G}\chi^*(A)\chi(A)=g \qquad \text{if } \chi \text{ is irreducible.}$$

In fact, this is an "if and only if" statement, for

$$\sum_{A\in G}\chi^*(A)\chi(A)=\sum_{ij}a_ia_j\sum_{A\in G}\chi^{i*}(A)\chi^j(A)=g\sum_{ij}a_ia_j\delta_{ij}=g\sum a_i^2.$$

If this is g, exactly one a_i is one and the others must be zero.

Let us return to the regular representation, which is a g x g matrix representation

$$\Gamma^{\text{reg}}_{\alpha\beta}(B) = \delta_{A_\alpha B,\, A_\beta}.$$

The diagonal elements are $\Gamma^{\text{reg}}{}_{\alpha\alpha}(\mathrm{B}) = \delta_{A_\alpha B,\, A_\alpha} = \delta_{B,\mathbb{1}}$, so

$$\chi^{\text{reg}}(\mathbb{1}) = \sum_{\alpha=1}^{g} (\mathbb{1})_{\alpha\alpha} = g, \qquad \chi^{\text{reg}}(B) = 0 \ \text{ for } B \neq \mathbb{1}.$$

For each irreducible representation Γ^{i} of G, $\chi^{\mathrm{i}}(\mathbb{1}) = \ell_{\mathrm{i}}$, so the number of times that representation occurs in the regular representation is

$$a_i = \frac{1}{g}\sum_A \chi^i(A)\chi^{\text{reg}}(A) = \frac{1}{g}\chi^i(\mathbb{1})\chi^{\text{reg}}(\mathbb{1}) = \ell_i,$$

as we already realized from establishing that our original inequality $g \geq \sum \ell^2{}_{\mathrm{i}}$ was in fact saturated, $g = \sum \ell^2{}_{\mathrm{i}}$.

This equality now reflects just that the dimension of the regular representation is the sum of the dimensions of its irreducible components:

$$\dim \Gamma^{\text{reg}} = g = \sum_i a_i \dim \Gamma^i = \sum_i \ell_i^2.$$

The characters also form a basis for functions on the conjugacy classes.

Let $\{\mathrm{C_k}\}$ be the conjugacy classes of G, and

$$f : \{\mathcal{C}_k\} \to \mathbb{C}$$

be an arbitrary function on the classes. Then $\tilde{f} : G \to \mathbb{C}$ defined by $\tilde{f}(A) = f(\mathcal{C}(A))$, where $\mathcal{C}(A)$ is the conjugacy class containing A, is a map from the group into the complex numbers. As any such function is a linear combination of the Γ^i_{ab}'s, $\tilde{f}(A) = \sum_{iab} c^i_{ba}\Gamma^i_{ab}(A)$ for some set of coefficients c^i_{ba}. Becaus $\tilde{f}(B^{-1}AB) = \tilde{f}(A)$, we have

$$\sum_{iab} c^i_{ba}\Gamma^i_{ab}(A) = \sum_{iab} c^i_{ba}\Gamma^i_{ab}(B^{-1}AB) = \sum_{iabrs} c^i_{ba}\Gamma^i_{ar}(B^{-1})\Gamma^i_{rs}(A)\Gamma^i_{sb}(B).$$

But the Γ^{i}_{ab} are a set of linearly independent functions, so

$$c^i_{sr} = \sum_{ab} c^i_{ba}\Gamma^i_{ar}(B^{-1})\Gamma^i_{sb}(B),$$

for any B, or in matrix form

$$c^i = \Gamma^i(B)c^i\Gamma^i(B^{-1}), \qquad \text{or } c^i\Gamma^i(B) = \Gamma^i(B)c^i.$$

But then Schur's first lemma tells us $c^i \propto \mathbb{I}$, or $c^i_{ba} = k^i\delta_{ab}$, and

$$\tilde{f}(A) = \sum_i k^i \sum_a \Gamma^i_{aa}(A) = \sum_i k^i\chi^i(A),$$

or $f = \sum_i k^i \chi^i$ is a linear combination of the characters, and consequently the number of irreducible representations is the same as the number of conjugacy classes.

The equation (2.8) can be thought of as an orthogonality statement for the characters.

As the character of each element in a conjugacy class is the same, if we define η_C to be the number of elements in the class C, we have

$$\sum_k \eta_{C_k} \chi^i(C_k)\chi^{j*}(C_k) = g\delta_{ij}.$$

The fact that any function on classes can be expressed as a linear combination of characters enables us to write

$$\delta_{kk'} = \sum_i a_i(C_{k'})\chi^i(C_k)$$

Multiply by $\eta_{C_k} \chi^{j*}(C_k)$ and sum over k,

$$\eta_{C_{k'}}\chi^{j*}(C_{k'}) = \sum_i a_i(C_{k'}) \sum_k \eta_{C_k}\chi^i(C_k)\chi^{j*}(C_k) = g a_j(C_{k'}),$$

so $a_j(C_{k'}) = \frac{\eta_{C_{k'}}}{g}\chi^{j*}(C_{k'})$, and

$$\delta_{kk'} = \sum_i \frac{\eta_{C_{k'}}}{g}\chi^{i*}(C_{k'})\chi^i(C_k).$$

This tells us

$$\sqrt{\frac{\eta_C}{g}}\,\chi^i(C)$$

is a complete orthonormal basis on the set of conjugacy classes.

2.7 Direct Products of Representations

Often a group acts separately on factorizable pieces of the wave function.

For example, if we wish to consider the states of two electrons in our lattice, the total wave function $\psi(x_1, y_1, x_2, y_2)$ could be written as a sum of products $\sum^i a_i\phi_i(x_1, y_1)\rho_i(x_2, y_2)$. ϕ_i and ρ_i can be separately decomposed into irreducible representations of the group, and the group acts on ψ by the direct product, because it acts together on both $\vec{r}_1$ and $\vec{r}_2$.

The vector space upon which the group is acting here is the tensor product of the space of all functions of $\vec{r}1 = (x1, y1)$ with the space of all functions of $\vec{r}2$.

Perhaps it would be better to consider first a finite-dimensional example.

The electron of a hydrogen atom in an $\ell = 1$ state is in a three-dimensional space with basis vectors $|m\propto$, $m=\rightarrow 1, 0, 1$.

The rotation group acts on this space in a three-dimensional representation to be further discussed later.

At the same time, the electron has a spin degree of freedom in a two-dimensional space with basis $\{|\uparrow\propto, |\downarrow\propto\}$, or $|s\propto$, $s= \pm 12$.

A full description of the state of the electron lies in a six-dimensional space whose basis is most naturally written as $\{|m, s\propto\}$ or

$$|-1,\uparrow\rangle\,, \quad |-1,\downarrow\rangle\,, \quad |0,\uparrow\rangle\,, \quad |0,\downarrow\rangle\,, \quad |1,\uparrow\rangle\,, \quad |1,\downarrow\rangle\,,$$

though of course we could still index these basis vectors as e1, ..., e6.

Now some operators, such as the orbital angular momentum $\vec{m}L$, may act only on part of the composite index, $\langle m\rightarrow, s\rightarrow| L |m, s\propto= Lm' m\delta s's$, while others might only act on the spin, $\langle m\rightarrow, s\rightarrow| S |m, s\propto= Ss's\delta m' m$, but in general an operator $\vec{O}$,e.g.$L\cdot S$, is a $6\leftrightarrow6$ matrix $O(m' s'),(ms)= \langle m\rightarrow, s\rightarrow| O |m, s\propto$.

The six dimensional space on which it acts is the tensor product of the three dimensional space of orbital angular momentum states and the two dimensional space of spin.

Under a rotation8 A, the state of the electron will be acted on by a six dimensional representation

$$\Gamma_{m's',ms}(A) = \Gamma^{\ell=1}_{m'm}(A)\Gamma^{s=\frac{1}{2}}_{s's}(A),$$

which is the **direct product**,

$$\Gamma(A) = \Gamma^{\ell=1}(A) \bigotimes \Gamma^{s=\frac{1}{2}}(A).$$

More generally we will consider the product of two irreducible representations of any group,

$$\Gamma(A) = \Gamma^i(A) \bigotimes \Gamma^j(A).$$

This is clearly a representation, because the matrix sum over the tensor product index is just an independent sum over the indices of each subspace.

If r, s, t are indices in the ℓi dimensional space on which "i acts and a, b, c indices for the ℓj dimensional space on which "j acts,

$$\begin{aligned}
\Gamma_{ra,sb}(AB) &= \Gamma^i_{rs}(AB)\Gamma^j_{ab}(AB) = \sum_t \Gamma^i_{rt}(A)\Gamma^i_{ts}(B) \sum_c \Gamma^j_{ac}(A)\Gamma^j_{cb}(B) \\
&= \sum \Gamma^i_{rt}(A)\Gamma^j_{ac}(A)\Gamma^i_{ts}(B)\Gamma^j_{cb}(B) = \sum \Gamma_{ra,tc}(A)\Gamma_{tc,sb}(B) \\
&= (\Gamma(A)\Gamma(B))_{ra,sb}\,.
\end{aligned}$$

The character of the product representation is simple:

$$\chi(A) = \mathrm{Tr}\,\Gamma(A) = \sum_{ra} \Gamma(A)_{ra,ra} = \left(\sum_r \Gamma^i_{rr}(A)\right)\left(\sum_a \Gamma^j_{aa}(A)\right) = \chi^i(A)\chi^j(A).$$

Now the product of two irreducible representations need not be irreducible, but it must be equivalent to a direct sum of irreducible representations.

For example, for the case of the electron with orbital angular momentum 1 and also spin, the full rotation group has representations labelled by j, and our six dimensional space is a sum of a j=12 and a j=32 representation,

$$\Gamma^{\ell=1} \bigotimes \Gamma^{s=\frac{1}{2}} \cong \Gamma^{j=\frac{1}{2}} \bigoplus \Gamma^{j=\frac{3}{2}}.$$

More generally,

$$\Gamma^i(A) \bigotimes \Gamma^j(A) \cong \bigoplus_k a_k^{ij} \Gamma^k(A),$$

where each representation may appear aijk times in this sum, called the Clebsch-Gordon series.

The non-negative integers aijk are called Clebsch-Gordon coe#cients by mathematicians but not by physicists, who mean something else.

From the formula (2.9) for the number of times a given irreducible representation occurs in an arbitrary one, the coeffients are easily calculated:

$$a_k^{ij} = \frac{1}{g} \sum_{A \in G} \chi^i(A) \chi^j(A) \chi^{k*}(A).$$

We reserve our discussion of the physics of this decomposition until after we have learned to construct irreducible representations of infinite groups.

Part 3 Infinite Groups

Most physical situations that have a symmetry group have an infinite group.

Some examples:

- Rotational invariance, $SO(3)$. Here we can rotate through an arbitrary angle specified by a continuous parameter θ, restricted to some finite range, say $[0, 2\pi)$. There are a infinite continuum of possible values of θ, even though its range is limited. There are also two continuous parameters necessary to specify the direction about which this rotation is to take place. This is a three-parameter group, and the space of these transformations is a three dimensional manifold[1].

- Translational invariance of the vacuum, $\vec{x} \to \vec{x} + \vec{a}$, for $\vec{a}$ an arbitrary three dimensional vector with a continuum of possibilities for each coefficient.

- A combination of the above, $\vec{x} \to \vec{x}\,'$ with $x_i' = \sum_j R_{ij} x_j + a_i$, with R_{ij} a rotation matrix.

- Translations on a lattice which leave the lattice unchanged. A perfect lattice in D dimensions has D linearly independent **lattice vectors** $\vec{a}_i, i = 1, ..., D$, such that the lattice is unchanged if the whole thing is translated by a vector $\sum_{i=1}^{D} n_i \vec{a}_i$, where the coefficients n_i are all arbitarary *integers*.

The last example differs from the others in an essential way — there are no group elements which do arbitrarily little, although of course there is one, the identity, which does nothing.

For the SO(3) rotations we can rotate through an arbitrarily small angle, for the translations of the vacuum we can translate by a femtometer (we theorists can — experimentalists might have a hard time).

But the symmetries on the lattice have a minimum nonzero distance for which a translation can be a symmetry.

A group that has elements which are infinitesimally different from is called a **continuous** group.

The others are called **discrete**.

A continuous group requires one or more continuous parameters to specify which element is being discussed.

For example, for the translation group of the vacuum, $\vec{a}$ = (ax, ay , az) is a set of three continuous parameters needed to specify the translation.

For the rotation group we can specify three Euler angles.

Later we will make a better choice, but it will still require three real parameters.

Notice that we have implicitly assumed some kind of topology on the group, for we have talked of elements arbitrarily close to the identity, which implies a sequence of elements converging to the identity.

For this reason these groups are also called **topological** groups

3.1 Connectedness

With topology comes the concept of connectedness — can any two elements of the group be connected by a continuous path of elements in the group.

The part of the group connected to the identity is called the connected component

Clearly the translations of the vacuum form a connected group, because, for any translation by a, the set of translations $\{ T_\lambda : \vec{x} \mapsto \vec{x} + \lambda\vec{a}, \text{ for } \lambda \in [0, 1]\}$, is a continuous path of translations starting from the identity at $\lambda = 0$ and ending at the translation by $\vec{a}$ at $\lambda = 1$.

The proper rotations are also connected by the same approach.

But if we consider the set of all transformations that preserve lengths, which is to sav the set of all orthogonal transformations, O(3), this includes the parity transformation P : $\vec{x} \rightarrow -\vec{x}$.

It is clear, however, that there is no path of orthogonal transformations which connects this parity transformation to the identity.

Parity in 3-D converts a left hand to a right hand, which can't be done continuously by orthogonal transformations.

More abstractly, and in arbitrary dimension, if A is an orthogonal matrix, $A^{-1} = A^T$.

Then $(\det A)^{-1} = \det(A^{-1}) = \det(A^T) = \det A$, so $\det A = \pm 1$.

The identity has determinant +1 while parity (in 3-D) has determinant -1.

But as nointermediate values are allowed for an orthogonal transformation, there is no path between them.

Thus this group, O(3), consists of two pieces, the connected component, called SO(3), which is the subgroup of orthogonal matrices with determinant +1, and the piece connected to P .

This second component is in fact the left coset of SO(3) in O(3) with respect to P , and SO(3) is a normal subgroup.

The connected component will always form a normal subgroup, and the factor group will always be discrete.

For the most part we will treat only connected groups.

The space of parameters describing the connected component of the group will form a manifold, that is, the neighborhood of each point can be described by Euclidean coordinates in n dimensions, though the metric may be only Euclidean in an infinitesimal neighborhood.

We will define a metric (or measure) on the parameter space of the group later, but for now I only comment on topological issues.

Besides connectedness, some other aspects of the topology of the group manifold which will come into play are whether or not it is **simply connected** and whether it is **compact**.

A manifold is simply connected if every closed path can be continuously shrunk to a point.

The surface of a sphere is simply connected, the surface of a donut, or a torus, is not.

The manifold is compact if it forms a closed and bounded set in the topology we are considering.

Only after defining a metric on the group manifold can we really answer the question of whether or not a sequence is a Cauchy sequence, which is necessary to define compactness.

Usually the metric will be uniformly continuous in the parameters, so a sequence of elements whose parameters approach a limit themselves approach a limit.

When this is true, compactness reduces to having a compact set in parameter space.

But it is not always true.

Examples:

- $SO(2)$, the set of rotations in two dimensions[4], where $g(\theta)$ is a counter-clockwise rotation through an angle $\theta \in [0, 2\pi)$. This is compact, but not simply connected. The group manifold is simply a circle.

- The translations in one dimension by an arbitrary amount, $T(a) : x \to x + a$. Here $a \in (-\infty, \infty)$ and the group is not bounded and not compact. However, we can't really tell just by the range of the parameter; we might have parameterized the group differently, $T'(\theta) : x \to x + \tan(\theta/2)$, for $\theta \in (-\pi, \pi)$. This is exactly the same group of transformations as the $\{T(a)\}$, so it is still noncompact, even though the parameters are on a bounded set. We will discuss very soon why a is a more valid parameterization than θ.

- $SO(3)$, rotations in three dimensions. Recall that the group elements can be described by a 3-D real vector $\vec{\omega}$ as a rotation through an angle $|\vec{\omega}|$ about the axis in the direction of $\vec{\omega}$. But only $\vec{\omega}$'s with $|\vec{\omega}| \leq \pi$ are needed, so the parameter space is a ball of radius π, and indeed the topology is strange, because for the points with $|\vec{\omega}| = \pi$, the group elements $g(\vec{\omega}) = g(-\vec{\omega})$, so opposite points of each diameter are identified, and points near the two ends of the diameter are close to each other.

3.2 Infinitesimal Generators

Let us concentrate on the connected component of the group.

Suppose that the group elements, at least those sufficiently near the identity, are parameterized by a D dimensional parameter ν_i, with $g(\nu_i = 0) = \mathbb{I}$.

Any representation Γ which respects the topology of the group will then have a power series expansion

$$\Gamma(g(\nu)) = \mathbb{1} + \sum_i \nu_i \Gamma_i + \mathcal{O}(\nu_i \nu_j).$$

The D matrices Γ_i are just

$$\Gamma_i = \frac{\partial}{\partial \nu_i} \Gamma(g(\nu)) \Big|_{\nu_j = 0}.$$

Of course the Γ_i depend on the representation and each is a matrix.

More abstractly, we can consider a function *f* in the space of all (sufficiently differentiable) functions defined on the group.

Let us define a set of differential operators L_i on this space by

$$[L_i f](g) = \frac{\partial}{\partial \nu_i} f(A(\nu)g)\Big|_{\nu_j=0} \quad \text{for } g \in G.$$

But representations are functions on the group, and if we consider an irreducible subspace $\Gamma^k{}_{ab}(A)$,

$$\begin{aligned} [L_i \Gamma^k_{ab}](g) &= \frac{\partial}{\partial \nu_i} \Gamma^k_{ac}(A(\nu))\Big|_{\nu_j=0} \Gamma^k_{cb}(g) \\ \text{so} \quad L_i \Gamma^k_{ab} &= \Gamma^k_{ac}(L_i)\Gamma^k_{cb} \\ \text{where} \quad \Gamma^k_{ac}(L_i) &:= \frac{\partial}{\partial \nu_i} \Gamma^k_{ac}(A(\nu))\Big|_{\nu_j=0} \end{aligned}$$

defines a representation not of the group but of the operators L_i.

Thus far we have considered only group elements in an infinitesimal neighborhood of the identity, i.e. $A(\nu)$ for infinitesimal ν.

Let us extend this parameterization by writing, for finite ν,

$$A(\nu) = \lim_{N\to\infty} \left[A\left(\frac{\nu}{N}\right)\right]^N.$$

For any representation,

$$\Gamma(A(\nu)) = \lim_{N\to\infty}\left[\Gamma\left(A\left(\frac{\nu}{N}\right)\right)\right]^N = \lim_{N\to\infty}\left[1+\frac{\nu}{N}\Gamma_i\right]^N = \exp\sum_i \nu_i\Gamma(L_i),$$

so we can write at least formally,

$$A(\nu) = e^{\Sigma_i \nu_i L_i}.$$

It can be shown that any element in the connected component of a compact Lie group is of this form, so we now have a good parameterization of the whole thing.

Example: SO(2) are the rotations in two dimensions.

Let us start with a very bad parameterization of these 2 x 2 matrices,

$$A(x) = \begin{pmatrix} \sqrt{1-x^2} & -x \\ x & \sqrt{1-x^2} \end{pmatrix} \qquad \text{(injudicious way of expressing } A\text{)}$$

Then the one L_i is

$$L = \frac{d}{dx}A(x)\Big|_{x=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Note $L^2 = -\mathbb{I}$, so

$$
\begin{aligned}
e^{\theta L} &= \sum_n \frac{1}{n!}\theta^n L^n = \mathbb{I}\sum_{\text{even } n}\left(\frac{1}{n!}(i\theta)^n\right) - iL\sum_{\text{odd } n}\frac{1}{n!}(i\theta)^n \\
&= \mathbb{I}\cos\theta + L\sin\theta = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}.
\end{aligned}
$$

So although we started with a deliberately poor parameterization in terms of x, we find a natural parameterization in terms of θ.

And any rotation in SO(2) can be written as $A(\theta) = e^{\theta L}$.

L is the only generator of SO(2).

Clearly $A(\theta_1)A(\theta_2) = A(\theta_1 + \theta_2)$ so the group is Abelian.

All of the irreducible representations of any Abelian group are one dimensional, because all the representatives commute with each other and therefore can be simultaneously diagonalized.

Thus for our SO(2),

$$\Gamma^{(m)}(A(\theta)) = e^{im\theta} = \chi^{(m)}(A(\theta)).$$

As $\theta + 2\pi$ describes the same group element as θ, $A(2\pi) = \mathbb{I}$ and we must have $\Gamma(A(2\pi)) = e^{2\pi i m} = 1$, so m must be an integer.

We have a countable infinity of representations.

The group-invariant volume for this group is just dθ, which is left invariant under left multiplication by A(ϕ) because

$$\int_0^{2\pi} d\theta\, f(A(\phi)A(\theta)) = \int_0^{2\pi} d\theta\, f(A(\theta+\phi)) = \int_\phi^{\phi+2\pi} d\theta'\, f(A(\theta'))$$
$$= \int_0^{2\pi} d\theta\, f(A(\theta)).$$

So we expect orthogonality of the characters in the continuous version:

$$\int_0^{2\pi} \chi^{m*}(\theta)\chi^n(\theta)\, d\theta = 2\pi\delta_{mn}.$$

The functions $e_m(\theta) = e^{im\theta}$, for m = -∞…∞ form a complete set of functions on the group, which as a set is just a circle.

SO(2) is clearly connected.

It is clearly multiply-, not simply-, connected, because the path $\lambda \rightarrow A(2\pi n\lambda)$ for $\lambda \in [0, 1]$ is a closed path which cannot be continuously deformed to a point because it wraps around the circle n times.

We have claimed that the group elements for the connected component of any compact Lie group can all be written as

$$g(\nu) = e^{\Sigma_i \nu_i L_i}$$

where the L_i are the generators of the group.

How does the multiplication law of the group manifest itself in properties of the generators L_i?

For small ν_1, ν_2,

$$g(\nu_1)g(\nu_2) \sim (1+\sum_i \nu_{1i}L_i)(1+\sum_j \nu_{2j}L_j) = 1+\sum_i(\nu_{1i}+\nu_{2i})L_i + \mathcal{O}(\nu_1\nu_2),$$

so a great deal of the group multiplication is built into our choice of parameters, where it is reflected additively.

To see more we need to go to higher order:

$$\begin{aligned} g(\nu_1)g(\nu_2) &= 1+\sum_i(\nu_{1i}+\nu_{2i})L_i + \sum_{ij}\nu_{1i}\nu_{2j}L_iL_j + \mathcal{O}(\nu_1^2,\nu_2^2), \\ g(\nu_2)g(\nu_1) &= 1+\sum_i(\nu_{1i}+\nu_{2i})L_i + \sum_{ij}\nu_{2j}\nu_{1i}L_jL_i + \mathcal{O}(\nu_1^2,\nu_2^2). \end{aligned}$$

We see that if the group is Abelian, $g(\nu_1)g(\nu_2) = g(\nu 2)g(\nu_2)$, then $[L_i, L_j] = 0$, and the generators commute.

If the generators do all commute, then

$$e^{\Sigma_i\nu_{1i}L_i}e^{\Sigma_i\nu_{2i}L_i} = e^{\Sigma_i(\nu_{1i}+\nu_{2i})L_i} = e^{\Sigma_i\nu_{2i}L_i}e^{\Sigma_i\nu_{1i}L_i}, \qquad \text{(Abelian group)}$$

and the group multiplication simply corresponds to addition in the vector space spanned by the L_i's, and the statement that the group elements commute is true, not just perturbatively.

If the generators do not commute, then the multiplication

$$e^{\Sigma_i \nu_{1i} L_i} e^{\Sigma_i \nu_{2i} L_i} = e^{\Sigma_i \nu_{3i} L_i}$$

will have an expression for $\nu_3(\nu_1, \nu_2)$ which is a more complicated function of its arguments.

We may, however, expand ν_3 in a power series in ν_1 and ν_2, and as we saw above it begins with $\nu_3 = \nu_1 + \nu_2 + \ldots$

Expanding to second order, (summations understood)

$$\begin{aligned} g(\nu_1)g(\nu_2) &= \left(1 + \nu_{1i}L_i + \frac{1}{2}\nu_{1i}\nu_{1j}L_iL_j\right)\left(1 + \nu_{2i}L_i + \frac{1}{2}\nu_{2i}\nu_{2j}L_iL_j\right) \\ &= \left(1 + \nu_{3i}L_i + \frac{1}{2}(\nu_{1i} + \nu_{2i})(\nu_{1j} + \nu_{2j})L_iL_j\right), \end{aligned}$$

where in the last term, which is quadratic in ν_3, the first order expression for $\nu_3(\nu_1, \nu_2)$ is sufficient.

Expanding the two sides we have

$$\begin{aligned} & 1 + (\nu_{1i} + \nu_{2i})L_i + \left(\frac{1}{2}\nu_{1i}\nu_{1j} + \frac{1}{2}\nu_{2i}\nu_{2j} + \nu_{1i}\nu_{2j}\right) L_iL_j \\ = \; & 1 + \nu_{3i}L_i + \left(\frac{1}{2}\nu_{1i}\nu_{1j} + \frac{1}{2}\nu_{2i}\nu_{2j} + \frac{1}{2}\nu_{1i}\nu_{2j} + \frac{1}{2}\nu_{2i}\nu_{1j}\right) L_iL_j \end{aligned}$$

Subtracting gives $0 = (\nu_{3i} - \nu_{1i} - \nu_{2i})L_i - \frac{1}{2}\nu_{1i}\nu_{2j}[L_i, L_j]$.

As the L_i are a complete set, this can only have a solution for ν_3, as it must, if

$$[L_i, L_j] = c_{ij}{}^k L_k$$

for some set of coefficients c^k_{ij} , which are called the s**tructure constants** of the group.

Clearly a group is Abelian if and only if all the structure constants are zero.

We see that the generators of the group form a Lie Algebra.

Definition: An r dimensional **Lie Algebra $\mathcal{L}$** is an r dimensional vector space together with a bilinear composition $[\cdot\ ,\cdot] : \mathcal{L}$ x $\mathcal{L} \rightarrow \mathcal{L}$ with the properties

$$[x, y] = -[y, x]$$

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

The second equation is called the **Jacobi identity**.

These requirements are automatically satisfied if the $[\cdot\ ,\cdot]$ law is defined as the commutator of an associative multiplication law, [x, y] = xy - yx.

But the commutator itself is not associative,

$$[x, [y, z]] - [[x, y], z] = [x, [y, z]] + [z, [x, y]] = -[y, [z, x]] \neq 0.$$

Note that the two laws of [· , ·] imply

$$c_{ij}{}^{k} = -c_{ji}{}^{k},$$

$$c_{i\ell}{}^{m} c_{jk}{}^{\ell} + c_{j\ell}{}^{m} c_{ki}{}^{\ell} + c_{k\ell}{}^{m} c_{ij}{}^{\ell} = 0.$$

Example 3: SU (2)

The group U(N) is the set of unitary N x N matrices under ordinary matrix multiplication.

As for O(N), the group of N x N real orthogonal matrices, it is useful to limit ourselves to those with determinant equal to 1, called SU(N) and SO(N) respectively.

So SU(2) is the group of 2 x 2 complex unitary matrices with determinant 1.

A unitary matrix U can always be written U= e^{iH} with H a hermitian matrix (proof: diagonalize first, prove it for the diagonalized version, then observe that the similarity transformation factors out).

We can also use the useful formula

$$\det U = e^{\mathrm{Tr}\ln U}$$

so det U = $e^{i\mathrm{Tr}\,H}$ = 1 implies Tr H = 0, so H is hermitian and traceless and is therefore a real linear combination of the Pauli matrices σ_j , H = 1/2 $\sum_j \omega_j \sigma_j$.

We have already parameterized the group in terms of its generators

$$L_j^P = \frac{1}{2}\sigma_j.$$

The factor of 1/2 is conventional, so that the L_i's are normalized so as to have the same structure constants as for SO(3)(I leave it to you to work out SO(3) on your own - see end of section).

For

$$[L_j^P, L_\ell^P] = \frac{1}{4}[\sigma_j, \sigma_\ell] = \frac{i}{2}\epsilon_{j\ell k}\sigma_k = i\epsilon_{j\ell k}L_k,$$

so the structure constants are

$$c_{j\ell}{}^k = \epsilon_{j\ell k} \quad \text{just as for } SO(3).$$

Thus the Lie algebra of SU(2) and the Lie algebra of SO(3) are the same.

The groups are locally isomorphic.

But a "rotation" through θ about the j axis gives

$$\begin{aligned} e^{i\theta\sigma_j/2} &= \sum_n \frac{(-1)^n}{2n!}\left(\frac{\theta}{2}\right)^{2n} + i\sigma_j \sum_n \frac{1}{(2n+1)!}(-1)^n \left(\frac{\theta}{2}\right)^{2n+1} \\ &= \cos(\theta/2) + i\sigma_j \sin(\theta/2), \end{aligned}$$

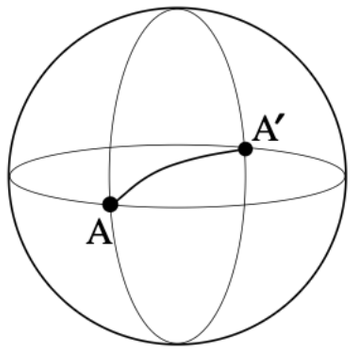
which are not the same.

In fact, the group space now consists of all $|\vec{\omega}| \leq 2\pi$ rather thar $|\vec{\omega}| \leq \pi$, but on the boundary $|\vec{\omega}| = 2\pi$

$$e^{-i\vec{\omega}\cdot\vec{\sigma}/2} = \cos(\omega/2) - i\vec{\omega}\cdot\vec{\sigma}\sin(\omega/2) \underset{|\omega|\to 2\pi}{\longrightarrow} -1,$$

so all the points on the surface of the ball, $|\vec{\omega}| = 2\pi$ are identified.

The group manifold is the closed ball $|\vec{\omega}| \leq \pi$, but with the opposite ends of each diameter of the ball identified with each other.



And we now have a simply connected manifold, because if we consider the path A A′ now, we can deform it by moving A′ without moving A, because all the points on the surface A are the same, and so we can bring A′ back to A and then shrink the rest of the path to a point.

Thus SU (2) is simply connected.

It is said to be the **covering group** of SO(3).

The subgroup {𝕀, -𝕀} is obviously a normal subgroup Z_2 of SU(2), and SU (2)/Z_2 ≅ SO(3).

Every point in SO(3) corresponds to two points in SU(2).

Every representation Γ of SU (2) thus provides two matrices for each element of SO(3), and the product of one of these for A and one for B will give one of the two matrices for AB, but in general there is no way to select, for each element g ∈ SO(3), a unique choice $\Gamma^{j}(g)$ such that the product for two elements will always give the correct choice for the product.

Instead, we have

$$\Gamma(g_1)\Gamma(g_2) = \pm\Gamma(g_1 g_2).$$

This is familiar from quantum mechanics.

Representations of the rotation group with j = (2n+1)/2 are not really representations at all, because they change sign under rotation by 2π.

These representations are used by fermions, and we escape the ill-definedness of the representation by insisting that only quadratic expressions in the fermions have physical meaning.

Most of our understanding of Lie groups comes from studying the Lie algebra of the generators.

The study of these will permit us to find the representations, and to classify all finite dimensional compact simply connected Lie groups.

We will do so following the book shortly.

3.3 Adjoint Representation, Killing Form, etc.

We will assume our algebra is finite dimensional and over the reals.

From now on we will use Physicist's generators, so

$$g = e^{i\omega^i L_i}, \qquad [L_i, L_j] = ic_{ij}{}^k L_k.$$

Every finite dimensional Lie group has an **adjoint representation**, given by

$$\Gamma^{\rm adj}_{jk}(L_i) = ic_{ji}{}^k.$$

Note

$$\begin{aligned}
\Gamma^{\rm adj}_{ab}\left([L_i, L_j]\right) &= ic_{ij}{}^k \Gamma^{\rm adj}_{ab}\left(L_k\right) = -c_{ij}{}^k c_{ak}{}^b \\
&= c_{ai}{}^k c_{jk}{}^b + c_{ja}{}^k c_{ik}{}^b = \left(ic_{ai}{}^k\right)\left(ic_{kj}{}^b\right) - \left(ic_{aj}{}^k\right)\left(ic_{ki}{}^b\right) \\
&= \Gamma^{\rm adj}_{ak}\left(L_i\right)\Gamma^{\rm adj}_{kb}\left(L_j\right) - \Gamma^{\rm adj}_{ak}\left(L_j\right)\Gamma^{\rm adj}_{kb}\left(L_i\right) \\
&= \left[\Gamma^{\rm adj}\left(L_i\right), \Gamma^{\rm adj}\left(L_j\right)\right]_{ab},
\end{aligned}$$

verifying that it is a representation of the Lie algebra.

The adjoint representation has the same dimension as the algebra.

It is used to construct a bilinear form on the algebra $\beta : \mathcal{L} \times \mathcal{L} \to \mathrm{R}$ given by its action on the generators

$$\beta(L_i, L_j) = \mathrm{Tr}\left(\Gamma^{\mathrm{adj}}(L_i)\,\Gamma^{\mathrm{adj}}(L_j)\right) = -c_{ai}{}^{b} c_{bj}{}^{a} = \beta_{ij}.$$

Note that although we have traced the product of the Γ's, we still have the indices i and j left over, and in these $\beta(L_i, L_j)$ is a symmetric real matrix, which can be diagonalized.

Doing so corresponds simply to a change in basis L_i of the vector space $\mathcal{L}$.

So $\beta(L_i, L_j) = k_i \delta_{ij}$ in this basis.

Furthermore, by changing the scale of the basis vectors L_i, we can change the magnitude of k_i

But we cannot change the sign or whether or not it is zero.

We could, however, normalize our L_i so that each k_i is ± 1 or 0.

The form β is called the **Killing form**.

The singularity of the matrix β (the existence of $k_i = 0$) is tied up with whether or not there is an abelian invariant subalgebra.

An **ideal** or **invariant subalgebra** $\mathcal{H}$ of $\mathcal{L}$ is a subspace such that $[\mathcal{H}, \mathcal{L}] \in \mathcal{H}$, that is, $\forall\, h \in \mathcal{H}$, $\forall \ell \in \mathcal{L}$, we have $[h, \ell] \in \mathcal{H}$.

An invariant subalgebra generates a *normal subgroup*.

$\mathcal{H}$ is abelian if $\forall\ h_1, h_2 \in \mathcal{H}$, $[h_1, h] = 0$.

If an algebra has no nontrivial invariant subalgebra it is called **simple**.

Here trivial means either the whole algebra L or the algebra {0} consisting only of the zero element.

If an algebra has no nontrivial abelian invariant subalgebra it is **semisimple**.

Theorem: β is a singular matrix if and only if $\mathcal{L}$ is not semisimple.

Theorem: $\mathcal{L}$ is semisimple if and only if it is the direct sum of simple ideals.

__Note__: __Ideals__ in Lie algebras correspond to normal subgroups of groups like subalgebras correspond to subgroups. A __simple ideal__ in Lie algebra is a non-abelian ideal that does not contain any nonzero proper ideals. In the context of Lie algebras, a simple Lie algebra itself is one that contains no nonzero proper ideals, making it a fundamental building block in the classification of Lie algebras.

Example: SO(3) or SU(2):

$$\begin{aligned} c_{ij}{}^{k} &= \epsilon_{ijk} \\ \beta_{ij} &= -\sum_{ab} \epsilon_{aib}\epsilon_{bja} = 2\delta_{ij} \end{aligned}$$

which is already diagonalized and all k_i normalized to be equal (although not to 1).

This is a nonsingular matrix, so the algebra is semisimple.

In fact, as any rotational direction can be rotated into any other, it is simple.

3.4 Quantum Operators

Let us return to compact semisimple algebras generating symmetry groups which do have unitary representations.

The states of a physical system having a symmetry group G transform under symmetry transformations according to some representation.

That is, we can find a basis of states | i > and the operators of the group are unitary operators on the states,

$$|i\rangle \to e^{i\vec{\nu}\cdot\vec{L}}\,|i\rangle = \sum_j |j\rangle\,\Gamma_{ji}\big(e^{i\vec{\nu}\cdot\vec{L}}\big).$$

Bras are the hermitian conjugates.

Assuming unitarity,

$$\langle i| \to \langle i|\, e^{-i\vec{\nu}\cdot\vec{L}} = \sum_j \Gamma_{ij}\big(e^{-i\vec{\nu}\cdot\vec{L}}\big)\,\langle j|$$

as Γ is unitary.

Note:

$$\begin{aligned}\langle k|\ell\rangle &\rightarrow \sum_{jm} \Gamma_{kj}\left(e^{-i\vec{\nu}\cdot\vec{L}}\right)\langle j|m\rangle\,\Gamma_{m\ell}\left(e^{i\vec{\nu}\cdot\vec{L}}\right)\\ &= \sum_{jm} \Gamma_{kj}\left(e^{-i\vec{\nu}\cdot\vec{L}}\right)\delta_{jm}\,\Gamma_{m\ell}\left(e^{i\vec{\nu}\cdot\vec{L}}\right) = \Gamma_{k\ell}(\mathbb{I}) = \delta_{k\ell}.\end{aligned}$$

so the scalar products are invariant under the action of the group, for unitary representations.

Consider an operator $\mathcal{O}$ which corresponds to a physical variable p.

If I measure p in a state ψ, I get various values averaging to p = $\langle\, \psi \mid \mathcal{O} \mid \psi \,\rangle$.

If I ask how a physical variable is changed under the action of a symmetry transformation, I measure the new value p′ by inserting $\mathcal{O}$ between the transformed states

$$p' = \langle\psi'|\,\mathcal{O}\,|\psi'\rangle = \langle\psi|\,e^{-i\vec{\nu}\cdot\vec{L}}\mathcal{O}e^{i\vec{\nu}\cdot\vec{L}}\,|\psi\rangle\,.$$

Thus I can equivalently think of the transformation as acting on the operator and leaving the states alone:

$$\mathcal{O} \rightarrow \mathcal{O}' = e^{-i\vec{\nu}\cdot\vec{L}}\mathcal{O}e^{i\vec{\nu}\cdot\vec{L}}.$$

Under an infinitesimal transformation ν,

$$\delta\mathcal{O} = \mathcal{O}' - \mathcal{O} = -i\nu_a\,[L_a, \mathcal{O}]\,.$$

There is some advantage to thinking of the symmetry operations as passive(as opposed to active), a kind of compensating transformation,.

This is especially useful for rotations, as the symmetry transformation can be thought of as not affecting the state at all, but only its description in terms of a rotated coordinate systems.

Suppose ψ and $\mathcal{O}$ are a wave function and an operator described with respect to an original coordinate system, and ψ' and $\mathcal{O}'$ the same state and operator described in a rotated one.

Then the eigenvalue $\langle \psi \mid \mathcal{O} \mid \psi >$ must change only as appropriate for an object of $\mathcal{O}$'s spin.

This leads to very powerful constraints on the matrix elements, known as the Wigner-Eckart theorem.

First we will work out the needed mathematics for SU(2), and then consider an arbitrary semisimple compact Lie group.

The things we must discuss are

- Irreducible representations of the group
- Completeness and orthogonality
- The Wigner-Eckart theorem

Example 2: $SO(3)$ (to help you work it out…..)

$SO(3)$ is the group of rotations in three dimensions. Consider a rotation about the z-axis, (viewed as an active transformation $\vec{r} \to \vec{r}\,'$) :

$$\begin{aligned} x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \\ z' &= z \end{aligned} \quad \text{so} \quad \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The infinitesimal generator is therefore

$$L_z = \frac{d}{d\theta} R_z(\theta) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Similarly,

$$L_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad L_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

To calculate the structure constants we expand

$$\begin{aligned} [L_x, L_y] &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = L_z \\ [L_y, L_z] &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = L_x \\ [L_x, L_z] &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = -L_y \end{aligned}$$

This should be familiar to you except for some i's. This is because, for the moment, I am using mathematician's notation for the generators. Physicists like to think of the group elements as unitary operators but the generators as hermitian, so we write

$$U = e^{-i\Sigma_j \theta_j L_j^P}$$

with Physicist's generators

$$L_j^P = iL_j, \qquad L_j = -iL_j^P,$$

so

$$[L_x^P, L_y^P] = iL_z^P, \qquad etc., \qquad \text{or better:} \quad [L_j^P, L_k^P] = i\epsilon_{jk\ell} L_\ell^P$$

which should be more familiar[11]. We also see for a Lie algebra in general that

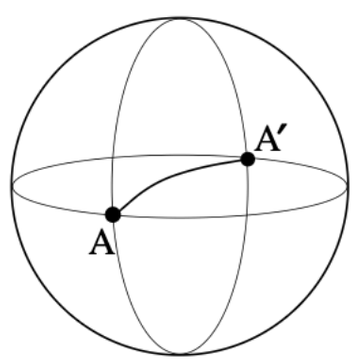
$$[L_j^P, L_\ell^P] = ic_{j\ell}{}^k L_k^P$$

with the same structure constants $c_{j\ell}{}^k$ the mathematicians use.

We see that for $SO(3)$, $c_{jk}{}^\ell = \epsilon_{jk\ell}$. Note that a rotation through angle θ about a general axis $\hat{\omega}$, (with $\hat{\omega}^2 = 1$) is given by $e^{-i\theta\hat{\omega}_j L_j^P}$. Then $\vec{\omega} = \theta\hat{\omega}$ can be used as the parameters for the group, $g(\vec{\omega}) = e^{-i\vec{\omega}\cdot\vec{L}^P}$, and the space of these parameters is a ball in three dimensions, $|\vec{\omega}| \leq \pi$.

Note that for any given axis, a rotation through π is the same transformation as a rotation about the same axis through $-\pi$. This means that the group manifold is the closed ball $|\vec{\omega}| \leq \pi$, but with the opposite ends of each diameter of the ball identified with each other.

Then the path shown is a closed path, because its ends, the points A and A' at the opposite ends of a diameter, are considered to be the same point. No matter how we try to continuously deform this path, the endpoints always stay opposite each other, and we cannot shrink the path to a point. Thus the manifold of $SO(3)$ is not simply connected.



In fact, the path AA' above can be continuously deformed into any other diameter, so any path on the $SO(3)$ manifold is deformable either into a point or into a particular diameter. In fact, AA' can be deformed into its "negative", the path taken in the reverse direction. Thus the path given by adding AA' to itself, in the sense of gluing the tail of the first path to the head of the second, is a closed path which **is** deformable to the identity[12].

Part 4 - SU(2)

4.1 Representations of SU(2)

We will now work out in detail the properties of SU(2) and its representations.

We have already seen that the generators may be chosen to be

$$L_i = \frac{1}{2}\sigma_i, \qquad \text{with } \sigma_i = \text{the Pauli matrices.}$$

Then $c^k{}_{ij} = \epsilon_{ijk}$ are the structure constants, and

$$\beta_{ij} = -\epsilon_{aib}\epsilon_{bja} = 2\delta_{ij}.$$

Thus our generators are not quite canonically normalized, but are all normalized equally, and β is positive definite.

This is related to the fact, which we have already seen, that the group is compact.

In writing down our generators, we have chosen, arbitrarily, one direction to make diagonal. them as diagonal.

Any rotation can be, by similarity transformation with at rotation, rotated into the z direction, so as we have a choice as to which of the infinite number of equivalent representations to choose, we may choose L_3 to be diagonal.

If we had several generators which commuted with each other, we could have chosen all of them to be diagonal.

In general, we will take a maximal set of commuting generators, called the Cartan subalgebra, and represent as diagonal.

For SU(2), no rotations about any other axis commute with L_3, so the Cartan subalgebra is one dimensional.

We want to look for finite dimensional irreducible representations, and we have chosen to make $\Gamma(L_3)$ diagonal.

Consider a basis vector em with $L_3 e_m = m e_m$, or $\Gamma_{m'm}(L_3) = m\delta_{m'm}$, where so far we are not restricting m other than to say it is real.

However, we know that $e^{4\pi i L_3} = \mathbb{I}$, so $e^{4\pi i m} = 1$, and m = n/2 for some integer n, for us to have a true representation of the SU(2) group.

But even without assuming this, we will find it anyway, from requiring the representation to be finite dimensional.

From L_1 and L_2 we can form the two operators

$$L_\pm = \frac{L_1 \pm iL_2}{\sqrt{2}}.$$

These are not actually part of the Lie algebra, because that is defined over the reals, but any representation of the algebra is automatically a representation of the complexification, i.e. $\{\sum v_a L_a\}$, $v_a \in C$.

$L_\pm$ are called raising and lowering operators respectively because

$$[L_3, L_\pm] = \frac{1}{\sqrt{2}}[L_3, L_1] \pm \frac{i}{\sqrt{2}}[L_3, L_2] = \frac{i}{\sqrt{2}}L_2 \pm \frac{1}{\sqrt{2}}L_1 = \pm L_\pm.$$

Thus if $L_3 e_m = m e_m$,

$$L_3 (L_\pm e_m) = (L_\pm L_3 \pm L_\pm) e_m = (m \pm 1) L_\pm e_m,$$

so $L_\pm e_m$ is a new basis vector of the representation, unless it vanishes.

Applying L_+ an arbitrary number of times generates an arbitrary number of vectors unless at some point it gives 0.

All of these vectors are linearly independent, so we would generate an infinite-dimensional space unless there exists some state $e_j \propto L_+^p e_m$ with $L_3 e_j = j e_j$ on which $L_+ e_j = 0$.

ej is called the **highest weight state**.

Let us form a normalized vector proportional to e_j calle $|j, j\rangle$, and write the inner product in quantum mechanical form

$$\langle j, j | j, j \rangle = 1.$$

Now we generate a sequence of orthonormal states

$$|j, m\rangle = N_{j,m} L_-^{j-m} |j, j\rangle ,$$

where $N_{j,m}$ is a real normalization constant.

The state $|j, m\rangle$ is an eigenstate of L_3 with eigenvalue m,

$$L_3 |j, m\rangle = m |j, m\rangle$$

because each application of L_- lowers the eigenvalue of L_3 by one.

The normalization factors can be found by observing that

$$|j, m\rangle = L_- |j, m + 1\rangle \quad \frac{N_{j,m}}{N_{j,m+1}}$$

tor $L_- |j, m + 1\rangle = N_{m+1} |j, m\rangle$.

On the other hand,

the the the

$$
\begin{aligned}
L_+ |j,m\rangle &= N_{j,m} L_+ L_-^k |j,j\rangle \quad \text{with } k = j-m \\
&= N_{j,m} \left(\sum_{r=0}^{k-1} L_-^{k-1-r} [L_+, L_-] L_-^r |j,j\rangle + L_-^k L_+ |j,j\rangle \right).
\end{aligned}
$$

But $[L_+, L_-] = 1/2\,[L_1 + iL_2, L_1 - iL_2] = i\,[L_2, L_1] = L_3$ and $L_3 L^r_- |j, j> = (j - r) L^r_- |j, j>$ as L_- is the lowering operator.

So as $L_+ |j, j> = 0$,

$$
\begin{aligned}
L_+ |j,m\rangle &= N_{j,m} \sum_{r=0}^{k-1} (j-r) L_-^{k-1} |j,j\rangle \\
&= \left[kj - \frac{k(k-1)}{2} \right] N_{j,m} L_-^{k-1} |j,j\rangle \\
&= k \left[j - \frac{k-1}{2} \right] N_{j,m} N_{j,m+1}^{-1} |j,m+1\rangle \\
&= \frac{1}{2}(j-m)(j+m+1) N_{m+1}^{-1} |j,m+1\rangle .
\end{aligned}
$$

Now $L_+ = L^\dagger_-$, so

$$
\begin{aligned}
\langle j,m+1| L_+ |j,m\rangle &= \langle j,m| L_- |j,m+1\rangle^* \\
= & = \\
\frac{1}{2}(j-m)(j+m+1) N_{m+1}^{-1} &= N_{m+1}
\end{aligned}
$$

or

$$
\begin{aligned}
N_{m+1} &= \frac{1}{\sqrt{2}}\sqrt{(j-m)(j+m+1)} \\
N_m &= \frac{1}{\sqrt{2}}\sqrt{(j-m+1)(j+m)}.
\end{aligned}
$$

Now the set of states |j, m > m = j, j-1, j-2, . . . must terminate somewhere if the representation is to be finite dimensional, so for some m < j, $N_m = 0$, so j + m = 0, or m = -j.

But j - m is an integer, so 2j is an integer, and hence 2m is an integer as well.

We have formed a representation with orthonormal basis {|j, m > , m = j, j-1, j-2,· · ·, -j}, which is 2j + 1 dimensional, with 2j an integer.

We found the representation of L_3 and $L_\pm$, from which L_1 and L_2 follow.

There is exactly one irreducible representation of SU(2) for each dimension and for each j = 0,1/2,1,3/2 , ...

Each representation has just one eigenstate of L_3 with each eigenvalue m ∈ [-j,-j + 1, . . . , j].

The representation is given by

$$
\begin{aligned}
L_3 \left|j,m\right\rangle &= m \left|j,m\right\rangle \\
L_+ \left|j,m\right\rangle &= \frac{1}{\sqrt{2}}\sqrt{(j-m)(j+m+1)}\left|j,m+1\right\rangle \\
L_- \left|j,m\right\rangle &= \frac{1}{\sqrt{2}}\sqrt{(j+m)(j-m+1)}\left|j,m-1\right\rangle
\end{aligned}
$$

the representation of the group elements is, of course, given by the exponential of the representation of the generators.

The 2j + 1 dimensional representation is usually denoted

$$\mathcal{D}^{j}_{m_1,m_2}(g)$$

The j = 1/2 representation is just the group elements themselves.

Therefore it is called the **defining representation**.

4.2 Reduction of Direct Products

If Γ^1 and Γ^2 are two representations of a Lie group, and Γ is their direct product, then

$$\Gamma_{(ix),(jy)}(g) = \Gamma^1_{ij}(g)\Gamma^2_{xy}(g).$$

Recalling that $\Gamma(L_a) = -i \left.\frac{d}{dv_a}\Gamma(g(v_a))\right|_{\substack{v=0\\g=1}}$,

$$\Gamma_{(ix),(jy)}(L) = \Gamma^1_{ij}(L)\Gamma^2_{xy}(\mathbb{I}) + \Gamma^1_{ij}(\mathbb{I})\Gamma^2_{xy}(L) = \delta_{xy}\Gamma^1_{ij}(L) + \delta_{ij}\Gamma^2_{xy}(L).$$

Now for SU(2), consider the direct product states with bases

$$|j_1, j_2; m_1, m_2\rangle = |j_1, m_1\rangle \otimes |j_2, m_2\rangle \,.$$

This representation can in principle be decomposed into a set of irreducible representations |j, m >.

To see which, we count the dimensionality of the eigenspaces of L_3,

$$\begin{aligned} L_3 \left|j_1, j_2; m_1, m_2\right\rangle &= \left(L_3 \left|j_1, m_1\right\rangle\right) \otimes \left|j_2, m_2\right\rangle + \left|j_1, m_1\right\rangle \otimes \left(L_3 \left|j_2, m_2\right\rangle\right) \\ &= \left(m_1 + m_2\right) \left|j_1, m_1; j_2, m_2\right\rangle . \end{aligned}$$

Fortunately our states are already eigenstates of L_3.

The largest eigenvalue is $j_1 + j_2$, which can occur in only one way.

So the representation $j = j_1 + j_2$ occurs exactly once in the direct product.

This representation accounts for one of the basis vectors for each m (= eigenvalue of L_3) with $| m | \leq j_1 + j_2$.

Now consider $m = m_1 + m_2 = j_1 + j_2 - 1$.

There are two ways to build it up, with $m_1 = j_1 - 1$, $m_2 = j_2$, or with $m_1 = j_1$, $m_2 = j_2 \rightarrow 1$, as long as each j is at least 1/2 .

The next higher m, $m = j_1 + j_2 - 2$, can be built three different ways.

The growth stops with $m = j_1 + j_2 - k$, which can be written in $k + 1$ ways, with $m_1 = j_1 - r$, $m_2 = j_2 - k + r$, $r = 0, \ldots, k$.

What stops it is that we must require $m_1 \geq -j_1$, and $m_2 \geq -j_2$, so $k \leq 2 \min(j_1, j_2)$.

After that, the m corresponding to the smaller j (say $j_{min} = k/2$) takes on all $2 j_{min} + 1$ values, until we get to $m < -|j_{max} - j_{min}|$, where the possible values of m for the smaller j are limited.

Example:

$$j_1 = 5/2$$
$$j_2 = 1$$

Thus the dimensionality of the m subspace is $\min(j_1 + j_2 - |m| + 1, 2j_{\min} + 1)$.

m	ways $m = m_1 + m_2$	number
$7/2$	$5/2 + 1$	1
$5/2$	$5/2 + 0$; $3/2+1$	2
$3/2$	$5/2 + (-1)$; $3/2+0$; $1/2+1$	3
$1/2$	$3/2 + (-1)$; $1/2+0$; $-1/2+1$	3
$-1/2$	$1/2 + (-1)$; $-1/2+0$; $-3/2+1$	3
$-3/2$	$-1/2 + (-1)$; $-3/2+0$; $-5/2+1$	3
$-5/2$	$-3/2 + (-1)$; $-5/2+0$	2
$-7/2$	$-5/2 + (-1)$	1

Each decrease of m gives a new state, which is the highest state of a new representation, up to the point where

$$j_1 + j_2 - m + 1 = 2j_{\min} + 1,$$

or $| m | = j_{max} - j_{min} = |j_1 \rightarrow j_2|$.

Thus

$$\Gamma^{j_1} \otimes \Gamma^{j_2} \cong \bigoplus_{j=|j_1-j_2|}^{j_1+j_2} \Gamma^j.$$

This statement is an equivalence, so there must be a unitary matrix U which connects the direct product to the direct sum.

This matrix has a right index which is a pair (m_1, m_2) and a left index which must specify which irreducible representation in the sum is referenced, and which m for that representation.

Because each j in the sum appears only once, we can use j to index the representation.

Thus the left index is (j, m).

Of course the whole matrix depends on j_1 and j_2.

It is generally written

$$(j_1, j_2, j, m|\ j_1, j_2, m_1, m_2)$$

with hermitian conjugate $(j_1, j_2, m_1, m_2 \mid j_1, j_2, j, m)$ and with

$$|j, m\rangle = \sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle\, (j_1, j_2, m_1, m_2|\ j_1, j_2, j, m)\,.$$

Both basis states are normalized, so $(j_1, j_2, m_1, m_2 \mid j_1, j_2, j, m)$ is a unitary matrix.

That does not completely define it, of course, because each of the invariant subspaces could be multiplied by an arbitrary phase $| j, m > \rightarrow \quad e^{i\phi_j} \ | j, m >$.

For each j, we choose reality and the sign convention by making this overlap for m = j and $m_1 = j_1$ be real and positive:

$$\text{SC:} \qquad (j_1, j_2, j, m = j \,|\, j_1, j_2, m_1 = j_1, m_2 = j - j_1) > 0.$$

A detailed example:

Consider the direct product of $j_1 = 1$ and $j_2 = 1/2$, as occurs if we ask about the total angular momentum of an electron (spin 1/2) in a p orbital (orbital angular momentum 1, in units of $\hbar$).

$$|j_1 = 1, m_1 = 1\rangle \otimes \left|j_2 = \frac{1}{2}, m_2 = \frac{1}{2}\right\rangle = \left|\frac{3}{2}, \frac{3}{2}\right\rangle,$$

so, abbreviating (j_1, j_2, m_1, m_2 | j_1, j_2, j, m) as (j_1, j_2, m_1, m_2 | j, m), we have (1,1/2,1,1/2 | 3/2,3/2) = 1.

Applying the lowering operator

$$\begin{aligned} L_- \left|\frac{3}{2}, \frac{3}{2}\right\rangle &= (L_- |1,1\rangle) \otimes \left|\frac{1}{2}, \frac{1}{2}\right\rangle + |1,1\rangle \otimes \left(L_- \left|\frac{1}{2}, \frac{1}{2}\right\rangle\right) \\ = \qquad\qquad & = \qquad\qquad = \\ \sqrt{\frac{1}{2}\cdot 3 \cdot 1}\left|\frac{3}{2}, \frac{1}{2}\right\rangle &= \sqrt{\frac{1}{2}\cdot 2 \cdot 1}\left|1, 0, \frac{1}{2}, \frac{1}{2}\right\rangle + \sqrt{\frac{1}{2}\cdot 1 \cdot 1}\left|1, 1, \frac{1}{2}, -\frac{1}{2}\right\rangle \\ \text{or} \qquad \left|\frac{3}{2}, \frac{1}{2}\right\rangle &= \sqrt{\frac{2}{3}}\left|1, 0, \frac{1}{2}, \frac{1}{2}\right\rangle + \sqrt{\frac{1}{3}}\left|1, 1, \frac{1}{2}, -\frac{1}{2}\right\rangle \end{aligned}$$

so (1,1/2, 0,1/2 | 3/2,1/2) = $\sqrt{\frac{2}{3}}$; , (1,1/2, 1,-1/2 | 3/2,1/2) = $\sqrt{\frac{1}{3}}$.

The state |1/2 ,1/2 > is orthogonal to |3/2 ,1/2 > but also must be made of the 2 m1 + m2 = 1/2 states, so

$$\left|\frac{1}{2},\frac{1}{2}\right\rangle = \alpha\left|1,\frac{1}{2},0,\frac{1}{2}\right\rangle + \beta\left|1,\frac{1}{2},1,-\frac{1}{2}\right\rangle .$$

Orthogonality tells us $\sqrt{2/3}\,\alpha + \sqrt{1/3}\,\beta = 0$ and normalization that $\alpha^2 + \beta^2 = 1$.

The sign convention SC says $\beta > 0$, so

$$\beta = \sqrt{\frac{2}{3}} = \left\langle \frac{1}{2},\frac{1}{2}\middle\| 1,\frac{1}{2},1,-\frac{1}{2}\right\rangle$$
$$\alpha = -\sqrt{\frac{1}{3}} = \left\langle \frac{1}{2},\frac{1}{2}\middle\| 1,\frac{1}{2},0,\frac{1}{2}\right\rangle$$

Applying L_- further to the states | 3/2 , m > and |1/2 , m > specifies the remaining coefficients without additional sign conventions.

As an exercise you should be able now to work out < 1, 1/2 , j, m | 1,1/2 , m_1, m_2 > for all nonzero elements.

These orthogonal matrices are called the **Vector coupling coefficients.**

They are more usually called, by physicists and incorrectly, **Clebsch-Gordon coefficients**.

The decomposition of the direct product of irreducible representations of a compact group can, of course, always be done, in a fashion analogous to this, as we shall see later.

There are, however, special features of SU(2) related to

1. Each representation j appears at most once in $j_1 \otimes j_2$, and all representations of a given dimension are equivalent. This is not true for SU(3), for example.

2. Because there is only one representation for each dimension, each representation is self-conjugate, $\Gamma^{j\,*} \cong \Gamma^j$. As a consequence, it is possible to treat $j_1 \otimes j_2 \rightarrow j_3$ in a symmetric fashion, defining the Wigner coefficient

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

   There is a great machinery for dealing with SU(2) representations, with such objects as 6-j and 9-j symbols, *etc.*. These are used in considering the overlaps of differing choices in how to combine the many angular momenta of components in an atom or in an atomic nucleus. References are Edmunds, Rose, and also Yutsis[6]. To see how crazy one can get, see J. Shapiro, Comp. Phys. Comm. **1**, (69) 207.

3. Starting with the defining representation, $j = \frac{1}{2}$, one can generate an arbitrary representation of spin j by taking the totally symmetric piece of the direct product of $2j$ defining representations $\frac{1}{2} \otimes \frac{1}{2} \otimes \ldots \otimes \frac{1}{2}$. Writing $\left|\frac{1}{2}, \frac{1}{2}\right\rangle$ as $\uparrow$ and $\left|\frac{1}{2}, -\frac{1}{2}\right\rangle$ as $\downarrow$, one can write such an expression, as, for example

$$|2,1\rangle = \frac{1}{2}\left(\uparrow\uparrow\uparrow\downarrow + \uparrow\uparrow\downarrow\uparrow + \uparrow\downarrow\uparrow\uparrow + \downarrow\uparrow\uparrow\uparrow\right).$$

4.3 Representations of Finite Rotations

Consider the action of a finite rotation on an irreducible representation

$$e^{i\nu L}\left|j,m,\alpha\right\rangle = \sum_{m'} \mathcal{D}^{j}_{m',m}\left(e^{i\nu L}\right)\left|j,m',\alpha\right\rangle ,$$

where α is some index describing all the other features of the states, not having to do with rotations, e.g. the principle quantum number.

Then

$$\left\langle j,m',\alpha\right| e^{i\nu L}\left|j,m,\alpha\right\rangle = \mathcal{D}^{j}_{m',m}(g), \qquad \text{where} \quad g = e^{i\nu L}.$$

Now a direct product state $|\, j_1, m_1, \alpha_1 > \otimes |\, j_2, m_2, \alpha_2 >$ can be decomposed into irreducible representations,

$$= \left|j_1,j_2,j,m,\alpha_1,\alpha_2\right\rangle \left\langle j_1,j_2,j,m|j_1,j_2,m_1,m_2\right\rangle$$

so

$$\begin{aligned}
&\left\langle j_1,m_1',\alpha_1\right| \otimes \left\langle j_2,m_2',\alpha_2\right| e^{i\nu L}\left|j_1,m_1,\alpha_1\right\rangle \otimes \left|j_2,m_2,\alpha_2\right\rangle \\
&= \mathcal{D}^{j_1}_{m_1',m_1}(g)\mathcal{D}^{j_2}_{m_2',m_2}(g) \\
&= \sum_{j,m,m'} \left\langle j_1,j_2,m_1',m_2'|j_1,j_2,j,m\right\rangle \mathcal{D}^{j}_{m',m}(g)\left\langle j_1,j_2,j,m|j_1,j_2,m_1,m_2\right\rangle
\end{aligned} \tag{4.1}$$

We have not yet shown that there exists an invariant integration measure on the group, and hence that we can prove a great orthogonality theorem.

We shall do so later, but for now let us assume it

$$\int dg\, \mathcal{D}^{j_1\,*}_{m_1',m_1}(g)\, \mathcal{D}^{j_2}_{m_2',m_2}(g) = N_{j_1}\delta_{j_1,j_2}\delta_{m_1,m_2}\delta_{m_1',m_2'}\,. \tag{4.2}$$

To get the normalization correct, normalize $\int dg\ 1 = 1$.

Now $\mathcal{D}$ is a unitary matrix, so

$$\sum_m \mathcal{D}^{j\,*}_{m,\mu}(g)\, \mathcal{D}^{j}_{m,\nu}(g) = \delta_{\mu\nu},$$

so integrating the left hand side gives

$$\delta_{\mu\nu} = N_j \delta_{\mu\nu} \underbrace{\sum_m \delta_{m,m}}_{2j+1}$$

so Nj =1/(2j + 1). [Compare this to g/li for the finite group representations.]

Applying this to (4.1)

$$\begin{aligned}
&\int dg\, \mathcal{D}^{j_3\,*}_{m_3',m_3}(g)\mathcal{D}^{j_1}_{m_1',m_1}(g)\mathcal{D}^{j_2}_{m_2',m_2}(g) \\
&= \sum_{j,m,m'} \langle j_1, j_2, m_1', m_2' | j_1, j_2, j, m' \rangle \langle j_1, j_2, j, m | j_1, j_2, m_1, m_2 \rangle \frac{\delta_{j,j_3}\delta_{m,m_3}\delta_{m',m_3'}}{2j+1} \\
&= \frac{1}{2j+1} \langle j_1, j_2, m_1', m_2' | j_1, j_2, j_3, m_3' \rangle \langle j_1, j_2, j_3, m_3 | j_1, j_2, m_1, m_2 \rangle
\end{aligned} \tag{4.3}$$

In atomic or nuclear physics, one very often wants to calculate the amplitude for emission of a photon from some excited state.

The photon emission operator is proportional to $e^{i\vec{k}\cdot\vec{r}}$, where $\vec{k}$ is the momentum of the emitted photon and $\vec{r}$ is the position operator acting on the constituents.

In fact, for nonrelativistic particles it is exactly

$$\langle\psi_f|\, e^{i\vec{k}\cdot\vec{r}}\, |\psi_i\rangle$$

which determines the transition amplitude.

Generally the transition energy corresponds to a photon momentum which is much less than the inverse size of the system, so $|\vec{k}\cdot\vec{r}| \ll 1$ whereever $\vec{r}$ has a significant matrix element.

So we can expand in a multipole expansion

$$e^{i\vec{k}\cdot\vec{r}} = 4\pi \sum_{\ell,m} i^\ell Y_\ell^{m\,*}(\hat{k}) Y_\ell^m(\hat{r}) j_\ell(|k||r|).$$

As the $j_\ell(kr)$ $(kr)^\ell$ for small kr, we can expect this expansion, plugged into our matrix element, to give a series of rapidly converging terms.

The spherical harmonics $Y_{m\ell}$, for a given ℓ, are a set of functions which transform under the spin ℓ representation of SU(2).

Consider a rotation of the particles $\vec{r}$ and their spins, but not the photon.

The states ψ_i and ψ_f are eigenstates of a spherically symmetric Hamiltonian so they have definite total angular momentum j_i and j_f , and z components m_i and m_f along with other quantum numbers α_i and α_f .

So with $| \psi_i > = | j_i , m_i , \alpha_i >$, if we consider a rotation $e^{i\nu L}$, we have

$$
\begin{aligned}
|j_i, m_i, \alpha_i\rangle \to e^{i\nu L} |j_i, m_i, \alpha_i\rangle &= \sum_{m_i'} |j_i, m_i', \alpha_i\rangle \mathcal{D}^{j_i}_{m_i', m_i} \left(e^{i\nu L}\right) \\
\langle j_f, m_f, \alpha_f| \to \langle j_f, m_f, \alpha_f| e^{-i\nu L} &= \sum_{m_f'} \mathcal{D}^{j_f *}_{m_f', m_f} \left(e^{i\nu L}\right) \langle j_f, m_f', \alpha_f| \\
e^{i\nu L} Y_{\ell m}(\hat{r}) e^{-i\nu L} &= \mathcal{D}^{\ell}_{m', m} \left(e^{i\nu L}\right) Y_{\ell m'}(\hat{r}).
\end{aligned}
$$

All taken together makes for no change, so

$$
\begin{aligned}
&\langle j_f, m_f, \alpha_f| Y_{\ell m}(\hat{r}) |j_i, m_i, \alpha_i\rangle \\
&\quad = \sum_{m' \, m_i' \, m_f'} \mathcal{D}^{j_f *}_{m_f', m_f} \mathcal{D}^{\ell}_{m', m} \mathcal{D}^{j_i}_{m_i', m_i} \langle j_f, m_f', \alpha_f| Y_{\ell m'} |j_i, m_i', \alpha_i\rangle
\end{aligned}
$$

for all ν.

Integrating over the invariant measure of the group,

$$
\langle j_f, m_f, \alpha_f| Y_{\ell m}(\hat{r}) |j_i, m_i, \alpha_i\rangle = \langle \ell j_i j_f m_f | \ell j_i m m_i\rangle \langle j_f \alpha_f || Y || j_i \alpha_i\rangle \, ,
$$

where

$$\langle j_f \alpha_f || Y || j_i \alpha_i \rangle := \frac{1}{2j_f + 1} \sum_{m' \, m'_i \, m'_f} \langle \ell j_i m' m'_i | \ell j_i j_f m'_f \rangle \langle j_f, m'_f, \alpha_f | \, Y_{\ell m'} \, | j_i, m'_i, \alpha_i \rangle \, .$$

This statement is called the **Wigner-Eckhart theorem**.

It is a nice division of the matrix elements into a piece which is pure group theory, the Clebsch-Gordon coefficient, and a piece which is more nearly pure physics, $< j_f \, \alpha_f \| Y \| j_i \, \alpha_i >$.

It is generally written for a set of operators $\mathcal{O}_{\ell m}$ which transform like $Y_{\ell m}$, for which it is written

$$\langle j_f, m_f, \alpha_f | \, \mathcal{O}^\ell_m \, | j_i, m_i, \alpha_i \rangle = \langle \ell j_i j_f m_f | \ell j_i m m_i \rangle \langle j_f \alpha_f \| \mathcal{O}^\ell \| j_i \alpha_i \rangle \, ,$$

The last factor is called the **reduced matrix element**.

4.3.1 Isospin

In elementary particle physics the number of particles is not conserved due to the possibility of pair creation.

For that reason one cannot work with N particle wave functions, but rather with a Hilbert space which contains subspaces having differing numbers of particles.

The various subspaces can be built up from the "vacuum" state $| 0 >$ by applying operators which create particles in a certain state.

Let $P^\dagger\alpha$ be the operators which creates a proton with properties specified by α (which includes momentum and spin value).

Then $P^\dagger_\alpha| 0 >$ is a single proton state, while $P^\dagger_\alpha P^\dagger_\beta |0 >$ is a state which has two protons, one with properties α and one with β.

This corresponds to the same physical state as $P^\dagger_\beta P^\dagger_\alpha| 0 >$ though there could be an arbitary phase difference.

Because they are fermions, states are "antisymmetric under interchange", which we insure by taking $P^\dagger_\alpha P^\dagger_\beta = -P^\dagger_\beta P^\dagger_\alpha$.

This automatically gives $(P^\dagger_\alpha)^2 = 0$, or the Pauli exclusion principle.

The hermitean conjugate operator P_α annihilates a proton in the state α, if it exists, or else it annihilates the state.

The algebra of the P 's is

$$\{P_\alpha, P_\beta\} = \left\{P^\dagger_\alpha, P^\dagger_\beta\right\} = 0, \qquad \left\{P_\alpha, P^\dagger_\beta\right\} = \delta_{\alpha\beta}.$$

A similar set of operators exist to create neutrons, $N^\dagger\alpha$, or electrons $E^\dagger\alpha$.

A state consisting on one proton and one neutron can be described by$P^\dagger_\alpha N^\dagger_\beta | 0 >$ or by $N^\dagger_\beta P^\dagger_\alpha | 0 >$.

These can differ by a phase, which is really a matter of convenience.

We choose anticommuting operators $\{P^\dagger_\alpha, N^\dagger_\beta\} = 0$ for all different fermion fields.

Representing different particles, the P and N operators must have trivial commutation relations, and as the creation operators have been chosen to anticommute, we also need $\{P^\dagger_\alpha, N^\dagger_\beta\} = 0$.

Now consider the operator

$$\sum_\alpha P^\dagger_\alpha N_\alpha.$$

It looks around for a neutron to annihilate and, when it finds one, replaces it with a proton of the same momentum and spin.

This is not a "physical" operator in the sense that there is no way one can do just this, but it is a mathematical object with an interesting consequence.

Before we consider this further, note that $\sum_\alpha P^\dagger_\alpha P_\alpha$ looks around for protons it can annihilate, does so but then recreates it and adds a notch to its belt, so in the end it does nothing but count the number of protons in the state.

In nuclear physics, the bulk of the Hamiltonian is the "strong" or nuclear interactions, which seem to treat the protons and neutrons as equivalent.

Of course protons and neutrons have different charges, so the electromagnetic (and electroweak) interactions do distinguish between them.

Let us write the hamiltonian as $H_S + H_{EW}$ and write the fact that H_S does not distinguish between protons and neutrons by

$$\left[H_S, \sum_\alpha P_\alpha^\dagger N_\alpha\right] = 0.$$

This is because it doesn't matter if you first replace a neutron by a proton and then let H_S act, or let H_S act first and then do the replacement, as H_S acts the same way on both particles.

The hermetian conjugate gives $[H_S\ , \sum_\alpha N^\dagger_\alpha P_\alpha\ ] = 0$.

Let's forget H_{EW} for now, and consider

$$T^+ = \frac{1}{\sqrt{2}}\sum_\alpha P_\alpha^\dagger N_\alpha \qquad \text{so} \qquad \left[H, T^+\right] = \left[H, T^-\right] = 0$$

$$T^- = \frac{1}{\sqrt{2}}\sum_\alpha N_\alpha^\dagger P_\alpha$$

Let

$$T_3 = \left[T^+, T^-\right] = \frac{1}{2}\sum_{\alpha\beta}\left[P_\alpha^\dagger N_\alpha, N_\beta^\dagger P_\beta\right] = \frac{1}{2}\sum_{\alpha\beta}\left(P_\alpha^\dagger N_\alpha N_\beta^\dagger P_\beta - N_\beta^\dagger P_\beta P_\alpha^\dagger N_\alpha\right)$$

$$= \frac{1}{2}\sum_{\alpha\beta}\Big(P_\alpha^\dagger\left\{N_\alpha, N_\beta^\dagger\right\}P_\beta - \left\{P_\alpha^\dagger, N_\beta^\dagger\right\}N_\alpha P_\beta + N_\beta^\dagger P_\alpha^\dagger\left\{N_\alpha, P_\beta\right\}$$

$$-N_\beta^\dagger\left\{P_\alpha^\dagger, P_\beta\right\}N_\alpha\Big)$$

$$= \frac{1}{2}\sum_{\alpha\beta}\left(P_\alpha^\dagger\delta_{\alpha\beta}P_\beta - 0 + 0 - N_\beta^\dagger\delta_{\alpha\beta}N_\beta\right)$$

$$= \frac{1}{2}\sum_\alpha\left(P_\alpha^\dagger P_\alpha - N_\alpha^\dagger N_\alpha\right).$$

We see that $2T_3$ counts the number of protons minus the number of neutrons.

As

$$[H, T_3] = [H, [T^+, T^-]] = -[T^+, [T^-, H]] - [T^-, [H, T^+]]$$

by the Jacobi identity, and as both terms vanish because $[H, T_\pm] = 0$, we have $[H, T_3] = 0$.

One can also show, as for angular momentum, that $[T_3, T^\pm] = \pm T^\pm$, so the T 's (or more precisely T_3, $T_1 = (T^+ + T^-)/\sqrt{2}$ i , and $T_2 = (T^+ - T^-)\sqrt{2}$) are the generators of an SU(2) symmetry group under which the strong interaction hamiltonian is invariant.

Note that T^+ increases the charge by 1|e|, so a multiplet consists of a sequence of charges differing by e.

To the extent that H_{EW} can be ignored, states should form multiplets (irreducible representations) of this isotopic spin symmetry.

For example:

The latter triple includes two unstable isotopes, C^{14} and O^{14}, and an excited state of N^{14}, all with spin 0. The ground state of N^{14} is a T = 0, J = 1 state.

$(p,\ n)$	$T = \frac{1}{2}$
deuteron	$T = 0$
H^3, He^3	$T = \frac{1}{2}$
C^{14}, N^{14*}, O^{14}	$T = 1$

Although isospin was introduced in nuclear physics, it is applicable in elementary particle physics as well.

In high energy interactions many particles, mostly unstable, are observed.

One set of three pseudoscalar particles, the pions, π^+, π^0 , π^-, form a T = 1 triplet, the lightest hadrons.

$\pi^\pm$ are unstable particles living for about 26 nanoseconds, but this is long enough to make a beam. π^0 lives only 10^{-16} s, so no beam is possible, and π^0's can be detected only from their decay into two photons.

This very different behavior is due to the electromagnetic interactions, which violate isospin symmetry.

Now if we consider scattering of π mesons o! nucleons (i.e. protons or neutrons), the conservation of isospin will relate different amplitudes.

The simplest application is the resonant state produced in π + p scattering, called the Δ^{++}, at a mass of 1232 MeV.

It must be a T = 3/2 state because it is $\pi^+ \otimes$ p or $|1, 1> \otimes |\ 1/2\ ,1/2 = |\ 3/2, 3/2>$.

This must be part of an irreducible representation of the (unphysical) rotations in isospace, with partners

Name	Q	rep
Δ^{++}	2e	$\left\|\frac{3}{2}, \frac{3}{2}\right\rangle$
Δ^{+}	e	$\left\|\frac{3}{2}, \frac{1}{2}\right\rangle$
Δ^{0}	0	$\left\|\frac{3}{2}, -\frac{1}{2}\right\rangle$
Δ^{-}	$-e$	$\left\|\frac{3}{2}, -\frac{3}{2}\right\rangle$

So observing the Δ^{++} resonance implies the existence of these three other particles, all of which are observed, though shortlived, particles.

When the Δ^{++} decays, it must always produce a p and a π^+, as there is no other choice.

On the other hand, the decay of the Δ^+ can give $p\pi^0$ or $n\pi^+$.

The decay amplitudes are given by the Wigner-Eckhart theorem:

$$\frac{\langle \pi^0 p|\Delta^+\rangle}{\langle \pi^+ n|\Delta^+\rangle} = \frac{\langle 1,\tfrac{1}{2},0,\tfrac{1}{2}|3/2,1/2\rangle}{\langle 1,\tfrac{1}{2},1,-\tfrac{1}{2}|3/2,1/2\rangle} = \frac{\sqrt{2/3}}{\sqrt{1/3}} = \sqrt{2}.$$

The decay probability is proportional to the amplitude squared, so

$$\begin{aligned} \Delta^+ &\rightarrow \pi^0 p \qquad 2/3 \text{ of the time} \\ \Delta^+ &\rightarrow \pi^+ n \qquad 1/3 \text{ of the time.} \end{aligned}$$

The Δ resonance is the most prominent feature of low energy πN scattering, and this prediction is the simplest.

But the whole scattering amplitude can be analyzed as well.

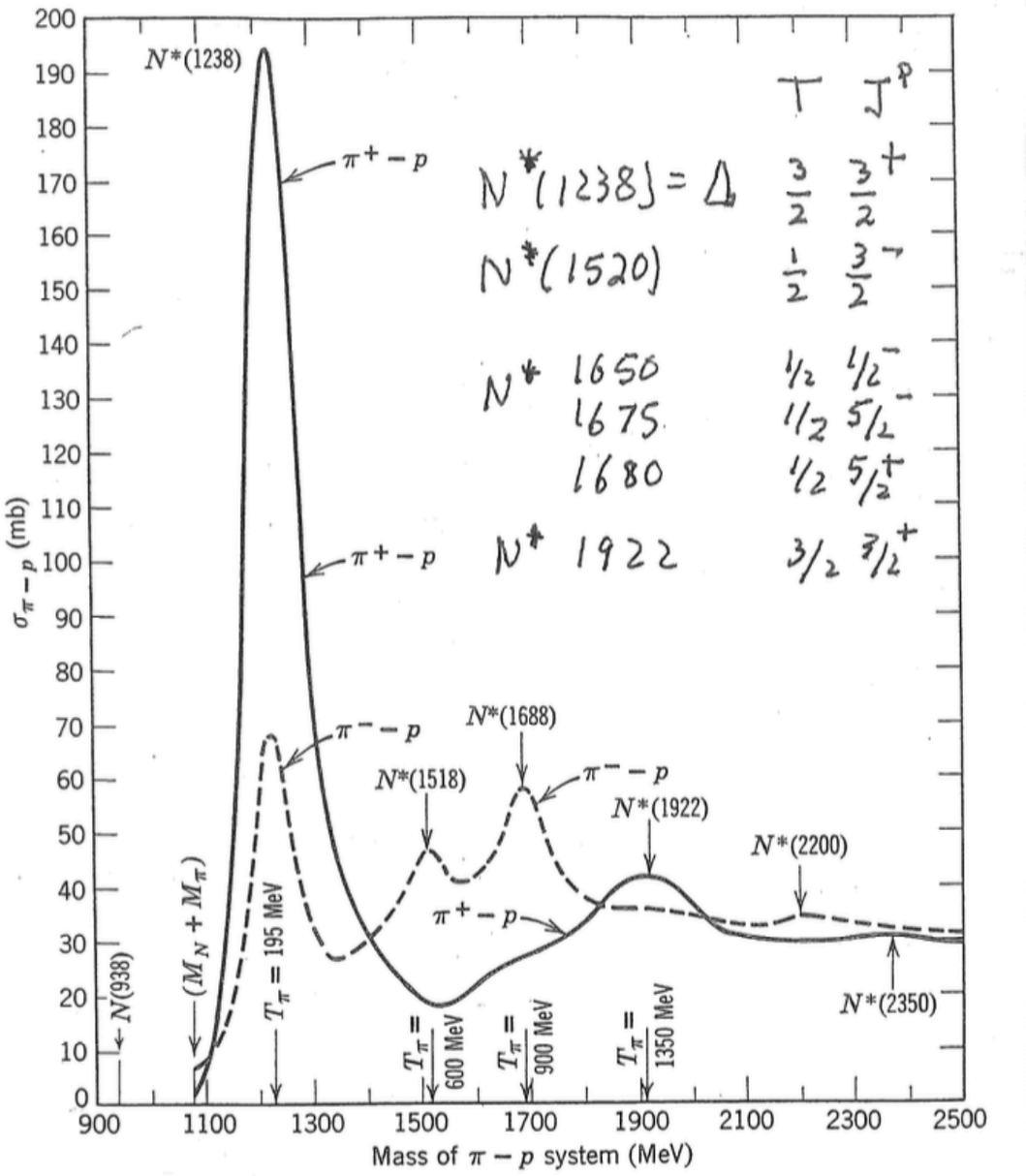


Fig. 19.3. Total cross sections for π^+p and π^-p scattering.

For example, the amplitude for $\pi^+ n \rightarrow \pi^0 p$, at a certain angle and with given polarization, is given by

$$\left\langle \pi^0 p \,\text{final} \right| S \left| \pi^+ n \,\text{init} \right\rangle$$

where the S or scattering matrix is isospin invariant.

This amplitude will be related to others, say $\pi^+ p \rightarrow \pi^+ p$, $\pi^- p \rightarrow \pi^- p$, etc.

If we write p = N, 1/2 and n = N, -1/2 we can think of the above as

$$\left\langle \pi, N, 0, \frac{1}{2} \right| S \left| \pi, N, 1, \frac{-1}{2} \right\rangle .$$

Inserting a complete set of isospin states, which can only have T = 1/2 or T = 3/2, we have

$$\begin{aligned}
&\left\langle \pi N 0 \frac{1}{2} \middle\| \pi N \frac{3}{2}\frac{1}{2} \right\rangle \left\langle \pi N \frac{3}{2}\frac{1}{2} \middle| S \middle| \pi N \frac{3}{2}\frac{1}{2} \right\rangle \left\langle \pi N \frac{3}{2}\frac{1}{2} \middle\| \pi N 1 \frac{-1}{2} \right\rangle \\
&\quad + \left\langle \pi N 0 \frac{1}{2} \middle\| \pi N \frac{1}{2}\frac{1}{2} \right\rangle \left\langle \pi N \frac{1}{2}\frac{1}{2} \middle| S \middle| \pi N \frac{1}{2}\frac{1}{2} \right\rangle \left\langle \pi N \frac{1}{2}\frac{1}{2} \middle\| \pi N 1 \frac{-1}{2} \right\rangle \\
&= \sum_{j=\frac{1}{2},\frac{3}{2}} \left\langle 1\frac{1}{2} j \frac{1}{2} \middle\| 1\frac{1}{2} 0 \frac{1}{2} \right\rangle \left\langle 1\frac{1}{2} j \frac{1}{2} \middle\| 1\frac{1}{2} 1 \frac{-1}{2} \right\rangle a_j = \frac{\sqrt{2}}{3} a_{3/2} - \frac{\sqrt{2}}{3} a_{1/2}
\end{aligned}$$

where the $a_j = <$ π N j 1/2 I S I πN j 12 $>$ are independent of m.

Is this progress?

We have replaced our scattering amplitude by a combination of two undetermined parameters, but we also have the other *measureable* amplitudes

$$
\begin{aligned}
\pi^+ p \to \pi^+ p &= a_{3/2} \\
\pi^- p \to \pi^- p &= \frac{1}{3} a_{3/2} + \frac{2}{3} a_{1/2} \\
\pi^+ n \to \pi^+ n &= \frac{1}{3} a_{3/2} + \frac{2}{3} a_{1/2} \\
\pi^- n \to \pi^- n &= a_{3/2} \\
\pi^- p \to \pi^0 n &= \frac{\sqrt{2}}{3} a_{3/2} - \frac{\sqrt{2}}{3} a_{1/2}
\end{aligned}
$$

So these five measurable cross sections should all be described by two scattering amplitudes.

Also determined are the undoable reactions

$$
\begin{aligned}
\pi^0 p &\to \pi^0 p \\
\pi^0 p &\to \pi^+ n \\
\pi^0 n &\to \pi^- p \\
\pi^0 n &\to \pi^0 n
\end{aligned}
$$

so it definitely is progress, has been tested, and agrees well with experiment.

If we examine the conserved quantum numbers of the particles we have discussed so far, we see $Q = 1/2\ B + T_3$, where B is the baryon number, one for nucleons and zero for pions.

Name	T_3	Q	Baryon number
n	$-1/2$	0	1
p	$1/2$	1	1
π^+	1	1	0
π^0	0	0	0
π^-	-1	-1	0

All of the particles we've considered so far are now considered to be made of two kinds of quarks, u and d, and their antiparticles $\bar{u}$ and $\bar{d}$.

u and d form an isospin 1/2 multiplet and thus have $T_3 = \pm 1/2$.

	T_3	Q	B
u	$1/2$	$2/3$	$1/3$
d	$-1/2$	$-1/3$	$1/3$

A Δ++ is made of three u quarks, so they have $Q = 2/3$ and $B = 1/3$.

A proton is uud and a neutron udd.

As the u and d quarks satisfy the $Q = B/2 + T_3$ relation, so will anything made up of them, as Q, T_3 and B are all arithmetically additive quantum numbers.

But not everything is made of u and d.

Physics learned about strangeness in 1953 and proposed the strange quark at the same time as u and d.

The strange quark has no isospin, Q = -1/3 and B = 1/3.

The isospin symmetry which rotated protons into neutrons, rotating u's into d's, can be extended to a bigger approximate symmetry group SU(3) by considering rotations into the strange (s) quark direction as well as the u and d directions

But before we launch into a big discussion of SU(3), we should note that the charmed quark was discovered in 1974 and the bottom quark in 1977.

Finally the sixth quark, the top, was discovered at Fermilab in 1994.

So it is time to consider groups more generally, so we can do all possibilities at once.

Part 5 Semisimple Compact Lie Groups

We return now to considering a general finite dimensional semisimple compact Lie group and its Lie algebra.

In the algebra there are many abelian subalgebras, though not invariant.

For example, any one dimensional subspace is an abelian subalgebra.

If there are several generators which commute with each other, then one can form a larger abelian subspace.

Let us take a maximal abelian subalgebra H $\subset \mathcal{L}$, of dimension m, which means there is no abelian subalgebra of larger dimension.

H is called a **Cartan subalgebra**, with basis H_i.

m is the rank of $\mathcal{L}$.

Consider a representation $\mathcal{D}$ of $\mathcal{L}$.

The matrices corresponding to the H_i may all be simultaneously diagonalized, as they commute, so we label the basis vectors μ, we have

$$H_i \, |\mu D\rangle = \mu_i \, |\mu D\rangle \, .$$

The eigenvalues μ_i are called the **weights** and the vector $\vec{\mu} = (\mu_1, \mu_2, \ldots, \mu_m)$ is called the **weight vector** corresponding to the basis vector μ.

Notice this is a vector in a m-dimensional space, where m is the rank, not the dimension of the full Lie algebra $\mathcal{L}$ or of the representation.

There is a *vector* $\vec{\mu}$ for each dimension μ of the *representation*.

But unlike for SU(2), there may be several basis vectors in the representation with the same μ, so we ought to include a subsidiary label x, $H_i |\mu, x, D> = \mu_i |\mu, x, D>$

We consider in particular the adjoint representation.

Then the basis vectors are also the Lie algebra generators L_a.

From compact semisimplicity we know we can choose Tr $L_a L_b = \lambda \delta_{ab}$ for the real basis vectors.

The generators act on the states by

$$L_a \left|L_b\right\rangle = \left|[L_a, L_b]\right\rangle = i f_{ab}{}^{c} \left|L_c\right\rangle .$$

We will consider states in the complexification of $\mathcal{L}$, so as to include things like $J_{\pm}$, so we can write in general

$$\langle X_1 | X_2 \rangle = \lambda^{-1} \operatorname{Tr}\left(X_1^{\dagger} X_2\right)$$

for X_1 and X_2 any elements of the complexification of $\mathcal{L}$, $X = \sum_a \nu_a L_a$ with $\nu_a \in C$.

From above, $H_i | H_j > = 0$.

The rest of the vector space $\mathcal{L}$ has a basis which, having diagonalized H, satisfies

$$H_i \left|E_{\vec{\alpha}}\right\rangle = \alpha_i \left|E_{\vec{\alpha}}\right\rangle .$$

The H_i are hermitean and diagonal, but the eigenvectors $E_{\vec{\alpha}}$ are not necessarily real combinations of the hermitean L_i, so $E_{\vec{\alpha}}$ is not necessarily equal to $E^\dagger_{\vec{\alpha}}$.

Nonetheless α_i is real, as all the eigenvalues of a hermitean matrix are, so, as

$$\begin{aligned} [H_i, E_{\vec{\alpha}}] &= \alpha_i E_{\vec{\alpha}}, \\ \left[H_i, E^\dagger_{\vec{\alpha}}\right] &= -\alpha_i E^\dagger_{\vec{\alpha}}. \end{aligned}$$

For each $\vec{\alpha}$, , the α_i, i = 1, . . . m, are not all zero because we assumed $E_{\vec{\alpha}}$ is not in the Cartan subalgebra.

So $E^\dagger_{\vec{\alpha}}$. is another eigenvector $E_{-\vec{\alpha}}$. .

Normalize the $E_{\vec{\alpha}}$'s so that $\langle E_{\vec{\alpha}}, E_{\vec{\beta}} \rangle = \delta_{\vec{\alpha},\vec{\beta}}$ as well as $\langle H_i, H_j \rangle = \delta_{ij}$.

The $\{\alpha_i\}$ for all the $E_{\vec{\alpha}}$ are the non-zero weights of the adjoint representation, and are called the roots of the algebra.

Again, each root is in the m-dimensional vector space.

The $E_{\vec{\alpha}}$ generators act as raising operators in some direction for any representation.

If we have a state of weight $\vec{\mu}$ in representation D, (so $H_i \mid \mu D \rangle = \mu_i \mid \mu D \rangle$) then $E_{\vec{\alpha}} \mid \mu D \rangle$ is a state of weight $\vec{\mu} + \vec{\alpha}$, for

$$H_i E_{\vec{\alpha}} |\mu D\rangle = \underbrace{[H_i, E_{\vec{\alpha}}]}_{\alpha_i E_{\vec{\alpha}}} |\mu D\rangle + E_{\vec{\alpha}} \underbrace{H_i |\mu D\rangle}_{\mu_i |\mu D\rangle} = (\alpha_i + \mu_i) E_{\vec{\alpha}} |\mu D\rangle ,$$

unless, of course, it vanishes with $E_{\vec{\alpha}} |\mu D\rangle = 0$.

In the adjoint representation consider the state $E_{\vec{\alpha}} |E_{-\vec{\alpha}}\rangle$.

It has weight 0, so commutes with each H_i and must be contained in the Cartan subalgebra, for otherwise it could have been added to it.

Thus $E_{\vec{\alpha}} |E_{-\vec{\alpha}}\rangle = \sum_i \beta_i |H_i\rangle$, or $[E_{\vec{\alpha}}, E_{-\vec{\alpha}}] = \sum_i \beta_i H_i$. As $[E_{\vec{\alpha}}, E_{-\vec{\alpha}}]$.

As $[E_{\vec{\alpha}}, E_{-\vec{\alpha}}]$ is hermitian, β_i is real.

Note

$$\beta_i = \langle H_i | E_{\vec{\alpha}} |E_{-\vec{\alpha}}\rangle = \langle E_{-\vec{\alpha}} | E_{-\vec{\alpha}} |H_i\rangle^* = \langle E_{-\vec{\alpha}} | [E_{-\vec{\alpha}}, H_i]\rangle^* = \alpha_i^* = \alpha_i,$$

so

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

Now consider an arbitrary finite dimensional representation D and a state of definite weight $\vec{\mu}$.

Just as we did with L_+ in SU(2), we will generate states with $E_{\pm\vec{\alpha}}$ until they vanish.

Consider a specific $E_{\vec{\alpha}}$, and apply $E_{\vec{\alpha}}$ until you get to a state $|\vec{\mu}'D\rangle$ which vanishes upon further application of $E_{\vec{\alpha}}$, so $E_{\vec{\alpha}}|\vec{\mu}'D\rangle = 0$, with $\vec{\mu}' = \vec{\mu} + p\vec{\alpha}$ for some integer $p \geq 0$.

Normalize the state $|\vec{\mu}'D\rangle$, and generate therefrom a sequence of normalized states

$$|\vec{\mu}' - n\vec{\alpha}, D\rangle = N_n^{-1} E_{-\vec{\alpha}} |\vec{\mu}' - (n-1)\vec{\alpha}, D\rangle$$

as long as $E_{-\vec{\alpha}}$ can be applied without killing the state.

The N_n is a real positive normalization factor so that all the states are normalized.

Thus

$$|\vec{\mu}' - n\vec{\alpha}, D\rangle = \left(\prod_{r=1}^{n} N_r^{-1}\right) E_{-\vec{\alpha}}^n |\vec{\mu}', D\rangle\,,$$

$$E_{\vec{\alpha}} |\vec{\mu}' - n\vec{\alpha}, D\rangle = \left(\prod_{r=1}^{n} N_r^{-1}\right) \sum_{r=0}^{n-1} E_{-\vec{\alpha}}^{n-r-1} [E_{\vec{\alpha}}, E_{-\vec{\alpha}}] E_{-\vec{\alpha}}^r |\vec{\mu}', D\rangle$$

$$= \left(\prod_{r=1}^{n} N_r^{-1}\right) \sum_{r=0}^{n-1} E_{-\vec{\alpha}}^{n-r-1} \underbrace{\sum_i \alpha_i H_i E_{-\vec{\alpha}}^r |\vec{\mu}', D\rangle}_{(\vec{\mu}' - r\vec{\alpha})\cdot\vec{\alpha}\, E_{-\vec{\alpha}}^r |\vec{\mu}', D\rangle}$$

$$= N_n^{-1} \left(n(\vec{\mu}'\cdot\vec{\alpha}) - \frac{1}{2} n(n-1)\alpha^2\right) \underbrace{\left(\prod_{r=1}^{n-1} N_r^{-1}\right) E_{-\vec{\alpha}}^{n-1} |\vec{\mu}', D\rangle}_{|\vec{\mu} - (n-1)\vec{\alpha}, D\rangle}\,.$$

So

$$\begin{aligned}\langle\vec{\mu}-(n-1)\vec{\alpha},D|\,E_{\vec{\alpha}}\,|\vec{\mu}'-n\vec{\alpha},D\rangle &= \frac{n(\vec{\mu}'\cdot\vec{\alpha})-\frac{1}{2}n(n-1)\alpha^2}{N_n}\\ =\langle\vec{\mu}'-n\vec{\alpha},D|\,E_{-\vec{\alpha}}\,|\vec{\mu}-(n-1)\vec{\alpha},D\rangle^* &= N_n,\end{aligned}$$

so $|N_n|^2 = n\vec{\mu}'\cdot\vec{\alpha}-\frac{1}{2}n(n-1)\alpha^2.$

For a finite dimensional representation, $N_{n+1} = 0$ for some $n \geq 0$, so

$$(n+1)\left[\vec{\mu}'\cdot\vec{\alpha}-\frac{1}{2}n\alpha^2\right]=0, \qquad \text{or} \quad \frac{\vec{\mu}'\cdot\vec{\alpha}}{\alpha^2}=\frac{n}{2}.$$

For the original vector $\vec{\mu}$,

$$\frac{\vec{\mu}\cdot\vec{\alpha}}{\alpha^2}=\frac{(\vec{\mu}'-p\vec{\alpha})\cdot\vec{\alpha}}{\alpha^2}=\frac{n}{2}-p.$$

If we had applied the lowering operator $E_{-\vec{\alpha}}$ to $|\vec{\mu},D\rangle$ directly, we would have gotten zero after some integer number q + 1 of applications of $E_{-\vec{\alpha}}$, with n + 1 = q + 1 + p, so we may write

$$\frac{\vec{\mu}\cdot\vec{\alpha}}{\alpha^2}=\frac{q-p}{2}.$$

the the the the

In the adjoint representation $\vec{\mu}$ is a root, say $\vec{\beta}$, so $\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = \frac{n}{2}$ for every pair of roots $\vec{\alpha}$ and $\vec{\beta}$, for some integer n.

Multiplying by the same relation with $\vec{\alpha} \leftrightarrow \vec{\beta}$,

$$\frac{(\vec{\alpha}\cdot\vec{\beta})^2}{\alpha^2\beta^2} = \cos^2\theta = \frac{m}{4},$$

with m an integer which must be 0, 1, 2, 3, or 4.

θ is the angle between the two roots.

Thus the possible angles between roots are 0, 30◦, 45◦, 60◦, 90◦,120◦, 135◦, 150◦, 180◦.

Of course every root has an angle of 0◦ with itself and 180◦ with its conjugate.

We now show that the angle between two different generators E_α and E_β cannot be 0.

For suppose $\vec{\alpha} \parallel \vec{\beta}$ with $|\vec{\beta}| \geq |\vec{\alpha}|$.

Then as

the only possibilities are $\vec{\beta} = 2\vec{\alpha}$ or $\vec{\beta} = \vec{\alpha}$.

First consider the possibility that $\vec{\alpha} = \vec{\beta}$, that two different vectors $E_{\vec{\alpha}}$ and $E'_{\vec{\alpha}}$ have the same root.

They can be chosen orthogonal.

Now $E_{-\vec{\alpha}} |E'_{\vec{\alpha}}\rangle = \sum_i \beta_i H_i$ as it has weight zero.

Then

$$\begin{aligned} \beta_i &= \langle H_i| E_{-\vec{\alpha}} |E'_{\vec{\alpha}}\rangle = \langle E'_{\vec{\alpha}}| E_{\vec{\alpha}} |H_i\rangle^* = \langle E'_{\vec{\alpha}}|[E_{\vec{\alpha}}, H_i]\rangle^* = -\langle E'_{\vec{\alpha}}|[H_i, E_{\vec{\alpha}}]\rangle^* \\ &= -\langle E'_{\vec{\alpha}}| H_i |E_{\vec{\alpha}}\rangle^* = -\alpha_i \langle E'_{\vec{\alpha}}|E_{\vec{\alpha}}\rangle^* = 0, \end{aligned}$$

so the number of times $E_{-\vec{\alpha}}$ can lower $|E'_{\vec{\alpha}}\rangle$ is zero, q = 0, and $\frac{\vec{\alpha}\cdot\vec{\alpha}}{\alpha^2} = 1 = \frac{q-p}{2} = -p/2$, which is impossible as p ≥ 0.

Therefore no root has two independent eigenvectors.

Now consider $\vec{\beta} = 2\vec{\alpha}$. $\vec{\alpha} + \vec{\beta}$ can't be a root, as it would be $3\vec{\alpha}$, so $E_{\vec{\alpha}}$ cannot raise $\left|E_{\vec{\beta}}\right\rangle$, the relevant p = 0, but n = 4, so $E_{-\vec{\alpha}} \left|E_{\vec{\beta}}\right\rangle \neq 0$, and it has weight $\vec{\alpha}$, , so must be proportional to $|E_{\vec{\alpha}}\rangle$, $E_{-\vec{\alpha}} \left|E_{\vec{\beta}}\right\rangle = k\,|E_{\vec{\alpha}}\rangle$.

Then $k = \langle E_{\vec{\alpha}}| E_{-\vec{\alpha}} \left|E_{\vec{\beta}}\right\rangle = \left\langle E_{\vec{\beta}}\right| E_{\vec{\alpha}} |E_{\vec{\alpha}}\rangle^* = \left\langle E_{\vec{\beta}}\right|[E_{\vec{\alpha}}, E_{\vec{\alpha}}]\rangle^* = 0$, so again we have a contradiction, which rules θ ≠ 0 for different roots, and θ ≠ 180° for non-conjugate roots.

Notice that $E_{\pm\vec{\alpha}}$ and $\sum \alpha_i H_i$ form an SU(2) subalgebra, and the string of n values generated about a given $\vec{\mu}$ is a $j = n/2$ representation of this SU(2), and the above investigation into the possible values is a rehash of what we did for SU(2).

Before we go on, we need another example, and a very important one in elementary particle physics, SU(3).

But first a recap:

We are in the process of classifying all finite dimensional compact semisimple Lie algebras and their finite-dimensional irreducible representations.

$\mathcal{H}$ is the Cartan subalgebra of the Lie algebra L we are considering, with generators H_i for $\mathcal{H}$ and $E_{\vec{\alpha}}$ for the rest of $\mathcal{L}$.

$$H_i |\vec{\mu}, D\rangle = \mu_i |\vec{\mu}, D\rangle , \quad \begin{cases} \vec{\mu} \text{ is the } \textbf{weight} \text{ of the state } |\vec{\mu}, D\rangle \\ \text{in representation } D \end{cases}$$

$$H_i |E_{\vec{\alpha}}\rangle = \alpha_i |E_{\vec{\alpha}}\rangle , \quad \begin{cases} \vec{\alpha} \text{ is the } \textbf{root}, \text{ the weight of a state in the} \\ \text{adjoint representation} \end{cases}$$

Considering $E_{\vec{\alpha}}$ as a raising operator acting on a state $|\vec{\mu}, D\rangle$, p = number of times $E_{\vec{\alpha}}$ can act on $|\vec{\mu}, D\rangle$ before vanishing, q = number of times $E_{-\vec{\alpha}}$ can act on $|\vec{\mu}, D\rangle$ before vanishing.

Then

$$q - p = 2\,\frac{\vec{\mu}\cdot\vec{\alpha}}{\alpha^2}.$$

If $\vec{\alpha}$ and $\vec{\beta}$ are roots, the angle θ between them has $4\cos^2\theta$ an integer, and $\theta \neq 0$ for different roots.

Part 6 SU(3)

SU(3) first hit the Physics world in 1961 through papers by Gell-Mann and Ne'eman which applied it to what we now call the flavor of hadrons, at a time when particles involving charm, top, or bottom were unknown.

In modern language, these hadrons are made up of quarks of three different "flavors", called up, down, and strange.

The SU(3) acts on these quarks exactly as isospin acts on p's and n's, which is in fact the same as isospin's action on the u and d quarks.

That is, the symmetry

$$\begin{pmatrix} u \\ d \\ s \end{pmatrix} \rightarrow U \begin{pmatrix} u \\ d \\ s \end{pmatrix}$$

where U is a unitary 3 x 3 matrix of determinant one, is an approximate symmetry of the Lagrangian or Hamiltonian for strong interactions.

The subset of unitary U matrices which leave the third component unchanged is just the isospin group.

This **flavor SU(3)** symmetry is, however, clearly not an exact symmetry, even for the strong interactions, as the particles involving strange quarks are considerably heavier than the others.

The interest in SU(3) as a flavor symmetry has largely faded away, after the discovery of 3 more quark flavors which act so differently from the first three that symmetry seems an inappropriate approximation.

But in the meantime the theory of QCD, which requires the quarks to have three colors as well as flavors, has become the standard way of understanding the strong interactions, and it is based on an exact color SU(3) symmetry, in fact a gauge symmetry.

So we still have a good reason to investigate SU(3) in detail, both for its own sake and as an example of more elaborate groups.

The generators of the group are clearly 3 x 3 hermitean traceless matrices.

This is an eight dimensional space, as the 9 real values have the one constraint of tracelessness.

The standard basis is Gell-Mann's original one,

$$\lambda_i = \begin{pmatrix} \sigma_i & & 0 \\ & & 0 \\ 0 \;\; 0 & & 0 \end{pmatrix} \text{ for } i = 1, 2, 3; \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}; \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}; \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

We take the generators to be $T_a = \lambda_a/2$.

Note T_1, T_2, and T_3 generate an SU(2) subalgebra which just rotates the isotopic spin doublet $\begin{pmatrix} u \\ d \end{pmatrix}$, leaving the isotopic singlet s invariant, so this is just the isospin group.

The structure constants for the Gell-Mann basis of SU(3) are always called $f_{jk\ell}$, with $[\lambda_j , \lambda_k] = 2if_{jk\ell}\lambda_\ell$.

As the first three T_i form an SU(2) subgroup, $f_{jk\ell} = \epsilon_{jk\ell}$ when all indices are ≤ 3, and $f_{jk\ell} = 0$ for $j, k \leq 3$ when $\ell > 3$.

Also T_4, T_5 and $1/2\ (\sqrt{3}\,T_8 + T_3)$form a canonical SU(2) basis, and so do T_6, T_7 and $1/2\ (\sqrt{3}\,T_8 - T3)$, so $f_{458} = 1/2\ \sqrt{3}$, $f_{453} = 1/2$, $f_{678} = 1/2\ \sqrt{3}$, $f_{673} = -1/2$

Finally $[\lambda_1, \lambda_4] = i\lambda_7$, so $f_{147} = 1/2$, and similarly $f_{246} = f_{257} = 1/2$, $f_{156} = -1/2$.

The others are given by total antisymmetry, or vanish.

The commutators are a general feature of a Lie algebra, but for SU(n) only there is also an anticommutator relation,

$$\{\lambda_i, \lambda_j\} = \frac{4}{3}\delta_{ij}\mathbb{I} + 2d_{ijk}\lambda_k$$

because the λ's and 𝕀 are a complete set of hermitean matrices.

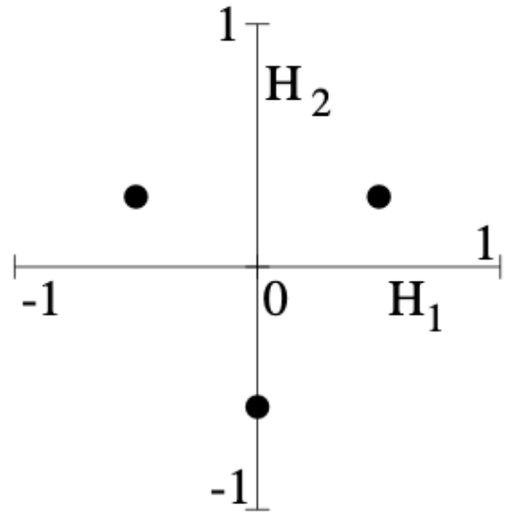
The matrices are themselves a representation (the defining representation, not the adjoint representation).

There can only be 2 independent simultaneously diagonalized traceless hermitean 3 x 3 matrices, and we already have them in λ_3 and λ_8, so the Cartan subalgebra is $\mathcal{H} = < T_3, T_8 >$.

Let $H_1 = T_3$, $H_2 = T_8$.

The weights in this representation are just the combinations (H_{1ii}, H_{2ii}) (no sum) or 1/2 (λ_{3ii}, λ_{8ii}) (no sum) or

$$(\mu_1, \mu_2) = \underbrace{\left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)}_{u}, \quad \underbrace{\left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right)}_{d}, \quad \underbrace{\left(0, -\frac{1}{\sqrt{3}}\right)}_{s}.$$

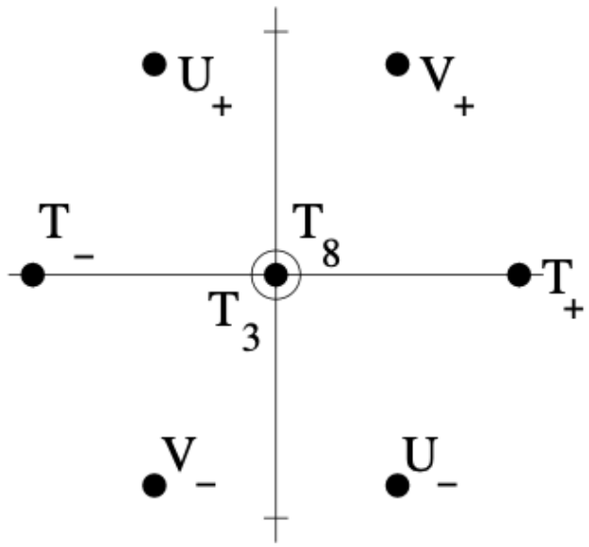


The three root vectors E_α and their conjugates $E_{-\alpha}$ span the six dimensions orthogonal to $\mathcal{H}$.

These make transitions between the weights.

The root which takes d → u, which we used to call T_+, is now $E_{(1,0)} = 1/\sqrt{2}\ (T_1 + iT_2)$.

We also have $E_{(1/2,\sqrt{3}/2)} = 1/\sqrt{2}\ (T_4 + iT_5)$, sometimes called V_+, which takes s → u, and $E_{(-1/2,\sqrt{3}/2)} = 1/\sqrt{2}\ (T_6 + iT_7)$, also called U_+, which takes s → d.



So the root space of the generators looks like the figure, with the roots forming a regular hexagon, with angles between them of 60◦ x n.

From the diagram we see T_- generates doublets on V_+ and U_- as required by $\dfrac{\vec{\alpha}_{T_+} \cdot \vec{\alpha}_{V_+}}{\vec{\alpha}_{T_+}^2} = \dfrac{n}{2} = \dfrac{1}{2}$.

Similarly U_- generates doublets (“U spin doublets”) starting with V_+ or T_-.

Part 7 Simple Roots

Now let us return to the general theory, using SU(3) and SU(2) as examples.

Consider a compact semisimple finite dimensional Lie algebra $\mathcal{L}$ (as usual) with a definite Cartan subalgebra and a fixed basis for it, $H_1, H_2, \ldots, H_m$.

Any representation can be written with a basis in which the H_i are diagonal, so the m eigenvalues belonging to a given basis vector in the representation is an m component vector $\mu_1, \ldots \mu_m$ which we have already defined as the weight.

We give an ordering to the weights by an alphabetic method, that is, $\mu > \mu'$ if the first nonzeroth element of the vector $\vec{\mu} - \vec{\mu}'$ is greater than zero; $\vec{\mu} = \vec{\mu}'$ only if the vectors are identical.

This definition is clearly arbitrary, because it depends on our choice of basis for H, even on the order in which we list the basis elements.

It is nonetheless useful.

A weight is positive if it is > 0.

The adjoint representation is also subject to this definition, so we have imposed an ordering on the roots.

Now all the roots are either positive or negative.

The positive roots can be considered raising operators, and the negative ones their conjugate lowering operators.

Note because we have chosen(we are making a standard choice) $H_1 = T_3$ and $H_2 = T_8$, instead of the other way around, the raising operators are T_+, V_+, and U_-, not U_+.

If we had interchanged H_1 and H_2, T_+, V_+, and U_+ would have been the raising operators as some authors choose.

Which order one chooses is not important as long as you stick to it.

When we want to know how the algebra acts on a representation vector space, we don't need to investigate how [A, B] operates if we already know how A and B work separately.

So if we already know the action of two roots $E_{\vec{\alpha}}$ and $E_{\vec{\beta}}$, , we don't need to worry about $\left[E_{\vec{\alpha}}, E_{\vec{\beta}}\right] \propto E_{\vec{\alpha}+\vec{\beta}}$, even if it is a root.

We therefore define:

A **simple root** is a positive root which cannot be written as a sum of two positive roots.

Lemma: If $\vec{\alpha}$ and $\vec{\beta}$ are unequal simple roots, $\vec{\alpha}-\vec{\beta}$ is not a root, and $\vec{\alpha}\cdot\vec{\beta} \leq 0$.

Proof: If $\vec{\alpha} - \vec{\beta}$ is a root, so is $\vec{\beta} - \vec{\alpha}$, and one of them is positive, say $\vec{\alpha} - \vec{\beta}$. Then $\vec{\alpha}$ is the sum of positive roots $\vec{\alpha} - \vec{\beta}$ and $\vec{\beta}$, and is not simple, which contradicts the hypothesis.

If $\vec{\beta} - \vec{\alpha}$ is not a root, the q for the $\vec{\alpha}$ multiplet starting from $\vec{\beta}$ is zero, and $\dfrac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = -\dfrac{p}{2} \leq 0$.

The set of nonnegative integers p for each ordered pair $\vec{\alpha}, \vec{\beta}$ determine not only the angles $\cos^2 \theta_{\alpha\beta} = p_{\alpha\beta}\, p_{\beta\alpha}/4$ but also the relative magnitudes $\beta^2/\alpha^2 = p_{\alpha\beta}/p_{\beta\alpha}$ for all the simple root vectors.

And each p can only be 0, 1, 2, or 3. Also either both $p_{\alpha\beta}$ and $p_{\beta\alpha}$ are zero or at least one of the two is 1.

Now if there aren't too many simple roots we should be able to determine everything.

We now show that the number of simple roots is equal to the rank m of the group.

First, we show that they are linearly independent.

Suppose not, so there exists a relation $\sum_i x_i \alpha^i = 0$ with some real $x_i \neq 0$.

Divide the expression into two parts with the coefficients positive or negative.

$$\sum_{i \ni x_i > 0} x_i \vec{\alpha}^{\,i} + \sum_{i \ni x_i < 0} x_i \vec{\alpha}^{\,i} = 0,$$

or, calling the first piece $y := \sum_{i \ni x_i > 0} x_i \vec{\alpha}^{\,i}$ and minus the second $z := \sum_{i \ni x_i < 0} |x_i| \vec{\alpha}^{\,i}$, , we have y = z.

Now y and z are both positive linear combinations of positive roots, so each is nonzero, but

$$y^2 = y \cdot z = \sum_{\substack{i \ni x_i > 0 \\ j \ni x_j < 0}} |x_i||x_j| \vec{\alpha}^{\,i} \cdot \vec{\alpha}^{\,j} \leq 0,$$

because all the included $\vec{\alpha}^{\,i} \cdot \vec{\alpha}^{\,j}$ are ≤ 0 and this contradicts $y^2 > 0$ as $y \neq 0$. QED.

Thus the number of positive roots cannot be greater than the dimensionality of the space they live in, which is the rank of the algebra.

Second, we show any positive root ϕ is a sum of one or more simple roots $\vec{\alpha}_i$ with nonnegative integer coefficients $K_{\vec{\alpha}_i}$. .

Suppose not.

Then there is a smallest positive root which cannot be so written, and it is not simple, so it can be written as a sum of two positive roots, each smaller than it, and hence expressible as such.

Third, we show that the simple roots span the whole m dimensional space, and hence there are exactly m of them.

For if not, let v be a vector in the m dimensional space perpendicular to all the simple roots.

Every root is a linear combination of these (as either it is positive or it is minus a positive root), so $\vec{v} \cdot \vec{\alpha} = 0$ for every $E_{\vec{\alpha}}$.

Then $\left[\vec{v} \cdot \vec{H}, E_{\vec{\alpha}}\right] = \vec{v} \cdot \vec{\alpha} = 0$ and $\vec{v} \cdot \vec{H}$ commutes with every generator of the algebra and generates a one-dimensional abelian invariant subalgebra.

But this is impossible because the whole algebra is semisimple.

Thus a compact semisimple finite dimensional Lie algebra has its structure determined by the dimensionality m of its Cartan subalgebra and the integers $p_{\alpha\beta}$ for each pair of the m simple roots, $\vec{\alpha}$ and $\vec{\beta}$.

These determine the $\vec{\alpha}\cdot\vec{\beta}$, and thus the root vectors which exist.

The procedure for determining which positive linear combinations of simple roots is described in many texts.

Instead of repeating it, I will work out an example.

Let us classify all the simisimple Lie algebras of rank 2.

There are only two simple roots, say $\vec{\alpha}$ and $\vec{\beta}$. , and we may as well choose them such that $|\alpha| \le |\beta|$.

We know that $-2\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = p,\ -2\frac{\vec{\alpha}\cdot\vec{\beta}}{\beta^2} = p',$ where p and p′ are nonnegative integers.

If $\vec{\alpha}\cdot\vec{\beta} = 0,$ both $p = p' = 0$, otherwise $\cos^2\theta = p\,p'/4$, so $p\,p' = 1, 2,$ or 3.

As we took $|\alpha| \le |\beta|$, $p \ge p'$, so the four possibilities are

$$\begin{aligned} &(a) \quad p = p' = 0 \\ &(b) \quad p = p' = 1 \\ &(c) \quad p = 2,\ p' = 1 \\ &(d) \quad p = 3,\ p' = 1 \end{aligned}$$

The first three possibilities should be done by you.

I will do case (d) now.

We have

$$-2\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = 3,\ -2\frac{\vec{\alpha}\cdot\vec{\beta}}{\beta^2} = 1,\ \text{so } \beta^2 = 3\alpha^2 \text{ and } \cos\theta = -\sqrt{3}/2, \quad \theta = 150°.$$

To choose our root vectors, we note that the requirements

1) the killing form is a multiple of the identity
2) the subspace $\mathcal{H}$ is abelian, and left invariant by overall scale changes and rotations of the H_i's.

This means that we can change the scale and rotate the whole root diagram.

The only proviso is that what had been positive may cease to remain so.

Choose $\vec{\alpha} = (0, 1)$ so $\vec{\beta} = \left(\sqrt{3}/2, -3/2\right)$.

All roots must be of the form $\vec{\gamma} = m\vec{\alpha} + n\vec{\beta}$ with m ≥ 0 and n ≥ 0, or the negative of such a root.

As p = 3, we are assured that $E_{\vec{\alpha}}$ can be applied three times to $E_{\vec{\beta}}$ before vanishing, giving

$$E_{\vec{\alpha}+\vec{\beta}},\ E_{2\vec{\alpha}+\vec{\beta}},\ E_{3\vec{\alpha}+\vec{\beta}},$$

but guaranteeing that $E_{4\vec{\alpha}+\vec{\beta}}$ does not exist.

As p′ = 1, can be applied only once to $E_{\vec{\alpha}}$, so $E_{\vec{\alpha}+2\vec{\beta}}$ does not exist.

We still need to ask if $E_{\vec{\beta}}$ can be applied to $E_{2\vec{\alpha}+\vec{\beta}}$ or $E_{3\vec{\alpha}+\vec{\beta}}$ and give a new state.

In both cases $E_{-\vec{\beta}}$ could not, because this would give a root parallel to $E_{\vec{\alpha}},$, so we know q = 0.

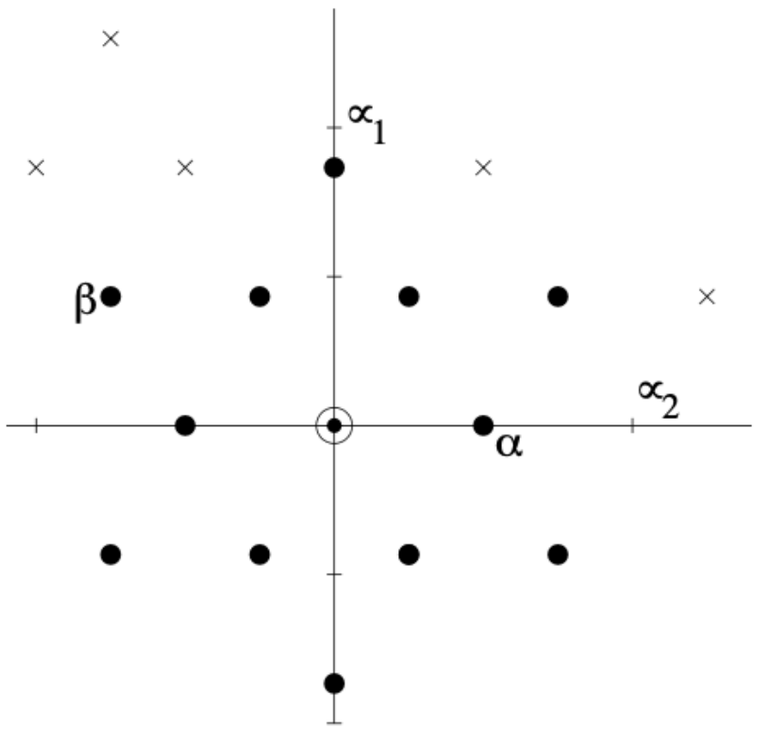
Then as for any $\vec{\mu},$ $\frac{\vec{\mu}\cdot\vec{\beta}}{\beta^2} = \frac{q-p}{2}$, in this case the appropriate p , is $-\frac{2}{3}\vec{\beta}\cdot(r\vec{\alpha}+\vec{\beta}) = r-2,$ so

$E_{2\vec{\alpha}+2\vec{\beta}}$ does not exist, but just one application of $E_{\vec{\beta}}$ is allowed, giving $E_{3\vec{\alpha}+\vec{\beta}}$, but not $E_{3\vec{\alpha}+2\vec{\beta}},$.

Can $E_{\vec{\alpha}},$ act on our new $E_{3\vec{\alpha}+3\vec{\beta}}.$?

$E_{-\vec{\alpha}}$ cannot, as $2\vec{\alpha}+2\vec{\beta}$ is not a root, so q = 0,
and $\vec{\alpha}\cdot(3\vec{\alpha}+2\vec{\beta}) = 0$, so p = 0 and it cannot.

So we have found all the positive roots, and adding their negatives we have this pretty diagram with two Cartan basis vectors and 12 roots, giving the 14 dimensional algebra called G_2.



Roots diagram for G_2. The two generators of the Cartan subalgebra are indicated at the origin. The 12 roots are shown by dots, and the weights excluded as being roots in the argument are shown as crosses.

Part 8 Dynkin Diagrams

We now describe how to use the constraints

$$\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = \frac{q-p}{2}$$

where p is the number of times $\vec{E}\alpha$ can hit $|\beta>$ without annihilating, and q is the number of times $E_{-\vec{\alpha}}$ can.

This is encoded in the **Cartan matrix** for the simple roots,

$$A_{ji} = \frac{2\vec{\alpha}_j\cdot\vec{\alpha}_i}{\alpha_i^2} = q-p.$$

We will draw a Dynkin diagram for each semisimple compact finite-dimensional Lie group in which each simple root will be represented by a dot or small circle, and they will be connected by bonds according to the p values.

As we saw, these p values are quite limited, but we will also find constraints on the way even these limited number of bonds can be put together.

We shall see that many conceivable diagrams are forbidden, and therefore that the set of all such groups is quite limited.

For each simple root draw a small open circle.

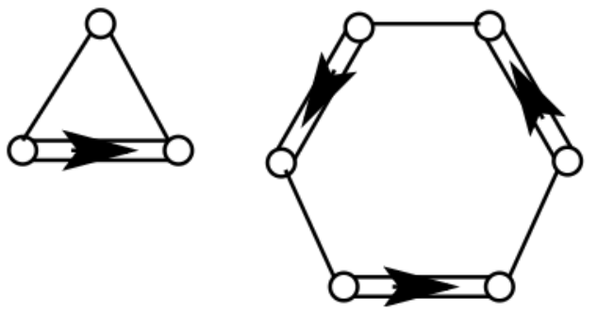
If α and β are two simple roots, and $-2\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2}$ = 2 or 3, draw 2 or 3 lines, respectively, between their corresponding circles, with an arrow from α to β.

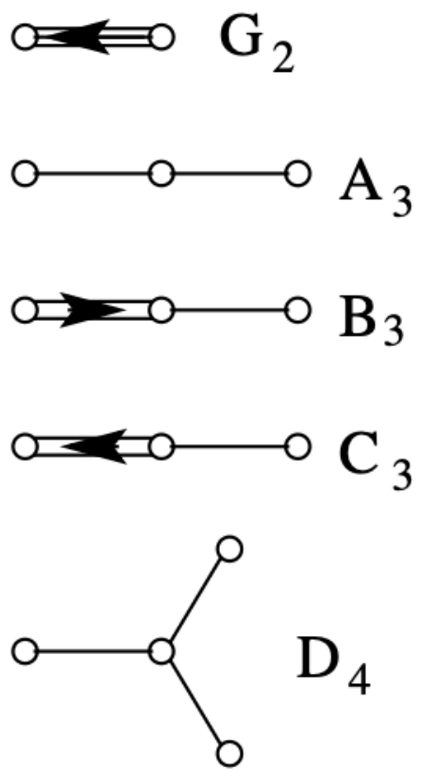
That is, from the shorter root to the longer.

If $-2\frac{\vec{\alpha}\cdot\vec{\beta}}{\alpha^2} = -2\frac{\vec{\alpha}\cdot\vec{\beta}}{\beta^2} = 1,$ draw a single line without an arrow between the two roots.

If $\vec{\alpha}\cdot\vec{\beta} = 0,$ connect the two circles corresponding to α and β.

With these rules we can draw diagrams such as these, and investigate if there is a group corresponding to that diagram. We will see, however, that there are restrictions on what is possible. It will turn out that the diagrams on the right do correspond to Lie groups, while those below do not.





8.1 Classification of Simple Lie Algebras

We have seen that given any algebra, there exist m simple roots, where m is the rank of the algebra, with relations that can be represented by a Dynkin diagram.

Let us first show any algebra with a disconnected diagrams cannot be simple.

Suppose that the simple roots can be divided into two sets, A and B, with the vertices of A disconnected from B.

Then, for $\alpha \in A$ and $\beta \in B$, $E_{-\alpha}| E_\beta > = 0$, which is always true for simple roots, so $q = 0$.

As there is no connection of α and β, $\vec{\alpha} \cdot \vec{\beta} = 0$, and $p = 0$, so $E_\alpha| E_\beta > = 0$.

Each simple root in A commutes with each in B and with its conjugate.

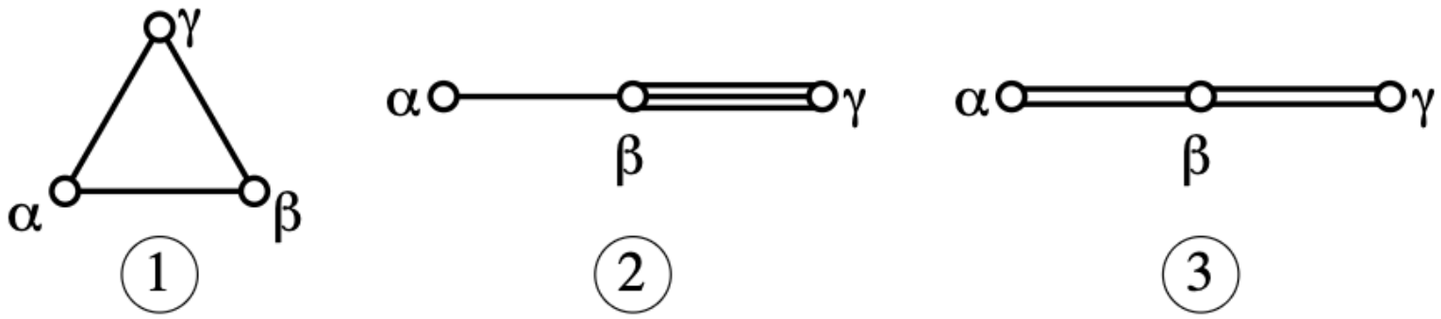
The other roots in A are commutators of simple ones, and similarly for B, so all roots in A commute with all roots in B.

The group is therefore the direct product of that generated by A and that generated by B.

Each of these is a nontrivial invariant subalgebra, so the whole group is not simple.

Of course the direct product of simple groups is semisimple, but not simple.

Next, we rule out diagrams with three simple roots connected as shown, whether or not there are additional connections to other roots.



The angles between these roots are as givenin the table, with the angles even bigger if any additional lines are added.

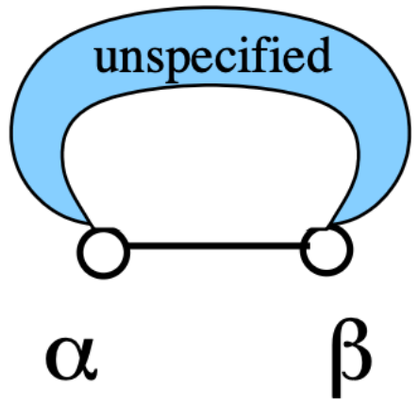
	(1)	(2)	(3)
$\theta_{\alpha\beta}$	120°	120°	135°
$\theta_{\beta\gamma}$	120°	150°	135°
$\theta_{\gamma\alpha}$	120°	90°	90°
$\sum \theta$	360°	360°	360°

Each is a vector with its tail at the origin.

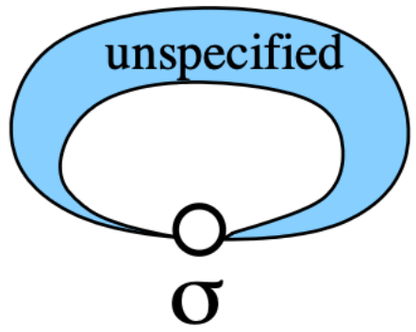
As the sum of three angles between three such vectors must be ≤ 360◦, and equal to 360◦ only if the three vectors lie in a plane, we see the three roots must lie in a plane, and therefore cannot be linearly independent, which is impossible.

Thus the only diagram with a “triple bond” is the diagram for G_2.

A set of vectors which corresponds to the Dynkin diagram rules, having the magnitudes and angles as indicated by the diagram, is called a Π system, without worrying about whether each Π system corresponds to an algebra.



Now consider a Π system containing two simple roots α and β connected to each other with a single line, with unspecified connections to the rest of the system.



Then the Π system which results from contracting the line, collapsing α and β to a single root $\vec{\sigma} = \vec{\alpha} + \vec{\beta}$. is also a Π system.

For $\sigma^2 = \alpha^2 = \beta^2$ because they were connected by a single line, so are of equal magnitude with an angle of 120° between them.

For each γ in the unknown, γ cannot be connected to both α and β, because diagram (1) above is excluded, and is not part of any Π system.

If γ was disconnected from β, $\vec{\gamma} \cdot \vec{\beta} = 0$, so,

$$\vec{\gamma} \cdot \vec{\sigma} = \vec{\gamma} \cdot (\vec{\alpha} + \vec{\beta}) = \vec{\gamma} \cdot \vec{\alpha},$$

so γ is connected to σ with the same connection that it had been to α.

Corrollaries:

- No Π system contains more than 1 double bond (or else collapsing whatever connected them would produce diagram (3) above).
- No Π system contains contains a closed loop, for if so it could be collapsed to diagram (1).

No Π system contains a root linked to four others with single links.

For then

$$\alpha^2 = \beta^2 = \gamma^2 = \delta^2 = \mu^2,$$
$$\text{and } \vec{\mu} \cdot \vec{\alpha} = -\mu^2/2 = \vec{\mu} \cdot \vec{\beta} = \vec{\mu} \cdot \vec{\gamma} = \vec{\mu} \cdot \vec{\delta},$$
$$\vec{\alpha} \cdot \vec{\beta} = \vec{\alpha} \cdot \vec{\gamma} = \vec{\alpha} \cdot \vec{\delta} = \vec{\beta} \cdot \vec{\gamma} = \vec{\beta} \cdot \vec{\delta} = \vec{\gamma} \cdot \vec{\delta} = 0,$$
$$\text{so } (\vec{\alpha} + \vec{\beta} + \vec{\gamma} + \vec{\delta} + 2\vec{\mu})^2 = \mu^2(4 + 4 - 8) = 0$$

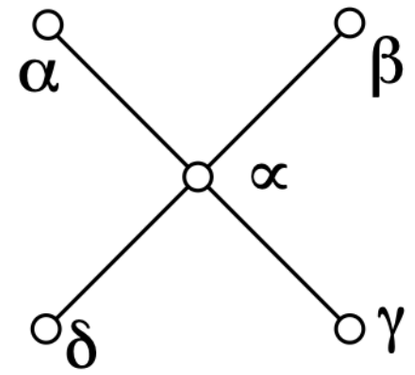
and the five roots are not all independent.

No Π system contains a root connected to three others, including one double bond, regardless of direction.

For

$$\alpha^2 = \beta^2 = \gamma^2,\ \vec{\alpha}\cdot\vec{\beta} = \vec{\beta}\cdot\vec{\gamma} = -\tfrac{1}{2}\beta^2,\ \ \vec{\alpha}\cdot\vec{\gamma} = \vec{\alpha}\cdot\vec{\delta} = \vec{\gamma}\cdot\vec{\delta} = 0.$$

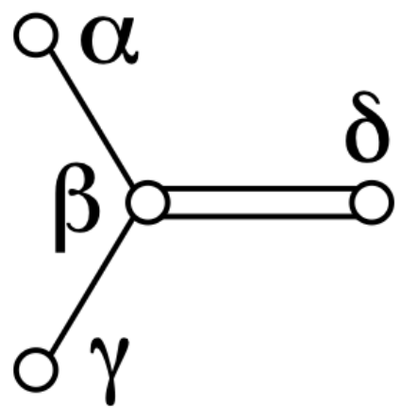
The angle between $\vec{\beta}$ and $\vec{\delta}$ is 135◦.



Let $\hat{\delta}$ be in the direction of $\vec{\delta}$ but with the length of $\vec{\beta}$ (changing it by $\sqrt{2}$ or its inverse).

Then

$$\vec{\beta}\cdot\hat{\delta} = \beta^2\cos 135^\circ = -\beta^2/\sqrt{2}$$



and

$$\begin{aligned}(\vec{\alpha}+\vec{\gamma}+2\vec{\beta}+\sqrt{2}\hat{\delta})^2 &= \alpha^2+\gamma^2+4\beta^2+2(\hat{\delta})^2+4\vec{\alpha}\cdot\vec{\beta}+4\vec{\gamma}\cdot\vec{\beta}+4\sqrt{2}\vec{\beta}\cdot\hat{\delta}\\ &= \beta^2(1+1+4+2-2-2-4) = 0,\end{aligned}$$

so again the roots are not linearly independent, which is impossible.

Corrolaries:

- No root is linked directly to four others. We have already shown that for four single bonds, but if one is double, contracting one of the others reduces to the case just excluded.
- No Π system contains both a branch point and a double bond, because contracting what is between them gives what we just excluded, and no Π system contains more than one branch point, because contracting what connects them leaves a root linked to four others.

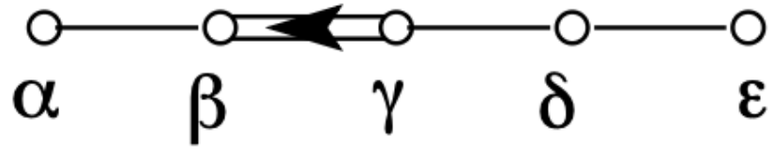
So thus far we see that all Π systems are either

(a) G_2
(b) a linear chain with at most one double bond
(c) a chain with one branch point and only single bonds.

Now we will limit choices (b) and (c).

First consider (b).

No diagram contains



$$\alpha^2 = \beta^2 = 2\gamma^2 = 2\delta^2 = 2\epsilon^2,\ \vec{\alpha}\cdot\vec{\beta} = -\tfrac{1}{2}\beta^2 = -\gamma^2,\ \vec{\beta}\cdot\vec{\gamma} = -\gamma^2,$$

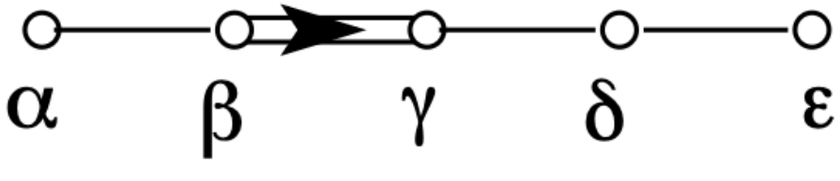
$$\vec{\gamma}\cdot\vec{\delta} = \vec{\delta}\cdot\vec{\epsilon} = -\tfrac{1}{2}\gamma^2$$

the other dot products vanish, and

$$(\vec{\alpha} + 2\vec{\beta} + 3\vec{\gamma} + 2\vec{\delta} + \vec{\epsilon})^2 = \gamma^2(2 + 8 + 9 + 4 + 1 - 4 - 12 - 6 - 2) = 0,$$

and again the five roots are not linearly independent.

If we reverse the arrow



$$2\alpha^2 = 2\beta^2 = \gamma^2 = \delta^2 = \epsilon^2,\ \vec{\alpha}\cdot\vec{\beta} = -\tfrac{1}{2}\alpha^2,\ \vec{\beta}\cdot\vec{\gamma} = -\beta^2,\ \vec{\gamma}\cdot\vec{\delta} = \vec{\delta}\cdot\vec{\epsilon} = \tfrac{1}{2}\gamma^2 = -\alpha^2$$

$$(2\vec{\alpha} + 4\vec{\beta} + 3\vec{\gamma} + 2\vec{\delta} + \vec{\epsilon})^2 = \alpha^2(4 + 16 + 18 + 8 + 2 - 8 - 24 - 12 - 4) = 0,$$

so once again the five roots are not linearly independent, and this is forbidden.

Thus the only double bond appears at the end of the chain or else we have only the group F4:

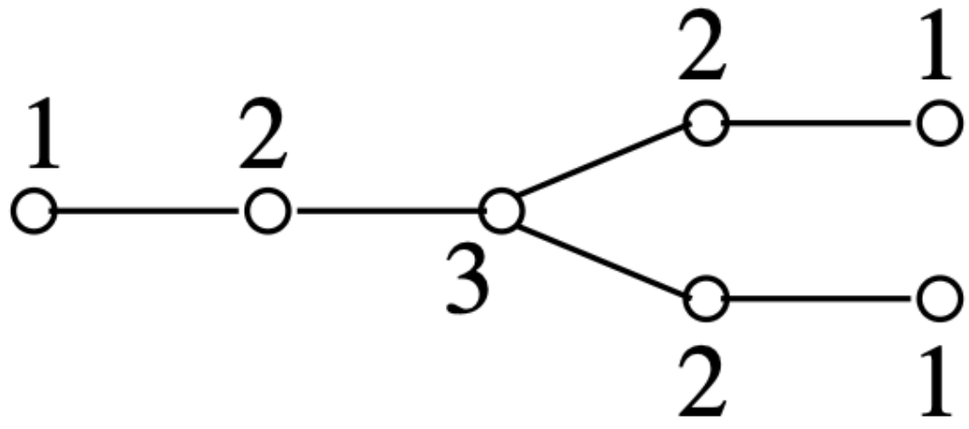
Now consider branches.

Here all the simple roots have the same magnitudes, say 1, and if connected have a dot product of -1/2.

If we add the roots α_i with weights w_i and square,

$$(\sum w_i\alpha_i)^2 = \sum w_i^2 - \sum_{\text{nn}} w_i w_j,$$

where $\sum_{\text{nn}}$ means sum over nearest neighbors (each pair once).

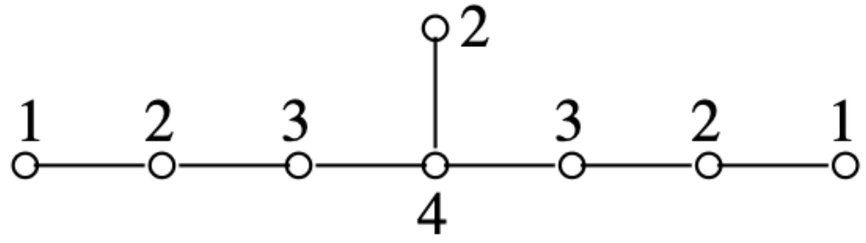


First consider seven roots connected as shown, with weights w_i as indicated.

$$\left(\sum w_i\alpha_i\right)^2 = 1 + 4 + 9 + 4 + 1 + 4 + 1 - (2 + 6 + 6 + 6 + 2 + 2) = 0.$$

So again this cannot be part of the Dynkin diagram of a group, and the shortest branch can have only one attached root.

Now consider whether both of the other branches can have three or more.

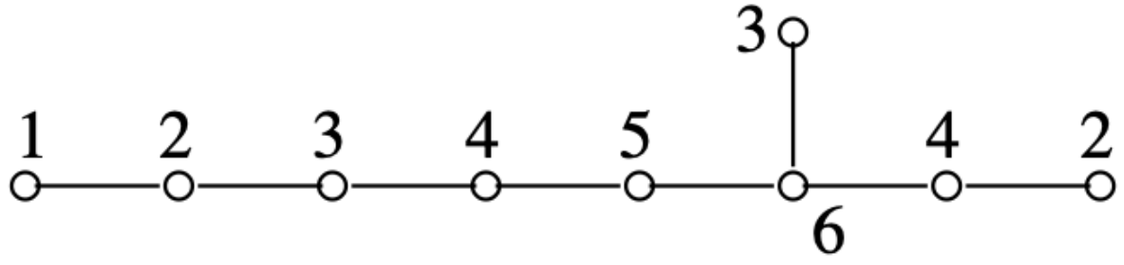


$$\left(\sum w_i\alpha_i\right)^2 = (1+4+9+16+4+9+4+1)-(2+6+12+8+12+6+2) = 48-48 = 0,$$

so the next shortest branch has at most 2 attached roots.

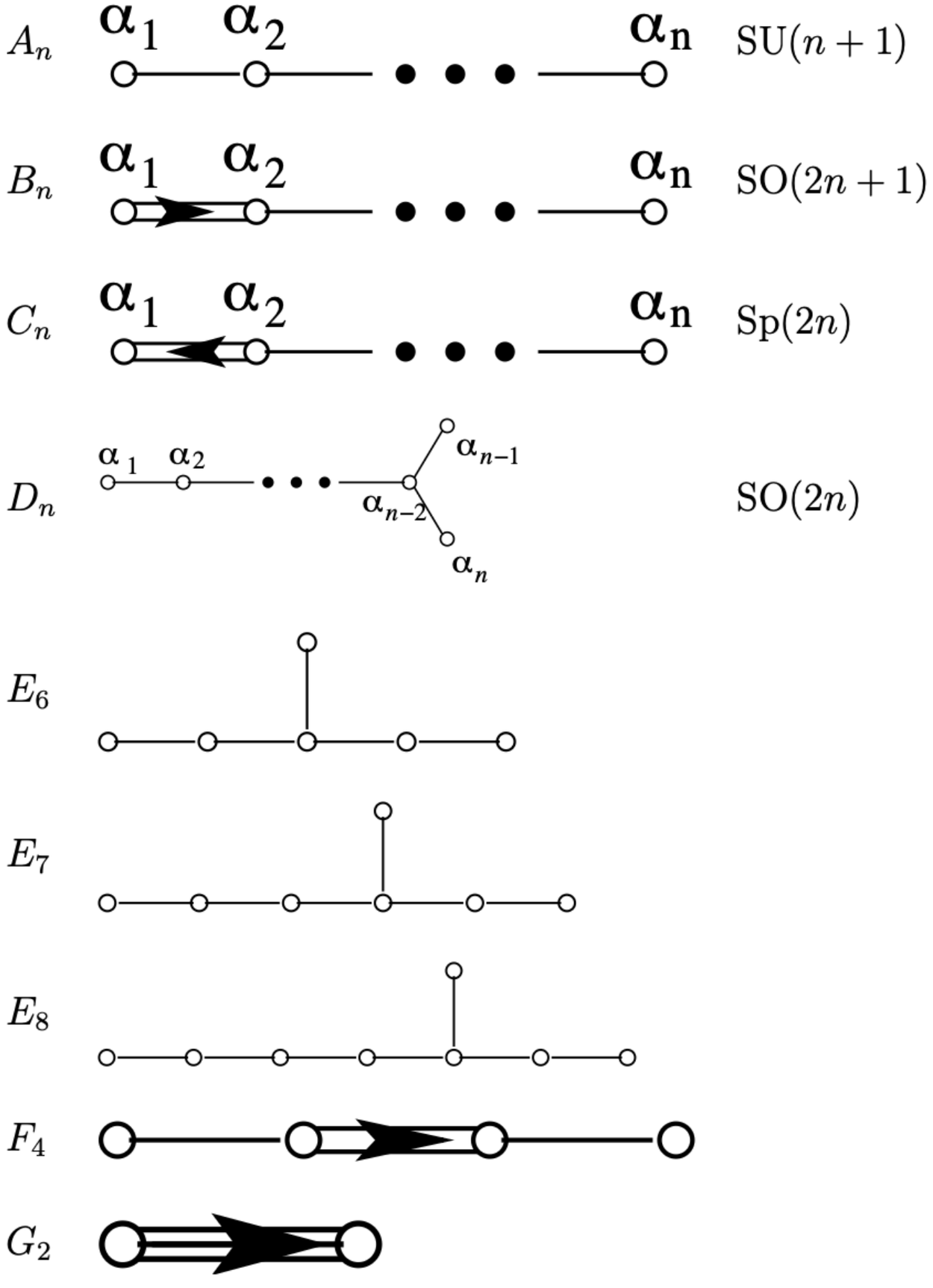
Next, if the two shortest are as long as possible, is there a limit to the third branch?

$$(1+4+9+16+25+36+9+16+4) - (2+6+12+20+30+18+24+8) = 120 - 120 = 0.$$



So the longest branch has at most 4 roots attached, if the second longest has two.

Thus we have the complete list of simple finite dimensional compact Lie groups



All of these Π systems are allowed and do correspond to a simple Lie algebra.

Part 9 Finding the Other Roots

We have seen that that master formula is very restrictive on the possible relations between the simple roots, and gives us the restrictions on how many times a raising or lowering operator can act on a given eigenvector of the Cartan subalgebra, though only the difference of these two numbers.

Recall that on a vector of weight $\vec{\mu}$, if a simple root $\vec{\alpha}^i$ can raise the weight p times and lower it q times,

$$2\frac{\vec{\mu}\cdot\vec{\alpha}_i}{\alpha_i^2} = q - p.$$

In particular, for the adjoint representation, where the weights are roots we will now call $\vec{\phi}_k$, and where every positive root is a sum $\vec{\phi}_k = \sum_j k_j \vec{\alpha}^j$, these are determined by the Cartan matrix for the simple roots

$$A_{ji} = \frac{2\vec{\alpha}_j\cdot\vec{\alpha}_i}{\alpha_i^2} = q - p.$$

Because we know that no simple root can lower another, so the relevant q = 0, when one acts on another, we can determine recursively the effective raising that can be done on each root.

As no two roots can have the same root vector, and as the root vectors can be found by raising with the simple roots, we can generate the whole algebra.

In the following explanation each box represents one positive root of the algebra, contains the q - p value for each simple root acting on the positive root.

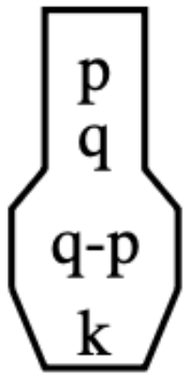
I will use a complex figure, a composite of bottles for each simple root,

Each bottle contains, in its widest part q^i - p^i Georgi puts in his box and the k^i at the bottom and the p^i above the q^i in the neck.

Of course this is redundant information, but it helps in tracing the diagram.

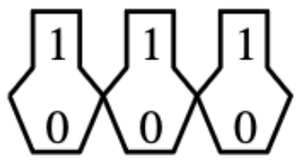
Each bottle contains four numbers:

1) the k_j for this root.
2) the $q - p$ for α^j on this root
3) the number of lowerings possible q
4) the number of raisings possible p



In the diagram, each positive root will consist of a string of touching bottles, one for each simple root.

The diagram starts with a different symbol which represents the entire Cartan subalgebra, but with only one set of m jugs, with just p and k, where all the $k^i = 0$ and all the p^i are 1 because each simple root can act only once ($E_{2\vec{\alpha}}$ is not a root).



There will be m of these truncated bottles for a rank m algebra.

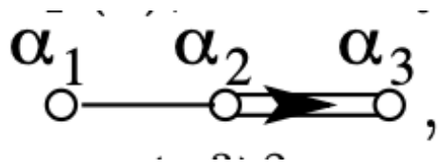
The simple roots correspond to a $k = k_i = 1$, one level up from the figure representing the Cartan subalgebra.

For each of these roots, E_{α^i}, we have $q_j = -2\delta_{ij}$ and $p_i = 0$, because $E_{-\alpha}$, $\alpha \cdot H$, and E_α are the only states parallel to α, and no other simple root can lower a simple root, because $[E_\alpha, E_{-\beta}] = 0$ for $\alpha \neq \beta$.

The q - p value for the i'th bottle in the j'th simple root is A_{ji} because only $k_j \neq 0$.

The p values are then determined.

Consider the example of C3 = Sp(6), with Dynkin diagram

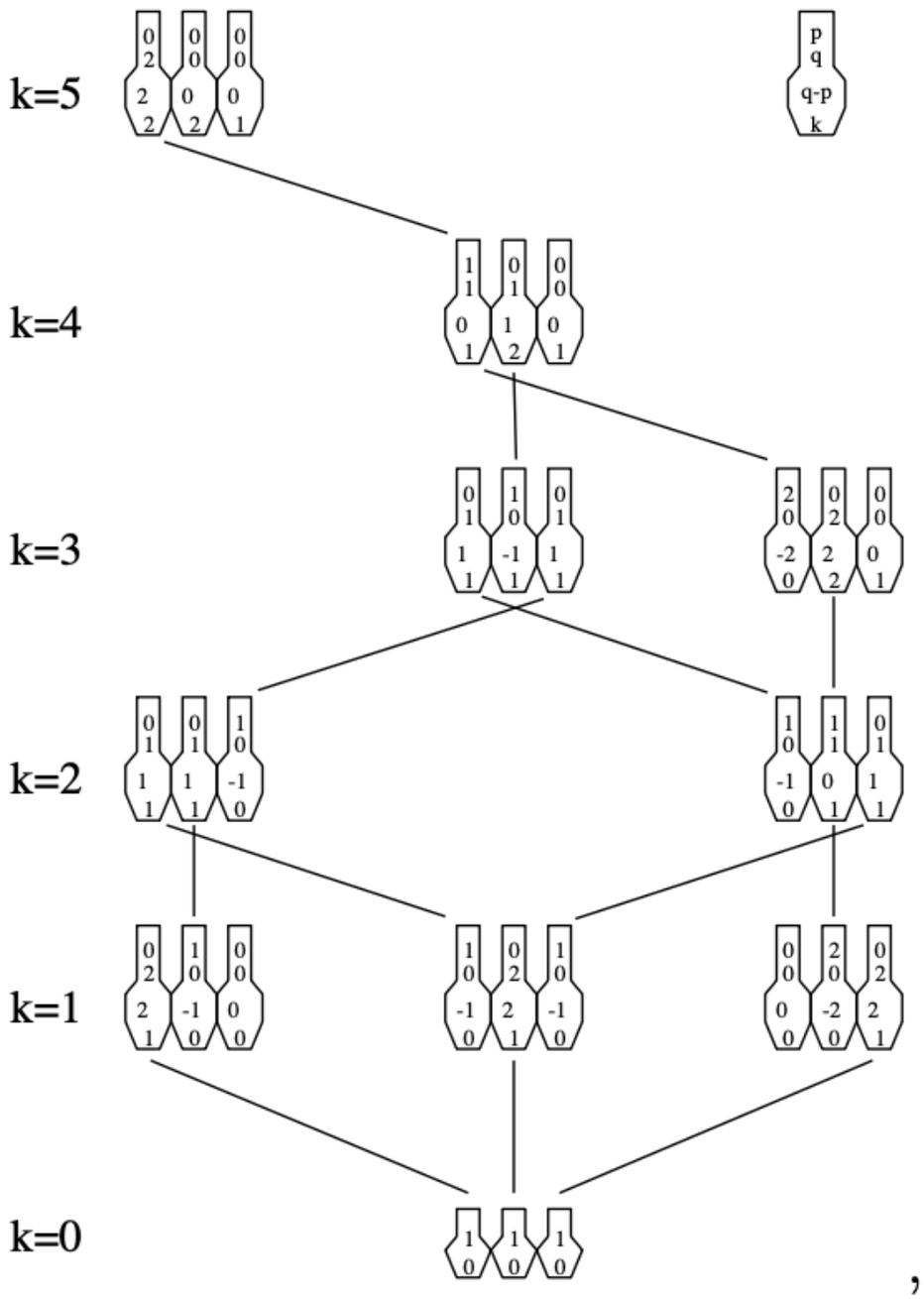


,

with

$$(\alpha^1)^2 = (\alpha^2)^2 = 1, \quad (\alpha^3)^2 = 2,$$

and $A_{ij} =$ $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{pmatrix}$.



,

The q - p values for the j'th root are just the entries in the j'th row of the Cartan matrix above.

Thus we have the k = 1 level in the diagram shown.

When $E_{\vec{\alpha}^j}$ acts on a state μ with q^i and p^i values for $\vec{\alpha}^i$ (with $p^i \neq 0$), the state $E_{\vec{\alpha}^j}|\mu\rangle$ has

$$k_i \to k_i + \delta_{ij},\ \vec{\mu} \to \vec{\mu} + \vec{\alpha}^j \text{ and } q^i - p^i \to q^i - p^i + A_{ji}.$$

So for each bottle with $p > 0$, there is a path up to one for $E_{\vec{\alpha}^j}$ acting on it with the the k^i and $q^i - p^i$ values incremented, and with its q^j one more than the q^j it came from.

A root may have more than one path leading up to it, with q^i's determined as just mentioned, and with $q^i = 0$ if there is no $E_{\vec{\alpha}^i}$ leading up to it.

With all the $q^i - p^i$ and q^i so determined, the p^i are also determined, and so we can continue generating or connecting new roots until all the highest k nodes have all $p^i = 0$.

This procedure will terminate after we have generated all the positive roots.

Of course the negative roots are just the negative of these, so we have generated the entire algebra.

Part 10 Representations of Lie Groups

10.1 Fundamental Weights

Let us label the simple roots $\vec{\alpha}^i$, i = 1, . . . , m.

They and their conjugates generate (not linearly) any element of the Lie algebra.

Thus any representation can be determined by how it behaves under these roots.

The weights have an ordering, so any finite dimensional representation has a greatest weight $\vec{\mu}_{\max}$.

The weight $\vec{\mu}$ of a basis element is $\vec{\mu}_{\max}$ if and only if $\vec{\mu}+\vec{\alpha}^i$ is not a weight for each $\vec{\alpha}^i$.

Define $q^i = 2\vec{\alpha}^i \cdot \vec{\mu}_{\max}/(\alpha^i)^2$, which can be arbitrary nonnegative integers.

As the $\vec{\alpha}^i$ are linearly independent and complete, the {q^i} and $\vec{\mu}_{\max}$ determine each other.

Define m vectors $\vec{\mu}^j$ such that

$$\frac{2\vec{\alpha}^i \cdot \vec{\mu}^j}{(\alpha^i)^2} = \delta_{ij}.$$

These are called the **fundamental weights** of the Lie algebra corresponding (as maximal weights) to a set of representations D^j called the **fundamental representations.**

Warning: $\vec{\mu}^{\,i}$ is a vector, the i'th fundamental weight, while μ_i is the i'th component of any old weight $\vec{\mu}$.

There is no connection, although each index takes on the same values $1, 2, \ldots, m$.

A tensor product of representations of a group has weights which are just the sum of the weights.

In particular, the maximum weight of the product is the sum of the maximum weights of the factors.

Thus we can find a representation of arbitrary maximal weight $\vec{\mu} = \sum q^i \vec{\mu}^{\,i}$ by reducing the tensor product

$$\bigotimes_i \left(D^i\right)^{q^i}.$$

10.1.1 SU(3) Multiplets

For SU(3), the positive roots are T_+, V_+, and U_-, of which the last two are simple.

$$V_+ = E_{1/2,\sqrt{3}/2}, \qquad U_- = E_{1/2,-\sqrt{3}/2},$$

$$\text{so} \quad \vec{\alpha}^{\,1} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad \vec{\alpha}^{\,2} = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), \qquad T_+ = E_{\vec{\alpha}^{\,1}+\vec{\alpha}^{\,2}} = E_{(1,0)}.$$

The corresponding fundamental weights with $2\dfrac{\vec{\alpha}^{\,i}\cdot\vec{\mu}^{\,j}}{(\alpha^i)^2} = \delta_{ij}$ is solved by

$$\vec{\mu}^{\,1} = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right), \quad \vec{\mu}^{\,2} = \left(\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right).$$

Let us generate the representation D^2. $\bar{d} := |\vec{\mu}^{\,2}\rangle$

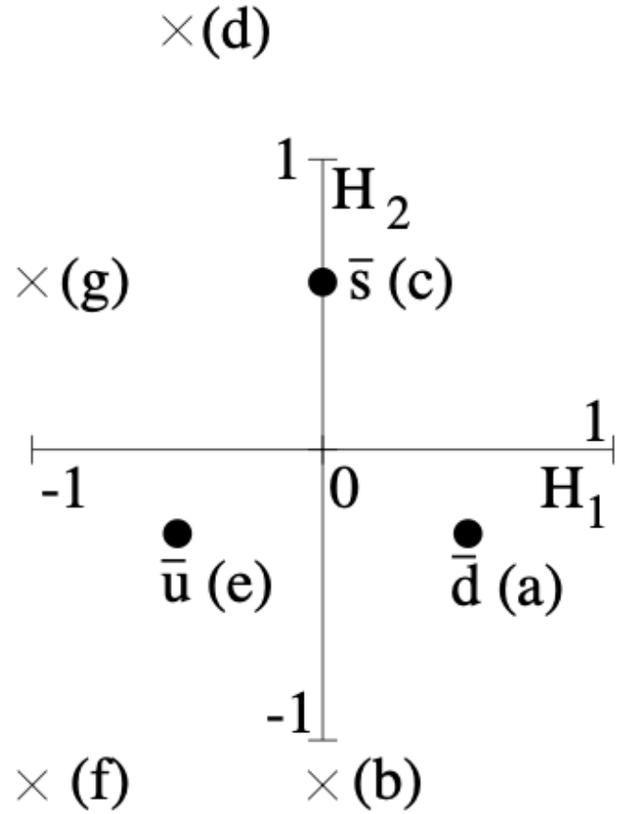
The q's are $q^1 = 0, q^2 = 1$, the highest weight state is (a).

Acting on this state $V_- = E_{-\vec{\alpha}^{\,1}}$ vanishes (b),

and $E_{-\vec{\alpha}^{\,2}} |\vec{\mu}^{\,2}\rangle$ can act once, giving a non-zero state, proportional to $\bar{s} := |0, 1/\sqrt{3}\rangle$ (c).

Now acting on this state again with $E_{-\vec{\alpha}^2}$ gives zero (d), as $2 > q^2 = 1$, but as $2\frac{\vec{\alpha}^{\,1}\cdot(0,1/\sqrt{3})}{(\alpha^1)^2} = 1$, the q for $\vec{\alpha}^{\,1}$ on $\bar{s}$ is 1, and so $E_{-\vec{\alpha}^{\,1}} |0, 1/\sqrt{3}\rangle \propto \bar{u} := |-\frac{1}{2}, -1/2\sqrt{3}\rangle$ (e)

is not zero, but a second application vanishes rather than giving a state at $(-1, 2/\sqrt{3})$ (f).



We must still check $\vec{\alpha}^{\,2}$ on $\bar{u}$, but $\vec{\alpha}^{\,2}(-\frac{1}{2}, -1/2\sqrt{3}) = 0$ and we already know p = 0, so, q = 0 and there is no state at $(-1, 2/\sqrt{3})$ (g).

So we are done, having found three basis vectors, as there is no other way to get another state.

In flavor SU(3), these are the antiquarks.

On the horizontal axis, $H_1 = T_3$ is the isospin component.

The highest weight state (rightmost) is the anti-d quark $|\mu^2\rangle$, part of an isospin doublet with the anti-u quark $T_- |\mu^2\rangle = E_{-\alpha^1} E_{-\alpha^2} |\mu^2\rangle$.

The antistrange quark $\bar{s} = E_{-\alpha^2} |\mu^2\rangle$ is an isosinglet, $T = T_3 = 0$.

The vertical axis is generally described in terms of strangeness or hypercharge.

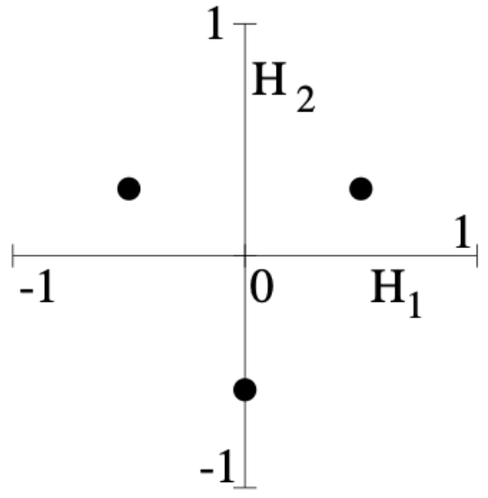
Strangeness S is defined as zero for the u and d quarks and their antiparticles, and 1 for the anti-strange quark.

Baryon number B is invariant under SU(3), and is defined as 1/3 the number of quarks minus the number of antiquarks, so is -1/3 for all the antiquark states.

The (strong2) hypercharge Y= S + B (in the absence of charm, topness and bottomness) is then 2/3 for $\bar{s}$ and -1/3 fo $\bar{u}$ and $\bar{d}$.

This representation is the conjugate of the representation D^1 corresponding to the fundamental weight μ^1, which is also called the **defining representation** for SU(3).

In flavor SU(3) this is the representation of the first three quarks, with the up quark u at the upper right, the down quark d to its left, and the strange quark s at the bottom.



The defining representation of SU(3)

The quarks have B = 1/3.

The electric charge is Q = T_3 + Y/2 times the positron charge e.

Then the quarks have the quantum numbers as shown.

The antiquarks have all the quantum numbers (except T) reversed.

quark	B	T	T_3	S	Y	Q/e
u	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{3}$	$\frac{2}{3}$
d	$\frac{1}{3}$	$\frac{1}{2}$	$-\frac{1}{2}$	0	$\frac{1}{3}$	$-\frac{1}{3}$
s	$\frac{1}{3}$	0	0	-1	$-\frac{2}{3}$	$-\frac{2}{3}$

In general, if we have a representation T_a of the generators of a Lie algebra, so that $[T_a,T_b] = ic^d_{ab}\, T_d$, with real structure constants c^d_{ab} , then $T'_a := -T^*_a$ satisfies $[T'_a, T'_b] = [T_a, T_b]^* = -ic^d_{ab}\, T^*_d = ic^d_{ab} T'_d$, so T' is also a representation, called the conjugate representation to T .

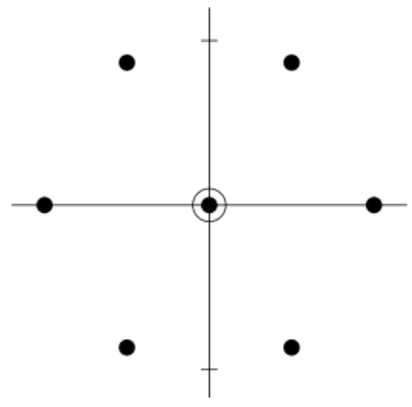
The weights are the eigenvalues of T (Hi), which are real, so $\vec{\mu}' = -\vec{\mu}.$.

Thus the conjugate representation has a weight diagram which is just a parity reversed $(i.e.\ \vec{\mu} \to -\vec{\mu})$ image of the original representation.

The highest weight of the conjugate representation is minus the lowest weight of the original.

The lowest weight of the defining representation for SU(3) is the d, and the $\bar{d} = |\mu^2\rangle$ is the highest weight of the antiquark representation.

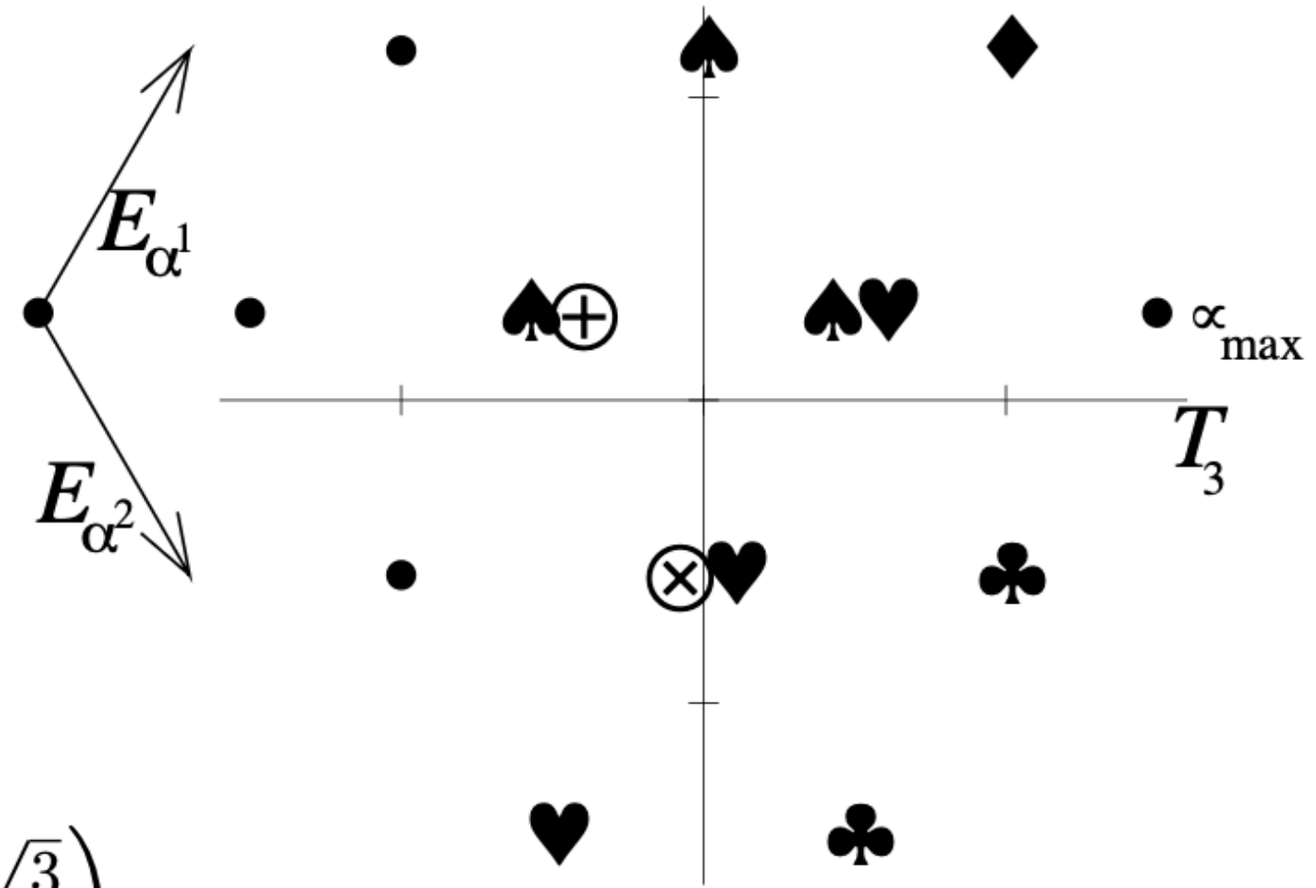
The adjoint representation of SU(3) is self-conjugate, that is, it is the same (equivalent, isomorphic) representation as its conjugate.



The adjoint representation of SU(3).

For another example, $q^1 = 2, q^2 = 1$, $\vec{\mu}_{\max} = 2\vec{\mu}^{\,1} + \vec{\mu}^{\,2} = (3/2, 1/2\sqrt{3})$.

Then to $\mu_{\max}$ we can apply $E_{-\alpha^1}$ twice to get 2 roots shown as ♣, and $E_{-\alpha^2}$ once to get ♦.



Now consider E_{α^1} on ♥ ⊗ ♦ = $\left|1, 2/\sqrt{3}\right\rangle$. p must be zero or we would get a root vector $\left|3/2, 7/2\sqrt{3}\right\rangle$ which is higher weight than $\mu_{\max}$.

But, $$2\frac{\vec{\mu}\cdot\vec{\alpha}^1}{(\alpha^1)^2} = q - p = 2\left(1, \frac{2}{\sqrt{3}}\right)\cdot\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = 3$$

so $E_{-\alpha^1}$ on ♦ generates three states, shown with ♥.

Next consider the higher ♣ = $\left|1, -1/\sqrt{3}\right\rangle$, acted on by E_{α^2} — the p is zero because

$$E_{\alpha^2}\left|1, -1/\sqrt{3}\right\rangle \propto E_{\alpha^2}E_{-\alpha^1}\left|\mu_{\max}\right\rangle = E_{-\alpha^1}\underbrace{E_{\alpha^2}\left|\mu_{\max}\right\rangle}_{=0}$$

where the first = is because different simple raising and lowering operators commute and the = 0 is because you can't raise the highest weight.

So $$q = 2\frac{(1, -1/\sqrt{3})\cdot(\frac{1}{2}, -\sqrt{3}/2)}{1} = 2,$$

Thus $E_{-\alpha^2}$ on $\left|1,-1/\sqrt{3}\right\rangle$ generates the two ♠ states.

Now the question arises whether the ♥ and ♠ at $\left(\frac{1}{2},\frac{1}{2\sqrt{3}}\right)$ are the same state or not.

This amounts to asking whether

$$E_{-\alpha^2}E_{-\alpha^1}\left|\mu_{\max}\right\rangle \quad \text{and} \quad E_{-\alpha^1}E_{-\alpha^2}\left|\mu_{\max}\right\rangle$$

Rather than continuing down and to the left, we will find that the symmetries of the root diagram will determine the rest.

Given any root, there is a reflection which one can perform analogous to $e^{i\pi L_2}$ which reflects the weight vector of any representation.

To see this, consider, for any root $\vec{\alpha}$, not necessarily simple,

$$\left[E_\alpha - E_{-\alpha}, H_i\right] = -\alpha_i\left(E_\alpha + E_{-\alpha}\right).$$

We now consider two cases.

If $\vec{\beta}\cdot\vec{\alpha} = 0$,

$$\left[E_\alpha - E_{-\alpha}, \vec{\beta}\cdot\vec{H}\right] = 0,$$

while

$$\left[E_\alpha - E_{-\alpha}, \vec{\alpha}\cdot\vec{H}\right] = -\alpha^2\left(E_\alpha + E_{-\alpha}\right).$$

Suppose we consider the state

$$|\psi\rangle = e^{-t\,(E_\alpha - E_{-\alpha})}\,|\vec{\mu}, D\rangle$$

where $\vec{\mu}$ is an arbitrary basis state with weight vector μ_i in an arbitrary representation D^i.

The Cartan generators

$$\begin{aligned}\vec{\beta}\cdot\vec{H}\,|\psi\rangle &= \vec{\beta}\cdot\vec{H}e^{-t\,(E_\alpha - E_{-\alpha})}\,|\vec{\mu}, D\rangle \\ &= e^{-t\,(E_\alpha - E_{-\alpha})}\vec{\beta}\cdot\vec{H}\,|\vec{\mu}, D\rangle = \vec{\beta}\cdot\vec{\mu}\,|\psi\rangle\end{aligned}$$

for $\vec{\beta}\cdot\vec{\alpha} = 0$,

To calculate

$$\vec{\alpha}\cdot\vec{H}\,|\psi\rangle = \vec{\alpha}\cdot\vec{H}e^{-t\,(E_\alpha - E_{-\alpha})}\,|\vec{\mu}, D\rangle$$

we use the general expression

$$e^{tA}Be^{-tA} = \sum_{n=0}^{\infty}\frac{t^n}{n!}\Omega_n(A,B),$$

where Ω_n means the n'th multiple commutator:

Let $A = E_\alpha - E_{-\alpha}$, $B = \vec{\alpha} \cdot \vec{H}$,

$$\begin{aligned} \Omega_1 &= -\alpha^2 \left(E_\alpha + E_{-\alpha}\right), \\ \Omega_2 &= -\alpha^2 \left[E_\alpha - E_{-\alpha}, E_\alpha + E_{-\alpha}\right] \\ &= -2\alpha^2 \left[E_\alpha, E_{-\alpha}\right] = -2\alpha^2 \vec{\alpha} \cdot \vec{H}. \end{aligned}$$

$$\text{Thus} \quad \Omega_n = \begin{cases} (-2\alpha^2)^{n/2} \vec{\alpha} \cdot \vec{H} & n \text{ even} \\ (-2\alpha^2)^{\frac{n-1}{2}} (-\alpha^2) \left(E_\alpha + E_{-\alpha}\right) & n \text{ odd} \end{cases},$$

and $e^{tA} \vec{\alpha} \cdot \vec{H} e^{-tA} = \vec{\alpha} \cdot \vec{H} \cos(t\sqrt{2\alpha^2}) - \sqrt{\frac{\alpha^2}{2}} \left(E_\alpha + E_{-\alpha}\right) \sin(t\sqrt{2\alpha^2})$. Let $t = \frac{\pi}{\sqrt{2\alpha^2}}$, so $e^{tA} \vec{\alpha} \cdot \vec{H} e^{-tA} = -\vec{\alpha} \cdot \vec{H}$. Now $\vec{\alpha} \cdot \vec{H} \left|\psi\right\rangle = \vec{\alpha} \cdot \vec{H} e^{-tA} \left|\vec{\mu}, D\right\rangle = e^{-tA}(-\vec{\alpha} \cdot \vec{H}) \left|\vec{\mu}, D\right\rangle = -\vec{\alpha} \cdot \vec{\mu} \left|\psi\right\rangle$. But for $\vec{\beta} \cdot \vec{\alpha} = 0$, $\vec{\beta} \cdot \vec{H} \left|\psi\right\rangle = \vec{\beta} \cdot \vec{\mu} \left|\psi\right\rangle$.

Now any vector $\vec{\gamma}$ can be written

$$\vec{\gamma} = \frac{\vec{\gamma} \cdot \vec{\alpha}}{\alpha^2} \vec{\alpha} + \vec{\beta}, \qquad \vec{\beta} = \vec{\gamma} - \frac{\vec{\gamma} \cdot \vec{\alpha}}{\alpha^2} \vec{\alpha}, \qquad \vec{\beta} \cdot \vec{\alpha} = 0.$$

$$\begin{aligned} \vec{\gamma} \cdot \vec{H} \left|\psi\right\rangle &= \left(\vec{\beta} - \frac{\vec{\gamma} \cdot \vec{\alpha}}{\alpha^2} \vec{\alpha}\right) \cdot \vec{\mu} \left|\psi\right\rangle \\ &= \left(\vec{\gamma} - 2\frac{\vec{\gamma} \cdot \vec{\alpha}}{\alpha^2} \vec{\alpha}\right) \cdot \vec{\mu} \left|\psi\right\rangle. \end{aligned}$$

Each generator H_i corresponds to $\gamma_j = \delta_{ij}$,

$$H_i |\psi\rangle = \mu_i - 2 \frac{\vec{\mu} \cdot \vec{\alpha}}{\alpha^2} \alpha_i |\psi\rangle ,$$

so ψ has weight vector $\vec{\mu}' = \vec{\mu} - 2 \frac{\vec{\mu} \cdot \vec{\alpha}}{\alpha^2} \vec{\alpha}$.

The transformation

$$e^{-\frac{\pi}{\sqrt{2\alpha^2}} (E_\alpha - E_{-\alpha})}$$

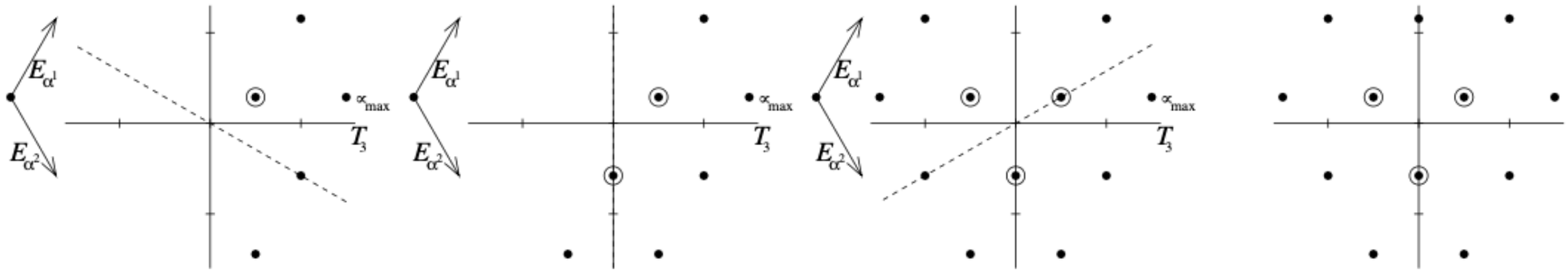
is a unitary transformation, so it makes a 1-1 correspondence between weights of weight $\vec{\mu}$ and those of weight $\vec{\mu} - \frac{2\vec{\alpha} \cdot \vec{\mu}}{\alpha^2} \vec{\alpha}$.

This corresponds to reflection in a plane (or hyperplane) perpendicular to .

Thus the weight diagram of any representation must be symmetric under such reflections.

From the part of the (2, 1) representation we have found so far, as shown in (a), we can reflect in the plane (line) perpendicular to α^1, to get the states in (b).

Then we reflect perpendicular to T_3 to get (c), then perpendicular to α^2 to get the full representation, or multiplet (d).



(a) (b) (c) (d)

This is the **15** representation of SU(3).

The reflection about the hyperplane perpendicular to $\vec{\alpha}^i$ is known as a *Weyl reflection*, and the group generated by all such reflections for a given algebra is called the *Weyl group*.

Any representation must be invariant under the Weyl group.

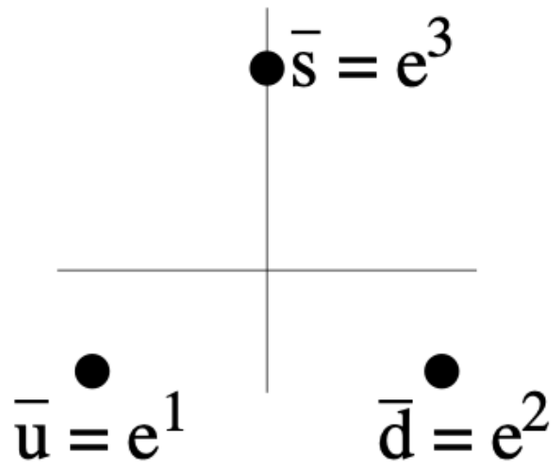
10.2 Tensor Methods

the

the the the the the

Consider the basis vectors e_1, e_2, e_3 of the defining representation of SU(3).

The algebra generators act by $T_a e_i = e_j (T_a)^j{}_i$ with $(T_a)^j{}_i = 1/2\ (\lambda_a)_{ji}$ (and we are using the summation convention that an index appearing once upstairs and once downstairs in a term is understood to be summed over).



From this point on we must be careful with upper and lower indices in another sense than in our previous discussion.

Here we are using them in the sense of co- and contra-variant quantities, as is done in relativity.

We consider the basis vectors of the conjugate representation .

The generators act here with the conjugate representation $T' = -T^*$, so

$$T_a e^i = e^j \left(T'_a\right)_j{}^i = -e^j T^*{}_j{}^i = -e^j T^i{}_j,$$

as λ is hermitean.

Now if we consider a tensor product, $e^{ij}{}_k = e^i \otimes e^j \otimes e_k$, the Lie algebra generators act as a sum of pieces (like a derivative does), so

$$T_a e^{ij}{}_k = e^{ij}{}_m T_a{}^m{}_k - e^{mj}{}_k T_a{}^i{}_m - e^{im}{}_k T_a{}^j{}_m.$$

A vector $\vec{v}$ can be specified in terms of its components, which we give indices to so as to contract with the basis vectors.

Thus a vector in the defining representation is

$$|v\rangle = v^i e_i, \quad T_a |v\rangle = v^i e_j T_a{}^j_i =: |\delta v\rangle = \delta v^j e_j,$$

with $\delta v^i = T_a{}^i{}_j\, v^j$.

A vector in the tensor product space has coefficients with several indices

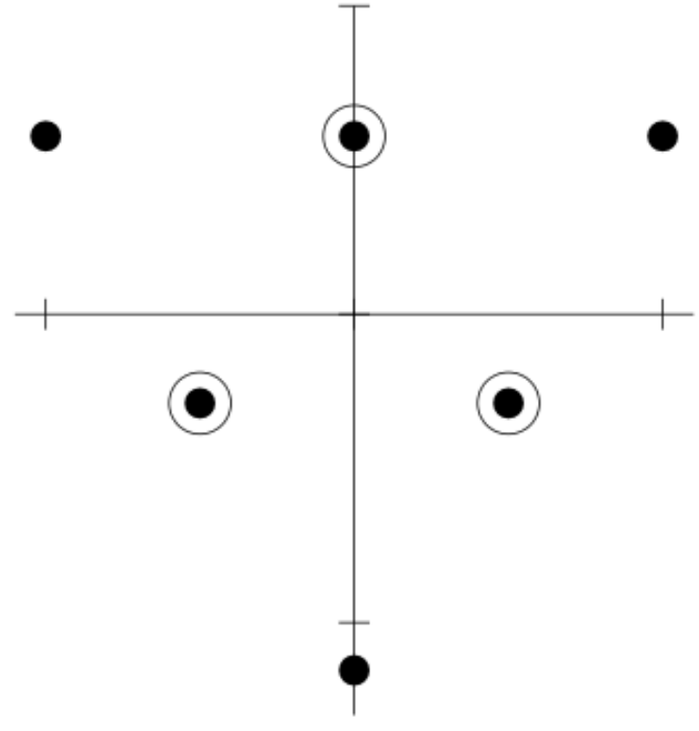
$$\begin{aligned} |v\rangle &= v_{ij}{}^k e^{ij}{}_k, \qquad T_a |v\rangle = |\delta v\rangle = \delta v_{ij}{}^k e^{ij}{}_k \\ \text{where} \quad \delta v_{ij}{}^k &= T^k{}_m v_{ij}{}^m - T^m{}_i v_{mj}{}^k - T^m{}_j v_{im}{}^k. \end{aligned}$$

The set of all states $\vec{v} = v_{ij}{}^k e^{ij}{}_k$, for arbitrary v, clearly form a representation, but it is also clearly not irreducible, because the operation of the group does nothing that alters the symmetry of the v's under $i \leftrightarrow j$).

That is, suppose we start off with a particular state $v^k{}_{ij}$, and divide it into parts $v^k{}_{ij} = s^k{}_{ij} + a^k{}_{ij}$ symmetric and antisymmetric under $i \leftrightarrow j$).

Then $s^k{}_{ij}$ is mapped into other symmetric coefficients under the group operations, while a is mapped into other antisymmetric ones, and they don't mix, so we have reduced the representation into two.

As an example, consider **3** ⊗ **3**, where **3** is the usual name for the defining representation with $q^i = (1, 0)$.



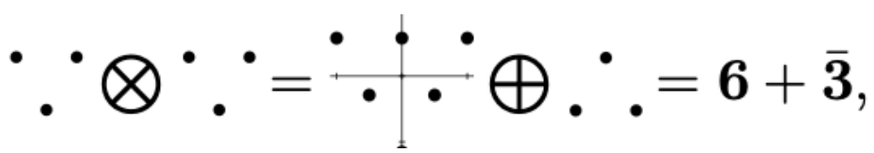
The tensor product is a 9 dimensional representation with nine basis vectors ei ⊗ej having weights μ(ei)+μ(ej), with a general vector |v∝= vij ei ⊗ ej .

The three basis vectors ei ⊗ ej with i= j have weights which can only be composed in one way, but the ones with i/= j have the same weight for i, j and for j, i.

So the weight diagram is as shown.

Dividing the space of 3 x 3 matrices v into symmetric ones s^{ij} and anti-symmetric ones a^{ij}, we see that s^{ij} forms a six dimensional space and a^{ij} a three dimensional one.

This divides



$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}},$$

each of which is irreducible.

Note that we can write the antiquark $\bar{\mathbf{3}}$, one of the fundamental representations, in terms of the quark, or defining, representation.

There is something else left invariant by the generators of SU(3).

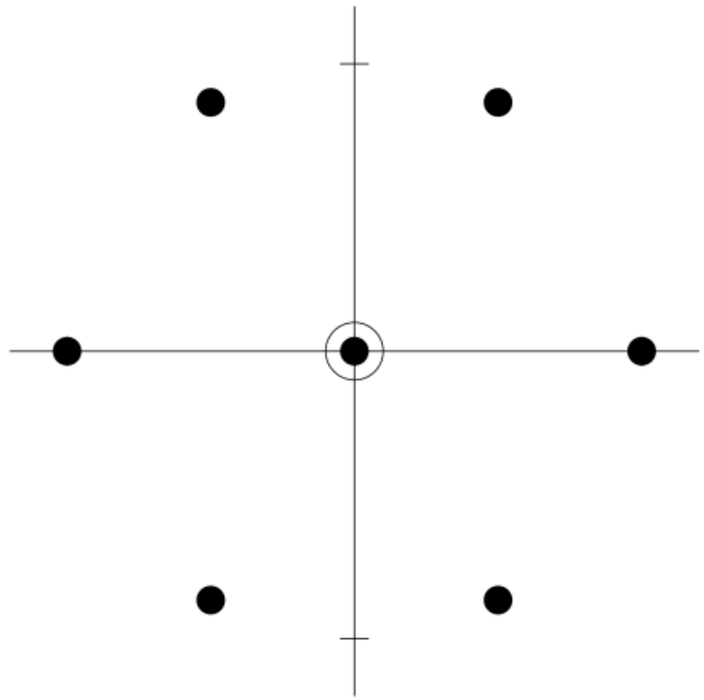
If we trace (i.e. set equal and sum) a lower with an upper index, the corresponding terms in δv cancel.

The simplest example is v_i^j of $\bar{\mathbf{3}} \otimes \mathbf{3}$.

The trace $v = \sum_i v_i^i$ is invariant under the generators,

$$\delta v_i^{\ i} = T^i_{\ m} v_i^{\ m} - T^m_{\ i} v_m^{\ i} = 0.$$

Writing $v^j_i = 1/3\, w\delta^j_i + w^j_i$ with $w = v^j_i$, which ensures that w^j_i is traceless, reduces the nine-dimensional representation $\bar{\mathbf{3}} \otimes \mathbf{3}$ into the irreducible representations **1** and **8**, where **1** is the identity representation (all group elements are represented by 1, and the generators by zero), and **8**, called the octet, is the adjoint representation.



The adjoint (or octet) representation.

Notice that $| v > := \delta^j_i\, e^i \otimes e_j$ is actually an invariant under the group, even though it looks like a tensor, i.e., both of its indices do transform under the group.

If we have a representation defined by some tensor coefficients v^k_{ij} which transform in the correct manner under the group, then the trace in any upper index paired with any lower index extracts a tensor, with fewer indices, which transforms appropriately: $v_j = \sum_i v_{ij}^{\ i}$ transforms properly as a $\bar{\mathbf{3}}$.

If we have two representations transforming properly, say u^{ij} and v^{k}_{ij} , the tensor product $w^{ij}{}_{\ell m}{}^{k} = u^{ij}\, v^{k}{}_{\ell m}$ is a tensor which transforms properly.

It can be contracted to form reduced representations, e.g. $w^{k} = w^{ij}{}_{ij}{}^{k} = u^{ij}\, v^{k}{}_{ij}$ is a **3**.

So this is one way of extracting a smaller representation from the tensor product.

But we can also impose symmetries to reduce it.

We will now construct the arbitrary SU(3) irreducible representation (n, m) from reducing a tensor product of n defining representations and m of its conjugate.

First construct w = D(n,0) . $\vec{\mu}_{\max} = n\vec{\mu}^{1}$, so $|\, \vec{\mu}_{\max} > = e_1 \otimes e_1 \otimes \cdots e_1$ (n of them).

This state corresponds to the tensor $w^{j_1 \cdots j_n} = \prod_i \delta_{j_i,1}$ which is clearly symmetric under interchange of any two indices.

How many components of w of the full representation are there?

One needs only to know how many 1's, 2's and 3's are picked to make a total of n.

If you choose r 1's, there are n - r + 1 choices of how many 2's to pick, so in total there are

$$\sum_{r=0}^{n}(n-r+1) = \sum_{1}^{n+1} j = \binom{n+2}{2} = \frac{(n+2)!}{n!\,2!} = \frac{1}{2}(n+1)(n+2)$$

choices, so w has (n + 1)(n + 2)/2 independent components, and the $D^{(n,0)}$ representation is (n + 1)(n + 2)/2 dimensional.

The same argument applies to the representation $D^{(0,m)}$ of weight $\vec{\mu}_{\max} = m\vec{\mu}^2$,

$$|\vec{\mu}_{\max}\rangle = e^2 \otimes e^2 \otimes \cdots e^2, \qquad (\text{with } m \text{ factors}).$$

So the $u_{k_1 \cdots k_m}$ corresponding to this representation is totally symmetric in all its indices, and is a (m + 1)(m + 2)/2 dimensional representation.

The tensor product of $w^{j_1 \cdots j_n}$ with $u_{k_1 \cdots k_m}$ can be reduced

$$w^{j_1 \cdots j_n} u_{k_1 \cdots k_m} = v^{j_1 \cdots j_n}_{k_1 \cdots k_m} + \sum_{rs} \delta^{j_r}_{k_s} X^{j_1 \cdots \hat{j}_r \cdots j_n}_{k_1 \cdots \hat{k}_s \cdots k_m},$$

where $\hat{j}_r$ means leave out the j_r index.

The division is arranged so that v is traceless,

$$\sum_{j_1} v^{j_1 j_2 \cdots j_n}_{j_1 k_2 \cdots k_m} = 0.$$

The $\binom{n+2}{2}\binom{m+2}{2}$ degrees of freedom in wu have had $\binom{n-1+2}{2}\binom{m-1+2}{2}$ degrees of freedom in X constrained out, leaving

$$\begin{aligned} \text{Dim } D^{(n,m)} &= \frac{(n+2)(n+1)(m+2)(m+1)}{4} - \frac{(n+1)n(m+1)m}{4} \\ &= \frac{(n+1)(m+1)(n+m+2)}{2} \end{aligned}$$

for the dimension of the (n, m) representation.

We have extracted the leading irreducible representation from the product, but we have not fully reduced the product.

Although v is irreducible, we could reduce X iteratively in the same way, extracting the traces and being left with traceless parts.

Thus we find

$$D^{(n,0)} \otimes D^{(0,m)} = D^{(n,m)} \oplus D^{(n-1,m-1)} \oplus \cdots \oplus D^{(n-m,0)}$$

if $n \geq m$, or ending with $D^{(0,m-n)}$ if $m > n$.

We have constructed an arbitrary representation of SU(3) by products of $\mathbf{3}$'s and $\bar{\mathbf{3}}$'s, but we have also seen that **3** is the antisymmetric part in the product of two **3**'s.

So any representation can be built of the defining representation alone.

This is a general feature of SU(n).

The states, however, will not correspond to some simple symmetry under permutations.

Consider $\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3}$, the states of three quarks.

The highest weight is 3 u's, so $(3,0) = \mathbf{10}$ is the symmetric part.

There is also a totally antisymmetric part w^{ijk}, but as there is only one choice for $\{i, j, k\}$ which doesn't vanish by antisymmetry, there is only one degree of freedom here, $w^{ijk} = w\epsilon^{ijk}$, so this is the one dimensional identity representation $\mathbf{1}$.

The remaining 16 degrees of freedom are in fact two octets (two $\mathbf{8}$'s).

We will have to show this.

For rank greater than 2, using $\otimes_i(D^i)q_i$ requires 3 or more sets of indices, and our notational skills are not up to that.

But, as we saw for SU(3), a fundamental representation may be extracted from a tensor product of copies of the defining one.

Many texts show that for SU(N), all fundamental representations can be extracted from tensor products of the *defining* representation, with mixed symmetries.

If we start with the tensor product of k defining representations of SU(n), we have an n^k dimensional space which is not only a representation of SU(n) but also of S_k, the permutation group on the indices.

In fact, these operations commute, so we may reduce the space into simultaneous representations.

So we must first learn something about representations of the permutation group.

Part 11 S_k and Tensor Representations

If we have an arbitrary tensor with k indices $W^{i_1,\cdots,i_k}$ we can act on it with a permutation

$$P = \begin{pmatrix} 1 & 2 & \cdots & k \\ a & b & \cdots & \ell \end{pmatrix}$$

so

$$(Pw)^{i_1,i_2,\cdots,i_k} = w^{i_a,i_b,\cdots,i_\ell}.$$

Consider the algebra $\mathcal{A}$ formed by taking arbitrary linear combinations of the different permutations, considered as operators acting on the space of k'th rank tensors.

This algebra can be constructed for any group, particularly finite groups, and is called the **group algebra**. (this is **not** the Lie algebra!).

Note that this sum of permutations makes sense only as operators on a vector space.

It is not the composition of permutations.

Also note that as $\mathcal{A}$ is an algebra (see definitions below), one can both add and multiply (by composition) elements in $\mathcal{A}$.

Definition: An *algebra* consists of a vector space V over a field F, together with a binary operation of multiplication on the set V of vectors, such that for all $a \in F$ and $\alpha, \beta, \gamma \in V$, the following are satisfied:

1. $(a\alpha)\beta = a(\alpha\beta) = \alpha(a\beta)$
2. $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$
3. $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$

V is an *associative algebra over* F if, in addition,

4 $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ for all $\alpha, \beta, \gamma \in V$.

The group algebra is useful because it can extract the tensors of specified symmetry.

First consider tensors of rank 2.

Writing 𝕀 = 1/2 (𝕀 + (1 2)) +1/2 (𝕀 - (1 2)) we can extract

$$
\begin{aligned}
s^{ij} &= \frac{1}{2}\left(\mathbb{I} + (1\,2)\right) w^{ij} \\
a^{ij} &= \frac{1}{2}\left(\mathbb{I} - (1\,2)\right) w^{ij}
\end{aligned}
$$

and $w^{ij} = s^{ij} + a^{ij}$ is a decomposition into a symmetric tensor and an anti-symmetric tensor.

The action of the permutations commutes with the SU(n) rotations on the tensors, so a constraint on a tensor of the form Aw = 0 for some A ∈ $\mathcal{A}$, if it holds for one state of an irreducible representation of SU(n), will hold on all states in that representation.

Thus s and a are separate representations.

Now consider a rank 3 tensor w^{ijk}, and define

$$s^{ijk} = \frac{1}{6} \sum_{P \in S_3} P w^{ijk}$$

$$a^{ijk} = \frac{1}{6} \sum_{P \in S_3} (\text{sign } P) P w^{ijk}$$

These are the totally symmetric and totally antisymmetric parts of w, but it is not all of w.

For example, suppose $w^{112} = w^{121} = 1$, $w^{211} = -2$, all other components zero.

Then s^{ijk} and a^{ijk} are both zero.

The rest is related to the two-dimensional representation of S_3.

In general, there will be operators in $\mathcal{A}$ associated with the different irreducible representations of S_k, which extract the corresponding irreducible representations of SU(n).

So we now turn to the problem of finding the irreducible representations of S_k.

11.1 Irreducible Representations of S_k

We know in general that the number of irreducible representations is the number of conjugacy classes.

So let us begin with that.

Any element of S_k can be written as a product of disjoint cycles.

For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix} = (1\,2\,3)(4\,5)$$

This factorization is unique (remember (1 2 3) = (2 3 1)) up to the order of the factors, which commute because they are disjoint cycles.

Under conjugation by

$$P = \begin{pmatrix} 1 & 2 & \cdots & k \\ P_1 & P_2 & \cdots & P_k \end{pmatrix}$$

a cycle simply has its elements permuted.

Thus $P(i\ j\ k)P^{-1} = (P_i\,P_j\,P_k)$.

This is true for products of cycles as well.

Thus two permutations whose descriptions in terms of disjoint cycles contain the same number of cycles of each length are conjugate, and only those are.

We describe the conjugacy class of elements describable in terms of disjoint cycles, I_k of length ℓ_k, as $\left(\ell_1^{i_1}\ \ell_2^{i_2}\ \cdots\right)$.

Including one-cycles for any element left unmoved, we have $\sum_m i_m \ell_m = k$.

Example: S_3

permutations		class	
$\mathbb{1}$	$\in$	(1^3)	
$(1\,2) = (1\,2)(3)$	$\in$	$(2,1)$	
$(2\,3),\ (1\,3)$	$\in$	$(2,1)$	as well
$(1\,2\,3),\ (1\,3\,2)$	$\in$	(3)	

There is one conjugacy class for each partition of k.

A partition of an integer k is an unordered set of positive integers, possibly with repeats, which add to k.

Example: How many classes are there in S_5?

$$(5);\quad (4,1);\quad (3,1,1) = (3,1^2);\quad (3,2);\quad (2^2,1);\quad (2,1^3);\quad (1^5)$$

answer: 7.

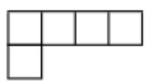
Thus we also know that there are that many irreducible representations, although there is not a straightforward correspondance between the representations and the conjugacy classes.

Define a **Young graph** for S_k as a set of k boxes arranged, left-justified, in rows each of which is no longer than the preceeding.

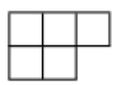
The lengths of the rows provide a partition of k.

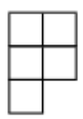
So

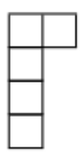
(5) 

$(4,1)$ 

$(3,1^2)$ 

$(3,2)$ 

$(2^2,1)$ 

$(2,1^3)$ 

(1^5) 

There is one irreducible representation of S_k corresponding to each Young graph.

A **Young tableau** is a Young graph with the numbers $1, 2, \cdots, k$ inserted in the boxes in some order, for example

$$\tau = \begin{array}{|c|c|c|}\hline 2 & 3 & 4 \\ \hline 5 & 1 \\ \cline{1-2}\end{array}.$$

For each tableau we define an element of the group algebra, $P_\tau = \sum P$, where the sum is over those permutations which permute the numbers within each row but do not move them from one row to another.

Here

$$P_\tau = [\mathbb{1} + (2\,3) + (3\,4) + (2\,4) + (2\,3\,4) + (2\,4\,3)]\,[\mathbb{1} + (5\,1)]\,,$$

which includes 12 of the 120 permutations in S_5.

We also associate $Q_\tau = \sum(\text{sign P})P$ where the sum includes only permutations which permute numbers in the same column but don't move numbers from one column to another.

Thus

$$Q_\tau = [\mathbb{I} - (2\,5)]\,[\mathbb{I} - (1\,3)]\,.$$

Finally we define the **Young operator** $Y_\tau = Q_\tau\, P_\tau$.

We see that the way to get a totally symmetric rank 5 tensor is to apply $Y_{\square\square\square\square\square}$ to an arbitrary one while you get a totally antisymmetric tensor by applying $Y_{\begin{smallmatrix}\square\\\square\\\square\\\square\\\square\end{smallmatrix}}$, with the numbers in any order in the boxes.

The Y_τ corresponding to any Young tableau τ is almost, but not quite, the element of the group algebra we want to extract irreducible representations.

We find a related set of basis vectors in the group algebra by using the representations of S_k.

Define

$$e^\eta_{ij} = \frac{\ell_\eta}{k!} \sum_{P \in S_k} \Gamma^\eta_{ji}(P^{-1})P,$$

where η is the Young graph corresponding to an irreducible representation of S_k, and ℓ_η is the dimension of that representation.

The sum is over all the permutations.

For $\square\square\square\square\square$, $\ell_\eta = 1$, $\Gamma = 1$, and $e^{\square\square\square\square\square}_{11} = Y_{\boxed{1}\boxed{2}\boxed{3}\boxed{4}\boxed{5}}$ which is also equal to any other Young operator for a tableau in $\square\square\square\square\square$.

For $\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}$, $\ell_\eta = 1$, $\Gamma = \mathrm{sign}\, P$, $e^{\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}}_{11} = Y_{\begin{smallmatrix}\boxed{1}\\\boxed{2}\\\boxed{3}\end{smallmatrix}}$. Other Young operators in $\begin{smallmatrix}\square\\\square\\\square\end{smallmatrix}$ differ only in sign, *e.g.* $Y_{\begin{smallmatrix}\boxed{2}\\\boxed{1}\\\boxed{3}\end{smallmatrix}} = \mathrm{sign}(1\,2)\cdot Y_{\begin{smallmatrix}\boxed{1}\\\boxed{2}\\\boxed{3}\end{smallmatrix}} = -Y_{\begin{smallmatrix}\boxed{1}\\\boxed{2}\\\boxed{3}\end{smallmatrix}}$.

The e's have some marvelous properties.

They form vector spaces transforming as irreducible representations under S_k separately from the right and from the left: For $Q \in S_k$,

$$\begin{aligned} Qe^\eta_{ij} &= \frac{\ell_\eta}{k!}\sum_P \Gamma^\eta_{ji}\left(P^{-1}\right) QP = \frac{\ell_\eta}{k!}\sum_m \Gamma^\eta_{mi}(Q)\sum_P \Gamma^\eta_{jm}\left(P^{-1}Q^{-1}\right) QP \\ &= \frac{\ell_\eta}{k!}\sum_m \Gamma^\eta_{mi}(Q)\sum_R \Gamma^\eta_{jm}\left(R^{-1}\right) R = \sum_m \Gamma^\eta_{mi}(Q)e^\eta_{mj} \end{aligned}$$

where we again used the rearrangement theorem.

Thus Q acts just the way you would expect for a basis vector e_i of representation η to transform, for each fixed j.

From the other side, $e^\eta_{ij}Q = \sum_m \Gamma^\eta_{jm}(Q)e^\eta_{im}$.

We say that the set e^{η}_{ij} is a two sided ideal (or invariant subalgebra) of the group algebra over S_k.

This gives the e's an interesting algebra:

$$
\begin{aligned}
e^{\eta}_{ij} e^{\eta'}_{mn} &= \frac{\ell_\eta \ell_{\eta'}}{(k!)^2} \sum_{P,P' \in S_k} \Gamma^{\eta}_{ji}\left(P^{-1}\right) \Gamma^{\eta'}_{nm}\left(P'^{-1}\right) PP' \\
&= \frac{\ell_\eta \ell_{\eta'}}{(k!)^2} \sum_{P,R \in S_k} \Gamma^{\eta}_{ji}\left(P^{-1}\right) \sum_{p} \Gamma^{\eta'}_{np}\left(R^{-1}\right) \Gamma^{\eta'}_{pm}\left(P\right) R \\
&= \frac{\ell_\eta}{k!} \sum_{P \in S_k} \Gamma^{\eta *}_{ij}(P) \sum_{p} \Gamma^{\eta'}_{pm}\left(P\right) e^{\eta'}_{pn} \qquad \text{by unitarity} \\
&= \delta_{\eta\eta'} \delta_{jm} e^{\eta}_{in} \qquad \text{by the great orthogonality theorem}
\end{aligned}
$$

We may also show that the diagonal elements e^{η}_{ii} form a decomposition of the identity.

From the great orthogonality theorem "transposed" ,

$$\delta_{GG'} = \sum_{ij\eta} \frac{\ell_\eta}{k!} \Gamma^{\eta *}_{ij}(G') \Gamma^{\eta}_{ij}(G)$$

we can write the identity element of S_k as

the the the the the

$$\begin{aligned}\mathbb{I} &= \sum_G \delta_{G^{-1},\mathbf{I}} G = \sum_{ij\eta G} \frac{\ell_\eta}{k!} \Gamma^{\eta *}_{ij}(\mathbb{I}) \Gamma^{\eta}_{ij}(G^{-1}) G \\ &= \sum_{i\eta G} \frac{\ell_\eta}{k!} \Gamma^{\eta}_{ii}(G^{-1}) G = \sum_{\eta\, i} e^{\eta}_{ii}\end{aligned}$$

so

$$\mathbb{I} = \sum_{\eta\, i} e^{\eta}_{ii}$$

Thus the whole algebra is spanned by these two sided ideals.

In particular, the Y_τ are contained in the corresponding e^{η}_{ij} (an ℓ^2_η dimensional algebra).

In fact, the space spanned by e^{η}_{ij} is also spanned by $Q_i\, s_{ij}\, P_j$, where Q_i and P_i are the antisymmetrizers and symmetrizers of a set of standard tableaux for η, which means tableaux in which the numbers increase left to right in each row, and also top to bottom in each column.

Thus $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array}$ and $\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array}$ are standard tableaux, but $\begin{array}{|c|c|}\hline 3 & 2 \\ \hline 1 \\ \cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array}$ and $\begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array}$ are not.

Here s_{ij} is the permutation such that $\tau_i = s_{ij}\, \tau_j$.

Each of these spaces has dimension ℓ^2_η , with ℓ_η equal to the number of standard tableaux, so

The dimension of Γ^η is the number of standard tableaux of η.

Counting all possibilities is tedious, so we have a magic formula in terms of hooks.

For each box b in a Young graph with k boxes, define the hook of b, g_b = 1 plus the number of boxes directly to the right plus the number of boxes directly beneath.

Then

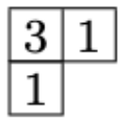
$$\ell_\eta = \frac{k!}{\prod_b g_b}.$$

Example: In the Young graph I have placed the corresponding hooks (this is not a Young tableau)

7	5	3	2
6	4	2	1
3	1		
1			

$$\ell = \frac{11!}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3^2 \cdot 2^2 \cdot 1^3} = 1320.$$

It would be hard to count this explicitly.

For our more reasonable case  ,  gives $\ell = 3!/3 = 2$.

11.2 Representations of SU(n)

We now turn to the extraction of arbitrary representations of SU(n).

An arbitrary representation can be found from a tensor product of an adequate number of defining representations.

The problem is to extract from the tensor product of k defining representations $\otimes N^k$ the irreducible pieces.

We have seen that this can be done by demanding that elements of the permutation group algebra vanish.

If we impose $e^{\eta}_{ii} w = 0$ for all η and i save one, that is equivalent to projecting out our representation

$$T^{a_1 a_2 \cdots a_k} = (e^{\eta}_{ii} w)^{a_1 a_2 \cdots a_k} \qquad \text{no sum on } i$$

for one particular representation η and one basis vector i.

The different i generate equivalent representations.

The different η's, however, each correspond to a different (inequivalent) representation of SU(n).

Before doing more formal arguments, we will do an example.

Consider three spin 1/2 objects, or the tensor product of three defining representations of SU(2).

We will extract from this 8 dimensional state space the piece $e_{\begin{smallmatrix}\square\square\\\square\end{smallmatrix}}$.

Some, not so easy calculations give

$$
\begin{aligned}
e^{\ydiagram{2,1}}_{11} &= \frac{2}{6}\sum_P \Gamma^{e^{\ydiagram{2,1}}_{11}}\left(P^{-1}\right) P \\
&= \frac{1}{3}\left(\mathbb{1} + (1\,2) - \frac{1}{2}(2\,3) - \frac{1}{2}(1\,3) - \frac{1}{2}(1\,2\,3) - \frac{1}{2}(1\,3\,2)\right),
\end{aligned}
$$

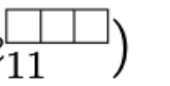
Let this act on the basis vectors which we expand as $\uparrow = e_1$, $\downarrow = e_2$.

v	$e^{\ydiagram{2,1}}_{11} v$
$\uparrow\uparrow\uparrow$	$\frac{1}{3}\uparrow\uparrow\uparrow + \frac{1}{3}\uparrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\uparrow = 0$
$\uparrow\uparrow\downarrow$	$\frac{1}{3}\uparrow\uparrow\downarrow + \frac{1}{3}\uparrow\uparrow\downarrow - \frac{1}{6}\uparrow\downarrow\uparrow - \frac{1}{6}\downarrow\uparrow\uparrow - \frac{1}{6}\downarrow\uparrow\uparrow - \frac{1}{6}\uparrow\downarrow\uparrow = \frac{2}{3}\uparrow\uparrow\downarrow - \frac{1}{3}\uparrow\downarrow\uparrow - \frac{1}{3}\downarrow\uparrow\uparrow$
$\uparrow\downarrow\uparrow$	$\frac{1}{3}\uparrow\downarrow\uparrow + \frac{1}{3}\downarrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\downarrow - \frac{1}{6}\uparrow\downarrow\uparrow - \frac{1}{6}\uparrow\uparrow\downarrow - \frac{1}{6}\downarrow\uparrow\uparrow = -\frac{1}{3}\uparrow\uparrow\downarrow + \frac{1}{6}\uparrow\downarrow\uparrow + \frac{1}{6}\downarrow\uparrow\uparrow$
$\downarrow\uparrow\uparrow$	$\frac{1}{3}\downarrow\uparrow\uparrow + \frac{1}{3}\uparrow\downarrow\uparrow - \frac{1}{6}\downarrow\uparrow\uparrow - \frac{1}{6}\uparrow\uparrow\downarrow - \frac{1}{6}\uparrow\downarrow\uparrow - \frac{1}{6}\uparrow\uparrow\downarrow = -\frac{1}{3}\uparrow\uparrow\downarrow + \frac{1}{6}\uparrow\downarrow\uparrow + \frac{1}{6}\downarrow\uparrow\uparrow$
$\uparrow\downarrow\downarrow$	$\frac{1}{3}\uparrow\downarrow\downarrow + \frac{1}{3}\downarrow\uparrow\downarrow - \frac{1}{6}\uparrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\uparrow - \frac{1}{6}\downarrow\uparrow\downarrow - \frac{1}{6}\downarrow\downarrow\uparrow = \frac{1}{6}\uparrow\downarrow\downarrow + \frac{1}{6}\downarrow\uparrow\downarrow - \frac{1}{3}\downarrow\downarrow\uparrow$
$\downarrow\uparrow\downarrow$	$\frac{1}{3}\downarrow\uparrow\downarrow + \frac{1}{3}\uparrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\uparrow - \frac{1}{6}\downarrow\uparrow\downarrow - \frac{1}{6}\downarrow\downarrow\uparrow - \frac{1}{6}\uparrow\downarrow\downarrow = \frac{1}{6}\uparrow\downarrow\downarrow + \frac{1}{6}\downarrow\uparrow\downarrow - \frac{1}{3}\downarrow\downarrow\uparrow$
$\downarrow\downarrow\uparrow$	$\frac{1}{3}\downarrow\downarrow\uparrow + \frac{1}{3}\downarrow\downarrow\uparrow - \frac{1}{6}\downarrow\uparrow\downarrow - \frac{1}{6}\uparrow\downarrow\downarrow - \frac{1}{6}\uparrow\downarrow\downarrow - \frac{1}{6}\downarrow\uparrow\downarrow = -\frac{1}{3}\uparrow\downarrow\downarrow - \frac{1}{3}\downarrow\uparrow\downarrow + \frac{2}{3}\downarrow\downarrow\uparrow$
$\downarrow\downarrow\downarrow$	$\frac{1}{3}\downarrow\downarrow\downarrow + \frac{1}{3}\downarrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\downarrow - \frac{1}{6}\downarrow\downarrow\downarrow = 0$

Notice this only results in one state of $J_z = 1/2$ and one of $J_z = -1/2$.

So e_{11}  projects out a 2-dimensional s = 1/2 state.

e_{22}  would project out an orthogonal spin 1/2 .

Thus the tensor product of three spin 1/2's is a spin 3/2 (the totally symmetric part, e_{11} ) and two spin 1/2 representations, 2 x 2 x 2 = 4 + 2 + 2.

Having completed this trivial but tedious example of the simple case of SU(2) and  , we are ready for some abstract reasoning.

Now we consider the general case of $\mathbf{N}^k$.

The basis vectors which are mixed by the permutations are only those with the same number of indices equal to 1, and the same number equal to 2, etc..

Consider the subspace with ri of the indices equal to i, with $\sum r_i = k$, each $r_i = 1, \ldots, N$.

This subspace $\mathcal{S}^{\vec{r}}$ is spanned by the basis vector

$$e = \underbrace{e_1 \otimes e_1 \cdots \otimes e_1}_{r_1 \text{ times}} \otimes \underbrace{e_2 \cdots \otimes e_2}_{r_2 \text{ times}} \cdots \otimes \underbrace{e_N \cdots \otimes e_N}_{r_N \text{ times}},$$

together with all permutations P e, for $P \in S_k$.

If all the indices are different, all $r_i = 0$ or 1, all of the permutations are inequivalent, and we get a k! dimensional space.

But if the r_i's are not all ≤ 1, there is a subgroup $\mathcal{P} \subset S_k$ with B e = e for $B \in \mathcal{P}$.

In fact, $\mathcal{P} = S_{r_1} \times S_{r_2} \times \cdots S_{r_N}$.

Let $P_{\mathcal{P}} = \sum_{B \in \mathcal{P}} B$ which is a element of the group algebra $\mathcal{A}$.

Then while the subspace $\mathcal{S}^{\vec{r}}$ is spanned by $\{P\, e | P \in S_k\}$ it is also spanned by $\{P\, P_{\mathcal{P}}\, e | P \in S_k\}$.

We now want to extract from $\mathcal{S}^{\vec{r}}$ the piece projected out by $e^{\eta_{ii}}$.

The products $\{e^{\eta_{ii}} P\}$ for all P are just sums of multiples of $e^{\eta_{ij}}$, for all j so we want to know the dimension of the space $\{e^{\eta_{ij}}\, P_{\mathcal{P}} | j = 1, \ell_\eta\}$.

As e^η is a two-sided ideal, this space is $\langle \sum_k e^{\eta_{ik}} b_k \rangle$, so the dimensionality depends on the constraints on b_k.

If they were all independent, they would form an ℓ_η dimensional space.

But there are constraints.

For $B \in \mathcal{P}$, $P_{\mathcal{P}}\, B = P_{\mathcal{P}}$ - let's be more explicit:

$$e^{\eta}_{ij} P_{\mathcal{P}} = \sum_n e^{\eta}_{in} b_{nj} = e^{\eta}_{ij} P_{\mathcal{P}} B = \sum_n b_{nj} e^{\eta}_{in} B = \sum_{nm} \Gamma^{\eta}_{nm}(B) e^{\eta}_{im} b_{nj}.$$

The e^{η}_{im} are linearly independent, so $b_{nj} = \sum_m b_{mj} \Gamma^{\eta}_{mn}(B)$, for $B \in \mathcal{P}$.

To find out how many degrees of freedom survive this constraint for each j, observe that $\Gamma^{\eta}_{mn}(B)$ forms a reducible representation of the subgroup $\mathcal{P}$.

So we can write

$$\Gamma^{\eta}(B) = U \bigoplus_{\epsilon} \Gamma^{\epsilon}(B)\, U^{-1}$$

where Γ^{ϵ} are irreducible representations of $\mathcal{P}$.

Now if $c = bU$,

$$b = b\Gamma^{\eta}(B) = bU \bigoplus \Gamma^{\epsilon}(B) U^{-1} \Longrightarrow c = c\left(\bigoplus \Gamma^{\epsilon}(B)\right).$$

The vector c breaks up into pieces for each representation ϵ, with $c^{\epsilon}\Gamma^{\epsilon}(B) = c^{\epsilon}$ for all $B \in \mathcal{P}$.

This is possible for nonzero c only if ϵ is the identity representation, as the representations are irreducible.

the the the the the

Therefore the dimensionality of the space $e^{\eta}_{ii}Pe$ is the number of times, γ_η, that the identity representation of $\mathcal{P}$ is contained in Γ^η.
But the number of times the representation i is contained in Γ^η is

$$\gamma_\eta = a_{i=\mathbf{1}} = \frac{1}{g_\mathcal{P}} \sum_{B\in\mathcal{P}} \chi^{i*}(B)\chi^\eta(B) = \frac{1}{g_\mathcal{P}} \sum_{B\in\mathcal{P}} \chi^\eta(B)$$

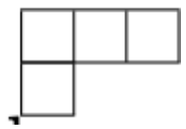
for i=identity, where $g_\mathcal{P}$ is the number of elements in $\mathcal{P}$, which is $\prod r_i!$

So, we now know how to extract the irreducible representations or just to count their dimensionality.

Now it is time for magic.

The number γ_η of states extracted by e^η_{ii} from $\mathcal{S}^{\vec{r}}$, the space spanned by $P(\bigotimes e_i^{r_i})$ by all $P \in S_k$, is given by the number of ways one can place r_1 1's, r_2 2's,$\cdots$ in the Young graph so that in each row the numbers do not decrease, and in each column they increase.

This is called a **permissible placement.**

To see how this works, let's check it out on  for SU(N).

For $r_1 = 4$ there is no way to avoid two 1's in the same column, so $\gamma = 0$.

For $r_1 = 3$ and $r_2 = 1$, the 2 has to be in the second row, so there is only one way, $\gamma = 1$.

For $r_1 = r_2 = 2$, the only possibility is $\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 2\\\hline\end{array}$, $\gamma = 1$.
For $r_1 = 2, r_2 = r_3 = 1$, we have $\begin{array}{|c|c|}\hline 1&1\\\hline 3\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 2\\\hline\end{array}$ and $\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 3\\\hline\end{array}$, so $\gamma = 2$.
For $r_1 = r_2 = r_3 = r_4 = 1$, we have $\begin{array}{|c|c|}\hline 1&2\\\hline 4\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 3\\\hline\end{array}$, $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 4\\\hline\end{array}$, $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 4\\\hline\end{array}$, and $\gamma = 3$.

Note in our previous method, it was clear that these numbers only depended on the set $\{r_i\}$ and not on the order.

This is now not obvious.

Consider $r_1 = 1, r_2 = 2, r_4 = 1$. Then we have $\begin{array}{|c|c|}\hline 1&2\\\hline 4\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 4\\\hline\end{array}$ and $\begin{array}{|c|c|}\hline 1&2\\\hline 4\\\cline{1-1}\end{array}\begin{array}{|c|}\hline 2\\\hline\end{array}$, so again $\gamma = 2$, as for $r_1 = 2, r_2 = r_3 = 1$.

Now to count the dimensionality of an irreducible representation of SU(N) belonging to the Young graph η, we must sum, over all choices r_i, the corresponding γ_η .

But for each choice of r_i the γ_η is the number of ways of placing the indices in the graph in a permissible fashion.

So the dimension of the full irreducible representation of SU(N) is the number of ways of placing k integers, chosen from $1, 2, , \ldots$ N (repeats allowed) in a permissible fashion in η.

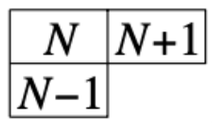
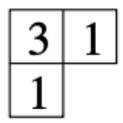
Example $\begin{array}{|c|c|}\hline &\\\hline &\\\cline{1-1}\end{array}$ for SU(3): $\begin{array}{|c|c|}\hline 1&1\\\hline 2\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 1&1\\\hline 3\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 1&2\\\hline 2\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 1&3\\\hline 3\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 2&2\\\hline 3\\\cline{1-1}\end{array}$, $\begin{array}{|c|c|}\hline 2&3\\\hline 3\\\cline{1-1}\end{array}$, for a total of 8 dimensions.

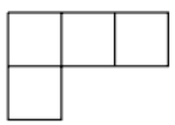
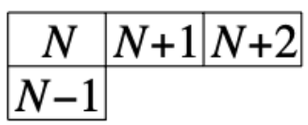
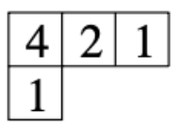
There is an easier method of finding the dimensionality.

For each box, associate the value (N + column number - row number).

Then divide by the hook of that box.

The dimension of the representation is the product of these quotients over all the boxes.

Examples: Dim  =  /  = $\frac{N(N^2-1)}{3}$

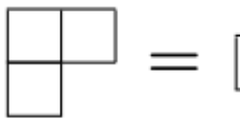
Dim  =  /  = $\frac{(N+2)!}{8(N-2)!}$

Note: If the first column of a graph has N boxes, the hook of each box in column 1 is equal to the (N + column number - row number) of the last box in that row.

Thus eliminating the first row does not change the dimension.

In fact, it does not change the representation either.

This is because a totally antisymmetric tensor with N indices is invariant.

Thus in SU(2),  , .

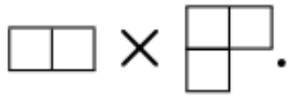
It also means that for SU(N), we needn't consider representations with N or more rows (except perhaps to indicate the identity representation by one column of N boxes).

Part 12 Tensor Products of Irreducible Representations

Consider two representations with Young Graphs η_1 and η_2, corresponding to tensors of rank k_1 and k_2.

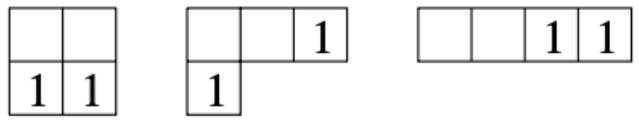
The tensor product is a tensor of rank $k_1 + k_2$, and must be decomposed into irreducible representations corresponding to Young Graphs of $S_{k_1+k_2}$.

Which ones?

I will give only the magic rule, and simultaneously we will do as an example  .

1) Take the second graph and place 1's in the first row of boxes, 2's in the second row, *etc.*:  .

2) Take the boxes of the top row of the second graph and add them to the first graph in all possible ways resulting in Young graphs. No two boxes may go in the same column.

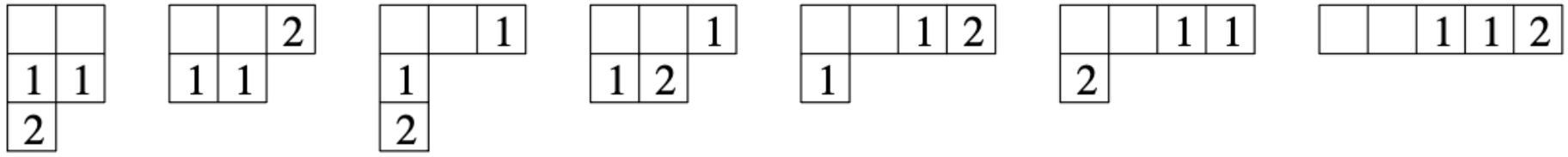


Successively add the boxes from the lower rows, one row at a time, working down. The configurations are constrained, however, by

(a) After adding each row, the graph must be a Young graph.

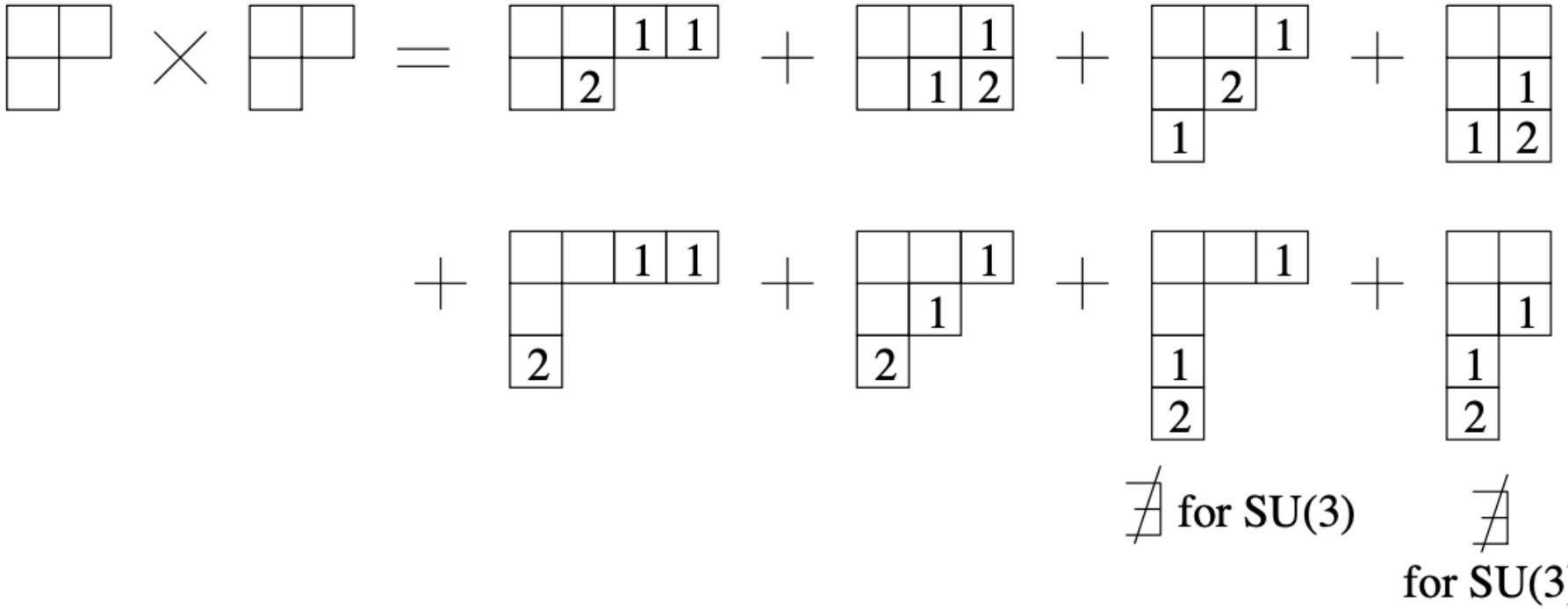
(b) No two boxes in a column can have the same number.

(c) Reading *right-to-left* and top-to-bottom (Hebrew fashion), at any point there must be no more $j+1$'s encountered than j's.

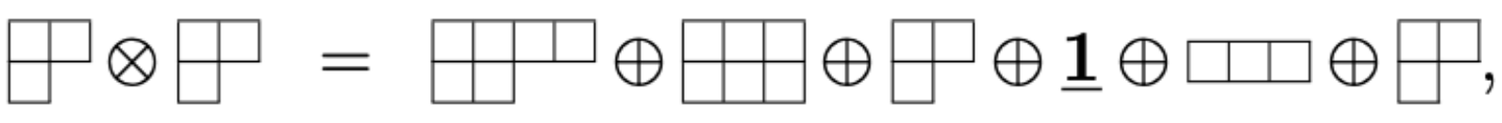


In SU(3), $\mathbf{6} \times \mathbf{8} = \overline{\mathbf{3}} + \mathbf{6} + \overline{\mathbf{15}} + \mathbf{24}$.

A very important example in SU(3) is octet ⊗ octet:



Recall that at most N boxes can be in a column, and a column of N boxes can be dropped, so for SU(3),



or $\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \overline{\mathbf{10}} \oplus \mathbf{8} \oplus \mathbf{1} \oplus \mathbf{10} \oplus \mathbf{8}$.

Notice that the octet appears twice in the product of two octets, as each different labelling counts.

This complicates the Clebsch-Gordonry.

For example, the when one considers πN scattering, we find that the S matrix was SU(2) invariant, so we can insert a set of intermediate projections onto states of total isospin T , with amplitudes $a_{1/2}$ and $a_{3/2}$.

But now let us generalize to the scattering of any baryon B in the same SU(3) multiple (the octet) as the proton (which means p or n or Λ or $\Sigma^{\pm}$ or Σ^0 or Ξ^0 or Ξ^-) scattering off any meson M in the meson octet: K, $\bar{K}$, π, η.

To the extent that we can ignore the breaking of SU(3) symmetry, we can combine initial two-particle states into states $| \eta, T_3, Y >$ for $\eta = \mathbf{27}, \overline{\mathbf{10}}, \mathbf{10}, \mathbf{1},$ and $\mathbf{8}$, but for the $\mathbf{8}$, there are two ways to combine B's and M 's into each state of the intermediate octet.

These are called $\mathbf{8}_s$ and $\mathbf{8}_a$.

The final state can be similarly combined.

Because the scattering matrix S is approximated by being SU(3) invariant, the initial and final states must be in the same SU(3) multiplet.

Thus there is an analog of the Wigner-Eckhart theorem which says the scattering amplitude is determined in terms of reduced amplitudes for the $\mathbf{27}$, for the $\mathbf{10}$, the $\mathbf{10}$, and the $\mathbf{1}$, together with four amplitudes $\mathbf{8}_{ss}$, $\mathbf{8}_{sa}$, $\mathbf{8}_{as}$, and $\mathbf{8}_{aa}$ which describe the four possibilities to combine initial states into a given state in the octet and then let it decompose into the final states.

There exists a classic paper Clebsch-Gordon coefficients for SU(3) is J. J. DeSwart, Rev. Mod. Phys. 35 (1963) 916.

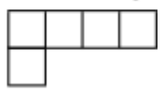
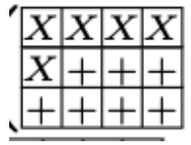
We will not go into this subject any further.

We have described the representations of SU(N) in terms of tensors made of quarks.

What are the representations made from antiquarks?

More generally, given a representation η, what is its conjugate representation η*?

The answer is found from the requirement that $\Gamma^{\eta} \times \Gamma^{\eta^*}$ contains the identity representation.

As we can lose boxes only by discarding columns of length N , we take our original graph η (say ), and extend it to a rectangle of height N , ( for SU(3)) , discard the original and rotate the rest by 180◦, to get $\eta*$ ( for SU(3)) .

Note that the $\eta*$ corresponding to each representation depends on N .

For SU(2), each representation is equivalent to its own conjugate, but this is not true for SU(N) for N > 2.

From this method, we see that  for SU(3) is the conjugate of the quark , so is the antiquark representation.

We also see that an initial column of N boxes doesn't affect η_*, so it also doesn't affect η.

Consider a state of three quarks.

Under color, the quarks are an SU(3) triplet.

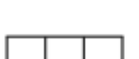
Let us consider only the "light flavors" u, d, and s, so they are also an SU(3) triplet under flavor.

Finally they are spin 1/2 fermions, so they are an SU(2) doublet in spin.

In the lowest bound state, the quarks all have the same spatial dependence, in the s state, $\ell = 0$.

The total wave function is therefore only dependent on color, flavor, and spin.

Fermi statistics requires that overall, the wave function must be antisymmetric.

Under flavor, we could have a decuplet  , an octet  , or a singlet  .

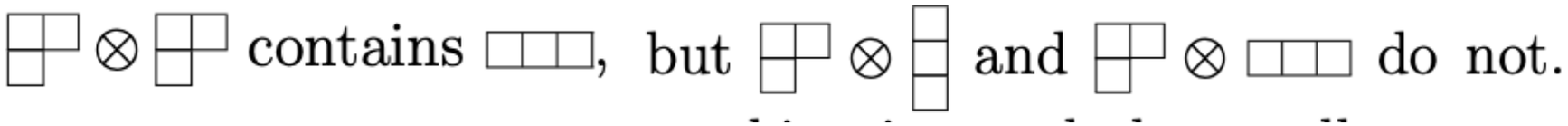
The same would be true for color, except that by "color confinement" only singlets under color can escape as particles.

Thus  is the only acceptable color representation, and is totally antisymmetric.

Finally, for SU(2),  doesn't exist, so there are only the J = 3/2 and the  J =1/2 possibilities.

This must be combined with the flavor representation to give total symmetry (as the color already provides the antisymmetry).

After some manipulation,



Of course the totally symmetric spin combination and the totally symmetric SU(3) combination have a totally symmetric composite.

Thus the quarks can form a decuplet of spin 3/2 particles or an octet of spin 1/2, but nothing else.

That this agrees with the lowest mass baryons was the basis of the quark model and the origin of the idea of color SU(3).

Without color, we would need flavor ⊗ spin to be antisymmetric.

This would give an **8** of J = 1/2 but a singlet of J = 3/2, which is phenomenologically disasterous.