Symmetries in Physics

J.Pearson

June 4, 2009

Abstract

These are a set of notes I have made, based on lectures given by M.Dasgupta at the University of Manchester Jan-May '09. Please e-mail me with any comments/corrections: jon@jpoffline.com. These notes may be found at www.jpoffline.com.

Contents

1 Introduction to Symmetries

We can state, very simply that a symmetry is something that is respected when we perform some operation upon an object, and the object is unchanged in some sense. That is, if something is invariant under reflection, if we reflect it, and the thing is still the same. This is rather vague, and will turn out to require very abstract mathematics to understand fully, but is a very powerful tool.

Another example of a symmetry, is the symmetry under interchange of the two hydrogen atoms of a water molecule. The "reason" the water molecule (for example) is symmetric

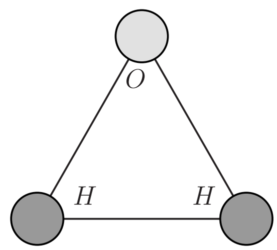


Figure 1.1: A water molecule is symmetric under interchange of the two hydrogen molecules.

under such an interchange, is that the potential is a function of "distances" only:

$$V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = V\left(|\mathbf{r}_1 - \mathbf{r}_2|, |\mathbf{r}_1 - \mathbf{r}_3|\right).$$

Such a symmetry translates into a set of very powerful mathematical constraints on the physical system.

Symmetries are intimately connected to the most fundamental laws of nature. For example, conservation of momentum is a result of some translational invariance within a system. Similarly, conservation of energy and angular momentum are consequences of time and rotational invariance.

Not all symmetries of a system are immediately obvious. For example, take a planetary orbit. A planet orbits in a spherically symmetric potential,

$$V \sim \frac{1}{r}.$$

Most planetary orbits are elliptical, and so dont seem to posses a spherical symmetry. However, the motion of a planet is confined to a plane, with the direction of angular momentum remaining "out of the plane", hence conserving angular momentum. Conservation of angular momentum is a consequence of the rotational invariance of motion in a spherically symmetric potential. Thus, even though the motion does not seem to outwardly posses a symmetry, the spherical symmetry is still there, but is somewhat "hidden" from view.

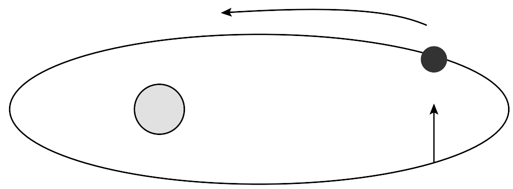


Figure 1.2: An elliptical orbit is not immediately spherically symmetric, until one considers the direction of angular momentum

Quantum field theory has the ideas of symmetry at its very heart; with the standard model unifying the strong, weak & EM forces, in a so-called symmetry gauge group,

$$SU(3) \times SU(2) \times U(1).$$

Such a consideration leads to the prediction of things like Baryon multiplets,

$$27 = 10 \oplus 8 \oplus 8 \oplus 1.$$

Infact, the idea of the existence of quarks stems from some symmetry consideration.

2 Group Theory

Before we start, let us just state some notation. The mathematical statement

$$\exists\, a : a \in G$$

should be read "there exists an element a, such that the element a is a part of the set G". The symbol $\forall$ should be read as "for all". We shall also use expressions such as

$$H = \{g \in G \,|\, g \circ a = 2\},$$

such should be read "the set H is the set of elements g, which are in the set G, such that g composed with a gives 2". We shall thus use both : and | to denote "such that". The usage of such notation simplifies expressions, but should be kept in mind what the "words are" behind the notation, to make an expression readable.

A group is a set of objects G that can be combined according to some combination law $\circ$, such that certain conditions are obeyed. The conditions are called the group axioms, and are:

- **G0: Closure**: for any $a, b \in G$, then the combination is always within the set G. That is,
$$\forall a, b \in G, a \circ b = c : c \in G.$$
- **G1: Identity**: there exists an identity element e in the set G, so that upon composition with any element $a \in G$, the element a is returned. That is,
$$\exists\, e \in G, a \in G : a \circ e = e \circ a = a.$$
- **G2: Inverse**: there exists an inverse element $a^{-1} \in G$, such that upon composition with $a \in G$, the identity element e is returned. That is,
$$\exists\, a^{-1} \in G, a \in G : a^{-1} \circ a = a \circ a^{-1} = e.$$
- **G3: Associativity**: for any $a, b, c \in G$, then
$$a \circ (b \circ c) = (a \circ b) \circ c$$
holds. Now, it is important to note that commutativity is not guaranteed, but, in the case that $a \circ b = b \circ a$, then G is said to be *Abelian*.

These fairly abstract sounding axioms must be checked, in order for the status of "group" to be ascribed to a set with composition rule. That is, suppose we have some *set* G, and composition rule $\circ$, then we must check that all four group axioms are satisfied before we call G a *group*. The members of the set G are called elements. Some examples will make this a little clearer.

2.1 Examples of Groups

2.1.1 Finite Groups

Finite groups are groups composed of a finite number of elements.

An example is the C_2 group, having two elements $\{e, a\}$, with composition rule "successive application". Then,

$$a \circ e = e \circ a = a, \quad a^2 = a \circ a = e.$$

Figure 2.1: Rotation of a directed line about π: this is the element a of the group C_2.

Now, from Figure (2.1), one could interpret the element a as being a rotation of a directed line through an angle π. Then, if we apply a to something that already has a applied to it, it has the effect of rotating the original object by 2π, hence rendering the object unchanged.

Another group is Z_n, which is the set of integers $0, \ldots, n-1$ with composition law being addition modulo n; that is,

$$Z_n \quad \Rightarrow \quad \{0, 1, 2, \ldots, n-1\}, \quad +_n.$$

The "modulo" bit of the composition rule means that after combining (addition in this case), if the result is greater than, or equal to, n, we can subtract off n enough times to bring it within the range (notice that without this, such a set of numbers would not be closed, and hence not a group). Such a group is Abelian, as addition is commutative. For example consider

$$Z_3 : \{0, 1, 2\}, +_3.$$

Then, it is clear that 0 is the identity element (add it to any element, and one finds the original element returned):

$$0 +_3 0 = 0, \quad 0 +_3 1 = 1, \quad 0 +_3 2 = 2.$$

So, let us see what happens if we combine element 1 with 2:

$$1 +_3 2 = 3 \operatorname{mod}(3) = 0 = e.$$

Hence, as $1 +_3 2 = 0 = e$, we see that 1 and 2 are inverses of each other. Table (2.1) allows us to immediately see that the set is closed, each element has an inverse, there exists an identity element, and the associativity holds. Hence, the set, under $+_3$, is a group. Also notice that

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Table 2.1: Multiplication table for the group Z_3, under addition modulo $+_3$.

the table is symmetric along the diagonal, which implies commutativity, which means that the group is Abelian.

Another example is S_n: the permutation group of n objects. So, for example, S_3 is the permutation group of the objects a, b, c (say). Now, suppose that the elements of the group are denoted as

$$a_3 = (1\,2), \quad a_4 = (2\,3), \quad a_5 = (1\,3), \quad a_1 = (1\,2\,3), \quad a_2 = (3\,2\,1), \quad e,$$

where e is the operation that returns the original object unaltered (i.e. the identity element). So, $a_3 = (1\,2)$ is the operation of swapping the first and second objects. Also, notice that

$$a_1 = (1\,2\,3) = (1\,2)(2\,3), \quad a_2 = (3\,2\,1) = (3\,2)(2\,1),$$

where we do the operation on the far right first. So, consider

$$a_3(a\,b\,c) = (1\,2)(a\,b\,c) = (b\,a\,c).$$

Also,

$$a_1(a\,b\,c) = (1\,2\,3)(a\,b\,c) = (1\,2)(2\,3)(a\,b\,c) = (1\,2)(a\,c\,b) = (c\,a\,b).$$

Consider acting a_3 twice,

$$a_3 a_3(a\,b\,c) = a_3(b\,a\,c) = (a\,b\,c).$$

Hence, the result of acting a_3 upon something that a_3 has already acted upon, has the result of returning the original object. Hence, $a_3^{-1} = a_3$ (i.e. it is self-inverse). Similarly,

$$(1\,2)^2 = e, \quad (1\,3)^2 = e \quad (2\,3)^2 = e.$$

We can also show that a_1 and a_2 are the inverses of each other. To see this, consider

$$\begin{aligned} a_1 a_2(a\,b\,c) &= (1\,2\,3)(3\,2\,1)(a\,b\,c) \\ &= (1\,2\,3)(3\,2)(2\,1)(a\,b\,c) \\ &= (1\,2\,3)(3\,2)(b\,a\,c) \\ &= (1\,2\,3)(b\,c\,a) \\ &= (1\,2)(2\,3)(b\,c\,a) \\ &= (1\,2)(b\,a\,c) \\ &= (a\,b\,c) \\ &= e. \end{aligned}$$

Going through everything is very tedious, but one can show that S_n is a group.

Now, C_3 (we just met C_2) can be represented as the symmetry group of an equilateral triangle, with directed sides; see Figure (2.2).

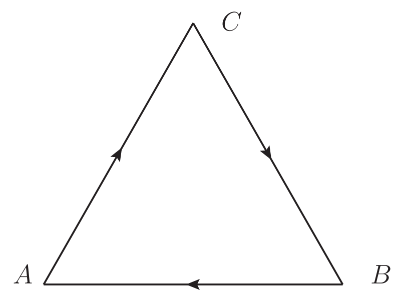


Figure 2.2: An equilateral triangle, with directed sides.

The equilateral triangle is symmetric under rotation through $120° = \frac{2\pi}{3}$ (about its centre). Hence, the group has 3 elements: the identity, this first rotation, and rotating again:

$$C_3 : \{e, c, c^2\}, \tag{2.1}$$

where c denotes rotation by $2\pi/3$ (and hence c^2 denotes rotation about $4\pi/3$), and e either no rotation or rotation by an integer multiple of 2π. So, let us consider the successive applications (i.e. successive rotations):

$$c^2 = c \circ c = 4\pi/3, \quad c^3 = c \circ c^2 = e(= 2\pi), \quad c^4 = c^2 \circ c^2 = c.$$

Notice that if we apply c 3 times (i.e. c^3), then we arrive at the identity element. Hence, we say that c is of order 3.

The **order** of an element is the power to which that element is raised (i.e. number of times it is applied) to obtain the identity element e. Every element in a finite group has a definite order.

The generalisation of the C_3 group is C_n, such that it is a group of n elements, which are represented by rotations of a regular n-gon (n-sided polygon); where rotations are through angles

$$\frac{2\pi}{n} r, \quad r = 0, 1, \ldots, n-1.$$

Hence, the elements of the group are

$$\{e, c, c^2, \ldots, c^{n-1}\},$$

where c is a rotation through $2\pi/n$. Thus, C_n is the symmetry group of a regular n-gon of n directed sides. Notice that the entire group is generated by the element c. That is, every element in the group may be obtained by applying a single rotation, enough times. Also note that the C_n group is isomorphic to Z_n.

An **isomorphism** means that there is a one-to-one relationship between the elements of two different groups. One notates such a relationship

$$G_1 \cong G_2.$$

Take our supposition that $Z_3 \cong C_3$. Then,

$$\{e, c, c^2\} \quad \cong \quad \{0, 1, 2\},$$

where the correspondence between the elements is

$$e \leftrightarrow 0, \quad c \leftrightarrow 1, \quad c^2 \leftrightarrow 2.$$

What we mean when we say that two groups are isomorphic, is that their elements play the same role. For example, combining the "same" elements (i.e. those we identified between the groups) gives the "same" result. Such as

$$c \circ c^2 = e, \quad 1 +_3 2 = 0, \quad c^2 \circ c^2 = c, \quad 2 +_3 2 = 1.$$

Infact, the isomorphism is between the general groups

$$C_n \cong Z_n \quad \Rightarrow \quad \{e, c, c^2, \ldots, c^{n-1}\} \cong \{0, 1, 2, \ldots, n-1\}.$$

One can think of a "reason" why the groups are isomorphic, is due to the way in which we get back into the range, upon combination.

A slightly more interesting example is the **dihedral group** D_n.

The group D_n is the symmetry group of the regular polygon, with n undirected sides. For example, D_3 is the group formed from symmetry operations performed on an equilateral triangle. With reference to Figure (2.3), we see that we can rotate anti-clockwise about O, through $2\pi/3$ (as we could in the directed sides case), in which case each application is denoted c. In addition, we are able to flip about the OX, OY and OZ axes (i.e. rotate about these axes, through an angle π). So, we have the extra 3 operations (elements of the group):

$$\begin{aligned} b_1 &: \quad \text{rotate about } OX \text{ by } \pi, \\ b_2 &: \quad \text{rotate about } OY \text{ by } \pi, \\ b_3 &: \quad \text{rotate about } OZ \text{ by } \pi. \end{aligned}$$

We think about the OX, OY, OZ axes as being fixed in space. Now, one can show (and indeed, we shall show) that $b_2 = cb_1c^{-1}$ – something we will later call an equivalence relation.

Now, suppose we notate the vertices (A, B, C) by $(1\,2\,3)$; i.e. fix the positions, so that the rotations described above can be written as

$$c = (2\,3\,1), \quad b_1 = (3\,2), \quad b_2 = (1\,3), \quad b_3 = (1\,2), \quad c^{-1} = (3\,2\,1).$$

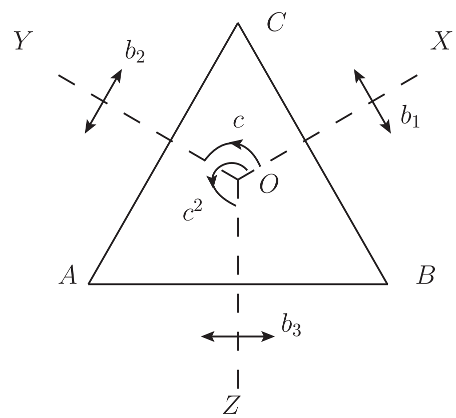


Figure 2.3: An equilateral triangle, with undirected sides. Note the definitions of the axes and vertices, as in D_3.

Hence, we see that written in this way, we have that b_2 (for example) swaps things in first and third places, and leaves the thing in the second place untouched. Hence,

$$b_2(A\,B\,C) = (1\,3)(A\,B\,C) = (C\,B\,A).$$

Also,

$$\begin{aligned} cb_1c^{-1}(A\,B\,C) &= cb_1(3\,2)(2\,1)(A\,B\,C) \\ &= cb_1(B\,C\,A) \\ &= c(3\,2)(B\,C\,A) \\ &= c(B\,A\,C) \\ &= (2\,3\,1)(B\,A\,C) \\ &= (C\,B\,A) \\ &= b_2(A\,B\,C) \end{aligned}$$

Hence, we see that the result of applying b_2 to the object, is the same as applying cb_1c^{-1} to the object. Therefore, we have an equivalence relation

$$b_2 = cb_1c^{-1}.$$

One may be able to start to see why we call this an equivalence relation – a rotation by π is a rotation by π, no matter where ones origin is. Hence, we should be able to show that b_2 and b_1 are equivalent (as they correspond to the same operation, only starting in a different place), which is indeed what we have done – although, we shall discuss equivalence in more detail, later. In a similar way, let us compute

$$b_2(A\,B\,C) = (1\,3)(A\,B\,C) = (C\,B\,A).$$

Also,

$$b_1c(A\,B\,C) = b_1(C\,A\,B) = (C\,B\,A).$$

Hence, we have another relation:

$$b_2 = b_1c.$$

Therefore, we see that elements of the group are related to each other. From hereon, we shall denote b_1 simply by b. So, we have derived that

$$b_2 = cbc^{-1}, \quad b_2 = bc.$$

Hence, equating the two,

$$cbc^{-1} = bc, \tag{2.2}$$

and multiplying by c from the right (noting that $c^{-1}c = e$), we get

$$cb = bc^2.$$

One can also show that

$$b_3 = c^{-1}bc.$$

Now, using (2.2) for bc, in this,

$$b_3 = c^{-1}bc = c^{-1}cbc^{-1} = bc^{-1}.$$

Now, as c is a rotation anti-clockwise by $2\pi/3$, then c^{-1} is a rotation by $2\pi/3$ clockwise, which is equivalent to rotating by $c^2 = 4\pi/3$ anti-clockwise. Thus, $c^{-1} = c^2$, and so

$$b_3 = bc^2.$$

Hence, we see that we have the relations

$$b_2 = bc, \quad b_3 = bc^2,$$

that is, the rotations about axes in the plane of the triangle can be generated by rotating perpendicular to the plane (i.e. c) and rotating about one of the axes (i.e. b). Hence, the elements of the dihedral group D_3 can be written

$$D_3 : \{e, c, c^2, b, bc, bc^2\}.$$

Thus, we see that D_3 is generated by just two elements. In general, the group D_n can be written

$$D_n : \{e, c, c^2, \ldots c^{n-1}, b, bc, bc^2, \ldots, bc^{n-1}\}.$$

One can notice that C_3 (2.1) is a subgroup of D_3 (indeed, C_n is a subgroup of D_n) – that is, the first half of the elements of D_n are exactly those elements of C_n.

A **subgroup** H of a group G is a set of elements $\{h\} \in G$, which is closed under the group composition law. That is, a subgroup is some subset of G that form a group on their own.

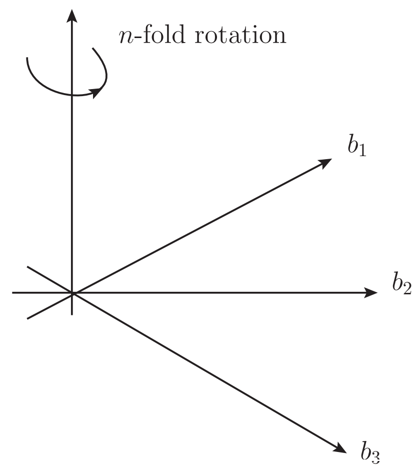


Figure 2.4: A representation of the elements of the dihedral group D_n. The elements c represent a rotation through $2\pi/n$ about an axis perpendicular to the figure; the b_i elements represent a rotation about π through axes in the plane of the figure.

2.2 Properties of Groups & Morphisms Between Groups

2.2.1 Conjugacy & Equivalence

We can establish equivalence relations between elements of groups. Note, two things that are equivalent are not necessarily equal.

We have already met conjugacy, in connection with the D_3 group. Recall we derived

$$b_2 = cbc^{-1}.$$

Hence, we say that elements b_2 and b are conjugates, with conjugating element c. Similarly, we stated $b_3 = c^{-1}bc$, in which case b_3 and b are conjugates, with conjugating element c.

Two elements $a, b \in G$ are said to be **conjugate** if $a = gbg^{-1}$, where $g \in G$. The conjugating element need not be unique.

For D_3, c and c^2 are also conjugate. To see this, recall (2.2), $cbc^{-1} = bc$. Then, multiplying from the right by c, we have $cb = bc^2$, and further multiplying from the right by b^{-1}, to give

$$c = bc^2b^{-1}.$$

Hence, c and c^2 are conjugates, with conjugating element b. Conjugacy is a special case of an equivalence relation.

A relation $\sim$ between two elements $a, b \in G$ is said to be an **equivalence relation** if it respects the following conditions:

- **Reflexive**: Every element should be equivalent to itself $a \sim a$.
- **Symmetry**: If $a \sim b$ then $b \sim a$.
- **Transitive**: If $a \sim b$ and $b \sim c$, then $a \sim c$.

We can check that conjugacy is an equivalence relation. So, let us perform the three checks required.

We require that a must be conjugate to a. So, let us have that $a = gag^{-1}$ where $a, g \in G$. Then, in order for this to hold, we need to find an element g that makes this statement true. Then, if we choose the conjugating element $g = e$, then $a = eae^{-1} = a$, which is indeed true. Thus, an element is conjugate to itself.

We next need to check that if $a \sim b$ then $b \sim a$. So, we have $a = gbg^{-1}$. Then, if we let $b = g^{-1}ag$, then $a = gbg^{-1} = gg^{-1}agg^{-1} = a$, which is true. Hence, we see that if a is conjugate to b, with conjugating element g, then b is conjugate to a with conjugating element g^{-1}.

Finally, we need to check that if $a \sim b$ and $b \sim c$, then $a \sim c$. So, suppose we have $a = g_1bg_1^{-1}$, and $b = g_2cg_2^{-1}$, where $g_1, g_2 \in G$ (that is, a and b are conjugates, with conjugating element g_1, and b and c are conjugates, with conjugating element g_2). Hence, $a = g_1bg_1^{-1} = g_1g_2cg_2^{-1}g_1^{-1} = g_1g_2c(g_1g_2)^{-1}$. Thus, as G is closed, and if $g_1, g_2 \in G$, then $g_1g_2 \in G$. Hence, a and c are conjugates, with conjugating element $g_1g_2 \in G$, by closure.

Hence, we see that upon comparison of our results for the D_3 group, then b, b_2, b_3 are equivalent. This is a specific example of the more general statement that rotations through the same angle must be equivalent.

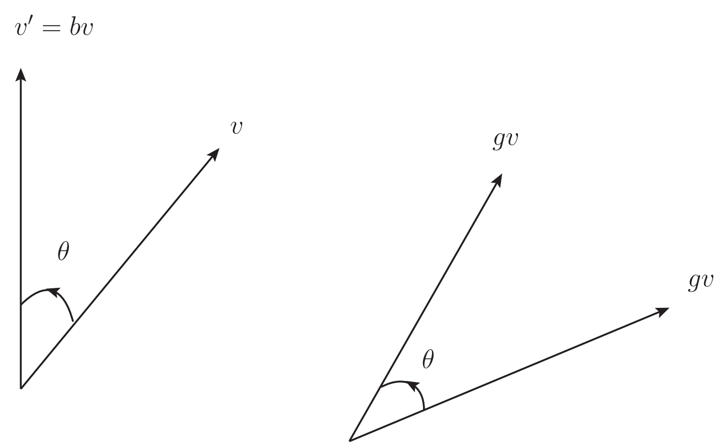


Figure 2.5: A rotation through θ in one place is equivalent to a rotation through θ in another place

Consider a vector v, and that $v' = bv$ is some rotated vector (i.e. b is some rotation operator). Now, suppose that another observer sees the vector gv (i.e. some transformation

g between the observers). Then, the rotated vector, according to the observer is gv'. Then,

$$gv' = gbv = (gbg^{-1})gv.$$

Hence, we see that the operator for rotating, for the observer is the bracketed quantity. This actually makes sense. The transformation g^{-1} will send the observer back to where the original was, then b will rotate the vector, then g will transform the observer back to his original (transformed) position. See Figure (2.5).

There is a caveat to this. If we are talking about the full rotation group through any angle, then rotations through a given angle are equivalent. But, this does not apply to a subgroup of the full rotation group. Take D_4 as an example. Let b_1 and b_2 be rotations by π through an axis through a side and vertex, respectively. They are not equivalent operations, as no member of D_4 can take one into the other.

2.2.2 Conjugacy Classes

These are a special case of equivalence classes.

Any equivalence relation partitions a group into disjoint set of elements, and those sets are called **equivalence classes**.

It is essentially a decomposition of a group. Thus, a conjugacy class of an element a, written (a) is the set of elements that are conjugate to a. We denote this as

$$(a) = \{b\,|b = gag^{-1} : b, g \in G\}. \tag{2.3}$$

This notation should be read "the conjugacy class of a is the set of elements b, such that b is conjugate to a, with conjugating element g, where b and g are in G". Let us consider some examples.

The C_n group,

$$C_n : \{e, c, c^2, \ldots, c^{n-1}\}$$

is Abelian, as $ag = ga$, $\forall a, g, \in C_n$. Hence, $gag^{-1} = agg^{-1} = a$ (which is true for any Abelian group). Hence, all elements are conjugate to itself, and no others. Hence, every element of C_n is in its own conjugacy class. So, the partitioning of C_n into its conjugacy classes is just

$$(e), \quad (c), \quad (c^2), \quad \ldots, \quad (c^{n-1}).$$

We can partition D_3 into conjugacy classes,

$$(e), \quad (c, c^2), \quad (b, bc, bc^2), \tag{2.4}$$

where each element of the set is conjugate to every other in that set (as we have shown).

Now, no distinct equivalence classes have common elements. To prove this, consider the equivalence classes of a and b; $(a), (b)$. Let us suppose that they have an element c in common.

Then, $a \sim c$ and $b \sim c$. Hence, $a \sim b$, and thus b is in the equivalence class of a. Hence, if two classes have common elements, then the classes are within the other, hence the classes are not distinct. Therefore, different classes have no common elements.

2.2.3 Subgroups

A subset H of a group G is a subgroup of G if the elements of H form a group. For finite groups, in order to ascertain that H is a subgroup, one only needs to check closure (remaining properties follow from those of G).

Suppose that $h_1 \in H$. Then, suppose that element h_1 has order r. Then, $h_1^r = e$. But, $h_1^r \in H$ and hence $e \in H$, by closure. Hence, closure of a subset automatically has the existence of the identity e in the subset. Similarly, every element has an inverse in H. Now, if $h_1 \in H$, where $h_1^r = e$, then $h_1^{r-1} \in H$. Thus, $h_1 h_1^{r-1} = h_1^r = e \in H$. Therefore, every element has an inverse in H, by closure, with the inverse of h_1 being h_1^{r-1}.

A **proper subgroup** is a subgroup that is not just the identity or the whole group itself.

For example, the D_3 group

$$D_3 : \{e, c, c^2, b, bc, bc^2\}$$

has two proper subgroups

$$C_3 : \{e, c, c^2\}, \quad C_2 : \{e, b\}.$$

2.2.4 Cosets

Consider a subgroup H of a group G, then one can form cosets $g_1 H, g_2, \ldots, g_s H$ such that

$$g_1 H : \{g_1 h_1, g_1 h_2, \ldots, g_1 h_r\},$$

where

$$H : \{h_1, h_2, \ldots h_r\}, \quad g_1 \in G.$$

Technically, we have formed the left coset of the subgroup H. The right coset is, for example, Hg_1. We say that $g_2 H$ is the coset of g_2 (where H is a subgroup of G, and $g_2 \in G$). Basically, to form a coset, we take any element of the group G, and multiply it from the left, onto every element of the subgroup H of G.

There is an equivalence relation between an element and the members of its coset. That is, $a \sim b$ if b is in the coset of a. Let us prove it.

First, the reflexive property. We require $a \sim a$, that is, every element is its own coset. If $e \in H$, and $a \in G$ (where H a subgroup of G), then $ae = a$. Hence, a in its own coset.

Second, the symmetry property. Let a be in the coset of b. Then $a = bh_k$, where $h_k \in H$. We also require that $b = ah_k^{-1}$, where $h_k^{-1} \in H$. So, $a = bh_k = ah_k^{-1} h_k = a$, which is true. Therefore, if a is in the coset of b, then b is in the coset of a.

Finally, the transitive property. Let a be in the coset of b, so that $a = bh_r$, where $h_r \in H$. Also, let b be in the coset of c, so that $b = ch_k$, where $h_k \in H$. Then, we want a to be in the coset of c. Hence, $a = ch_k h_r$, where $h_k h_r \in H$, by closure. Thus, $a = ch_n$ where $h_n \in H$. Therefore, a is in the coset of c, which is what we require.

Hence, we see that elements of a coset are equivalent.

Let us consider some examples.

The dihedral group

$$D_3 : \{e, c, c^2, b, bc, bc^2\}$$

has a subgroup

$$C_2 : \{e, b\} : H.$$

Let us then form the coset of all elements of D_3. So,

$$\begin{aligned} eH &: e\{e, b\} = \{e, b\}, \\ cH &: c\{e, b\} = \{c, cb\} = \{c, bc^2\}, \\ c^2H &: \{c^2, c^2b\} = \{c^2, bc\}. \end{aligned}$$

What we see here, is that a coset can be labelled by any of its elements (as the elements are equivalent).

Consider a coset we call gH (i.e. the coset of g),

$$gH : \{gh_1, \ldots, gh_r\}, \quad h_i \in H.$$

Let us pick any of the elements in the coset gH, say gh_k, and form its coset, gh_kH (where $gh_k \in gH$). But, $h_kH \in H$ by closure. Therefore, $gh_kH = gH$.

2.2.5 Normal Subgroups

We start with the definition that a **normal subgroup** are those subgroups H, of a group G, for which

$$gHg^{-1} = H. \tag{2.5}$$

An alternative way of notating this, which is perhaps more transparent, is

$$gh_ig^{-1} = h_j, \quad \forall g \in G, \quad \forall h_i, h_j \in H.$$

This definition essentially says that the left and right cosets are equal, so that $gH = Hg$. Suppose we have

$$H = \{h_1, h_2, \ldots, h_\ell\},$$

then (2.5) does not have to hold on an element-by-element basis. That is, gh_kg^{-1}, where $h_k \in H$, need not be equal to h_k, but does have to be equal to some element $h_i \in H$.

The general idea behind a normal subgroup, is that if we have a subgroup H (which is a subgroup of G), and if we take all elements $g \in G$, then by forming gHg^{-1}, we regenerate that subgroup H, if H is a normal subgroup. That is, if this relation holds, then H is a normal subgroup.

Example: $G = D_3$ Recall the D_3 group,

$$D_3 : \{e, c, c^2, b, bc, bc^2\},$$

where c is a rotation by $2\pi/3$. Then, let us consider the subgroup

$$C_2 : \{e, b\}.$$

So, let us take a "random element" of D_3, c say, and form cbc^{-1}, where b is an element of the subgroup C_2. Now, in our previous discussions (2.2), we showed that $cbc^{-1} = bc$. But, we see that $bc \notin C_2$, and therefore C_2 cannot be a normal subgroup of D_3. Let us consider another subgroup of D_3,

$$C_3 : \{e, c, c^2\}.$$

Recall that c and c^2 were in their own conjugacy classes, which meant that for any $g \in D_3$, one could form gcg^{-1} and gc^2g^{-1} and remain in the conjugacy class. Therefore, including the identity element e trivially keeps this closure. Therefore, we see that C_3 is a normal subgroup of D_3.

2.2.6 Quotient Group

For a normal subgroup H, the set of cosets $\{g_1H, g_2H, \ldots, g_sH\}$, where $g \in G$, which is denoted G/H, forms a group under suitable multiplication law – this group is called the quotient group. Such a multiplication law is

$$(g_1H) \circ (g_2H) = (g_1 \circ g_2)H. \tag{2.6}$$

Let us confirm that G/H is infact a group. So, we need to check the validity of the group axioms. So that we dont have to keep stating it, we have $g_1, g_2 \in G$, where G is a group, and H is a normal subgroup of G.

So, closure. If we compose two elements of G/H, using the above rule,

$$(g_1H) \circ (g_2H) = (g_1 \circ g_2)H = g_3H,$$

where $g_1 \circ g_2 \in G$, but as G is a group, G is closed, and hence $g_3 \in G$, and therefore $g_3H \in G/H$. We require an identity element. Consider that the identity of G is e; hence,

$$(eH) \circ (gH) = (g \circ e)H = gH, \quad (gH) \circ (eH) = (e \circ g)H = gH.$$

Therefore, we see that eH is the identity element of G/H. We also need to check that there is an inverse element, $\forall gH \in G/H, \exists g^{-1}H \in G/H$. So,

$$(gH) \circ (g^{-1}H) = (g \circ g^{-1})H = eH \in G/H,$$

which follows as the inverse of g, in G, is g^{-1}. Hence, the inverse of $gH \in G/H$ is $g^{-1}H \in G/H$. The associative law is inherited from the group G.

Now, we still need to check that our multiplication law is unique, and for this, we shall need to use that H is a normal subgroup of G. Consider forming a coset

$$g_aH : \{g_ah_a, g_ah_2, \dots g_ah_r\} = g_ah_kH, \quad h_k \in H, g_a \in G.$$

which follows from our discussion at the end of §2.2.4. In a similar way,

$$g_bH = g_bh_lH, \quad h_l \in H, g_b \in G.$$

Then, g_aH and g_bH will be elements of the quotient group G/H. So, their composition,

$$\begin{aligned}(g_aH) \circ (g_bH) &= (g_a \circ g_b)H \\ &= (g_ag_b)H \\ &= (g_ah_kg_bh_l)H.\end{aligned}$$

Now, by closure of the subgroup H, $h_lH = H$. Therefore,

$$(g_ah_kg_bh_l)H = (g_ah_kg_b)H.$$

We now use the fact that H is a normal subgroup, which lets us say that $g_bH = Hg_b$, which means that we write the above as

$$(g_ah_kg_b)H = g_ah_kHg_b = g_aHg_b = g_ag_bH.$$

Therefore, we have found that the multiplication law is unique.

For a normal subgroup, the multiplication of two costes $g_aH \circ g_bH$ turns out to be just element-by-element multiplication.

Let us consider a quotient group G/A, and assume that it is isomorphic to a group B,

$$G/A \cong B.$$

We may ask then, can we write $G \cong A \times B$? It turns out that this is sometimes possible, but not always. Infact, if it is true that $G \cong A \times B$, then one does indeed have that $G/A \cong B$; however, one is not guaranteed to have the converse true.

2.2.7 Direct Products of Groups

A group G can be expressed as a direct product of its subgroups A and B if, and only if,

- All elements of B commute with those in A,
- Every element $g \in G$ can be expressed in a unique way, as $g = a_i b_j$ with $a_i \in A$ and $b_j \in B$.

As a consequence of this, A and B are normal subgroups of G. Let us prove this consequence.

Consider gag^{-1}, where $a \in A$. Now, $g = a_i b_j$, then,

$$gag^{-1} = a_i b_j a b_j^{-1} a_i^{-1}.$$

But, elements of A commute with those in B, so that

$$a_i b_j a b_j^{-1} a_i^{-1} = a_i a b_j b_j^{-1} a_i^{-1} = a_i a a_i^{-1} \in A,$$

by closure. Therefore,

$$gag^{-1} = a_i a a_i^{-1} \in A,$$

which is the statement that A is a normal subgroup of G. The proof is analogous for B.

It is always true that if one has

$$G \cong A \times B,$$

then

$$G/A \cong B.$$

To prove this, consider that G/A is the set of cosets $g_i A$, where $g_i \in G$. Now, g must be of the form $g = b_j a_k$, where $b_j \in B, a_k \in A$. So, $g_i A = b_j a_k A = b_j A$, by closure. So, every coset in the quotient group G/A is identified with some distinct element b_j (say) of B. These cosets faithfully follow the multiplication law of the group B, because

$$(b_j A) \circ (b_k A) = (b_j \circ b_k) A,$$

and if $b_j b_k = b_l \in B$, then $(b_j A) \circ (b_k A) = b_l A$. Therefore, a one-to-one correspondence is preserved after multiplication. Thus, an isomorphism.

Exercise Let us show that $C_6 \cong C_3 \times C_2$, and comment upon C_6/C_3.

Now, the group is

$$C_6 : \{e, c, c^2, c^3, c^4, c^5\},$$

where c is a rotation by $2\pi/6 = \pi/3$. The subgroups are

$$C_2 : \{e, c^3\}, \quad C_3 : \{e, c^2, c^4\}.$$

Notice that C_2 has elements the identity e and rotation $c^3 = 3 \times \pi/3 = \pi$, as expected. It is fairly obvious that all elements commute, and one can check that all elements of C_6 can be uniquely expressed as an element of C_2 times C_3:

$$\begin{aligned}
e &= e \times e, \\
c &= c^4 \times c^3 = c^7 = c, \\
c^2 &= c^2 \times e, \\
c^3 &= e \times c^3, \\
c^4 &= c^4 \times e, \\
c^5 &= c^2 \times c^3.
\end{aligned}$$

Hence, we have shown that $C_6 \cong C_3 \times C_2$. Now, we have to show that $C_6/C_3 \cong C_2$. Then, let us take the subgroup $H = C_3$, and find the set of cosets,

$$\begin{aligned}
eH &= \{e, c^2, c^4\} = H \equiv E, \\
cH &= \{c, c^3, c^5\} \equiv B, \\
c^2H &= \{c^2, c^4, e\} = E, \\
c^3H &= \{c^3, c^5, c\} = B, \\
c^4H &= \{c^4, e, c^2\} = E, \\
c^5H &= \{c^5, c, c^3\} = B.
\end{aligned}$$

Therefore, the quotient group, which is the set of cosets of the subgroup C_3 is

$$C_6/C_3 = \{E, B\}.$$

Notice that $B^2 = BB = cHcH = c^2H = E$, and $EB = eHcH = ecH = cH = B$. Therefore, we see that the quotient group C_6/C_3 is isomorphic to C_2 (notice that the elements of C_6/C_3 are not identical to those of C_2, but they act in the same way – which is an isomorphism).

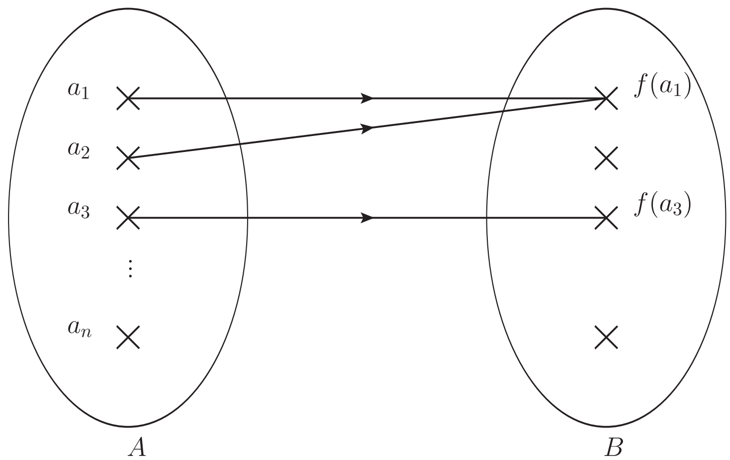
This exercise has provided us with our first example illustrating the main point of group theory. In the exercise, we *identified* the set of objects $\{e, c^2, c^4\}$ as a single object, which we labelled E; and the set $\{c, c^3, c^5\}$ as the single object B. We then went on to see that no matter which of the elements of the set B, say, that we used in a composition with either another element of B or of E, the result is the same as if we had just used E or B themselves, under the same rules as the elements of C_2. We could also say that the elements of the set E or B form a triplet: there is a kind of internal symmetry between all elements within the set. Hence, we also see why we say that C_6/C_3 is isomorphic to C_2, and not equal to; the result is not identical to C_2, but everything works in the exact same way as the elements of C_2.

2.3 Homomorphisms

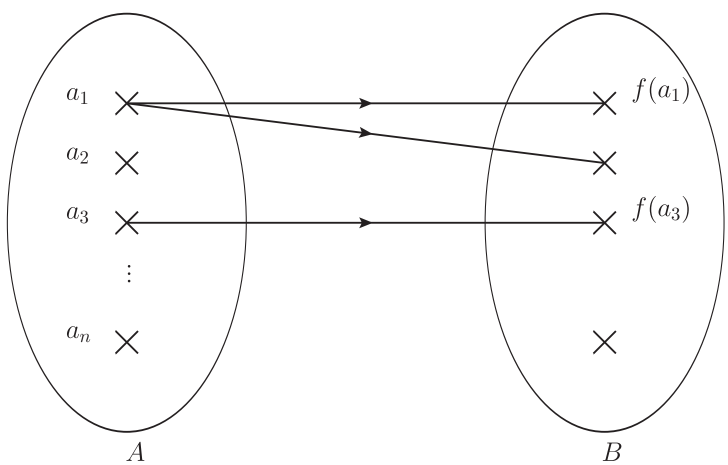
A homomorphism between groups is a mapping such that group multiplication is preserved. A mapping is denoted

$$f : A \longmapsto B,$$

where we say that the mapping f takes an element from A and uses some rule to turn it into some element in B.



(a) A valid homomorphism



(b) An invalid homomorphism

Figure 2.6: Schematic of a homomorphism between two groups A and B.

Figure (2.6)a depicts a valid homomorphism where two (or more) elements of a group A can be mapped into the same element of B, but an element of A cannot be mapped into two different elements of B, as in Figure (2.6)b. Elements in B need not be the image of any element of A.

Group multiplication is preserved,

$$f(a_1 \circ a_2) = f(a_1) \star f(a_2).$$

This notation means that $f(a_i)$ is the image in B, of the element $a_i \in A$. We have that elements in A have multiplication law $\circ$, and elements in B the law $\star$.

The set of elements in B, that are images of elements in the group A, form a subgroup of B. Closure can be checked, for example, by noting that $f(a_i) \star f(a_j)$ must be the image of some other element in A, as $f(a_i) \star f(a_j) = f(a_i \circ a_j)$, but A is closed, so that $a_i \circ a_j \in A$.

2.3.1 The Kernel of a Mapping or Homomorphism

The Kernel of a mapping, or homomorphism, is defined to be the set of elements in A that are mapped into the identity element e_B of group B.

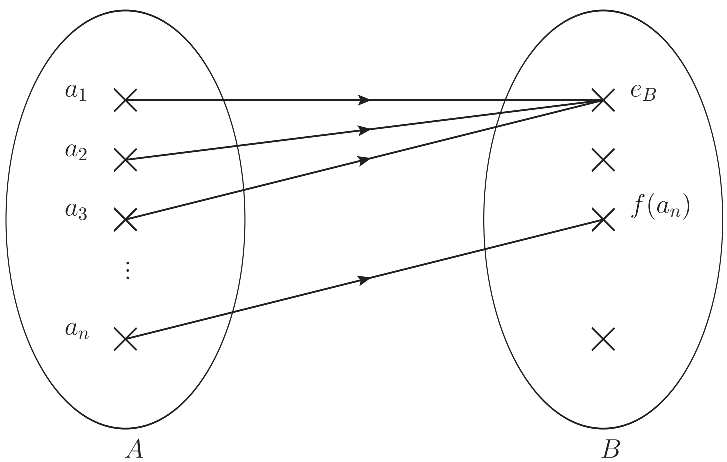


Figure 2.7: Schematic of the Kernel of a mapping from A to B.

One denotes the Kernel of a mapping f as Ker$f$, and is the set

$$\text{Ker}f = \{a \in A \,|\, f(a) = e_B\} \tag{2.7}$$

The Kernel infact forms a normal subgroup of A, which we shall prove.

First, closure of the Kernel. Consider $a_i, a_j \in \text{Ker}f$. Then, $f(a_i a_j) = f(a_i)f(a_j)$. But, as $a_i, a_j \in \text{Ker}f$, $f(a_i) = f(a_j) = e_B$. Therefore, $f(a_i a_j) = e_B e_B = e_B$, and thus the Kernel is closed.

The existence of the identity of A in the Kernel is proved by considering $f(ae_A) = f(a)f(e_A) =$

e_B, as $ae_A = a$ and $f(a) = e_B$, as $a \in \text{Ker}f$. Therefore, $f(e_A) = e_B$, and hence $e_A \in \text{Ker}f$. The inverse requires that $\forall a \in \text{Ker}f, \exists a^{-1} \in \text{Ker}f$. So, consider $f(aa^{-1}) = f(e_A) = e_B$, and that $f(aa^{-1}) = f(a)f(a^{-1})$ is therefore equal to e_B. But, as $a \in \text{Ker}f$, then $f(a) = e_B$ and hence $f(a^{-1}) = e_B$, and thus $a^{-1} \in \text{Ker}f$.

Finally, to prove that $\text{Ker}f$ is a normal subgroup, consider $k \in \text{Ker}f$, so that $f(k) = e_B$. Then, we must have that $aka^{-1} \in \text{Ker}f$, $\forall a \in A$. So,

$$\begin{aligned} f(aka^{-1}) &= f(a)f(k)f(a^{-1}) \\ &= f(a)f(a^{-1}) \\ &= f(aa^{-1}) \\ &= f(e_A) \\ &= e_B. \end{aligned}$$

Therefore, $aka^{-1} \in \text{Ker}f$, and hence the Kernel is a normal subgroup.

3 Representations of Groups

We can move from having abstract elements of groups, to having some sort of concrete realisation of group elements, in terms of physical operations. We start by giving an example, and then will give a more concrete definition of a representation.

Consider the rotation of a vector. The coordinates (x, y, z) map onto the spherical polar system (r, θ, ϕ) via

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta. \tag{3.1}$$

Let us consider the rotation of a vector about the z-axis, through an angle β,

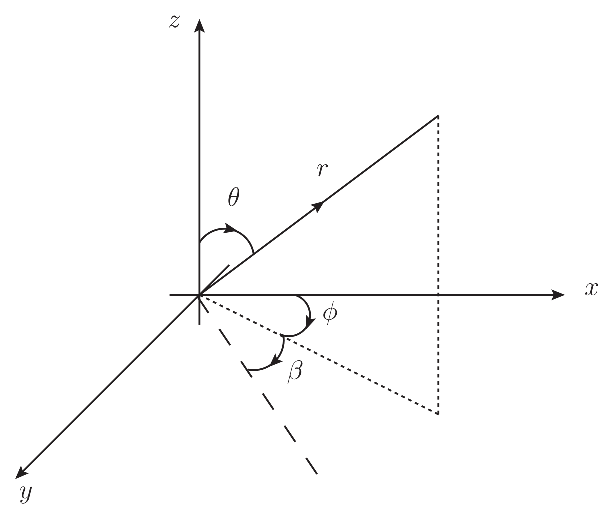


Figure 3.1: Rotating a vector.

$$\phi \longmapsto \phi + \beta.$$

So, using this in (3.1),

$$\begin{aligned} x \longmapsto x' &= r\sin\theta\cos(\phi+\beta) \\ &= r\sin\theta\cos\phi\cos\beta - r\sin\theta\sin\phi\sin\beta \\ &= x\cos\beta - y\sin\beta, \\ y' &= y\cos\beta + x\sin\beta, \\ z' &= z. \end{aligned}$$

Hence, we see that we can write this transformation as a matrix equation,

$$\mathbf{x}' = R(\beta)\mathbf{x},$$

where

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad R(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.2}$$

Notice that the new coordinates are linear superpositions of the old coordinates. We can now represent the elements of the C_3 group by matrices. That is, we can express e, c, c^2 which are rotations through $0, 2\pi/3, 4\pi/3$, using $R(\beta)$:

$$\begin{aligned} e &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ c \longmapsto D(c) = R(2\pi/3) &= \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ c^2 \longmapsto D(c^2) = R(4\pi/3) &= \begin{pmatrix} -1/2 & \sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

The matrices should multiply like the group elements; so that one can verify

$$D(c)D(c) = D(c^2), \quad D(c)D(c^2) = e, \quad D(c^2)D(c^2) = D(c).$$

Also, this preserves group multiplication,

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

Let us define a representation formally.

A **representation** of dimension $[n]$ of the abstract group G is defined as a homomorphism $G \to GL(n, \mathbb{C})$, to the group of non-singular $n \times n$ matrices with complex entries.

Notice that as it is a homomorphism, the group multiplication

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

is preserved. The matrices must be non-singular, due to the essential existence of the inverse,

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

Rotation matrices are orthogonal, as they preserve the length of the vector they rotate. Consider

$$x \longmapsto x' = Rx,$$

and that the length of the vector x is given by $x^T x$. Now, $x'^T = x^T R^T$. So,

$$x'^T x' = x^T R^T R x = x^T x,$$

as required, provided

$$R^T R = 1 \quad \Rightarrow \quad R^{-1} = R^T.$$

Therefore, R is orthogonal.

Example Let us write the matrices representing the elements of the finite group D_3, which is the group of symmetry operations on an equilateral triangle having undirected sides. Now, the elements of the group are

$$D_3 : \{e, c, c^2, b, bc, bc^2\},$$

which can be noted from Figure (2.3). The element c denotes rotation by an angle $\theta = 2\pi/3$, about the z-axis, and b a reflection in the 2D plane (hence, this allows us to see why $b^2 = bc$: first rotate by $2\pi/3$, then reflect). We can use the 2D version of the matrix (3.2) to see that

$$\begin{aligned} D(c) &= R(2\pi/3) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \\ D(c^2) &= R(4\pi/3) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}. \end{aligned}$$

Before we proceed, it is worth noting that these two matrices are transposes of each other. Also, recall that $c \circ c^2 = e$, so that c and c^2 are inverses. Hence, that the two matrices representing these elements are merely transposes of each other should not be a surprise: the inverse of an orthogonal matrix (i.e. distance preserving) is just its transpose: $M^{-1} = M^T$. Therefore, we should expect the matrices representing mutually inverse elements to be transposes of each other. The matrix representing b can be seen to be just

$$D(b) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where we are at liberty to choose the reflection to be about the x-axis. The representations of the other elements are easily computed

$$\begin{aligned} D(bc) &= D(b)D(c) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}, \\ D(bc^2) &= D(b)D(c^2) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}. \end{aligned}$$

Hence, we have representations of the rotating elements. The final element of the group that needs representing, is the identity element. This is just the unit matrix,

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

Now, if we compute the traces (i.e. the characters) of each of these 6 matrices, we find

$$\chi_e = 2, \quad \chi_c = \chi_{c^2} = -1, \quad \chi_b = \chi_{bc} = \chi_{bc^2} = 0.$$

Hence, we see that matrices representing elements in the same conjugacy class have the same character; in accord with the classes in (2.4).

3.1 Induced Transformation of the Wavefunction in Quantum Mechanics

Let us consider how to start to use this rather abstract group formalism.

Consider the electron wavefunction of a $2p$ state, whereby

$$|\psi|^2 \propto \cos^2\theta.$$

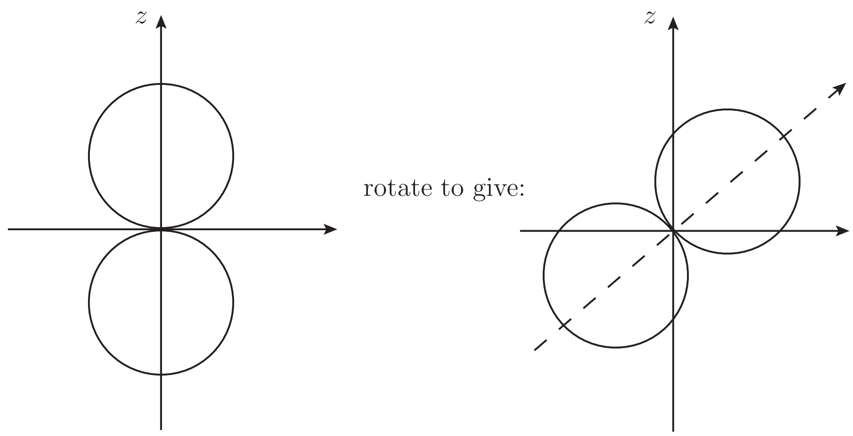


Figure 3.2: Rotating a $2p$ electron wavefunction.

After rotation, the wavefunction is concentrated about a new axis. So, the new wavefunction is ψ', but, the value of the new wavefunction, in the new coordinate system, is the same as that of the old wavefunction, in the old coordinate system,

$$\psi'(x') = \psi(x),$$

where $x' = Rx$, and hence, $x = R^{-1}x'$. Therefore,

$$\psi'(x') = \psi(R^{-1}x'),$$

dropping the dummy-prime on the coordinates,

$$\psi'(x) = \psi(R^{-1}x). \tag{3.3}$$

Recall that the Hamiltonian for the hydrogen atom is of the form

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(r),$$

and eigenfunctions are $U_{n\ell m}(r,\theta,\phi)$, where the quantum numbers n,ℓ and m correspond to the energy, orbital angular momentum and projection respectively;

$$E_n = \frac{E_1}{n^2}, \quad L^2 = \ell(\ell+1)\hbar^2, \quad L_z = m\hbar, \quad |m| \leq \ell.$$

Just as in the rotation of a vector case, the new $U'_{n\ell m}$ will be a superposition of the old $U_{n\ell m}$,

$$U'_{n\ell m}(x) = \sum_{m'} D_{m'm} U_{n\ell m'}(x).$$

As m lies between $-\ell$ and ℓ, there are $2\ell+1$ values of m for a given ℓ. Therefore, $D_{m'm}$ is a $(2\ell+1)\times(2\ell+1)$ matrix, and it will be a representation of the rotation group, just as

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

was. We can show that

$$U'_{n\ell m}(x) = \sum_{m'} D_{m'm}(R) U_{n\ell m'}(x)$$

is a representation by considering successive rotations R_1 and R_2. Let us consider a concrete example.

Consider the $n=2, \ell=1$ wavefunction. So, we have $m=-1,0,1$, and $U_{n\ell m}(r,\theta,\phi)$ which correspond to the spherical harmonics via

$$U_{n\ell m}(r,\theta,\phi) = F(r) Y_{\ell m}(\theta,\phi),$$

where

$$Y_{11} = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi}, \quad Y_{10} = -\sqrt{\frac{3}{4\pi}}\cos\theta, \quad Y_{1-1} = \sqrt{\frac{3}{4\pi}}\sin\theta e^{-i\phi}.$$

We shall ignore the radial part, and set $r=1$, so that $U_{n\ell m}(r,\theta,\phi) = Y_{\ell m}(\theta,\phi)$.

Let us consider $U_{210} = Y_{10}$, and its transformed version, by (3.3),

$$U'_{210}(x) = U_{210}(R^{-1}x) = (R^{-1}x)_z,$$

where we resolve in the z-direction. Now, let R^{-1} be a rotation by β about the y-axis,

$$R^{-1}x = R^{-1}\begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} x' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}\begin{pmatrix} x \\ z \end{pmatrix}.$$

Hence, multiplying out the component in the z-direction,

$$(R^{-1}x)_z = x\sin\beta + z\cos\beta = z',$$

if we use (3.1) for x and z, then

$$\begin{aligned} z' &= \sin\beta\sin\theta\cos\phi + \cos\beta\cos\theta \\ &= \cos\beta U_{210} + \frac{1}{\sqrt{2}}\sin\beta\left(U_{21-1} - U_{211}\right) \\ &= U'_{210}. \end{aligned}$$

Hence, we see that the new wavefunction is indeed a combination of the old components. We can read off the matrix elements,

$$D_{00} = \cos\beta, \quad D_{-10} = \frac{1}{\sqrt{2}}\sin\beta, \quad D_{10} = -\frac{1}{\sqrt{2}}\sin\beta = -D_{-10}.$$

By considering the rotation of the other basis vectors, we could build up the full rotation matrix.

3.2 Equivalent Representations

Even though we have seen that we can represent elements of a group by matrices, there is an infinite set of matrices which we can use to define to represent the group.

If D is a representation, then so is SDS^{-1}. That is, two representation, D_1 and D_2 are equivalent if they are related by the *similarity transformation*

$$D_2 = SD_1S^{-1}.$$

To prove this, consider that as D is a representation,

$$D(g_1)D(g_2) = D(g_1g_2).$$

Then, consider

$$D'(g_1) = SD(g_1)S^{-1}, \quad D'(g_2) = SD(g_2)S^{-1}, \quad D'(g_1g_2) = SD(g_1g_2)S^{-1}.$$

But, the last expression can be written as

$$\begin{aligned} D'(g_1g_2) &= SD(g_1g_2)S^{-1} \\ &= SD(g_1)D(g_2)S^{-1} \\ &= SD(g_1)S^{-1}SD(g_2)S^{-1} \\ &= D'(g_1)D'(g_2). \end{aligned}$$

Therefore, we see that D and D' are equivalent representations. What this means, for example, is that if we have an operation that rotates by a certain angle, it does not matter where we do the rotation, the object will always be rotated by that angle. If D takes a vector v to v', then SDS^{-1} takes the vector Sv to the vector Sv'.

So, we need some way to distinguish genuinely different representations, which is called the character χ of a representation.

The **character** of a representation D of a group G is the set

$$\chi = \{\chi(g) \,|\, g \in G\}, \tag{3.4}$$

where $\chi(g)$ is the trace of $D(g)$. Then, equivalent representations will have the same character, as

$$\text{Tr } (SDS^{-1}) = \text{Tr } (SS^{-1}D) = \text{Tr } (D),$$

after using the cyclic invariance of the trace,

$$\text{Tr } (ABC) = \text{Tr } (BCA) = \text{Tr } (CAB).$$

3.3 Reducibility of Representations

Consider the 3D (i.e. [3]-dim) representation of rotations about the z-axis,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Notice that it is only the x and y coordinates which "mix" under the transformation. We write the above in *block diagonal form*,

$$\begin{pmatrix} \mathbf{x}' \\ z' \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ z \end{pmatrix} \quad \Rightarrow \quad \mathbf{x}' = A\mathbf{x}, \quad z' = z,$$

where the correspondence is obvious;

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}.$$

Therefore, we say that the [3]-dim representation has been broken into $D^{(2)} \oplus D^{(1)}$, a [2]-dim and [1]-dim representation that do not mix; these are the invariant subspaces $\binom{x}{y}$ and z.

A representation of dimension $n+m$ is said to be reducible if $D(g)$ takes the form

$$D(g) = \begin{pmatrix} A(g) & C(g) \\ 0 & B(g) \end{pmatrix},$$

where $A(g)$ is an $m \times m$ matrix, $C(g)$ is an $m \times n$ matrix, $B(g)$ is an $n \times n$ matrix and 0 is an $n \times m$ null matrix. Thus, A and B constitute $[m]$ and $[n]$ dimensional representations of G. To prove this, notice,

$$D(gg') = \begin{pmatrix} A(gg') & C(gg') \\ 0 & B(gg') \end{pmatrix}, \tag{3.5}$$

but,

$$\begin{aligned} D(gg') &= D(g)D(g') \\ &= \begin{pmatrix} A(g) & C(g) \\ 0 & B(g) \end{pmatrix} \begin{pmatrix} A(g') & C(g') \\ 0 & B(g') \end{pmatrix} \\ &= \begin{pmatrix} A(g)A(g') & A(g)C(g') + C(g)B(g') \\ 0 & B(g)B(g') \end{pmatrix}. \end{aligned} \tag{3.6}$$

Now, if we equate the elements of (3.5) and (3.6), we see that

$$A(gg') = A(g)A(g'), \quad B(gg') = B(g)B(g').$$

This confirms that A is an $[m]$-dim representation, and B is an $[n]$-dim representation.

3.3.1 Complete Reducibility

Consider the case where $C(g) = 0$, so that

$$D(g) = \begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}.$$

Then, the $[n+m]$-dim representation is totally specified in terms of $[n]$ and $[m]$-dim representations, which we write as

$$D(g) = A(g) \oplus B(g).$$

In analogy with the rotation of a vector matrix, the n components of A do not allow a mixing with the m components of B.

It may be that A and B can be reduced further, until we reach the irreducible representations which are the fundamental building blocks in terms of which we can write any representation. We will use the direct sum $\oplus$ a lot more in subsequent sections (where its use will explain its meaning), but it essentially combines two matrices, by putting a second matrix along the diagonal, following a first matrix (the direct sum does not add the elements).

3.4 Groups Acting on Vector Spaces

So far, we have mapped group elements into matrices, so that

$$g \longmapsto D(g), \quad D(g_1 g_2) = D(g_1)D(g_2).$$

And thus far, the objects the matrices act upon have been mapped to vectors. If we consider our discussion on page 5, we had group elements a_i which acted upon the triangle with vertices (a, b, c). In this example, we would map the group elements a_i onto matrix operators, and the triangles onto vectors, so that the matrices act upon a vector to produce a new vector.

We must now talk about what we mean by a vector.

3.4.1 Vector Space Axioms

A vector is a member of the vector space, which has the following properties. A vector space V is defined over the field of complex numbers, and is a set of elements $\{v\}$, with two operations defined – addition and multiplication by a scalar. The axioms (or, properties) that define a vector space are:

- **A: Addition**

  A0: Closure: adding two vectors in the vector space V gives a result within V

  $$\forall \mathbf{u}, \mathbf{v} \in V, \quad \mathbf{u} + \mathbf{v} \in V.$$

  A1: Associativity: we have that

  $$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}, \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V.$$

  A2: Commutative: addition in V is commutative,

  $$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}, \quad \forall \mathbf{u}, \mathbf{v} \in V.$$

  A3: Null Vector: there exists a null vector,

  $$\exists\, \mathbf{0} \in V : \mathbf{u} + \mathbf{0} = \mathbf{u}, \quad \forall \mathbf{u} \in V.$$

  A4: Inverse: every element $\mathbf{u}$ has an inverse $-\mathbf{u}$,

  $$\forall \mathbf{u} \in V, \exists - \mathbf{u} \in V : \mathbf{u} + (-\mathbf{u}) = \mathbf{0}.$$

- Hence, we see that the vector space V forms an Abelian group under addition

- **B: Multiplication by a Scalar**

  B0: Closure: multiplying a vector by a scalar returns a vector,

  $$\forall a \in \mathbb{C}, \mathbf{u} \in V : a\mathbf{u} \in V.$$

  B1: Distributive: we have

  $$a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}, \quad \forall a \in \mathbb{C}, \mathbf{u}, \mathbf{v} \in V.$$
  $$(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}, \quad \forall a, b \in \mathbb{C}, \mathbf{u} \in V.$$

  B2: Associative: we have

  $$a(b\mathbf{u}) = (ab)\mathbf{u}, \quad \forall a, b \in \mathbb{C}, \mathbf{u} \in V.$$

  B2: Identity element: there is an identity scalar, such that

  $$1\mathbf{u} = \mathbf{u}, \quad 1 \in \mathbb{C}, \forall \mathbf{u} \in V.$$

Pretty much all of these definitions are obvious, but one tends to use them without realising that one only can use them because the space is defined to allow their use.

Examples A rather trivial example of a vector space, are the vectors in 3D. Another example, are function spaces. Consider the space of functions of a real variable x, over the interval $0 \leq x \leq 1$, such that $f(0) = f(1) = 0$. For example, this could be a wavefunction in an infinite potential outside the range. We can express such functions as

$$f(x) = \sum_{n=0}^{\infty} f_n u_n(x),$$

where, for example, $u_n(x) = \sin n\pi x$. These u_n can be thought of as a basis, and the f_n as components. As the sum is infinite, the space has infinite dimension – called a Hilbert space.

3.4.2 Dimension of a Vector Space

In order to define the dimension of a vector space, we need to define the concepts of linear independence and of a basis.

Linear Independence A set of vectors $\{\mathbf{e}_i\}$, $i = 1, \dots m$, is linearly independent if there is no non-trivial combination of them which yields the null vector. That is, if we cannot write one in terms of the others. This means that if we have $\lambda_i \mathbf{e}_i = \mathbf{0}$, then the only solution, if the $\mathbf{e}_i$ are linearly independent, is $\lambda_i = 0$.

Basis A linearly independent set of vectors $\{\mathbf{e}_i\}$, $i = 1, \dots m$, forms a basis of the vector space, if they span the space. That is, $\forall \mathbf{u} \in V$, one can express $\mathbf{u}$ in terms of a linear superposition of the basis – we can decompose any vector in terms of the basis, $\mathbf{u} = u_i \mathbf{e}_i$.

Dimension of a Vector Space The dimensionality of the vector space is simply the number of basis vectors.

3.4.3 Groups as Linear Transformations on a Vector Space

A linear transformation acts upon a vector space according to

$$\hat{T}\left(\alpha \mathbf{u} + \beta \mathbf{v}\right) = \alpha \hat{T} \mathbf{u} + \beta \hat{T} \mathbf{v}.$$

Now, for a linear transformation to represent a group (i.e. be a representation), it must satisfy

$$\hat{T}(g_1 g_2) = \hat{T}(g_1) \hat{T}(g_2).$$

We could connect the linear transformation to matrices, but, there are an infinite set of matrices with which we could make the connection. Let us show this link.

First, one chooses a basis $\{\mathbf{e}_i\}$, and then we can write a vectors as

$$\mathbf{v} = \mathbf{e}_i u_i,$$

so that the vector $\mathbf{v}$ has components u_i. Then, consider the action of the linear transformation on the vector,

$$\hat{T}\mathbf{v} = \left(\hat{T}\mathbf{e}_i\right) u_i.$$

Now, as $\hat{T}\mathbf{e}_i \in V$, but is some "new vector" in V, we can write it as a linear combination of the basis within V,

$$\hat{T}\mathbf{e}_i = \mathbf{u}'_i = \mathbf{e}_j D_{ji}.$$

Therefore,

$$\hat{T}\mathbf{v} = \mathbf{e}_j D_{ji} u_i,$$

or, defining $u'_j \equiv D_{ji} u_i$, we see that

$$\hat{T}\mathbf{v} = \mathbf{e}_j u'_j.$$

Hence, we see that the action of a linear transformation upon a vector gives a new vector with different components – which are linear combinations of the old components – but uses the same basis. Hence, we say that D_{ji} is a matrix which represents the action of the linear transformation. So, the action of $\hat{T}$ is described by a matrix multiplying the coordinates in a given basis; for example

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

We now want to show the relation between two different matrix representations of $\hat{T}$ in two different bases. Let us consider $\hat{T}$ acting on a basis vector in the $\mathbf{f}$-basis, given knowledge of the action of $\hat{T}$ on the $\mathbf{e}$-basis. We are also given the relation between the $\mathbf{e}$- and $\mathbf{f}$-bases.

Let us take

$$\hat{T}\left(\mathbf{f}_i\right) = \hat{T}\left(\mathbf{e}_j S_{ji}\right) = \hat{T}\left(\mathbf{e}_j\right) S_{ji},$$

so that S_{ji} is the relation between the bases. Now, we know that the action of $\hat{T}$ on the $\mathbf{e}$-basis, via the D_{ji}, so that

$$\hat{T}\left(\mathbf{e}_j\right) = \mathbf{e}_k D_{kj}.$$

So, combining these,

$$\hat{T}\left(\mathbf{f}_i\right) = \mathbf{e}_k D_{kj} S_{ji}. \tag{3.7}$$

Now, the relation between the bases is given by the matrix equation

$$\mathbf{f} = \mathbf{e}S \quad \Rightarrow \quad \mathbf{e} = \mathbf{f}S^{-1} \quad \Rightarrow \quad \mathbf{e}_k = \mathbf{f}_l S^{-1}_{lk},$$

and so using in (3.7)

$$\hat{T}\left(\mathbf{f}_i\right) = \mathbf{f}_l S^{-1}_{lk} D_{kj} S_{ji} = \mathbf{f}_l \left(S^{-1} D S\right)_{li}.$$

Therefore,

$$\hat{T}(\mathbf{f}_i) = \mathbf{f}_j \left(S^{-1}DS\right)_{ji},$$

that is, the matrix representing $\hat{T}$ in the $\mathbf{f}$-basis is $S^{-1}DS$.

Now let us derive the relation between components of vectors in the $\mathbf{e}$- and $\mathbf{f}$-bases. Consider the two vectors,

$$\mathbf{f}v = \mathbf{e}u,$$

where the vector is the same, but has different components in a different basis. Now, we have that $\mathbf{f} = \mathbf{e}S$, so that the above is just

$$\mathbf{e}Sv = \mathbf{e}u \quad \Rightarrow \quad Sv = u,$$

or,

$$v = S^{-1}u.$$

Hence, we have that the components in $\mathbf{f}$ relate to those in $\mathbf{e}$ using the inverse of the matrix that relates the basis vectors.

Example Show that the [3]-dim representation of the group of rotations about the z-axis is completely reducible when reffered to

$$\left(-\frac{x+iy}{\sqrt{2}}, \frac{x-iy}{\sqrt{2}}, z\right)$$

as the basis. Also calculate the character of the representation.

So, we are given the transformation

$$\begin{aligned} x' &= -\frac{x}{\sqrt{2}} - i\frac{y}{\sqrt{2}}, \\ y' &= \frac{x}{\sqrt{2}} - i\frac{y}{\sqrt{2}}, \\ z' &= z; \end{aligned}$$

from which we can construct the matrix

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Hence,

$$S = \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The operator D corresponds to rotation $\hat{T}(\theta)$, in the (x, y, z)-basis;

$$D = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We now want to compute $S^{-1}DS$. The inverse of S is just the Hermitian conjugate $S^\dagger$ (but a little more on this later), so that

$$\begin{aligned} SDS^{-1} &= \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Therefore, a rotation by θ in the (x', y', z')-basis is represented by the matrix

$$SDS^{-1} = \begin{pmatrix} e^{i\theta} & 0 & 0 \\ 0 & e^{-i\theta} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Finally, we want to compute the character of the two representations. Recall that the character of a representation is the trace of the matrix. So, we have

$$\begin{aligned} \mathrm{Tr}\ D &= 1 + 2\cos\theta, \\ \mathrm{Tr}\ SDS^{-1} &= 1 + e^{i\theta} + e^{-i\theta} \\ &= 1 + 2\cos\theta. \end{aligned}$$

Hence, the characters of the representations are the same, and hence the two representations are equivalent. This is obvious, as the action of rotating by θ should always been the same; if one starts at a different place, then one gets a different answer, but the angle by which one rotates is the same.

In doing this, we exploited the unitarity of S. Consider that

$$\mathbf{u} = S\mathbf{v} \quad \Rightarrow \quad \mathbf{u}^\dagger = \mathbf{v}^\dagger S^\dagger,$$

then, the length of a vector is just

$$\mathbf{u}^\dagger \mathbf{u} = \mathbf{v}^\dagger S^\dagger S \mathbf{v},$$

and the length is preserved if

$$S^\dagger S = I,$$

the identity. Therefore,

$$S^{-1} = S^\dagger$$

for length preserving transformations.

3.4.4 Reducibility

In the language of linear transformations, acting on a vector space amounts to the existence of two or more invariant subspaces:

$$\hat{T}\left(\mathbf{u}+\mathbf{v}\right)=\hat{T}\mathbf{u}+\hat{T}\mathbf{v}.$$

Now, as $\hat{T}\mathbf{u}\in U$ and $\hat{T}\mathbf{v}\in V$, then $\hat{T}$ is completely reducible. Thus, in matrix form, we will have something like

$$\begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}=\begin{pmatrix} u' \\ v' \end{pmatrix}=\begin{pmatrix} Au \\ Bv \end{pmatrix}.$$

where u is an $[m]$-dim vector and v is an $[n]$-dim vector. Basically, if the representation is reducible, the action of the transformation will not mix between the u and v.

3.5 Direct Products of Representations & the Clebsch-Gordan Decomposition

Here we will develop a formalism that will allow us to understand how composite systems can have their quantum numbers read off.

Consider a composite system, such as an electron and something else. Let the components of the electrons wavefunction transform under a rotation $g\in G$ by

$$\psi'_a=D^{(\mu)}_{ba}(g)\psi_b,$$

and the other object that rotates via

$$\chi'_c=D^{(\nu)}_{dc}(g)\chi_d.$$

We thus have two wavefunctions, ψ_a,χ_c that transform. The indices μ,ν are labels for the representation, so that we have two different representations for the two different "types" of thing we are transforming. Then, combining the two, the composite system transforms as

$$\psi'_a\chi'_c=D^{(\mu)}_{ba}(g)D^{(\nu)}_{dc}(g)\psi_b\chi_d. \tag{3.8}$$

Now, consider the composite system as having wavefunction Ψ, so that this becomes

$$\Psi'_{ac}=D^{(\mu)}_{ba}(g)D^{(\nu)}_{dc}(g)\Psi_{bd}.$$

Now, we can consider the ordered pair of indices (bd) as a single label B, and likewise (ac) as A. Thus, the above becomes

$$\Psi'_A=D^{(\mu\times\nu)}_{BA}(g)\Psi_B,$$

where

$$D^{(\mu\times\nu)}_{BA}(g)=D^{(\mu)}_{ba}(g)D^{(\nu)}_{dc}(g).$$

These matrices $D^{(\mu\times\nu)}_{BA}(g)$ are a representation of G, since they satisfy

$$D^{(\mu\times\nu)}(g_1g_2) = D^{(\mu\times\nu)}(g_1)D^{(\mu\times\nu)}(g_2).$$

Now, if there is no interaction of the particles, or wavefunctions, when we bring them together into a composite system, then $D^{(\mu\times\nu)}$ actually represents $G\times G$ (i.e. independent rotations of ψ and χ), and $D^{(\mu\times\nu)}$ will be an irreducible representation. That is, we bring the two particles together, where each is represented in some way by the symmetry group G, and the composite system just has $G\times G$ – the symmetry of each particle is preserved.

Conversely, if we do have interaction, then $G\times G\to G$, in which case the two symmetry groups become one, and some symmetry is lost. In this case, the $D^{(\mu\times\nu)}$ are reducible, and can be written as a direct sum of irreducible representations of G,

$$D^{(\mu)}\otimes D^{(\nu)} = D^{(\mu\times\nu)} = \sum_{\oplus\sigma} a_\sigma D^{(\sigma)}. \tag{3.9}$$

Basically, what the direct sum does, is not to add the elements of the irreducible representations (which is meaningless in the case that the representations have a different dimension), but to put the irreducible representation matrices along the diagonals of a bigger matrix – construct a bigger matrix from irreducible representations – each of which may have different dimension, to make a block-diagonal matrix. Then, we say that $D^{(\sigma)}$ are the irreps of the group G (we use the shorthand that irreps stands for irreducible representations).

Let us consider an example. The states ψ,χ could be electronic wavefunctions, rotated in spin-space. Then, each has two components, and hence the composite $\psi\chi$ is a 4-component tensor. The composite then transforms under a 4x4 matrix $D^{(2\times2)}$, where

$$D^{(2\times2)} = D^{(3)}\oplus D^{(1)} = \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix},$$

where A is a 3x3 matrix and C is a 1x1 number (scalar). Initially, we have two electrons, each of which can be in one of two states (hence why ψ,χ had 2-components). Then, when we bought them together, we have an object which can have 3 different spins, but there are 4 ways of doing this. Hence, the final representation matrix is 4x4, and is made up of an irrep of dimension-3, and an irrep of dimension-1.

The direct sum over irreps (3.9) is called a Clebsch-Gordan decomposition.

4 Continuous Groups

Thus far, we have only considered groups with a finite number of elements, or, rotations through a specific set of angles. Now we shall consider groups with an infinite number of elements, and rotations through any angle. Such groups are also called *Lie groups* (pronounced "lee").

4.1 $SO(N)$ Groups

These groups represent rotations through any angle, in N-dimensions.

One of these groups is the $SO(2)$ group. This is the group of special (hence the S) orthogonal (hence the O) matrices R, in 2-dimensions (hence the 2), representing a 2D rotation. That we specify special means that the determinant of R is unity, and that R is orthogonal means that its inverse is its transpose;

$$\begin{aligned} \text{special} \quad &\Rightarrow \quad \det R = 1, \\ \text{orthogonal} \quad &\Rightarrow \quad R^T = R^{-1}. \end{aligned}$$

In the usual Cartesian basis, we can write the matrix $R(\phi)$ representing rotations through ϕ, about the z-axis (notice that we keep the "thing" we are rotating "in" the 2D plane), as

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \tag{4.1}$$

and this matrix acts upon the usual 2D position vector $\binom{x}{y}$. Now, as there are many ways of representing a group, we may as well single out one to be the *defining representation*. Therefore, we say that (4.1) is the defining representation of $SO(2)$. It is an irrep over the field of real numbers $\mathbb{R}$.

Let us change basis to a complex basis, as we did in our example on p34,

$$(x, y, z) \longmapsto \left(-\frac{x + iy}{\sqrt{2}}, \frac{x - iy}{\sqrt{2}}, z \right).$$

Then, this can be represented in polar coordinates,

$$\left(re^{i\theta}, re^{-i\theta}, z \right).$$

Now consider a rotation by an angle ϕ,

$$\begin{pmatrix} re^{i\theta} \\ re^{-i\theta} \\ z \end{pmatrix} \longmapsto \begin{pmatrix} re^{i(\theta+\phi)} \\ re^{-i(\theta+\phi)} \\ z \end{pmatrix} = \begin{pmatrix} re^{i\theta}e^{-\phi} \\ re^{-i\theta}e^{-i\phi} \\ z \end{pmatrix},$$

and hence, we see that we can represent the rotation as

$$\begin{pmatrix} re^{i\theta} \\ re^{-i\theta} \end{pmatrix} \longmapsto \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix} \begin{pmatrix} re^{i\theta} \\ re^{-i\theta} \end{pmatrix},$$

that is, with rotation matrix

$$R_2(\phi) \equiv \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}. \tag{4.2}$$

Now, notice that the trace of $R_2(\phi)$ is the same as that for the defining representation (4.1),

$$\text{Tr } R = \text{Tr } R_2 = 2\cos\phi,$$

and therefore, R and R_2 are equivalent representations of a rotation through an angle ϕ. Now, notice that $R_2(\phi)$ is of block-diagonal form, and its entries are the irreps. Therefore, we see that the irreps of $SO(2)$ are just 1-dimensional, and are of the form $e^{-im\phi}$, where m is an integer, $m \in \mathbb{Z}$. The Clebsch-Gordan decomposition – that is, how to bring two rotations together – is just

$$e^{-im\phi}e^{-im'\phi} = e^{-i(m+m')\phi}.$$

Hence, we see that the Clebsch-Gordan series has only one term, and is of the form

$$D^{(m)} \oplus D^{(m')} = D^{(m+m')}.$$

We restrict m to $m \in \mathbb{Z}$, so that $D^{(m)}(\phi + 2\pi) = D^{(m)}(\phi)$. In quantum mechanics, we lift this restriction, and allow m to be a half-integer as well.

4.2 Generators of a Lie Group

The generators are the fundamental objects of a Lie group, and are the matrices responsible for infinitesimal rotations. Let us obtain the generator for the defining representation of $SO(2)$, (4.1);

$$R(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

Let us take a very small rotation angle, $\phi \ll 1$, and expand the sines and cosines about $\phi = 0$;

$$R(\phi) = \begin{pmatrix} 1 - \frac{\phi^2}{2!} + \ldots & -\phi + \ldots \\ \phi + \ldots & 1 - \frac{\phi^2}{2!} + \ldots \end{pmatrix},$$

and, to first order only,

$$\begin{aligned} R(\phi) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\phi \\ \phi & 0 \end{pmatrix} + \mathcal{O}(\phi^2) \\ &= I - iX\phi. \end{aligned} \tag{4.3}$$

where we have made the identification

$$-iX = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{4.4}$$

Now, infact notice that

$$\frac{dR}{d\phi} = \begin{pmatrix} -\sin\phi & -\cos\phi \\ \cos\phi & -\sin\phi \end{pmatrix} \quad \Rightarrow \quad \left.\frac{dR}{d\phi}\right|_{\phi=0} = -iX.$$

Therefore, we have what we call the *generator*,

$$-iX = \left.\frac{dR}{d\phi}\right|_{\phi=0}. \tag{4.5}$$

We must have that X is Hermitian. This follows from the unitarity of R, so that $R^\dagger R = I$. So, using (4.3),

$$R^\dagger R = (I + i\phi X^\dagger)(I - i\phi X) = I - i\phi(X - X^\dagger) = I,$$

which gives us the requirement that $X = X^\dagger$, and hence that X is Hermitian. Finally, notice that $\det R = 1$, which means that the trace of X is zero,

$$\mathrm{Tr}\ X = 0.$$

This follows as the determinant of a diagonal matrix is the product of its entries (which are those on the diagonal). Therefore, as $\det R = 1$, and upon expansion we find the unit matrix (which is diagonal, and whose components multiply to unity), and therefore, the second matrix in the expansion (i.e. the X-matrix) must have zero trace (which is the sum of its diagonal components). We shall state, and then show, that finite rotations can be built up from the generators

$$R(\phi) = e^{-iX\phi}. \tag{4.6}$$

Let us check this, given (4.4), by expanding $e^{-iX\phi}$. So, the expansion is just

$$R(\phi) = \sum_{n=0}^{\infty} (-i\phi)^n \frac{X^n}{n!}.$$

Note that, as X is unitary, $X^2 = I$ and $X^3 = X$. Thus, $X^{2n} = I$ and $X^{2n+1} = X$; hence, using these, the expansion is just

$$R(\phi) = I\left(1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!}\ldots\right) - iX\left(\phi - \frac{\phi^3}{3!} + \ldots\right).$$

We now notice that the first bracket is just the expansion of cosine, and the second bracket of sine. Therefore,

$$\begin{aligned} R(\phi) &= I\cos\phi - iX\sin\phi \\ &= \begin{pmatrix} \cos\phi & 0 \\ 0 & \cos\phi \end{pmatrix} + \begin{pmatrix} 0 & -\sin\phi \\ \sin\phi & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \end{aligned}$$

Hence, we see that we have recovered the defining representation using the generator. We have, however, relied upon a specific form of X, (4.4); what about an arbitrary form? Does this still work? A more general way that works for an arbitrary representation $D(\phi)$ is by noting that

$$D(\phi)D(\phi') = D(\phi + \phi'),$$

where this is the homomorphism requirement. Then, we differentiate both sides with respect to ϕ, and set $\phi' = 0$. To do this to the RHS, notice that

$$\frac{d}{d\phi'}D(u) = \frac{dD(u)}{du}\frac{du}{d\phi'}, \quad u \equiv \phi + \phi',$$

so that

$$\left.\frac{d}{d\phi'}D(u)\right|_{\phi'=0} = \left.\frac{dD(u)}{du}\right|_{u=\phi} = \frac{dD(\phi)}{d\phi}.$$

Therefore, using this, we have

$$D(\phi)\left.\frac{dD(\phi')}{d\phi'}\right|_{\phi'=0} = \frac{dD(\phi)}{d\phi}.$$

Now, using (4.5), which is the general form of the generator, we have

$$D(\phi)(-iX) = \frac{dD(\phi)}{d\phi},$$

which easily integrates to give

$$D(\phi) = e^{-iX\phi}. \tag{4.7}$$

Therefore, we have a general expression for the generator, without resorting to a specific representation.

We have seen that the generator is always given by

$$-iX = \left.\frac{dD(\phi)}{d\phi}\right|_{\phi=0}.$$

For example, for the 1-dimensional representations of $SO(2)$, $D(\phi) = e^{-im\phi}$, the generator is

$$-iX = -im \quad \Rightarrow \quad X = m,$$

for the representation labelled m. It is the object X which we term the generator.

4.3 Role of Generators in Quantum Mechanics

Generators correspond to a differential operator, which we shall show.

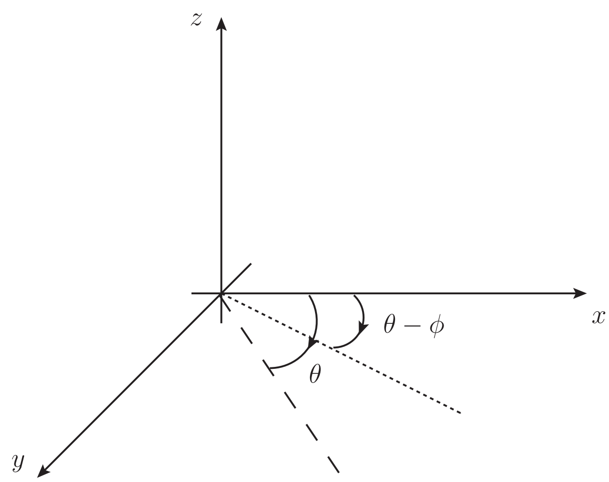


Figure 4.1: Rotating a vector by an angle ϕ.

Consider rotations about the z-axis through an angle ϕ, as in Figure (4.1). Recall (3.3),

$$\psi'(x) = \psi(R^{-1}x),$$

then,

$$\psi'(r,\theta) = \psi(r,\theta-\phi). \tag{4.8}$$

Now, in terms of operators, we can write

$$\psi'(r,\theta) = \hat{U}_{\mathrm{R}}\psi(r,\theta).$$

If we consider small rotation angles, $\phi \ll 1$, we make the expansion to see that

$$\hat{U}_{\mathrm{R}} = 1 - i\hat{X}\phi,$$

and hence that

$$\psi'(r,\theta) = (1 - i\hat{X}\phi)\psi(r,\theta) \tag{4.9}$$

where $\hat{X}$ is the generator of the rotation. Hence, equating (4.8) and (4.9),

$$\psi(r,\theta-\phi) = \psi(r,\theta) - i\hat{X}\phi\psi(r,\theta),$$

which easily rearranges into

$$\frac{\psi(r,\theta-\phi)-\psi(r,\theta)}{\phi} = -i\hat{X}\psi(r,\theta).$$

If we take the limit of $\phi \to 0$ (i.e. an infinitesimal rotation), then the LHS becomes a differential, so that

$$-\frac{d}{d\theta}\psi(r,\theta) = -i\hat{X}\psi(r,\theta),$$

and hence we can make the identification of the generator with a differential operator,

$$\hat{X} = -i\frac{d}{d\theta}. \tag{4.10}$$

Notice that this is just like the quantum mechanical operators for angular momentum, or linear momentum,

$$\hat{X} = \frac{\hat{J}_z}{\hbar}, \quad \hat{P}_x = -i\hbar\frac{d}{dx}.$$

4.4 The $SO(3)$ Group

The $SO(3)$ group is the group of all proper rotations in 3-dimensions. The defining representation consists of a 3x3 matrix that we have already met. Let us start with rotations about the z-axis (which we shall call the "3" axis),

$$R_3(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

and hence, the generator $-iX_3$ is

$$-iX_3 = \left.\frac{dR_3}{d\phi}\right|_{\phi=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notice that X_3 is both traceless and Hermitian. We can elegantly express X_3 in terms of the Levi-Civita tensor ϵ_{ijk},

$$(X_3)_{ij} = -i\epsilon_{ij3}. \tag{4.11}$$

Just to be clear, ϵ_{ijk} has the value zero if any of its indices are the same, value $+1$ if its indices are an even permutation of ijk, and -1 if an odd permutation. For example, $\epsilon_{112} = 0, \epsilon_{312} = 1, \epsilon_{132} = -1$. We can do the same thing for rotations about the x- and y-axes. For example,

$$R_1(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}.$$

Hence, the other two generators are

$$(X_1)_{ij} = -i\epsilon_{ij1}, \quad (X_2)_{ij} = -i\epsilon_{ij2}. \tag{4.12}$$

Hence, one can easily see that the general expression for the generators is

$$(X_k)_{ij} = -i\epsilon_{ijk}. \tag{4.13}$$

Then, the full rotation matrix is given by

$$D(\phi) = e^{-i\hat{\mathbf{n}}\cdot\mathbf{X}\phi}, \tag{4.14}$$

where $\hat{\mathbf{n}}$ is a unit vector along one of the $(1,2,3)$-axes, and $\mathbf{X} = (X_1, X_2, X_3)$, the vector of generators.

4.4.1 $SO(3)$ Rotation About an Arbitrary Axis

We have derived the rotation matrix for rotations about the axes defining a Cartesian coordinate system. Now we want to use the generators to derive the rotation matrix for rotations about any axis.

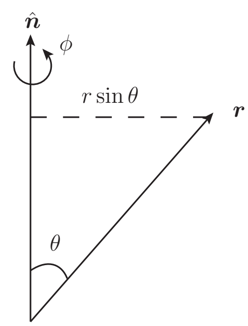


Figure 4.2: Rotating a vector $\mathbf{r}$ by an angle ϕ, about an arbitrary axis $\hat{\mathbf{n}}$.

With reference to Figure (4.2), we see how things are defined in our problem. Notice that the distance from the axis of rotation $\hat{\mathbf{n}}$ to the vector $\mathbf{r}$ is $r\sin\theta$, which is just $\hat{\mathbf{n}} \times \mathbf{r}$.

Now consider rotating about the axis, by some small angle $\phi \ll 1$. Then, there will be a change in the vector $\mathbf{r}$, which will just be its current distance from the axis times the angle it was rotated by (i.e. a chord length),

$$\delta\mathbf{r} = \hat{\mathbf{n}} \times \mathbf{r}\phi.$$

Therefore, we take our vector $\mathbf{r}$ to some new vector $\mathbf{r}'$, say, by adding on $\delta\mathbf{r}$,

$$\begin{aligned} \mathbf{r} \longmapsto \mathbf{r}' &= \mathbf{r} + \delta\mathbf{r} \\ &= \mathbf{r} + \phi\left(\hat{\mathbf{n}} \times \mathbf{r}\right). \end{aligned}$$

In terms of components, this just reads

$$r_i \longmapsto r_i' = r_i + \phi\epsilon_{ikj}n_k r_j.$$

We can use the anti-symmetry of the last two indices of the Levi-Civita tensor to write

$$\begin{aligned} r_i' &= r_i - \phi \epsilon_{ijk} n_k r_j \\ &= \left(\delta_{ij} - \phi \epsilon_{ijk} n_k\right) r_j . \end{aligned}$$

Therefore, notice that by writing the transformation like this, we have terms to zeroth and first order correction in ϕ (i.e. the Kronecker-delta and Levi-Civita terms, respectively). Hence, we can read off that

$$\begin{aligned} \delta r_i &= -\phi \epsilon_{ijk} n_k r_j \\ &= -i\phi \left(X_n\right)_{ij} r_j . \end{aligned}$$

Thus, the generator is

$$\left(X_n\right)_{ij} = -i\epsilon_{ijk} n_k = n_k \left(X_k\right)_{ij} ,$$

where $\left(X_k\right)_{ij}$ is the ij^{th}-component of the rotation generator about the k^{th}-axis (which refer to the $(1,2,3)$-axes of the Cartesian system). Hence, we can easily see that this is just a scalar product,

$$\left(X_n\right)_{ij} = \left(\hat{\mathbf{n}} \cdot \mathbf{X}\right)_{ij} ,$$

where $\mathbf{X} = (X_1, X_2, X_3)$. Therefore, the generator for a rotation about any axis $\hat{\mathbf{n}}$ is just $\hat{\mathbf{n}} \cdot \mathbf{X}$, where $\mathbf{X}$ is the vector composed of the generators about the $(1,2,3)$-axes of the Cartesian coordinate system (essentially, what we do, is to determine the projection of the new axis onto the Cartesian system). Hence, the full rotation matrix is just

$$R(\phi) = e^{-i\hat{\mathbf{n}} \cdot \mathbf{X} \phi}, \tag{4.15}$$

where $\hat{\mathbf{n}} \cdot \mathbf{X}$ is also traceless and Hermitian. Therefore, $R(\phi)$ is a special orthogonal matrix;

$$\det R = 1, \quad R^{-1} = R^T .$$

4.4.2 Commutation Relations

The generators form an algebra, or vector space, so that linear combinations of generators produces generators. There is another combination law for this space, commutation: the commutator of two generators is also a generator.

Specifically, for $SO(3)$, the commutator is

$$[X_a, X_b] = i\epsilon_{abk} X_k . \tag{4.16}$$

This can be proved directly, using the Levi-Civita tensor, and the identity

$$\epsilon_{ilk} \epsilon_{jmk} = \delta_{ij} \delta_{lm} - \delta_{im} \delta_{lj} . \tag{4.17}$$

And we shall also present a slightly more physical proof. Let us prove (4.16) using this identity then. So, in index notation, (4.16) reads

$$[X_a, X_b]_{ij} = i\epsilon_{abk}\left(X_k\right)_{ij}.$$

So, the commutator, in index notation, is just

$$[X_a, X_b]_{ij} = \left(X_a X_b\right)_{ij} - \left(X_b X_a\right)_{ij},$$

where we must remember that the X_a are matrices, and that this is matrix multiplication. Now, the first term is just

$$\left(X_a X_b\right)_{ij} = \left(X_a\right)_{ik}\left(X_b\right)_{kj}$$

Now, by (4.13), we have an expression for each of the elements of the generators, $\left(X_k\right)_{ij} = -i\epsilon_{ijk}$. Hence,

$$\begin{aligned}\left(X_a\right)_{ik}\left(X_b\right)_{kj} &= (-i)\epsilon_{ika}(-i)\epsilon_{kjb} \\ &= -\epsilon_{ika}\epsilon_{kjb}.\end{aligned}$$

We now want to get this into a form from which we can apply the identity (4.17); hence, we want the same index to be at the far RHS of both Levi-Civita tensors (i.e. the k-index). So, we can use the anti-symmetry of the last two indices on the first tensor to write $\epsilon_{ika} = -\epsilon_{iak}$. We can then use the cyclic symmetry to write that $\epsilon_{kjb} = \epsilon_{jbk}$. Hence using this,

$$\begin{aligned}\left(X_a\right)_{ik}\left(X_b\right)_{kj} &= \epsilon_{iak}\epsilon_{jbk} \\ &= \delta_{ij}\delta_{ab} - \delta_{ib}\delta_{aj}.\end{aligned}$$

In a very similar fashion, one can show that

$$\left(X_b X_a\right)_{ij} = \delta_{ij}\delta_{ba} - \delta_{ia}\delta_{bj}.$$

Hence, putting these lat two results together, we have the component form of the commutator,

$$\begin{aligned}[X_a, X_b]_{ij} &= \left(X_a X_b\right)_{ij} - \left(X_b X_a\right)_{ij} \\ &= \delta_{ij}\delta_{ab} - \delta_{ib}\delta_{aj} - \delta_{ij}\delta_{ba} + \delta_{ia}\delta_{bj}.\end{aligned}$$

The symmetry of the Kronecker-delta allows us to see that the first and third terms cancel, leaving

$$[X_a, X_b]_{ij} = \delta_{ia}\delta_{bj} - \delta_{ib}\delta_{aj}.$$

We can then use the identity (4.17) to restore the RHS to a product of two Levi-Civita tensors,

$$[X_a, X_b]_{ij} = \epsilon_{ijk}\epsilon_{abk}.$$

Now, $\epsilon_{ijk} = i\left(X_k\right)_{ij}$, by (4.13), and therefore,

$$[X_a, X_b]_{ij} = i\left(X_k\right)_{ij}\epsilon_{abk}.$$

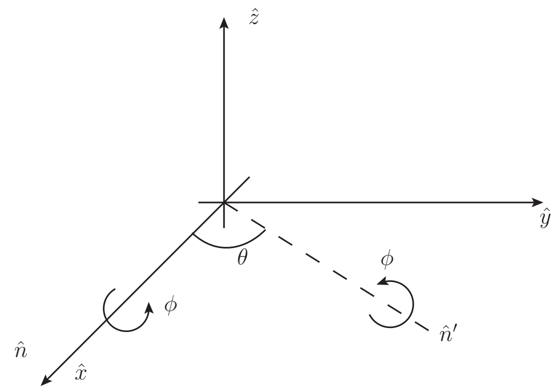


Figure 4.3: Rotating by ϕ about two axes, $\hat{n}$ and $\hat{n}'$. The first axis is aligned along the x-axis, whilst the second is an arbitrary axis in the $x-y$-plane.

We can now take ourselves out of component notation (just to be clear, the above is the ij^{th}-element of the commutator generators a and b, giving some generator k),

$$[X_a, X_b] = iX_k\epsilon_{abk},$$

which is exactly what we set out to prove, (4.16).

With reference to Figure (4.3), we see that the actions of rotating about two different axes, by the same angle, are equivalent. Consider a rotation by ϕ about two axes related by a rotation θ about the z-axis. Then, we have some rotation matrix S that transfers between the two axes of rotation;

$$S : \hat{n} \longmapsto \hat{n}'.$$

This matrix is just the usual rotation matrix about the z-axis, by an angle θ,

$$S(\theta) = e^{-iX_3\theta}.$$

We have a second matrix D, say, that rotates about $\hat{n}$ (i.e. the x-axis) by an angle ϕ,

$$D(\phi) = e^{-i\hat{\mathbf{n}}\cdot\mathbf{X}\phi} = e^{-iX_1\phi}.$$

Now, a rotation by ϕ about the $\hat{n}'$-axis is equivalent to rotating by ϕ about the $\hat{n}$ axis; and their matrices are related just by $S_2 = SDS^{-1}$;

$$S_2 = e^{-iX_3\theta}e^{-iX_1\phi}e^{iX_3\theta}.$$

Now, we can also write the S_2 rotation matrix directly, via

$$S_2 = e^{-i\hat{\mathbf{n}}'\cdot\mathbf{X}\phi} = e^{-i(X_1\cos\theta + X_2\sin\theta)\phi}.$$

Therefore, equating these two representations of S_2,

$$e^{-iX_3\theta}e^{-iX_1\phi}e^{iX_3\theta} = e^{-i(X_1\cos\theta + X_2\sin\theta)\phi}.$$

Expanding the middle exponential and far RHS exponential, to first order in ϕ,

$$e^{-iX_3\theta}\left(I - iX_1\phi\right)e^{iX_3\theta} = I - i\left(X_1\cos\theta + X_2\sin\theta\right)\phi,$$

which is just

$$e^{-iX_3\theta}X_1e^{iX_3\theta} = X_1\cos\theta + X_2\sin\theta.$$

Let us now differentiate this with respect to θ,

$$-iX_3e^{-iX_3\theta}X_1e^{iX_3\theta} + e^{-iX_3\theta}X_1iX_3e^{iX_3\theta} = -X_1\sin\theta + X_2\cos\theta,$$

and set $\theta = 0$, to find

$$i\left(X_1X_3 - X_3X_1\right) = X_2,$$

which is just

$$[X_1, X_3] = -iX_2.$$

This provides a rather physical "proof" of one of the components of (4.16). To do the others, one just rotates about a different axis.

4.4.3 Irreps of $SO(3)$

We now want to find irreducible matrices corresponding to the generators. Now, the generators satisfy the commutation relation $[X_i, X_j] = i\epsilon_{ijk}X_k$, so, we must look for matrices that satisfy this relation. We shall exploit a connection with our knowledge of quantum mechanics: the angular momentum operators satisfy this commutation relation.

We can diagonalise $\hat{J}^2$ and one of $\hat{J}_x, \hat{J}_y, \hat{J}_z$ in the same basis; we choose $\hat{J}_z$ by convention. That we can diagonalise two matrices in the same basis means that we can know their eigenvalues and eigenstates simultaneously. We can label the eigenvalues of $\hat{J}^2$ by j, and of $\hat{J}_z$ by m, so that

$$\hat{J}^2\left|j,m\right\rangle = j(j+1)\left|j,m\right\rangle, \quad \hat{J}_z\left|j,m\right\rangle = m\left|j,m\right\rangle.$$

Now, for each value of j, there are $2j+1$ eigenvalues m and eigenstates $\left|j,m\right\rangle$ of $\hat{J}_z$. Therefore, the eigenstates of $\hat{J}_z$ form a $(2j+1)$-dim basis; a representation of $SO(3)$.

Let us take $j = 1$ as an example; we have that $m = -1, 0, 1$. The matrix elements of the generator for rotations about the z-axis are just

$$\left(X_3\right)_{m'm} = \left\langle j,m'\right| X_3\left|j,m\right\rangle = m\delta_{m'm},$$

and hence, X_3 looks like

$$X_3 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

and thus, the rotation matrix,

$$R_3^{j=1}(\phi) = \begin{pmatrix} e^{-i\phi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\phi} \end{pmatrix}.$$

Infact, one can see that the general matrix, for arbitrary j, is

$$R_3^j(\phi) = \begin{pmatrix} e^{-ij\phi} & 0 & 0 & 0 & 0 \\ 0 & e^{-i(j-1)\phi} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & e^{i(j-1)\phi} & 0 \\ 0 & 0 & 0 & 0 & e^{ij\phi} \end{pmatrix}, \quad |m| \leq j.$$

We can also work out the other generators by introducing the ladder operators,

$$X_+ \equiv X_1 + iX_2, \quad X_- \equiv X_1 - iX_2,$$

where

$$\langle j, m'| X_\pm |j, m\rangle = \sqrt{j(j+1) - m(m \pm 1)}\delta_{m',m\pm 1}.$$

Now, the value $2j + 1$ must be an integer, as it corresponds to the dimensionality of the representation. However, in quantum mechanics, we can have that j is an integer or half-integer. Therefore, we must exclude the half-integer values from the group $SO(3)$. This is due to the requirement of rotations being in real space, and that $D(\phi + 2\pi) = D(\phi)$.

If one exponentiates the half-integer values (i.e. includes them), one arrives at a different, but related, group $SU(2)$.

4.5 The $SU(N)$ Groups

Now, the commutation relations for $SO(3)$ can also be satisfied by 2x2 Hermitian (complex) matrices,

$$X_i = \frac{1}{2}\sigma_i,$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.18}$$

These matrices are the Pauli matrices. Notice that all σ_i are traceless. We can show that any 2x2 traceless Hermitian matrix M can be written as a linear combination of these Pauli matrices. Now, such a matrix M has the form

$$M = \begin{pmatrix} a & b - ic \\ b + ic & -a \end{pmatrix}, \tag{4.19}$$

where $a, b, c \in \mathbb{R}$. Now, we can write the matrix as

$$M = a\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

which are just the Pauli matrices. This proof is very simple, once one notices that (4.19) is the only form that such a matrix M could take (which is verifiable by recalling that $M^\dagger = (M^*)^T$).

Now, the Pauli matrices act on spin states, corresponding to up- and down-spin

$$\left|\tfrac{1}{2}, \tfrac{1}{2}\right\rangle \longleftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \left|\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle \longleftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and are eigenstates of σ_3 (and therefore of X_3),

$$X_3\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad X_3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

This factor of $\frac{1}{2}$ in the eigenvalues corresponds to the $m = \frac{1}{2}$ of $SO(3)$ which we disallowed.

A general rotation in $SU(2)$-space is given by the matrix

$$U(\theta) = e^{-\frac{1}{2}i\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\theta}. \tag{4.20}$$

Hence, notice that $U^\dagger = U^{-1}$, so that U is unitary. Another way of writing this matrix is

$$U(\theta) = I\cos\frac{\theta}{2} - i\left(\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\right)\sin\frac{\theta}{2}. \tag{4.21}$$

4.5.1 Relation Between $SO(3)$ and $SU(2)$

We may notice that the rotation matrices for $SO(3)$ and $SU(2)$ are very similar;

$$\begin{aligned} SU(2) &: \quad e^{-\frac{1}{2}i\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}\theta}, \\ SO(3) &: \quad e^{-i\mathbf{X}\cdot\hat{\mathbf{n}}\theta}. \end{aligned}$$

Thus, we could consider a mapping of $SU(2) \to SO(3)$; for each element of $SU(2)$ we have a corresponding element of $SO(3)$;

$$M : SU(2) \longmapsto SO(3).$$

Now, notice that $U(0) = I$ and $U(2\pi) = -I$ (from the $SU(2)$ rotation matrix), but $R(0) = R(2\pi) = I$ in the $SO(3)$ group; that is, two elements of $SU(2)$ are mapped into the identity element of $SO(3)$.

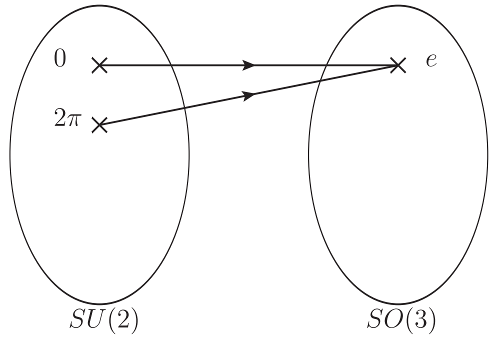


Figure 4.4: A schematic of the mapping of $U(0), U(2\pi)$ into the identity element of $SO(3)$; this is the Kernel of the mapping.

With reference to Figure (4.4), we see that the Kernel of the mapping M is non-trivial (i.e. contains more than just the identity element of $SU(2)$). Using (4.21), we can write the elements of the Kernel into a set;

$$\text{Ker}M = \left\{ \left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), \quad \left(\begin{array}{cc} -1 & 0 \\ 0 & -1 \end{array} \right) \right\}.$$

These elements form a group (as the Kernel is a (normal) subgroup of the image), and infact correspond to the rotations of a line by $0, \pi$. Thus, the Kernel is a Z_2 normal subgroup (with Z_2 just denoting a reflection operation).

The Isomorphism Theorem The image of a mapping f, from $A \mapsto B$ is $A/\text{Ker}f$, so that $B \cong A/\text{Ker}f$.

Hence, by the isomorphism theorem, we see that

$$SO(3) \cong SU(2)/Z_2. \tag{4.22}$$

4.5.2 Example: Isospin in $SU(2)$

Another example of $SU(2)$ transformations is isospin. This example is much more based upon particle physics, and looks at the states themselves, rather than the rotation matrices. Isospin is also called flavour.

One can think of the u and d quarks as an isospin doublet,

$$q = \begin{pmatrix} u \\ d \end{pmatrix},$$

so that a u quark has $I_3 = \frac{1}{2}$ and a d quark has $I_3 = -\frac{1}{2}$ (in complete analogy with electrons having spin up and down).

Now, combining two such quarks gives

$$\tfrac{1}{2} \oplus \tfrac{1}{2} = 1 \oplus 0.$$

The "1" refers to a isospin triplet with $I_3 = -1, 0, 1$, and the "0" to a scalar with $I_3 = 0$. Hence, this predicts four different types of "thing" (infact, two of them – with $I_3 = 0$ – are formed by a linear combination); these are the pions,

$$\begin{aligned} I_3 = 1 \quad &: \quad \pi^+ = u\bar{d}, \\ I_3 = 0 \quad &: \quad \pi^0 = \frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right), \\ I_3 = -1 \quad &: \quad \pi^- = d\bar{u}. \end{aligned}$$

We can also write the combination in terms of the dimensionality of the objects involved. For example, a spin-$\frac{1}{2}$ quark has dimension two (i.e. two degrees of freedom: spin up and down); hence, the above combination can be written

$$2 \otimes 2 = 3 \oplus 1.$$

Notice that we use the direct product symbol on the LHS, and that the scalar state has unit dimension (which is pretty obvious).

Likewise, we can combine 3 quarks,

$$\begin{aligned} \tfrac{1}{2} \oplus \tfrac{1}{2} \oplus \tfrac{1}{2} &= (1 \oplus 0) \oplus \tfrac{1}{2} \\ &= \tfrac{3}{2} \oplus \tfrac{1}{2} \oplus \tfrac{1}{2}. \end{aligned}$$

In words; we combine three particles, each with two degrees of freedom. Straight away, we expect $2 \times 2 \times 2 = 8$ degrees of freedom in our composite particle. So, we first combine two particles, to give $1 \oplus 0$, then combine this with our other $\frac{1}{2}$, which gives both "1" and "0" an extra two degrees of freedom. This then results in a set of three independent things, one having 4 degrees of freedom (from the "$\frac{3}{2}$" having 4 projections $I_3 = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}\frac{3}{2}$), and two having two degrees of freedom (each "$\frac{1}{2}$" has two projections $I_3 = \pm\frac{1}{2}$). Therefore, we have found our eight degrees of freedom, in a quadruplet and two doublet states. One of the doublets is the $\binom{p}{n}$-doublet. In terms of dimension, this combination of three quarks can be written as the decomposition

$$2 \otimes 2 \otimes 2 = 4 \oplus 2 \oplus 2.$$

Now, in writing the composition $\frac{1}{2} \oplus \frac{1}{2} \oplus \frac{1}{2} = \frac{3}{2} \oplus \frac{1}{2} \oplus \frac{1}{2}$, where the end-products are the direct-sums, we are writing the combined system in terms of things which dont interact. That is, the doublets and quadruplet dont interact; there is no symmetry between them. Therefore, as we stated that one of the doublets is the $\binom{p}{n}$ doublet, we expect that p and n are able to be rotated into each other, by a rotation in $SU(2)$-space (i.e. are related by an isospin rotation), and hence should have identical masses. This is not physically the case, as it is not just the $SU(2)$ operations on must consider when thinking about protons and neutrons. The symmetry between protons and neutrons is broken by the constituent quark masses (the quarks dont have exactly the same mass), and electroweak effects. This is because quarks get their mass from a Higgs mechanism, not the strong force. If we tried to extend $SU(2)$ flavour symmetry to $SU(N_{\mathrm{f}})$, where N_{f} is the number of flavours, then the symmetry is badly broken.

This has merely been an illustrative example, and we shall discuss these concepts in more detail later; but, it is worth noting that this example takes a composite system and works out its Clebsch-Gordan decomposition into irreducible representations.

4.6 $SU(N)$ Tensors

This considers the vectors in the vector space itself, rather than the rotation matrices (as did our example on isospin).

We shall now work in terms of objects that carry (or, transform under) the representations of $SU(N)$. Under $SU(N)$, the wavefunctions are vectors, and combining them yields a tensor, and we can study the transformation properties of that tensor. Reducibility will correspond to the existence of invariant subspaces that do not mix under rotation. As we shall see, tensors in $SU(N)$ and $SU(N')$ do not act in an obviously similar way, so we shall begin by discussing $SU(2)$ and $SU(3)$ separately, and then make the discussion more general.

4.6.1 Tensors in $SU(2)$

In $SU(2)$, the fundamental object is a two component vector, or spinor, ψ_a where $a = 1, 2$. These two components correspond to spin-up and spin-down;

$$\psi_1 \quad \Rightarrow \quad |\uparrow\rangle \, , \quad \psi_2 \quad \Rightarrow \quad |\downarrow\rangle \, .$$

The rotation of the spinor under $SU(2)$ is just given by

$$\psi'_a = U_{ab}\psi_b, \quad U \in SU(2). \tag{4.23}$$

There are two types of spinor: spinors with upper and lower indices. Spinors with lower indices (such as ψ_a) correspond to particles, and spinors with upper indices (such as ψ^a) to anti-particles. The transformation above is obviously for a spinor with lower indices; for a spinor with upper indices we have

$$\psi'^a = U^*_{ab}\psi^b. \tag{4.24}$$

Hence, we have that upper-index spinors transform according to the complex conjugate representation U^*. Notice that we have not "made" the indices match up (i.e. contracting lower with upper) as we would with Lorentz indices. We can prove that the U^* form a representation because we can define a mapping from U to U^* that preserves group multiplication. Consider the images of two elements under such a mapping,

$$D(U_1) = U_1^*, \quad D(U_2) = U_2^*.$$

Now, we also have that

$$D(U_1U_2) = (U_1U_2)^* = U_1^*U_2^*,$$

which is just the same as

$$D(U_1)D(U_2) = U_1^*U_2^*.$$

Hence, group multiplication is preserved and therefore U^* is a representation. This essentially follows as the U rotation matrix comes from the σ_i matrices, and that the complex conjugate σ_i^* satisfy the same commutation relations as the σ_i.

For $SU(2)$, U and U^* are equivalent representations. To prove this statement, we introduce a matrix

$$C = i\sigma_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad C^2 = -I. \tag{4.25}$$

Then, we can easily verify that

$$\begin{aligned} C\sigma_1C^{-1} &= (i\sigma_2)\sigma_1(-i\sigma_2), \\ &= -\sigma_1, \\ C\sigma_2C^{-1} &= \sigma_2, \\ C\sigma_3C^{-1} &= -\sigma_3. \end{aligned}$$

Now, as σ_1 and σ_3 are completely real, and σ_2 is completely imaginary, we can thus summarise these three equations by

$$C\boldsymbol{\sigma}C^{-1} = -\boldsymbol{\sigma}^*, \quad \boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3). \tag{4.26}$$

Now, recall (4.21),

$$U(\theta) = I\cos\frac{\theta}{2} - i\,(\boldsymbol{\sigma}\cdot\hat{\mathbf{n}})\sin\frac{\theta}{2},$$

then, we can use (4.26) to easily see that

$$\begin{aligned} CUC^{-1} &= I\cos\frac{\theta}{2} + i\,(\boldsymbol{\sigma}^*\cdot\hat{\mathbf{n}})\sin\frac{\theta}{2} \\ &= U^*, \end{aligned} \tag{4.27}$$

which means that U and U^* are equivalent representations (with conjugating element C). This proves that U and U^* are equivalent representations, for the $SU(2)$ group.

Upper and lower index spinors are related to each other, via

$$\psi^a = C_{ab}\psi_b. \tag{4.28}$$

Again, we note that upper and lower indices on the C-matrix are equivalent (but they are not on the spinors).

Invariance of C We can show that the matrix C (4.25) is invariant under an $SU(2)$ transformation. As $C^2 = -I$, we can trivially write $UC^2 = C^2U$ (note that this is just $-UI = -IU$). Multiplying from the left and right (on both sides of the equation) by C^{-1} gives

$$C^{-1}UC^2C^{-1} = C^{-1}C^2UC^{-1} \quad \Rightarrow \quad C^{-1}UC = CUC^{-1}.$$

We can replace the term on the far RHS by (4.27), so that

$$C^{-1}UC = U^*. \tag{4.29}$$

Now, as the Hermitian conjugate of a matrix is the transpose of the complex conjugate, the complex conjugate of a matrix is the transpose of its Hermitian conjugate; hence,

$$U^\dagger = (U^*)^T \quad \Rightarrow \quad U^* = (U^\dagger)^T.$$

Now, U is unitary, so that $U^\dagger = U^{-1}$, hence

$$U^* = (U^{-1})^T = (U^T)^{-1}.$$

We can use this in the RHS of (4.29), to see that

$$C^{-1}UC = (U^T)^{-1}.$$

We now multiply from the left by C,

$$UC = C(U^T)^{-1},$$

and from the right by U^T,

$$UCU^T = C. \tag{4.30}$$

Similarly, one can show that

$$U^TCU = C.$$

Therefore, what this and (4.30) show, is that C is an invariant under the $SU(2)$ transformation. We can write (4.30) in index notation, to see this in a way we may recognise from special relativity. So,

$$\begin{aligned} C_{ij} &= (UCU^T)_{ij} \\ &= U_{ia}C_{ab}U^T_{bj} \\ &= U_{ia}C_{ab}U_{jb} \\ &= U_{ia}U_{jb}C_{ab}. \end{aligned}$$

Hence, we have that $C_{ij} = U_{ia}U_{jb}C_{ab}$. Now, the RHS of this is the transformation rule for a Cartesian tensor of rank-2. Usually, this would give a new matrix C'_{ij}, but, we have that $C'_{ij} = C_{ij}$, which is the statement of invariance. In special relativity, we would say that the U_{ij} have the same role as the Jacobians $J^\mu{}_\nu$.

As we have seen, the upper and lower indices are related via

$$\psi^a = C^{ab}\psi_b,$$

and $C_{ab} = \epsilon_{ab}$, where ϵ_{ab} is the completely anti-symmetric tensor. That is, we define its elements to be such that

$$\epsilon_{ij} = \begin{cases} 1 & \text{even permutation of } i, j, \\ -1 & \text{odd permutation e.g. } i = 2, j = 1, \\ 0 & \text{else, e.g. } i = j = 1. \end{cases}$$

Hence, we shall write

$$\psi^a = \epsilon^{ab}\psi_b, \tag{4.31}$$

In writing the relation like this, we now see an analogue from special relativity, where we change between upper and lower indices via the metric $x^\mu = g^{\mu\nu}x_\nu$ (also recall that in special, *not general*, relativity, the metric $g_{\mu\nu}$ and inverse metric $g^{\mu\nu}$ have the same components).

Just like we form invariants (i.e. scalars) in special relativity, $g_{\mu\nu}x^\mu x^\nu = x^\mu x_\mu$ (from vectors), we can form invariants from spinors. That is, we contract upper and lower indices to form an invariant;

$$\text{invariant scalar :} \quad \psi^a\phi_a.$$

Let us prove that this construction is infact an invariant. Consider a transformation

$$\psi^a\phi_a \longmapsto \psi'^a\phi'_a.$$

We now use (4.31) to express ψ'^a in terms of a lower-index spinor, so that

$$\psi'^a\phi'_a = \epsilon^{ab}\psi'_b\phi'_a.$$

We can then use (4.23) on both of the primed spinors to transfer to the unprimed frame

$$\psi'^a\phi'_a = \epsilon^{ab}U_b{}^c\psi_c U_a{}^d\phi_d.$$

Now, notice that the first and second terms on the RHS form the ac^{th} element of the matrix ϵU, so that

$$\psi'^a\phi'_a = (\epsilon U)^{ac}\psi_c U_a{}^d\phi_d.$$

Now,

$$U_a{}^d = (U^T)^d{}_a,$$

so that

$$\psi'^a\phi'_a = (\epsilon U)^{ac}\psi_c(U^T)^d{}_a\phi_d,$$

where we now notice that the first and third terms form the cd^{th} element of the matrix $U^T\epsilon U$, so that

$$\psi'^a\phi'_a = (U^T\epsilon U)^{cd}\psi_c\phi_d.$$

Finally, we use (4.30) to see that

$$\psi'^a \phi'_a = \epsilon^{cd} \psi_c \phi_d = \psi^d \phi_d.$$

Therefore, we have proved that $\psi^a \phi_a$ is an $SU(2)$-invariant scalar. In this proof we have been careful about matching upper and lower indices, as in Lorentz indices, purely for clarity.

Hence, just to summarise a little; we have that U is a transformation matrix (just like the Jacobian from special relativity), and C is a matrix that sends lower index spinors to upper index spinors (just like the metric of special relativity). We make the identification $C_{ij} = \epsilon_{ij}$, which is the rank-2 version of the Levi-Civita anti-symmetric tensor.

4.6.2 Constructing Higher Dimensional Irreps & Clebsch-Gordan Series

Let us combine two spinors ψ_a and ϕ_b,

$$(\chi)_{ab} = \psi_a \phi_b.$$

Then, as each of the spinors ψ_a and ϕ_b have two components, the resultant object χ will be a four component tensor. To write the irreps of the decomposition, we shall symmetrise and antisymmetrise the 4-component tensor. That is, we write

$$\psi_a \phi_b = \frac{1}{2} \left(\psi_{(a} \phi_{b)} + \psi_{[a} \phi_{b]} \right), \tag{4.32}$$

where the symmetric and antisymmetric terms are

$$\psi_{(a} \phi_{b)} = \psi_a \phi_b + \psi_b \phi_a, \tag{4.33}$$
$$\psi_{[a} \phi_{b]} = \psi_a \phi_b - \psi_b \phi_a. \tag{4.34}$$

Notice that the symmetric tensor has 3 components:

$$\text{symmetric}: \quad \psi_1 \phi_1, \quad \psi_2 \phi_2, \quad \psi_1 \phi_2 + \psi_2 \phi_1.$$

These correspond to the cases $|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, |\uparrow\downarrow + \downarrow\uparrow\rangle$ (in terms of spin). The antisymmetric tensor has only one component

$$\text{antisymmetric}: \quad \psi_1 \phi_2 - \psi_2 \phi_1,$$

which corresponds to $|\uparrow\downarrow - \downarrow\uparrow\rangle$. This single component "tensor" is infact a scalar (a tensor with a single component is obviously a scalar). Hence, we say that the antisymmetric tensor is an $SU(2)$ scalar;

$$\psi_a \phi^a = \epsilon^{ab} \psi_a \phi_b.$$

The general rule is that the symmetric and antisymmetric parts of the rank-2 tensor transform independently. That is, they form invariant subspaces.

Although we shall explain this more later on, we pictorially depict this as

$$\boxed{a}\otimes\boxed{b}=\begin{array}{|c|c|}\hline a & b\\\hline\end{array}\oplus\begin{array}{|c|}\hline a\\\hline b\\\hline\end{array}.$$

The first box denotes ψ_a, the second box ϕ_b. The third term, the row of two boxes, denotes the symmetric component; the fourth term, the column of two boxes, denotes the antisymmetric component. Thus, we see that $\begin{array}{|c|c|}\hline a & b\\\hline\end{array}$ has three components, or, is a [3]-dim irreducible representation. Similarly, $\begin{array}{|c|}\hline a\\\hline b\\\hline\end{array}$ is a [1]-dim scalar. Also, as $\boxed{a}$ represents ψ_a, then $\boxed{a}$ is a [2]-dim spinor. So, let us write this decomposition again, without the labels within the boxes, and with the relevant dimensions below. So, for $SU(2)$,

$$\boxed{}\otimes\boxed{} = \begin{array}{|c|c|}\hline & \\\hline\end{array}\oplus\begin{array}{|c|}\hline \\\hline \\\hline\end{array} \tag{4.35}$$

$$2\otimes 2 = 3\oplus 1. \tag{4.36}$$

This is the Clebsch-Gordan series for the composition of two spinors.

Let us now consider the composition of three spinors,

$$2\otimes 2\otimes 2 : \psi_a\phi_b\chi_c.$$

First, we can symmetrise and antisymmetrise the a and b indices, to get

$$\psi_{(a}\phi_{b)}\chi_c, \quad \psi_{[a}\phi_{b]}\chi_c.$$

We saw that for the two-spinor composition a completely antisymmetric tensor $\psi_{[a}\phi_{b]}$ is just a [1]-dim scalar. Thus, the second term above, $\psi_{[a}\phi_{b]}\chi_c$ is a [2]-dim spinor. For the first term above, $\psi_{(a}\phi_{b)}\chi_c$, we can symmetrise or antisymmetrise with respect to c. If we antisymmetrise, we create a one-index object which is again a spinor. If we symmetrise all three indices, we get a completely symmetric tensor, $\xi_{(abc)}$, say. So, to summarise, we first decompose the two spinor object,

$$\psi_a\phi_b \longrightarrow \psi_{(a}\phi_{b)}+\psi_{[a}\phi_{b]},$$

then we compose this with the remaining spinor, χ_c,

$$\begin{aligned}\psi_a\phi_b\chi_c \longrightarrow \quad & \left(\psi_{(a}\phi_{b)}+\psi_{[a}\phi_{b]}\right)\chi_c\\ & =\psi_{(a}\phi_{b)}\chi_c+\psi_{[a}\phi_{b]}\chi_c.\end{aligned}$$

In the second term, $\psi_{[a}\phi_{b]}$ is a scalar, and so $\psi_{[a}\phi_{b]}\chi_c$ is a spinor. In the first term, we either form a totally symmetric tensor, $\xi_{(abc)}=\psi_{(a}\phi_b\chi_{c)}$, or antisymmetrise a or b, with respect to c. Hence, this will produce a spinor. It is worth noting that $\xi_{(abc)}$ has four components,

$$(\xi_{111},\xi_{112},\xi_{221},\xi_{222},) \quad\Rightarrow\quad (|\uparrow\uparrow\uparrow\rangle, |\uparrow\uparrow\downarrow\rangle, |\downarrow\downarrow\uparrow\rangle, |\downarrow\downarrow\downarrow\rangle),$$

where we have given the correspondence to spin. Notice that as $|\uparrow\rangle$ corresponds to spin $+\frac{1}{2}$, and $|\downarrow\rangle$ to spin $-\frac{1}{2}$, the components of the tensor correspond to spins $+\frac{3}{2},+\frac{1}{2},-\frac{1}{2},-\frac{3}{2}$.

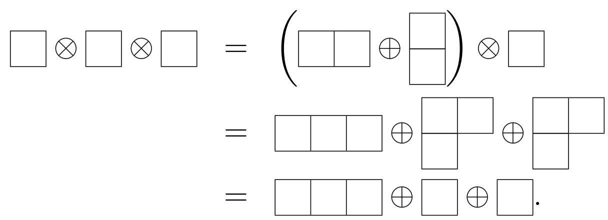
Hence, we see that we have found out where the rule for adding spins comes from. So, we have formed the spin compositions,

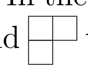
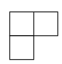
$$\begin{aligned}\frac{1}{2}\oplus\frac{1}{2}\oplus\frac{1}{2} &= (1\oplus 0)\oplus\frac{1}{2}\\ &= \frac{3}{2}\oplus\frac{1}{2}\oplus\frac{1}{2},\end{aligned}$$

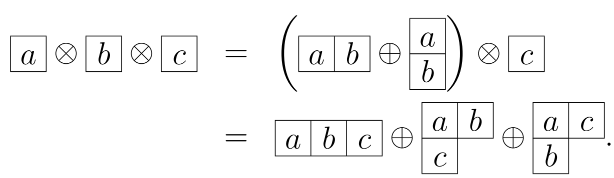
which correspond in terms of dimension to

$$\begin{aligned}2\otimes 2\otimes 2 &= (3\oplus 1)\otimes 2\\ &= 4\oplus 2\oplus 2.\end{aligned}$$

Pictorially, in terms of our little boxes, we have



To explain this: in the first equality we symmetrised and antisymmetrised the first two boxes (exactly as we did in our two spinor case). In the second equality we then either symmetrised or antisymmetrised (which gave the row and  terms respectively), and we also get a spinor. In the third equality we notice that  is a scalar times a spinor (the column bit times the single box on the right). Hence, we remove the scalar part to leave the spinor. Just to note, the $\boxed{}\boxed{}\boxed{}$-term corresponds to the completely symmetric tensor $\xi_{(abc)}$. Basically, what we do, is compose things term by term, with any direct-producted boxes being put both alongside and below an existing diagram. Let us repeat the first two lines of this calculation, with labels, so that we can track where boxes go;



More generally speaking, we can say that the irreps are carried by tensors with n symmetric indices $\phi_{(ab\ldots n)}$ with dimensionality $(n+1)$ for $SU(2)$. When we multiply two such tensors, we get a higher dimension representation, which is then reducible;

$$\boxed{m}\otimes\boxed{n}=\boxed{n}\boxed{+}\boxed{m},\quad (m+1)\otimes(n+1)=m+n+1.$$

Here, we have that $\boxed{m}$ denotes a row of m boxes. One by one we start to antisymmetrise, whereby antisymmetrising causes us to loose two boxes (equivalently, two indices). Then general result is a sum

$$\sum_{\oplus\ell=m-n}^{m+n}\boxed{\ell},$$

where the sum increments in steps of two (as we loose two indices on each antisymmetrisation). If we identify $m = 2j_1$ and $n = 2j_2$, then the above rule is just

$$(2j_1 + 1) \otimes (2j_2 + 1) = \sum_{\oplus j=|j_1-j_2|}^{j_1+j_2} 2j + 1,$$

which is the rule for adding spins. Infact, being more specific and taking $j_1 = j_2 = \frac{1}{2}$, then this is just

$$2 \otimes 2 = 3 \oplus 1;$$

Again, a result we have previously stated.

4.6.3 $SU(3)$ Tensors

In $SU(3)$, the basic object is a three component tensor. Continuing our analogy with particle physics (i.e. in $SU(2)$ we had each of the two components of the spinor corresponding to spin up and down), we represent the tensors as

$$\psi_a = \begin{pmatrix} r \\ g \\ b \end{pmatrix},$$

where r, g, b represent colour. This tensor transforms as

$$\psi'_a = U_{ab}\psi_b,$$

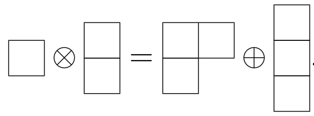
which is just like the $SU(2)$ tensor (recall that a tensor with lower indices corresponds to particles, and with upper indices to antiparticles). Here, the complex conjugate representation corresponding to the rotation of antiparticles is not an equivalent representation to that for particles. The complex conjugate representation transforms not like a lower spinor, but as a pair of antisymmetric lower indices. Hence, a quark is represented as $\square$ and an anti-quark as $\begin{smallmatrix}\square\\\square\end{smallmatrix}$. There is still an invariant tensor, ϵ^{abc}, with $\epsilon^{123} = 1$. This invariant tensor converts an antisymmetric pair of lower indices to an upper index,

$$\psi^a = \epsilon^{abc}\phi_{[bc]}. \tag{4.37}$$

One can form scalar invariants, as before, by contraction with ϵ^{abc},

$$\chi_a\psi^a = \epsilon^{abc}\chi_a\phi_{[bc]}. \tag{4.38}$$

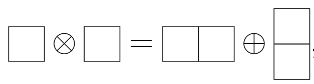
So, we still break a muti-index tensor into irreps by symmetrising/antisymmetrising with the ϵ tensor. Let us combine a quark transforming as a fundamental [3]-dim representation of $SU(3)$ with an anti-quark transforming in the conjugate representation $\bar{3}$. Essentially, what we are about to do, is to combine $q\bar{q}$, and decompose into irreps. So, combining the quark and anti-quark,



The first term on the RHS is a mixed symmetry object, and is of [8]-dim (more on how to compute this dimension, later); the second term on the RHS is a totally antisymmetric tensor (i.e. a scalar), $\epsilon^{abc}\chi_a\phi_{[bc]}$. So, in terms of dimension, this decomposition is

$$3 \otimes \bar{3} = 8 \oplus 1.$$

We will note, and later comment on the significance of, that there is a singlet present in this decomposition (i.e. the scalar with unit dimension). We can also compose qq, to find



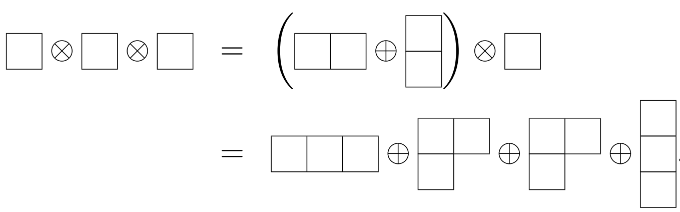
or, in terms of dimension,

$$3 \otimes 3 = 6 \oplus \bar{3}.$$

This time, we note that there is no singlet present.

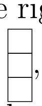
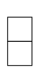
If a singlet state is formed in the decomposition, then the system can be observed (i.e. found) in nature. Given our two above examples, we see that $q\bar{q}$ can be observed, but qq cannot (something which has been verified by experiment).

We can add a third quark,

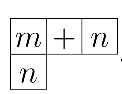


In terms of dimension, this decomposition is

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1.$$

Let us make some observations. Firstly, the dimension on the LHS is $3 \times 3 \times 3 = 27$, and on the right is $10 + 8 + 8 + 1 = 27$ (i.e. they match). Secondly, in $SU(3)$, a scalar is represented by , and not by  as in $SU(2)$; this is because we need complete antisymmetrisation for a scalar.

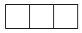
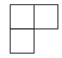
The irreps of $SU(3)$ are carried by diagrams (the Young Tableaux "boxes" that we have been using) with $(m + n)$ upper boxes and m lower boxes;

$SU(3)$ irreps : 

. (4.39)

The dimensionality of such a diagram is

$$\text{dim}: \quad \frac{1}{2}(m+1)(n+1)(m+n+2). \tag{4.40}$$

So, for example,  has $m = 3, n = 0$, and hence has dimension $\frac{1}{2}(3+1)(0+1)(3+0+2) = 10$, as stated above. Also, a diagram such as  has $m = n = 1$, and thus dimension $\frac{1}{2}(1+1)(1+1)(1+1+2) = 8$ (again, as stated above).

4.7 $SU(N)$ and Young Tableaux

So far we have looked at combining spinors in various dimensions, to build composite tensor objects; we then briefly looked at how to decompose the resulting object into its irreps (this effectively will break up a, say, 56×56 transformation matrix into a set of lower dimension transformation matrices, which will be easier to handle). We formed $\psi_a \phi_b \chi_c \ldots$, and asked how the composite object transformed under a transformation matrix which was an element of the $SU(N)$ group. The general result is a direct sum over the irreps,

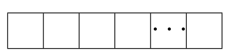
$$D(\psi_a) \otimes D(\phi_b) \otimes D(\chi_c) = \sum_{\oplus \sigma} a_\sigma D^\sigma,$$

which is the Clebsch-Gordan decomposition.

For $SU(2)$, an antisymmetric combination of two spinors was a scalar,

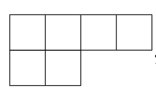
$$\epsilon^{ab} \psi_a \phi_b = \psi_a \phi^a = \text{scalar};$$

and hence the irreps of $SU(2)$ are carried by tensors that are symmetric on all indices, which we represent by a series of boxes in one row

$SU(2)$:  $\phi_{(ab\ldots)}$ $\quad \text{dim} = n+1$,

where n is the number of indices (i.e. boxes).

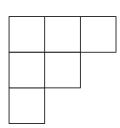
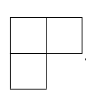
In $SU(3)$, the irreps are carried by diagrams with up to two rows,

$SU(3)$: ,

and the totally anti-symmetric tensor construction

$$\epsilon^{abc} \psi_a \phi_b \chi_c = \text{scalar}.$$

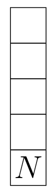
Hence, in $SU(3)$, any column with three boxes is a scalar, so that we just erase the column,

$SU(3)$:  $=$ .

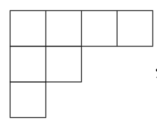
The dimensionality of a diagram in $SU(3)$, with $(m+n)$ upper boxes and m lower boxes is

$$SU(3): \quad \dim = \frac{1}{2}(m+1)(n+1)(m+n+2).$$

Thus, we say that in $SU(N)$, a column with N boxes is a scalar,

$SU(N)$:  = scalar.

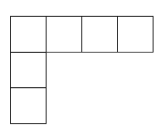
General Young Tableaux for $SU(N)$ A Young tableau is an arrangement of boxes as below,

 ,

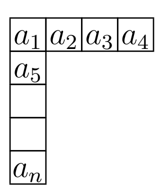
such that the following rules are adhered to:

- each row contains no more boxes that the one above it,
- the number of rows should not exceed N for a diagram in $SU(N)$, so that one cannot have $N+1$ boxes in any given column,
- every Young tableau represents a tensor $\psi_{a_1 \ldots a_n}$ with n indices, each of which can take N values for $SU(N)$,
- the Young tableau represents this tensor after a definite process of symmetrisation and antisymmetrisation.

The correspondence between a tableau and tensor can be found as follows. Suppose we have a tableau



with n boxes, then, the tensor that this corresponds to can be found by first filling the indices into the boxes working from left to right, and top to bottom, in an obvious manner.



We then must apply two operations. First, symmetrisation of indices in each row and column, and sum over rows. This is achieved by an operator P which denotes sum of permutations of indices in each row and then sum over rows. The second operation is antisymmetrisation of all indices in a column and sum over column; an operation denoted by the operator Q. Once the operation QP has been applied to the tableau, the diagram represents the tensor $\psi_{a_1 \ldots a_n}$.

Let us consider a few examples; the process is much easier done than said!

First consider the diagram

$$\begin{array}{|c|c|c|}\hline a_1 & a_2 & a_3 \\ \hline \end{array},$$

then, the tensor is just a symmetric sum

$$\psi_{(a_1 a_2 a_3)} = \psi_{a_1 a_2 a_3} + \psi_{a_1 a_3 a_2} + \ldots$$

Second, consider the diagram

$$\begin{array}{|c|c|}\hline a_1 & a_2 \\ \hline a_3 \\ \cline{1-1} \end{array}$$

and the tensor $\psi_{a_1 a_2 a_3}$. First, we apply P,

$$P(\psi_{a_1 a_2 a_3}) = \psi_{a_1 a_2 a_3} + \psi_{a_2 a_1 a_3},$$

where we symmetrise in the rows a_1, a_2, and secondly apply Q,

$$QP(\psi_{a_1 a_2 a_3}) = \psi_{a_1 a_2 a_3} - \psi_{a_3 a_2 a_1} + \psi_{a_2 a_1 a_3} - \psi_{a_2 a_3 a_1},$$

where this result is the tensor that corresponds to the tableau.

Thirdly, consider the diagram

$$\begin{array}{|c|}\hline a_1 \\ \hline a_2 \\ \hline a_3 \\ \hline \end{array}$$

then the tensor is just $\psi_{[a_1 a_2 a_3]}$.

4.7.1 Dimensionality of Young Tableau in $SU(N)$

The dimensionality of a Young tableau can be obtained as follows:

- Fill the tableau with numbers starting with N for $SU(N)$ at the top left, and increase by 1 for each successive column, and decrease by 1 for each successive row. Consider for example a diagram in $SU(5)$,

  $$\begin{array}{|c|c|c|c|}\hline 5 & 6 & 7 & 8 \\ \hline 4 & 5 & 6 \\ \cline{1-3} 3 & 4 \\ \cline{1-2} \end{array}$$

  then work out the product of all numbers, and call it $\mathcal{N}$.

- Fill the tableau with numbers again, but this time the entry should be the number of boxes to the right and below, plus one for itself. Again, our example in $SU(5)$ would have

$$\begin{array}{|c|c|c|c|}\hline 6 & 5 & 3 & 1 \\ \hline 4 & 3 & 1 & \\ \cline{1-3} 2 & 1 & & \\ \cline{1-2}\end{array}$$

 then work out the product of all entries, call it $\mathcal{D}$.

- The dimensionality is then just $\dim = \mathcal{N}/\mathcal{D}$;

Let us do some examples.

Consider the following diagram in $SU(2)$, with n boxes in a line

$$SU(2): \quad \begin{array}{|c|c|c|c|c|c|}\hline \ & \ & \ & \ \cdots & \ & n \\ \hline\end{array}.$$

Then, we compute $\mathcal{N}$ by putting N (i.e. 2) in the far left box, and incrementing (i.e. the entries will just be $2, 3, 4, \ldots, n+1$); we then multiply all these numbers, which will just give $\mathcal{N} = (n+1)! = n!(n+1)$. We compute $\mathcal{D}$ by writing in each box how many are to the right and below, plus one. Hence, the entries are just $n, n-1, \ldots, 1$, and thus $\mathcal{D} = n!$. Hence the dimension of such a diagram is $\mathcal{N}/\mathcal{D} = n+1$, which we knew already! So, the dimension of $\begin{array}{|c|c|}\hline \ & \ \\ \hline\end{array}$ is 3 and $\begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline\end{array}$ is 4.

Let us consider a column of N boxes, in $SU(N)$,

$$SU(N): \quad \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline \ \\ \hline N \\ \hline\end{array}.$$

Then, $\mathcal{N}$ is just $N!$ (i.e. we put $N, N-1, \ldots, 1$ in the boxes, and multiply together), and $\mathcal{D} = N!$. Therefore, this diagram has dimension $N!/N! = 1$, which is a scalar; again, something we knew.

Now consider

$$SU(6): \quad \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline\end{array}.$$

So, we first write

$$\begin{array}{|c|c|c|}\hline 6 & 7 & 8 \\ \hline\end{array} \quad \Rightarrow \quad \mathcal{N} = 6 \times 7 \times 8,$$

and then we write

$$\begin{array}{|c|c|c|}\hline 3 & 2 & 1 \\ \hline\end{array} \quad \Rightarrow \quad \mathcal{D} = 3 \times 2 \times 1.$$

Then, the dimensionality is just (where it is generally easier to keep the factors "unmultiplied", so that cancellations are easy to see),

$$\dim = \frac{6 \times 7 \times 8}{3 \times 2 \times 1} = 56.$$

Now let us consider in $SU(3)$, three diagrams, where we shall find something interesting about the dimension of the last two. Firstly,

$$SU(3): \quad \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline \end{array},$$

so that its dimension is

$$\dim = \frac{3 \times 4 \times 5}{3 \times 2 \times 1} = 10.$$

Second, consider

$$SU(3): \quad \begin{array}{|c|c|}\hline \ & \ \\ \hline \end{array},$$

which has dimension

$$\dim = \frac{3 \times 4}{2 \times 1} = 6.$$

Thirdly, consider

$$SU(3): \quad \begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \ \\ \hline \end{array},$$

where we fill numbers as

$$\mathcal{N}: \quad \begin{array}{|c|c|}\hline 3 & 4 \\ \hline 2 & 3 \\ \hline \end{array}, \qquad \mathcal{D}: \quad \begin{array}{|c|c|}\hline 3 & 2 \\ \hline 2 & 1 \\ \hline \end{array},$$

so that the dimension is

$$\dim = \frac{3 \times 4 \times 2 \times 3}{3 \times 2 \times 2 \times 1} = \bar{6}.$$

Therefore, we find that the dimensionality of $\begin{array}{|c|c|}\hline \ & \ \\ \hline \end{array}$ and $\begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \ \\ \hline \end{array}$, in $SU(3)$, are the same. One will notice that we labelled the first just as 6, and the second with an over-bar, $\bar{6}$. The over-bar denotes the conjugate representation. A similar thing occurs with the diagrams $\begin{array}{|c|}\hline \ \\ \hline \end{array}$ and $\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \end{array}$ in $SU(3)$, which have dimension 3 and $\bar{3}$ respectively.

It is worth noting the dimensionalities of various diagrams in $SU(3)$, as we may come across them regularly;

$$\begin{array}{|c|}\hline \ \\ \hline \end{array} = 3, \qquad \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline \end{array} = 10,$$
$$\begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \end{array} = \bar{3}, \qquad \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline \end{array} = 1,$$
$$\begin{array}{|c|c|}\hline \ & \ \\ \hline \ \\ \cline{1-1} \end{array} = 8, \qquad \begin{array}{|c|c|}\hline \ & \ \\ \hline \end{array} = 6. \tag{4.41}$$

We know that the diagram in $SU(N)$ with dimension N is a single box. Then, we could ask what the diagram is that has dimension $\bar{N}$. The diagram is simply a vertical column of $(N-1)$-boxes – a result which is very easy to prove by just considering that $N!/(N-1)! = N$.

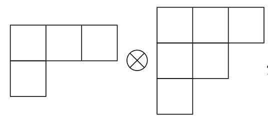
Now we know how to compute the dimensionality of any diagram in $SU(N)$, let us discuss the rules by which diagrams can be combined.

4.7.2 Forming the Clebsch-Gordan Series

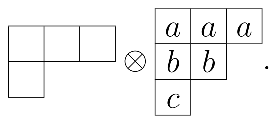
We want to use our Young tableau formalism to combine irreps of $SU(N)$, such as $8 \otimes 8$ in $SU(3)$, and work out the decomposition of the composite system into irreps (i.e. the Clebsch-Gordan series). We do so by following a set of rules (as we did for finding the dimension of a given diagram).

1. First, write down the tableau we are taking the products of, for example

   $$T_1 \otimes T_2 \quad \Longleftrightarrow$$

   and label successive rows of T_2 with indices,

2. We then add boxes from the first row of T_2, one at a time, to T_1; once all of the first row, then second, then third etc, such that

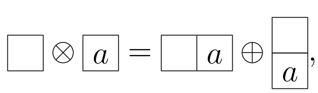
   - the augmented T_1 diagram is always a legal Young tableau,
   - boxes with the same label, e.g. a, must not appear in the same column of the augmented diagram,
   - if n_a is the number of boxes with label a, n_b the number of b, above and to the right of any box in the augmented diagram, then $n_a \geq n_b \geq n_c \geq \ldots$ (this includes everything in the upper-right quadrant).

3. If there are multiple tableau of the same shape, one keeps all except those where the labeling is identical (i.e. only one copy of each absolutely identical tableau are left).

4. Excise columns from any diagrams in $SU(N)$, which have N boxes, as they correspond to scalars. Also note that there should not be any columns with more than N boxes, due to their illegality.

Again, this is much easier done than said, so let us do a few examples.

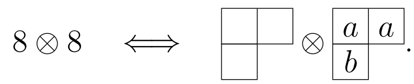
First consider the easy $3 \otimes 3$ in $SU(3)$ (which could correspond to bringing together two quarks, and the decomposition will tell us what the combined colour state is). So, their tableau are just



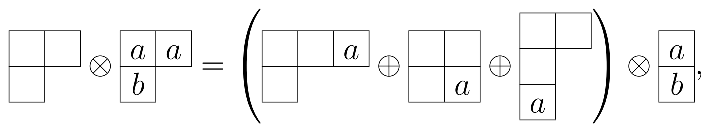
where we have jumped to the answer due to its triviality. In terms of dimensionality, this is

$$3 \otimes 3 = 6 \oplus \bar{3}.$$

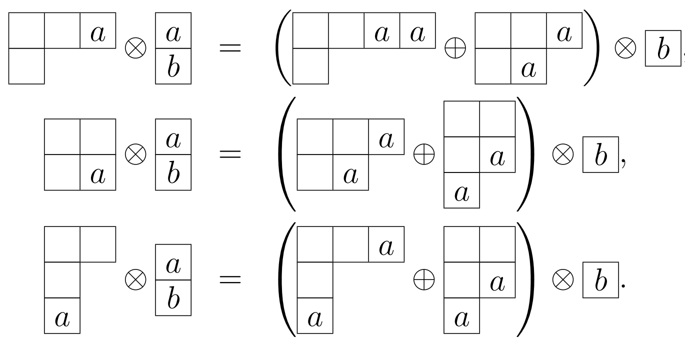
Let us now consider a non-trivial example. Consider $8 \otimes 8$ in $SU(3)$. Their tableau are



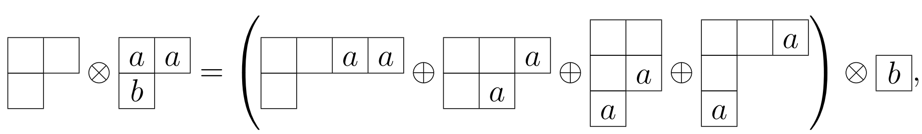
Let us first take the far right $\boxed{a}$-box, and put it in each of the legal places on the left-hand-diagram,



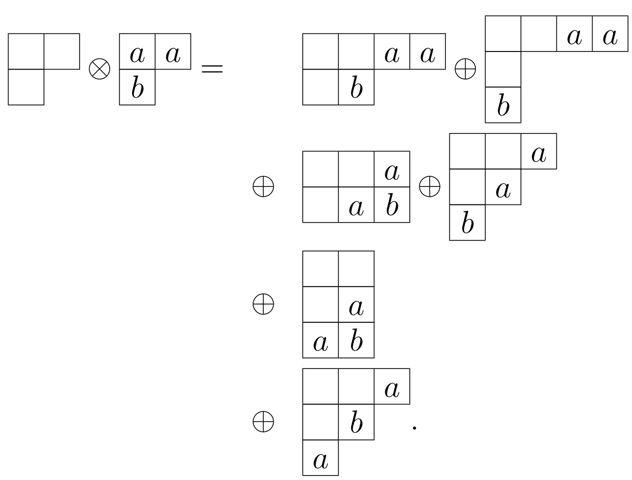
then we put the last $\boxed{a}$ in each of the legal positions of each of the three bracketed diagrams.



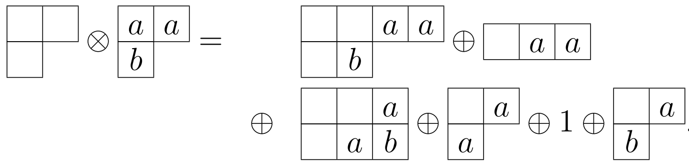
Before we continue, we notice that there are a number of duplicate diagrams. The second in the first row, and first in the second row, are the same (hence we remove one of them), and the second from the second and third rows. Hence, after removing duplicates we have



and then putting $\boxed{b}$ in all possible places,

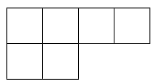


Finally, we excise all columns with 3 boxes, as they corresponds to scalars; hence,



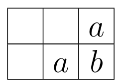
(4.42)

Let us compute the dimensions for those diagrams we did not state in (4.41). Firstly,



$$: \quad \dim = \frac{3 \times 4 \times 5 \times 6 \times 2 \times 3}{5 \times 4 \times 2 \times 1 \times 2 \times 1} = 27.$$

Secondly,

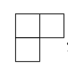


$$: \quad \dim = \frac{3 \times 4 \times 5 \times 2 \times 3 \times 4}{4 \times 3 \times 2 \times 3 \times 2 \times 1} = \bar{10},$$

where we decide that this is the conjugated representation merely by convention. Hence, we have the dimensionalities of all diagrams in (4.42), so that we can write

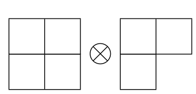
$$8 \otimes 8 = 27 \oplus 10 \oplus \bar{10} \oplus 8 \oplus 8 \oplus 1. \tag{4.43}$$

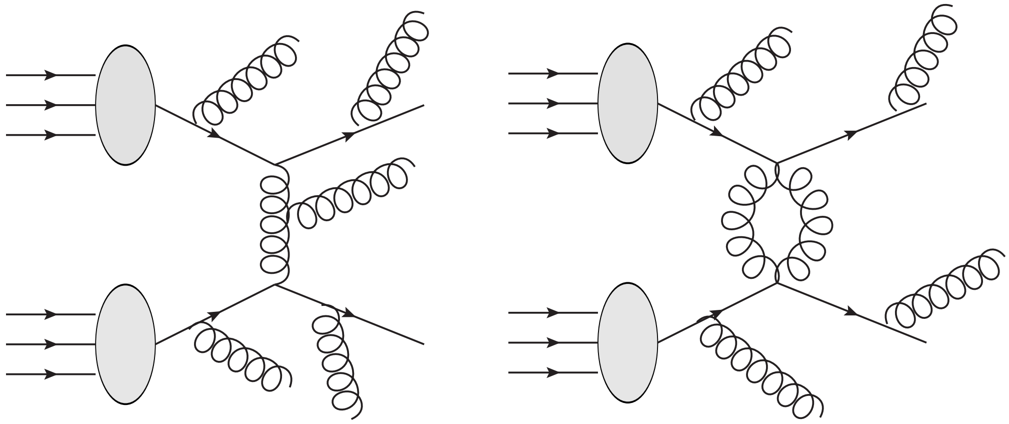
Therefore, we have the Clebsch-Gordan decomposition (4.42), and the dimensions of the various terms involved. Notice that if we had not done the decomposition, all we could say about the resulting system is that it transformed via a $8 \times 8 = 64$-dimensional matrix – instead, we have found that the resulting system has 6 invariant subspaces, each of which transform under a matrix of different dimension.

As something of a sideline, the physical manifestation of the $SU(3)$ octet, , is the gluon. Therefore, the "physics-content" of (4.42) and (4.43) is that a pair of gluons can be in a variety of colour states, including a singlet.

With reference to Figure (4.5), we see two particle-physics events. The first only is a single gluon exchange, which radiates gluons everywhere. The second event is a gluon-pair exchange, which can exist by (4.43), in a colour singlet state, so that there is no gluon radiation between the jets; such a clean channel is a promising event candidate for new physics at the LHC at CERN.

Let us consider a second non-trivial example, where this time we shall present the whole calculation as one. Let us, in $SU(3)$, compose

$SU(3)$:  $\iff \bar{6} \otimes 8.$

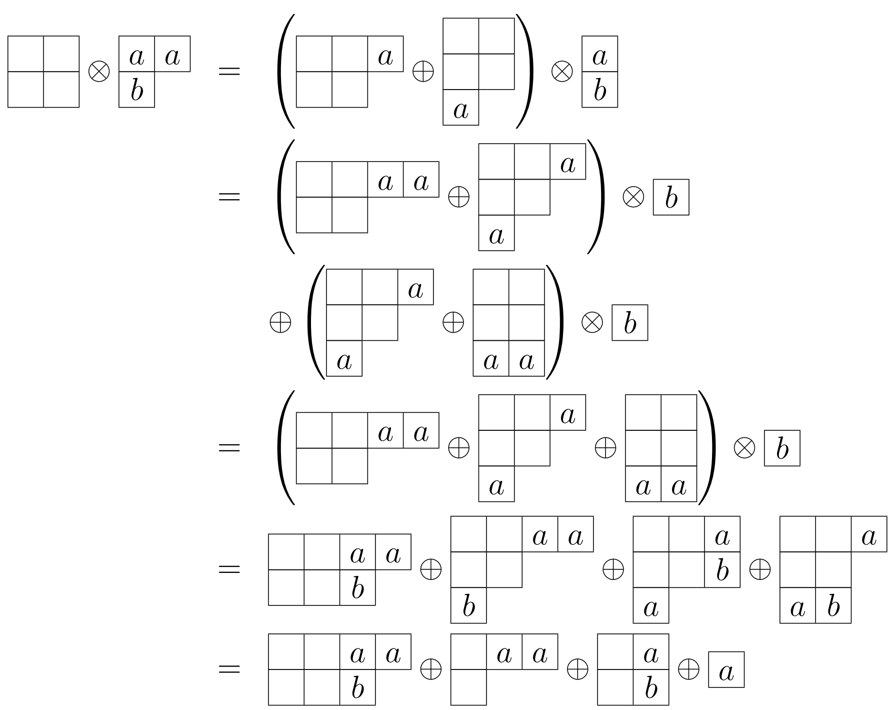


(a) Radiation everywhere.

(b) Gap between jet events, for gluon-pair exchange

Figure 4.5: Two possibilities for qq-scattering from, for example, two incoming protons. In the first process there is gluon radiation everywhere, including between the gluon-exchange region. The second process is much "cleaner", as the two-gluon interaction between the qq is a singlet state, hence no radiation from this region is possible.

So,



or, in terms of dimension,

$$\bar{6} \otimes 8 = 24 \oplus \bar{15} \oplus \bar{6} \oplus 3.$$

This somewhat concludes our discussion on Young tableaux, and we shall now move onto discussing some of the general points of Lie algebras

4.8 Lie Algebras

The generators of groups form an algebra, or a vector space, closed under addition and commutation. As we have seen, the commutator of two generators is itself a generator,

$$[X_\alpha, X_\beta] = i f_{\alpha\beta\gamma} X_\gamma, \tag{4.44}$$

where the X_α are the generators, and the $f_{\alpha\beta\gamma}$ are the structure constants of the group, which can be made antisymmetric on all three indices. Now, the structure constants themselves generate a representation of the group, and is known as the adjoint representation. This commutation relation describes a Lie algebra.

Figure 4.6: Gluons (i.e. on the left) interact using the adjoint representation, using $f_{\alpha\beta\gamma}$, whereas quarks interact using the fundamental representation of $SU(3)$, via the Gell-mann matrices $(t^a)_{ij}$.

For example, gluons transform in the adjoint representation, using the structure factors; whereas quarks transform in the fundamental representation; as in Figure (4.6).

4.8.1 The Adjoint Representation

We can define matrices

$$(T_\alpha)_{\beta\gamma} = -i f_{\alpha\beta\gamma}, \tag{4.45}$$

where $-i f_{\alpha\beta\gamma}$ is the $\beta\gamma^{\text{th}}$-component of the matrix T_α. These matrices generate the adjoint representation, and hence satisfy the commutator

$$[T_\alpha, T_\beta] = i f_{\alpha\beta\gamma} T_\gamma. \tag{4.46}$$

The structure factors satisfy the Jacobi identity,

$$f_{\beta\gamma\delta} f_{\alpha\delta\epsilon} + f_{\alpha\beta\delta} f_{\gamma\delta\epsilon} + f_{\gamma\alpha\delta} f_{\beta\delta\epsilon} = 0, \tag{4.47}$$

which corresponds to the commutator (4.46). To prove this, consider the commutator

$$[T_\alpha, [T_\beta, T_\gamma]]\,. \tag{4.48}$$

Now, we can write this in two ways, and equating the two ways will give us the requirement that the structure constants satisfy the Jacobi identity. The first way is to use (4.46) on the "inside" commutator, to write

$$[T_\alpha, [T_\beta, T_\gamma]] = if_{\beta\gamma\delta}[T_\alpha, T_\delta],$$

and use (4.46) again, so that

$$\begin{aligned} if_{\beta\gamma\delta}[T_\alpha, T_\delta] &= if_{\beta\gamma\delta} if_{\alpha\delta\epsilon} T_\epsilon \\ &= -f_{\beta\gamma\delta} f_{\alpha\delta\epsilon} T_\epsilon. \end{aligned} \tag{4.49}$$

The second way is to expand the "inside" commutator of (4.48), directly,

$$\begin{aligned} [T_\alpha, [T_\beta, T_\gamma]] &= [T_\alpha, T_\beta T_\gamma - T_\gamma T_\beta] \\ &= [T_\alpha, T_\beta T_\gamma] - [T_\alpha, T_\gamma T_\beta]. \end{aligned}$$

To go further, we use the identity

$$\begin{aligned} [A, BC] &= ABC - BCA \\ &= ABC - BAC + BAC - BCA \\ &= [A, B]C + B[A, C], \end{aligned}$$

so that we can write

$$\begin{aligned} [T_\alpha, T_\beta T_\gamma] &= [T_\alpha, T_\beta]T_\gamma + T_\beta[T_\alpha, T_\gamma] \\ &= if_{\alpha\beta\delta} T_\delta T_\gamma + if_{\alpha\gamma\delta} T_\beta T_\delta. \end{aligned}$$

Analogously, we have

$$[T_\alpha, T_\gamma T_\beta] = if_{\alpha\gamma\delta} T_\delta T_\beta + if_{\alpha\beta\delta} T_\gamma T_\delta.$$

Hence, we have

$$\begin{aligned} [T_\alpha, [T_\beta, T_\gamma]] &= [T_\alpha, T_\beta T_\gamma] - [T_\alpha, T_\gamma T_\beta] \\ &= if_{\alpha\beta\delta} T_\delta T_\gamma + if_{\alpha\gamma\delta} T_\beta T_\delta - if_{\alpha\gamma\delta} T_\delta T_\beta - if_{\alpha\beta\delta} T_\gamma T_\delta \\ &= if_{\alpha\beta\delta}[T_\delta, T_\gamma] + if_{\alpha\gamma\delta}[T_\beta, T_\delta]. \end{aligned}$$

We then use (4.46) on the remaining commutators, to see that

$$if_{\alpha\beta\delta}[T_\delta, T_\gamma] + if_{\alpha\gamma\delta}[T_\beta, T_\delta] = -f_{\alpha\beta\delta} f_{\delta\gamma\epsilon} T_\epsilon - f_{\alpha\gamma\delta} f_{\beta\delta\epsilon} T_\epsilon.$$

We then equate this to the result of the first method, (4.49), so that

$$-f_{\beta\gamma\delta} f_{\alpha\delta\epsilon} T_\epsilon = -f_{\alpha\beta\delta} f_{\delta\gamma\epsilon} T_\epsilon - f_{\alpha\gamma\delta} f_{\beta\delta\epsilon} T_\epsilon.$$

We can obviously cancel off the T_ϵ, and take everything to a single side,

$$-f_{\alpha\beta\delta}f_{\delta\gamma\epsilon} - f_{\alpha\gamma\delta}f_{\beta\delta\epsilon} + f_{\beta\gamma\delta}f_{\alpha\delta\epsilon} = 0.$$

To get this in the form of the Jacobi identity required, (4.47), we need to remove two minus signs. To do this, we use the antisymmetry of the first two indices $f_{\alpha\beta\gamma} = -f_{\beta\alpha\gamma}$. This follows from (4.46), as swapping the first two indices on f corresponds to reversing the commutator. So, we swap the indices on the second f on the first term and the first f on the first term, giving

$$f_{\alpha\beta\delta}f_{\gamma\delta\epsilon} + f_{\gamma\alpha\delta}f_{\beta\delta\epsilon} + f_{\beta\gamma\delta}f_{\alpha\delta\epsilon} = 0.$$

Hence, we have the Jacobi identity we stated the structure constants must satisfy.

4.8.2 The Casimir Operator

We define the Casimir operator

$$T^2 = \sum_\alpha T_\alpha T_\alpha, \tag{4.50}$$

which is the sum of the squares of the generators, and it commutes with all the generators,

$$[T^2, T_\alpha] = 0. \tag{4.51}$$

Schur's lemma is that $T^2 \propto I$, the identity.

An example of a Casimir operator is the familiar angular momentum operator, $J^2 = J_x^2 + J_y^2 + J_z^2$, from quantum mechanics, whereby

$$[J^2, J_x^2] = [J^2, J_y] = [J^2, J_z] = 0.$$

5 Symmetries of the Lagrangian in QFT

Here we shall discuss the symmetries of the fundamental objects of quantum field theory: Lagrangians. The principle of least action is as applicable to a quantum theory of fields as it is to classical Newtonian theory – one of the very few concepts that bridges the gap, infact. Quantum field theories (QFTs) describe all laws of natures, except gravity. We shall study the symmetries of the Lagrangians of a QFT, where we shall generally impose some sort of invariance after action of an element of a group. We shall look at global symmetries of the Lagrangian, where we rotate the fields by a constant phase (i.e. the phase is independent of its spacetime position), and we will find that such symmetries give rise to conservation laws. The second type of symmetry we will look at is a local phase transformation (also called a gauge transformation), which give rise to forces.

5.1 Maxwell's Equations

To introduce the idea, let us investigate the invariance of Maxwell's equations under given transformations. So, recalling Maxwell's equations,

$$\nabla\cdot\mathbf{E} = \frac{\rho}{\varepsilon_0}, \quad \nabla\cdot\mathbf{B} = 0, \quad \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}, \quad \nabla\times\mathbf{B} = \mu_0\mathbf{J} + \mu_0\varepsilon_0\frac{\partial\mathbf{E}}{\partial t}.$$

We introduce a vector and scalar potential, via

$$\mathbf{B} = \nabla\times\mathbf{A}, \quad \mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \tag{5.1}$$

where the first follows as div curl is automatically zero, and the second follows simply by noting that curl grad is zero. Now, we are free to choose ϕ and $\mathbf{A}$ (i.e. they are not unique), within certain restrictions, without changing the physical fields $\mathbf{E},\mathbf{B}$. Now, as curl grad is zero, we can add the gradient of a scalar onto the vector potential, without changing the magnetic field,

$$\mathbf{A} \longmapsto \mathbf{A}' = \mathbf{A} + \nabla\Omega. \tag{5.2}$$

That is,

$$\mathbf{B}' = \nabla\times\mathbf{A}' = \nabla\times(\mathbf{A}+\nabla\Omega) = \nabla\times\mathbf{A} = \mathbf{B}.$$

Likewise, we can also change the scalar potential,

$$\phi \longmapsto \phi' = \phi - \frac{\partial\Omega}{\partial t}, \tag{5.3}$$

so that

$$\mathbf{E}' = -\nabla\phi' - \frac{\partial\mathbf{A}'}{\partial t} = -\nabla\phi + \frac{\partial}{\partial t}\nabla\Omega - \frac{\partial\mathbf{A}}{\partial t} - \frac{\partial}{\partial t}\nabla\Omega = \mathbf{E}.$$

Hence, by making the simultaneous transformations (5.2) and (5.3), we leave the $\mathbf{E}$ and $\mathbf{B}$ field unchanged. These transformations are called gauge transformations, as the parameter depends upon its spacetime position.

5.1.1 Gauge Choice & the Wave Equation

A standard procedure is to derive a wave equation from Maxwell's equations. Let us substitute (5.1) into Amperes law (i.e. the one concerning the curl of the magnetic field). Hence, doing so trivially gives

$$\nabla \times \nabla \times \mathbf{A} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial}{\partial t}\left(-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}\right).$$

We can then expand the LHS using the vector identity,

$$\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A},$$

and hence we see that

$$\nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial}{\partial t}\left(-\nabla \phi - \frac{\partial \mathbf{A}}{\partial t}\right).$$

We now use the relation $\mu_0 \varepsilon_0 = 1/c^2$, and rearrange to easily see that

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \mu_0 \mathbf{J} - \nabla\left(\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A}\right).$$

Now, we can exploit gauge freedom to redefine the potentials $\mathbf{A}, \phi$ such that the bracketed term on the far RHS vanishes;

$$\frac{1}{c^2}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} = 0. \tag{5.4}$$

This is called the Lorentz gauge. To see how we can do this, suppose we start with potentials $\mathbf{A}, \phi$ that do not satisfy this condition, and let us transform to potentials that do, via the gauge transformations (5.2), (5.3);

$$\begin{aligned} \mathbf{A} \longmapsto \mathbf{A}' &= \mathbf{A} + \nabla \Omega, \\ \phi \longmapsto \phi' &= \phi - \frac{\partial \Omega}{\partial t}. \end{aligned}$$

Now, as we stipulated, $\mathbf{A}'$ and ϕ' do satisfy the Lorentz gauge, so that

$$\frac{1}{c^2}\frac{\partial \phi'}{\partial t} + \nabla \cdot \mathbf{A}' = 0,$$

and inserting our transformation,

$$\frac{1}{c^2}\frac{\partial}{\partial t}\left(\phi - \frac{\partial \Omega}{\partial t}\right) + \nabla \cdot (\mathbf{A} + \nabla \Omega) = 0.$$

Therefore, so that the Lorentz gauge holds, we must have that the scalar that we added, Ω, satisfies

$$\nabla^2 \Omega - \frac{1}{c^2}\frac{\partial^2 \Omega}{\partial t^2} = -\nabla \cdot \mathbf{A} - \frac{1}{c^2}\frac{\partial \phi}{\partial t}. \tag{5.5}$$

Hence, supposing we have chosen a suitable Ω, the wave equation under the Lorentz gauge reads

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathbf{A} = \mu_0 \mathbf{J}. \tag{5.6}$$

We can derive a similar equation for the scalar potential,

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi = \frac{\rho}{\varepsilon_0}. \tag{5.7}$$

Hence, in writing the wave-equations in this way, we have decoupled them, in the Lorentz gauge.

5.1.2 Covariant Formulation

All of our above discussion can be cast into a very elegant covariant form. We introduce a 4-potential and 4-current density,

$$A^\mu = (\phi/c, \mathbf{A}), \quad J^\mu = (\rho c, \mathbf{J}),$$

as well as the partial derivative contravariant 4-vector,

$$\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right).$$

It is also worth noting that the Minkowski metric $\eta^{\mu\nu}$ can be used to convert between covariant and contravariant tensors, $x^\mu = \eta^{\mu\nu}x_\nu$, where we must notice that the indices balance on both sides, and that a repeated index denotes summation. In Minkowski spacetime, the effective difference between the components of contravariant and covariant vectors is a minus sign in the space parts, as the metric is $\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$.

Hence, using this notation, the wave-equations (5.6) and (5.6) can be written as a single covariant equation,

$$\partial_\mu \partial^\mu A^\nu = \mu_0 J^\nu. \tag{5.8}$$

Similarly, the Lorentz gauge (5.4) can be written in covariant form,

$$\partial^\mu A_\mu = 0. \tag{5.9}$$

Also, (5.5) can be written

$$\partial_\nu \partial^\nu \Omega = \partial_\mu A^\mu.$$

Using the definitions given above, one can see that the gauge transformations (5.2), (5.3) can also be written in a covariant form,

$$A_\mu \longmapsto A'_\mu = A_\mu - \partial_\mu \Omega. \tag{5.10}$$

The electric and magnetic fields can be derived from the field strength tensor,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{5.11}$$

where

$$E_i = cF_{0i}, \quad B_i = -\frac{1}{2}\epsilon_{ijk}F_{jk}.$$

One can notice that the gauge transformation (5.10) leaves the field strength tensor (and hence the fields) unchanged;

$$\begin{aligned} F'_{\mu\nu} &= \partial_\mu A'_\nu - \partial_\nu A'_\mu \\ &= \partial_\mu A_\nu - \partial_\mu\partial_\nu\Omega - \partial_\nu A_\mu + \partial_\nu\partial_\mu\Omega \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}. \end{aligned}$$

This provides a good example of how a seemingly arbitrary choice in potential can be used, within a certain transformation that leaves observable quantities unchanged (i.e. invariant).

5.2 Non-relativistic Quantum Mechanics

Let us begin by looking at the role of gauge transformations, first in non-relativistic quantum mechanics, and then in a relativistic quantum field theory.

So, let us start with Schrodinger's equation for a particle moving in an electrostatic potential ϕ (where the particle has charge e). Then, the Schrodinger equation (SE) is just

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)\psi. \tag{5.12}$$

Now, it is worth noting that the wavefunction ψ itself is not a physical quantity – it is the probability density $\psi^*\psi$ which is the physical observable. So, let us consider a change in the phase of ψ ,

$$\psi \longmapsto \psi' = e^{i\alpha}\psi, \tag{5.13}$$

so that the probability density $\psi^*\psi$ and operator expectation values $\langle\psi|\,\hat{O}\,|\psi\rangle$ are left unchanged (i.e. invariant). We can now consider the three possibilities.

First, consider the case where α is a constant. Then, upon substitution of (5.13) into (5.12), writing $\psi = e^{-i\alpha}\psi'$, we see that

$$i\hbar\frac{\partial\psi'}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)\psi',$$

where we have been able to cancel out the common factors of $e^{-i\alpha}$. Therefore, the SE is unchanged by a global phase transformation, as the same equation is satisfied by both the original and transformed wavefunction.

Let us consider the case where the phase is time dependent, $\alpha = \alpha(t)$, so that $\psi = e^{-i\alpha(t)}\psi$, and upon substitution into the SE (5.12), we find

$$i\hbar\frac{\partial}{\partial t}\left(e^{-i\alpha(t)}\psi'\right) = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)e^{-i\alpha(t)}\psi',$$

where we must now differentiate the LHS by the product rule, to find

$$i\hbar e^{-i\alpha(t)}\frac{\partial\psi'}{\partial t} + \hbar\psi' e^{-i\alpha(t)}\frac{\partial\alpha}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)e^{-i\alpha(t)}\psi'.$$

Now we can cancel off the common factor of $e^{-i\alpha(t)}$,

$$i\hbar\frac{\partial\psi'}{\partial t} + \hbar\psi'\frac{\partial\alpha}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi\right)\psi'.$$

Let us now take the "extra bit" we have picked up by the time derivative of the phase factor to the RHS,

$$i\hbar\frac{\partial\psi'}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\left(\phi - \frac{\hbar}{e}\frac{\partial\alpha}{\partial t}\right)\right)\psi'. \tag{5.14}$$

We can absorb this extra term by redefining the potential,

$$\phi' = \phi - \frac{\hbar}{e}\frac{\partial\alpha}{\partial t}, \tag{5.15}$$

so that (5.14) reads

$$i\hbar\frac{\partial\psi'}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + e\phi'\right)\psi'.$$

Hence, by simultaneously making the gauge transformation (5.15) and transforming the phase of the wavefunction, the SE is left invariant.

Thirdly, let us consider the most general case where the phase is dependent upon its spacetime coordinate, $\alpha = \alpha(x^\mu)$. Then, notice that

$$\nabla\psi = \nabla\left(e^{-i\alpha}\psi'\right) = e^{-i\alpha}\nabla\psi' - i\psi'\nabla\alpha. \tag{5.16}$$

This then generates a new gradient of a scalar term, which we know can be incorporated into a redefinition of $\mathbf{A}$ via a gauge transformation. Therefore, in order that we have a gauge invariant SE, we must introduce a vector potential $\mathbf{A}$ which we do via "minimal substitution",

$$-i\hbar\nabla\psi \longmapsto \left(-i\hbar\nabla - e\mathbf{A}\right)\psi, \tag{5.17}$$

which effectively changes the gradient operator. Hence, by introducing this minimal substitution, we see that the extra term we picked up in (5.16) can be absorbed by a transformation;

$$\begin{aligned}\left(-i\hbar\nabla - e\mathbf{A}\right)\psi &= \left(-i\hbar\nabla - e\mathbf{A}\right)\left(e^{-i\alpha}\psi'\right)\\ &= e^{-i\alpha}\left(-i\hbar\nabla - \hbar\nabla\alpha - e\mathbf{A}\right)\psi'\\ &= e^{-i\alpha}\left(-i\hbar\nabla - e\mathbf{A}'\right)\psi',\end{aligned}$$

where we have made the transformation

$$\mathbf{A}' = \mathbf{A} + \frac{\hbar}{e}\nabla\alpha. \tag{5.18}$$

Therefore, the full invariant SE is

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\left(\nabla - \frac{ie}{\hbar}\mathbf{A}\right)^2 + e\phi\right)\psi, \tag{5.19}$$

where we see that we had to modify the differential operator to include a gauge field so that we can absorb with gauge freedom, any gauge transformations. For this reason, we call local gauge phase transformations *gauge transformations*.

5.3 Quantum Fields & Interactions

By way of a quick introduction, quantum field theory is the theory that results from making quantum mechanics consistent with the concepts of special relativity – mainly the equivalence of energy and mass. Fields are useful as the formalism that comes with them easily allows the production of particles from the vacuum – creation operators act upon the vacuum state to produce particles (this sort of formalism simply is not present in a non-relativistic quantum theory), so that the number of particles in a system is not constant.

The fundamental object, from which a "theory" comes, is the Lagrangian,

$$L = \int d^3x\, \mathcal{L}\left(\phi(x), \partial_\mu\phi(x)\right),$$

where $\mathcal{L}$ is the Lagrange density. We shall use the notation that an argument x denotes dependence on the full 4-position vector x^μ and that $\partial_\mu\phi = \partial\phi/\partial x^\mu$. One can make an analogy between $\phi(x)$ in field theory and the position $\mathbf{x}$ of Newtonian theory; and $\partial_\mu\phi$ with the momentum of Newtonian theory. It is quite important to note that once we have written a Lagrangian, for it to be of any use, we must produce observables from it – it is this task which is non-trivial in QFT.

The fundamental principle that tells us what paths systems follow is the action principle, whereby the variation of the action is zero,

$$\delta\int_{t_1}^{t_2} L\, dt = 0.$$

A rather standard derivation produces the equation of motion for a quantum field (i.e. the Euler-Lagrange equations). So, the action is just

$$S = \int d^4x\, \mathcal{L}(\phi, \partial_\mu\phi),$$

and given these dependencies of the Lagrangian density, the variation of the Lagrangian is

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi).$$

Thus, if we put this variation into the action, the variation of the action is just

$$\delta S = \int d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta(\partial_\mu\phi) \right\}.$$

We can commute $\delta(\partial_\mu\phi) = \partial_\mu\delta\phi$, so that

$$\delta S = \int d^4x \left\{ \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) \right\}.$$

Now, we integrate the second term by parts,

$$\int d^4x \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\delta\phi) = \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\right] - \int d^4x \left\{ \delta\phi\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \right\}.$$

Hence, inserting this into the variation of the action, but letting $\delta\phi = 0$ at the boundaries,

$$\delta S = \int d^4x\delta\phi \left\{ \frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} \right\}.$$

If we now suppose that the variation of the action is zero, and noting the arbitrariness of $\delta\phi$, we set the integrand to zero, giving the Euler-Lagrange equations

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0. \tag{5.20}$$

An equivalent derivation in Newtonian dynamics produces an entirely analogous equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}}\right) - \frac{\partial L}{\partial x} = 0.$$

In Newtonian dynamics, one "writes" the Lagrangian for a particle of mass x in potential $V(\mathbf{x})$,

$$L = \frac{1}{2}m\dot{x}^2 - V(\mathbf{x}),$$

so that upon insertion into the equation of motion, one easily finds

$$m\ddot{x} = -V'(x) = F(x),$$

which is just the Newtonian expression for the equation of motion $F = ma$.

We shall deal with three kinds of particle (equivalently, fields – a given type of field will describe a given type of particle): spin zero scalar particles correspond to scalar fields, half-integer spin fermions (such as electrons or quarks) and gauge bosons (such as photons or gluons) which have spin 1. We want to write down Lagrange densities for each of the above "types" and derive the appropriate equation of motion (this process is essentially done with trial and error – one tends to know the equation of motion one wants, such as the Klein-Gordon equation, and so one must deduce the Lagrange density that gives such an equation of motion – however, we shall somewhat present it simply from a "we know the answer" point of view).

5.3.1 Scalar Fields

Let us postulate the Lagrange density

$$\mathcal{L}_{\mathrm{KG}} = \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - \frac{1}{2}m^2\phi^2, \tag{5.21}$$

and insert it into (5.20) to find the resulting equation of motion. So, the first term needs us to compute

$$\begin{aligned}
\frac{\partial\mathcal{L}_{\mathrm{KG}}}{\partial(\partial_\mu\phi)} &= \frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}(\partial_\nu\phi)(\partial^\nu\phi) \\
&= \frac{1}{2}\eta^{\nu\alpha}\frac{\partial}{\partial(\partial_\mu\phi)}(\partial_\nu\phi)(\partial_\alpha\phi) \\
&= \frac{1}{2}\eta^{\nu\alpha}\left(\partial_\alpha\phi\delta^\mu{}_\nu + \partial_\nu\phi\delta^\mu{}_\alpha\right) \\
&= \eta^{\mu\alpha}\partial_\alpha\phi.
\end{aligned}$$

The second term we must compute is

$$\frac{\partial\mathcal{L}_{\mathrm{KG}}}{\partial\phi} = -m^2\phi.$$

Therefore, putting these results into the equation of motion, we find

$$\partial_\mu\eta^{\mu\alpha}\partial_\alpha\phi + m^2\phi = 0,$$

where the first term can obviously be written

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

Hence,

$$\left(\partial^\mu\partial_\mu + m^2\right)\phi = 0. \tag{5.22}$$

Therefore, we have used the Lagrangian (5.21) to derive that the corresponding equation of motion for a field not interacting with other fields, is the Klein-Gordon equation.

5.3.2 The Photon Field

Let us state the equation of motion corresponding to the Lagrange density

$$\mathcal{L}_{\mathrm{M}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{5.23}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

Hence, rather than the scalar ϕ of the Klein-Gordon field, we work with the vector field A^μ. The equation of motion is

$$\partial^\mu F_{\mu\nu} = \partial^\mu \partial_\mu A_\nu - \partial_\nu(\partial^\mu A_\mu) = 0,$$

where the second term is zero in the Lorentz gauge. Hence, in the Lorentz gauge,

$$\partial^\mu \partial_\mu A_\nu = 0, \tag{5.24}$$

which is exactly (5.8), in free space. Although we shall not go into it, in QFT one actually needs an additional term to "fix the gauge",

$$\mathcal{L}_{\text{gauge-fixing}} = -\frac{(\partial^\mu A_\mu)^2}{2\xi},$$

where the choice of ξ is the choice of gauge – this extra piece does not have any physical consequences.

As another example, which we shall work through, let us consider the equations of motion corresponding to the Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + e A_\mu \bar{\psi} \gamma^\mu \psi(x). \tag{5.25}$$

Then, we must run this Lagrangian through the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial A_\mu} - \partial_\nu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} \right) = 0.$$

So, the first term is just

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = e \delta^\mu_\nu \bar{\psi} \gamma^\nu \psi = e \bar{\psi} \gamma^\mu \psi.$$

The second term requires a little more care,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu)} &= -\frac{1}{4} \frac{\partial}{\partial(\partial_\nu A_\mu)} F_{\alpha\beta} F^{\alpha\beta} \\ &= \partial^\mu A^\nu - \partial^\nu A^\mu = -F^{\nu\mu}. \end{aligned}$$

Hence, plugging these into the Euler-Lagrange equation, we have

$$e \bar{\psi} \gamma^\mu \psi + \partial_\nu F^{\nu\mu} = 0,$$

or,

$$\partial_\nu F^{\nu\mu} = j^\mu_{\text{em}}, \tag{5.26}$$

where we have defined

$$j^\mu_{\text{em}} \equiv -e \bar{\psi} \gamma^\mu \psi.$$

Therefore, we have found that the Lagrangian (5.25) gives Maxwell's equations (5.26) with a source (the expression we have given is completely irrespective of gauge).

5.3.3 Dirac Fields

Such a Lagrangian will describe fields/particles with half-integer spin. The Lagrangian is

$$\mathcal{L}_{\mathrm{D}} = \bar{\psi}\left(i\gamma \cdot \partial - m\right)\psi, \tag{5.27}$$

where ψ are 4-component spinors, $\gamma \cdot \partial \equiv \gamma^{\mu}\partial_{\mu}$ and γ^{μ} are the γ-matrices. The Dirac adjoint is $\bar{\psi} \equiv \psi^{\dagger}\gamma^{0}$. The resulting equation of motion is

$$\left(i\gamma \cdot \partial - m\right)\psi = 0, \tag{5.28}$$

which is the Dirac equation.

5.4 Symmetries of the Lagrangian

Let us explore the consequences of a symmetry of a Lagrangian $\mathcal{L}$, under transformations of the fields. Let us begin with global transformations of the fields, induced by a Lie group with generators T_i. Hence, we can recall that such a group has rotation matrix

$$U = e^{i\theta_a T^a}.$$

In the small rotation angle case, $\theta_a \ll 1$, we can make a simple expansion of the rotation matrix,

$$U = \left(1 + i\theta_a T^a\right), \quad \theta_a \ll 1. \tag{5.29}$$

Now, if we have that

$$\phi \longmapsto \phi' = U\phi,$$

or, in component notation,

$$\phi_i \longmapsto \phi_i' = U_{ij}\phi_j;$$

and that the new field is the old field plus a small part,

$$\phi_a' = \phi_a + \delta\phi_a,$$

then via (5.29),

$$\delta\phi_i = i\theta_a \left(T^a\right)_{ij}\phi_j. \tag{5.30}$$

Now, suppose we demand that $\mathcal{L}(\phi_i, \partial_\mu\phi_i)$ should be invariant under this transformation, then the variation $\delta\mathcal{L} = 0$. Given the dependencies on ϕ_i and $\partial_\mu\phi_i$, we have

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta(\partial_\mu\phi_i).$$

Writing $\delta(\partial_\mu \phi_i) = \partial_\mu(\delta\phi_i)$, we have

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial_\mu(\delta\phi_i). \tag{5.31}$$

Suppose we wrote

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)\delta\phi_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial_\mu(\delta\phi_i),$$

then we can rearrange to find

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial_\mu(\delta\phi_i) = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)\delta\phi_i.$$

So, inserting this for the second term on the RHS of (5.31), we see that

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\phi_i}\delta\phi_i + \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right) - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)\delta\phi_i,$$

and taking out a common factor,

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right) + \delta\phi_i\left(\frac{\partial\mathcal{L}}{\partial\phi_i} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\right)\right).$$

Now, the bracketed term on the far RHS is infact the Euler-Lagrange equation of motion (5.20), and is hence zero. Thus, we are left with

$$\delta\mathcal{L} = \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right).$$

Requiring no variation in the Lagrangian gives

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\delta\phi_i\right) = 0, \tag{5.32}$$

but, by (5.30), we have

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}i\theta_a\left(T^a\right)_{ij}\phi_j\right) = 0. \tag{5.33}$$

For global transformations, θ_a is a constant, so that we are left with

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}i\left(T^a\right)_{ij}\phi_j\right) = 0. \tag{5.34}$$

Let us now define the 4-current (not the same as the 4-current discussed above – this one is a symmetry current)

$$J^\mu \equiv \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}i\left(T^a\right)_{ij}\phi_j, \tag{5.35}$$

then (5.34) gives us a conservation equation,

$$\partial_\mu J^\mu = 0. \tag{5.36}$$

Expanded out of index notation, this equation reads

$$\partial_\mu J^\mu = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

Associated with this conservation equation is a conserved charge: integrating this equation,

$$\int_V \nabla \cdot \mathbf{J}\, dV = -\frac{\partial}{\partial t}\int_V \rho\, dV,$$

where by Gauss' law we send the LHS volume integral to a surface integral,

$$\int_S \mathbf{J}\cdot d\mathbf{S} = -\frac{d}{dt}\int_V \rho\, dV.$$

By exploiting our freedom to choose the surface over which to integrate, we choose a surface at infinity, so that the surface integral vanishes. Hence, we are simply left with

$$\frac{dQ}{dt} = 0, \quad Q \equiv \int_V \rho\, dV, \tag{5.37}$$

that is, a conserved charge.

Therefore, we have seen that by requiring a Lagrangian to be invariant under global transformations from a Lie group, there is a conserved 4-current J^μ and a conserved charge Q. This is called Noether's theorem.

5.4.1 The $SO(2)$ Symmetry Current

Let us give an example of a symmetry current associated with the invariance of a Lagrangian, under $SO(2)$ transformations of the fields. Consider a real field $\Phi = (\phi_1, \phi_2)$, with Lagrangian of the form

$$\mathcal{L} = \frac{1}{2}\left[(\partial_\mu\phi_1)^2 + (\partial_\mu\phi_2^2)\right] - V(\phi_1,\phi_2).$$

Then, via transformations by an element of the $SO(2)$ Lie group, upon the fields, we have

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \longmapsto \begin{pmatrix} \phi_1' \\ \phi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}.$$

The first quantity we must compute is the generator. The rotation matrix $R(\theta)$ in $SO(2)$ is just the matrix above, and the generator is just given by

$$iT = \left.\frac{dR}{d\theta}\right|_{\theta=0}.$$

Hence, we can easily see that the generator of $SO(2)$ rotations is

$$iT = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{5.38}$$

We must note that we have a single generator, T. We now appeal to the definition of the 4-current as in (5.32), as the bracketed quantity (as it is more insightful to talk in terms of variations of the field, which we must still compute),

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \delta\phi_i. \tag{5.39}$$

The variations of the field for a Lie group are just computed from (5.30),

$$\delta\phi_i = i\theta_a \left(T^a\right)_{ij} \phi_j.$$

To compute this for our model, we must explicitly do the sum. So, the transformations of the fields is,

$$\begin{aligned} \delta\phi_1 &= i\theta T_{11}\phi_1 + i\theta T_{12}\phi_2, \\ \delta\phi_2 &= i\theta T_{21}\phi_1 + i\theta T_{22}\phi_2. \end{aligned}$$

It is worth noting that we have been somewhat flippant in using i or $-i$: it does not matter, as J^μ satisfies the same continuity equation as $-iJ^\mu$. So, noting our generator (5.38), we see that the diagonal components are zero $T_{11} = 0 = T_{22}$ and the off-diagonals are $T_{12} = -1, T_{21} = 1$. Hence, our variations become just

$$\delta\phi_1 = -\theta\phi_2, \quad \delta\phi_2 = \theta\phi_1. \tag{5.40}$$

We now perform the summation over i in the 4-current (5.39), to see that

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_1)} \delta\phi_1 + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_2)} \delta\phi_2.$$

By the form of our Lagrangian, we can easily note that

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_1)} = \partial_\mu\phi_1, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_2)} = \partial_\mu\phi_2.$$

Hence, inserting this, the 4-current becomes

$$J_\mu = \partial_\mu\phi_1\delta\phi_1 + \partial_\mu\phi_2\delta\phi_2,$$

and further inserting our variations of the field (5.40), we see that

$$J_\mu = -\theta\left(\phi_2\partial_\mu\phi_1 - \phi_1\partial_\mu\phi_2\right).$$

Hence, as θ is a constant (we are making global transformations of the field), we see that the conserved symmetry current associated with $SO(2)$ transformations of the fields, is

$$J_\mu = \phi_2\partial_\mu\phi_1 - \phi_1\partial_\mu\phi_2. \tag{5.41}$$

The 0-component of this, J_0 is a conserved charge density, when integrated over all space, gives a conserved charge.

5.5 Gauging QED & QCD

Here we will discuss how to make a Lagrangian invariant under transformation associated with the groups $U(1)$ and $SU(3)$; making Lagrangians invariant in this way produces the theories of quantum electrodynamics (QED) and quantum chromodynamics (QCD). QED is the theory of interacting photons and matter, QCD the theory of interacting quarks and matter.

We will find that the Abelian $U(1)$ group, with the generator unity, leads to the absence of photon self-interaction, and induces electron-photon interaction. In a non-Abelian gauge group, such as $SU(3)$, in making a Lagrangian invariant, we find that gluon self-interactions are a direct consequence. In all cases we will consider, we shall start from the free-electron Dirac Lagrangian,

$$\mathcal{L} = \bar{\psi}\left(i\gamma \cdot \partial - m\right)\psi, \tag{5.42}$$

from which one can obtain the Dirac equation – the Dirac equation describes the motion of relativistic fermions. The general "recipe" for what we are about to do is quite simple. We take a free Lagrangian, transform the fields under a given group,

$$\psi \longmapsto \psi' = M\psi, \quad M \in G,$$

and see what needs to be redefined to keep the Lagrangian invariant.

5.5.1 QED: Invariance Under an $U(1)$ Abelian Group

Let us consider initially transforming the fields ψ under global $U(1)$ transformations,

$$\psi \longmapsto \psi' = e^{ie\Lambda}\psi, \quad \bar{\psi} \longmapsto \bar{\psi}' = e^{-ie\Lambda}\bar{\psi}, \tag{5.43}$$

where Λ is a constant. Notice that such transformations are just transformations of the phase – so we shall just call them phase transformations. We put the electronic charge e into the phase transformation term, by convention, as it is useful to see where it ends up. So, let us insert these transformations of the field into the free Lagrangian (5.42):

$$\begin{aligned}\mathcal{L} \longmapsto \mathcal{L}' &= \bar{\psi}'\left(i\gamma \cdot \partial - m\right)\psi' \\ &= e^{-ie\Lambda}\bar{\psi}\left(i\gamma \cdot \partial - m\right)e^{ie\Lambda}\psi,\end{aligned}$$

the factor of $e^{ie\Lambda}$ can "travel backwards" through the derivative (as it is a constant), to cancel with the $e^{-ie\Lambda}$ on the far LHS. This just leaves the Lagrangian invariant. Therefore, the free Lagrangian is invariant under global $U(1)$ transformation.

Now let us consider a local phase transformation,

$$\psi \longmapsto \psi' = e^{ie\Lambda(x)}\psi, \quad \bar{\psi} \longmapsto \bar{\psi}' = e^{-ie\Lambda(x)}\bar{\psi}, \tag{5.44}$$

where $\Lambda(x)$ – a function of spacetime. Then, let us substitute these transformations into the free Lagrangian, being careful to differentiate the argument of the exponential,

$$\begin{aligned} \mathcal{L} \longmapsto \mathcal{L}' &= \bar{\psi}' (i\gamma \cdot \partial - m) \psi' \\ &= e^{-ie\Lambda(x)} \bar{\psi} (i\gamma \cdot \partial - m) e^{ie\Lambda(x)} \psi \\ &= \bar{\psi} (i\gamma \cdot \partial - m - e\gamma \cdot \partial\Lambda) \psi. \end{aligned}$$

So, upon comparison of this with the free Lagrangian, we see that there is an extra term, $(-e\gamma \cdot \partial\Lambda)$, which was due to the position dependence of the phase transformation.

To obtain invariance, we need to introduce a 4-potential A_μ (which we call a gauge field), such that under its transformation law (which we must specify), it can absorb such terms. So, we introduce a field which has transformation rule

$$A_\mu \longmapsto A'_\mu \equiv A_\mu + \partial_\mu \Lambda. \tag{5.45}$$

We must also specify how to get this field into the Lagrangain. We do this by modifying the definition of the partial derivative,

$$\partial_\mu \longmapsto \mathcal{D}_\mu \equiv \partial_\mu - ieA_\mu. \tag{5.46}$$

We call $\mathcal{D}_\mu$ the covariant derivative (those familiar with general relativity will recall that we define a covariant derivative that takes care of derivatives of position dependent bases). Thus, we propose (and then show) that

$$\mathcal{L}_{\text{QED}} = \bar{\psi} (i\gamma \cdot \mathcal{D} - m) \psi \tag{5.47}$$

is invariant under local $U(1)$ phase transformation. So, to show the invariance of this Lagrangian, consider

$$(\partial_\mu - ieA_\mu) \psi = \mathcal{D}_\mu \psi,$$

and transform it to

$$\mathcal{D}'_\mu \psi' = \left(\partial_\mu - ieA'_\mu\right) \psi'.$$

Hence, inserting our transformations of the field (5.44) and gauge field (5.45),

$$\begin{aligned} \mathcal{D}'_\mu \psi' &= \left(\partial_\mu - ieA'_\mu\right) \psi' \\ &= \left(\partial_\mu - ieA'_\mu\right) e^{ie\Lambda(x)} \psi \\ &= \left(e^{ie\Lambda(x)} \partial_\mu + e^{ie\Lambda(x)} ie\partial_\mu\Lambda - e^{ie\Lambda(x)} ieA'_\mu\right) \psi \\ &= e^{ie\Lambda(x)} \left(\partial_\mu + ie\partial_\mu\Lambda - ie(A_\mu + \partial_\mu\Lambda)\right) \psi, \end{aligned}$$

we now see that the second and fourth terms will cancel, leaving

$$\mathcal{D}'_\mu \psi' = e^{ie\Lambda(x)} (\partial_\mu - ieA_\mu) \psi = e^{ie\Lambda} \mathcal{D}_\mu \psi.$$

The content of this expression is that the covariant derivative is gauge invariant – the derivative itself absorbs terms which come about from finding the derivative of the rotation term. Therefore, we see that $\bar{\psi} \mathcal{D}_\mu \psi$ is invariant, just like $\bar{\psi}\psi$.

The gauge invariant Lagrangian (5.47) now contains an interaction term, between ψ and the gauge field A_μ. To see this, we shall expand out the Lagrangian,

$$\begin{aligned}\mathcal{L}_{\text{QED}} &= \bar{\psi}\left(i\gamma^\mu\left(\partial_\mu - ieA_\mu\right) - m\right)\psi \\ &= i\bar{\psi}\gamma^\mu\partial_\mu\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - m\bar{\psi}\psi,\end{aligned}$$

where the three-point interaction term is the middle one, $e\bar{\psi}\gamma^\mu A_\mu\psi$; the strength of the interaction is just proportional to e; see Figure (5.1) for a pictorial description of this interaction.

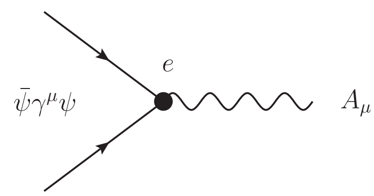


Figure 5.1: Interaction of photon with electron, due to the interaction term in the gauge invariant QED Lagrangian.

This interaction is of the form of a current,

$$e\bar{\psi}\gamma^\mu A_\mu\psi = J^\mu A_\mu,$$

which will be familiar from relativistic quantum mechanics.

We generally add on the photon kinetic term to the Lagrangian, so that

$$\mathcal{L} = \bar{\psi}\left(i\gamma\cdot\mathcal{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{(\partial\cdot A)^2}{2\xi}. \tag{5.48}$$

It is interesting to note that there are no photon-photon interaction terms (i.e. no $A^\mu A_\mu$ terms – this infact corresponds to the massless nature of the photon, which we now see is a perfectly natural feature of the theory). We added on the photon term to the QED gauge invariant Lagrangian, without checking that this new piece is also gauge invariant (if it is not, then we shall have to invent new terms). So, let us transform the field strength tensor

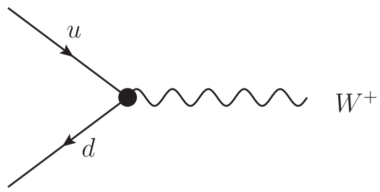
$$\begin{aligned}F_{\mu\nu} \longmapsto F'_{\mu\nu} &= \partial_\mu A'_\nu - \partial_\nu A'_\mu \\ &= \partial_\mu\left(A_\nu + \partial_\nu\Lambda\right) - \partial_\nu\left(A_\mu + \partial_\mu\Lambda\right) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ &= F_{\mu\nu},\end{aligned}$$

where the third equality follows as partial derivatives commute. Therefore, we have shown that the field-strength tensor is gauge invariant. Hence, we say that (5.48) is the gauge invariant QED Lagrangian, such that the Lagrangian is invariant under local Abelian transformations of the field.

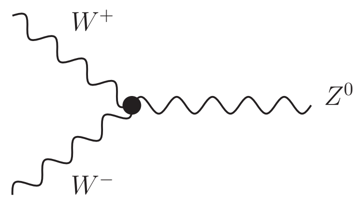
5.5.2 Weak Force: Invariance Under an $SU(2)$ Group

We shall not discuss transformations under the $SU(2)$ group in detail, as we can make inferences backwards, when we discuss $SU(3)$ gauge invariance. The salient point of the $SU(2)$ group is that it is non-Abelian – whereas the $U(1)$ group is Abelian.

There are two sorts of interactions which get generated as a result of making the Lagrangian invariant, displayed in Figure (5.2).



(a) Cross-interaction.



(b) Self-interaction.

Figure 5.2: Interactions in the $SU(2)$ gauge field, which arise from the invariant Lagrangian.

In the QED theory, there were no terms which allowed an interaction of the gauge field with itself. In making a Lagrangian gauge invariant under $SU(2)$, one finds terms which allow the interaction of the gauge field with itself. As one can see from Figure (5.2)a, the interaction of the ud-quarks produces a W^+ gauge boson. This can be thought about in terms of rotations in isospin space, where one should recall from our previous discussions that these two quarks inhabit the same isospin doublet; one can think about rotations in isospin space being compensated by the emission of W^+.

5.5.3 QCD: Invariance Under an $SU(3)$ Group

Let us now consider how to modify the Lagrangian to make it invariant under $SU(3)$ gauge (i.e. local) transformations. Recall that $SU(3)$ is non-Abelian – which we will find has some interesting consequences. The fundamental objects analogous to the electrons of QED, are coloured quarks,

$$\psi = \begin{pmatrix} \psi_r \\ \psi_b \\ \psi_g \end{pmatrix}.$$

These objects are transformed by the 3x3 matrices of $SU(3)$,

$$\psi'(x) = U(x)\psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x)U^\dagger(x).$$

We must now be very careful when ordering things: in our QED discussion we did not, as the group elements were Abelian. By the unitarity of $U \in SU(3)$ we have

$$U^\dagger U = UU^\dagger = I,$$

where I is the identity matrix.

Before we continue, it is worth our working out how many generators the general $SU(N)$ group has. A matrix $M \in SU(N)$ has $N \times N = N^2$ elements, each of which is (generally) complex. Hence, each matrix has $2N^2$ degrees of freedom (the two comes from the two degrees of freedom per complex number). Now, as M is unitary, $M^\dagger M = I$, which gives N^2 constraints. Also, $\det M = 1$ (recalling that $SU(N)$ denotes the group of unitary matrices with unit determinant), which is another constraint. Hence, there are $2N^2 - (N^2+1) = N^2 - 1$ degrees of freedom left. Therefore, we can take these to be rotation angles ω_a, each of which we associate with a generator t_a. Hence, we see that the group $SU(N)$ has $N^2 - 1$ generators. So, we write the general matrix as

$$U = e^{i\omega_a t_a},$$

with an implied summation over the repeated index a. The generators t_a are traceless Hermitian matrices.

Thus, we see that the $SU(3)$ group has 8 generators, which we call gluons, and we use them to absorb phases. Notice that the $SU(2)$ group hence has three generators, which corresponds to W^+, W^-, Z^0.

Let us consider infinitesimal rotations (i.e. small angle $\delta\omega_a$) of quarks in colour space,

$$\psi \longmapsto \psi' = U\psi = e^{i\delta\omega_a t_a}\psi.$$

So, writing the new field as the old field plus a small change,

$$\psi' = \psi + \delta\psi,$$

and noting the expansion of the exponential,

$$e^{i\delta\omega_a t^a} = I + i\delta\omega_a t_a,$$

we read off that

$$\delta\psi = i\delta\omega_a t^a \psi. \tag{5.49}$$

The t_a are the generators of rotations of quarks, which transform in the fundamental representation. So, for anti-quarks,

$$\begin{aligned} \bar{\psi} \longmapsto \bar{\psi}' &= \bar{\psi} U^\dagger \\ &= \bar{\psi} e^{-i\delta\omega_a t_a} \\ &= \left(\bar{\psi} + \delta\bar{\psi}\right), \end{aligned}$$

and hence we see that

$$\delta\bar{\psi} = -i\delta\omega_a \bar{\psi} t_a, \tag{5.50}$$

where we have been careful in preserving the order of the spinor and generator. Hence, we see that the generators for rotations of anti-quarks are $-t_a$. In $SU(3)$, the t_a are called the Gell-Mann matrices.

Now, the generators satisfy the commutator

$$[t_a, t_b] = if_{abc}t_b, \tag{5.51}$$

where the f_{abc} are the totally anti-symmetric structure constants.

Following this brief review, we are now able to construct a Lagrangian invariant under $SU(3)$ gauge transformations.

Let us consider the free Dirac term, as we did for QED. So, ignoring the γ^μ as they play no role in our discussions, we have

$$\mathcal{L} = \bar{\psi}\partial_\mu\psi.$$

Let us then transform the fields,

$$\psi \to U\psi, \quad \bar{\psi} \to \bar{\psi}U^\dagger = \bar{\psi}U^{-1},$$

where the last equality follows from unitarity. Then, the Dirac term becomes

$$\begin{aligned}
\mathcal{L} \longmapsto \mathcal{L}' &= \bar{\psi}U^\dagger\partial_\mu(U\psi) \\
&= \bar{\psi}U^{-1}\partial_\mu U\psi \\
&= \bar{\psi}U^{-1}U(\partial_\mu\psi) + \bar{\psi}U^{-1}(\partial_\mu U)\psi \\
&= \bar{\psi}\partial_\mu\psi + \bar{\psi}U^{-1}(\partial_\mu U)\psi.
\end{aligned}$$

Hence, we see an extra term appearing, $\bar{\psi}U^{-1}(\partial_\mu U)\psi$, meaning that the Lagrangian, as it stands, is not invariant. Let us try the same method as before, and introduce a compensating field, by replacing the partial derivative by the covariant derivative,

$$\partial_\mu \longmapsto \mathcal{D}_\mu \equiv \partial_\mu - igA_\mu, \tag{5.52}$$

where we now have "g" rather than "e", which denotes the colour charge. We must also specify the transformation rule of this gauge field; which we do via

$$A_\mu \longmapsto A'_\mu = UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1}. \tag{5.53}$$

There are some important differences between this rule and the rule presented for the $U(1)$ case, (5.45), most notably the U-terms. With this transformation, let us check that the new Dirac term is gauge invariant. So,

$$\begin{aligned}
\bar{\psi}\mathcal{D}_\mu\psi \longmapsto \bar{\psi}'\mathcal{D}'_\mu\psi' &= \bar{\psi}'\left(\partial_\mu - igA'_\mu\right)\psi' \\
&= \bar{\psi}'\left(\partial_\mu - ig\left(UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1}\right)\right)U\psi \\
&= \bar{\psi}'\left((\partial_\mu U) + U\partial_\mu - igUA_\mu\psi - (\partial_\mu U)\right)\psi \\
&= \bar{\psi}'U\left(\partial_\mu - igA_\mu\right)\psi,
\end{aligned}$$

and when we put in the transformation $\bar{\psi}' = \bar{\psi}U^{-1}$, we see that $U^{-1}U = I$, and hence that

$$\begin{aligned}\bar{\psi}\mathcal{D}_\mu\psi \longmapsto \bar{\psi}'\mathcal{D}'_\mu\psi' &= \bar{\psi}'\left(\partial_\mu - igA'_\mu\right)\psi' \\ &= \bar{\psi}\left(\partial_\mu - igA_\mu\right)\psi \\ &= \bar{\psi}\mathcal{D}_\mu\psi.\end{aligned}$$

Therefore, we have shown the gauge invariance of the Dirac term.

There is an interesting subtlety which we have not yet considered. In QED the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ was gauge invariant, and infact contained observables in the electric/magnetic fields. In QCD, the "normal" field strength tensor is not invariant, and so cannot correspond to an observable. That is, we cannot observe the colour field. So, we modify using the covariant derivatives,

$$F_{\mu\nu} = \mathcal{D}_\mu A_\nu - \mathcal{D}_\nu A_\mu, \tag{5.54}$$

This is the non-Abelian generalisation of the field strength tensor. Another, completely equivalent, way of writing the field strength tensor is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]. \tag{5.55}$$

To see this equivalence, let us simply insert the definition of the covariant derivative (5.52) into (5.54) and expand out

$$\begin{aligned}F_{\mu\nu} &= (\partial_\mu - igA_\mu)A_\nu - (\partial_\nu - igA_\nu)A_\mu \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu - ig\left(A_\mu A_\nu - A_\nu A_\mu\right),\end{aligned}$$

where the final term is just the commutator expressed in (5.55). We must now take into account the different gluons, which we do by introducing

$$A_\mu = t_a A_{\mu a}, \quad F_{\mu\nu} = t_a F_{\mu\nu a},$$

with an implied summation over a. Hence the field-strength tensor can be written

$$F_{\mu\nu a} = \partial_\mu A_{\nu a} - \partial_\nu A_{\mu a} + g f_{abc} A_{\mu b} A_{\nu c}, \tag{5.56}$$

which can be written as

$$\mathbf{F}_{\mu\nu} = \partial_\mu \mathbf{A}_\nu - \partial_\nu \mathbf{A}_\mu + g\mathbf{A}_\mu \times \mathbf{A}_\nu.$$

Then,

$$\sum_a F_{\mu\nu a} F^{\mu\nu}{}_a$$

is gauge invariant.

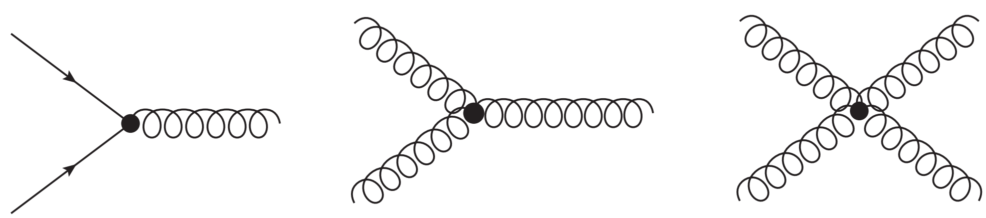
So, the full non-Abelian gauge invariant Lagrangian is

$$\mathcal{L} = \bar{\psi}\left(i\gamma \cdot \mathcal{D} - m\right)\psi - \frac{1}{4}F_{\mu\nu a}F^{\mu\nu}{}_a, \tag{5.57}$$

where (to recap),

$$\mathcal{D}_\mu = \partial_\mu - igA_\mu, \tag{5.58}$$
$$A_\mu = t_a A_{\mu a}, \tag{5.59}$$
$$A'_\mu = UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1}. \tag{5.60}$$

As in the QED Lagrangian, we can read off various interaction terms. If one expands out all of the terms in the Lagrangian, one finds three types of interaction as displayed in Figure (5.3). The first interaction, between the fermionic field and the gauge field has strength g, as in QED. However, there are extra terms which arise from squaring the field strength tensor.



(a) Quark emitted from gluon.

(b) Gluon self-interaction: g.

(c) Gluon self-interaction: g^2.

Figure 5.3: Interactions in the $SU(3)$ gauge field, which arise from the invariant Lagrangian.

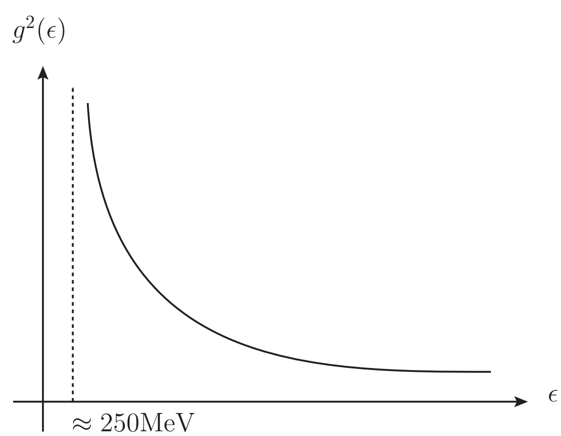


Figure 5.4: Running of the coupling "constant" in QCD, due to the self-interaction terms in the Lagrangian.

As Figure (5.4) shows, the coupling "constant" g changes with energy ϵ, which are a consequence of the self-interactions present in the field-strength tensor. The coupling strength increases with distance (i.e. with decreasing energy), up to $\epsilon \approx 250\text{MeV}$, where there is a confinement region. Above the confinement region, is asymptotic freedom. That is, at small distances, the colour field is free, but at large distances, the self-interactions become much stronger.