IV. SU(3)

A. SU(3)

The three continuous groups that play the most prominent role in particle physics are U(1),. SU(2), and SU(3). We now turn our attention to the group SU(3)

SU(3) is the set of 3×3 unitary matrices U with determinant one. As always, we can write U in terms of generators:

$$U = \exp\left(-ix_a T_a\right) \tag{4.1}$$

where we have introduced for the first time the *Einstein summation convention*, which means that any index that is repeated in a single term (in this case, a) has an implied sum. This notation will prove very useful later in the chapter. For U to be unitary, T_a must be Hermitian. For it to be determinant one, we must have

$$1 = \left|U\right| = \left|\exp\left(-ix_a T_a\right)\right| = \exp\left[-i\mathrm{tr}\left(x_a T_a\right)\right] \tag{4.2}$$

and therefore $\mathrm{Tr}\left(T_a\right) = 0$. There are nine linearly independent Hermitian matrices, but the traceless constraint reduces this to eight. By convention, these matrices are given by

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$
$$T_5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad T_6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad T_7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad T_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{4.3}$$

These have been normalized so that $\mathrm{tr}\left(T_a T_b\right) = \frac{1}{2}\delta_{ab}$.

We note first that among these generators, we can find *two* generators that commute with each other, and therefore we can simultaneously diagonalize two of them. We have already done so, and the two diagonal matrices T_3 and T_8 will play a special role in SU(3). This three dimensional representation will be called the "3", and ultimately it is the only irrep of SU(3) we will explicitly write down. The three basis vectors in this representation are called $\left|_1\right\rangle$, $\left|_2\right\rangle$, and $\left|_3\right\rangle$, and the eigenvalues of these three operators under T_3 and T_8 are evident from (4.3):

$$\begin{aligned} T_3\left|_1\right\rangle &= +\tfrac{1}{2}\left|_1\right\rangle, & T_3\left|_2\right\rangle &= -\tfrac{1}{2}\left|_2\right\rangle, & T_3\left|_3\right\rangle &= 0, \\ T_8\left|_1\right\rangle &= +\tfrac{1}{2\sqrt{3}}\left|_1\right\rangle, & T_8\left|_3\right\rangle &= +\tfrac{1}{2\sqrt{3}}\left|_2\right\rangle & T_8\left|_3\right\rangle &= -\tfrac{1}{\sqrt{3}}\left|_3\right\rangle \end{aligned} \tag{4.4}$$

These eigenvalues, taken as pairs, are called the *weights* of this representation. They are plotted in two dimensions, as illustrated in Fig. 4-1, with each basis vector plotted as a point in the two-

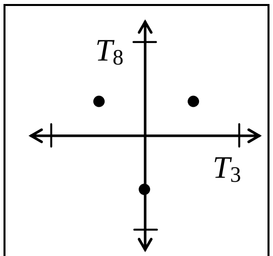

Figure 4-1: The weights of the 3 irrep of SU3).

dimensional space of eigenvalues under the two operators. They form an equilateral triangle of side one, centered on the origin.

Given a representation, we can always find the complex conjugate representation. The complex conjugate of the 3 representation is called the "$\bar{3}$", and the generators are given by

$$T_a^{(\bar{3})} = -T_a^{(3)T} \tag{4.5}$$

Let's denote the basis vectors for the $\bar{3}$ as $\left|{}^1\right\rangle$, $\left|{}^2\right\rangle$, and $\left|{}^3\right\rangle$. The eigenvalues of these three basis vectors will therefore be given by the negatives of the corresponding basis vectors of the 3 representation, and therefore a plot of the weights will simply be the negatives of the 3 representation, as sketched in Fig. 4-2. Since the weights are different, the 3 is not equivalent to its complex conjugate, and therefore the 3 is a complex representation. For completeness, I have also included a sketch of the trivial 1 representation in Fig. 4-3.

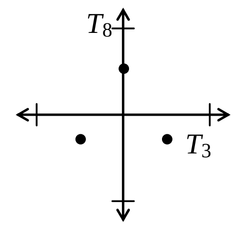

Figure 4-2: The weights of the $\bar{3}$ irrep of SU(3).

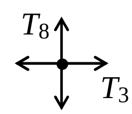

Figure 4-3: The weight of the 1 irrep of SU(3).

B. Tensor Notation

It will prove helpful to develop a tensor notation which keeps track of what happens to an arbitrary vector as we act on it with various generators. Our first step is a minor notational one: whenever we encounter a matrix, we will write the row index elevated ("staggered indices"); for example, the components of a unitary matrix U would be written $U^a{}_b$, with a the row index and b the column index. The product of two unitary matrices would look like $U^a{}_b V^b{}_c$; note that the index that is summed over ends up with one index up and one down. This will be a general feature of how we combine indices.

Suppose we are looking at an arbitrary element of the 3 irrep of SU(3). Such an arbitrary element could be written in the form

$$\left|\mathbf{v}\right\rangle = \left|{}_i\right\rangle v^i \tag{4.6}$$

where v^i are the three "components" of this vector. If we let a generator T_a act on this, for example, we would have

$$T_a\left|\mathbf{v}\right\rangle = T_a\left|{}_i\right\rangle v^i = \left|{}_j\right\rangle \left(T_a^{(3)}\right)^j{}_i v^i = \left|{}_i\right\rangle \left(T_a^{(3)}\right)^i{}_j v^i \tag{4.7}$$

In terms of components, then the effect on **v** is

$$T_a : v^i \to \left(T_a^{(3)}\right)^i{}_j v^j \tag{4.8}$$

In a similar manner, we can write an arbitrary vector in the $\bar{3}$ representation as

$$\left|\mathbf{u}\right\rangle = u_i\left|{}^i\right\rangle \tag{4.9}$$

With the help of (4.10), it is easy to figure out the effect of the generators on these vectors:

$$\left(T_a u\right)_i = -u_j \left(T_a^{(3)}\right)^j{}_i \tag{4.10}$$

In each case, it is easy to see that these objects are three-dimensional representations of the group SU(3), because there are three independent components of v^i or u_i.

Now, suppose we have a tensor product representation, for example, the $3 \otimes 3$ representation of SU(3). The basis states in this case look like $\left| {}_{ij} \right\rangle$, with i and j running from 1 to 3. A general vector in this basis state would look like $w^{ij} \left| {}_{ij} \right\rangle$. It would transform under the action of a generator according to

$$\left(T_a w\right)^{ij} = \left(T_a^{(3)}\right)^i{}_k w^{kj} + \left(T_a^{(3)}\right)^j{}_k w^{ik} \tag{4.11}$$

This concept can be generalized. We might write the most general element of the tensor product of N 3's and $M\,\overline{3}$'s as

$$w^{i_1 i_2 \ldots i_N}_{j_1 j_2 \ldots j_M} \left| {}^{j_1 j_2 \ldots j_M}_{i_1 i_2 \ldots i_N} \right\rangle \tag{4.12}$$

We will find within these very general representations of SU(3) all of the irreps. The action of a generator on this tensor will be

$$\begin{aligned}\left(T_a w\right)^{i_1 i_2 \ldots i_N}_{j_1 j_2 \ldots j_M} \rightarrow &\left(T_a^{(3)}\right)^{i_1}{}_k w^{k i_2 \ldots i_N}_{j_1 j_2 \ldots j_M} + \left(T_a^{(3)}\right)^{i_2}{}_k w^{i_1 k \ldots i_N}_{j_1 j_2 \ldots j_M} + \cdots + \left(T_a^{(3)}\right)^{i_N}{}_k w^{i_1 i_2 \ldots k}_{j_1 j_2 \ldots j_M} \\ &- w^{i_1 i_2 \ldots i_N}_{k j_2 \ldots j_M} \left(T_a^{(3)}\right)^k{}_{j_1} - w^{i_1 i_2 \ldots i_N}_{j_1 k \ldots j_M} \left(T_a^{(3)}\right)^k{}_{j_2} - \cdots - w^{i_1 i_2 \ldots i_N}_{j_1 j_2 \ldots k} \left(T_a^{(3)}\right)^k{}_{j_M}\end{aligned} \tag{4.13}$$

The great thing about this approach is that any representation that the effects of the generators on these complicated tensors can be worked out using *only* the explicit form of the T_a's in the 3 representation; no more complicated representations are ever needed.

C. Invariant Tensors and Irreps

The representations we have found so far are *not* irreps. For example, consider the $3 \otimes \overline{3}$ representation, which has in general nine components. Consider, in particular, the component given by $w^i_j = \delta^i_j$. The action of a generator on this particular tensor is

$$\left(T_a \delta\right)^i_j \rightarrow \left(T_a^{(3)}\right)^i{}_k \delta^k_j - \delta^i_k \left(T_a^{(3)}\right)^k{}_j = 0 \tag{4.14}$$

This means that this particular component of the $3 \otimes \overline{3}$ representation is unchanged under *all* of the generators. δ^i_j is called an *invariant* tensor. Basically, it is invariant because, under a general unitary transformation U, it would change to

$$U : \delta^i_j \rightarrow U^i{}_k \delta^k_l \left(U^\dagger\right)^l{}_j = \left(U U^\dagger\right)^i{}_j = \delta^i_j \tag{4.15}$$

This is not the only invariant tensor. It is not hard to show that the Levi-Civita tensor ε_{ijk} and the corresponding tensor ε^{ijk} are also invariant; for example

$$(T_a\varepsilon)_{ijk} = -\varepsilon_{ljk}\left(T_a^{(3)}\right)^l{}_i - \varepsilon_{ilk}\left(T_a^{(3)}\right)^l{}_j - \varepsilon_{ijl}\left(T_a^{(3)}\right)^l{}_k = 0 \tag{4.16}$$

This can be shown to be the result of the T_a's being traceless, which is in turn related to the fact that the U's have determinant one.

We can now proceed to trying to find a large number of irreps of SU(3); in fact, we will find all of them, though we won't know that until the next chapter. Our goal is to write them as tensors $w^{i_1 i_2 \cdots i_N}_{j_1 j_2 \cdots j_M}$ while keeping the number of indices $N + M$ to a minimum. We'll start with a simple example. Consider a general tensor of the form w^i_j containing one index up and one down, so $(N, M) = (1,1)$. This has nine independent components, but it is not an irrep, because it contains a portion proportional to δ^i_j. The trace of w^i_j, does not change as you act on it with SU(3) operators. Hence the $3 \otimes \bar{3}$ representation contains the trivial representation (under our current notation, the $(0,0)$ representation). We must eliminate this, simply by demanding that $w^i_i = 0$, so that there is no trace.

Demanding that the trace of w vanish reduces the number of independent components of w from nine to eight. Hence this will be an eight-dimensional irrep of SU(3). It is easy to work out the weights. The weights of the $3 \otimes \bar{3}$ are simply the sum of all the weights of the 3 and $\bar{3}$ representation, but then subtract out one of the zero weights, leaving the eight weights

$$\left\{(\pm 1, 0), \left(\pm\tfrac{1}{2}, \pm\tfrac{\sqrt{3}}{2}\right), \left(\pm\tfrac{1}{2}, \mp\tfrac{\sqrt{3}}{2}\right), 2\times(0,0)\right\}. \tag{4.17}$$

and sketched in Fig. 4-4. This representation is called the 8. It isn't hard to show that it is a real representation; indeed, it is the adjoint representation, defined in chapter 1. The weights form a regular hexagon with side 1 centered on the origin, plus a pair of zero weights (designated by a dot and a circle).

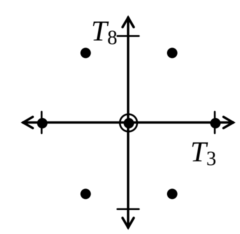

Figure 4-4: The weights of the 8 irrep of SU3).

Let's try again, this time working with the $3 \otimes 3$ representation. The components of an arbitrary vector will be of the form w^{ij}, or nine independent components. However, if we imagine pulling out the anti-symmetric part of it by defining

$$w_k \equiv w^{ij}\varepsilon_{ijk} \tag{4.18}$$

we would end up with a part with only one down index, transforming as the $\bar{3}$ representation. To avoid having this piece, we demand that this vanish, which will occur if the tensor has no anti-symmetric part, so we demand $w^{ij} = w^{ji}$. This reduces the number of components to six. This irrep, which is called the 6, will have weights again arranged in an equilateral triangle, as sketched in Fig. 4-5.

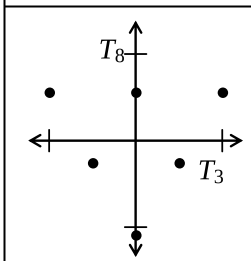

Figure 4-5: The weights of the 6 irrep of SU3).

Let's try to make this as general as possible. A general representation with N indices up and M indices down will be denoted by a tensor of type $w^{i_1 i_2 \ldots i_N}_{j_1 j_2 \ldots j_M}$. To make sure we can't play a trick like Eq. (4.22) and replace two up indices with a single down index, this tensor must be completely symmetric in all of its up indices. Similarly, it must be completely symmetric in the down indices as well. Finally, to keep you from eliminating one up and one down index, we must also demand

$$w^{k i_2 \ldots i_N}_{k j_2 \ldots j_M} = 0 \tag{4.19}$$

Because we have already demanded that this tensor be symmetric on all up and all down indices, it doesn't matter which index we choose from the top and bottom to sum over.

We have now figured out how to construct basis vectors chosen from $3^N \otimes \overline{3}^M$ to make what I am calling the (N,M) irrep. I haven't really proven that these will be irreps (they are), nor that we got all of them (we did). I would like to explain just a little more clearly how we can work out what the weights are in any given representation. For example, the $(3,0)$ irrep corresponds to tensors of type w^{ijk} that are completely symmetric. Therefore, for example, $w^{123} = w^{132}$, etc. The weights of this will be simply the sum of the weights corresponding to any three of the weights corresponding to $\left|{}_1\right\rangle$, $\left|{}_2\right\rangle$, and $\left|{}_3\right\rangle$, allowing repeats but not counting separately 1+2+3 and !+3+2, etc. This works out to form a ten-dimensional representation of SU(3), as illustrated in Fig. 4-6. This representation, called the 10, will prove important in later sections.

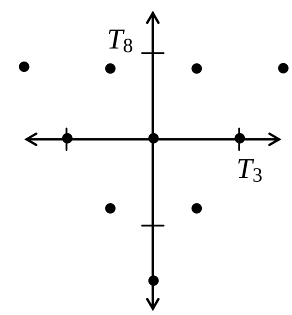

Figure 4-6: The weights of the 10 irrep of SU3).

The general pattern is not too hard to figure out. The weights of the (N,M) irrep always turn out to make a hexagon pattern, with all the outer angles 120 degrees, and the sides alternately being of size N and M (if one of these is zero, then it will be an equilateral triangle instead). The outermost layer of weights will be singular, but as we work our way in, the weights will become doubly and then triply degenerate, etc., until the pattern is triangular, at which the weights will no longer increase their degeneracy. In Fig. 4-7, I've sketched the weights for the $(2,1)$ irrep, which I believe is called the 15 representation.[1]

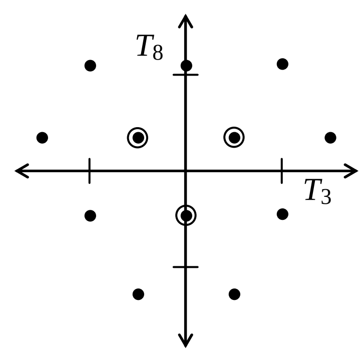

Figure 4-7: The weights of the 15 irrep of SU3).

The dimensions of the (N,M) irrep is a bit tricky to work out, but not too bad. It turns out that the number of ways of picking the indices $i_1 i_2 \ldots i_N$ out of the set $\{1,2,3\}$, if order doesn't matter but repetition is allowed, is $\frac{1}{2}(N+2)(N+1)$. Since we have also to

[1] The (4,0) irrep is called the 15'.

choose lower indices, you might naively expect the dimension to be the product of this with the same formula for *M*, or $\frac{1}{4}(N+2)(N+1)(M+2)(M+1)$. But this ignores the traceless condition, Eq. (4.19), which adds $\frac{1}{4}(N+1)N(M+1)M$ constraints. The dimension is the difference between these numbers, which works out to

$$\dim(N,M)=\tfrac{1}{2}(N+1)(M+1)(N+M+2) \tag{4.20}$$

D. Tensor Products of Irreps

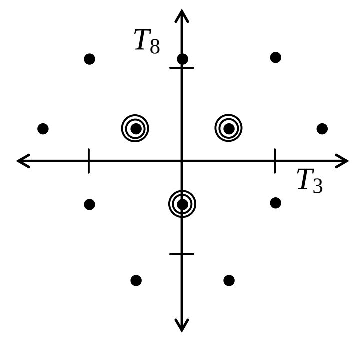

Figure 4-8: The weights of the $6\otimes\overline{3}$ representation of SU(3). By inspection, the result is the weights of the 15 and the weights of the 3.

Tensor products are easily worked out if you have the weights of the various irreps. For example, suppose we want to take the tensor product $6\otimes\overline{3}$. This representation is eighteen-dimensional. The weights are simply the sum of all the weights of the 6 and the $\overline{3}$. I have sketched the results in Fig. 4-8. It bears a strong resemblance to the fifteen-dimensional irrep in Fig. 4-7. Indeed, it includes all of these weights, and the three remaining weights are simply those of the 3. Hence we conclude

$$6\otimes\overline{3}=15\oplus 3 \tag{4.21}$$

We can get the same result another way. Recall that the 6 irrep is written in terms of symmetric tensors of the form w^{ij}, and the $\overline{3}$ is written like u_k. When we multiply them together, we can combine them to make a tensor with two indices up and one down, like this: $w^{ij}u_k$, which is nothing but the 15 irrep. However, we must remove the trace part, which would be a tensor like $w^{ik}u_k$, which has only one up index.

Let's do a more complicated case to make sure we understand it. What is $8\otimes 8$? This would be represented by two tensors u_i^j and v_i^j. Then can be multiplied together to make a tensor like $u_i^j v_k^l$, an irrep of type $(2,2)$, which according to equation (4.20) is 27 dimensional. You can also combine indices, connecting them up and down, to make something with only two indices left over, like $u_i^j v_j^l$ or $u_i^j v_k^i$, which are two 8's (since they have one index up and one index down), or even get rid of all the indices: $u_i^j v_j^i$ to make a 1. Finally, you can turn two of the up indices into a down index, or vice versa, to make $u_i^j v_k^l \varepsilon_{jlm}$ or $u_i^j v_k^l \varepsilon^{ikm}$, therefore making an irrep of type (0,3) (the $\overline{10}$) or of type (3,0) (the 10). Putting it all together, we conclude

$$8\otimes 8=27\oplus 10\oplus\overline{10}\oplus 8\oplus 8\oplus 1 \tag{4.22}$$

This technique can be systematized using a variety of techniques, but we won't worry about it too much. Suffice it to say there are straightforward ways of finding tensor products of any pair of representations in SU(3), and they can be generalized to SU(*N*).

D. Hypercharge and Gell-Mann SU(3)

It's time to start making some connections with particle physics. We start by defining hypercharge Y for strongly interacting particles as

$$Y = 2\left(Q - I_3\right) \qquad (4.23)$$

where Q is electric charge and I_3 is the third component of isospin. Because charge increases by one unit for each unit of increase in isospin, hypercharge Y will be the same across an isospin multiplet. Since Q and I_3 are preserved by strong interactions (and electromagnetic interactions, for that matter), isospin will be conserved by strong interactions. Hence the kaons, for example, are long lived (by particle physics standards), because there is nothing lighter with hypercharge that it can decay into.

What's interesting is if we make a plot of I_3 vs. Y for the eight lightest mesons, which are all spin zero, as sketched in Fig. 4-9. The plot looks extremely similar to the weight diagram for the 8 irrep of SU(3). The plot is not *exactly* a regular hexagon, because the vertical scale is off by a factor of $2/\sqrt{3}$, which suggests perhaps $Y = 2T_8/\sqrt{3}$. An identical plot occurs if we consider the eight lightest spin half baryons, or the eight lightest spin-1 mesons. It looks as if these particles might somehow be connected by an SU(3) symmetry! Also consider a plot of the nine lightest spin-3/2 baryons. These fit perfectly into the 10 irrep of SU(3), as sketched in Fig. 4-10, except there is one point missing. This missing state, if it existed, should have hypercharge -2 (something that had never been seen before), and charge -1. It was named the Ω^- particle.

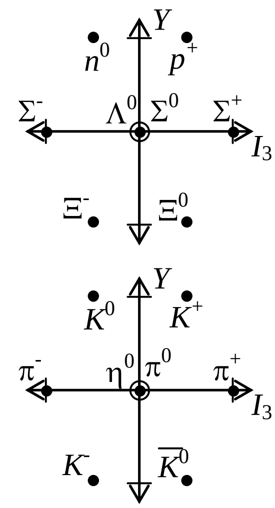

Figure 4-9: The third component of isospin vs. hypercharge for the eight lightest spin-0 mesons (bottom) and eight lightest spin-1/2 baryons (top).

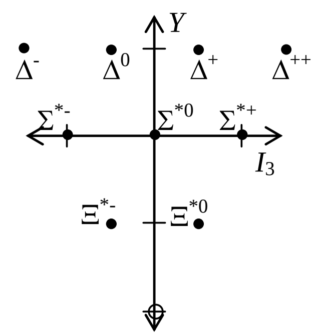

Figure 4-10: The third component of isospin vs. hypercharge for the nine lightest spin-3/2 baryons (filled circles). The open circle at the bottom would complete the 10 irrep of SU(3), and was predicted by Gell-Mann. It was named the Ω^-.

A problem occurs almost immediately. If isospin symmetry is really valid, then it is easy to show that all the particles in a multiplet should have the same mass. This doesn't work particularly well for the eight spin-1/2 baryons, nor for the spin-3/2 baryons, and is positively awful for the eight spin-0 mesons. At first glance, this seems like a disaster. But there is, perhaps a couple of reasons to think things aren't as disastrous as they seem. First of all, it is possible for the *interactions* of particles to more or less respect a symmetry, even if the *mass* does not. For example, even though the proton and deuteron have very different masses, their interactions with the electrons is nearly identical. In particle physics, because we are so often dealing with relativistic particles, there is reason to believe there should be less distinction between "masses" and "interactions," but still, it is too soon to give up. Secondly, it is commonly the case that even when *predictions* from group theory no longer are accurate, they can commonly still be useful as an organizing principle. For example, in multi-electron atoms, we commonly speak of having a certain number of 1s electrons, 2s electrons, 2p electrons, and so on. This way of describing electrons is not technically accurate, because each electron can be identified with a certain orbital only if we ignore the interactions between electrons, a terrible approximation. Nonetheless, this way of describing the electrons composing an atom actually works pretty well. Note that the generators T_1, T_2, and T_3 are the generators of an SU(2) subgroup of SU(3); this is simply the isospin group discussed in the previous chapter.

We would like to describe all the baryons or mesons that occur together in a multiplet as if they are components of a common object, using our vector notation. For example, to describe a spin-0 meson, we would write it in the form

$$u_j^i \left| M_i^j \right\rangle \tag{4.24}$$

where the components u_j^i tells us which of the eight particles we are dealing with. The values of the various u-components, up to sign conventions, can be worked out from the weights, together with some knowledge about isospin. The non-zero ones are given by

$$\begin{array}{llll} K^0 : u_3^2 = 1, & K^+ : u_3^1 = 1, & K^- : u_1^3 = 1, & \bar{K}^0 : u_2^3 = 1, \\ \pi^- : u_1^2 = 1, & \pi^0 : u_1^1 = -u_2^2 = \frac{1}{\sqrt{2}}, & \pi^+ : u_2^1 = 1, & \eta^0 : u_1^1 = u_2^2 = \frac{1}{\sqrt{6}}, u_3^3 = -\frac{2}{\sqrt{6}} \end{array} \tag{4.25}$$

Hence, for example, we would write $\left|K^+\right\rangle = \left|M_1^3\right\rangle$, or $\left|\pi^0\right\rangle = \frac{1}{\sqrt{2}}\left|M_1^1\right\rangle - \frac{1}{\sqrt{2}}\left|M_2^2\right\rangle$. Note that the factors in front are always chosen to be normalized to one (so that $u_i^j u_i^{j*} = 1$) and they are also traceless in every case, as they must be.

Similarly, we will write the eight lightest spin-1/2 baryons, and the ten lightest spin-3/2 baryons, in the form

$$v_j^i \left| B_i^j \right\rangle \quad \text{and} \quad w^{ijk} \left| B_{ijk}^* \right\rangle \tag{4.26}$$

The values for the v's are very similar to (4.25):

$$\begin{array}{llll} n^0 : v_3^2 = 1, & p^+ : v_3^1 = 1, & \Xi^- : v_1^3 = 1 & \Xi^0 : v_2^3 = 1 \\ \Sigma^- : v_1^2 = 1, & \Sigma^0 : v_1^1 = -v_2^2 = \frac{1}{\sqrt{2}}, & \Sigma^+ : v_2^1 = 1, & \Lambda^0 : v_1^1 = v_2^2 = \frac{1}{\sqrt{6}}, v_3^3 = -\frac{2}{\sqrt{6}} \end{array} \tag{4.27}$$

For the spin-3/2 baryons, we must make w completely symmetric. The appropriate assignments are:

$$\begin{gathered} \Delta^- : w^{222} = 1, \quad \Delta^0 : w^{221} = w^{212} = w^{122} = \tfrac{1}{\sqrt{3}}, \quad \Delta^+ : w^{112} = w^{121} = w^{112} = \tfrac{1}{\sqrt{3}}, \quad \Delta^{+++} : w^{111} = 1, \\ \Sigma^{*-} : w^{223} = w^{232} = w^{322} = \tfrac{1}{\sqrt{3}}, \quad \Sigma^{*+} : w^{113} = w^{131} = w^{311} = \tfrac{1}{\sqrt{3}}, \\ \Sigma^{*0} : w^{123} = w^{132} = w^{213} = w^{231} = w^{312} = w^{321} = \tfrac{1}{\sqrt{6}}, \\ \Xi^{*-} : w^{233} = w^{323} = w^{332} = \tfrac{1}{\sqrt{3}}, \quad \Xi^* : w^{133} = w^{313} = w^{331} = \tfrac{1}{\sqrt{3}}, \quad \Omega^- : w^{333} = 1. \end{gathered} \tag{4.28}$$

For example, we could write $\left|\Sigma^{*+}\right\rangle = \frac{1}{\sqrt{3}}\left(\left|B_{113}^*\right\rangle + \left|B_{131}^*\right\rangle + \left|B_{311}^*\right\rangle\right)$.

E. Masses and the Gell-Mann – Okubo Formulas

We would like to make predictions about the masses and interactions of the various particles that fit together in an SU(3) multiplet. Let's start with the masses. For example, suppose we want to know the mass of a spin-3/2 baryon. We would expect this mass to be governed by some term or terms in the Hamiltonian, so, for a fermion, we expect there to have

$$m_{B^*} = \left\langle B^* \middle| H \middle| B^* \right\rangle \tag{4.29}$$

Now, if we write this baryon in the form $\left|B^*\right\rangle = w^{ijk}\left|B^*_{ijk}\right\rangle$, so that $\left\langle B^*\right| = \left\langle B^*_{ijk}\right| w^{ijk*}$. Now, when we put it all together, we would anticipate that we could write something like

$$m_{B^*} = a \cdot w^* \cdot w \tag{4.30}$$

where a is, for the moment, a vaguely defined constant, and the dot product means we are somehow multiplying the components of w by components of the complex conjugate. We'll get more specific in a moment.

Now, in expression (4.30), we would like this matrix element to be some product of the corresponding tensors that is invariant under SU(3). Now, w has three up indices, but w^* will transform like the complex conjugate representation. Basically, since the complex conjugate of the 3 is the $\overline{3}$, this means that w^* acts as if it has three down indices. I'll denote this by writing this expression in the form

$$m_{B^*} = a^{ijk}_{lmn} w^\dagger_{ijk} w^{lmn} \tag{4.31}$$

where $w^\dagger_{ijk}$ just means the same thing as w^{ijk*}. Now, if we knew the coefficients a^{ijk}_{lmn} (there are only 729 of them), we'd be all set. But we now argue that a^{ijk}_{lmn} must be an invariant tensor, which means that it can only be the product of δ_i^j's, ε_{ijk}'s or ε^{ijk}'s. In this case, because the number of indices up and down are equal, we won't use the ε's. If we write a^{ijk}_{lmn} as the product of three δ's, we can attach the up indices to the down indices in six different ways, something like this:

$$m_{B^*} = aw^\dagger_{ijk} w^{ijk} + bw^\dagger_{ijk} w^{jik} + cw^\dagger_{ijk} w^{ikj} + dw^\dagger_{ijk} w^{kij} + ew^\dagger_{ijk} w^{jki} + fw^\dagger_{ijk} w^{kji} \tag{4.32}$$

However, recall that the w's are completely symmetric, so in fact all these terms are identical. We therefore have

$$m_{B^*} = aw^\dagger_{ijk} w^{ijk} \tag{4.33}$$

We can now predict the mass of any of the B^*'s. For example, for Δ^+ we would have

$$m_{\Delta^+} = a\left(w^\dagger_{112} w^{112} + w^\dagger_{121} w^{121} + w^\dagger_{211} w^{211}\right) = a\left(\tfrac{1}{3} + \tfrac{1}{3} + \tfrac{1}{3}\right) = a \tag{4.34}$$

We can now repeat this computation for every particle in the SU(3) multiplet. The conclusion is simple, they all have the same mass, since they all have $w^\dagger_{ijk} w^{ijk} = 1$. In fact we already knew this. But it was worth showing how it works out in this simple case.

Fortunately, we can do better. Although we knew going in that SU(3) symmetry isn't really a good symmetry for the masses, we *do* expect isospin to be a good symmetry for the masses. We therefore conjecture that the mass part of the Hamiltonian has an additional term which commutes with isospin, but not with all of SU(3). If we look over our eight SU(3) generators, and recognize that T_1, T_2, and T_3 generate isospin, then we realize there is only one generator that commutes with isospin, namely T_8. We therefore conjecture that the *real* formula for the mass looks something like

$$m_{B^*} = a^{ijk}_{lmn} w^\dagger_{ijk} w^{lmn} + b^{ijkp}_{lmno} w^\dagger_{ijk} w^{lmn} \left(T_8\right)^o_p \tag{4.35}$$

Now, we've already worked out the form for the a term. For the b term, we have to connect the two indices on the T_8 with something. We can't connect them with each other (that would yield the trace of T_8, which is zero), so clearly one of the indices must attach itself to w and the other to $w^\dagger$. It doesn't matter which one connects to which, because the indices on w and $w^\dagger$ are completely symmetric. The remaining indices must now connect with each other so that

$$m_{B^*} = a w^\dagger_{ijk} w^{ijk} + b w^\dagger_{ijk} w^{ijl} \left(T_8\right)^k_l \tag{4.36}$$

We can use (4.36) to predict relations between the various components of the 10. For example, to find the mass of the Δ^+, we would have

$$\begin{aligned} m_{\Delta^+} &= a w^\dagger_{ijk} w^{ijk} + b\left(T_8\right)^k_l w^\dagger_{ijk} w^{ijl} = a + b\tfrac{1}{2\sqrt{3}}\left(w^\dagger_{ij1} w^{ij1} + w^\dagger_{ij2} w^{ij2} - 2 w^\dagger_{ij3} w^{ij3}\right) \\ &= a + b\tfrac{1}{2\sqrt{3}}\left(w^\dagger_{121} w^{121} + w^\dagger_{211} w^{211} + w^\dagger_{112} w^{112}\right) = a + \tfrac{1}{2\sqrt{3}} b \end{aligned} \tag{4.37}$$

Now, because isospin is still a valid symmetry, the mass of the Δ^0, Δ^-, and Δ^{++} work out to exactly the same thing. In retrospect, it would have been easier to work with the Δ^- or Δ^{++}, since these would have had only one term in the sum. We can proceed to find masses of some of the other particles in the SU(3) multiplet. We find

$$m_{\Sigma^*} = a, \quad \text{and} \quad m_{\Xi^*} = a - \tfrac{1}{2\sqrt{3}} b. \tag{4.38}$$

Of course, we have no idea what a and b are. But we can find one relation between them. Combining (4.37) and (4.38), it is clear that

$$2 m_{\Sigma^*} = m_{\Xi^*} + m_\Delta . \tag{4.39}$$

This is one of the Gell-Mann – Okubo mass formulas. Looking up the masses in the previous chapter, and using the average for the multiplet, you will find that the left side is about 2770 MeV/c^2, and the right about 2767 MeV/c^2, or an error around 0.1%. Not bad!

Using (4.36), Gell-Mann and Okubo also managed to calculate a formula for the as-yet undiscovered Ω^- particle. Based on this mass, they predicted that it would not be able to decay strongly. Predicting a new strongly interacting particle together with its spin, mass, and some aspects of its decay was considered a great triumph for the theory.

Can a similar formula be found for octets, such as the spin-1/2 baryons, or the spin-0 mesons? It can, but the computation is a bit more difficult. Since the mesons are bosons, the matrix element will describe the *square* of the mass. Including from the start the T_8 mass terms, we would expect something akin to

$$m_M^2 = a u^{\dagger j}_{i} u^{i'}_{j'} + b \left(T_8\right)^{l'}_{l} u^{\dagger j}_{i} u^{i'}_{j'} \tag{4.40}$$

where $u^{\dagger j}_{i} \equiv u^{i*}_{j}$. We now need to figure out all the possible ways of connecting the indices. On the first term, we can't connect an index to itself because $u^i_i = 0$. So there is only one way to connect the indices. The first term ends up just contributing a constant *a* to the mass squared, and if this were the only term, all components of the meson 8 would have the same mass. The second term turns out to be more complicated. No index can be connected to the other index on the same factor, but this still leaves more than one way of connecting them, so in fact the *b* term actually stands for more than one term, each with their separate coefficient. It is nonetheless possible to eliminate all the parameters, and to find a simple relationship between the various isospin multiplets of the mesons, which turns out to work pretty well. A similar expression can be found relating the spin-1/2 baryons, though in this case it will be a relationship between the masses, not their squares.

F. Interactions

Although SU(3) is not a very good symmetry to describe the *masses* of the baryons and mesons, it would be expected to work better for the interactions. For example, consider the matrix elements describing the decays of the heavy spin-3/2 baryons to the lighter spin-1/2 baryons plus the spin-0 mesons. We would expect such decays to be governed by matrix elements of the form $\left\langle BM \middle| H \middle| B^* \right\rangle$. Naively, we would then anticipate that the decay rates would be proportional to these matrix elements, so that

$$\Gamma\left(B^* \to BM\right) \propto \left|\left\langle BM \middle| H \middle| B^* \right\rangle\right|^2 \tag{4.41}$$

The problem with this argument is that the computation of this decay rate, accomplished, say, with the help of Fermi's Golden rule, would result in various kinematic factors which would depend on the masses of the particles involved. In some cases, a matrix element may be non-zero, but the masses make the decay impossible. Even when a decay is allowed, the matrix element may have some momentum dependence, which group theory cannot predict, and hence we cannot proceed all the way to ratios of decay rates. Nonetheless, particle physics can give us some good guesses as to the momentum dependence, even when we do not understand the underlying theory, and combined with group theory, we can make good estimates of various decay rates.

Since this is a group theory course, we will focus *exclusively* on the matrix elements, and leave to particle physicists the work of interpreting the resulting ratios. Let us attempt to write down the most general form of this matrix element $\left\langle BM \middle| H \middle| B^* \right\rangle$. If, in the usual way, we write $\left|B^*\right\rangle = w^{ijk}\left|B^*_{ijk}\right\rangle$, $\left|M\right\rangle = u^j_i\left|M^i_j\right\rangle$, and $\left|B\right\rangle = v^j_i\left|B^i_j\right\rangle$. Then the matrix elements will be of the form

$$\left\langle BM \middle| H \middle| B^* \right\rangle \sim \left(v^\dagger\right)^i_n \left(u^\dagger\right)^j_p w^{klm} \tag{4.42}$$

We will assume the interaction is SU(3) invariant, which means that we need to get rid of all the indices in (4.42) using invariant tensors. There are too many indices up, so to get rid of them, we need to include ε_{abc}. Because *w* is completely symmetric on its three indices, we can attach no more than one of the three indices of ε_{abc} to it, and by the symmetry of *w*, it doesn't matter which one. It follows that the remaining indices must attach to the upper indices on $v^\dagger$ and $u^\dagger$. It doesn't matter which index attaches to which, because ε_{abc} is completely anti-symmetric.

This leaves still two lower indices, one each on $v^\dagger$ and $u^\dagger$, and two upper indices on *w*. These must be connected together, and again, since *w* is completely symmetric, it doesn't matter which one connects to which. So in summary, there is only one type of term in (4.42) which doesn't vanish, namely

$$\left\langle BM \middle| H \middle| B^* \right\rangle = a\left(v^\dagger\right)^i_l \left(u^\dagger\right)^j_m w^{klm}\varepsilon_{ijk} \tag{4.43}$$

We can then compute, in a straightforward way, the corresponding amplitudes.

For example, let's find the relative size of the matrix elements for the two processes $\Delta^{++} \to p^+\pi^+$ and $\Sigma^{*+} \to \Lambda^0\pi^+$. These are both sensible decay rates to consider, since they conserve both charge and hypercharge. Looking at equations (4.27), (4.29) and (4.30), we can extract the relevant components in tensor notation. For the Δ^{++}, the p^+, and the π^+, we have only one non-vanishing component, which makes the computation easy, and we quickly find

$$\left\langle p^+\pi^+ \middle| H \middle| \Delta^{++} \right\rangle = a\left(v^\dagger\right)^3_1 \left(u^\dagger\right)^2_1 w^{111}\varepsilon_{3j1} = a\left(v^\dagger\right)^3_1 \left(u^\dagger\right)^2_1 w^{111}\varepsilon_{321} = -a1\cdot 1\cdot 1 \tag{4.44}$$

For the Σ^{*+} there are three components, as there are for the Λ^0, but it isn't too hard to figure out.

$$\begin{aligned}
\left\langle \Lambda^0\pi^+ \middle| H \middle| \Sigma^{*+} \right\rangle &= a\left(v^\dagger\right)^i_l \left(u^\dagger\right)^j_m w^{klm}\varepsilon_{ijk} = a\left(v^\dagger\right)^i_l \left(u^\dagger\right)^2_1 w^{kl1}\varepsilon_{i2k} \\
&= a\left[\left(v^\dagger\right)^1_1 \left(u^\dagger\right)^2_1 w^{311}\varepsilon_{123} + \left(v^\dagger\right)^3_3 \left(u^\dagger\right)^2_1 w^{131}\varepsilon_{321}\right] \\
&= a\left[\tfrac{1}{\sqrt{6}}\cdot 1\cdot \tfrac{1}{\sqrt{3}} - \left(-\tfrac{2}{\sqrt{6}}\right)\cdot 1\cdot \tfrac{1}{\sqrt{3}}\right] = \tfrac{1}{\sqrt{2}}a
\end{aligned} \tag{4.45}$$

We therefore have

$$\frac{\left|\left\langle \Lambda^0\pi^+ \middle| H \middle| \Sigma^{*+} \right\rangle\right|^2}{\left|\left\langle p^+\pi^+ \middle| H \middle| \Delta^{++} \right\rangle\right|^2} = \frac{1}{2} \tag{4.46}$$

Unfortunately, we cannot then conclude anything about the relative decay rates, because we don't know how to do the remaining kinematic integrals.

G. Quarks

Gell-Mann SU(3) helped relate and organize the vast number of strongly interacting particles that were at the time being discovered, but it left several patterns unexplained. It turns out that all of the particles discovered fit into just a few simple irreps of SU(3). The mesons always occur in either the 8 or 1 irreps of SU(3). The baryons occur in either the 8, 10, or 1 irrep of SU(3). The anti-baryons, being the anti-particles of the baryons, occur in either the 8, $\overline{10}$ or 1 irrep of SU(3). Why these irreps, and no others? The baryons are always fermions, while the mesons are always bosons. At the time, these were simply disconnected facts, with no underlying explanation. In 1964, presumably while playing with SU(3), Murray Gell-Mann and George Zweig independently came up with a simpler explanation. They proposes that all strongly interacting particles are actually made of fundamental particles called *quarks*. There would be three quarks, which they named up, down, and strange (u, d, or s for short). Together they would form a 3 irrep of SU(3), as sketched in Fig. 4-11. In other words, everything is built out of quarks q_i, where

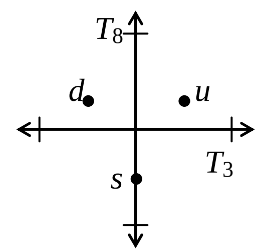

Figure 4-11: The three quarks form a 3 representation of SU(3).

$$|q_1\rangle = |u\rangle, \quad |q_2\rangle = |d\rangle, \quad |q_3\rangle = |s\rangle. \tag{4.47}$$

The rules for combining quarks were simple. Three quarks could be combined together to make a baryon. Three anti-quarks could be combined together to make an anti-baryon. And a quark and an anti-quark could be combined to make a meson. Following these rules, it is easy to understand why only certain combinations occur. For example, because $3 \otimes \overline{3} = 8 \oplus 1$, it follows that one should only find the 8 and 1 irreps for mesons. If we further assume that quarks are spin ½, it follows that the lightest mesons (with no orbital angular momentum) should be spin 0 or 1. Similarly, a combination of three quarks should fall into a $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$, and hence must be in the 1, 8, or 10 irrep. They should also have spin ½ or spin 3/2.

The quarks were a bit odd. Their charges were easy to work out: the up had charge $+\frac{2}{3}$, while the down and strange each had charge $-\frac{1}{3}$. Such fractionally charged particles had never been seen, and nearly half a century later a quark in isolation has still never been observed. Despite these and other problems, we now strongly believe that quarks are real. The three quarks are assumed to have identical strong interactions. Isospin symmetry relates the up and down quarks, both of which have very small inherent mass, making this symmetry a relatively good approximation. The strange quark, in contrast, is 150 MeV/c^2 or so heavier, which is why, for example, there is an increasing progression of masses from the Δ's to the Ω^- as we increase the fraction of strange quarks inside the various spin 3/2 baryons.

We have since discovered there are other, heavier quarks as well: the charm, bottom, and top quarks (charge $+\frac{2}{3}$, $-\frac{1}{3}$, and $+\frac{2}{3}$ respectively). Although these quarks are expected to have the same strong interactions as the three light ones, the masses are so much greater that they do not increase the apparent symmetries of particle physics.

H. Color

There is one glaring problem with quarks. Consider, for example, the Δ^{++}, which must actually consist of three up quarks, $\left|\Delta^{++}\right\rangle = \left|uuu\right\rangle$. Since it is a "light" baryon, it must have no internal angular momentum, which suggests the spatial wave function is completely symmetric. Since it is spin 3/2, the spins must all be aligned, which tells us the spin wave function is also completely symmetric. And there are three up quarks. The conclusion is that the three quarks in the Δ^{++} are in a completely symmetric state. Indeed, you can very well account for *all* the light baryons by simply assuming that there is an arbitrary rule that all quarks must be in a completely symmetric state. This contradicts fundamental principles of particle physics, that spin-1/2 particles *must* be fermions, and have anti-symmetric wave functions. We also have an additional problem: there is no explanation of why quarks combine only as three quarks, three anti-quarks, or quark plus anti-quark.

The solution to the anti-symmetry problem is simple. Let's assume that there is not one up quark, but three, u_1, u_2 and u_3. This additional index is called a *color* index, and the three colors are sometimes called red, green, and blue. All three will be assumed to have *exactly* the same mass, charge, strong interactions, etc. Then we can make a Δ^{++} by assuming that the three up quarks are different, so we make our wave function by writing

$$\left|\Delta^{++}\right\rangle = \tfrac{1}{\sqrt{6}}\left(\left|u_1u_2u_3\right\rangle + \left|u_2u_3u_1\right\rangle + \left|u_3u_1u_2\right\rangle - \left|u_1u_3u_2\right\rangle - \left|u_2u_1u_3\right\rangle - \left|u_3u_2u_1\right\rangle\right). \tag{4.48}$$

This expression can be written more succinctly by writing

$$\left|\Delta^{++}\right\rangle = \tfrac{1}{\sqrt{6}}\left|u_iu_ju_k\right\rangle \varepsilon^{ijk}. \tag{4.49}$$

Since the "quark" part of the wave function is now anti-symmetric, the rest of the wave function (spin, for example) can be completely symmetric, without violating the rule that fermions must be in anti-symmetric wave functions. Color is assumed to apply to all three (or six) quarks, so there are a total of nine (or eighteen) quarks.

Although this works, it seems terribly awkward. We have suddenly tripled the number of particles. And, come to think of it, why should the three up quarks have exactly the same mass? They shouldn't – unless there is actually some sort of symmetry connecting them. If this were a course on group theory, we would now discuss this symmetry, . . . oh wait, this is a course on group theory!

Color is going to be assumed to be not just a label, but also a symmetry, specifically, the symmetry SU(3). It cannot be emphasized too strongly that color SU(3) is different from Gell-Mann SU(3). Gell-Mann SU(3) is an approximate symmetry that relates the up, down, and strange quarks. It is violated by the mass terms for the three quarks, as well as electromagnetic and weak interactions. In contrast, color SU(3) is an *exact* symmetry that relates the three colors of up (or down, or . . .) quarks. The masses are exactly equal, the charges of these three up quarks is exactly the same, and so on. The three up quarks lie together in the 3 irrep of SU(3).

Now, why is the particular combination (4.49) favored for the color part of the wave function? What is special about this combination? Recall that ε^{ijk} is an invariant tensor of SU(3). As such, this particular combination of colors lies in the 1 irrep of color SU(3). We are now ready to state the fundamental assumption about color: **The only combinations allowed are those that lie in the trivial (1) irrep of color SU(3)**.

Now, if each type of quark has three colors, and lies in the 3 irrep of SU(3), then the corresponding anti-quark lies in the $\overline{3}$ irrep of SU(3). Now, according to the colorless assumption, what combinations of quarks and anti-quarks are allowed? There will be one combination for every invariant tensor of SU(3). This suggests three combinations of colors:

$$\left|B\right\rangle = \tfrac{1}{\sqrt{6}}\left|q_i q_j q_k\right\rangle \varepsilon^{ijk}, \quad \left|\overline{B}\right\rangle = \tfrac{1}{\sqrt{6}}\left|\overline{q}^i \overline{q}^j \overline{q}^k\right\rangle \varepsilon_{ijk}, \quad \left|M\right\rangle = \tfrac{1}{\sqrt{3}}\left|q_i \overline{q}^j\right\rangle \delta^i_j \tag{4.50}$$

Hence, our hypothesis leads *naturally* to the following conclusion: quarks bind only in the combinations three quarks (baryon), three anti-quarks (anti-baryon), or quark plus anti-quark (meson). These are exactly the combinations observed in nature.

Of course, it may seem like we are still running in circles. We started with Gell-Mann SU(3), which seemed to work pretty well, then we showed how mathematically we could get everything to work by assuming things are made of quarks, though we had to make up some arbitrary rules (symmetric wave functions, only certain numbers of quarks, etc.). We fixed a significant problem by then tripling the number of quarks, and added another symmetry (color SU(3)) to explain some patterns we saw. Finally, we added a brand new rule (only colorless combinations), without any justification other than, "it works." And yet we are at the threshold of explaining the *real* nature of the strong force.

Let me lead you to it by analogy. Suppose someone had never heard of electromagnetic forces, but was doing experiments with atoms. She finds certain quantities are conserved, and starts assigning a "charge" of -1 to the electron, +2 to the helium nucleus, etc. She describes this in terms of a U(1) symmetry. Upon further experimentation, she discovers that atoms naturally tend to fall in the trivial representation of U(1); they tend to prefer to be chargeless. Puzzled, she thinks about this, and suddenly the idea hits her that perhaps charge is *not* just some abstract concept, but actually has a force associated with it. Chargeless combinations are preferred, not because of some arcane rule, but because "opposites attract" and therefore combinations in the trivial representation of U(1) tend to have the least energy. Hence electromagnetism is discovered.

The analogy is as follows. Imagine that color SU(3) is not just a symmetry, but a gauge force, analogous to electromagnetism. Color-neutral combinations; that is, combinations which are in the trivial irrep of color SU(3) will naturally have the lowest energy. This gauge force is much stronger than electromagnetism; in electromagnetism, the strength of the force is described in terms of the fine structure constant $\alpha = \frac{1}{137}$, but the strong force has $\alpha_S \approx 10$.

Just as the electromagnetic field has a particle (the photon) associated with it, there will be particles associated with color SU(3). The way it works out is that there will be one strongly interacting gauge particle (gluon) associated with each of the eight generators of SU(3). Each of the gluons has spin 1, like the photon. *Unlike* the photon, because the generators of SU(3) do *not* commute with each other, the gluons interact with each other in complicated ways. This, together with the non-perturbative strength of the strong coupling, make detailed calculations very difficult. For example, suppose a quark and anti-quark of opposite colors were near each other. We would expect there to be color flux lines (akin to electric field lines for electromagnetism) connecting them. However, color flux lines (unlike electric field lines) actually have color themselves, and are attracted to each other. This causes the color flux lines to bunch together to form a flux tube, as sketched in Fig. 4-12 at right. As you attempt to pull a quark and anti-quark apart, the flux tube does not get larger and more spread out, instead, it stays constant size, and hence the force attracting the quark and anti-quark (estimated to be about one ton!) does *not* diminish with distance. Eventually, the color flux tube has sufficient energy associated with it that it can simply create a quark anti-quark pair, and we simply end up with two separate mesons, as sketched in Fig. 4-13 at left. Hence it is believed to be impossible to actually separate a quark by itself. Quarks and anti-quarks only appear in colorless combinations.

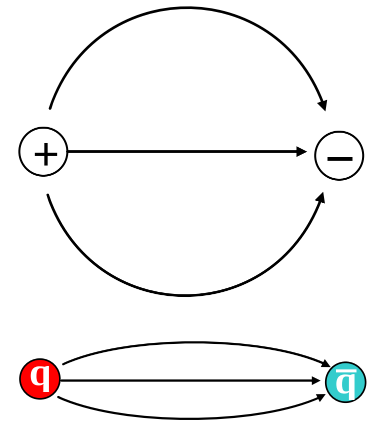

Figure 4-12: Two electric charges have widely separated electric field lines (top). In contrast, the color field lines tend to have field lines tightly bundled into a "flux tube."

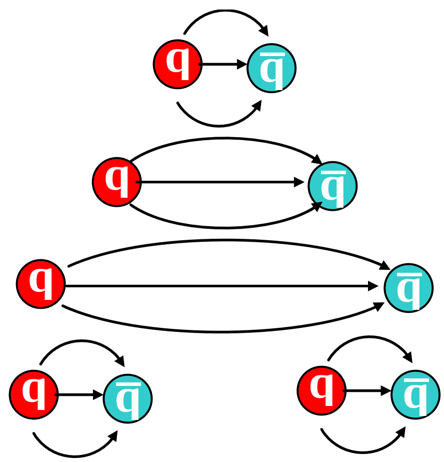

Figure 4-13: An attempt is made to separate a pair of quarks in a meson. An initially small meson (top) has its two quarks separated by increasing distance (middle), but the force between then remains roughly constant. A flux tube connects them whose strength is roughly independent of distance. Eventually, the flux tube breaks, (bottom) producing a new quark anti-quark pair, resulting ultimately in two mesons.

This color SU(3) is believed to be the *true* strong force. The "strong force" that we encounter in nuclear physics, for example, is simply a short-range effect caused by the proximity of two color-neutral combinations, akin to the van der Waal's interaction between a pair of neutral atoms.