High Energy Physics

Ananda Dasgupta
Lecture notes taken by Anindya Sengupta

April 27, 2016

Department of Physical Sciences
Indian Institute of Science Education and Research, Kolkata

I have tried to follow ADG's lectures as closely as possible. In a few places, I have taken the freedom to elaborate some concepts spending more words than ADG did in the class. I was absent in a few lectures. Sohitri has kindly provided me with excellent handwritten notes for those lectures. Thank you Sohitri.

Please notify me of typographical/conceptual errors if you find any. You can contact me at tuhin039@gmail.com and/or +91-9836874337.

– Anindya

1 Week 1

1.1 Lecture 1 : January 4, 2016

High Energy Physics (HEP) is the study of elementary particles and their interactions. Elementary particles are those which are made of no sub-particles. Interactions of particles could be of several kinds : particle decay, scattering, pair creation or pair annihilation etc. We shall be studying the kinematics of decay and scattering processes in the first few lectures.

Suppose that a particle may decay in several ways, so that its different decay processes give birth to different species of product particles. By specifying the identities of the decay products we can be sure which one of these processes we are talking abut. A specific decay process of a particle, identified by the products, is called a **decay channel**. In a channel, momenta of the product particles may vary.

Two Body Decay :

Let's denote the identities of the particles by the letter a with numerical subscripts as and when necessary. A channel of a two body decay would be

$$a(p) \rightarrow a_1(p_1) + a_2(p_2) \tag{1}$$

The momenta of the particles are given in the brackets. Conservation of 4-momentum gives

$$p = p_1 + p_2 \tag{2}$$

Taking a Lorentz invariant square on both sides gives

$$p^2 = p_1^2 + p_2^2 + 2p_1.p_2$$

$$\Rightarrow m^2 = m_1^2 + m_2^2 + 2\left(E_1E_2 - \vec{p}_1.\vec{p}_2\right) \tag{3}$$

We can now go to the rest frame of the mother particle in which

$$\vec{p} = 0 = \vec{p}_1 + \vec{p}_2 \tag{4}$$

Let,

$$\vec{p}_1 = -\vec{p}_2 = \vec{P} \tag{5}$$

$$\Rightarrow E_1E_2 - \vec{p}_1.\vec{p}_2 = \frac{1}{2}\left(m^2 - m_1^2 - m_2^2\right)$$

$$\Rightarrow \sqrt{\left|\vec{P}\right|^2 + m_1^2}.\sqrt{\left|\vec{P}\right|^2 + m_2^2} + \left|\vec{P}\right|^2 = \frac{1}{2}\left(m^2 - m_1^2 - m_2^2\right)$$

Squaring both sides, and a bit of algebra, yield

$$E_1^2 = \left(\frac{m^2 + m_1^2 - m_2^2}{2m}\right)^2$$

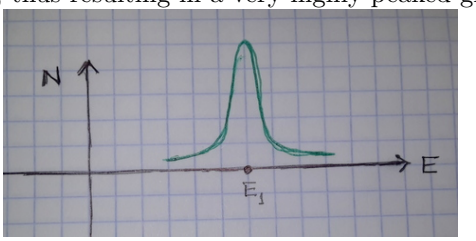
Taking the positive[1] square root, we find

$$E_1 = \frac{1}{2m}\left(m^2 + m_1^2 - m_2^2\right) \tag{6}$$

Using $m = E = E_1 + E_2$(or by simply interchanging the subscripts 1 and 2), one finds,

$$E_2 = \frac{1}{2m}\left(m^2 + m_2^2 - m_1^2\right) \tag{7}$$

So we find that, given the masses of the particles involved in a two body decay, the total energy is shared in a precise fashion between the daughter particles. If we perform a decay experiment and keep measuring the energy of one of the daughter particles, say a_1, we shall see that almost all of them have the energy E_1, thus resulting in a very highly peaked graph.



The small spread originates from what is known as the **Doppler broadening**. This occurs because of the thermal motion of the mother particle a.

If we are not in the rest frame of a, we can do one of two things – employ $p = p_1 + p_2$ in the new frame and solve a complicated quadratic equation to get $E_{1,2}$; or we can solve the problem in the rest frame of a and then boost the solution to go back to the new frame. If we are following the first approach, the calculations will be a little easier if we square[2] $p - p_1 = p_2$.

$$(p - p_1)^2 = p_2^2 \Rightarrow m^2 + m_1^2 - 2\left(EE_1 - \vec{p}.\vec{p_1}\right) = m_2^2 \tag{8}$$

When a_1 is a massless particle, say, a neutrino, this formula gives

$$m^2 - 2\left(EE_1 - |\vec{p}|E_1\cos\theta\right) = m_2^2 \tag{9}$$

where θ is the angle between $\vec{p}$ and the direction of motion of the initial particle. As an example, consider α decay :

$$X \to Y + \alpha \tag{10}$$

[1] $$E = E_1 + E_2$$
, where E, E_1, $E_2$are all positive numbers. $E = m$ in the rest frame of a. Now, $E_2 \geq m_2 \Rightarrow m \geq m_2$. It follows, $m^2 + m_1^2 - m_2^2 \geq 0$.

[2] In a large variety of problems in special relativity, one can simplify the calculations by cleverly choosing to compute some Lorentz invariant squares. However, sometimes it is better to take the brute force approach rather than split one's hair looking for a nice Lorentz scalar.

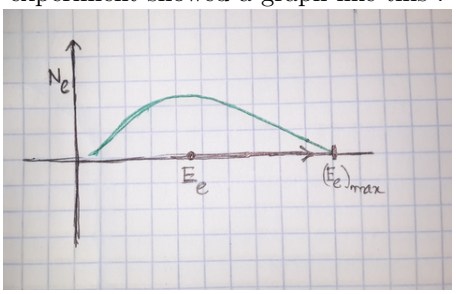
In the rest frame of the mother particle X, rest energy of X will equal the sum of the rest energies and kinetic energies of the product particles. We define a quantity, called the Q-**factor** of the process :

$$Q = m_X - m_Y - m_\alpha \tag{11}$$

Q-factor measures the excess energy in X compared to the combined rest mass of Y and $\alpha \Rightarrow$ the energy available for "motion of Y and α".

If the N_{a_1} vs. E_1 curve of a 2-body decay has multiple peaks, one argues that the product particles appear via intermediate steps : $X \to Y^* \to Y$. In such cases, α decay is usually accompanied by γ-decay.

In the beginning, it was assumed that β-decay was a 2-body decay. However, the experiment showed a graph like this :



This was contradictory to the assumption. Many propositions came forward in order to explain it, some as outlandish as to demand violation of 4-momentum conservation for β-decays. Pauli proposed (1930) that β-decay is not a 2-body decay and proposed the existence of a yet to be discovered particle – the anti-neutrino.

$$X \to Y + e + \bar{\nu}_e \tag{12}$$

We will see soon that the sharing of energy between three daughter particles cannot exactly be fixed and that contributes to the shape of the graph above. Antineutrino was experimentally detected[3] for the first time 26 years after Fermi's prediction!

Three body decay :

In case of two body decays, momentum conservation in the rest frame of the mother particle imply that the 3-momenta of the two daughters have to be exactly equal and opposite to each other. However, if we have three daughter particles, then in the rest frame of the mother particle three 3-momentum vectors will have to add up to zero. $\Rightarrow$ The 3-momenta of the daughters will have to form a triangle. A triangle can be constructed out of three independent parameters – lengths of two sides and the angle between them, for example. One of these will be taken away by the equation for energy $\left(p^0\right)$ conservation. We are still left with two independent parameters which can be varied, as a result

[3] In 1956, by Cowan and Reines.

of which the energy distribution between the daughter particles is not as boring as in the two body case.

$$a\,(p) \rightarrow a_1\,(p_1) + a_2\,(p_2) + a_3\,(p_3) \tag{13}$$

In the rest frame of a,

$$\vec{p} = 0 = \vec{p}_1 + \vec{p}_2 + \vec{p}_3 \tag{14}$$

With this much information, we cannot solve exactly for $E_{1,2,3}$, but we can find out the allowed ranges of energies of the daughter particles. Let's start with a_3.

$$(E_3)_{min} = m_3 \tag{15}$$

For a two body decay in the rest frame of the mother particle, none of the daughter particles can be at rest[4]. But this can happen in a three body decay – a_3 **can be** at rest (in the rest frame of a), while the 3-momenta of a_2 and a_3 are equal and opposite to each other. In such a case a_3 will achieve the minimum energy stated above. E_3 will be maximum when $|\vec{p}_3|$ is maximum. Now, $\vec{p}_3 = -\left(\vec{p}_1 + \vec{p}_2\right)$ implies that $|\vec{p}_3|$ is maximum when $\vec{p}_1$ and $\vec{p}_2$ point in the same direction. In such a case,

$$(E_3)_{max} = \frac{m^2 + m_3^2 - (m_1 + m_2)^2}{2m} \tag{16}$$

Derive this result yourself as an exercise. If there were a two body decay in which a daughter particle of mass m_3 goes in one direction and another of mass $(m_1 + m_2)$ goes in the opposite direction, obviously with momenta of equal magnitude, then the daughter particle with mass m_3 would come out with exactly the amount of energy stated in equation (16).

In special relativity, kinetic energy is defined as $T \equiv E - m$, that is the excess in total energy over the mass. For the beta decay considered in (12), we use the approximation that the masses of X and Y are much bigger than the Q-factor of the process and also the masses of the electron and the anti-neutrino. In such a case, equation (16) implies that the maximum kinetic energy achievable by the electron is given by

$$(T_e)_{max} \simeq Q - m_{\bar{\nu}_e} \tag{17}$$

Since anti-neutrino is very difficult to detect but electrons are not, one might hope to experimentally measure $(T_e)_{max}$, Q and from there derive an estimate for the mass of an anti-neutrino. However, this cannot be done in practice because the mass of anti-neutrino $\left(\sim 0.320 \pm 0.081 \text{ eV/c}^2\right)$ is very small compared to the error-bars of β-decay experiments. The energy equations we discussed so far correspond to the rest frame of the mother particle. These energies will of course differ in other frames. To derive results or conclusions that are frame

[4]Otherwise, the other daughter would also have to be at rest and that would mean $m = m_1 + m_2$. It is very unlikely that there will exist three particles whose masses have this kind of an exact relation.

independent, one uses some specific Lorentz invariant scalars which we define here :

$$\begin{aligned} m_{12}^2 &\equiv (p_1 + p_2)^2 = (p - p_3)^2 \\ m_{23}^2 &\equiv (p_2 + p_3)^2 = (p - p_1)^2 \\ m_{31}^2 &\equiv (p_3 + p_1)^2 = (p - p_2)^2 \end{aligned} \tag{18}$$

Note that these are denoted by the letter m, but they are not masses of some particles. Not all three of them are independent – two are. Because,

$$\begin{aligned} m_{12}^2 + m_{23}^2 + m_{31}^2 &= 3m^2 + m_1^2 + m_2^2 + m_3^2 - 2p^2 \\ m_{12}^2 + m_{23}^2 + m_{31}^2 &= m^2 + m_1^2 + m_2^2 + m_3^2 \end{aligned} \tag{19}$$

1.2 Lecture 2 : January 6, 2016

Reminder : A given channel of a decay is identified by the identities of the particles being produced.

In two body decays such as $a\,(p) \to a_1\,(p_1) + a_2\,(p_2)$, we have a total of six unknowns and four conservation equations. So, there are two independent unknowns. In the rest frame of a these would be the two angles that determine the direction of the straight line along which a_1 and a_2 move, in opposite directions. Since everything else is determined, we know exactly what amount of energy each of $a_{1,2}$ carry. In three body decays such as $a\,(p) \to a_1\,(p_1) + a_2\,(p_2) + a_3\,(p_3)$, there are nine unknowns and four equations, leaving five independent unknowns. This is not enough information to exactly pinpoint $E_{1,2,3}$. Therefore, we try to derive the range of energies the product particles are allowed to carry. In the last class we derived an upper bound and a lower bound for E_3. Today we shall investigate whether all energy values lying between those bounds are accessible to a_3 or not, and also find out the allowed energy values of $a_{1,2}$.

A very important thing to remind ourselves of : these energy values and their bounds etc. have been derived in the rest frame of a. If we keep on working in the same frame, we shall derive the allowed ranges of the energies in this particular frame. However, we would like to derive the bounds of some Lorentz invariant quantities so that they hold irrespective of the frame we choose to be in. To that end, we defined the three quantities m_{12}, m_{23}, m_{31}, of which only two are independent because the sum of their squares is fixed for a given channel :

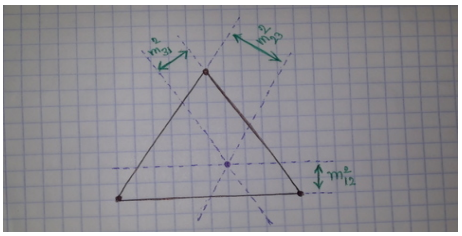
$$m_{12}^2 + m_{23}^2 + m_{31}^2 = m^2 + m_1^2 + m_2^2 + m_3^2$$

Today we shall work out the allowed ranges of the quantities m_{12}^2 etc. and visualize the bounds by plotting them on a plane[5].

[5] Because only two of them are independent, we plot two of them on a plane. The third one gets fixed by equation (19). In other words, if we plot all three of $m_{12,23,31}^2$ in a 3-d plot, the allowed values will lie on a 2-plane by virtue of equation (19).

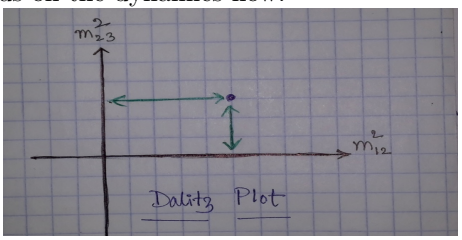
The Dalitz triangle :

There is a nice geometric way of plotting on a plane three non-negative numbers m_{12}^2, m_{23}^2 and m_{31}^2 whose sum is fixed. Consider an equilateral triangle. From any point lying inside (meaning an interior point) the triangle, draw perpendiculars on the three sides. The sum of the lengths of these three perpendiculars is a constant $= \frac{\sqrt{3}}{2}$ (length of a side of the triangle). For a given channel, we choose an equilateral triangle with a side length[6] equal to $\frac{2}{\sqrt{3}}\left(m^2 + m_1^2 + m_2^2 + m_3^2\right)$. Any point inside the triangle now gives us a possible set of values for $\left(m_{12}^2, m_{23}^2, m_{31}^2\right)$:



Note that once we fix the distances of an interior point from two of the sides (say, m_{12}^2 and m_{23}^2), its distance from the third side (m_{31}^2) gets automatically fixed. This triangle is called the Dalitz triangle.

The Dalitz Plot : This is another way of plotting the invariant scalar. Here, we only plot two of them, say m_{12}^2 and m_{23}^2 along two perpendicular axes. For a given channel, the products may come out with various combinations of momenta, resulting in various values of m_{12}^2, m_{23}^2 and m_{31}^2. However, there is a region in the plot, called the **Dalitz region**, outside which the points will not wander, barring experimental errors. Using only kinematics, we will figure out the shape of the boundary of this region. How the points will be distributed inside the Dalitz region is governed by the dynamics of the process. We will not focus on the dynamics now.



m_{12}^2 is a Lorentz scalar. Let's compute its value in the rest frame of a. In the calculations below, frame-dependent quantities will bear the superscript $^{(0)}$ to denote that we are computing in a's rest frame.

$$m_{12}^2 = (p_1 + p_2)^2 = (p - p_3)^2 = m^2 + m_3^2 - 2mE_3^{(0)}$$

[6]With a scaling convention that 1 mass unit of the particles $\equiv$1 length unit of the triangle

$$\Rightarrow \left(m_{12}^2\right)_{max} = m^2 + m_3^2 - 2m\left(E_3^{(0)}\right)_{min}$$

We obtained $\left(E_3^{(0)}\right)_{min}$ in equation (15). Using that,

$$\left(m_{12}^2\right)_{max} = m^2 + m_3^2 - 2mm_3 = (m - m_3)^2 \tag{20}$$

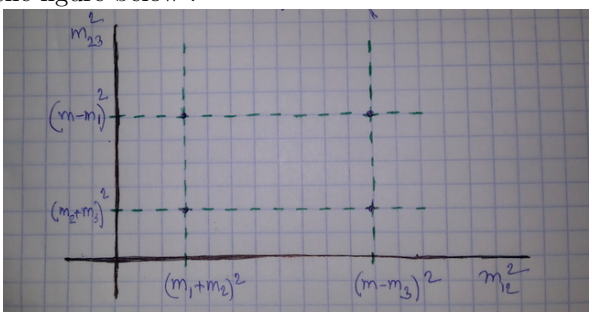
Similarly,

$$\left(m_{12}^2\right)_{min} = m^2 + m_3^2 - 2m\left(E_3^{(0)}\right)_{max} = (m_1 + m_2)^2 \tag{21}$$

Similarly,

$$\begin{aligned} \left(m_{23}^2\right)_{max} &= (m - m_1)^2 \\ \left(m_{23}^2\right)_{min} &= (m_2 + m_3)^2 \end{aligned} \tag{22}$$

Hence, the Dalitz region has to be contained inside the dotted boundary shown in the figure below :



All the points inside this dotted rectangle will not be accessible to the decay events because in determining the bounds of m_{12}^2 we have not put any restrictions on m_{23}^2 and vice versa. So, let us fix a value of m_{12} and see which values of m_{23} are then allowed. We shall choose to do this computation in the CM[7] frame of a_1 and a_2. Frame-dependent quantities computed in this frame will be bear a superscript *.

$$m_{12} = E_1^* + E_2^* \ \left[\because \text{ in the CM frame } (\vec{p}_1 + \vec{p}_2 = 0)\right] \tag{23}$$

$$\therefore p_1^* + p_2^* = \left(m_{12}, \vec{0}\right) \tag{24}$$

In the * frame, we need to calculate

$$m_{23}^2 = \left(p_2^* + p_3^*\right)^2 = m_2^2 + m_3^2 + 2\left(E_2^* E_3^* - \vec{p}_2^* . \vec{p}_3^*\right) \tag{25}$$

[7] CM frame in special relativity does not mean "center of mass" frame. In special relativity, the mass that enters in the dynamics is frame-dependent. Hence, in any frame one has two masses to worry about, the frame-dependent mass that enters the dynamical equations and the rest mass of the particle. However, the center of mass frame used in classical mechanics is also the frame in which the total 3-momentum is zero. And this is the concept of prime importance. Therefore, in special relativity, we replace the phrase "center of mass" with the more appropriate phrase "center of momentum". Hence, CM frame in special relativity $\Rightarrow$ **Center of Momentum frame**, the frame in which the total 3-momentum is zero.

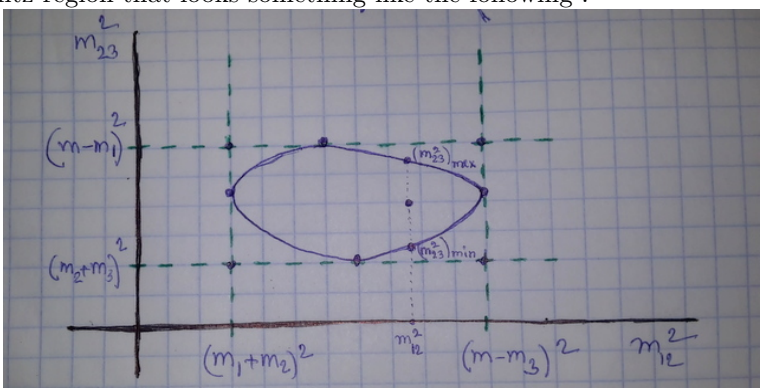
To find $E^*_{2,3}$, start from

$$\begin{aligned} m^2 = (p^*)^2 &= ((p_1^* + p_2^*) + p_3^*)^2 = m_{12}^2 + m_3^2 + 2m_{12}E_3^* \\ m_1^2 = (p_1^*)^2 &= ((p_1 + p_2) - p_2)^2 = m_{12}^2 + m_2^2 - 2m_{12}E_2^* \end{aligned} \tag{26}$$

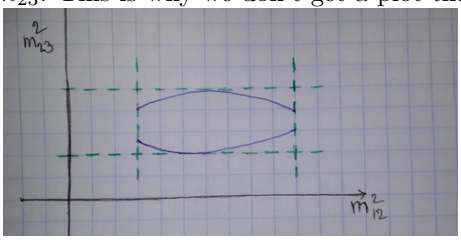
From kinematics alone, we figured out $E^*_{2,3}$ in terms of m_{12}, m_2 and m_3. Now we have everything we need to put into equation (25) and get an expression for m_{23} in terms of m_{12}, m_1, m_2 and m_3. Now, m_{12} fixes $E^*_{2,3}$, or equivalently $|\vec{p}_2^*|$, $|\vec{p}_3^*|$. The only parameter in equation (25) that is not fixed by m_{12} is the angle θ between $\vec{p}_2^*$ and $\vec{p}_3^*$. Since $-1 \leq \cos\theta \leq 1$, therefore

$$(m_{23})^2_{\substack{max \\ min}} = m_2^2 + m_3^2 + 2\left(E_2^* E_3^* \pm |\vec{p}_2^*||\vec{p}_3^*|\right) \tag{27}$$

For an arbitrary choice of the masses of the particles, this would result in a Dalitz region that looks something like the following :



I have deliberately drawn a graph that merges smoothly on the edges (dotted lines) determined by the bounds we obtained earlier. This is exactly what will happen in a real plot. Let's understand why. First of all, if the top edge and the bottom edge be tangents to the boundary of the Dalitz region, then so would be the right and the left edges. Because, one can flip the order of the axes by relabeling the particles and the nature of the plot should not change. In other words, if bounds of m_{23} behave in a certain way for a fixed value of m_{12}, then there is no reason why bounds of m_{12} should behave differently for a given value of m_{23}. This is why we don't get a plot that looks like



Now, the fact that all the edges are tangents to the boundary can be argued this way. The curves for the upper and lower bounds of m_{23}^2 will merge when

$|\vec{p}_2^*||\vec{p}_3^*| = 0$. This can happen when either of the two factors vanish. Carry forward the calculation from here to convince yourself that $|\vec{p}_2^*| = 0$ or $|\vec{p}_3^*| = 0$ happens on the left and the right edges respectively.

The Dalitz plot is a very important tool for experimentalists. Suppose in a decay process you have discovered a new particle whose mass you do not know. You can look at the Dalitz region whose shape depends on the masses of all the particles involved and find an estimate for the mass of the new particle.

Tomorrow we will talk about the kinematics of scattering. Let us clear up a few things about some nomenclature. In the past, scattering experiments used to be performed with fixed targets. This frame in which the target would be kept fixed used to be called the Lab frame. Modern experiments are so designed that we can create the CM frame in the lab itself. However, the previous nomenclature has stuck and many people still call the fixed target frame by the name "Lab frame". We shall take use the same nomenclature. In our convention, **Lab frame** $\Leftrightarrow$ **fixed target frame**. Palash Pal's book is more accurate that way – he uses the names "fixed target frame" and "CM frame".

1.3 Lecture 3 : January 7, 2016

Kinematics of scattering :

Scatterings can be of many types. **Elastic** scatterings are those in which the particles do not change identities – outgoing particles are the same as the incoming particles. **Inelastic** scatterings are those in which the particle content changes after scattering. For simplicity, we will consider 2-to-2 scatterings.

$$a_1\left(p_1\right) + a_2\left(p_2\right) \rightarrow a_1^{'}\left(p_1^{'}\right) + a_2^{'}\left(p_2^{'}\right) \tag{28}$$

We work in the center of momentum frame of the incoming particles. In this frame the formulas look simple. Hence,

$$\vec{p}_1 + \vec{p}_2 = 0 = \vec{p}_1^{'} + \vec{p}_2^{'} \tag{29}$$

The second half of the above equality follows from momentum conservation. Define :

$$\vec{P}_{CM}^{'} = \vec{p}_1^{'} = -\vec{p}_2^{'} \tag{30}$$

Energy conservation implies

$$E_1 + E_2 = E_1^{'} + E_2^{'} \equiv E_{tot} \tag{31}$$

$$\Rightarrow \sqrt{|\vec{P}_{CM}^{'}|^2 + m_1^{'2}} + \sqrt{|\vec{P}_{CM}^{'}|^2 + m_2^{'2}} = E_{tot} \tag{32}$$

Now,

$$E_2^{'} = E_{tot} - E_1^{'}$$

$$\Rightarrow E_2^{'2} = E_{tot}^2 + E_1^{'2} - 2E_1^{'}E_{tot} \tag{33}$$

Also, in the CM frame :

$$E_2^{'2} - E_1^{'2} = m_2^{'2} - m_1^{'2} \tag{34}$$

Comparing equations (33) and (34), we get

$$E_1^{'} = \frac{E_{tot}^2 + m_1^{'2} - m_2^{'2}}{2E_{tot}} \tag{35}$$

Similarly, one can calculate an expression for $E_2^{'}$. Or we can interchange the labels 1 and 2 to get

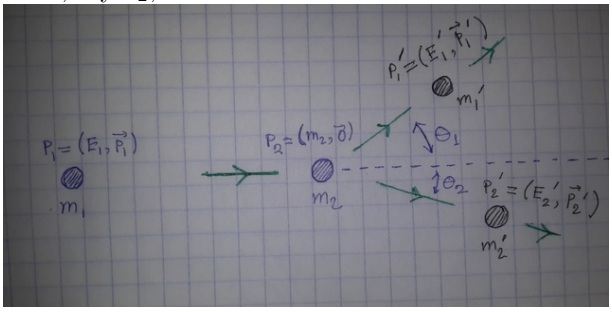
$$E_2^{'} = \frac{E_{tot}^2 + m_2^{'2} - m_1^{'2}}{2E_{tot}} \tag{36}$$

We see that in a $2 \to 2$ scattering, given the total energy of the particles in the initial state, the energies of the particles in the final state are determined. Thus, the magnitudes of the 3-momenta are also determined, but the directions are not. In the CM frame we need to specify the direction of the 3-momentum of one of the final particles. This requires two angular parameters – like the polar and the azimuthal angles on the surface of a sphere.

Consider an **endergonic** reaction for which the total mass of the final particles is bigger than that of the initial particles. An endergonic scattering is obviously an inelastic scattering.

$$m_1 + m_2 > m_1^{'} + m_2^{'} \Rightarrow T_1 + T_2 < T_1^{'} + T_2^{'} \tag{37}$$

For example, consider a two body decay which is endergonic. In the rest frame of the mother particle with mass m_1, energy conservation gives $m_1 = m_1^{'} + m_2^{'} + T_1^{'} + T_2^{'}$. The condition for it being endergonic implies $T_1^{'} + T_2^{'} < 0$, which can never happen (because kinetic energy is non-negative). Clearly, endergonic decays can never occur. You might protest that this conclusion has been drawn in the rest frame of the decaying particle and may not be true in all frames. But, the decay of a particle is a frame invariant phenomenon. Hence, even if we sit in a frame where the mother particle has an enormous amount of energy E_1, it will not decay endergonically. The next question to ask would be whether endergonic scatterings occur or not; and if they do, then what conditions have to be met. Let us investigate this in the lab frame where one of the initial particles, say a_2, is at rest.



Since $\vec{p}_2 = 0$, therefore, $\vec{p}_1 = \vec{p}_1^{'} + \vec{p}_2^{'}$, and $\vec{p}_1, \vec{p}_1^{'}, \vec{p}_2^{'}$ lie on the same plane. We choose that plane to be the plane of the paper. The 4-momenta of the particles are

$$\begin{aligned} p_1 &= (E_1, |\vec{p}_1|, 0, 0) \\ p_2 &= (m_2, 0, 0, 0) \\ p_1^{'} &= \left(E_1^{'}, |\vec{p}_1^{'}| \cos\theta_1, |\vec{p}_1^{'}| \sin\theta_1, 0\right) \\ p_2^{'} &= \left(E_2^{'}, |\vec{p}_2^{'}| \cos\theta_2, |\vec{p}_1^{'}| \sin\theta_2, 0\right) \end{aligned} \tag{38}$$

When we put a detector to the right at a fixed angle θ, we can measure the energies of particles getting scattered in that direction. Varying θ then gives us the angular distribution. We will now employ 4-momentum conservation and perform a bit of algebra.

$$p_1 + p_2 = (E_1 + m_2, |\vec{p}_1|, 0, 0)$$

$$\Rightarrow p_1 + p_2 - p_1^{'} = \left(E_1 + m_2 - E_1^{'}, |\vec{p}_1| - |\vec{p}_1^{'}| \cos\theta_1, -|\vec{p}_1^{'}| \sin\theta_1, 0\right)$$

$$\Rightarrow m_2^{'2} = p_2^{'2} = \left(p_1 + p_2 - p_1^{'}\right)^2 = \left(E_1 + m_2 - E_1^{'}, |\vec{p}_1| - |\vec{p}_1^{'}| \cos\theta_1, -|\vec{p}_1^{'}| \sin\theta_1, 0\right)^2$$

$$\Rightarrow m_2^{'2} = (E_1 + m_2)^2 - |\vec{p}_1|^2 + m_1^{'2} - 2\,(E_1 + m_2)\,E_1^{'} - 2|\vec{p}_1||\vec{p}_1^{'}| \cos\theta_1 \tag{39}$$

From this, one can determine the energy $E_1^{'}$ of $a_1^{'}$ as a function of the scattering angle θ_1– the angular distribution of energy. The masses of the particles and the energy E_1 of the incoming particle a_1 will appear as parameters in the expressions for $E_{1,2}^{'}$. Since $E_1^{'}$ has to be positive, that leads to a constraint in the parameters. This constraint implies that there is a threshold energy E_1 of the incoming beam of particles below which endergonic scattering will not occur. Details of that calculation is straightforward and you can do it as an exercise. So, we see that endergonic scatterings are possible, although endergonic decays are not.

Symmetry and Groups :

Group theory is the mathematical structure of symmetries. Symmetries play an all-pervading role in physics. Particle physics is one of the areas where symmetry becomes the most important guiding star. We will assume that the very basics of group theory are known to you. Let us summarize a few results before we wrap up today.

- Given a group G and $H \subset G$, H is a subgroup of G (written $H \leq G$) iff $\forall a, b \in H$, $ab^{-1} \in H$.
- $GL\,(3)$ is the set of all 3×3 invertible matrices and forms a group under matrix multiplication.

- $O(3)$ is the set of all 3×3 orthogonal matrices and forms a group under matrix multiplication. $O(3) \leq GL(3)$.

$SO(3)$ is the set of orthogonal 3×3 matrices with determinant 1. This set also forms a group. This is the group of proper rotations, meaning, it doesn't contain inversions and reflections. All elements in $SO(3)$ are continuously connected to the identity matrix. In the next class, we will take up a detailed study of rotations and derive results which apply generally to a lot of other groups of our interest.

2 Week 2

2.1 Lecture 4 : January 11, 2016

We started talking about symmetries and groups. Let us consider rotations in 3-d space. Rotation about the z-axis by an angle ϕ is given by the matrix

$$R_z(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \simeq \begin{bmatrix} 1 & -\phi & 0 \\ \phi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \equiv \mathbb{I} - i\phi T_z \tag{40}$$

where the approximation holds for **small** ϕ. The above equation defines the matrix T_z. It is defined with a factor of i to ensure that T_z is Hermitian. Mathematicians do not mind working with anti-Hermitian matrices and they will define it without the factor of i. Similarly we can define the matrices $T_{x,y}$. Let us list these matrices.

$$T_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \; T_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \; T_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{41}$$

We can now compute the matrix commutators of these three matrices. It yields,

$$[T_x, T_y] = iT_z \;\; \text{(and cyclic permutations)} \tag{42}$$

This is the angular momentum algebra. We notice that the matrices $T_{x,y,z}$, called **generators** of rotation, give infinitesimal rotation matrices. We would like to know if the matrices corresponding to finite (meaning not infinitesimal) rotations can also be expressed in terms of these generators. Let ϕ be finite. We can achieve a rotation about the z-axis by an angle ϕ by breaking it into N smaller rotations, each by an angle $\frac{\phi}{N}$ about the z-axis. Then we take the limit $N \to \infty$ such that each step becomes smaller and smaller and we can use the expression for infinitesimal rotation given in equation (40).

$$R_z(\phi) = \lim_{N\to\infty} \left(\mathbb{I} - i\frac{\phi}{N}T_z\right)^N = \exp(-i\phi T_z) \tag{43}$$

A general rotation $\vec{\phi}$ $\left(\text{angle} = \left|\vec{\phi}\right|, \text{ axis of rotation}=\widehat{\phi}\right)$ will be given by $\exp(-i\phi_j T_j)$. Recall that, in order to exponentiate matrices, you should diagonalize it and then proceed.

We have already seen explicitly that the commutators of two of the generators yields a linear combination of the generators [equation (41)]. We can also see that from the general theory of Lie groups and Lie algebras[8]. Closure property of the rotation group implies that multiplying two rotations, $\exp(-i\alpha.T)$ and $\exp(-i\beta.T)$, gives another rotation, say, $\exp(-i\gamma.T)$.

$$\exp(-i\alpha.T)\,.\,\exp(-i\beta.T) = \exp(-i\gamma.T) \tag{44}$$

[8] We shall learn what these things are in course of time.

Expanding up to first order,

$$\text{LHS} = (\mathbb{I} - i(\alpha.T))\,(\mathbb{I} - i(\beta.T)) \simeq \mathbb{I} - i\,(\alpha.T + \beta.T)$$

This expansion holds when both α and β are small, in which case the resultant rotation will also be small. Therefore, the RHS can also be expanded to first order.

$$\text{RHS} = \mathbb{I} - i\,(\gamma.T)$$

Comparing the two sides, $\alpha+\beta = \gamma$. This also tells us that order of two infinitesimal rotations do not matter, since $\alpha+\beta = \beta+\alpha$. $\Rightarrow$ Infinitesimal rotations commute.For large rotations, we need more terms in the expansions. Expanding the LHS up to second order gives

$$\text{LHS} = \exp(-i\alpha.T)\,.\exp(-i\beta.T) \simeq \mathbb{I} - i\,(\alpha.T + \beta.T) - \frac{1}{2}\,(\alpha.T + \beta.T)^2 - \frac{1}{2}\,[\alpha.T,\,\beta.T] \tag{45}$$

$$\text{RHS} \simeq \mathbb{I} - i\,(\gamma.T) - \frac{1}{2}\,(\gamma.T)^2 \tag{46}$$

Comparing the two equations above order by order, it is evident that the commutator of any two generators yields a linear combination of the generators themselves. This approach, where we start from a Lie group and derive the Lie algebra underlying it, we can see that the algebra is closed automatically[9].

Digression :

Definition of an algebra over a field : Let F be a field and V be a vector space over F. Let V be equipped with a binary operation $[\,,\,] : V \times V \to V$. Then V is an algebra over the field F if the following hold

1. $\forall x, y, z \in V$, $[x + y, z] = [x, z] + [y, z]$ and $[x, y + z] = [x, y] + [x, z]$
2. $\forall x, y \in V$, and $\forall a, b \in F$, $[ax, by] = ab\,[x, y]$

Definition of Lie Algebra : A Lie algebra is an algebra V with the binary operation $[\,,\,] : V \times V \to V$ satisfying the following properties :

1. $\forall x, y \in V$, $[x, y] = -\,[y, x]$ (antisymmetry)
2. $\forall x_1, x_2, y \in V$, $[x_1 + x_2, y] = [x_1, y] + [x_2, y]$ (linearity in prefactor). [This and antisymmetry implies linearity in the post-factor.]
3. $\forall x, y \in V$, $\forall \alpha \in F$, $[\alpha x, y] = \alpha\,[x, y]$ (distributive scalar multiplication)
4. $\forall x, y, z \in V$, $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ (Jacobi identity)

[9] Read this statement once again after we have discussed the definitions of Lie groups and Lie algebras.

This binary operation [,] is therefore a bilinear product which is called the **Lie product**. Vectors forming a basis of the vector space V are called **generators** of the Lie algebra.

The span of the matrices $T_{x,y,z}$ form a vector space. We notice the fact that this vector space along with the binary operation of matrix commutation defined on it, forms a Lie algebra! This is a concrete example of the abstract definition of the Lie algebra we have given above. In this concrete example, the abstract Lie product is commutators of matrices. Associativity of matrix multiplication ensures that the Jacobi identity is satisfied in this algebra. However, an abstract Lie algebra is given by defining an abstract Lie product on a vector space. There is no reason to suspect that this Lie product can always be written as a "commutator". In order to be able to do that, one has to have another product, say $*$, defined on the space such that $[A, B] = A * B - B * A$ makes sense. On top of that, this product has to be associative so that the Jacobi identity is satisfied, and distributive over addition so that bilinearity of [,] is satisfied. In other words, the vector space along with the operation $*$ has to form an associative algebra. In the matrix Lie algebras, matrix multiplication has all these properties and hence induces a Lie algebra on the space.

So, in our case, $V = \text{span}\{T_x, T_y, T_z\}$ and a general matrix belonging to V has the form $\alpha_x T_x + \alpha_y T_y + \alpha_z T_z$. Given a basis (set of generators) $\{T_a\}$ of the Lie algebra, the Lie products of the basis vectors carry all the information abut the Lie product of two arbitrary vectors (because of bilinearity of Lie products). Therefore, it is sufficient to provide the Lie products

$$[T_a, T_b] = f_{ab}^c T_c \tag{47}$$

The numbers $f_{ab}^c$are called the **structure constants** of the Lie algebra. Structure constants are antisymmetric in the two lower indices. Evidently, they depend on the choice of basis (generators) of the Lie algebra. Equation (42) gives us the structure constants of the angular momentum Lie algebra for the basis $\{T_x, T_y, T_z\}$. Let us illustrate by one example how structure constants depend on basis.

Consider a Lie algebra generated by t_x, t_y, t_z, with the non-zero structure constants $f_{xy}^z = f_{yz}^x = f_{zx}^y = i$. That is, $[t_x, t_y] = it_z$ and cyclic permutations. Let's choose a different basis $\{t_z, t_+, t_-\}$ where

$$t_\pm = t_x \pm it_y \tag{48}$$

This yields,

$$[t_z, t_\pm] = \pm t_\pm, \quad [t_+, t_-] = 2t_z \tag{49}$$

That is, the non-zero structure constants in this basis are $f_{z\pm}^\pm = \pm 1$, $f_{+-}^z = 2$.

Representation :

Suppose we have an abstract Lie algebra $\mathcal{A}$ generated by t_x, t_y, t_z. The idea of representation is to look for three matrices, T_x, T_y, T_z, such that the commutation relations of these matrices mimic the Lie product relations of the

generators t_x, t_y, t_z. To be precise, the mapping $t_i \to T_i$, $i = z, +, -$, is called a representation of the Lie algebra if

$$[t_z, t_\pm] = \pm t_\pm, [t_+, t_-] = 2t_z \;\Rightarrow [T_z, T_\pm] = \pm T_\pm, [T_+, T_-] = 2T_z \tag{50}$$

For one last time, let me mention that the square brackets on the LHS of the above implication stand for abstract Lie products, while those on the RHS stand for matrix commutators, a concrete realization of the Lie product. The above mapping establishes a one-to-one correspondence between the abstract Lie algebra $\mathcal{A}$ and the matrix Lie algebra $\mathcal{M}$ generated by T_x, T_y, T_z.

Now we shall work out the entries of the matrices T_+, T_-, T_z in the eigenbasis of T_z. Suppose that these are $N \times N$ matrices and they act as linear operators on the vector space of $N \times 1$ column vectors. Let v_k be the eigenvector of T_z corresponding to the eigenvalue k (which can be any complex number).

$$T_z v_k = k v_k \tag{51}$$

Now, $[T_z, T_\pm] = \pm T_\pm$ implies

$$T_z \left(T_\pm v_k\right) = \left(T_\pm T_z \pm T_\pm\right) v_k = \left(k \pm 1\right) \left(T_\pm v_k\right) \tag{52}$$

This means that $(T_\pm v_k)$ is either an eigenstate of T_z with eigenvalue $(k \pm 1)$ or the zero vector. Remember, $(k \pm 1)$ are complex numbers, as of now, and are different from k. Eigenvectors corresponding to different eigenvalues are linearly independent. If we keep on applying T_+on v_k, we shall keep on getting more and more linearly independent vectors unless for some complex number $j = k + r$ (r is a non-negative integer)

$$T_+ v_j = 0 \tag{53}$$

Since the space on which the generators act is finite dimensional, such a j is guaranteed to exist. Similar arguments tell us that there must exist a complex number $q = k - s$ (s is a non-negative integer) so that

$$T_- v_q = 0 \tag{54}$$

These numbers j and q are not completely independent and arbitrary. One thing is certain :

$$j - q \in \{\text{non-negative integers}\} \tag{55}$$

We shall later prove that $q = -j$. Sometimes we shall use the notations

$$\begin{aligned} v_j &\to |j, j\rangle \\ v_k &\to |j, k\rangle \end{aligned}$$

We do not need q to label the states because of the result $q = -j$ which we shall derive shortly.

We have seen that $T_\pm v_k$ are eigenstates of T_z with eigenvalues $(k \pm 1)$. Hence, $T_\pm v_k = \# v_{k\pm1}$ where $\#$ is a number that fixes the normalization. Our convention of normalization will be the following.

$$\begin{aligned} T_- v_k &= v_{k-1} \\ T_+ v_k &= r_k v_{k+1} \end{aligned} \tag{56}$$

The first one of the equations above is the choice of normalization. In this normalization, r_k are some numbers yet to be determined. We shall do that in the next class. Before concluding today, let us learn a nice fact. In the matrix Lie algebra that we have been working with today, define

$$T^2 = T_x^2 + T_y^2 + T_z^2 \tag{57}$$

and notice

$$\left[T^2, T_{x,y,z}\right] = 0 \tag{58}$$

That is, a particular combination of second order operators (second order in the generators) gives zero Lie product (or "commutes") with all the elements of the Lie algebra. For another Lie algebra, this same combination might not have the same property. However, **Casimir's theorem** ensures that for any Lie algebra we can always find a second order operator which commutes with all the elements of the Lie Algebra.

Another fact : $T_{x,y,z}$ have been so defined that they are Hermitian. This implies that $T_{x,y,z}^2$ are positive definite. This implies that

$$\text{eigenvalues of } T^2 \geq \text{eigenvalues of } T_{x,y,z}^2$$

2.2 Lecture 5 : January 13, 2016

Equation (55) from the last class implies that we must be able to go from v_j to v_q by a finite number of applications of T_-. Today we shall learn how many times we have to use T_- to get there. Start by deriving a formula for r_k.

$$\begin{aligned} T_+ v_k = r_k v_{k+1} = T_+ \left(T_- v_{k+1}\right) = \left(T_- T_+ + 2T_z\right) v_{k+1} \\ = \left(r_{k+1} + 2\left(k+1\right)\right) v_{k+1} \end{aligned} \tag{59}$$

We get a recursion relation

$$r_{k+1} - r_k = -2\left(k+1\right) \tag{60}$$

A boundary condition for the recursion relation would be $r_j = 0$, since $T_+ v_j = 0$. Therefore,

$$\begin{aligned} r_j - r_k = & \sum_{i=k}^{j-1} \left(r_{i+1} - r_i\right) = -2 \sum_{i=k}^{j-1} \left(i+1\right) = -2 \sum_{i=j}^{j} i \\ & = -j\left(j+1\right) + k\left(k+1\right) \end{aligned} \tag{61}$$

Therefore,

$$r_k = j\left(j+1\right) - k\left(k+1\right) \tag{62}$$

This is consistent with $r_j = 0$, of course. Now we shall derive the relation between j and q.

$$T_+ T_- v_q = 0 \Rightarrow (T_- T_+ + 2T_z)\, v_q = 0$$

$$\Rightarrow (r_q + 2q)\, v_q = 0 \Rightarrow (r_q + 2q) = 0$$

Entering the expression for r_q,

$$q\,(q-1) = j\,(j+1) \tag{63}$$

This is a quadratic equation in q and the two roots are

$$q = -j, \text{ or } q = j+1 \tag{64}$$

The second root is not possible since we demanded that v_j would be the highest rung in the ladder of states obtained by applying T_+on v_k. Hence, we get the sought after relation

$$q = -j \tag{65}$$

Equation (55) now implies that $2j$ must be a non-negative integer, and hence

$$j \text{ must be an integer or a half odd integer} \tag{66}$$

Eigenvalues of T_z are

$$-j,\, (-j+1)\,,\, \ldots\, (j-1)\,,\, j \;\Rightarrow \#\,\text{distinct eigenvalues } = (2j+1) \tag{67}$$

Therefore, the matrices $\{T_z, T_\pm\}$ give a $(2j+1)$-dimensional representation. The vector space is

$$V = \text{span}\,\{v_j, v_{j-1}, \ldots v_{-j}\} \tag{68}$$

In this **ordered basis**, the matrices look like the following.

$$T_z = \begin{bmatrix} j & 0 & \cdots & 0 \\ 0 & j-1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -j \end{bmatrix}, T_+ = \begin{bmatrix} 0 & r_{j-1} & 0 & \cdots & 0 \\ 0 & 0 & r_{j-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & r_{-j} \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}, T_- = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{69}$$

Recall that the matrix representation D of an operator T , in an ordered basis $\{e_i\}$is obtained from the following :

$$Te_i = D_{ji} e_j \tag{70}$$

That is, one has to read off the coefficients and transpose them. This convention is adopted because, as a consequence, the matrix representation of the composite operator $T_1 T_2$ would be $D_1 D_2$, where $D_{1,2}$ are the matrix representations of $T_{1,2}$.

Note that the matrices T_x, T_y, obtained from $T_\pm$, will not always be Hermitian. The reason for that can be traced back to our knowledge of linear algebra/quantum mechanics. The matrix representation of a Hermitian operator does not necessarily have to be a Hermitian matrix. When the ordered basis being used is orthonormal, Hermitian operators do give Hermitian matrices. In the normalization convention we are using, the basis vectors are orthogonal alright (because they correspond to distinct eigenvalues of T_z), but they are not normalized.

Irreducible representations :

Consider the subspace span $\{v_j, v_{j-1}\}$. T_z and T_+ act on an arbitrary element of this subspace to produce another element in this subspace. However, T_- may act on an element in this subspace and produce a result outside of this subspace. Similarly, no proper subspace of V is **invariant** under the action of all the three generators in the above sense. Since no such **invariant subspace** exists in V, therefore this representation is called an **irreducible representation** (or **irrep** in short).

Equivalent representations :

Given T_i's, define $T_i^{'} = ST_iS^{-1}$ for an arbitrary invertible matrix S (with proper dimensions). $T_i^{'}$'s also form a representation of the algebra. However, these two representations are related by a similarity transformation and are called **equivalent representations**. Quite naturally, we shall be interested in inequivalent representations.

We conclude today by finding out the eigenvalues of the quadratic operator T^2.

$$T^2 = T_x^2 + T_y^2 + T_z^2 = T_z^2 + \frac{1}{2}\left(T_+T_- + T_-T_+\right)$$

$$\Rightarrow T^2 v_k = \left(k^2 + \frac{1}{2}\left(r_{k-1} + r_k\right)\right) v_k = j\left(j+1\right) v_k \tag{71}$$

So, all the eigenstates v_k of T_z are simultaneously eigenstates of T^2 with a common eigenvalue $j\left(j+1\right)$. That is, T^2 has degeneracy in its spectrum! We started by analyzing symmetries and it led us to degeneracy of an observable. This is a common feature of systems with symmetry.

2.3 Lecture 6 : January 14, 2016

In the last class, we saw that the generators of rotation form the Lie algebra $su\,(2)$. You might think that the algebra should be called $so\,(3)$ since it underlies the Lie group $SO\,(3)$. The fact is that these two Lie algebra are one and the same. That is, the groups $SU\,(2)$ and $SO\,(3)$ have the same algebra. The reason for this will be discussed later. As a teaser, let's just make a fancy statement which we will understand properly in due time – $SU\,(2)$ is the universal (double) cover of $SO\,(3)$.

So, we saw that an abstract Lie algebra $su\,(2)$, generated by three generators, say, t_z, $t_\pm$, with the Lie products $[t_z, t_\pm] = \pm t_\pm$, $[t_+, t_-] = 2t_z$, may be put in a one-to-one correspondence with a set of matrices T_z,$T_\pm$ if these matrices have commutation relations $[T_z, T_\pm] = \pm T_\pm$, $[T_+, T_-] = 2T_z$. This would be a concrete realization of the abstract algebra. We call this a matrix (or linear) representation of the Lie algebra. We also saw that the representations of this algebra are $(2j+1)$ dimensional where j is of the form $\frac{n}{2}$, $n \in \mathbb{N}$.

The $j = 0$ representation :

Mathematicians would call it the 1-representation because of its dimension $2j + 1 = 1$. This is a trivial representation where the vector space is one dimensional, hence the representative matrices are just numbers. Each of the generators $t_{z,\pm}$ will have to be mapped to the number 0 since commutator of two numbers is always 0.

The $j = \frac{1}{2}$ representation :

Mathematicians call it the 2-representation. The vector space W is 2-dimensional. Let's choose two basis vectors and call them $w_{1/2}$ and $w_{-1/2}$. $W = \text{span}\left\{w_{1/2}, w_{-1/2}\right\}$.

$$T_z : w_{1/2} \to \frac{1}{2} w_{1/2},\ w_{-1/2} \to -\frac{1}{2} w_{-1/2}$$

$$T_+ : w_{1/2} \to 0,\ w_{-1/2} \to w_{1/2}$$

$$T_- : w_{1/2} \to w_{-1/2},\ w_{-1/2} \to 0$$

We already stated our normalization convention. This is a little different than what is mostly used in quantum mechanics texts. With the above results, we have

$$T_z = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix},\ T_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ T_- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{72}$$

$$T_x = \frac{1}{2}\left(T_+ + T_-\right) = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ T_y = \frac{1}{2i}\left(T_+ - T_-\right) = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{73}$$

These are nothing but the Pauli matrices with a factor of $\frac{1}{2}$ that arises due to the difference in normalization. However, the T matrices are Hermitian in the $j = \frac{1}{2}$ representation. We shall see that this will no longer hold for $j = 1$ and above.

The matrices $T_{x,y,z}$ are traceless and Hermitian. They form a basis for the real vector space of 2×2 traceless Hermitian matrices. The complex vector space spanned by these three matrices will contain traceless matrices but an arbitrary matrix in that space won't be Hermitian (because of complex coefficients). So, with real coefficients a_1, a_2, a_3, an arbitrary 2×2 traceless Hermitian matrix will take the form $a_1 T_x + a_2 T_y + a_3 T_z$. This is an element of the algebra. Let's exponentiate it. Exponential of a Hermitian matrix gives a unitary matrix. Exponential of a traceless matrix gives a unideterminantal matrix[10]. Therefore, the exponentiation of the elements of the algebra of the $j = \frac{1}{2}$ representation of $su\,(2)$ gives rise to 2×2 unitary unideterminantal matrices $\in SU\,(2)$. This is the reason why we have been calling this Lie algebra $su\,(2)$ all along. The above discussion is one way to see how generators of $SO\,(3)$ rotations also generate $SU\,(2)$ matrices.

Before moving on to discuss the $j = 1$ representation, let's digress a bit and talk about $SU\,(3)$ and $su\,(3)$.

Let $\exp(iA) \in SU\,(3)$. $\Rightarrow A$ is a 3×3 traceless Hermitian matrix having the most general form[11]

$$A = \begin{bmatrix} a_3 + \frac{1}{\sqrt{3}} a_8 & a_1 - i a_2 & a_4 - i a_5 \\ a_1 + i a_2 & -a_3 + \frac{1}{\sqrt{3}} a_8 & a_6 - i a_7 \\ a_4 + i a_5 & a_6 + i a_7 & -\frac{2}{\sqrt{3}} a_8 \end{bmatrix} \equiv \sum_{i=1}^{8} (a_i \lambda_i) \tag{74}$$

where a_i's are real and the matrices λ_i are called the Gell-Mann matrices.

$$\begin{aligned} \lambda_1 &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \lambda_2 &= \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \lambda_3 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \lambda_4 &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} & \lambda_5 &= \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} & \lambda_6 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ \lambda_7 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} & \lambda_8 &= \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & -\frac{2}{\sqrt{3}} \end{bmatrix} \end{aligned} \tag{75}$$

Notice that the first three matrices are exactly the $su\,(2)$ generators except for the third row and column of zeroes augmented to them. We can also explicitly calculate the commutators of these matrices to find that the commutation relations are closed. However, we can see that on two different counts. Firstly, the general theory of groups tells us that combining two elements of the Lie group $SU\,(3)$ will produce another element in $SU\,(3)$. Closure of the commutation relations of the algebra will follow from the closure of the group. Secondly, commutator of two traceless Hermitian matrices is a traceless anti-Hermitian matrix (commutators of matrices are always traceless). Therefore, in $[\lambda_1, \lambda_2] = i\#$, the

[10] $\mathrm{Tr}\,(\log A) = \log(\det A)$

[11] The parametrization used here of the entries in A is purely conventional.

has to be a traceless Hermitian matrix, and hence a (real) linear combination of the λ matrices.

Also note that λ_3 and λ_8 commute. The subalgebra spanned by λ_3 and λ_8 is Abelian. This is called the **Cartan subalgebra** of the Lie algebra[12]. One can find essentially all information about a Lie algebra from its Cartan subalgebra.

The $j = 1$ representation :

This is a 3-dimensional representation. The eigenvalues of T_z are $1, 0, -1$ and let the corresponding eigenvectors be v_1, v_0, v_{-1}.

$$T_z : v_1 \to v_1,\ v_0 \to 0,\ v_{-1} \to -v_{-1}$$

$$T_+ : v_1 \to 0,\ v_0 \to 2v_1,\ v_{-1} \to 2v_0$$

$$T_- : v_1 \to v_0,\ v_0 \to v_{-1},\ v_{-1} \to 0$$

We can read off the matrix elements :

$$T_z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, T_+ = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, T_- = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{76}$$

$$\Rightarrow T_x = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \end{bmatrix}, T_y = \begin{bmatrix} 0 & -i & 0 \\ \frac{i}{2} & 0 & -i \\ 0 & \frac{i}{2} & 0 \end{bmatrix} \tag{77}$$

This is what we were talking about earlier – the matrices $T_{x,y}$ are not Hermitian in this representation. However, the operators $t_{x,y}$ are Hermitian and therefore the linear operators $T_{x,y}$ represent are Hermitian.

An important observation is that all these representations characterized by half-integral j-values are irreducible. Irreducible representations are important because all representations can be broken down in their irreducible components. Let's see what this breaking down means.

The direct sum representation :

Let $V_{1,2}$ be two vector spaces over the same field F. Let $\left\{e_1^1, e_2^1, \ldots, e_{n_1}^1\right\}, \left\{e_1^2, e_2^2, \ldots, e_{n_2}^2\right\}$ be two ordered bases of the two spaces respectively. The direct sum space $V \equiv V_1 \oplus V_2$ is another vector space (over F) of dimension $(n_1 + n_2)$with the natural ordered basis : $\left\{e_1^1, e_2^1, \ldots, e_{n_1}^1, e_1^2, e_2^2, \ldots, e_{n_2}^2\right\}$. An arbitrary vector in the direct sum space is denoted by $v_1 \oplus v_2$. Suppose we have two representations $\rho_{1,2}$ of the group G on the two vector spaces $V_{1,2}$ respectively. The direct sum representation $\rho_1 \oplus \rho_2$ is defined to be a representation on the space $V_1 \oplus V_2$ given by

$$(\rho_1 \oplus \rho_2)(g) : V_1 \oplus V_2 \to V_1 \oplus V_2 \text{ such that } (\rho_1 \oplus \rho_2)(g)(v_1 \oplus v_2) = \rho_1(v_1) \oplus \rho_2(v_2) \tag{78}$$

[12] Proper definition will follow soon.

In the natural ordered basis, the matrix representations of the direct sum operators are going to be block diagonal.

$$(\rho_1 \oplus \rho_2)(g)\left(e_1^1\right) = e_1^1 D_{11}^1(g) + e_2^1 D_{21}^1(g) + \ldots + e_{n_1}^1 D_{n_1 1}^1(g) + e_1^2.0 + \ldots + e_{n_2}^2.0 \tag{79}$$

and so on. This gives the following block diagonal matrix for the operator $(\rho_1 \oplus \rho_2)(g)$:

$$(\rho_1 \oplus \rho_2)(g) = \left[\begin{array}{c|c} D^1(g)_{n_1 \times n_1} & 0_{n_1 \times n_2} \\ \hline 0_{n_2 \times n_1} & D^2(g)_{n_2 \times n_2} \end{array}\right] \tag{80}$$

The nice block diagonal form arises because we have used the natural ordered basis in the direct sum space. If one were to change basis in $V_1 \oplus V_2$, one would not get a block diagonal form in general. If, for a representation, **all the matrices** corresponding to the group elements can be brought to **the same block diagonal form at once** (by choosing one particular basis), then it is evident that the representation is reducible[13]. The invariant subspaces would correspond to the blocks. For example, if there is a 2×2 square block starting from the 5th diagonal element of each of the group representative matrices, then the subspace spanned by $\{e_5, e_6\}$ will be an invariant subspace.

Before ending today, let's briefly state the motivation for studying group representations in Physics. It should be common knowledge that symmetry transformations of a physical system form a group (eg., Lorentz transformations, Poincare transformations, rotations etc.). Let's take the Lorentz group as an example. In 4-d space, choose a set of Lorentz coordinates, and consider another Lorentz frame whose coordinates are obtained by Lorentz-transforming your frame. As a result, numerical values of many physical objects differ in the two frames. Rules by which components of 4-vectors such as 4-velocity etc. change are easy to deduce – they change in the same way as the 4-vector (t, x, y, z). However, there are physical objects such as F, the electromagnetic field tensor, whose components $F_{\mu\nu}$ change in a systematic way which is different from how 4-vectors change. The components of this tensor can be put in the form of an antisymmetric 4×4 matrix. The space of such matrices is 6-dimensional. The way the components $F_{\mu\nu}$ change due to a Lorentz transformation must be a linear transformation on this space of 4×4 antisymmetric matrices. These transformations will naturally have to form a representation of the Lorentz group on the said space. Therefore, the various representations of symmetry groups may essentially be implementations of the symmetry transformations on the spaces of physical objects. All physical objects will have to "belong to" one of the representations of the total symmetry group of the system.

[13] Necessary and sufficient conditions for reducibility of representations exist. Teaching those will require half a semester worth of time, so we won't do that now.

3 Week 3

3.1 Lecture 7 : January 18, 2016

A representation of a group or an algebra is called reducible if the vector spce on which the representation is defined has one or more non-trivial invariant subspace(s). The null space and the total vector space which are trivial invariant subspaces. In a reducible representation, all the matrices may be brought to a block diagonal form by choosing a particular basis. However, changing the basis (or even the order of the basis) is likely to destroy the block diagonal form.

For the $su\left(2\right)$ Lie algebra, we can find irreps of every dimension n by setting $j=\frac{n-1}{2}$, so that $\left(2j+1\right)=n$.

Today we will talk about the another way of combining two representations to give a higher dimensional representation – the direct product representation. As a motivation for this exercise, let's start with the following question. In quantum mechanics, angular momentum $\left(\vec{L}\right)$ and spin $\left(\vec{S}\right)$ operators both satisfy the same algebra, namely $su\left(2\right)$. They are different representations of $su\left(2\right)$ acting on different spaces. However, a particle may have both angular momentum and spin, in which case we defined the total angular momentum as $\vec{J}=\vec{L}+\vec{S}$. What do we mean by this?

Direct product representation :

Let $V_{1,2}$ be two vector spaces over a field F, having respective ordered bases $\left\{e_1^1,e_2^1,\ldots,e_{n_1}^1\right\}$ and $\left\{e_1^2,e_2^2,\ldots,e_{n_2}^2\right\}$. The direct product space $V_1\otimes V_2$ is defined to be a vector space over F of dimension n_1n_2 with a natural ordered basis $\left\{e_i^1\otimes e_j^2\right\}_{(i,j)=(1,1)}^{(n_1,n_2)}$. An arbitrary vector in the direct product space is $v_1\otimes v_2$ where $v_{1,2}$ are arbitrary vectors in $V_{1,2}$ respectively. For example,

$$\left(2e_1^1+3e_2^1\right)\otimes\left(-1e_1^2+3e_7^2\right)=-2e_1^1\otimes e_1^2+6e_1^1\otimes e_7^2-3e_2^1\otimes e_1^2+9e_2^1\otimes e_7^2$$

Now, given two smaller representations $\rho_{1,2}$ of a group G living on $V_{1,2}$, we can define a bigger representation $\rho_1\otimes\rho_2$, called the direct product representation in the following way :

$$\left(\rho_1\otimes\rho_2\right)\left(g\right)\left(v_1\otimes v_2\right)\equiv\left(\rho_1\left(g\right)\left(v_1\right)\right)\otimes\left(\rho_2\left(g\right)\left(v_2\right)\right)\tag{81}$$

Let us now consider the rotation group. An arbitrary element of the group will be rotation about the unit vector $\hat{n}$ through an angle ϕ. If we have two representations $R_{\hat{n}}^1\left(\phi\right)$, $R_{\hat{n}}^2\left(\phi\right)$ of the rotation group, then the product representation $R_{\hat{n}}\left(\phi\right)$ acts as follows

$R_{\hat{n}}\left(\phi\right)\left(v_1\otimes v_2\right)=\left(R_{\hat{n}}^1\left(\phi\right)\left(v_1\right)\right)\otimes\left(R_{\hat{n}}^2\left(\phi\right)\left(v_2\right)\right)$, which, for small angle ϕ becomes

$$\begin{aligned}&\simeq\left(\mathbb{I}-i\phi\hat{n}.\vec{T}^1\right)v_1\otimes\left(\mathbb{I}-i\phi\hat{n}.\vec{T}^2\right)v_2\\&=v_1\otimes v_2-i\phi\hat{n}.\left[\left(\vec{T}^1v_1\right)\otimes v_2+v_1\otimes\left(\vec{T}^2v_2\right)\right]\end{aligned}$$

$$= \left(\mathbb{I} - i\phi\hat{n}.\left[\vec{T}^1 \otimes \mathbb{I} + \mathbb{I} \otimes \vec{T}^2\right]\right)(v_1 \otimes v_2) \tag{82}$$

where the identities in the different spaces are all written by the same symbol $\mathbb{I}$ with the assumption that you will figure out which identity belongs to which space from the context.

Let's answer the question we used as a motivation to study the direct product representations. Angular momentum and spin – these two sets of operators act on two different Hilbert spaces, say H_1 and H_2. States of a particle with both angular momentum and spin will be represented by vectors in the direct product space $H_1 \otimes H_2$. In this space, the angular momentum operator is really $\overrightarrow{L} \otimes \mathbb{I}$ and the spin operator is $\mathbb{I} \otimes \overrightarrow{S}$. Therefore, what we really mean by $\vec{J} = \vec{L} + \vec{S}$ is

$$\vec{J} = \vec{L} \otimes \mathbb{I} + \mathbb{I} \otimes \vec{S} \tag{83}$$

As an exercise, we will build a direct product representation of the $su\,(2)$ Lie algebra from the $j = \frac{1}{2}$ and the $j = 1$ representations. The product representation is going to be 6-dimensional. Let the ordered bases of the relevant vector spaces be $\{v_1, v_0, v_{-1}\}$,$\{w_{1/2}, w_{-1/2}\}$ and

$\{v_1 \otimes w_{1/2}, v_0 \otimes w_{1/2}, v_1 \otimes w_{-1/2}, v_{-1} \otimes w_{1/2}, v_0 \otimes w_{-1/2}, v_{-1} \otimes w_{-1/2}\}$.

We have

$$T_z :\, v_1 \otimes w_{1/2} \to 1.v_1 \otimes w_{1/2} + v_1 \otimes \frac{1}{2}w_{1/2} = \frac{3}{2}\left(v_1 \otimes w_{1/2}\right) \tag{84}$$

Thus, the chosen basis vectors of the product space are all eigenvectors of T_z with respective eigenvalues $\frac{3}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$ - just what we get in addition of angular momenta. Matrix representation of T_z in the product space in the chosen basis is

$$T_z = \begin{bmatrix} \frac{3}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} \end{bmatrix} \tag{85}$$

However, $T_\pm$ will not be diagonal in this basis, quite evidently. You can work out the matrix forms of these operators as an exercise.

Question : Is this product representation an irrep?

Had this been an irrep, its dimension suggests that it would correspond to $j = \frac{n-1}{2} = \frac{5}{2}$. However, the highest eigenvalue of T_z obtained above is $\frac{3}{2}$. Hence, the product representation is reducible. In fact, it is isomorphic to the direct sum representation of the two irreps $j = \frac{3}{2}$ (dimension 4) and $j = \frac{1}{2}$ (dimension 2). Symbolically, we write the following :

$$3 \otimes 2 = 4 \oplus 2 \tag{86}$$

Let us explicitly find out the invariant subspaces. Define the following vectors :

$$\begin{array}{cc} a_{3/2} \equiv v_1 \otimes w_{1/2}; & a_{1/2} \equiv T_- a_{3/2} = v_0 \otimes w_{1/2} + v_1 \otimes w_{-1/2} \\ a_{-1/2} \equiv T_- a_{1/2} = v_{-1} \otimes w_{1/2} + 2v_0 \otimes w_{-1/2;} & a_{-3/2} \equiv T_- a_{-1/2} = 3v_{-1} \otimes w_{-1/2} \end{array} \tag{87}$$

These four vectors are eigenvectors of T_z with respective eigenvalues $\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$. Also,$T_\pm$ acting on a linear combination of $a_{3/2,1/2,-1/2,-3/2}$ gives rise to another linear combination of the same 4-vectors. Therefore, span $\left\{a_{3/2}, a_{1/2}, a_{-1/2}, a_{-3/2}\right\}$ is an invariant subspace (4-dimensional) under the actions of $T_z, T_\pm$. We shall now look for another invariant subspace, 2-dimensional, whose basis vectors will be linearly independent of the basis vectors of the invariant subspace already obtained above. My claim is that there exists another linear combination of $v_0 \otimes w_{1/2}$ and $v_1 \otimes w_{-1/2}$ which is linearly independent of $a_{1/2}$ and is the highest state of a 2-dimensional invariant subspace. Let the linear combination be $\alpha v_0 \otimes w_{1/2} + \beta v_1 \otimes w_{-1/2}$. We note that

$$T_z\left(\alpha v_0 \otimes w_{1/2} + \beta v_1 \otimes w_{-1/2}\right) = \frac{1}{2}\left(\alpha v_0 \otimes w_{1/2} + \beta v_1 \otimes w_{-1/2}\right) \tag{88}$$

This implies that, if this state indeed were a highest state, then the corresponding irrep will be 1-dimensional. This being a highest state means

$$\begin{gathered} T_+\left(\alpha v_0 \otimes w_{1/2} + \beta v_1 \otimes w_{-1/2}\right) = 0 \\ \Rightarrow (2\alpha + \beta)\left(\alpha v_0 \otimes w_{1/2} + \beta v_1 \otimes w_{-1/2}\right) = 0 \\ \Rightarrow (2\alpha + \beta) = 0 \end{gathered} \tag{89}$$

Not bothering about the normalization of the state, let's take $\alpha = 1$, $\beta = -2$. Thus, define

$$\begin{gathered} b_{1/2} \equiv v_0 \otimes w_{1/2} - 2v_1 \otimes w_{-1/2} \\ b_{-1/2} \equiv T_- b_{-1/2} = b_{1/2} \equiv v_{-1} \otimes w_{1/2} - v_0 \otimes w_{-1/2} \end{gathered} \tag{90}$$

It is straightforward to see that $b_{-1/2}$ is killed by T_-, and span $\left\{b_{1/2}, b_{-1/2}\right\}$ is a 2-dimensional invariant subspace of the product space. Let's also see explicitly that $b_{1/2}$ and $a_{1/2}$ are linearly independent.

$$\begin{gathered} \alpha a_{1/2} + \beta b_{1/2} = 0 \Rightarrow (\alpha + \beta)\, v_0 \otimes w_{1/2} + (\alpha - 2\beta)\, v_1 \otimes w_{-1/2} = 0 \\ \Rightarrow \alpha = 0 = \beta \quad \square \end{gathered}$$

Similarly, $b_{-1/2}$ and $a_{-1/2}$ are also linearly independent. Therefore the set $\left\{a_{3/2}, a_{1/2}, a_{-1/2}, a_{-3/2}, b_{1/2}, b_{-1/2}\right\}$ contains 6 linearly independent vectors and hence is a basis of the product space. In this basis, the reducibility of the product space is obvious.

Upshot : by taking a direct product of a 3-dimensional irrep and a 2-dimensional irrep of $su\,(2)$, we get a reducible product representation made up of two irrep components – one 4-dimensional and another 2-dimensional. This is exactly what we did while studying addition of angular momentum in quantum mechanics. The process of breaking down the product space into its irrep-components is termed Clebsch-Gordan decomposition.

When in quantum mechanics we add two angular momenta, say, $j_1 = 2$ (5-dimensional irrep of $su(2)$) and $j_2 = 1$ (3-dimensional irrep of $su(2)$), we find that it results in the following set of angular momenta :

$$j = (j_1 + j_2), (j_1 + j_2 - 1), \ldots, (|j_1 - j_2| + 1), (|j_1 - j_2) \rightarrow 3, 2, 1$$

Writing the dimensions of the irreps, we get

$$5 \otimes 3 = 7 \oplus 5 \oplus 3$$

You can count the number of basis states on both sides and see for yourself that they match.

Teaser for the next few lectures :

In $su(3)$, we will carry out exactly the same exercise. Conceptually there won't be any leap forward, but unlike $su(2)$, $su(3)$ has two generators that commute with each other. We shall choose simultaneous eigenvectors of those two operators as our basis vectors. That will result in a little more work on our part, and some extra features in the spectrum of $su(3)$. For example,

- Unlike $su(2)$, $su(3)$ does not have irreps of every dimension.
- $su(3)$ has two inequivalent irreps of dimension 3, and they are denoted by 3 and $\bar{3}$. Combining these two gives $3 \otimes \bar{3} = 8 \oplus 1$. We also have $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$. These two decompositions have special significance in the context of elementary particle physics as we will learn soon.

3.2 Lecture 8 : January 20, 2016

Let $\exp(iA)$ be an arbitrary $SU(3)$ matrix. This implies that A is a 3×3 traceless Hermitian matrix. The most general 3×3 traceless Hermitian matrix can be written in the following form :

$$A = \begin{bmatrix} a_3 + \frac{a_8}{\sqrt{3}} & a_1 - ia_2 & a_4 - ia_5 \\ a_1 + ia_2 & -a_3 + \frac{a_8}{\sqrt{3}} & a_6 - ia_7 \\ a_4 + ia_5 & a_6 + ia_7 & -\frac{2a_8}{\sqrt{3}} \end{bmatrix} \tag{91}$$

where $a_i \in \mathbb{R}$, $i = 1(1)8$. It is evident[14] that 3×3 traceless Hermitian matrices form a real vector space of dimension 8. We consider the real vector space spanned by λ's and not the complex vector vector space because the space is that of Hermitian matrices. The above has been a standard choice of parametrization

[14] 3×3 complex matrices have 9 complex entries $\Rightarrow$ 18 real parameters. The condition for being Hermitian gives 3 complex equations and 3 real equations $\Rightarrow$ 9 real equations in total. Tracelessness gives another real equation. Hence we are left with $18 - 9 - 1 = 8$ independent real parameters.

of the $SU(3)$ matrices since pretty early. From the above parametrization, eight (independent) basis vectors can be read off :

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$
$$\lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix} \quad \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \tag{92}$$

These are known as the Gell-Mann matrices, as we mentioned earlier. In the space of complex matrices one can define a norm[15] $\mathrm{Tr}\left(A^\dagger A\right)$. All the Gell-Mann matrices have the same norm, namely 2. The first three matrices $\lambda_{1,2,3}$ are nothing but twice $T_{1,2,3}$ with an extra row and an extra column of zeroes appended at the end. It is straightforward to see that $\mathrm{span}\left\{\lambda_1, \lambda_2, \lambda_3\right\}$ is a subalgebra which is isomorphic to $su(2)$. It is customary to choose the following as generators of $su(3)$:

$$T_x \equiv \frac{\lambda_1}{2}, T_y \equiv \frac{\lambda_2}{2}, T_z \equiv \frac{\lambda_3}{2}, V_x \equiv \frac{\lambda_4}{2}, V_y \equiv \frac{\lambda_5}{2}, U_x \equiv \frac{\lambda_6}{2}, U_y \equiv \frac{\lambda_7}{2}, Y \equiv \frac{\lambda_8}{\sqrt{3}} \tag{93}$$

Define

$$T_\pm = T_x \pm iT_y,\ U_\pm = U_x \pm iU_y,\ V_\pm = V_x \pm iV_y \tag{94}$$

You might wonder why there is a T_z operator and no U_z or V_z operator. We already stated without proof that $su(3)$ has a maximum of two commuting generators. In our choice, these are T_z and Y, the diagonal matrices. As a result, $T_{x,y,z}$ form the $su(2)$ subalgebra. If we change basis and diagonalize[16] $-\frac{1}{2}T_z + \frac{3}{4}Y$ (call it U_z), then $U_{x,y,z}$ will form the $su(2)$ subalgebra. So, the reason why we call the first three operators $T_{x,y,z}$ is essentially because they give the $su(2)$ subalgebra and we choose to name them with the letter T.

An abstract $su(3)$ Lie algebra will have generators $t_+, t_-, t_z, u_+, u_-, v_+, v_-, y$

[15] This satisfies all the defining properties of a norm, namely

- $\| A \| \equiv \mathrm{Tr}\left(A^\dagger A\right) \geq 0$ with equality holding iff $A = 0$.
- $\| zA \| = |z| \| A \|, \forall z \in \mathbb{C}$
- $\| A + B \| \leq \| A \| + \| B \|$

[16] The reason for taking this particular linear combination will become apparent soon.

with the table for Lie products given by

	t_+	t_-	t_z	u_+	u_-	v_+	v_-	y
t_+	0	$2t_z$	$-t_+$	v_+	0	0	$-u_-$	0
t_-	$-2t_z$	0	t_-	0	$-v_-$	u_+	0	0
t_z	t_+	$-t_-$	0	$-\frac{1}{2}u_+$	$\frac{1}{2}u_-$	$\frac{1}{2}v_+$	$-\frac{1}{2}v_-$	0
u_+	$-v_+$	0	$\frac{1}{2}u_+$	0	$\frac{3}{2}y - t_z$	0	t_-	$-u_+$
u_-	0	v_-	$-\frac{1}{2}u_-$	$-\frac{3}{2}y + t_z$	0	$-t_-$	0	u_-
v_+	0	$-u_+$	$-\frac{1}{2}v_+$	0	t_-	0	$\frac{3}{2}y + t_z$	$-v_+$
v_-	u_-	0	$\frac{1}{2}v_-$	$-t_-$	0	$-\frac{3}{2}y - t_z$	0	v_-
y	0	0	0	u_+	$-u_-$	v_+	$-v_-$	0

(95)

You know everything about $su\,(3)$ from this table. Here, let us briefly state the motivation for studying $su\,(3)$, or any algebra for that matter.

Suppose that a physical system is theoretically described by some internal[17] coordinates as well as coordinates in space-time. Further suppose that there are three internal coordinates and an $SU\,(3)$ transformation in the internal space does not change the Physics of the system $\Rightarrow SU\,(3)$ is a symmetry of the system. Now suppose that a physical object associated with this system has n-components and an $SU\,(3)$ transformation in the internal space leads to a change of values of these components. Therefore, we need to figure out how n-component objects change when we effect an $SU\,(3)$ transformation. In other words, we have to look for representations of $SU\,(3)$. We are doing this because we assume $SU\,(3)$ is an exact symmetry of the system under consideration. We will soon see that $SU\,(3)$ is an approximate symmetry of nature.

Upshot : we will now look for representations of $su\,(3)$ or $SU\,(3)$.

1. The trivial representation : Let all the generators map to the real number 0. The vector space on which this representation is defined is $\mathbb{R}$. Linear transformations on $\mathbb{R}$ is nothing but multiplication by numbers. All the generators mapping onto 0 implies that all the group elements map on to $e^0 = 1$. This is obviously a representation but is trivial. Every group/algebra has this representation.

2. The Adjoint representation : This is another representation that every Lie algebra has. A representation of $SU\,(3)$ is, by definition, a homomorphic map from the group $SU\,(3)$ to invertible linear transformations on a vector space V. In what we discussed so far, we have already seen a vector space – the Lie algebra itself! $su\,(3)$ is an 8-dimensional vector space at heart. The adjoint representation of $su\,(3)$, defined in the following way, is therefore

[17] As an example of internal coordinates, let's think of a system described by a Lagrangian which is made out of two scalar fields ϕ_1, ϕ_2 and their spacetime derivatives. One solves for the field configurations $\phi_{1,2}$ that satisfy the equations of motion derived from the Lagrangian. Further assume that a general field configuration satisfying the e.o.m is of the form $\alpha\phi_1^0 + \beta\phi_2^0$ for arbitrary numbers α, β. Then we say that α, β are coordinates of this system in the "internal" space that is spanned by the particular fields ϕ_1^0, ϕ_2^0.

an 8-dimensional representation on the vector space $su(3)$.

$$\forall x \in su(3),\ x \mapsto ad(x) :\ \forall y \in su(3),\ ad(x)(y) \equiv [x,y] \tag{96}$$

Evidently, $ad(x)$ is a linear transformation on $su(3)$ because the Lie product is bilinear.

$$ad(x)\{ay+bz\} = a\{ad(x)(y)\} + b\{ad(x)(z)\}, \forall x,y,z \in su(3) \tag{97}$$

$ad(x)$, being a linear transformation, can be expressed in the matrix form. The Lie algebra of the operators $ad(x)$ have the natural Lie product – matrix commutators. Whenever we write $[ad(x), ad(y)]$, it is understood that this is a commutator and is also the Lie product. We need to check that the map $x \mapsto ad(x)$ is a homomorphism, i.e., structure preserving

$$[ad(x), ad(y)](z) = [ad(x)\,ad(y) - ad(y)\,ad(x)](z) = ad(x)[y,z] - ad(y)[x,z]$$

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

$$= ad([x,y])(z) \tag{98}$$

In the second line we have used two properties of the Lie bracket – antisymmetry and the Jacobi identity. Since the above holds $\forall z \in su(3)$, therefore the operators are equal :

$$[ad(x), ad(y)] = ad([x,y]) \tag{99}$$

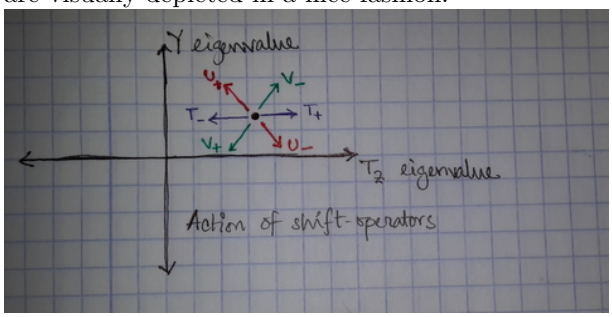
Hence the map is structure preserving. In establishing the adjoint representation above, we have only used the general properties of Lie brackets and not any property particular to $su(3)$. This tells us that every Lie algebra has an adjoint representation.
Everything about the adjoint operator $ad(x)$ corresponding to an arbitrary element x of the Lie algebra $su(3)$ will be known if we know the adjoint operators $ad(t_i)$ corresponding to a set of basis vectors t_i of $su(3)$. This is how : let $x = \alpha_i t_i$. Then, for arbitrary $y \in su(3)$ such that $y = \beta_j t_j$, $ad(x)(y) = [x,y] = [\alpha_i t_i, \beta_j t_j] = \alpha_i\beta_j [t_i,t_j] = \alpha_i\beta_j ad(t_i)(t_j)$. Therefore, due to bilinearity of Lie products, all we need to know is the action of $ad(t_i)$ on arbitrary t_j. This information is provided in table (95). Let us work out the matrix form of $ad(t_+)$ with respect to the ordered basis $\{t_+, t_-, t_z, u_+, u_-, v_+, v_-, y\}$:

$$ad(t_+) = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{100}$$

Let us now notice something interesting that happens, of course not by accident :

$$\begin{array}{cc} ad\left(t_z\right)\left(t_+\right) = t_+ & ad\left(y\right)\left(t_+\right) = 0 \\ ad\left(t_z\right)\left(t_-\right) = -t_- & ad\left(y\right)\left(t_-\right) = 0 \\ ad\left(t_z\right)\left(t_z\right) = 0 & ad\left(y\right)\left(t_z\right) = 0 \\ ad\left(t_z\right)\left(u_+\right) = -\frac{1}{2}u_+ & ad\left(y\right)\left(u_+\right) = u_+ \\ ad\left(t_z\right)\left(u_-\right) = \frac{1}{2}u_- & ad\left(y\right)\left(u_-\right) = -u_- \\ ad\left(t_z\right)\left(v_+\right) = -\frac{1}{2}v_+ & ad\left(y\right)\left(v_+\right) = -v_+ \\ ad\left(t_z\right)\left(v_-\right) = \frac{1}{2}v_- & ad\left(y\right)\left(v_-\right) = v_- \\ ad\left(t_z\right)\left(y\right) = 0 & ad\left(y\right)\left(y\right) = 0 \end{array} \tag{101}$$

That is, the basis vectors $t_+, t_-, t_z, u_+, u_-, v_+, v_-, y$ are simultaneous eigenvectors of $ad\left(t_z\right)$ and $ad\left(y\right)$. t_z and y commute, therefore so do $ad\left(t_z\right)$ and $ad\left(y\right)$. That implies that $ad\left(t_z\right)$ and $ad\left(y\right)$ can be simultaneously diagonalized. In the chosen basis, we see that simultaneous diagonalization of $ad\left(t_z\right)$ and $ad\left(y\right)$ has been achieved. In an arbitrary representation of $su\left(3\right)$, we can choose simultaneous eigenvectors of t_z and y as the basis of the vector space. The basis states will then be labeled by the t_z and y eigenvalues (a, b), say. Now, the commutations in equation (101) tell us how on application of the various generators $t_+, t_-, u_+, u_-, v_+, v_-$ on a state with eigenvalues (a, b) we obtain states with shifted eigenvalues $\left(a^{'}, b^{'}\right)$. Therefore the operators $t_+, t_-, u_+, u_-, v_+, v_-$ are called **shift operators**. In a diagram where we plot the eigenvalues of t_z and y of the simultaneous eigenstates, the effect of application of the shift operators are visually depicted in a nice fashion.



Let us quickly lay down the steps for what we are going to do tomorrow.

Killing Form :

Killing form on a Lie algebra $\mathcal{L}$ is a map $g : \mathcal{L} \times \mathcal{L} \to F$, where F is the field over which $\mathcal{L}$ is a vector space, defined in the following way :

$$g\left(a, b\right) \equiv \left(a, b\right) \equiv g_{ab} = Tr\left[ad\left(a\right) ad\left(b\right)\right] \tag{102}$$

g_{ab} is clearly symmetric in a, b because of cyclicity of trace. It is like an inner product but not quite. The field F is either $\mathbb{R}$ or $\mathbb{C}$ for almost all the cases of interest to us.

Cartan Subalgebra :

The maximal Abelian subalgebra of a Lie algebra is is called its Cartan subalgebra.

We will define tomorrow what a subalgebra is.

3.3 Lecture 9 : January 21, 2016

We will start with a few definitions.

Subalgebra :

$\mathcal{M}$ is a subalgebra of a Lie algebra $\mathcal{L}$ if it is a subspace of $\mathcal{L}$ and $\forall x, y \in \mathcal{M}$, $[x, y] \in \mathcal{M}$.

Recall the connection between Lie groups and Lie algebras. If $x, y \in$some Lie algebra then $e^{ix}, e^{iy} \in$the associated Lie group. Let $e^{\mathcal{M}}$ denote the set of elements of the Lie group that are generated by the elements of the subalgebra $\mathcal{M}$. For $x, y \in \mathcal{M}$, $e^{ix}.e^{iy} \in e^{\mathcal{M}}$, because commutator of x, y is in $\mathcal{M}$. Therefore, $\mathcal{M}$ being a subalgebra of the Lie algebra $\mathcal{L} \Rightarrow e^{\mathcal{M}}$ is a subgroup of the Lie group $e^{\mathcal{L}}$.

Normal Subgroup :

A subgroup N of a group G is said to be a normal subgroup of G, denoted $N \lhd G$, if $gNg^{-1} = N \ \forall g \in G$.

Suppose $N \lhd G$, where G is a Lie group. Let $\mathcal{N}$ be the subset of the Lie algebra $\mathcal{L}$ of G, elements of which generate elements of N. $\mathcal{N}$ is a subalgebra for sure, but it generates a subgroup of G with a special property (N is normal). Hence, one expects that $\mathcal{N}$ will have some special properties as well. It indeed does, and Lie subalgebras which generate normal subgroups are called **Ideals**.

Ideal :

$\mathcal{I}$ is an ideal of a Lie algebra $\mathcal{L}$ if it is a subalgebra of $\mathcal{L}$ and obeys the following

$$\forall x \in \mathcal{I}, \forall y \in \mathcal{L}, [x, y] \in \mathcal{I}$$

Every Lie algebra has two trivial ideals – span $\{0\}$ and $\mathcal{L}$. Ideals generate normal subgroups in the Lie group. The subgroups generated by span $\{0\}$ and $\mathcal{L}$ are the trivially normal subgroups $\{e\}$ and the full Lie group. Ideals are subalgebras. They can be abelian or non-abelian. span $\{0\}$is an abelian ideal that every Lie algebra has.

Simple Lie Algebra :

A simple Lie algebra is a Lie algebra which is non-abelian and has no non-trivial ideal.

Semisimple Lie Algebra :

A Lie algebra is semisimple if it does not contain any non-zero abelian ideals.

Semisimple Lie algebras can be built by taking direct products of simple lie algebras. The theory of which we are going to scratch the surface is that of semisimple Lie algebras. Some examples :

- Let $\mathcal{L}$ be a Lie algebra and $x \in \mathcal{L}$ be such that x commutes with all elements of $\mathcal{L}$. Then span $\{x\}$ is an ideal.
- $su(2)$: This does not even have a nontrivial subalgebra. Hence it is simple.
- $su(3)$: This does have a subalgebra : span $\{t_z, y\}$. However, this subalgebra is not an ideal. Therefore, $su(3)$ is simple. The subalgebra span $\{t_z, y\}$ is a Lie algebra in its own right. And this is an abelian algebra. Hence span $\{t_z, y\}$ is not simple.
- $SU(3) \times SU(2) \times U(1)$: The Lie algebra corresponding to this Lie group is not semisimple. We shall work out the details in one of the assignments.

Killing Form :

Killing form on a Lie algebra $\mathcal{L}$ is a map $g : \mathcal{L} \times \mathcal{L} \to F$, where F is the field over which $\mathcal{L}$ is a vector space, defined in the following way :

$$g(a,b) \equiv (a,b) \equiv g_{ab} = Tr\left[ad(a)\, ad(b)\right] \tag{103}$$

The field F is either $\mathbb{R}$ or $\mathbb{C}$ for almost all the cases of interest to us. Properties of the Killing form :

- $(a,b) = (b,a)$ because of cyclicity of trace.
- $(a, b+c) = (a,b) + (a,c)$ because matrix multiplication is distributive over addition and trace is additive.

If we also had the property $(a,a) \geq 0$, with $(a,a) = 0$ iff $a = 0$, then $(,)$ would define a nice inner product on the Lie algebra. The Killing form does not have this property, unfortunately. $Tr\left(A^\dagger A\right)$ for complex matrices A and $Tr\left(A^T A\right)$ for real matrices A has this property and all the other properties necessary for qualifying as an inner product.

Let X_a's be the generators of a Lie algebra $\mathcal{L}$ and $[X_a, X_b] = f_{ab}^c X_c$. $ad\,(X_a)$ is a linear operator on the vector space $\mathcal{L}$ which has X_a's as a set of basis vectors. In this basis, we can look for the matrix form D_a of $ad\,(X_a)$.

$$ad\,(X_a)\,X_b = (D_a)_b^c\,X_c$$

$$\Rightarrow (D_a)_b^c = f_{ab}^c \tag{104}$$

Matrices representing operators $ad\,(X_a)$ that form the adjoint representation of $\mathcal{L}$ are constructed out of the structure constants. If the generators are all Hermitian, then their commutators will be anti-Hermitian, in which case the structure constants will be imaginary. Physicists often prefer real structure constants and therefore define them as follows :

$$[X_a, X_b] = i f_{ab}^c X_c \tag{105}$$

Now,

$$g_{ab} = (a,b) = Tr\,[D_a D_b] = (D_a)_d^c\,(D_b)_c^d = f_{ad}^c f_{bc}^d \tag{106}$$

The structure constants are anti-symmetric in the two lower indices. However, they are not totally anti-symmetric. We studied two different sets of basis vectors of $su\,(2)$. The structure constants obtained with the basis $\{t_x, t_y, t_z\}$ did come out to be totally anti-symmetric – ϵ_{ijk}. But those obtained with the basis $\{t_\pm, t_z\}$ were not. In general, therefore, structure constants are not totally anti-symmetric. But there exists a basis for every Lie algebra for which f_{ab}^c is totally anti-symmetric. In that basis, we do not use an upper index and two lower indices for f_{ab}^c because all three indices have the same stature. Also, the Killing form becomes diagonal in that basis.

Basis dependence of structure constants :

Let us choose another set of basis vectors $X_a^{'}$, which can be expressed as linear combinations of the old basis vectors X_i, and of course vice versa :

$$X_a^{'} = L_a^i X_i;\ X_i = L_i^a X_a^{'} \tag{107}$$

where $L_a^i \equiv (L)_{ai}$, the $(ai)^{th}$ element of the basis transformation matrix L and $L_i^a \equiv \left(L^{-1}\right)_{ia} = \left(\left(L^{-1}\right)^T\right)_{ai}$. Now,

$$\left[X_a^{'}, X_b^{'}\right] = \left[L_a^i X_i, L_b^j X_j\right] = L_a^i L_b^j f_{ij}^k X_k = L_a^i L_b^j f_{ij}^k L_k^c X_c^{'}$$

$$\Rightarrow f_{ab}^{'c} = L_a^i L_b^j f_{ij}^k L_k^c \tag{108}$$

In terms of matrices of the operators in the adjoint representation :

$$\left(D_a^{'}\right)_b^c = L_a^i L_b^j\,(D_i)_j^k\,L_k^c$$

$$\Rightarrow D_a^{'} = L\left(L_a^i D_i\right)L^{-1} \tag{109}$$

Interpreting the nice equation (109) : $\left(L_a^i D_i\right)$ corresponds to the operator $X_a^{'}$ in the adjoint representation. Since D_i's are the matrix forms of the operators in the adjoint representation given that the basis being used is $\{X_a\}$, therefore, $X_a^{'}$'s matrix form is precisely $\left(L_a^i D_i\right)$ given the basis $\{X_a\}$. This matrix will change its form if we change the basis from $\{X_a\}$ to $\left\{X_a^{'}\right\}$ via the transformation L. The effect is the similarity transformation $A \rightarrow LAL^{-1}$. In particular, $D_a^{'}|_{\text{old basis}} \equiv \left(L_a^i D_i\right) \rightarrow L\left(L_a^i D_i\right) L^{-1} \equiv D_a^{'}|_{\text{new basis}}$. Given this,

$$g_{ab}^{'} = Tr\left(D_a^{'} D_b^{'}\right) = Tr\left(L\left(L_a^i D_i\right)\left(L_b^j D_j\right) L^{-1}\right) = Tr\left(\left(L_a^i D_i\right)\left(L_b^j D_j\right)\right) \text{ from cycility of trace}$$

$$= L_a^i L_b^j Tr\left(D_i D_j\right) = L_a^i L_b^j g_{ij} \tag{110}$$

We can think of g_{ab} as being elements of a matrix g. Thus,

$$g^{'} = LgL^T \tag{111}$$

Since g is a symmetric matrix, it can be orthogonally diagonalized. Hence, we can find a matrix L such that LgL^T is a diagonal matrix. Using this matrix L to find the new set of bases therefore renders the new Killing form $g^{'}$ diagonal. With such a choice of basis, let

$$g = \begin{bmatrix} \lambda_1 & & \\ & \ddots & 0 \\ & 0 & \lambda_n \end{bmatrix} \tag{112}$$

An interesting result : the basis in which g is diagonal is the one in which the structure constants become totally antisymmetric. In this basis, all the three indices of $f_{ab}^{\ c}$ are written downstairs – f_{abc}. A further basis transformation (scaling) such as $X_i \rightarrow X_i^{'} = \frac{1}{\sqrt{|\lambda_i|}} X_i$ (where i is such that $\lambda_i \neq 0$) turns all the non-zero diagonal elements into $+1$ or -1. The corresponding transformation matrix would be

$$L = \begin{bmatrix} \frac{1}{\sqrt{|\lambda_1|}} & & \\ & \ddots & 0 \\ & 0 & \frac{1}{\sqrt{|\lambda_n|}} \end{bmatrix} \tag{113}$$

Some of the eigenvalues λ_i may be zero.

Theorem :

A Lie algebra is semisimple iff g has no vanishing eigenvalues.

We shall probably do half the proof in one of the assignments. The full proof is quite involved.

4 Week 4

4.1 Lecture 10 : January 25, 2016

The following is the conventional choice of generators for $su\,(3)$:

$$\frac{\lambda_1}{2} \mapsto t_1, \frac{\lambda_2}{2} \mapsto t_2, \frac{\lambda_3}{2} \mapsto t_3, \frac{\lambda_4}{2} \mapsto v_1, \frac{\lambda_5}{2} \mapsto v_2$$

$$\frac{\lambda_6}{2} \mapsto u_1, \frac{\lambda_7}{2} \mapsto u_2, \frac{\lambda_8}{2} \mapsto \frac{\sqrt{3}}{2} y \equiv m \tag{114}$$

The last generator m is defined this way because this leads to a nicer symmetry in the weight diagrams. We shall soon explain what that means. With this definition of generators, we can check (from the explicit forms of the Gell-Mann matrices), that the first three generators satisfy $[t_i, t_j] = i\epsilon_{ijk} t_k$. This is a subalgebra of $su\,(3)$ isomorphic to $su\,(2)$. Define

$$\begin{aligned} t_\pm &= t_1 \pm i t_2 \\ u_\pm &= u_1 \pm i u_2 \\ v_\pm &= v_1 \pm i v_2 \end{aligned} \tag{115}$$

Now we look for irreps of $su\,(3)$. Let the operators in an irrep of $su\,(3)$ be denoted by capital Latin alphabets T, U, V, M etc. $[t_3, m] = 0 \Rightarrow [T_3, M] = 0$. Thus, simultaneous eigenstates of T_3 and M exist. We choose these simultaneous eigenstates as basis vectors. For notational brevity, define

$$\vec{G} = (T_3, M) \tag{116}$$

with eigenvalues $\vec{g} \equiv (t_3, m)$, also known as **weights**. Let the simultaneous eigenvectors be denoted by $|\alpha, \vec{g}\rangle$:

$$\vec{G}|\alpha, \vec{g}\rangle = \vec{g}|\alpha, \vec{g}\rangle \tag{117}$$

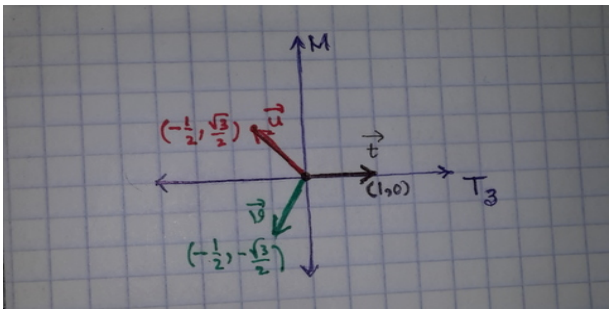
where α denotes the additional label(s) that may be required to identify the states. In this notation,

$$\begin{aligned} \left[\vec{G}, T_\pm\right] &= \pm \vec{t} T_\pm \qquad \vec{t} = (1, 0) \\ \left[\vec{G}, U_\pm\right] &= \pm \vec{u} U_\pm \quad \vec{u} = \left(-\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\right) \\ \left[\vec{G}, V_\pm\right] &= \pm \vec{v} V_\pm \quad \vec{v} = \left(-\tfrac{1}{2}, -\tfrac{\sqrt{3}}{2}\right) \end{aligned} \tag{118}$$

The vectors $\vec{t}, \vec{u}, \vec{v}$ are called **root vectors** or simply **roots**. Notice that roots are all unit vectors and satisfy

$$\vec{t} + \vec{u} + \vec{v} = 0 \tag{119}$$

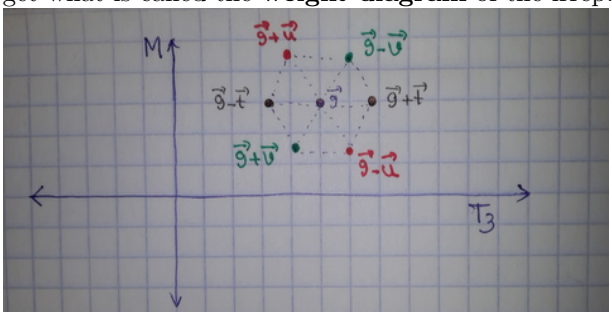
This is a consequence of defining m (or M) the way did. Without the factor of $\frac{\sqrt{3}}{2}$ in the definition $\frac{\sqrt{3}}{2} y \equiv m$, $\vec{u}, \vec{v}$ would not be unit vectors.



The commutation relations in (118) imply

$$\begin{array}{l} T_\pm|\alpha,\vec{g}\rangle, \text{if non-zero, has weight } \vec{g}\pm\vec{t} \\ U_\pm|\alpha,\vec{g}\rangle, \text{if non-zero, has weight } \vec{g}\pm\vec{u} \\ V_\pm|\alpha,\vec{g}\rangle, \text{if non-zero, has weight } \vec{g}\pm\vec{v} \end{array} \tag{120}$$

Therefore, if we have one state with weight $\vec{g}$, we are going to have six more states with "shifted weights" $\vec{g}\pm\vec{t}, \vec{g}\pm\vec{u}, \vec{g}\pm\vec{v}$ unless one or more of these "shifted states" happen to vanish. If we plot the weights of the states of an irrep on a plane with eigenvalues of T_3 and M plotted along the x and y axes respectively, we get what is called the **weight diagram** of the irrep.



The discussion above tells us that weight diagrams of irreps of $su\,(3)$ have a regular hexagonal symmetry. Again note that, the symmetry is regular because the roots are all unit vectors, which is a consequence of the peculiar definition of m. Now, if we are looking for finite dimensional irreps, we cannot have infinitely many states. Therefore, the weight diagram of a finite dimensional irrep cannot be an infinite hexagonal lattice – it **must have boundaries/edges**.

Ordering of the weights :

In irreps of $su\,(2)$, the eigenstates of T_3 are chosen as basis vectors. That is, every basis state is labeled by the eigenvalue of one operator because no two linearly independent $su\,(2)$ generators commute. As a result, the weights in $su\,(2)$ are just numbers, and the weight diagrams are composed of points lying on a line – the T_3 axis. The T_3 eigenvalues, or weights, can be compared and we have one state with the maximum weight for a given irrep. This is the so called **highest-weight state**. In $su\,(3)$, however, weight diagrams are two

dimensional, and points in 2-d do not have a natural order defined on them – one cannot simply say $(x_1, y_1) > (x_2, y_2)$ without first defining what one means by saying one point on a 2-plane is larger than another. So, we define what is known as the **dictionary order** :

$$(x_1, y_1) > (x_2, y_2) \text{ iff either } y_1 > y_2 \text{ or } y_1 = y_2 \text{ AND } x_1 > x_2 \tag{121}$$

Given two words in a dictionary, we first compare the first letter. The word which has a "smaller" first letter (that is, its first letter appears sooner in the alphabet than the first letter of the other word does) is "smaller" (that is, appears sooner in the dictionary). If the first letters match, then we compare the second letter and so on. This is exactly what we are doing here. Given two states, we first compare their M eigenvalues first, and then T_3 eigenvalues. This defines an order and we can find the highest weight state of an irrep of $su\,(3)$ unambiguously – climb vertically to reach the highest M eigenvalue, and then keeping the M eigenvalue fixed, go to your extreme right to find the highest weight state. We shall use this order as a convention. One might define another unambiguous order and work with that convention consistently.

Let the highest weight, according to dictionary order, be $\vec{g}_{\max}$ and the[18] highest weight state $|\alpha, \vec{g}_{max}\rangle$. Since T_+, U_+ both increase at least one of the eigenvalues of T_3 and M, therefore

$$\begin{array}{c} T_+|\alpha, \vec{g}_{max}\rangle = 0 \\ U_+|\alpha, \vec{g}_{max}\rangle = 0 \end{array} \tag{122}$$

V_+ need not kill $|\alpha, \vec{g}_{max}\rangle$ since V_- lowers eigenvalues of both T_3 and M. But, $V_-|\alpha, \vec{g}_{max}\rangle = 0$.

Equation (118) gives the commutation relations between $\vec{G}$ and the shift operators. We would also need to know commutation relations between the various shift operators. These can be derived very easily using (118), and some clever tricks. Of course these can also be figured out from the commutation relations of the Gell-Mann λ matrices which are the generators (modulo scaling) in the defining or fundamental[19] representation of $su\,(3)$. But the tricks we are going to learn will probably enhance our familiarity with the structure of the Lie algebra and teach us newer (and sometimes more beautiful) methods to solve problems.

Commutations $[T_+, T_-]\,, [U_+, U_-]\,, [V_+, V_-]$ **:**

Start by noticing that, in the basis $\{X_a\}_{a=1}^{8} \equiv \{T_x, T_y, T_z, U_x, U_y, V_x, V_y, M\}$ of our choice,

$$Tr\,(ad\,(X_a)\,.ad\,(X_b)) = k\delta_{ab} \tag{123}$$

[18] For an irrep, the highest weight state will not have any degeneracy, as we will soon prove. Hence we can talk about "the" highest weight state.

[19] $su\,(3)$ is by definition the algebra of generators of $SU\,(3)$ matrices. This is a concrete example of a Lie algebra. These are 3×3 matrices, and therefore can be thought of as linear operators on the vector space of 3×1 column vectors. On this vector space, $su\,(3)$ is its own representation. This is called the fundamental representation of $su\,(3)$.

where k is a fixed number $\forall a, b$ and δ_{ab} is the Kronecker delta. The number k can be easily found by direct computation, but we won't need its value. We will only use the fact that $Tr\left(ad\left(X_i\right).ad\left(X_i\right)\right) = Tr\left(ad\left(X_j\right).ad\left(X_j\right)\right)$, for all pairs i, j, with no summation implied on eiher side, which is a consequence of our careful scaling of the basis vectors. In particular,

$$Tr\left(ad\left(T_+\right).ad\left(T_-\right)\right) = Tr\left(ad\left(T_x\right)ad\left(T_x\right) + ad\left(T_y\right)ad\left(T_y\right) + iad\left(T_y\right)ad\left(T_x\right) - iad\left(T_x\right)ad\left(T_y\right)\right)$$

$$= 2k \tag{124}$$

With this in mind, let's compute $[T_+, T_-]$:

$$\text{Jacobi identity} \Rightarrow [A,[B,C]] = [B,[A,C]] + [[A,B],C]$$

$$\Rightarrow \left[\vec{G},[T_+,T_-]\right] = \left[T_+,\left[\vec{G},T_-\right]\right] + \left[\left[\vec{G},T_+\right],T_-\right]$$

$$= -\vec{t}[T_+,T_-] + \vec{t}[T_+,T_-] = 0$$

Since $[T_+, T_-]$ commutes with $\vec{G}$, therefore it commutes with all elements of the Cartan subalgebra $C \equiv \text{span}\{T_3, M\}$. Therefore, $[T_+, T_-]$ must also be a member of C. Thus,

$$[T_+,T_-] = \alpha_i G_i \ (= \alpha_1 T_3 + \alpha_2 M) \tag{125}$$

$$\Rightarrow ad\left([T_+,T_-]\right).ad\left(G_j\right) = \alpha_i ad\left(G_i\right).ad\left(G_j\right)$$

$$\Rightarrow Tr\left(ad\left([T_+,T_-]\right).ad\left(G_j\right)\right) = \alpha_i Tr\left(ad\left(G_i\right).ad\left(G_j\right)\right) = \alpha_i k\delta_{ij}, \text{ from (123)}$$

$$\Rightarrow Tr\left(\left[ad\left(T_+\right),ad\left(T_-\right)\right].ad\left(G_j\right)\right) = \alpha_j k, \text{ since } ad\left([T_+,T_-]\right) = \left[ad\left(T_+\right),ad\left(T_-\right)\right]$$

$$\Rightarrow Tr\left(\left[ad\left(G_j\right),ad\left(T_+\right)\right].ad\left(T_-\right)\right) = k\alpha_j \tag{126}$$

where we have used the fact that $Tr\left(\left[A,B\right].C\right) = Tr\left(\left[C,A\right].B\right)$ for three matrices A, B, C. This can be easily proved using cyclicity of trace. Now,

$$[G_j,T_+] = t_j T_+ \Rightarrow \left[ad\left(G_j\right),ad\left(T_+\right)\right] = ad\left([G_j,T_+]\right) = t_j ad\left(T_+\right)$$

Therefore, equation (126) implies

$$t_j Tr\left(ad\left(T_+\right).ad\left(T_-\right)\right) = k\alpha_j$$

Now, (124) implies

$$2kt_j = k\alpha_j \Rightarrow 2t_j = \alpha_j \tag{127}$$

Thus, $[T_+, T_-] = \alpha_i G_i = 2t_i G_i$. Similarly we can find $[U_+, U_-]$ and $[V_+, V_-]$.

$$\begin{aligned} [T_+,T_-] &= 2\vec{t}.\vec{G} \\ [U_+,U_-] &= 2\vec{u}.\vec{G} \\ [V_+,V_-] &= 2\vec{v}.\vec{G} \end{aligned} \tag{128}$$

Commutations $[T_+, U_\pm]$ and the like :

Adjoint representation of $su\,(3)$ is 8-dimensional. Hence, the two components of $ad\left(\vec{G}\right)$ are operators on an eight dimensional vector space. We note that $ad\left(\vec{G}\right)$ has precisely 8 eigenvalues corresponding to 8 linearly independent eigenvectors :

$$\begin{aligned} ad\left(\vec{G}\right): \qquad T_3 &\mapsto \left[\vec{G}, T_3\right] = 0 \\ M &\mapsto \left[\vec{G}, M\right] = 0 \\ T_+ &\mapsto \left[\vec{G}, T_+\right] = \vec{t} T_+ \\ T_- &\mapsto \left[\vec{G}, T_-\right] = -\vec{t} T_- \\ U_+ &\mapsto \left[\vec{G}, U_+\right] = \vec{u} U_+ \\ U_- &\mapsto \left[\vec{G}, U_-\right] = -\vec{u} U_- \\ V_+ &\mapsto \left[\vec{G}, V_+\right] = \vec{v} V_+ \\ V_- &\mapsto \left[\vec{G}, V_-\right] = -\vec{v} V_- \end{aligned} \tag{129}$$

Hence, the eight eigenvalues of $ad\left(\vec{G}\right)$ are $0, 0, \pm\vec{t}, \pm\vec{u}, \pm\vec{v}$. There can be no other eigenvalues of $ad\left(\vec{G}\right)$ since it is a linear operator on an eight dimensional space. With this is mind, let's compute $[T_+, U_-]$ using a trick.

$$ad\left(\vec{G}\right)([T_+, U_-]) = \left[\vec{G}, [T_+, U_-]\right] = \left[\left[\vec{G}, T_+\right], U_-\right] + \left[T_+, \left[\vec{G}, U_-\right]\right]$$

$$\Rightarrow ad\left(\vec{G}\right)([T_+, U_-]) = \left(\vec{t} - \vec{u}\right)[T_+, U_-] \tag{130}$$

That is, $[T_+, U_-]$, if non-zero, is an eigenvector of $ad\left(\vec{G}\right)$ with eigenvalue $\left(\vec{t} - \vec{u}\right)$! But, $\left(\vec{t} - \vec{u}\right)$ is not one of the eigenvalues of $ad\left(\vec{G}\right)$. Therefore, the only way in which equation (130) can be true is $[T_+, U_-] = 0$. This way, one can show that

$$[T_\pm, U_\mp] = [U_\pm, V_\mp] = [V_\pm, T_\mp] = 0 \tag{131}$$

Now we will employ the same kind of trick to compute $[T_+, U_+]$.

$$\left[\vec{G}, [T_+, U_+]\right] = \left[T_+, \left[\vec{G}, U_+\right]\right] + \left[\left[\vec{G}, T_+\right], U_+\right] = \left(\vec{t} + \vec{u}\right)[T_+, U_+] = -\vec{v}\,[T_+, U_+]$$

Now, $-\vec{v}$ **is** an eigenvalue of $ad\left(\vec{G}\right)$ and the corresponding eigenvectors are proportional to V_-. Hence, we can conclude that

$$[T_+, U_+] = \xi V_- \tag{132}$$

This is all we can say about the commutator $[T_+, U_+]$ using just the information about eigenvalues and eigenvectors of $ad\left(\vec{G}\right)$. The proportionality factor ξ has

to be figured out by explicit calculation and it turns out to be 1. Similar commutation relations are the following : $[T_-, U_-], [U_\pm, V_\pm], [V_\pm, T_\pm]$. In the equation below, I am summarizing all the non-zero commutators for $su(3)$:

$$\begin{aligned}
\left[\vec{G}, T_\pm\right] &= \pm\vec{t}T_\pm \\
\left[\vec{G}, U_\pm\right] &= \pm\vec{u}U_\pm \\
\left[\vec{G}, V_\pm\right] &= \pm\vec{v}V_\pm \\
[T_+, T_-] &= 2\vec{t}.\vec{G} \\
[U_+, U_-] &= 2\vec{u}.\vec{G} \\
[V_+, V_-] &= 2\vec{v}.\vec{G} \\
[T_\pm, U_\pm] &= \pm V_\pm \\
[U_\pm, V_\pm] &= \pm T_\pm \\
[V_\pm, T_\pm] &= \pm U_\pm
\end{aligned} \tag{133}$$

4.2 Lecture 11 : January 27, 2016

At the outset, let me give an overview by stating facts without proof, so that you understand the motivation behind the calculations we are going to perform soon. Today we are going to discuss $su(3)$ **super-multiplets** or the set of eigenstates in an irrep of $su(3)$. The eigenvalues corresponding to these states are plotted in the weight diagrams. We have already seen that $su(3)$ has as its subalgebra $su(2)$. In our conventional basis, span $\{T_x, T_y, T_z\}$ is a subalgebra isomorphic to $su(2)$. Also, span $\left\{U_x, U_y, \vec{u}.\vec{G}\right\}$, span $\left\{V_x, V_y, \vec{v}.\vec{G}\right\}$ are two other subalgebras, both isomorphic to $su(2)$. In a given irrep of $su(3)$, of dimension m, say, there will be m states. We will soon see that, some of these m states will span an invariant subspace[20] for the elements of the subalgebra span $\{T_x, T_y, T_z\}$. In fact, all the states belong to at least one such invariant subspace - invariant with respect to actions of elements of span $\{T_x, T_y, T_z\}$. The states belonging to one such invariant subspace form what is called an $su(2)$ **multiplet**. These different T-multiplets do not mix under the actions of $T_{x,y,z}$. They do mix under the actions of U and V operators though. These T-multiplets make up the whole of the $su(3)$ representation, that is why the set of all states is called a **super-multiplet** of $su(3)$. There is nothing special about the T operators – span $\left\{U_x, U_y, \vec{u}.\vec{G}\right\}$, span $\left\{V_x, V_y, \vec{v}.\vec{G}\right\}$ are also isomorphic to $su(2)$. Therefore, $su(3)$ super-multiplets can also be partitioned into U-multiplets or V-multiplets equivalently.

[20] Note that this invariant subspace I am talking about is not an invariant subspace of all elements of the algebra. Had that been the case, then the representation would become reducible!

In the brief discussion following equation (120) we learned that, given one state $|\alpha, \vec{g}\rangle$ with weight $\vec{g}$ of an irrep, we immediately get six more (shifted) states $T_{\pm}|\alpha, \vec{g}\rangle, U_{\pm}|\alpha, \vec{g}\rangle, V_{\pm}|\alpha, \vec{g}\rangle$ with weights $\left(\vec{g} \pm \vec{t}\right), (\vec{g} \pm \vec{u}), (\vec{g} \pm \vec{v})$ respectively. Since we are looking for finite dimensional irreps of $su(3)$, the weight diagrams will have **edges**. For a state on one of the edges, some of these shifted states will vanish. Let us look more closely at the **symmetries of the weight diagrams**.

Weyl Symmetry :

We start by defining three new unitary operators

$$\begin{array}{c} P_{\vec{t}} = e^{-i\pi T_2} \\ P_{\vec{u}} = e^{-i\pi U_2} \\ P_{\vec{v}} = e^{-i\pi V_2} \end{array} \tag{134}$$

Using Baker-Hausdorff lemma, one can explicitly calculate the operator $P_{\vec{t}}^{-1}\vec{G}P_{\vec{t}}$ (and also $P_{\vec{u}}^{-1}\vec{G}P_{\vec{u}}$,$P_{\vec{v}}^{-1}\vec{G}P_{\vec{v}}$). For your convenience, let me state the lemma here :

$$e^{A}Be^{-A} = \sum_{n=0}^{\infty} \frac{[A,B]_n}{n!} \tag{135}$$

where the symbol $[A,B]_n$ is defined recursively as follows

$$[A,B]_n = \left[A, [A,B]_{n-1}\right] \text{ with } [A,B]_o \equiv B \tag{136}$$

In the language of operators in the adjoint representation,

$$e^{A}Be^{-1} = e^{ad(A)}B \tag{137}$$

Equation (137) is neater and much more economical. We state the result without proof :

$$P_{\vec{t}}^{-1}\vec{G}P_{\vec{t}} = P_{\vec{t}}^{-1}\left(T_3, M\right)P_{\vec{t}} = \left(-T_3, M\right) \tag{138}$$

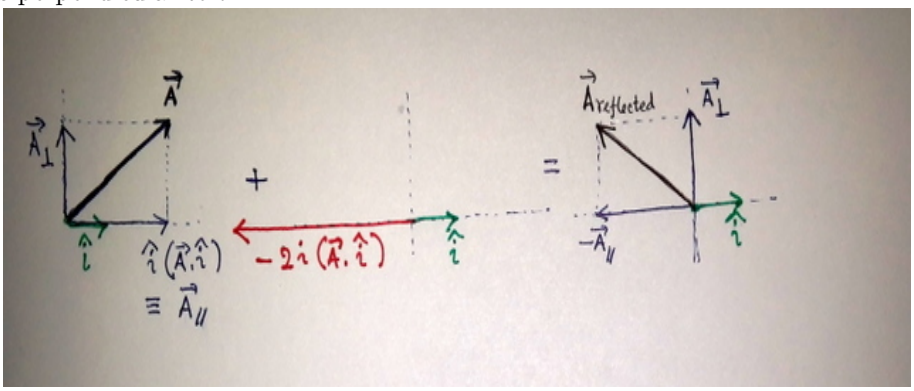
An alternative form which is going to be more illuminating for us is the following :

$$\begin{array}{c} P_{\vec{t}}^{-1}\vec{G}P_{\vec{t}} = \vec{G} - 2\vec{t}\left(\vec{t}.\vec{G}\right) \\ P_{\vec{u}}^{-1}\vec{G}P_{\vec{u}} = \vec{G} - 2\vec{u}\left(\vec{u}.\vec{G}\right) \\ P_{\vec{v}}^{-1}\vec{G}P_{\vec{v}} = \vec{G} - 2\vec{v}\left(\vec{v}.\vec{G}\right) \end{array} \tag{139}$$

Let us now see why these operators are useful. Suppose, $\vec{G}|\alpha, \vec{g}\rangle = \vec{g}|\alpha, \vec{g}\rangle$. Then,

$$\begin{aligned} \vec{G}P_{\vec{t}}|\alpha, \vec{g}\rangle &= P_{\vec{t}}\left(P_{\vec{t}}^{-1}\vec{G}P_{\vec{t}}\right)|\alpha, \vec{g}\rangle = P_{\vec{t}}\left(\vec{G} - 2\vec{t}\left(\vec{t}.\vec{G}\right)\right)|\alpha, \vec{g}\rangle \\ &= \left(\vec{g} - 2\vec{t}\left(\vec{t}.\vec{g}\right)\right)|\alpha, \vec{g}\rangle \end{aligned} \tag{140}$$

That is, $P_{\vec{t}}|\alpha,\vec{g}\rangle$ is either zero or is another eigenvector of $\vec{G}$ with weight $\left(\vec{g}-2\vec{t}\left(\vec{t}.\vec{g}\right)\right)$. However, $P_{\vec{t}}$ being a unitary operator cannot change the norm of a state. Since $|\alpha,\vec{g}\rangle$ is a non-zero state to begin with, therefore $P_{\vec{t}}|\alpha,\vec{g}\rangle$ cannot be zero. Similarly, $P_{\vec{u}}|\alpha,\vec{g}\rangle$, $P_{\vec{v}}|\alpha,\vec{g}\rangle$ are eigenstates of $\vec{G}$ with respective weights $(\vec{g}-2\vec{u}\,(\vec{u}.\vec{g}))$ and $(\vec{g}-2\vec{v}\,(\vec{v}.\vec{g}))$. Now, $\vec{t},\vec{u},\vec{v}$ are all unit vectors on a 2-plane. Suppose $\vec{A}$ is an arbitrary vector on a 2-plane and $\hat{i}$ is an arbitrary unit vector on the same plane. Then $\vec{A}=\vec{A}_{\perp}+\vec{A}_{\parallel}$ where $\vec{A}_{\perp}$ and $\vec{A}_{\parallel}$ are projections of $\vec{A}$ perpendicular and parallel to $\hat{i}$. Clearly, $\vec{A}_{\parallel}=\hat{i}\left(\hat{i}.\vec{A}\right)$. Therefore, $\vec{A}-2\hat{i}\left(\hat{i}.\vec{A}\right)=\vec{A}_{\perp}-\vec{A}_{\parallel}$, which is the vector obtained by reflecting $\vec{A}$ about the axis perpendicular to $\hat{i}$.



This tells us that, if a given weight $\vec{g}$ in the (T_3,M) plane is reflected about the line perpendicular to $\vec{t}$ (this is the M-axis) and going through the origin, the line perpendicular to $\vec{u}$ and going through the origin (call this $\vec{u}_{\perp}$), and the line perpendicular to $\vec{v}$ and going through the origin (call this $\vec{v}_{\perp}$), we get three more weights. This implies that the weight diagram is symmetric about the lines $\vec{t}_{\perp}(\equiv$ unit vector along the M axis), $\vec{u}_{\perp}$, $\vec{v}_{\perp}$. This is known as the **Weyl symmetry**. Notice that

$$P_{\vec{t}}^2=P_{\vec{u}}^2=P_{\vec{v}}^2=\mathbb{I} \tag{141}$$

which is consistent with $P_{\vec{t},\vec{u},\vec{v}}$ being reflection operators. Also, if a weight $\vec{g}$ is located **on one of the lines of symmetry**, namely the lines along $\vec{t}_{\perp},\vec{u}_{\perp},\vec{v}_{\perp}$, then it is its own reflection about that line. Using these pieces of information, we shall construct the shapes of the weight diagrams of a super-multiplet and their various properties.

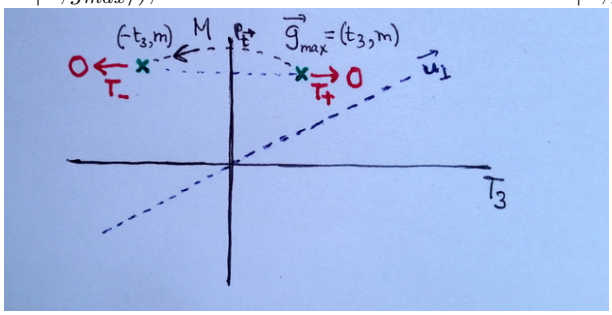
Exercise : Show that, $e^{i\theta T_2}\,(T_3,M)\,e^{-i\theta T_2}=((\cos\theta)\,T_3+(i\sin\theta)\,T_1,M)$.

Building and characterizing an $su\,(3)$ super-multiplet step by step :

- Let $|\alpha,\vec{g}_{max}\rangle$ be the highest weight state belonging to an irrep of $su\,(3)$. We are using the dictionary order for the weights, which is almost universally used. We shall assume that $|\alpha,\vec{g}_{max}\rangle$ is non-degenerate. Had

it been doubly degenerate, say, then we could orthogonalize the two linearly independent states with weight $\vec{g}_{max}$ by the Gram-Schmidt process. The various shift operators then would act on the two orthogonal, highest weight, states (denote them by $|1\rangle, |1^{'}\rangle$) to produce a bunch of states. However, it would turn out that the states obtained from $|1\rangle$ would form an invariant subspace under the action of all $su\,(3)$ operators. So would the states obtained from $|1^{'}\rangle$. Thus, we would end up with two copies of the same irrep direct summed with each other, which serves no useful purpose. Therefore, we choose to work with non-degenerate highest weight states.

- The highest weight state has a $P_{\vec{t}}$ partner – a state whose weight is obtained by reflecting $\vec{g}_{max}$ about the M axis – call it $|\alpha, \vec{g}_{max}^{partner}\rangle$. And since there is no state whose weight lies to the right of $\vec{g}_{max}$ (i.e., with T_3 eigenvalue greater than that of $|\alpha, \vec{g}_{max}\rangle$ and M eigenvalue equal to that of $|\alpha, \vec{g}_{max}\rangle$), there will be no state to the left of $|\alpha, \vec{g}_{max}^{partner}\rangle$.



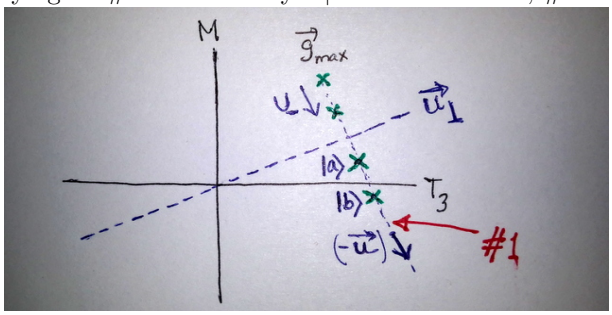
$$T_+|\alpha, \vec{g}_{max}\rangle = 0 = U_+|\alpha, \vec{g}_{max}\rangle \Rightarrow T_-|\alpha, \vec{g}_{max}^{partner}\rangle = 0$$

However, unless $\vec{g}_{max}$ lies on the line along $\vec{u}_\perp$, $U_-|\alpha, \vec{g}_{max}\rangle \neq 0$. One can reach $|\alpha, \vec{g}_{max}^{partner}\rangle$ by starting from $|\alpha, \vec{g}_{max}\rangle$ and applying on it T_- repeatedly until $T_-|\alpha, \vec{g}_{max}^{partner}\rangle = 0$. This set of states between $|\alpha, \vec{g}_{max}^{partner}\rangle$ and $|\alpha, \vec{g}_{max}\rangle$, including the two, form the top boundary, or the "roof", of the weight diagram and their weights lie on the straight line $M = m$. There cannot lie any weight above this line because otherwise that weight would be the highest weight and not (t_3, m).

- The states obtained by repeated actions of U_-starting from $|\alpha, \vec{g}_{max}\rangle$ will lie on the line along $-\vec{u}$. Of course this string of states has to end somewhere down the line (i.e., $(U_-)^q\,|\alpha, \vec{g}_{max}\rangle = 0$ for some $q \in \mathbb{W}$[21]) since we are looking for a finite dimensional irrep. I claim that these states will form another boundary of the weight diagram. We already know that T_+ kills $|\alpha, \vec{g}_{max}\rangle$. However, it requires proof that T_+ also kills all the states obtained by repeated application of U_- on $|\alpha, \vec{g}_{max}\rangle$. In fact, this is a general result for all the boundaries/edges of the weight diagram.

[21] $\mathbb{W} = \mathbb{N} \cup \{0\}$

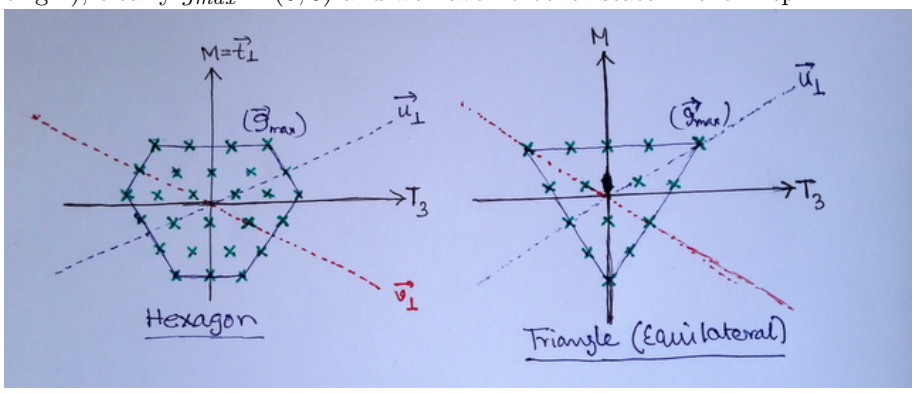
- The boundaries/edges of a weight diagram have to be straight lines. We established that the top boundary or the "roof" of the weight diagram has to be a straight line. We will now prove that the edge labeled by #1 in the diagram below also has to be a straight line. That is, all the states lying on #1 are killed by T_+– in other words, #1 is not jagged.



Let $|a\rangle$ be a state on #1 that gets killed by T_+ : $T_+|a\rangle = 0$. At least one such state exists – $|\alpha, \vec{g}_{max}\rangle$. Let $|b\rangle = U_-|a\rangle$. Now,

$$T_+|b\rangle = T_+U_-|a\rangle = U_-T_+|a\rangle = 0$$

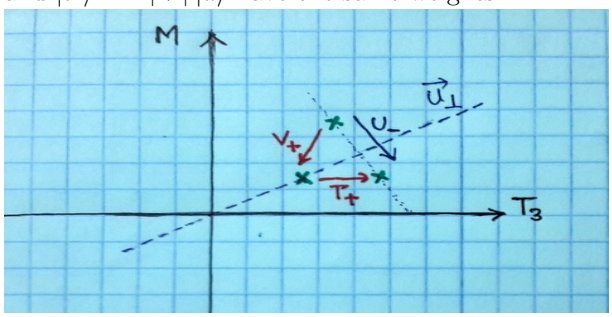
because $[T_+, U_-] = 0$. Hence the proof. This easily extends to all the edges. Therefore, the boundary of a weight diagram will be a closed polygon. The three lines of symmetry ensure that this polygon is a hexagon in general. A special case is when $\vec{g}_{max}$ lies on one of the lines of symmetry in which case the boundary of the diagram becomes triangular. When $\vec{g}_{max}$ lies on the point of intersection of $\vec{t}_\perp, \vec{u}_\perp, \vec{v}_\perp$(all of them pass through the origin), clearly $\vec{g}_{max} = (0, 0)$ and we have no other state in the irrep.



- Degeneracies : The next question to address is the following. We agreed to start with a non-degenerate highest weight state. Starting from$|\alpha, \vec{g}_{max}\rangle$, one gets all the states in the irrep by actions of various shift operators. In particular, we get all the boundary states starting from $|\alpha, \vec{g}_{max}\rangle$. Are these boundary states degenerate? Let us examine this for the states on #1. Suppose $|a\rangle$ be non-degenerate[22]. Now, the two states $|b^{'}\rangle \equiv U_-|a\rangle$

[22] Again, at least one such state exists – $|\alpha, \vec{g}_{max}\rangle$

and $|b^{'}\rangle \equiv T_+ V_+ |a\rangle$ have the same weights.



$[T_+, V_+] = -U_-$ implies

$$|b^{'}\rangle = T_+ V_+ |a\rangle = (V_+ T_+ - U_-)\,|a\rangle = -U_- |a\rangle = -|b\rangle$$

Hence, $|b\rangle$ and $|b^{'}\rangle$ are linearly dependent. Note that, there are many more ways to get from $|a\rangle$ to a state whose weight is equal to that of $|b\rangle$. Without providing a more general proof, we state that all these apparently different ways give rise to the same state modulo an overall phase which is physically irrelevant. So, our conclusion is the following : Boundary states have no degeneracy. How about the inner states (meaning states whose weights lie not on but inside the boundary)? Again, without a proof, let us state the result : Boundary states of a weight diagram are non-degenerate. If the boundary is hexagonal, then the next inner layer of states will have two-fold degeneracy. In subsequent inner layers, the degeneracy count will keep on increasing by 1 until the first triangular layer of states is reached. Once the first triangular layer, call it $\triangle_1$, is reached, then the subsequent inner layers will have the same degeneracy as $\triangle_1$. That is, degeneracy count will not increase any more. The layers of states inside $\triangle_1$ are all triangular, except for the case when the inner-most layer becomes a point.

- Finally, we note that all the boundary states, and in turn the entire super-multiplet, may be obtained from the highest weight state. Therefore, if we know $\vec{g}_{max}$, we know the entire super-multiplet.

We have skipped a few general proofs because they take a lot of effort and we are under strict time constraint. If you are interested, you can take a look at any good textbook on Lie algebra.

4.3 Lecture 12 : January 28, 2016

Yesterday we saw how we obtain the entire super-multiplet from the the highest weight state using Weyl symmetries. Today we shall work out a few examples and observe some other special features of the super-multiplets.

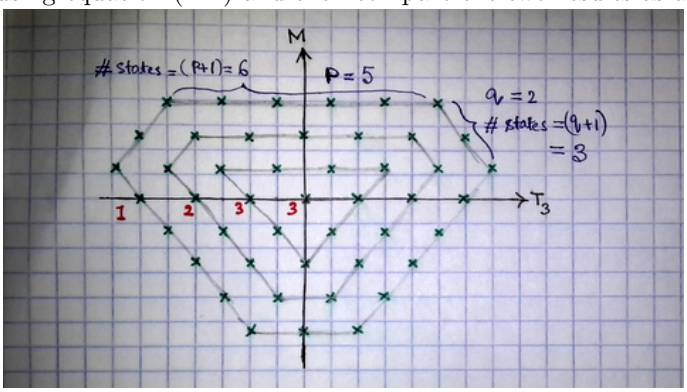
We learned how the knowledge of $\vec{g}_{max} = (t_3, m)$ fixes the entire super-multiplet. Therefore, we can uniquely label the $su(3)$ irreps by (t_3, m), the highest weight of that irrep. Equivalently, we can label an irrep by a set of two non-negative integers, (p, q), where p and q are, respectively, the numbers of times T_- and U_- can act on $|\alpha, \vec{g}_{max}\rangle$ without killing it. The numbers p, q tell us the perpendicular distances of the point $\vec{g}_{max}$ from the lines $\vec{t}_\perp, \vec{u}_\perp$ respectively. It is an easy problem in coordinate geometry to explicitly solve for

$$\vec{g}_{max} = \vec{g}_{max}(p, q)$$
$$(p, q) = f(\vec{g}_{max})$$

More often than not, the (p, q) labeling is adopted for $su(3)$ super-multiplets. We shall do the same. Starting from (p, q), one can draw the entire weight diagram and figure out the degeneracy count of each weight using the rules stated and partially proved yesterday. Thus, we can calculate the total number of states, or the dimension D, of a (p, q) super-multiplet. The result is stated here and left for you to prove in an exercise :

$$D = \frac{1}{2}(p+1)(q+1)(p+q+2) \tag{142}$$

Let us draw the weight diagram for $(p, q) = (5, 2)$, count the states, compute D using equation (142) and then compare the two results as a sanity check.

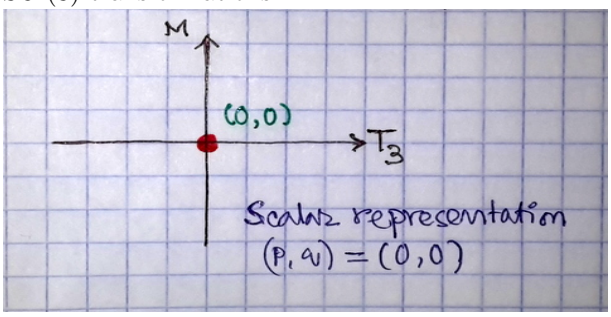


In the figure, the layers have been explicitly drawn, and the degeneracy count of each layer written adjacent to it in red ink. The outer-most layer is hexagonal, has 21 weights with degeneracy count of 1. The next inner layer is hexagonal, has 15 weights with degeneracy count of 2. The next inner layer is triangular (this one is $\triangle_1$, following the nomenclature from yesterday's class), has 9 states with degeneracy count of 3. Degeneracy count will not increase inside this layer. The inner-most layer is a point, i.e., 1 weight with degeneracy count of 3. Hence, we have a total of $21 + 2(15) + 3(9) + 3(1) = 81$ states. Using equation (142), $D = \frac{1}{2}(5+1)(2+1)(5+2+2) = 81$. Consistency checked! So, here we have an 81 dimensional irrep of $su(3)$. The representative matrices are 81×81. However, these matrices are going to be very sparse. Nonetheless, it is much more convenient to list the actions of the operators on the states instead of writing down the 81×81 matrices.

Notice that (142) is symmetric in p and q. For instance, $(2,1)$ and $(1,2)$ are two irreps of the same dimension, namely 15. These are two irreps of the same dimension, but they are inequivalent. These two super-multiplets have the same number of states, but different weight diagrams. That means, the eigenvalues of T_3 and M for these two representations are different. Had these been equivalent representations, the eigenvalues would be the same in both of them. This feature is unlike $su\,(2)$ which had only one (inequivalent) irrep of every dimension. $su\,(3)$ does not have an irrep of every dimension and, for some dimensions, has more than one inequivalent irrep.

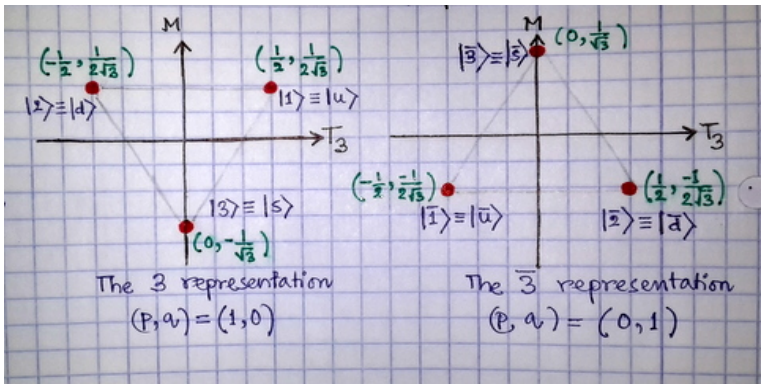
Our next program is to study some particular $su\,(3)$ super-multiplets that have a great role to play in particle physics.

- $(p,q) = (0,0)$:
  This is a 1-dimensional representation : $D = 1$. The only (linearly independent) state in this irrep sits at the origin of the weight diagram; its eigenvalues are $(t_3, m) = (0,0)$. This irrep is called the $su\,(3)$ **scalar representation**, since the state does not change[23] under the action of $SU\,(3)$ transformations.



- $(p,q) = (1,0)$:
  Here, $D = 3$. This is often referred to as the 3 representation. The three states are labeled $|1\rangle, |2\rangle, |3\rangle$, as depicted in the figure below. In the context of particle physics, they are often equivalently labeled $|u\rangle, |d\rangle, |s\rangle$, the letters being representative of the **up**, **down** and the **strange** quarks. We will make the connection with particle physics later.

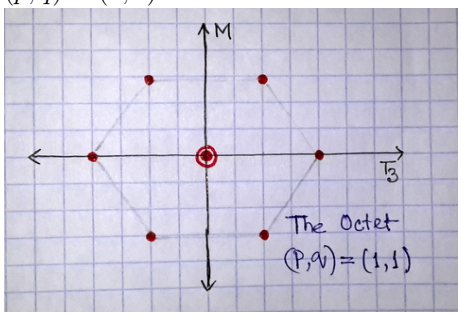
[23] There is still the provision of a state gaining an overall phase under the action of an $SU\,(3)$ transformation, but, overall phases being irrelevant, we choose the $SU\,(3)$ representative operators to be unity.



- $(p, q) = (0, 1)$
  Here also, $D = 3$. This irrep and the previous one are both 3-dimensional but are certainly inequivalent. One way to see it is that the M eigenvalues of the states in the previous irrep are $\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, -\frac{1}{\sqrt{3}}$, whereas those in this irrep are $\frac{1}{\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}$. Had these two irreps been equivalent, then the M eigenvalues would have been exactly the same. This irrep is called the $\bar{3}$ irrep and the states are labeled $|\bar{1}\rangle, |\bar{2}\rangle, |\bar{3}\rangle$, equivalently[24] $|\bar{u}\rangle, |\bar{d}\rangle, |\bar{s}\rangle$ as depicted in the figure above.

- $(p, q) = (2, 0)$
  This has $D = 6$ and is called the 6 irrep.

- $(p, q) = (0, 2)$
  This is another irrep with $D = 6$ and is inequivalent to the previous irrep. It is called the $\bar{6}$ irrep.

- $(p, q) = (1, 1)$
  

  This is the only 8-dimensional irrep of $su\,(3)$. This is the famous **octet** of the **eight-fold-way** in particle physics. We shall spend some time on it soon.

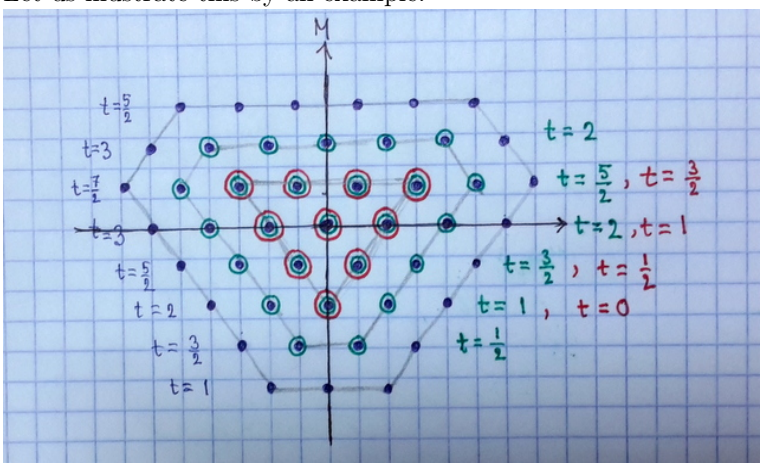
[24] Sometimes you will find texts in which the state $|\bar{2}\rangle$ is labeled $-|\bar{d}\rangle$. This is a choice or convention and the physics does not change because of it since overall phase factors are irrelevant in the description of a physical state.

We have already seen that $su\,(3)$ irreps have degenerate states in general. We need to find a way to distinguish degenerate states. Eigenvalues of T_3, M don't help. Following what is the standard refuge in quantum mechanics in such situations, we should find another operator that commutes with both T_3 and M, and has distinct eigenvalues for the states that have the same set of (t_3, m) eigenvalues. We notice that

$$\left[T^2, T_3\right] = 0 = \left[T^2, M\right] \tag{143}$$

We can use T^2 eigenvalues to distinguish the degenerate states. Given a weight diagram, focus on the states belonging to one of the horizontal levels. These states are connected by $T_\pm$. Under the actions of the three operators $T_{x,y,z}$ and their linear combinations ($\sim$ an $su\,(2)$ subalgebra of $su\,(3)$), these states form an invariant subspace. Therefore, each horizontal level (corresponding to a given M eigenvalue) contains $su\,(2)$ multiplets and can therefore be uniquely labeled by T^2 and T_3 eigenvalues. For states that form a $(2t+1)$ irrep of the $su\,(2)$ subalgebra, the T^2 eigenvalues will be $t\,(t+1)$– a standard $su\,(2)$ result.

Let us illustrate this by an example.



- In the weight diagram above (we have seen this once already), the topmost horizontal level, corresponding to the highest M eigenvalue, contains 6 weights, each with degeneracy one. All of these 6 states mix under actions of $T_{x,y,z}$. Hence, they form a 6-dimensional irrep of the $su\,(2)$ subalgebra. These states have $t = \frac{5}{2}$ (so that $(2t+1) = 6$).

- The next horizontal level has 12 states in 7 weights. Two of the weights lie on the outer boundary of the weight diagram and hence are non-degenerate. Each of the remaining 5 weights has a two-fold degeneracy. The state to the extreme right with weight (t_3, m), say, when acted upon by T_-, gives rise to a state with weight $(t_3 - 1, m)$ which is the second weight from the right on this horizontal level, and this weight has a degeneracy count of two. Therefore, starting from the extreme right state, we get one of the two states that have weight $(t_3 - 1, m)$ by application

of T_-. Continuing this way, by repeatedly applying T_-, we get a string of 7 states, one for every weight on this horizontal level, that form a basis of a 7-dimensional irrep of the $su\,(2)$ subalgebra. These states have $t = 3$, so that $(2t + 1) = 7$. Now we have 5 states remaining, each non-degenerate. Of these, the one on the extreme right has weight $(t_3 - 1, m)$. We assume that this state is orthogonal to the other state that has the same weight $(t_3 - 1, m)$ and belongs to the 7-dimensional irrep that we just constructed. Now we start from this state, keep on applying T_-, and get a string of 5 states that mix under actions of $T_{x,y,z}$. The states thus obtained are orthogonal to the states that have the same weights but belong to the $t = 3$ irrep on this level. These 5 states have $t = 2$, so that $(2t + 1) = 5$. Thus we exhaust all the 12 states on this level.

- We continue doing this for all the levels, eventually exhausting the entire super-multiplet.

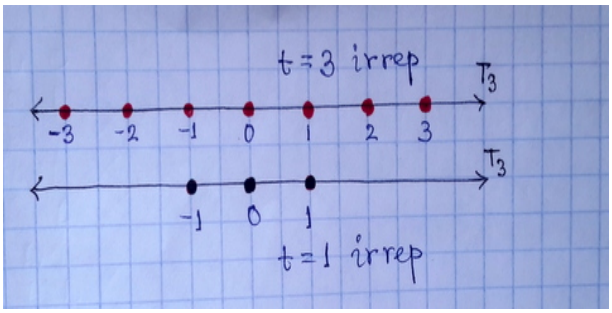
Having more or less understood the details of $su\,(3)$ super-multiplets and their weight diagrams, we would like to study bigger representations formed by taking direct products of irreps. In fact, we would be interested in breaking down the product representations, which are reducible, in terms of their component irreps. We did the same exercise[25] for $su\,(2)$ and there is no conceptual leap forward in $su\,(3)$. However, taking the same approach as in $su\,(2)$ would result in quite lengthy and tedious calculations. There is an alternative, geometric way of finding out the dimensions of the component irreps of a product representation. We shall first illustrate this method for $su\,(2)$ and then straightforwardly generalize for $su\,(3)$. Note, however, that this method only tells us the dimensions of the invariant subspaces of a bigger space that carries a product representation. It does not tell us which particular subspaces are invariant. In order to find that detail, we need to follow the complete Clebsch-Gordan prescription and find the states that span the invariant subspaces.

$su\,(2)$ Direct Products and their Decomposition :

Consider taking the direct product of the two $su\,(2)$ irreps : $t = 3$ (7-dimensional) and $t = 1$ (3-dimensional). The weight diagrams[26] of the two irreps would be the following :

[25] The Clebsch-Gordan series and decomposition

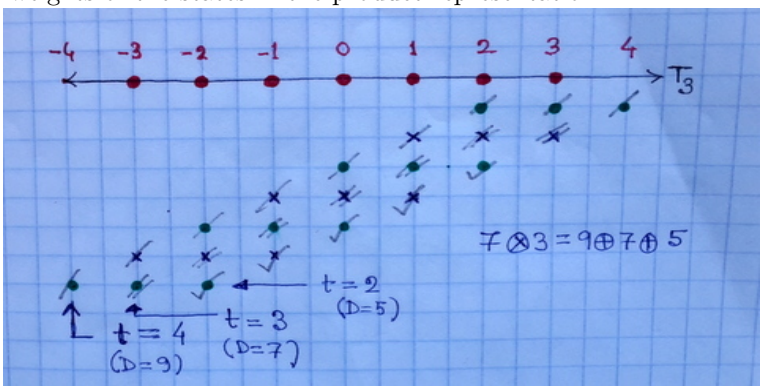
[26] $su\,(2)$ weight diagrams are one dimensional.



The states in the two irreps are, respectively, $\{|3;\pm3\rangle, |3;\pm2\rangle, |3;\pm1\rangle, |3;0\rangle\}$ and $\{|1;\pm1\rangle, |1;0\rangle\}$, where we adopt the familiar $su\,(2)$ labeling scheme $|j;m\rangle$. In the direct product space, consider the states $|3;2\rangle\otimes|1;1\rangle, |3;2\rangle\otimes|1;0\rangle, |3;2\rangle\otimes|1;-1\rangle$. These are eigenvalues of the direct product operator $T_3 \equiv T_3^{(3)} \otimes T_3^{(1)}$ with respective eigenvalues $3, 2, 1$:

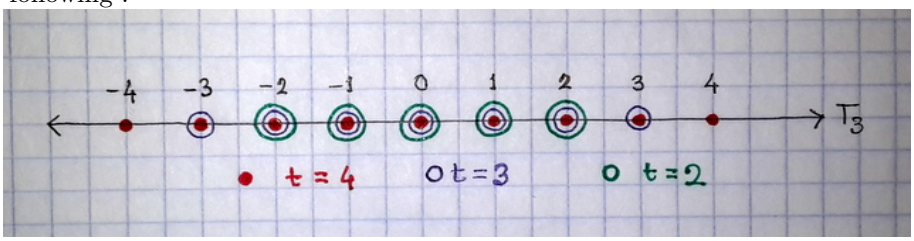
$$\begin{aligned} T_3\left(|3;2\rangle \otimes |1;1\rangle\right) &= 3\left(|3;2\rangle \otimes |1;1\rangle\right) \\ T_3\left(|3;2\rangle \otimes |1;0\rangle\right) &= 2\left(|3;2\rangle \otimes |1;1\rangle\right) \\ T_3\left(|3;2\rangle \otimes |1;-1\rangle\right) &= 1\left(|3;2\rangle \otimes |1;1\rangle\right) \end{aligned}$$

We can depict this in the weight diagram in the following manner. Make a copy, call it C_2, of the weight diagram of the $t = 1$ irrep and superimpose it on top of the weight diagram of the $t = 3$ irrep such that the center of C_2 (i.e., the 0 eigenvalue in C_2) lies on top of the weight of $|3;2\rangle$ (i.e., the 2 eigenvalue of the $t = 3$ irrep). Now do this for all 7 weights of the $t = 3$ irrep – superimpose a copy of the weight diagram of the $t = 1$ irrep on each of them. This gives us the weights of the states in the product representation.



Evidently, the product representation has states with degeneracy. In an irrep of $su\,(2)$, states are never degenerate, so we deduce that the product representation is reducible and we intend to find the dimensions of the irreps whose direct sum gives the product representation we have. The state with the highest T_3 eigenvalue has to belong to and be the highest state of the irrep with $t = 4$. This irrep has 9 states with weights $\pm4, \pm3, \pm2, \pm1, 0$. Suppose that the weights that have been marked with a single slash in the diagram above correspond to these states. Having enumerated these weights, we eliminate them from the diagram

and are left with 12 more states. Among these, the one with the highest T_3 eigenvalue has to belong to and be the highest state of the irrep with $t = 3$. Once we count the weights of this irrep (marked by double slashes), we are left with weights (marked with ticks) that have to belong to the irrep with $t = 2$. The weight diagram of the product representation, therefore, looks like the following :



Thus, the direct product of a 7 dimensional irrep and a 3 dimensional irrep of $su\,(2)$ is reducible and can be expressed as the direct sum of a 9 dimensional irrep, a 7 dimensional irrep, and a 5 dimensional irrep. This is the Clebsch-Gordan decomposition and is often written as

$$7 \otimes 3 = 9 \oplus 7 \oplus 5 \tag{144}$$

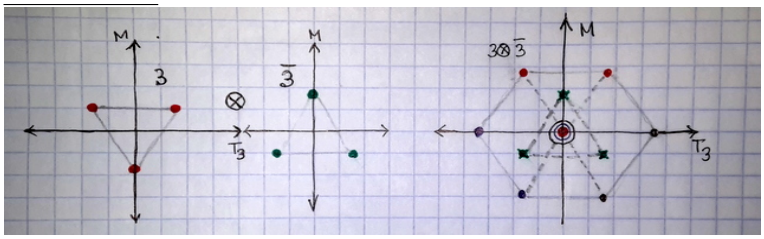
Notice that, when we copied the weight diagrams of the $t = 1$ irrep and pasted them on the weights of the $t = 3$ irrep, all we were doing was vector addition in one dimension. The T_3 eigenvalue of a product state $|j_1; m_1\rangle \otimes |j_2; m_2\rangle$ is the sum of the T_3 eigenvalues of $|j_1; m_1\rangle$ and $|j_2; m_2\rangle$. Think of eigenvalues of T_3 being one dimensional vectors. When you add two of them, by the rule of vector addition, you have to place the tail of one of them on the tip of the other. That is exactly what we have been doing!

This idea will help us a great deal when we perform the same exercise with $su\,(3)$ for which weight diagrams are two dimensional. There, too, T_3, M eigenvalues of product states $|\alpha, \vec{g}_1\rangle \otimes |\alpha, \vec{g}_2\rangle$ are sums of the T_3, M eigenvalues of the states $|\alpha, \vec{g}_1\rangle$ and $|\alpha, \vec{g}_2\rangle$. In other words, the weight $\vec{g}$ of the product state is simply $\vec{g}_1 + \vec{g}_2$ – vector addition! Therefore, in order to find the weights corresponding to a direct product representation of two irreps of $su\,(3)$, we shall make copies of one of them and superimpose the centers (or zeros) of the copies on top of each weight of the other irrep. Having thus found all the weights of the product representation, we shall take the highest weight and treat it as the highest weight state of one of the component irreps that build the product representation. Then we shall enumerate all the weights[27] that belong to this irrep and set them aside. Then we shall take the highest of the remaining weights and treat it as the highest weight state of another component irrep. We continue this process until all the weights of the product representation are exhausted.

Let us illustrate this by a few examples.

[27] If a weight $\vec{g}$ of that irrep has a two-fold degeneracy, say, then two $\vec{g}$'s from the set of product weights (in other words, two of the degenerate product states with weight $\vec{g}$) are included in that irrep and so on.

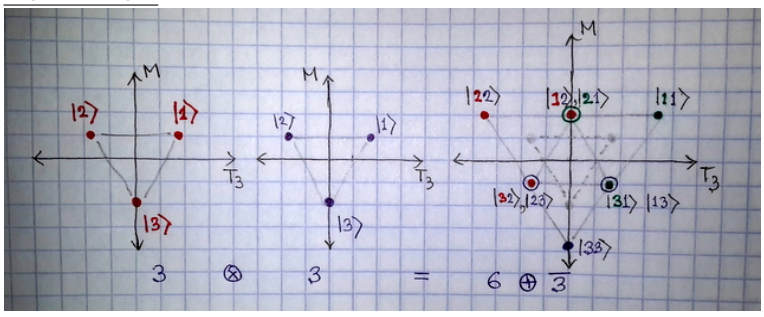
- $3 \otimes \bar{3} = 8 \oplus 1$



The weight diagram of the product representation has 6 weights on a hexagonal boundary and 3 states at the center $(0, 0)$. The 6 boundary weights and 2 of the $(0, 0)$ weights together form the weight diagram of the octet which is a component irrep of $3 \otimes \bar{3}$. The state corresponding to the remaining weight $(0, 0)$ forms, by itself, the scalar irrep or the singlet. FYI, let me make a connection with particle physics here. You shall probably better understand it when we will have a full-fledged discussion on elementary particles. Quarks belong to the 3 representation of $su\,(3)$ and antiquarks to the $\bar{3}$ representation. When we combine a quark and an antiquark, we get a product state that lives in the $3 \otimes \bar{3}$ representation. This product representation, as we have seen, has 9 states, 8 of which belong to the octet irrep and the remaining state belongs to the scalar irrep. In nature, we do find this happening. There are 8 particles, each consisting of a quark and an antiquark, that have almost the same mass and mix under $su\,(3)$ transformations like the states of the octet. There is also another particle, formed by a quark and an antiquark, that stays the same under $su\,(3)$ transformations. We shall learn the names of these particles later.

Now, where is mass in all this? Notice that we have so far assumed $su\,(3)$ to be a symmetry of nature. Because of this, the generators of $su\,(3)$ commute with the Hamiltonian H whose eigenvalues give nothing but mass! This implies that all the eigenstates of T_3, M are eigenstates of H with the same eigenvalue – hence the same mass.

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



The diagram explains it, I hope. The product states $|a\rangle \otimes |b\rangle$ are simply denoted by $|ab\rangle$, $a, b \in \{1, 2, 3\}$. Although $3 \otimes 3$ is mathematically as interesting (or not) as any other product representation, it doesn't appear

very often in physics, because we do not have bound states of two quarks. But we have bound states of three quarks. It would be a nice exercise for you to derive the following Clebsch-Gordan decomposition

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1 \tag{145}$$

An historical note : In nature, we do see manifestation of the 10 irrep (aka the decuplet) of $su\,(3)$. There are 10 particles (baryons), of almost similar masses, that mix under $su\,(3)$ transformations. But when Gell-Mann proposed $su\,(3)$ as a symmetry of nature, one of these particles was yet to be discovered. Gell-Mann predicted $su\,(3)$ being a symmetry, observed that 9 particles have very similar masses and hypothesized that they belong to an $su\,(3)$ decuplet. Naturally, he predicted the existence of the 10^{th} which was discovered later. This is the particle $\triangle_{++}$.
Another note : Had $su\,(3)$ been an exact symmetry of nature, the states belonging to a given irrep of $su\,(3)$ would have the same mass. Also, one would interpret, in keeping with standard principles of symmetry in quantum mechanics, these states being the states of the same particle. In that case, one would not say, e.g., that the states in the decuplet are those of 10 different particles. But $su\,(3)$ is not an exact symmetry of nature.

5 Week 5

5.1 Lecture 13 : February 1, 2016

Today we shall start by giving a quick review of $su\,(2)$ that you typically learn in non-relativistic quantum mechanics.

We shall borrow all the results from the $su\,(2)$ algebra of angular momentum. We learned that, for a given irrep, states are labeled by their J_3 eigenvalues. And different irreps are labeled by their j values; j gives the total angular momentum of the states : the J^2 eigenvalue of all the states of an irrep is $j\,(j+1)$. The ladder operators $J_\pm$ take you between states. Assuming that the usual inner product is defined, and that the J_3 eigenstates $|j;m\rangle$ are orthonormalized, we have

$$\begin{aligned} J_3|j;m\rangle &= m|j;m\rangle \\ J_\pm|j;m\rangle &= \sqrt{(j\mp m)\,(j\pm m+1)}|j;m\pm 1\rangle \end{aligned} \tag{146}$$

We derive the second equation the following way. We suppose that the factor is N_- , a complex number.

$$J_-|j;m\rangle = N_-|j;m-1\rangle$$

$$\Rightarrow |N_-|^2 = \langle j;m|J_+J_-|j;m\rangle = (j+m)\,(j-m+1)$$

This gives us $|N_-|$, not N_-. When we demand that

$$N_- = \sqrt{(j+m)\,(j-m+1)} \tag{147}$$

we are making a choice of phase. We could very well have chosen N_- to have an arbitrary phase factor. Let me elaborate it a little further, because this is an assumption (valid, but an assumption nonetheless) of extreme importance and one often takes it for granted. J_-, acting on $|j;m\rangle$, produces another eigenstate of J_3 with eigenvalue $(m-1)$. If $|\psi\rangle$ is such a state, then so is $e^{i\phi}|\psi\rangle$. The state for which $\phi = 0$ is being **defined** to be $|j;m-1\rangle$. From among the infinitely many eigenstates of J_3 with eigenvalue $(m-1)$, one for every choice of phase ϕ relative to the state $|j;m\rangle$, we are choosing to include that one in the eigenbasis which corresponds to $\phi = 0$, and we name it $|j;m-1\rangle$. The relative phase between two states is physically significant. That is why, when we are making these phase choices, we are specializing in choosing the basis states. Starting with the highest state $|j;j\rangle$ (which can have an arbitrary overall phase that is physically insignificant), we keep on making this choice so that the relative phases between two successive states is fixed. Note that, having fixed the relative phase between $|j;m\rangle$ and $|j;m-1\rangle$, and that between $|j;m-1\rangle$ and $|j;m-2\rangle$, the relative phase between $|j;m\rangle$ and $|j;m-2\rangle$ get automatically fixed. The upshot is that the choice of N_- being real positive fixes the relative phases between every pair of states in the eigenbasis we work with. Having defined the states in the basis, the next question then is, whether

N_+ in $J_+|j;m\rangle = N_+|j;m+1\rangle$ is real positive or has a phase. The commutation relations of the $su\,(2)$ algebra determines that for you. Fortunately for us, N_+ turns out to be real and positive as well. We could do it the other way round. We could choose N_+ to be real and positive and N_- would then become real and positive as a consequence. We provide an example to illustrate why the choice of N_+ also determines N_-. Take $j = \frac{1}{2}$, and let the two states be denoted by $|+\rangle, |-\rangle$. Suppose we choose N_+ to be real and positive :

$$J_+|-\rangle = \sqrt{\left(\frac{1}{2} - \left(-\frac{1}{2}\right)\right)\left(\frac{1}{2} + \left(-\frac{1}{2}\right) + 1\right)}|+\rangle = |+\rangle$$

Now,

$$J_-|+\rangle = (|-\rangle\langle -| + |+\rangle\langle +|)\, J_-|+\rangle = |-\rangle\langle -|J_-|+\rangle + |+\rangle\langle +|J_-|+\rangle$$

The second term in the sum is 0, because $\langle +|J_- = (J_+|+\rangle)^\dagger = 0$. Therefore,

$$J_-|+\rangle = (\langle -|J_-|+\rangle)\,|-\rangle \equiv \alpha|-\rangle$$

The number α is N_- and can be determined as follows :

$$\alpha = \langle -|J_-|+\rangle = (\langle +|J_+|-\rangle)^\dagger = (\langle +|+\rangle)^\dagger = 1$$

Now we can claim that we understand (146) fully.

The above analysis is for an $su\,(2)$ multiplet and we shall extend the ideas to $su\,(3)$ now. On an $su\,(3)$ super-multiplet, states lying on a horizontal line (fixed M eigenvalue) belong to some $su\,(2)$ multiplet. That is, they span an invariant subspace of the subalgebra span $\{T_+, T_-, T_3\} \cong su\,(2)$. Therefore, we can exercise the freedom of phase choice discussed above and set the relative phase of the states $T_\pm|\alpha,(t_3,m)\rangle$ and $|\alpha,(t_3 \pm 1, m)\rangle$ to be 0. This makes the coefficients $N_\pm$ real and positive as shown above. What makes $su\,(3)$ different from $su\,(2)$ is the fact that $su\,(3)$ has two more subalgebras isomorphic to $su\,(2)$. States lying on lines parallel to $\vec{u}$ (or, $\vec{v}$) span invariant subspaces of the subalgebra span $\left\{U_+, U_-, \vec{u}.\vec{G}\right\}$ (or, span $\left\{V_+, V_-, \vec{v}.\vec{G}\right\}$). Therefore, one state in the $su\,(3)$ super-multiplet is part of three different $su\,(2)$ multiplets. Keeping this in mind, let us consider the (p,q) irrep, start with the highest weight state and label it $|1\rangle$ – this is the first state in our ordered eigenbasis. Then the next state in the ordered basis $|2\rangle$ is chosen so that N_- is real and positive in $T_-|1\rangle = N_-|2\rangle$. Continuing this way, we label all the states $\{|1\rangle, ..., |p+1\rangle\}$ with the highest M eigenvalue – states on the top-most horizontal line. Unless $q = 0$, $U_-|\alpha, \vec{g}_{max}\rangle$ is the highest state (maximum T_3 eigenvalue) in the next horizontal level. We choose $U_-|\alpha, \vec{g}_{max}\rangle \equiv U_-|1\rangle$ to be the next state in the ordered basis, $|p+2\rangle$, modulo a real and positive normalization factor. That is, R_- is real and positive in $U_-|1\rangle = R_-|p+2\rangle$. We can do this because we have a freedom to choose the relative phase between the two states $|1\rangle$ and $|p+2\rangle$. Then we label all the states on this horizontal level according to the set rules.

The upshot is this : by exercising the freedom to choose the relative phases of the states in the eigenbasis, we can make the coefficients $R_\pm, N_\pm$ real and positive. That is, the T and U multiplets are chosen to follow the simplest $su\,(2)$ "ladder" (or "shift") relations. Once we have done that, the relative phases between any pair of states in the eigenbasis are fixed. We cannot simplify the ladder relations corresponding to $V_\pm$ any more. That is, if two states $|f\rangle, |g\rangle$ in the chosen eigenbasis are such that $V_-|f\rangle = S_-|g\rangle$, then there is no guarantee that S_- will be real (let alone positive). The relative phase between $|f\rangle$ and $|g\rangle$ has already been fixed to 0 by choosing to have real and positive values for $R_\pm, N_\pm$ in the U and T multiplets and there is no reason why $V_-|f\rangle$ and $|g\rangle$ should also have 0 relative phase. These phases will be dictated by the commutation relations of $su\,(3)$. So, we need to pay close attention to the actions of the $V_\pm$ shift operators on the chosen ordered eigenbasis.

We have already said that we shall label the states in one horizontal level of an $su\,(3)$ weight diagram by their j, m values (we shall use j and t interchangeably). For the irreps we will work with, we shall need the following :

- $j = \frac{1}{2}$

$$\begin{array}{c} J_+|\frac{1}{2};\frac{1}{2}\rangle = 0 \\ J_+|\frac{1}{2};-\frac{1}{2}\rangle = 1.|\frac{1}{2};\frac{1}{2}\rangle \leftarrow \text{a phase choice between } |\frac{1}{2};\frac{1}{2}\rangle, |\frac{1}{2};-\frac{1}{2}\rangle \\ J_-|\frac{1}{2};\frac{1}{2}\rangle = 1.|\frac{1}{2};-\frac{1}{2}\rangle \leftarrow \text{no phase choice, consequence of } su\,(2) \\ J_-|\frac{1}{2};-\frac{1}{2}\rangle = 0 \end{array} \tag{148}$$

- $j = 1$

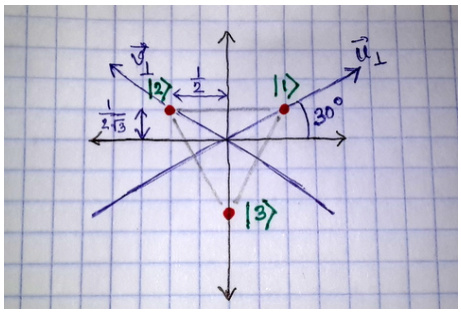
$$\begin{array}{c} J_-|1;1\rangle = \sqrt{2}|1;0\rangle \leftarrow \text{phase choice} \\ J_-|1;0\rangle = \sqrt{2}|1;-1\rangle \leftarrow \text{phase choice} \\ J_-|1;-1\rangle = 0 \\ J_+|1;1\rangle = 0 \\ J_+|1;0\rangle = \sqrt{2}|1;1\rangle \leftarrow \text{not a phase choice} \\ J_+|1;-1\rangle = \sqrt{2}|1;0\rangle \leftarrow \text{not a phase choice} \end{array} \tag{149}$$

Now we shall do the same exercise for the 3, $\bar{3}$ and 8 irreps of $su\,(3)$. The plan is to start from the highest weight state and then make the shift operators act on it. Just like in $su\,(2)$, the basis states will be so defined such that the results of the action of shift operators are as simple as possible – in other words, relative phases will be chosen to be 0 as far as possible. In what follows, I shall explicitly indicate whenever an equation has a phase choice involved in it. I shall also indicate which ones do not involve phase choices and simply follow from the choices already made and the $su\,(3)$ algebra.

- The 3 irrep :
  To begin with, let us agree that we shall include orthonormal states in the eigenbasis. But also notice that, if $|\psi_1\rangle$ and $|\psi_2\rangle$ are orthonormal, then so are $e^{i\theta_1}|\psi_1\rangle$ and $e^{i\theta_2}|\psi_2\rangle$. Orthonormality of a set of states does not uniquely determine everything about the states. Their relative phases

could still be arbitrarily chosen.

$$\begin{array}{cccc} T_3: & |1\rangle \mapsto \frac{1}{2}|1\rangle & M: & |1\rangle \mapsto \frac{1}{2\sqrt{3}}|1\rangle \\ & |2\rangle \mapsto -\frac{1}{2}|2\rangle & & |2\rangle \mapsto \frac{1}{2\sqrt{3}}|2\rangle \\ & |3\rangle \mapsto 0 & & |3\rangle \mapsto -\frac{1}{\sqrt{3}}|3\rangle \end{array} \tag{150}$$



$$\begin{array}{cccc} T_-: & |1\rangle \mapsto |2\rangle \leftarrow \text{phase choice} & T_+: & |1\rangle \mapsto 0 \\ & |2\rangle \mapsto 0 & & |2\rangle \mapsto |1\rangle \leftarrow \text{not a choice} \\ & |3\rangle \mapsto 0 & & |3\rangle \mapsto 0 \\ U_-: & |1\rangle \mapsto 0 & U_+: & |1\rangle \mapsto 0 \\ & |2\rangle \mapsto |3\rangle \leftarrow \text{phase choice} & & |2\rangle \mapsto 0 \\ & |3\rangle \mapsto 0 & & |3\rangle \mapsto |2\rangle \leftarrow \text{not a choice} \\ V_-: & |1\rangle \mapsto 0 & V_+: & |1\rangle \mapsto |3\rangle \leftarrow \text{not a choice} \\ & |2\rangle \mapsto 0 & & |2\rangle \mapsto 0 \\ & |3\rangle \mapsto |1\rangle \leftarrow \text{not a choice} & & |3\rangle \mapsto 0 \end{array} \tag{151}$$

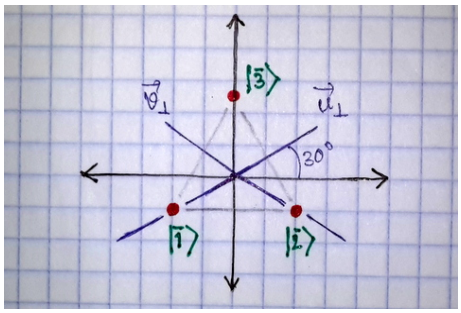
It stands to reason that we have only two relative phases that we can choose since there are three states in this irrep. The rest of the operator actions, ones indicated with the phrase "not a choice", are insisted on by the algebra. Let us show that for one of those relations. Since the relative phase between $|2\rangle$ and $|3\rangle$ is fixed by the equation $U_-|2\rangle = |3\rangle$

$$V_-|3\rangle = V_-U_-|2\rangle = (U_-V_- + T_+)\,|2\rangle = T_+|2\rangle = |1\rangle$$

Just as we promised! Having fixed the relative phases, we have the freedom to change the overall phases of all the states by the same amount.

- <u>The $\bar{3}$ irrep</u> :

$$\begin{array}{cccc} T_3: & |\bar{1}\rangle \mapsto -\frac{1}{2}|\bar{1}\rangle & M: & |\bar{1}\rangle \mapsto -\frac{1}{2\sqrt{3}}|\bar{1}\rangle \\ & |\bar{2}\rangle \mapsto \frac{1}{2}|\bar{2}\rangle & & |\bar{2}\rangle \mapsto -\frac{1}{2\sqrt{3}}|\bar{2}\rangle \\ & |\bar{3}\rangle \mapsto 0 & & |\bar{3}\rangle \mapsto \frac{1}{\sqrt{3}}|\bar{3}\rangle \end{array} \tag{152}$$



$$
\begin{array}{lcc|lcc}
T_- : & |\bar{1}\rangle \mapsto 0 & & T_+ : & |\bar{1}\rangle \mapsto |\bar{2}\rangle \leftarrow \text{not a choice} \\
& |\bar{2}\rangle \mapsto |\bar{1}\rangle \leftarrow \text{phase choice} & & & |\bar{2}\rangle \mapsto 0 \\
& |\bar{3}\rangle \mapsto 0 & & & |\bar{3}\rangle \mapsto 0 \\
U_- : & |\bar{1}\rangle \mapsto 0 & & U_+ : & |\bar{1}\rangle \mapsto 0 \\
& |\bar{2}\rangle \mapsto 0 & & & |\bar{2}\rangle \mapsto |\bar{3}\rangle \leftarrow \text{not a choice} \\
& |\bar{3}\rangle \mapsto |\bar{2}\rangle \leftarrow \text{phase choice} & & & |\bar{3}\rangle \mapsto 0 \\
V_- : & |\bar{1}\rangle \mapsto -|\bar{3}\rangle \leftarrow \text{not a choice} & & V_+ : & |\bar{1}\rangle \mapsto 0 \\
& |\bar{2}\rangle \mapsto 0 & & & |\bar{2}\rangle \mapsto 0 \\
& |\bar{3}\rangle \mapsto 0 & & & |\bar{3}\rangle \mapsto |\bar{1}\rangle \leftarrow \text{not a choice}
\end{array}
\tag{153}
$$

Notice here that $V_-|\bar{1}\rangle \mapsto -|\bar{3}\rangle$, we have a negative sign and we cannot avoid it. Let's see how the negative sign comes about.

$$V_-|\bar{1}\rangle = V_-T_-|\bar{2}\rangle = (T_-V_- - U_+)\,|\bar{2}\rangle = -U_+|\bar{2}\rangle = -|\bar{3}\rangle$$

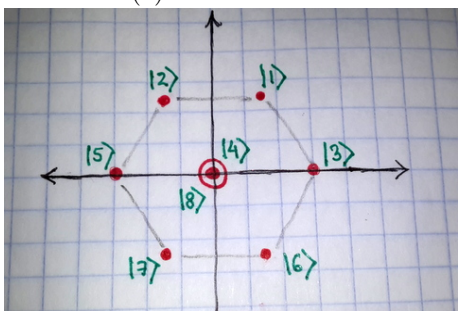
A note : In physics, sometimes one makes a different phase choice between $|\bar{1}\rangle$ and $T_-|\bar{2}\rangle$. One chooses $T_-|\bar{2}\rangle = -|\bar{1}\rangle$. This is done because, as a consequence, a final result appears nicer. But this is a choice of basis after all and the physics does not depend on it. In both the irreps discussed above, we have three states and can choose two relative phases as per our convenience. There is only so much freedom we have. Once we use it up, we don't have any freedom left any longer.

- <u>The 8 irrep</u> :

$$
\begin{array}{lcc|lcc}
T_3 : & |1\rangle \mapsto \frac{1}{2}|1\rangle & & M : & |1\rangle \mapsto \frac{\sqrt{3}}{2}|1\rangle \\
& |2\rangle \mapsto -\frac{1}{2}|2\rangle & & & |2\rangle \mapsto \frac{\sqrt{3}}{2}|2\rangle \\
& |3\rangle \mapsto |3\rangle & & & |3\rangle \mapsto 0 \\
& |4\rangle \mapsto 0 & & & |4\rangle \mapsto 0 \\
& |5\rangle \mapsto -|5\rangle & & & |5\rangle \mapsto 0 \\
& |6\rangle \mapsto \frac{1}{2}|6\rangle & & & |6\rangle \mapsto -\frac{\sqrt{3}}{2}|6\rangle \\
& |7\rangle \mapsto -\frac{1}{2}|7\rangle & & & |7\rangle \mapsto -\frac{\sqrt{3}}{2}|7\rangle \\
& |8\rangle \mapsto 0 & & & |8\rangle \mapsto 0
\end{array}
\tag{154}
$$

Notice that, although we said we shall be labeling states on a horizontal level by their t and m values, we have not done that. We could do it,

of course, but the labels we are using here are simpler. However, we are going to use the $su\,(2)$ "ladder" relations in that we will not compute the factors $N_\pm$ afresh and will borrow the results from what we have already done in $su\,(2)$.



$$
\begin{array}{lcc}
T_- : & |1\rangle \mapsto |2\rangle \leftarrow \text{phase choice} & U_- : \quad |1\rangle \mapsto |3\rangle \leftarrow \text{phase choice} \\
& |2\rangle \mapsto 0 & |2\rangle \mapsto \alpha|4\rangle + \beta|8\rangle \\
& |3\rangle \mapsto \sqrt{2}|4\rangle \leftarrow \text{phase choice} & |3\rangle \mapsto 0 \\
& |4\rangle \mapsto \sqrt{2}|5\rangle \leftarrow \text{phase choice} & |4\rangle \mapsto \gamma|6\rangle \\
& |5\rangle \mapsto 0 & |5\rangle \mapsto |7\rangle \leftarrow \text{phase choice} \\
& |6\rangle \mapsto |7\rangle \leftarrow \text{phase choice} & |6\rangle \mapsto 0 \\
& |7\rangle \mapsto 0 & |7\rangle \mapsto 0 \\
& |8\rangle \mapsto 0 & |8\rangle \mapsto \delta|6\rangle
\end{array}
\tag{155}
$$

Notice again, that there are infinitely many states with the weight $(0,0)$. But only two of them are linearly independent (since $(0,0)$ is doubly degenerate) and we can take one of them to be $T_-|3\rangle$. We have decided to label that state as $|4\rangle$, modulo the factor $\sqrt{2}$. That is, $|3\rangle, |4\rangle, |5\rangle$ form an $su\,(2)$ triplet (corresponding to $T, T_\pm$). This justifies the factors of $\sqrt{2}$ in their $T_\pm$ ladder relations. The other state with weight $(0,0)$ is of course linearly independent of $|4\rangle$, and we choose to call it $|8\rangle$. However, since these two states belong to the same degenerate eigenbasis, therefore $|8\rangle$ is not uniquely determined by simply demanding that it be linearly independent of $|4\rangle$. Because, for any such $|8\rangle$, $|8\rangle' = |8\rangle + \xi|4\rangle, \xi \neq 0$, is also linearly independent of $|4\rangle$ and has weight $(0,0)$. We shall be able to uniquely specify $|8\rangle$ when we specify the factor β in

$$U_-|2\rangle = \alpha|4\rangle + \beta|8\rangle, \tag{156}$$

because, in this equation, the states $|2\rangle$ and $|4\rangle$ have already been chosen and U_- is also a known operator. An immediate observation is the following : if $|\psi\rangle$ is the next normalized state in the U-triplet, then $U_-|2\rangle = \sqrt{2}|\psi\rangle$. Comparing this with (156):

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(\alpha|4\rangle + \beta|8\rangle\right), \tag{157}$$

the next normalized state in the U-triplet that has $|2\rangle$ as its highest weight state. The third member of this triplet has to be $|6\rangle$. Since neither $|4\rangle$ nor

$|8\rangle$ belong to the U-triplet starting with $|2\rangle$, therefore γ and δ in (155) are $\sqrt{2}$. They are not even guaranteed to be real. Notice that, we have made 6 phase choices in (155), and there are 8 states. Hence, we have leftover freedom to choose one more relative phase. We could use this to choose either γ or δ to be real. But we shall not do that. We shall use the phase freedom to fix the relative phase of $|2\rangle$ and $|8\rangle$ so that β is real. Therefore, γ and δ are complex numbers as of now.

5.2 Lecture 14 : February 3, 2016

Today we shall calculate the numbers $\alpha, \beta, \gamma, \delta$, exploiting our phase freedom to the fullest extent so as to make as many of them real and positive as possible. Recall that the states $|1\rangle, |2\rangle, ..., |8\rangle$ are orthonormal and that we have used up 6 of the 7 phase choices. Equation (156) implies

$$\alpha = \langle 4|U_-|2\rangle = \frac{1}{\sqrt{2}}\langle 3|T_+U_-|2\rangle \ \left[\because T_-|3\rangle = \sqrt{2}|4\rangle\right]$$

$$= \frac{1}{\sqrt{2}}\langle 3|U_-T_+|2\rangle = \frac{1}{\sqrt{2}}\langle 3|U_-|1\rangle = \frac{1}{\sqrt{2}}\langle 3|3\rangle$$

$$\therefore \alpha = \frac{1}{\sqrt{2}} \tag{158}$$

Hence, α is real and positive as a consequence of the algebra (and the phase choices already made). So we still have the freedom to choose one relative phase. Now, equation (157) from last class told us that $\frac{1}{\sqrt{2}}(\alpha|4\rangle + \beta|8\rangle)$ has a norm of 1. Hence,

$$|\alpha|^2 + |\beta|^2 = 2 \Rightarrow |\beta|^2 = \frac{3}{2}$$

This implies that $\beta = \sqrt{\frac{3}{2}}e^{i\phi}$. We shall use our last phase freedom to absorb the phase factor $e^{i\phi}$ in the definition of $|8\rangle$. Hence,

$$\beta = \sqrt{\frac{3}{2}} \tag{159}$$

We have used up all the freedom we had and will have to accept whatever values for γ and δ the algebra throws at us.

$$U_-|4\rangle = U_-T_+\left(\frac{1}{\sqrt{2}}|5\rangle\right) = T_+U_-\left(\frac{1}{\sqrt{2}}|5\rangle\right) = \frac{1}{\sqrt{2}}T_+|7\rangle = \frac{1}{\sqrt{2}}|6\rangle$$

$$\therefore \gamma = \frac{1}{\sqrt{2}} \tag{160}$$

Luckily, γ turns out to be real and positive. To compute δ, note that the first two equations in (149) imply that $J_-J_-|1;1\rangle = 2|1;-1\rangle$. For the U-triplet, $|2\rangle$ corresponds to $|1;1\rangle$ and $|6\rangle$ corresponds to $|1;-1\rangle$. Therefore,

$$U_-\left(U_-|2\rangle\right) = U_-\left(\sqrt{\frac{1}{2}}|4\rangle + \sqrt{\frac{3}{2}}|8\rangle\right)$$

$$\Rightarrow 2|6\rangle = \sqrt{\frac{1}{2}}U_-|4\rangle + \sqrt{\frac{3}{2}}U_-|8\rangle = \frac{1}{2}|6\rangle + \sqrt{\frac{3}{2}}\delta|6\rangle$$

$$\therefore \delta = \sqrt{\frac{3}{2}} \tag{161}$$

δ also turns out to be real and positive! Therefore, we have fixed all the relative phases in the octet and figured out the following :

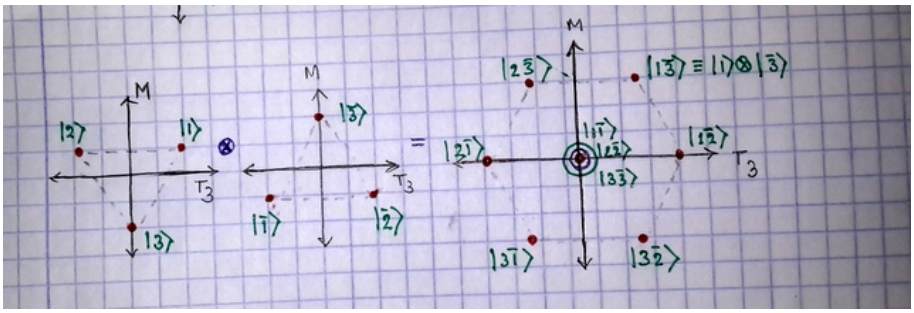
$$\begin{array}{lcc}
 & T_- : & U_- : \\
 & |1\rangle \mapsto |2\rangle \leftarrow \text{phase choice} & |1\rangle \mapsto |3\rangle \leftarrow \text{phase choice} \\
 & |2\rangle \mapsto 0 & |2\rangle \mapsto \sqrt{\frac{1}{2}}|4\rangle + \sqrt{\frac{3}{2}}|8\rangle \leftarrow \text{phase choice} \\
 & |3\rangle \mapsto \sqrt{2}|4\rangle \leftarrow \text{phase choice} & |3\rangle \mapsto 0 \\
 & |4\rangle \mapsto \sqrt{2}|5\rangle \leftarrow \text{phase choice} & |4\rangle \mapsto \sqrt{\frac{1}{2}}|6\rangle \\
 & |5\rangle \mapsto 0 & |5\rangle \mapsto |7\rangle \leftarrow \text{phase choice} \\
 & |6\rangle \mapsto |7\rangle \leftarrow \text{phase choice} & |6\rangle \mapsto 0 \\
 & |7\rangle \mapsto 0 & |7\rangle \mapsto 0 \\
 & |8\rangle \mapsto 0 & |8\rangle \mapsto \sqrt{\frac{3}{2}}|6\rangle
\end{array} \tag{162}$$

You should compute the actions of the rest of the shift operators on the chosen basis states.

Now that we have chosen a convenient basis for the octet, let us remind ourselves that, for any vector space, basis is not unique. We encountered the following Clebsch-Gordan decomposition earlier :

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

The product representation $3 \otimes \bar{3}$ has the octet as one of its component irrep. In the product representation, product states form the natural basis. The natural question to ask now is whether 8 of these 9 natural basis product states have one-to-one correspondence with the basis we have chosen for the octet. We shall find soon that in general the answer is no. The product states lying on the edges of the weight diagram of $3 \otimes \bar{3}$ have a one-to-one correspondence with the "edge-states" of our chosen basis for the octet, but not the product states with weight $(0,0)$. Certain linear combinations of these "center-states" have one-to-one correspondences with $|4\rangle$ and $|8\rangle$, as we show below. We shall use the crisp notation $|ab\rangle, |a\bar{b}\rangle$ for the product states $|a\rangle \otimes |b\rangle, |a\rangle \otimes |\bar{b}\rangle$ etc. We choose eight linear combinations of the product states that will serve as basis vectors of the 8-dimensional invariant subspace carrying the octet irrep, and label these states $|1\rangle_8, \ldots, |8\rangle_8$.



1. The weight of the product state $|1\bar{3}\rangle$ is equal to the highest weight of the octet.
2. The highest weight of the octet is non-degenerate.
3. We have the liberty to arbitrarily choose the overall phase of the first state $|1\rangle_8$ of the octet.

Because of these reasons, we can choose

$$|1\rangle_8 \equiv |1\bar{3}\rangle$$

We choose a trivial overall phase for convenience, of course. Now, in the octet, we made use of phase freedom to demand that $|2\rangle_8 \equiv T_-|1\rangle_8$. This gives us

$$|2\rangle_8 = T_-|1\bar{3}\rangle = \left(T_-^{(3)}|1\rangle\right) \otimes |\bar{3}\rangle + |1\rangle \otimes \left(T_-^{(\bar{3})}|\bar{3}\rangle\right) = |2\bar{3}\rangle$$

Similarly, $|3\rangle_8 = U_-|1\rangle_8$ gives $|3\rangle_8 = U_-|1\bar{3}\rangle = |1\bar{2}\rangle$. Continuing this way, one finds that the edge states of the octet are exactly the product states sitting on the edge of the product representation. We only need to find which linear combinations to choose for $|4\rangle_8$ and $|8\rangle_8$. That is also easy to find. $|4\rangle_8 = \frac{1}{\sqrt{2}}T_-|3\rangle_8$ implies that

$$|4\rangle_8 = \frac{1}{\sqrt{2}}T_-|1\bar{2}\rangle = \frac{|1\bar{1}\rangle + |2\bar{2}\rangle}{\sqrt{2}}.$$

And, $U_-|2\rangle_8 = \sqrt{\frac{1}{2}}|4\rangle_8 + \sqrt{\frac{3}{2}}|8\rangle_8$ implies that $|8\rangle_8 = \frac{1}{\sqrt{3}}\left(\sqrt{2}U_-|2\rangle_8 - |4\rangle_8\right)$.

$$\therefore |8\rangle_8 = \frac{1}{\sqrt{3}}\left[\sqrt{2}\left(|3\bar{3}\rangle + |2\bar{2}\rangle\right) - \frac{|1\bar{1}\rangle + |2\bar{2}\rangle}{\sqrt{2}}\right] = \frac{2|3\bar{3}\rangle + |2\bar{2}\rangle - |1\bar{1}\rangle}{\sqrt{6}}$$

Therefore,

$$\begin{array}{cc} |1\rangle_8 = |1\bar{3}\rangle & |5\rangle_8 = |2\bar{1}\rangle \\ |2\rangle_8 = |2\bar{3}\rangle & |6\rangle_8 = |3\bar{2}\rangle \\ |3\rangle_8 = |1\bar{2}\rangle & |7\rangle_8 = |3\bar{1}\rangle \\ |4\rangle_8 = \frac{|1\bar{1}\rangle + |2\bar{2}\rangle}{\sqrt{2}} & |8\rangle_8 = \frac{2|3\bar{3}\rangle + |2\bar{2}\rangle - |1\bar{1}\rangle}{\sqrt{6}} \end{array} \tag{163}$$

Let us notice in passing that the states we have obtained are all normalized, which is just as it should be! So, we have found eight linear combinations, namely $\left\{|1\bar{3}\rangle, |2\bar{3}\rangle, |1\bar{2}\rangle, \frac{|1\bar{1}\rangle+|2\bar{2}\rangle}{\sqrt{2}}, |2\bar{1}\rangle, |3\bar{2}\rangle, |3\bar{1}\rangle, \frac{2|3\bar{3}\rangle+|2\bar{2}\rangle-|1\bar{1}\rangle}{\sqrt{6}}\right\}$, of the natural basis states $\{|1\bar{1}\rangle, |1\bar{2}\rangle, |1\bar{3}\rangle, |2\bar{1}\rangle, |2\bar{2}\rangle, |2\bar{3}\rangle, |3\bar{1}\rangle, |3\bar{2}\rangle, |3\bar{3}\rangle\}$ of the product space that span an 8-dimensional invariant subspace. This invariant subspace carries the octet irrep. All we are left with is to find another linear combination of the basis states $|1\bar{1}\rangle, |2\bar{2}\rangle, |3\bar{3}\rangle$ that belongs to the singlet, or the 1-dimensional irrep of $su\,(3)$. We choose it to be orthogonal to the states of the octet. Let us denote the singlet state as $|1\rangle_1$, and let

$$|1\rangle_1 = \alpha|1\bar{1}\rangle + \beta|2\bar{2}\rangle + \gamma|3\bar{3}\rangle \tag{164}$$

It is easy to see that the singlet state cannot be a linear combination of anything other than the states $|1\bar{1}\rangle, |2\bar{2}\rangle, |3\bar{3}\rangle$, because, otherwise, $|1\rangle_1$ would not have the desired weight $(0,0)$. As a consequence, $|1\rangle_1$ is automatically orthogonal to $|1\rangle_8, |2\rangle_8, |3\rangle_8, |5\rangle_8, |6\rangle_8, |7\rangle_8$. Now, demanding orthogonality with $|4\rangle_8$ and $|8\rangle_8$ yields

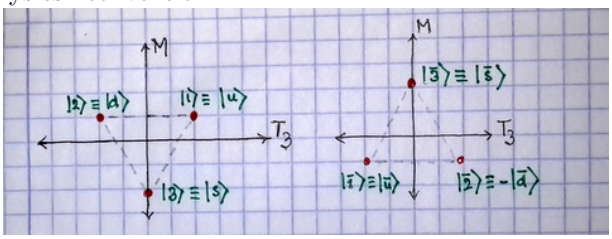
$$\begin{aligned} {}_8\langle 4|1\rangle_1 &= \tfrac{\alpha+\beta}{\sqrt{2}} = 0 \Rightarrow \alpha = -\beta \\ {}_8\langle 8|1\rangle_1 &= \tfrac{2\gamma+\beta-\alpha}{\sqrt{6}} = 0 \Rightarrow \gamma = \alpha \end{aligned}$$

Hence, we have $\alpha = -\beta = \gamma$. Normalization requires $\alpha = \frac{1}{\sqrt{3}}e^{i\theta}$. **Choosing** the simplest phase,

$$|1\rangle_1 = \frac{|1\bar{1}\rangle - |2\bar{2}\rangle + |3\bar{3}\rangle}{\sqrt{3}} \tag{165}$$

This somewhat long exercise shows that $3\otimes\bar{3} = 8\oplus 1$, and the basis of the product representation $3\otimes\bar{3}$ in which this Clebsch-Gordan decomposition is explicitly seen is $\left\{|1\bar{3}\rangle, |2\bar{3}\rangle, |1\bar{2}\rangle, \frac{|1\bar{1}\rangle+|2\bar{2}\rangle}{\sqrt{2}}, |2\bar{1}\rangle, |3\bar{2}\rangle, |3\bar{1}\rangle, \frac{2|3\bar{3}\rangle+|2\bar{2}\rangle-|1\bar{1}\rangle}{\sqrt{6}}, \frac{|1\bar{1}\rangle-|2\bar{2}\rangle+|3\bar{3}\rangle}{\sqrt{3}}\right\}$, i.e., $\{|1\rangle_8, |2\rangle_8, |3\rangle_8, |4\rangle_8, |5\rangle_8, |6\rangle_8, |7\rangle_8|8\rangle_8, |1\rangle_1\}$. Check for yourself that the singlet state indeed spans an invariant subspace. Computing the Clebsch-Gordan series for the decomposition $3\otimes 3\otimes 3 = 10\oplus 8\oplus 8\oplus 1$ would be a nice homework exercise.

Before wrapping up today, let me mention the consequence of the following "Physics" convention :



Here, the phase convention is that $U_-|\bar{s}\rangle = -|\bar{d}\rangle$. That is, what we have been calling $|\bar{2}\rangle$ so far is being called $-|\bar{2}\rangle$ in this convention. The rest of it is

the same as before. With this convention,

$$\begin{array}{c} |1\rangle_1 = \frac{|u\bar{u}\rangle+|d\bar{d}\rangle+|s\bar{s}\rangle}{\sqrt{3}} \\ |8\rangle_8 = \frac{2|s\bar{s}\rangle-|d\bar{d}\rangle-|u\bar{u}\rangle}{\sqrt{6}} \end{array} \tag{166}$$

All the coefficients in the linear combination for $|1\rangle_1$ become positive. And the numerators of the coefficients in the linear combination for $|8\rangle_8$ become $+2, -1, -1$, which is a little bit better[28] than $+2, +1, -1$, what we had before. These coefficients being easier to remember is the reason why physicists often adopt this phase convention. You might object by saying that this destroys the niceness of the coefficients elsewhere, for instance, $|3\rangle_8 = -|u\bar{d}\rangle$, $|4\rangle_8 = \frac{|1\bar{1}\rangle-|2\bar{2}\rangle}{\sqrt{2}}$ etc. Then again, you can't have everything in life! Choose for yourself which convention is easier for you to remember and stick to it.

5.3 Lecture 15 : February 4, 2016

Symmetry in Quantum Mechanics :

Consider a transformation (unitary) U_R of the Hilbert space $\mathcal{H}$ of a system where R corresponds to a change of the reference system. Under this transformation, the states $|\psi\rangle \in \mathcal{H}$ change : $|\psi\rangle \to |\psi^{'}\rangle = U_R|\psi\rangle$: while the operators A remain unchanged. As a result, the expectation values of states change in the following way : $\langle A\rangle_{|\psi\rangle} \equiv \langle\psi|A|\psi\rangle \to \langle\psi|U_R^{-1}AU_R|\psi\rangle$. This way of looking at the transformation in which states change is called the active view. The passive view is an alternative (and equivalent) way of effecting the same transformation in which the states are kept fixed and the operators are changed in the following manner : $A \to A^{'} = U_R^{-1}AU_R$. Both the approaches bring about the same change in the expectation values which are the physically important numbers carrying all the information about the states. Hence, these two views are equivalent.
Now, such a transformation U_R is said to be a symmetry of the system if the Hamiltonian H of the system does not change even according to the passive view, i.e., if

$$U_R^{-1}HU_R = H \iff [H, U_R] = 0 \tag{167}$$

Let us consider a system with rotational symmetry. Therefore, its Hamiltonian H must commute with U_R corresponding to all rotations R of the coordinate axes. Therefore, R can be a rotation about any axis through any angle. $U : R \mapsto U_R$ is a representation of the rotation group on $\mathcal{H}$. The necessary and sufficient condition for H to commute with all possible U_R is that H has to commute with the generators $J_i, i \in \{1, 2, 3\}$ of the rotation matrices U_R:

$$[H, J_i] = 0;\ i \in \{1, 2, 3\} \tag{168}$$

[28] Better in the sense that this is perhaps easier to remember.

Even when we do not know the intricate details of the interactions in the system and, therefore, are ignorant of the exact form of the H, we have at least this information about H based on its symmetry.

The system of our interest is a particle which has isospin as its symmetry. Isospin is probably a new term, but we have already seen it. Isospin operators are nothing but the first three generators T_1, T_2, T_3 of $su(3)$. The isospin algebra is $su(2)$ which is the same as the rotation algebra $so(3)$. Hence, $[H, J_i] = 0$ holds for our system as well. Since H commutes with all the isospin operators, in particular with J_3, we can choose simultaneous eigenvectors of H and J_3 as our basis. For a particle with the total isospin j, the eigenbasis is $\{|j;j\rangle, |j;j-1\rangle, \ldots, |j;-j\rangle\}$.

Now, the Hamiltonian H is the generator of time translation. For very small time t, the time translation operator is

$$U = \mathbb{I} - iHt \tag{169}$$

The following inner product is called the transition amplitude for a particle to transition (evolve in time) from the state $|j;m\rangle$ to the state $|j^{'};m^{'}\rangle$:

$$\langle j^{'};m^{'}|H|j;m\rangle \tag{170}$$

If this vanishes, then transition does not occur between the states $|j;m\rangle$ and $|j^{'};m^{'}\rangle$. Otherwise, there is non-zero probability of the transition happening. We are suddenly interested in the transition amplitudes because these are also the matrix elements of the Hamiltonian in the chosen eigenbasis. We shall now show that, for the rotationally symmetric Hamiltonian, the transition amplitude is zero whenever $j \neq j^{'}$ **and** $m \neq m^{'}$. Not only that, for $j = j^{'}, m = m^{'}$, the non-zero value of the transition amplitude depends only on the value of j and not on the value of m. This result is wonderful because it is a consequence of the symmetry of H, not of its explicit form which is unknown to us at the moment.

Since $[H, J_i] = 0$, therefore $\left[H, J^2\right] = 0$. This implies

$$\langle j^{'};m^{'}|\left[H, J^2\right]|j;m\rangle = 0 \Rightarrow \left(j\left(j+1\right) - j^{'}\left(j^{'}+1\right)\right)\langle j^{'};m^{'}|H|j;m\rangle = 0$$

Hence, $j \neq j^{'} \Rightarrow \langle j^{'};m^{'}|H|j;m\rangle = 0$; only non-zero matrix elements are those with $j = j^{'}$. Also, $[H, J_3] = 0$ implies

$$\langle j^{'};m^{'}|\left[H, J_3\right]|j;m\rangle = 0 \Rightarrow \left(m - m^{'}\right)\langle j^{'};m^{'}|H|j;m\rangle = 0$$

Thus, $m \neq m^{'} \Rightarrow \langle j^{'};m^{'}|H|j;m\rangle = 0$. Therefore, we can write

$$\langle j^{'};m^{'}|H|j;m\rangle = \delta_{jj^{'}}\delta_{mm^{'}}\langle j;m|H|j;m\rangle \tag{171}$$

Finally, $[H, J_+] = 0$ implies

$$\langle j;m|\left[H, J_+\right]|j;m-1\rangle = 0$$

$$\Rightarrow \sqrt{(j-m+1)\,(j+m)}\langle j;m|H|j;m\rangle = \sqrt{(j+m)\,(j-m+1)}\langle j;m-1|H|j;m-1\rangle$$

$$\Rightarrow \langle j;m|H|j;m\rangle = \langle j;m-1|H|j;m-1\rangle \tag{172}$$

Therefore, not only the Hamiltonian is diagonal, its matrix elements being non-zero only for the diagonal entries $\langle j;m|H|j;m\rangle$, these values are the same for all m for a given j. So we can label these values by j:

$$\langle j;m|H|j;m\rangle \equiv H_j \;\forall m \in \{j, j-1, \ldots, -j\} \tag{173}$$

With this notation,

$$\langle j^{'};m^{'}|H|j;m\rangle = \delta_{jj'}\delta_{mm'}H_j \tag{174}$$

Equation (174) is nothing but the **Wigner-Eckart theorem**[29] for scalar operators.

A little digression :

The classification of operators as scalars, vectors etc. or, most generally as tensors of different ranks (or types), is always on the basis of some transformations of the coordinates (or change of reference frames). Suppose we are concerned with rotations of Cartesian axes in three dimensional Euclidean space. Things that do not change under this set of transformations are called scalars under rotations – example : Euclidean distances between points. The set of three Cartesian coordinates of points do change under rotations in a definite fashion. This set of coordinates of a point is by definition called a vector. Any other set of three numbers that also change under rotations in precisely the same way as a vector is also called a vector. So, the family of vectors is defined in terms of the way in which the members transform under rotations. Similarly, there are other families of objects, tensors of several types, members of which transform differently from vectors, but all tensors of one type (or family) transform similarly under rotations. Take note that whenever I am saying that something transforms this way or that way, I have been constantly following that up with the phrase "under rotations". Very often for brevity we do not repeat this phrase when we are talking about only one kind of coordinate transformation, here rotations. But when we have many types of coordinate transformations to worry about, say rotations, Lorentz transformations etc., and we are saying that something is a vector or a scalar or a tensor of a certain type, we have to mention "under which transformation" it is so. If an object A is a vector under rotations (transforms similarly to the coordinates (x, y, z) of a point), it will most likely not be a vector under Lorentz transformation (will not transform similarly to the coordinates (t, x, y, z) of an event). Having told you the importance of mentioning which type of transformation one is dealing with, I shall give in, like everybody else, to the laziness of not mentioning it.
In the context of quantum mechanics, a three component operator V_j is a vector

[29] Brush up on the general Wigner-Eckart theorem for tensor operators of arbitrary ranks.

operator under rotation if it satisfies the following commutation relation with angular momentum operators :

$$[J_i, V_j] = i\epsilon_{ijk} V_k \tag{175}$$

I will not elaborate on why a vector operator is defined this way. Look it up, or better think about it. Similarly, a scalar operator is one which commutes with all the J_is. The rotationally symmetric Hamiltonian H is a scalar operator.

We shall now take up the case of the neutron and the proton. These two particles have nearly the same mass. The major difference between them is that neutron is electrically neutral whereas proton has an electric charge of one unit. In fact, the slight difference in the masses of these two particles comes from electromagnetism. In a world without electromagnetism, there would be virtually no difference between a neutron and a proton. This prompts us to hypothesize that neutron and proton are not two different particles but two states of the same particle. We shall assume that, in our imaginary world without electromagnetism, isospin is a symmetry and proton and neutron form an isospin doublet. They belong to $j = \frac{1}{2}$, with the neutron and the proton having up spin and down spin respectively. The states are denoted by

$$|n\rangle \equiv |p, \uparrow\rangle, \quad |p\rangle \equiv |p, \downarrow\rangle \tag{176}$$

The neutron and the proton are nucleons and the dominant force acting on them is the strong force. The full Hamiltonian of interactions between elementary particles have three pieces (in the approximation of weak gravitational field)

$$H = H_{strong} + H_{weak} + H_{em} \tag{177}$$

where the terms correspond, in order, to the strong interaction, the weak interaction and the electromagnetic interaction. We are assuming that isospin is a symmetry of the strong interaction. Hence, under transformations generated by isospin operators, $H_{strong} \to H_{strong}$. These transformations act not on the configuration space of the particles, obviously, and are said to act on the **internal space** of the system. Such symmetries of the internal space are called **internal symmetries**. Now, isospin operators generate $SU(2)$ transformations. So, according to our hypothesis, the symmetry transformations of the internal space are $SU(2)$ transformations. We could take an alternative approach and derive the symmetry transformations of the internal space. It goes like the following. Suppose that the internal space is spanned by $|n\rangle$ and $|p\rangle$. Consider a symmetry transformation (linear)

$$\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \to \begin{pmatrix} |p'\rangle \\ |n'\rangle \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \tag{178}$$

This being a symmetry really means that norms do not change under such a transformation. That implies $U \equiv \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is unitary : $U^\dagger U = \mathbb{I}$. Hence,

$|\det U|^2 = 1 \Rightarrow \det U = e^{i\theta}$, for some $\theta \in [0, 2\pi)$. Every 2×2 unitary matrix U can be expressed in the form $U = e^{i\theta/2}U^{'}$ where $\det U^{'} = 1, \therefore U^{'} \in SU(2)$. In the language of group theory, $U(2) \cong SU(2) \times U(1)$ (this is a direct product). Thus we have

$$\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \rightarrow \begin{pmatrix} |p^{'}\rangle \\ |n^{'}\rangle \end{pmatrix} = e^{i\theta/2} \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} \text{ where } |\alpha|^2 + |\beta|^2 = 1 \qquad (179)$$

Recall that $\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix}$ and $\begin{pmatrix} |p^{'}\rangle \\ |n^{'}\rangle \end{pmatrix}$ represent the same physical state. One observer (A) represents the state with the vector $\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix}$, and the U-transformed observer (B) represents it with $\begin{pmatrix} |p^{'}\rangle \\ |n^{'}\rangle \end{pmatrix}$. If observer B chooses to represent all the states with an overall phase of $e^{i\theta/2}$, then this factor will be absorbed in the state $\begin{pmatrix} |p^{'}\rangle \\ |n^{'}\rangle \end{pmatrix}$. Then, one can live with just an $SU(2)$ transformation matrix $\begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}$.

We studied in quantum mechanics that the total wave function of a particle with spin has two factors : $\psi = \psi_{space}\psi_{spin}$. The total wave function should be antisymmetric for fermions. Now that we have isospin in the picture, $\psi = \psi_{space}\psi_{spin}\psi_{isospin}$ should be antisymmetrized. Deuteron is a fermion that has zero orbital angular momentum, so its ψ_{space} is symmetric. It has spin 1, so its ψ_{spin} is symmetric (triplet). Therefore, it has to have an antisymmetric $\psi_{isospin}$. It actually turns out that it is an isospin 0 particle, hence a singlet.

6 Week 6

6.1 Lecture 16 : February 8, 2016

We saw that the T operators of $su(3)$ are called isospin operators. They are often denoted by I instead of T. We could just as well work with the $su(2)$ algebra of the U (or the V) operators and call that isospin. But it is customary to work with the T operators. We also saw that the matrix elements $\langle j; m|H|j^{'}; m^{'}\rangle$ of a rotationally invariant Hamiltonian are non-zero only when $j = j^{'}, m = m^{'}$. Since these matrix elements are also the transitions amplitudes, therefore the result $\langle j^{'}; m^{'}|H|j; m\rangle = \delta_{jj'}\delta_{mm'}H_j$ is essentially a statement of isospin conservation.

Consider the process in which a pion and a deuteron collide to produce two nucleons. Pion, denoted Π, is a particle with $I = 1$, so it has three states Π^+, Π^0, Π^- corresponding to the I_3 values $1, 0, -1$ respectively. Deuteron is an isospin singlet, and has only one state d. Nucleons proton and neutron are two states of an $I = \frac{1}{2}$ particle. Charge conservation[30] tells us that these are the allowed channels :

$$\begin{array}{c} \Pi^+ d \to pp \\ \Pi^0 d \to pn \text{ or } np \\ \Pi^- d \to nn \end{array} \tag{180}$$

In a scattering of a pion and a deuteron, all these three channels are probable outcomes. Using isospin conservation we shall figure out the relative scattering cross-sections of these channels.

In order to calculate the ratios of the scattering cross-sections of the three channels, we need to know the formula for the scattering cross-section of one channel. This is an exercise in quantum mechanics and I shall just quote the result here without deriving it. For a scattering process with the Hamiltonian H, the scattering cross section σ of a channel where the initial and the final states are respectively $|i\rangle$ and $|f\rangle$is given by

$$\sigma \sim |\langle f|H|i\rangle|^2\rho \tag{181}$$

where ρ is a phase factor that depends on the masses of the particles involved in the channel. For the channels of our interest, the phase factors are going to be nearly the same since all the pions have nearly the same mass and so do the nucleons. The difference between the scattering cross-sections will arise from the factor $|\langle f|H|i\rangle|^2$. Note that we do not have any clue about the explicit form of the Hamiltonian except for the fact that it is rotationally invariant. The magic is that rotational invariance is all you need to know about H in order to find the ratios of the cross-sections. Let's see how.

[30] You might object here. We had earlier said that we would switch off electromagnetism so that isospin would be a symmetry. Now we are using charge conservation to see which channels are allowed! Your objection is justified. However, since we know that electromagnetism is there (and therefore isospin is not an exact symmetry), all physical processes must obey charge conservation. Therefore, to ease the present dilemma, let's say that we are "interested" (for no good reason) in the channels mentioned above and take it from there.

The initial states of the three channels belong to the $3 \otimes 1$ product representation of (isospin) $su(2)$. And the final states belong to the $2 \otimes 2 = 3 \oplus 1$ representation. We would like to express the initial states $|\Pi^{0,\pm} d\rangle$ in terms of the natural basis states of the product representation. It is easy to see that

$$\begin{array}{c} |\Pi^+ d\rangle = |1;1\rangle \otimes |0;0\rangle = |1;1\rangle \\ |\Pi^0 d\rangle = |1;0\rangle \\ |\Pi^- d\rangle = |1;-1\rangle \end{array} \tag{182}$$

The states on the right hand side should really be written in the format $|j_1, j_2; j, m\rangle$. For instance, for the first state : $|j_1; m_1\rangle \otimes |j_2; m_2\rangle \equiv |1;1\rangle \otimes |0;0\rangle = |1,0;1,0\rangle \equiv |j_1, j_2; j, m\rangle$. Since all the states on the right hand side have the same $j_1$and j_2 labels, so we drop them and only explicitly write the j and m labels. We do the same for the final states with two nucleons. This is just like adding two $\frac{1}{2}$ spins. The Clebsch-Gordan series for this simple exercise gives the three triplet states

$$\begin{array}{c} |1;1\rangle = |pp\rangle \\ T_-|1;1\rangle = \sqrt{2}|1;0\rangle = T_-|pp\rangle = (|np\rangle + |pn\rangle) \Rightarrow |1;0\rangle = \frac{|np\rangle + |pn\rangle}{\sqrt{2}} \\ |1;-1\rangle = |nn\rangle \end{array} \tag{183}$$

and the singlet state

$$|0;0\rangle = \frac{|pn\rangle - |np\rangle}{\sqrt{2}} \tag{184}$$

Having described the incoming and outgoing states, let us compute

$$\langle pp|H|\Pi^+ d\rangle = \langle 1;1|H|1;1\rangle \equiv H_1 \tag{185}$$

Using isospin conservation we have concluded that $\langle 1;1|H|1;1\rangle$ is non-zero, depends only on the j value which is 1 and is therefore labeled as H_1. Notice that the bra $\langle 1;1|$ and the ket $|1;1\rangle$ in the above equation do not correspond to the same state (because they represent two very different pairs of particles!). However, the isospin labels (or eigenvalues) of these two states are the same and isospin is all we are concerned about right now. So, in that sense, they are the same state. Next,

$$\langle nn|H|\Pi^- d\rangle = \langle 1;-1|H|1;-1\rangle = \langle 1;1|H|1;1\rangle \equiv H_1 \tag{186}$$

Here we have used $\langle 1;-1|H|1;-1\rangle = \langle 1;1|H|1;1\rangle$, the fact that the matrix element $\langle j;m|H|j;m\rangle$ is the same for all m for a given value of j. Finally,

$$\begin{array}{l} \langle \Pi^0 d|H|pn\rangle = \langle 1;0|H\left(\frac{|1;0\rangle + |0;0\rangle}{\sqrt{2}}\right) = \frac{\langle 1;0|H|1;0\rangle}{\sqrt{2}} = \frac{H_1}{\sqrt{2}} \\ \langle \Pi^0 d|H|np\rangle = \langle 1;0|H\left(\frac{|1;0\rangle - |0;0\rangle}{\sqrt{2}}\right) = \frac{\langle 1;0|H|1;0\rangle}{\sqrt{2}} = \frac{H_1}{\sqrt{2}} \end{array} \tag{187}$$

Since both the processes are possible in this channel, we add the probability amplitudes. So, the net probability amplitude is

$$\frac{H_1}{\sqrt{2}} + \frac{H_1}{\sqrt{2}} = \sqrt{2} H_1 \tag{188}$$

Hence,

$$\sigma\left(\Pi^+ d \to pp\right) : \sigma\left(\Pi^0 d \to pn\right) : \sigma\left(\Pi^- d \to nn\right) = |H_1|^2 : 2|H_1|^2 : |H_1|^2 = 1 : 2 : 1 \tag{189}$$

If isospin symmetry is assumed to be exact, which is what we did above, then the masses of the three pions would be identical, masses of the nucleons would be identical, and ρ factors would exactly cancel from the ratio of cross-sections. And we get the exact ratio given in equation (189). We can experimentally verify this pretty easily. Hence, our proposition of isospin being a symmetry is falsifiable. Note that we have only been able to calculate the ratios of the scattering cross-sections using symmetry. The exact values of the cross-sections cannot be calculated unless one knows the details of the dynamics – the Hamiltonian.

Let us master the technique by repeating the exercise at another process. This one would require slightly more effort calculation-wise. Consider the scattering of a pion (Π) and a nucleon (N) that produces a pion and a nucleon : $\Pi N \to \Pi N$. Again, the channels allowed by charge conservation are the ones we should be interested in. There are many such channels. Let us focus on two of them and calculate the ratio of their scattering cross-sections.

$$\begin{array}{c} \Pi^+ p \to \Pi^+ p \\ \Pi^- p \to \Pi^- p \text{ or } \Pi^0 n \end{array} \tag{190}$$

The states on the left belong to the $3 \otimes 2$ product representation of isospin $su\,(2)$, and[31] $3 \otimes 2 = 4 \oplus 2$. Now, $|\Pi^+ p\rangle$ is the state with a total J_3 eigenvalue of $\frac{3}{2}$, and must be the highest weight state of the 4 dimensional component irrep. Hence, suppressing the j_1 and j_2 labels in the coupled basis state $|j_1, j_2; j, m\rangle$, $|\Pi^+ p\rangle = |\frac{3}{2}; \frac{3}{2}\rangle (= |1, \frac{1}{2}; \frac{3}{2}, \frac{3}{2}\rangle)$. The other two product states of interest to us are $|\Pi^- p\rangle$ and $|\Pi^0 n\rangle$. Both these states have a J_3 eigenvalue of $-\frac{1}{2}$. There are two states in the direct sum decomposition $4 \oplus 2$ with J_3 eigenvalue of $\frac{1}{2}$, namely $|\frac{3}{2}; -\frac{1}{2}\rangle$ and $|\frac{1}{2}; -\frac{1}{2}\rangle$. Therefore, we must have $|\frac{3}{2}; -\frac{1}{2}\rangle = \alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle$, with normalization imposed by $|\alpha|^2 + |\beta|^2 = 1$. $|\frac{1}{2}; -\frac{1}{2}\rangle$ is also a linear combination of the states $|\Pi^- p\rangle$ and $|\Pi^0 n\rangle$. Orthogonality of $|\frac{3}{2}; -\frac{1}{2}\rangle$ and $|\frac{1}{2}; -\frac{1}{2}\rangle$ demands that $|\frac{1}{2}; -\frac{1}{2}\rangle = \beta|\Pi^- p\rangle - \alpha|\Pi^0 n\rangle$. We could easily obtain the coefficients α and β by making the shift/ladder operators acting on the various states following the prescription of the Clebsch-Gordan series. There is another way of computing these coefficients. We shall demonstrate it now. Take $|\frac{3}{2}; -\frac{1}{2}\rangle = \alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle$ and make I^2 act on both sides. The action on the L.H.S gives

$$I^2|\frac{3}{2}; -\frac{1}{2}\rangle = \frac{3}{2}\left(\frac{3}{2}+1\right)|\frac{3}{2}; -\frac{1}{2}\rangle = \frac{15}{4}|\frac{3}{2}; -\frac{1}{2}\rangle = \frac{15}{4}\left(\alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle\right) \tag{191}$$

To compute the action on the right hand, notice that $I^2 = \left(I^{(1)}\right)^2 + \left(I^{(2)}\right)^2 +$

[31] In other words, adding the angular momenta $j_1 = 1$ and $j_2 = \frac{1}{2}$ gives us one total angular momentum of $j = j_1 + j_2 = \frac{3}{2}$ and another total angular momentum of $j = j_1 + j_2 - 1 = \frac{1}{2} = |j_1 - j_2|$.

$2I_3^{(1)}I_3^{(2)} + I_+^{(1)}I_-^{(2)} + I_-^{(1)}I_+^{(2)}$. Therefore,

$$I^2|\Pi^- p\rangle = \left(\left(I^{(1)}\right)^2 + \left(I^{(2)}\right)^2 + 2I_3^{(1)}I_3^{(2)} + I_+^{(1)}I_-^{(2)} + I_-^{(1)}I_+^{(2)}\right)|\Pi^- p\rangle$$

$$= \left(2 + \frac{3}{4} + 2\,(-1)\left(\frac{1}{2}\right)\right)|\Pi^- p\rangle + \sqrt{2}|\Pi^0 n\rangle = \frac{7}{4}|\Pi^- p\rangle + \sqrt{2}|\Pi^0 n\rangle$$

$$I^2|\Pi^0 n\rangle = \left(\left(I^{(1)}\right)^2 + \left(I^{(2)}\right)^2 + 2I_3^{(1)}I_3^{(2)} + I_+^{(1)}I_-^{(2)} + I_-^{(1)}I_+^{(2)}\right)|\Pi^0 n\rangle$$

$$= \left(2 + \frac{3}{4} + 2\,(0)\left(-\frac{1}{2}\right)\right)|\Pi^0 n\rangle + \sqrt{2}|\Pi^- p\rangle = \frac{11}{4}|\Pi^0 n\rangle + \sqrt{2}|\Pi^- p\rangle$$

$$\therefore I^2\left(\alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle\right) = \frac{7\alpha + 4\sqrt{2}\beta}{4}|\Pi^- p\rangle + \frac{4\sqrt{2}\alpha + 11\beta}{4}|\Pi^0 n\rangle \qquad (192)$$

Comparing equations (191) and (192),

$$\begin{array}{c} 15\alpha = 7\alpha + 4\sqrt{2}\beta \Rightarrow \sqrt{2}\alpha = \beta \\ 15\beta = 4\sqrt{2}\alpha + 11\beta \Rightarrow \sqrt{2}\alpha = \beta \end{array} \qquad (193)$$

Are you surprised that both equations give the same solution? Of course, had they given inconsistent solutions then we would have been in a soup. So, the solution above being consistent is a relief. But weren't we supposed to get the values of α and β out of these equations, whereas we get the ratio $\frac{\alpha}{\beta}$ instead? The answer is simple. All we demanded is that $\alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle$ be equal to $|\frac{3}{2}; -\frac{1}{2}\rangle$, thus having a an I^2 eigenvalue of $\frac{15}{4}$ and an I_3 eigenvalue of $-\frac{1}{2}$. The fact that it has I_3 eigenvalue of $-\frac{1}{2}$ is automatically satisfied for all α, β, and does not tell us anything about them. So we moved on to the I^2 eigenvalue and that gave us the equations above. However, if $\alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle$ has an I^2 eigenvalue of $\frac{15}{4}$, then so does $k\left(\alpha|\Pi^- p\rangle + \beta|\Pi^0 n\rangle\right)$ for any arbitrary complex number k. So, the I^2 eigenvalue also does not tell us everything about the coefficients and gives only their ratio. Now we make use of the normalization condition $|\alpha|^2 + |\beta|^2 = 1$ and the phase freedoms to choose simplest phases for α, β. We get : $\alpha = \frac{1}{\sqrt{3}}; \beta = \sqrt{\frac{2}{3}}$.

$$\begin{array}{c} |\frac{3}{2}; \frac{3}{2}\rangle = |\Pi^+ p\rangle \\ |\frac{3}{2}; -\frac{1}{2}\rangle = \frac{1}{\sqrt{3}}|\Pi^- p\rangle + \sqrt{\frac{2}{3}}|\Pi^0 n\rangle \\ |\frac{1}{2}; -\frac{1}{2}\rangle = \sqrt{\frac{2}{3}}|\Pi^- p\rangle - \frac{1}{\sqrt{3}}|\Pi^0 n\rangle \end{array} \qquad (194)$$

The method we employed here to find out α, β works as efficiently as the Clebsch-Gordan method. Although the full blown Clebsch-Gordan computation gives you the entire set of states, it is often too much work if you do not need the Clebsch-Gordan (aka CG) coefficients of all the states. Our approach takes less time to calculate the CG coefficients for a few states of interest. You might think that the full blown Clebsch-Gordan computation for the above example

would not have taken more time than our trick did. That is true. But for larger dimensions of product representations, the economy of time gained by this trick becomes significant.

Now we need to invert these to express $|\Pi^+ p\rangle, |\Pi^- p\rangle, |\Pi^0 n\rangle$ as linear combinations of $|\frac{3}{2};\frac{3}{2}\rangle, |\frac{3}{2};-\frac{1}{2}\rangle, |\frac{1}{2};-\frac{1}{2}\rangle$. This is also easy.

$$\begin{aligned} |\Pi^+ p\rangle &= |\tfrac{3}{2};\tfrac{3}{2}\rangle \\ |\Pi^- p\rangle &= \tfrac{1}{\sqrt{3}}|\tfrac{3}{2};-\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}|\tfrac{1}{2};-\tfrac{1}{2}\rangle \\ |\Pi^0 n\rangle &= \sqrt{\tfrac{2}{3}}|\tfrac{3}{2};-\tfrac{1}{2}\rangle - \tfrac{1}{\sqrt{3}}|\tfrac{1}{2};-\tfrac{1}{2}\rangle \end{aligned} \tag{195}$$

Finally we compute the transition amplitudes for the channels $|i\rangle \to |f\rangle$, denoted $a\left(|i\rangle \to |f\rangle\right)$. These are the matrix elements $\langle f|H|i\rangle$. We have, $\langle \Pi^+ p|H|\Pi^+ p\rangle = \langle \frac{3}{2};\frac{3}{2}|H|\frac{3}{2};\frac{3}{2}\rangle \equiv H_{3/2}$. Next,

$$\begin{aligned} \langle \Pi^- p|H|\Pi^- p\rangle &= \left(\frac{1}{\sqrt{3}}\langle\frac{3}{2};-\frac{1}{2}| + \sqrt{\frac{2}{3}}\langle\frac{1}{2};-\frac{1}{2}|\right) H \left(\frac{1}{\sqrt{3}}|\frac{3}{2};-\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}|\frac{1}{2};-\frac{1}{2}\rangle\right) \\ &= \frac{1}{3}H_{3/2} + \frac{2}{3}H_{1/2} \end{aligned}$$

Finally,

$$\begin{aligned} \langle \Pi^0 n|H|\Pi^- p\rangle &= \left(\sqrt{\frac{2}{3}}\langle\frac{3}{2};-\frac{1}{2}| - \frac{1}{\sqrt{3}}\langle\frac{1}{2};-\frac{1}{2}|\right) H \left(\frac{1}{\sqrt{3}}|\frac{3}{2};-\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}|\frac{1}{2};-\frac{1}{2}\rangle\right) \\ &= \frac{\sqrt{2}}{3}H_{3/2} - \frac{\sqrt{2}}{3}H_{1/2} \end{aligned}$$

Hence,

$$\begin{aligned} a\left(\Pi^+ p \to \Pi^+ p\right) &= H_{3/2} \\ a\left(\Pi^- p \to \Pi^- p\right) &= \tfrac{1}{3}H_{3/2} + \tfrac{2}{3}H_{1/2} \\ a\left(\Pi^- p \to \Pi^0 n\right) &= \tfrac{\sqrt{2}}{3}H_{3/2} - \tfrac{\sqrt{2}}{3}H_{1/2} \end{aligned} \tag{196}$$

6.2 Lecture 17 : February 10, 2016

Let's take off from where we were in the last class. It seems that all we need to compute is the matrix elements such as $\langle \Pi^- p|H|\Pi^- p\rangle$. Why then did we go through all the trouble of expressing the states $|\Pi N\rangle$ as linear combinations of eigenstates of the total isospin operator I^2? In other words, why did we compute the CG coefficients at all instead of just putting in the states $|\Pi N\rangle$ for themselves? The answer is simple. We do not know anything about the Hamiltonian H other than the fact that it is isospin symmetric and, consequently, is diagonal

in the so called coupled basis, $\left\{|i^{(1)}, i^{(2)}; i, i_3\rangle\right\}$, which is the simultaneous eigenbasis of $\left(I^{(1)}\right)^2$, $\left(I^{(2)}\right)^2$, I^2 and I_3. The states $|\Pi N\rangle$ (which form the so called uncoupled basis) are simultaneous eigenstates of $\left(I^{(1)}\right)^2$, $I_3^{(1)}$, $\left(I^{(2)}\right)^2$,$I_3^{(2)}$ and not of the total angular momentum operator I^2. In this basis, therefore, H is not diagonal. In order to exploit the diagonal structure of H in the coupled basis, we had to express the states $|\Pi N\rangle$ in their Clebsch-Gordan series expansion. In a quantum field theoretic approach where the Hamiltonian is known and is expressed in terms of the creation and annihilation field operators of the particles involved, we would not need to carry out the CG expansions of the $|\Pi N\rangle$ states. In that case we would be able to compute the individual cross-sections, not just their ratios, since the dynamics would be completely known. There is also a physical significance of the fact that H is not diagonal in the $|\Pi N\rangle$ basis. Had it been so, then there could only have been elastic scatterings of pions and nucleons since the transition amplitudes for the inelastic channels would have vanished. H is not diagonal in this basis precisely because of the interaction terms in the Lagrangian of the pion fields and the nucleon fields. Rather, since we (experimentalists) know that pions and nucleons do scatter inelastically, H must contain such interaction terms.

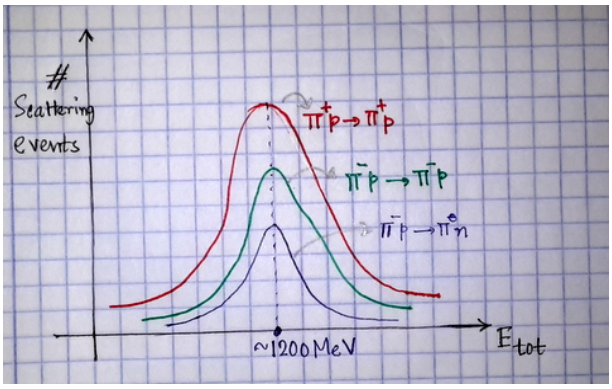
Now, equation (196) gives the scattering amplitudes. We need to take the squares of the moduli of the amplitudes to compute the cross-sections, denoted $\sigma\left(|i\rangle \to |f\rangle\right)$.

$$\begin{aligned} \sigma\left(\Pi^+ p \to \Pi^+ p\right) &= |H_{3/2}|^2 \\ \sigma\left(\Pi^- p \to \Pi^- p\right) &= \tfrac{|H_{3/2}|^2}{9}|1 + \tfrac{2H_{1/2}}{H_{3/2}}|^2 \\ \sigma\left(\Pi^- p \to \Pi^0 n\right) &= \tfrac{2|H_{3/2}|^2}{9}|1 - \tfrac{H_{1/2}}{H_{3/2}}|^2 \end{aligned} \tag{197}$$

Therefore,

$$\sigma\left(\Pi^+ p \to \Pi^+ p\right) : \sigma\left(\Pi^- p \to \Pi^- p\right) : \sigma\left(\Pi^- p \to \Pi^0 n\right) = 1 : \frac{|1 + \frac{2H_{1/2}}{H_{3/2}}|^2}{9} : \frac{2|1 - \frac{H_{1/2}}{H_{3/2}}|^2}{9} \tag{198}$$

This ratio can be experimentally verified like the previous one only if we know the theoretical values of $H_{1/2}$ and $H_{3/2}$. These values can be computed using QFT and experimental plots can be matched with the theoretical predictions. We can perform the experiment of scattering a nucleon and a pion at various energies, meaning that the total energy E_{tot} in the CM frame of the incoming particles may be varied. An experimentalist plots the scattering cross-sections of the three channels at all these energies to get one σ versus E_{tot} curve for each channel and checks whether the theoretically predicted ratio holds true at each value of E_{tot}. Such an experimental plot schematically looks like this :



All the three channels have unmistakable peaks at $E_{tot} \sim 1200$ MeV. The heights of peaks are almost 60 times the heights of the flatter part of the curves. This peak indicates that at this particular value of the total energy in the CM frame, $H_{3/2} \gg H_{1/2}$. This should not happen in the $\Pi^+ p \to \Pi^+ p$ channel since it seems to be an elastic scattering. But you cannot deny the experimental results. That means that $\Pi^+ p \to \Pi^+ p$ must not be a simple elastic collision. This scattering has an intermediate step in which a very short lived-particle, later identified as the particle $\triangle^{++}$. This is the experiment that led to the discovery of $\triangle^{++}$. Experimentally, its track shows that it hardly moves, confirming that it is indeed a short-lived particle. Its fast production and consequent decay is a signature of the robustness of the strong interaction through which it is created and annihilated.

At this point, for the sake of giving you a flavor of the history of the development of particle physics, let me quote a formula known as the Gell-Mann - Nishijima formula :

$$Q = I_3 + \frac{Y}{2} \tag{199}$$

Here Q is the electric charge of an elementary particle (say A), I_3 is its I_3-eigenvalue, and $\frac{Y}{2}$ is $\langle Q \rangle$, the average of the electric charges of all other elementary particles belonging to the same isospin multiplet as that of A.

$$\frac{Y}{2} \equiv \langle Q \rangle \tag{200}$$

Therefore, all particles belonging to the same isospin multiplet have the same Y value by virtue of the definition of Y. The Gell-Mann - Nishijima formula relates the electric charge of a particle with its isospin value. This was an empirical formula that was arrived at by simply observing the charges of the particles in different isospin multiplets. Today we can prove the result using sophisticated QFT calculations. One could easily define $Y \equiv \langle Q \rangle$, without the factor of $\frac{1}{2}$. But the usual definition is preferred because that results in integer values of Y. We shall take in all this as mere facts for now, without much justification, keeping in mind that all this can be derived rigorously using QFT.

Now, we have seen that existence of $\triangle^{++}$ was predicted to explain the peak in the σ vs. E_{tot} curve for $\Pi^+ p \to \Pi^+ p$. Charge conservation demands that

$\triangle^{++}$ have an electric charge of 2 units, hence the superscript $++$. Similarly, existence of the particle $\triangle^0$ was predicted to explain the peak in the curve for $\Pi^- p \to \Pi^- p$. This way, existence of four $\triangle$ particles were predicted with $\triangle^{++}$ having the highest amount of positive charge. These are : $\triangle^{++}, \triangle^+, \triangle^0, \triangle^-$. These four particles form an isospin multiplet among themselves. The following table gives the Y values of certain isospin multiplets :

$\triangle^{++}$	$\triangle^+$	$\triangle^0$	$\triangle^-$	$Y = 1$
Π^+		Π^0	Π^-	$Y = 0$
p		n		$Y = 1$

(201)

The particles in this table are all **hadrons**. Hadron is a name given to particles that take part in strong interaction. Heavy hadrons such as p, n, $\triangle$'s are called **baryons**. Hadrons with weight in the medium range are called mesons, e.g. $\Pi^{0,\pm}$. Notice in the table that the two multiplets with $Y = 1$ are composed of baryons having half-integral isospin, while that with $Y = 0$ has mesons with integral isospin. So people thought that Y is nothing but B, the **baryon** number, which takes the value 1 for baryons and the value 0 otherwise. So, the original Gell-Mann - Nishijima formula was written as $Q = I_3 + \frac{B}{2}$. Later people found the much heavier **strange particles**[32], and this simple interpretation $Y = B$ did not suffice anymore. Thus, the formula kept getting modified over time.

6.3 Lecture 18 : February 11, 2016

Today we shall give a very brief overview of scattering processes in quantum mechanics. You should already be familiar with the topic. Our purpose is just to brush up on the results we would be using later on.

[32]Strange particles are much heavier, so it takes more energy to find them. It took much time to improve the experimental setups to operate on such high energy scales as that needed to observe strange particles.

7 Week 7

7.1 Lecture 19 : February 22, 2016

Discussion of midsem question paper

7.2 Lecture 20 : February 24, 2016

Continuing discussion of scattering up to Breit-Wigner.

We shall now start a discussion on Poincare symmetry and the effect of Poincare transformations on the states of a Hilbert space. A Poincare transformation (Λ, a) is a transformation of the coordinates $(t, \vec{x})^\mu \equiv x^\mu$

$$x^\mu \to x^{'\mu} = \Lambda^\mu_\nu x^\nu + a^\mu \tag{202}$$

where Λ is a Lorentz transformation matrix satisfying $\Lambda^T \eta \Lambda = \eta$, η being the Minkowski metric $\eta = \text{diag}\,(1, -1, -1, -1)$, and a^μ's are components of a translation 4-vector. The set of all Poincare transformations, denoted $\mathscr{P} = \left\{(\Lambda, a) : \Lambda^T \eta \Lambda = \eta\right\}$ is a group under the group composition rule

$$(\Lambda, a)\left(\bar{\Lambda}, \bar{a}\right) = \left(\Lambda\bar{\Lambda}, \Lambda\bar{a} + a\right) \tag{203}$$

The two important subgroups of $\mathscr{P}$ are $\mathscr{L} = \left\{(\Lambda, 0) : \Lambda^T \eta \Lambda = \eta\right\}$, the Lorentz subgroup, and $\mathscr{T} = \{(\mathbb{I}, a)\}$, the translation subgroup, where $\mathbb{I}$ is the identity matrix. You can check that these are subgroups by verifying $(\Lambda, 0)\left(\bar{\Lambda}, 0\right) = \left(\Lambda\bar{\Lambda}, 0\right)$ and $(\mathbb{I}, a)\,(\mathbb{I}, \bar{a}) = (\mathbb{I}, a + \bar{a})$. Clearly, these subgroups are isomorphic to the Lorentz group and the translation group respectively. You might wonder if $\mathscr{P}$ is a direct product of the subgroups $\mathscr{L}$ and $\mathscr{T}$. In fact, it is not. Instead, $\mathscr{P} = \mathscr{L} \rtimes \mathscr{T}$, that is, $\mathscr{P}$ is a semidirect product of $\mathscr{L}$ and $\mathscr{T}$. We shall not go into the definition and detailed discussion of the semidirect product now. Instead, let us investigate the action of Poincare transformations on Hilbert spaces.

We have seen that the Poincare group, by definition, consists of transformations of the coordinates x^μ. If we demand that these transformations be symmetries of nature, then we can conclude that each Poincare transformation can be represented as either a unitary and linear operator on a Hilbert space, or as an anti-unitary and anti-linear operator on a Hilbert space. This conclusion is the result of Wigner's theorem. For the time being, we focus on Proper[33] Orthochronous Poincare transformations for which $\det\Lambda = 1$ and $\Lambda^0_0 \geq 1$. It can be shown that representations of proper orthochronous Poincare transformations can be chosen to be unitary and linear. From here onward, we shall

[33] The defining requirement $\Lambda^T \eta \Lambda = \eta$ on Λ for it to be a Lorentz transformation implies that $\det\Lambda = \pm 1$ and $|\Lambda^0_0| \geq 1$. These conditions are used to classify Lorentz transformations. Lorentz transformations with $\Lambda^0_0 \geq 1$ are called orthochronous, those with $\Lambda^0_0 \leq -1$ are called non-orthochronous. Lorentz transformations with $\det\Lambda = 1$ are called proper and those with $\det\Lambda = -1$ are called improper. The set of Lorentz transformations that are both proper and orthochronous forms a subgroup and is called P.O.L.T. group or the group of Proper Orthochronous Lorentz transformations. The set of Poincare transformations whose Lorentz transformation component is proper and orthochronous is a subgroup and is called the Proper Orthochronous Poincare transformation group or the P.O.P.T group.

not explicitly qualify a Poincare transformation with the adjectives proper and orthochronous. Unless otherwise mentioned, we will be considering proper and orthochronous transformations only. So, the effect of Poincare transformations in quantum mechanics is the following :

$$|\psi\rangle \rightarrow |\psi^{'}\rangle = U(\Lambda, a)\,|\psi\rangle \tag{204}$$

where $U(\Lambda, a)$ is chosen to be unitary and linear[34]. Evidently, the operators $U(\Lambda, a)$ form a representation of the Poincare group. We have,

$$U(\Lambda, a)\,U\left(\bar{\Lambda}, \bar{a}\right) = U\left(\Lambda\bar{\Lambda}, \Lambda\bar{a} + a\right) \tag{205}$$

Consider an infinitesimal proper Poincare transformation :

$$\begin{array}{c} \Lambda^{\mu}_{\ \nu} = \delta^{\mu}_{\ \nu} + \omega^{\mu}_{\ \nu} \\ a^{\mu} = \epsilon^{\mu} \end{array} \tag{206}$$

where $\omega^{\mu}_{\ \nu}$ and ϵ^{μ} are infinitesimal. The condition $\Lambda^T\eta\Lambda = \eta$ has to be satisfied b the infinitesimal transformation, and that implies $\omega_{\mu\nu} + \omega_{\nu\mu} = 0$. That is, ω, with both indices lowered, is antisymmetric. The antisymmetric matrix which has $\omega^{\mu}_{\ \nu}$ as the entry in the intersection of the μth row and the νth column is denoted by Ω. Similarly, the column vector with ϵ^{μ} as its μth entry is denoted by ϵ. Thus, the infinitesimal Poincare transformation (206) can be expressed in the matrix form as $(\mathbb{I} + \Omega, \epsilon)$. We have 6 independent Lorentz parameters $(\omega_{\mu\nu})$, 3 of them correspond to boosts and 3 to rotations. In addition, there are 4 translation parameters. So the Poincare group is a 10 parameter group, and therefore has 10 generators. Corresponding to the infinitesimal Poincare transformation (206), we have a unitary linear operator $U(\mathbb{I} + \Omega, \epsilon)$, which is different from the identity operator on the Hilbert space only slightly.

$$U(\mathbb{I} + \Omega, \epsilon) = \mathbb{I} + \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} - i\epsilon_{\mu}P^{\mu} \tag{207}$$

The i's are there to ensure that $J^{\mu\nu}$ and P^{μ}, the generators (on the Hilbert space) of Lorentz transformations and translations respectively, are Hermitian. From equation (205), we deduce that

$$U(\Lambda, a)\,U\left(\bar{\Lambda}, \bar{a}\right)\left(U(\Lambda, a)\right)^{-1} = U\left(\Lambda\bar{\Lambda}\Lambda^{-1}, \Lambda\bar{a} + a - \Lambda\bar{\Lambda}\Lambda^{-1}a\right) \tag{208}$$

In particular, for $\left(\bar{\Lambda}, \bar{a}\right) = (\mathbb{I} + \Omega, \epsilon)$,

$$U(\Lambda, a)\,U(\mathbb{I} + \Omega, \epsilon)\left(U(\Lambda, a)\right)^{-1} = U\left(\mathbb{I} + \Lambda\Omega\Lambda^{-1}, \Lambda\epsilon - \left(\Lambda\Omega\Lambda^{-1}\right)a\right) \tag{209}$$

We notice that the net Poincare transformation $\left(\mathbb{I} + \Lambda\Omega\Lambda^{-1}, \Lambda\epsilon - \left(\Lambda\Omega\Lambda^{-1}\right)a\right)$ is also infinitesimal because both Ω and ϵ are infinitesimal. Therefore, the right hand side can be expanded, following equation (207), as $\mathbb{I} + \frac{i}{2}\omega_{\mu\nu}\left(\Lambda\Omega\Lambda^{-1}\right)^{\mu\nu} -$

[34]The fact that it can be done is guaranteed by Wigner's theorem. For a detailed proof of the theorem, look up Weinberg, volume 1.

$i\left(\Lambda\epsilon - \Lambda\Omega\Lambda^{-1}a\right)_\mu P^\mu$. The middle factor on the left hand side has the familiar expansion (207). Thus,

$$U(\Lambda,a)\left(\mathbb{I}+\frac{i}{2}\omega_{\mu\nu}J^{\mu\nu}-i\epsilon_\mu P^\mu\right)(U(\Lambda,a))^{-1}=\mathbb{I}+\frac{i}{2}\omega_{\mu\nu}\left(\Lambda\Omega\Lambda^{-1}\right)^{\mu\nu}-i\left(\Lambda\epsilon-\Lambda\Omega\Lambda^{-1}a\right)_\mu P^\mu$$

$$\Rightarrow \mathbb{I}+\frac{i}{2}\omega_{\rho\sigma}U(\Lambda,a)\,J^{\rho\sigma}U^{-1}(\Lambda,a)-i\epsilon_\rho P^\rho=\mathbb{I}+\frac{i}{2}\omega_{\mu\nu}\left(\Lambda\Omega\Lambda^{-1}\right)^{\mu\nu}-i\left(\Lambda\epsilon-\Lambda\Omega\Lambda^{-1}a\right)_\mu P^\mu$$

Equating the antisymmetric part of the coefficients of $\omega_{\rho\sigma}$ and also coefficients of ϵ_ρ on both sides, we get

$$\begin{gathered}U(\Lambda,a)\,J^{\rho\sigma}U^{-1}(\Lambda,a)=\Lambda^\rho_\mu\Lambda^\sigma_\nu\left(J^{\mu\nu}-a^\mu P^\nu+a^\nu P^\mu\right)\\ U(\Lambda,a)\,P^\rho U^{-1}(\Lambda,a)=\Lambda^\rho_\mu P^\mu\end{gathered}\tag{210}$$

Here, $\Lambda^\rho_\mu \equiv \left(\Lambda^{-1}\right)^\rho_{\ \mu}$ = the $\mu\rho$th element of $\left(\Lambda^{-1}\right)^T$. Then, we choose (Λ,a) to be infinitesimal as well. Putting $(\Lambda,a)=(\mathbb{I}+\Omega,\epsilon)$ in (210), we get the commutators of the generators of Poincare transformations. These commutator relations give the Poincare algebra :

$$\begin{gathered}i\left[J^{\mu\nu},J^{\rho\sigma}\right]=\eta^{\nu\rho}J^{\mu\sigma}+\eta^{\mu\sigma}J^{\nu\rho}-\eta^{\mu\rho}J^{\nu\sigma}-\eta^{\nu\sigma}J^{\mu\rho}\\ i\left[P^\mu,J^{\rho\sigma}\right]=\eta^{\mu\rho}P^\sigma-\eta^{\mu\sigma}P^\rho\\ \left[P^\mu,P^\nu\right]=0\end{gathered}\tag{211}$$

The commutators become familiar when we switch to the one-index notation for the Lorentz generators $J^{\mu\nu}$. Define

$$\begin{gathered}\vec{J}=\left(J^1,J^2,J^3\right)\equiv\left(J^{23},J^{31},J^{12}\right)\leftarrow\text{rotation generators}\\ \vec{K}=\left(K^1,K^2,K^3\right)\equiv\left(J^{01},J^{02},J^{03}\right)\leftarrow\text{boost generators}\\ \vec{P}=\left(P^1,P^2,P^3\right)\leftarrow\text{space translation generators}\\ H=P^0\leftarrow\text{time translation generator, or the Hamiltonian}\end{gathered}\tag{212}$$

With this notation, the Poincare algebra[35] takes the form

$$\begin{gathered}\left[J_i,J_j\right]=i\epsilon_{ijk}J_k\qquad\left[K_i,K_j\right]=-i\epsilon_{ijk}J_k\\ \left[J_i,K_j\right]=i\epsilon_{ijk}K_k\qquad\left[K_i,P_j\right]=iH\delta_{ij}\\ \left[J_i,P_j\right]=i\epsilon_{ijk}P_k\qquad\left[H,K^i\right]=-iP^i\end{gathered}\tag{213}$$

This completes a quick review of the Poincare algebra that we learned in last semester's QFT course. Now, since $[P^\mu,P^\nu]=0$, therefore single particle states can be chosen to be simultaneous eigenstates of the 4-momentum operators P^μ. Let, $|p,\sigma\rangle$ be a simultaneous eigenstate of P^μ with respective eigenvalues p^μ. Here, σ denotes other labels that may be required to identify the states.

$$P^\mu|p,\sigma\rangle=p^\mu|p,\sigma\rangle\tag{214}$$

[35]This form of the commutation relations explicitly shows that $\vec{K}$ and $\vec{P}$ are vector operators under rotation.

Now that we have chosen one particle states, the next task is to compute the actions of Poincare transformations on these one particle states. Action of the translation operators is easy to deduce. We know that $U(\mathbb{I}, a) = e^{-iP.a}$, where $P.a = P^\mu a_\mu$. Therefore,

$$U(\mathbb{I}, a)\,|p, \sigma\rangle \equiv e^{-iP.a}|p, \sigma\rangle = e^{-ip.a}|p, \sigma\rangle \tag{215}$$

That is, $|p, \sigma\rangle$ is an eigenstate of $U(\mathbb{I}, a)$ with eigenvalue $e^{-ip.a}$. Therefore the translation operators are diagonal in the basis of one particle states $|p, \sigma\rangle$.

Before finishing off, let me briefly explain how we have $U(\mathbb{I}, a) = e^{-iP.a}$. An infinitesimal translation operator looks like the following : $\mathbb{I} - i\epsilon_\mu P^\mu$. To achieve a finite translation by $a \equiv \left(a^0, a^1, a^2, a^3\right)$, we can break it up into N smaller translations, each by an amount $\frac{a}{N}$. As we increase N, the smaller translations become infinitesimal, and we can approximate them by $\mathbb{I} - i\frac{a_\mu}{N}P^\mu$. In the limit $N \to \infty$, this becomes exact. Hence,

$$U(\mathbb{I}, a) = \lim_{N\to\infty} \left(\mathbb{I} - i\frac{a_\mu P^\mu}{N}\right)^N = e^{-ia_\mu P^\mu} \tag{216}$$

The last equality follows from the fact that the exponential function is defined as $e^x \equiv \lim_{N\to\infty} \left(1 + \frac{x}{N}\right)^N$. You may protest, that x is a number whereas $a_\mu P^\mu$ is an operator. That's true, but all the P^μ's commute with each other. The above limit involves only the operations of addition and multiplication, and x and P^μ's add and multiply exactly the same way.

7.3 Lecture 21 : February 25, 2016

We chose the simultaneous eigenstates $|p,\sigma\rangle$ of P^μ as the basis one particle states and figured out the action of a translation operator $U(\mathbb{I},a) = e^{-iP.a}$ on them. We found that the translation operators are diagonal in this basis. Now we would inspect the action of a pure[36] Lorentz transformation $U(\Lambda,0) \equiv U(\Lambda)$ on the states $|p,\sigma\rangle$. Start by observing

$$\begin{aligned} P^\mu\left(U(\Lambda)|p,\sigma\rangle\right) &= U(\Lambda)\left[U^{-1}(\Lambda)P^\mu U(\Lambda)\right]|p,\sigma\rangle \\ &= U(\Lambda)\Lambda^\mu_\rho P^\rho|p,\sigma\rangle = \left(\Lambda^\mu_\rho p^\rho\right)\left(U(\Lambda)|p,\sigma\rangle\right) \\ \therefore P^\mu\left(U(\Lambda)|p,\sigma\rangle\right) &= (\Lambda p)^\mu\left(U(\Lambda)|p,\sigma\rangle\right) \end{aligned} \tag{217}$$

Here, we have used the equation $U(\Lambda,a)P^\rho U^{-1}(\Lambda,a) = \Lambda^\rho_\mu P^\mu$ of (210) with Λ^{-1} in place of Λ and $a = 0$. This tells us that $U(\Lambda)|p,\sigma\rangle$ belongs to the eigenspace of P^μ corresponding to the eigenvalue $(\Lambda p)^\mu$. Thus,

$$U(\Lambda)|p,\sigma\rangle = \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle \tag{218}$$

Our job of computing the action of $U(\Lambda)$ on $|p,\sigma\rangle$ essentially boils down to finding out the coefficients $C_{\sigma\sigma'}(\Lambda,p)$. Even before setting out to find the $C_{\sigma\sigma'}$'s, let us observe the following. Consider the hypothetical situation where there is only one label σ and it can take 5 different values, namely $1,2,3,4,5$. Suppose now that we take the state $|p,1\rangle$ and make $U(\Lambda)$ act on it. The resultant state is a linear combination $C_{11}(\Lambda,p)|\Lambda p,1\rangle + C_{21}(\Lambda,p)|\Lambda p,2\rangle + C_{31}(\Lambda,p)|\Lambda p,3\rangle + C_{41}(\Lambda,p)|\Lambda p,4\rangle + C_{51}(\Lambda,p)|\Lambda p,5\rangle$. The coefficients $C_{\sigma'\sigma}$ can be put into a matrix form with σ' being the row index and σ the column index. Call this matrix C. Now, if we can choose the σ labels in such a way that C assumes a block diagonal form, say of the following kind,

$$C = \begin{bmatrix} \boxed{}_{3\times 3} & & \\ & \boxed{}_{2\times 2} & \\ & & \end{bmatrix} \tag{219}$$

then we can easily see that $U(\Lambda)$ mixes states with σ values $1,2$ and 3 among themselves and states with σ values 4 and 5 among themselves. You might think that such a choice for the label σ breaks the Hilbert space up in a direct sum of the component irreps of $U(\Lambda)$. But that is simply wrong. Because, even with such a block diagonal C, $U(\Lambda) : |p,1\rangle \mapsto C_{11}(\Lambda,p)|\Lambda p,1\rangle + C_{21}(\Lambda,p)|\Lambda p,2\rangle + C_{31}(\Lambda,p)|\Lambda p,3\rangle$. That is, $|p,(1,2,3)\rangle$ go to a linear combination of $|\Lambda p,(1,2,3)\rangle$ and not of $|p,(1,2,3)\rangle$. In addition to mixing of the σ values, the value of p

[36] Meaning a Poincare transformation with zero translation.

changes to Λp. Notice one more thing. $U(\Lambda)$ is a symmetry transformation. It relates two ways of describing/observing the same physical system – precisely a change of frames. That means, in the hypothetical example, if one observer describes a particle with momentum p by the state $|p,1\rangle$, then a Λ-transformed observer will describe the same particle with a linear combination of the states $|\Lambda p,(1,2,3)\rangle$. This means that $|\Lambda p,1\rangle, |\Lambda p,2\rangle, |\Lambda p,3\rangle$ are all different states of the same particle. And, if C could be brought in the block diagonal form as above, then the states $|\Lambda p,(1,2,3)\rangle$ never mix with the states $|\Lambda p,(4,5)\rangle$, no matter which Lorentz transformation you use. Therefore, $|\Lambda p,(4,5)\rangle$ have to be states of a different particle. With this in mind, let us try to decipher as much as we can about the structure of $C_{\sigma'\sigma}$. The ensuing discussion will teach us what is known as the Wigner's method of induced representation.

To begin with, I shall state, without proof, two properties of proper orthochronous Lorentz transformations. The proofs can be found at Weinberg's first volume, but this would be a nice exercise for you to try on your own.

Proper orthochronous Lorentz transformations leave the following invariant :

- Norm of a state.
- Algebraic sign[37] of a^0 if a^μ is a time-like 4-vector.

We have been using the mostly negative Minkowski metric. With this signature, time-like vectors have positive norm. Now, massive particles have $p^2 = m^2 > 0$, massless particles have $p^2 = 0$. Finally, the vacuum has $p \equiv (0,0,0,0)^T$.

Weinberg uses a notation in which the zeroth component of a 4-vector is written at the bottom of the column vector : $p = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \\ p^0 \end{pmatrix}$. We shall also use this convention for the present discussion.

Consider a massive particle, with $p^2 = m^2 > 0$. The momentum 4-vector of the particle has the same norm m^2 in all Lorentz frames. In an instantaneous rest frame of the particle, it has $p = (0,0,0,M)^T$. Therefore, the sign of p^0 stays positive under every P.O.L.T. So, we have an infinite family of 4-momentum vectors p, each with $p^2 = m^2$ and $\text{sgn}\left(p^0\right) > 0$. From this infinite family of momentum 4-vectors related to each other through proper orthochronous Lorentz transformations, let us choose one as a standard, and call it k. Having chosen the standard 4-momentum k, we can make a P.O.L.T act on it to get an arbitrary 4-momentum p belonging to the same family. In fact, given a particular p in the family, there are infinitely[38] many proper orthochronous

[37] We shall denote it by $\text{sgn}\left(a^0\right)$

[38] Let me explain why or how there are infinitely many Lorentz transformations that take k to p. Let Λ_1 be a Lorentz transformation that takes k to $p = \left(p^0, \vec{p}\right)$. Consider now a rotation

Lorentz transformations that take k to p. We shall choose one of these Lorentz transformations as a standard Lorentz transformation and name it $L(p)$. The conventional choice of standard momenta are listed below :

Category	Choice of standard momentum	Little group
$p^2 = m^2 > 0$; massive particle	$k = \begin{pmatrix} 0 \\ 0 \\ 0 \\ m \end{pmatrix}$	$SO(3)$
$p^2 = 0$; massless particle	$k = \begin{pmatrix} 0 \\ 0 \\ \kappa \\ \kappa \end{pmatrix}$	$ISO(2)$
Vacuum	$k = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$SO(3,1)$

The meaning of the third column will be apparent soon. We also have some conventional choices for the standard Lorentz transformations. We shall quote them and justify the choices as we go along. Suppose, for now, that we have a standard momentum k, and a standard Lorentz transformation $L(p)$ that takes k to p :

$$p^\mu = [L(p)]^\mu_{\ \nu} k^\nu; \quad p^2 = k^2; \quad \text{sgn}\left(p^0\right) = \text{sgn}\left(k^0\right) \tag{220}$$

Also suppose that we have assigned the labels σ for the states with 4-momentum eigenvalue k. Pick the state $|k,\sigma\rangle$ and make $U(L(p))$ act on it. Since $L(p)$ is a Lorentz transformation that takes k to p, therefore equation (217) implies that $U(L(p))|k,\sigma\rangle$ has p as its 4-momentum eigenvalue. Since we haven't decided on the σ labeling of the states with 4-momentum eigenvalue p yet, we have the liberty to fix that by the following definition of $|p,\sigma\rangle$:

$$|p,\sigma\rangle \equiv N(p)\, U(L(p))\, |k,\sigma\rangle \tag{221}$$

Equation (221) connects the σ values for the different values of the 4-momentum. The factor $N(p)$ stands there as a normalization[39] factor. Now,

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= U(\Lambda)\, U(L(p))\, |k,\sigma\rangle . N(p) = N(p)\, U(\Lambda L(p))\, |k,\sigma\rangle \\ &= N(p)\, U(L(\Lambda p))\, U^{-1}(L(\Lambda p))\, U(\Lambda L(p))\, |k,\sigma\rangle \\ \therefore U(\Lambda)|p,\sigma\rangle &= N(p)\, U(L(\Lambda p))\, U\left((L(\Lambda p))^{-1} \Lambda L(p)\right) |k,\sigma\rangle \end{aligned} \tag{222}$$

R by an arbitrary angle θ about the axis pointing along the direction of $\vec{p}$. Evidently, $Rp = p$. Therefore, $(R\Lambda)\, k = R(\Lambda k) = Rp = p$, proving the claim.

[39] The momentum eigenstates are so defined that $\langle k,\sigma|k',\sigma'\rangle = \delta_{\sigma\sigma'}\delta^3\left(\vec{k} - \vec{k}'\right)$. This normalization is not Lorentz invariant. Therefore, the action of $U(L(p))$ on $|k,\sigma\rangle$ produces a state that no longer has a unit norm. We multiply $U(L(p))|k,\sigma\rangle$ with the factor $N(p)$ so that the product has unit norm, and define that product to be $|p,\sigma\rangle$.

Focus now on the Lorentz transformation

$$\left(L\left(\Lambda p\right)\right)^{-1}\Lambda L\left(p\right)\equiv W\left(\Lambda,p\right) \tag{223}$$

The action of $W\left(\Lambda,p\right)$ on an arbitrary 4-momentum is not known. But its action on k is easily found out : $W\left(\Lambda,p\right)k=\left[\left(L\left(\Lambda p\right)\right)^{-1}\Lambda L\left(p\right)\right]k=\left[\left(L\left(\Lambda p\right)\right)^{-1}\Lambda\right]p=\left[L^{-1}\left(\Lambda p\right)\right]\left(\Lambda p\right)=k$. That is, $\mathcal{W}=\{W\left(\Lambda,p\right)\}$ is the set of Lorentz transformations that leave the standard 4-momentum k invariant. It's easy to verify that $\mathcal{W}$ is a group. It is called the **little group** of k.

Definition : Little group of k is the group of all proper orthochronous Lorentz transformations that leave k invariant.

Every k comes with a little group $\mathcal{W}_k$ of its own. $\mathcal{W}_k$ is a subgroup of the entire P.O.L.T group. That is probably the reason why it is called the little group, signifying the fact that $\mathcal{W}_k$ is a "smaller" set of symmetries. The three standard 4-momenta that will be of our interest are those for the massive particles, the massless particles and the vacuum respectively. The choices of k for these three have been supplied already. Their corresponding little groups have also been tabulated. We shall soon work out the details. Let me first point out an interesting detail. The standard momentum for massive particles has $SO\left(3\right)$ as its little group. As a consequence, the knowledge of $SO\left(3\right)$ is all we need to treat the quantum theory of a relativistic particle. This significantly cuts down our effort to find the structure of $C_{\sigma'\sigma}$ for massive particles since we already know everything there is to know about $SO\left(3\right)$.

Let $U\left(W\left(\Lambda,p\right)\right)$ be the operator on the Hilbert space carrying out the effect of $W\left(\Lambda,p\right)$. Since we know that the action of $W\left(\Lambda,p\right)$ on k is to leave it invariant, therefore

$$U\left(W\left(\Lambda,p\right)\right)\left|k,\sigma\right\rangle=\sum_{\sigma'}D_{\sigma'\sigma}\left(W\left(\Lambda,p\right)\right)\left|k,\sigma'\right\rangle \tag{224}$$

This result is nicer than (218) because it makes the fact obvious that the span of the states $\left|k,\sigma\right\rangle$ for different σ's carries an irrep of the little group $\mathcal{W}_k$. How does all this help in determining the action of $U\left(\Lambda\right)$ on $\left|p,\sigma\right\rangle$? Equation (222) implies

$$\begin{aligned}U\left(\Lambda\right)\left|p,\sigma\right\rangle&=N\left(p\right)U\left(L\left(\Lambda p\right)\right)U\left(\left(L\left(\Lambda p\right)\right)^{-1}\Lambda L\left(p\right)\right)\left|k,\sigma\right\rangle\\&=N\left(p\right)U\left(L\left(\Lambda p\right)\right)\sum_{\sigma'}D_{\sigma'\sigma}\left(W\left(\Lambda,p\right)\right)\left|k,\sigma'\right\rangle\\&=N\left(p\right)\sum_{\sigma'}D_{\sigma'\sigma}\left(W\left(\Lambda,p\right)\right)U\left(L\left(\Lambda p\right)\right)\left|k,\sigma'\right\rangle\\\therefore U\left(\Lambda\right)\left|p,\sigma\right\rangle&=\frac{N\left(p\right)}{N\left(\Lambda p\right)}\sum_{\sigma'}D_{\sigma'\sigma}\left(W\left(\Lambda,p\right)\right)\left|\Lambda p,\sigma'\right\rangle\end{aligned} \tag{225}$$

Equations (225) and (218) both give the action of $U(\Lambda)$ on $|p,\sigma\rangle$. Comparing the two, we get $C_{\sigma'\sigma}(\Lambda,p)$ in terms of $D_{\sigma'\sigma}(W(\Lambda,p))$. Since the D matrix is that of an irrep of the little group and is easily found out, therefore our problem is solved.

8 Week 8

8.1 Lecture 22 : February 29, 2016

Today we shall start off from equation (225). Let me first throw in a fact without proof : the ratio $\frac{N(p)}{N(\Lambda p)}$ equals $\sqrt{\frac{(\Lambda p)^0}{p^0}}$. This is a consequence of $N(p) = \sqrt{\frac{k^0}{p^0}}$. You shall prove it in your next assignment. It basically follows from $\langle p|p^{'}\rangle = (2\pi)^3 2E_{\vec{p}}\delta^3\left(\vec{p}-\vec{p}^{'}\right) = (2\pi)^2 2p^0\delta^3\left(\vec{p}-\vec{p}^{'}\right)$, the proof of which you saw in the QFT course last semester. Hence, the end result of the last lecture was as follows :

$$C_{\sigma^{'}\sigma}(\Lambda,p) = \frac{N(p)}{N(\Lambda p)} D_{\sigma^{'}\sigma}(W(\Lambda,p)) = \sqrt{\frac{(\Lambda p)^0}{p^0}} D_{\sigma^{'}\sigma}(W(\Lambda,p)) \tag{226}$$

This is the central result of Wigner's method of induced representations.

We shall now make the somewhat general discussion of the method of induced representations particular by considering concrete applications.

- Massive particles :
  Norm of the 4-momentum of a particle of mass M is M^2. We are adopting Weinberg's notation – 4-vectors : $x = \left(x^1, x^2, x^3, x^0\right)^T$, metric : diag $(-1,-1,-1,1)$. Our choice of standard momentum for this case is

$$k = \begin{pmatrix} 0 \\ 0 \\ 0 \\ M \end{pmatrix} \tag{227}$$

  A Lorentz transformation that leaves k invariant cannot involve boosts, because a boost changes the zeroth component of a 4-vector. Any rotation in space leaves k invariant. Therefore, the little group of this standard momentum k is $\mathcal{W}_k = SO(3)$. Equation (226) essentially tells us that how a state changes under a general Lorentz transformation is determined by how states in the irreps of $\mathcal{W}_k$ mix under little group elements. In this case, the little group is $SO(3)$ and we know all about the irreps of $SO(3)$, that the states are labeled by j and m_j, and how the labels m_j mix under rotations. This is great news, except for the fact that we haven't exactly specified $L(p)$ and hence $W(\Lambda,p) = (L(\Lambda p))^{-1}\Lambda L(p)$. We shall do it now. Firstly, $L(p)$ is not unique. All we know about it is $L(p)k = p$. Let's see what that tells us about the matrix elements of $L(p)$.

$$L(p)\begin{pmatrix} 0 \\ 0 \\ 0 \\ M \end{pmatrix} = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \\ \gamma M \end{pmatrix} \tag{228}$$

Also, $|\vec{p}| = \beta\gamma M \Rightarrow \frac{|\vec{p}|}{M} = \beta\gamma \Rightarrow \frac{\vec{p}}{M} = \frac{|\vec{p}|}{M}\hat{p} = \beta\gamma\hat{p}$. Therefore,

$$L(p) = \left[\begin{array}{c|c} & \beta\gamma\hat{p} \\ \hline & \gamma \end{array}\right] : \begin{pmatrix} 0 \\ 0 \\ 0 \\ M \end{pmatrix} \mapsto \begin{pmatrix} \vec{p} \\ \\ \gamma M \end{pmatrix} \tag{229}$$

Therefore, the 4^{th} column of $L(p)$ is determined from its action on k. The first three columns are not completely arbitrary, because $L(p)$ has to be a Lorentz transformation after all. But there are infinitely many choices for these three columns. If you find one $L(p)$ that works, you immediately get infinitely many $L^{'}(p) = (\text{rotation about } \hat{p})\, L(p)\, (\text{any rotation})$ that work just as well. Therefore, we shall have to exercise a choice of the standard Lorentz transformation that we will use for all massive particles. Notice that, a Lorentz transformation that maps $(0,0,0,M)^T$ to $(0,0,\beta\gamma M,\gamma M)^T$ is the following boost :

$$B(|\vec{p}|) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & \beta\gamma \\ 0 & 0 & \beta\gamma & \gamma \end{bmatrix} \tag{230}$$

This is a boost along the z direction. To obtain a boost along an arbitrary direction $\hat{p}$, we shall just have to follow $B(|\vec{p}|)$ up with a rotation $R(\hat{p})$ that rotates the z-axis onto the $\hat{p}$ axis. So, we might say that we have found what we were after :

$$L(p) = R(\hat{p})\, B(|\vec{p}|) \tag{231}$$

However, this is not a unique choice because of two reasons. First off, we could begin with an arbitrary rotation about an arbitrary axis which would not change $(0,0,0,M)^T$ at all. We shall indeed use this freedom and choose the following as our standard Lorentz transformation :

$$L(p) = R(\hat{p})\, B(|\vec{p}|)\, R^{-1}(\hat{p}) \tag{232}$$

The first rotation has been chosen to be $R^{-1}(\hat{p})$. This is a choice and you could have chosen any other rotation just as well. We choose this because it is customary, and this choice is customary because of a good reason that will become apparent soon. There is a second reason why (231) was not unique and that reason still persists. The rotation $R(\hat{p})$ has been defined as one that rotates the z-axis onto the $\hat{p}$ direction. There are infinitely many rotations that do that. If you come up with one rotation that does the job, we can follow it up with an arbitrary rotation about the $\hat{p}$ axis and the resultant $L(p)$ will have the same action on $(0,0,0,M)^T$. So, we shall have to choose one $R(\hat{p})$. Surprisingly, we shall see that any choice of $R(\hat{p})$ gives rise to the same $L(p)$! Now, notice that (232) is a

symmetric choice for $L(p)$. The rotation $R(\hat{p})$ is a 4×4 matrix of the form $\left[\begin{array}{c|c} R_{(3)}(\hat{p}) & \begin{matrix}0\\0\\0\end{matrix} \\ \hline \begin{matrix}0 & 0 & 0\end{matrix} & 1 \end{array}\right]$, where $R_{(3)}(\hat{p})$ is an arbitrary rotation that rotates $\hat{e}_3 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$, the unit vector along the z-axis, to $\hat{p}$.

$$R_{(3)}(\hat{p})\,\hat{e}_3 = \hat{p} \tag{233}$$

Therefore,

$$L(p) = \left[\begin{array}{c|c} R_{(3)}(\hat{p}) & \begin{matrix}0\\0\\0\end{matrix} \\ \hline \begin{matrix}0 & 0 & 0\end{matrix} & 1 \end{array}\right] \left[\begin{array}{c|c} \begin{matrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \gamma\end{matrix} & \gamma\beta\hat{e}_3 \\ \hline \gamma\beta\hat{e}_3^T & \gamma \end{array}\right] \left[\begin{array}{c|c} R_{(3)}^T(\hat{p}) & \begin{matrix}0\\0\\0\end{matrix} \\ \hline \begin{matrix}0 & 0 & 0\end{matrix} & 1 \end{array}\right] \tag{234}$$

Now we use the following :

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \gamma\end{bmatrix} = \mathbb{I}_3 + (\gamma-1)\,\hat{e}_3\hat{e}_3^T \tag{235}$$

Therefore,

$$L(p) = \left[\begin{array}{c|c} R_{(3)}(\hat{p}) & 0 \\ \hline 0 & 1 \end{array}\right] \left[\begin{array}{c|c} \mathbb{I}_3 + (\gamma-1)\,\hat{e}_3\hat{e}_3^T & \gamma\beta\hat{e}_3 \\ \hline \gamma\beta\hat{e}_3^T & \gamma \end{array}\right] \left[\begin{array}{c|c} R_{(3)}^T(\hat{p}) & 0 \\ \hline 0 & 1 \end{array}\right]$$

$$\therefore L(p) = \left[\begin{array}{c|c} \mathbb{I}_3 + (\gamma-1)\,\hat{p}\hat{p}^T & \gamma\beta\hat{p} \\ \hline \gamma\beta\hat{p}^T & \gamma \end{array}\right] \tag{236}$$

The final result shows us that the choice of $R_{(3)}(\hat{p})$ is immaterial – any choice gives rise to the same $L(p)$. Comparing this to (229) we see that the 4^{th} columns match, which is no surprise. So, our choice/convention of $L(p) = R(\hat{p})\,B(|\vec{p}|)\,R^{-1}(\hat{p})$ yields

$$\begin{aligned} L(p)^0_{\ 0} &= \gamma \\ L(p)^0_{\ i} &= \beta\gamma\hat{p}_i \\ L(p)^i_{\ j} &= \delta_{ij} + (\gamma-1)\,\hat{p}_i\hat{p}_j \end{aligned} \tag{237}$$

Having fixed the standard Lorentz transformation, we now have to compute $W(\Lambda,p) = L^{-1}(\Lambda p)\,\Lambda L(p)$. Let me stress here that a rotation is a Lorentz transformation. We already know that the little group for the standard k chosen for the massive particles is $SO(3)$. An $SO(3)$ matrix is, by definition, 3×3. But elements of the little group of k have to be 4×4, since they are Lorentz transformations. Therefore, when we say that the little group is $SO(3)$, we really mean that it is the group of 4×4 matrices which have the following form

$$\mathcal{R} = \begin{bmatrix} R\in SO(3) & 0 \\ 0 & 1 \end{bmatrix} \tag{238}$$

Consider $\Lambda = \mathcal{R}$.

$$W(\Lambda, p) = W(\mathcal{R}, p) = L^{-1}(\mathcal{R}p)\,\mathcal{R}L(p) \tag{239}$$

Now, $L(p) = R(\hat{p})\,B(|\vec{p}|)\,R^{-1}(\hat{p})$ implies, $L(\mathcal{R}p) = R(\mathcal{R}\hat{p})\,B(|\mathcal{R}\vec{p}|)\,R^{-1}(\mathcal{R}\hat{p})$ and $L^{-1}(\mathcal{R}p) = R(\mathcal{R}\hat{p})\,B^{-1}(|\mathcal{R}\vec{p}|)\,R^{-1}(\mathcal{R}\hat{p})$. Furthermore, $|\mathcal{R}\vec{p}| = |\vec{p}|$. Hence,

$$W(\mathcal{R}p) = \left[R(\mathcal{R}\hat{p})\,B^{-1}(|\vec{p}|)\,R^{-1}(\mathcal{R}\hat{p})\right]\mathcal{R}\left[R(\hat{p})\,B(|\vec{p}|)\,R^{-1}(\hat{p})\right] \tag{240}$$

There are two boosts in the above expression. If we could somehow commute $B^{-1}(|\vec{p}|)$ across $R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}R(\hat{p})$, then the two boosts would negate each other. Let us investigate the action of $R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}R(\hat{p})$ on $\hat{e}_3$. $\left[R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}R(\hat{p})\right]\hat{e}_3 = \left[R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}\right]\hat{p} = \left[R^{-1}(\mathcal{R}\hat{p})\right](\mathcal{R}\hat{p}) = \hat{e}_3$. This means that $R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}R(\hat{p})$ is a rotation that leaves $\hat{e}_3$ unchanged. Hence, it must be a rotation about $\hat{e}_3$. And, $B(|\vec{p}|)$ is a boost along $\hat{e}_3$. Therefore, these two must commute. Therefore, $B^{-1}(|\vec{p}|)$ can indeed be commuted across $R^{-1}(\mathcal{R}\hat{p})\,\mathcal{R}R(\hat{p})$. We end up with $W(\mathcal{R}p) = \left[R(\mathcal{R}\hat{p})\,R^{-1}(\mathcal{R}\hat{p})\right]\mathcal{R}\left[R(\hat{p})\,R^{-1}(\hat{p})\right]$. Clearly,

$$W(\mathcal{R}, p) = \mathcal{R} \tag{241}$$

The end result is that the little group element $W(\mathcal{R}, p)$ corresponding to the Lorentz transformation $\mathcal{R}$ (a rotation) is $\mathcal{R}$ itself. This is a consequence of the choice of $L(p)$ we made in (232). Had we not chosen the standard Lorentz transformation $L(p)$ to be symmetric by starting with the rotation $R^{-1}(\hat{p})$, this nice result would not have followed. This is exactly the reason why we chose such an $L(p)$. Keep in mind that (241) holds for rotations only. For a Lorentz transformation Λ which is not a rotation, $W(\Lambda, p)$ will not be equal to Λ. So, our choice of $L(p)$ favors rotations over other Lorentz transformations in that W's for $\mathcal{R}$'s equal $\mathcal{R}$'s themselves. However, this is not an unfair partiality because rotations have a special significance for massive particles. They form the little group of the standard k chosen for massive particles.

Let me mention an important point here. We said before that if there are some one particle states which mix under Lorentz transformations then those states must be the states of the same particle. Because, otherwise, merely going to a different Lorentz frame would change the identity of a particle, which is absurd. This is impeccable logic. However, we tend to extend this argument to also claim that states which do not mix under any Lorentz transformation belong to different particles. Is this claim necessarily true? Let me mention a counter-example. Photons are massless particles. There are two independent polarization states of photons, left and right circularly polarized states. These states do not mix under any boosts or rotations whatsoever. Does that mean that these two states are basically of two different particles and there are really

two different photons in the picture? The answer is a resounding no! Because, boosts and rotations are not the only symmetry transformations of photons. Electromagnetism also has parity symmetry which is a Lorentz transformation that belongs outside the P.O.L.T group. And parity mixes these two polarization states. A left circularly polarized photon would be observed to be right circularly polarized from a parity reversed frame. Since a change of frame cannot alter the identity of a particle, therefore we say that both the states are the states of the same particle – the photon. The lesson is the following. If S is the set of ALL the symmetry transformations of a one particle system and no element of S mixes the states $|\alpha\rangle$ and $|\beta\rangle$, then there is no reason to believe that these two states are of the same particle. In such a situation, we conclude that these are states of two different particles and try to look for some intrinsic quantities (e.g., mass, spin etc.) with respect to which they differ.

8.2 Lecture 23 : March 2, 2016

The idea of our present discussion is the following. If we know what the little group does to states then we know what the entire P.O.L.T group does to states. So, we look for the action of the little group elements. In the last class, we finished discussing the choice of $L(p)$ for massive particles that conveniently results in $W(\mathcal{R}, p) = \mathcal{R}$ where $\mathcal{R} \in SO(3)$, the little group for the standard momentum $k = \begin{pmatrix} 0 & 0 & 0 & M \end{pmatrix}^T$ chosen for massive particles. Today we shall take up the vacuum and the massless particles.

- The Vacuum :
  The standard momentum vector is

$$k = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{242}$$

  and the little group is $\mathcal{W}_k = SO(3,1)$, the entire P.O.L.T group. The prescription of looking for irreps of the little group does not help because the little group is by no means littler than the entire group. But wait! The point of studying the little group was that its action on the states would determine the action of an arbitrary $U(\Lambda)$ on the states. We already know what an arbitrary $U(\Lambda)$ does to the vacuum states.

$$\Lambda \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \implies U(\Lambda) |\text{vacuum}\rangle = |\text{vacuum}\rangle \tag{243}$$

  So we are done here. The span of the vacuum state carries the trivial representation of the Lorentz group.

- Massless particles :
  The standard momentum vector is

$$k = \begin{pmatrix} 0 \\ 0 \\ \kappa \\ \kappa \end{pmatrix} \tag{244}$$

  ... to be continued...