ELEMENTS OF MODERN PHYSICS

Prof. Saurabh Basu
Department of Physics
IIT Guwahati

Lec 9: Spin Angular Momentum, Perturbation Theory

We will finish first two things that were pending from the earlier discussion. One is the continuity equation and Pauli matrices. So, we have talked about angular momentum in very brief. We will talk about Pauli matrices and some of the properties of Pauli matrices. And then we will get ahead with a new topic called as approximate methods in quantum mechanics, which involves learning perturbation theory and corrections, sorry about that, the corrections to energy and eigenfunction. We'll talk about non-degenerate perturbation theory and then degenerate perturbation theory and apply it to a few examples and then probably go over to the time-dependent perturbation theory. This, what I'm talking about here is a time-independent perturbation theory. Okay, so let me do continuity equation and It has been written earlier, but let us derive it. This is also called as charge continuity equation and so on. So, let us just assume a region in space which has no source of charge or current. So, no source of charge and current and will sort of derive a continuity equation. So, what we have is that the Hamiltonian is given by minus h-cross squared over 2m and a del squared plus V of r, and this V of r could be a function of r, theta, and phi, but mostly it depends only on r, that is, the radial variable is a real quantity, okay.

So, for the Schrödinger equation, we have H psi equal to i h-cross d psi dt. Let us call this equation 1, and this is equation 2. Now, let us take the conjugate of equation 2. So, we have H psi star equal to minus i h-cross d psi star dt. Remember that H is equal to H star because H is a Hermitian operator. Let us call this equation 3. Now, multiply equation 2 by psi star and equation 3 by psi, and then finally subtract. and if you do that then what happens is that so you have a psi star h psi minus psi h psi star which is equal to ih cross psi star d psi dt and minus so this will be So, plus ih cross psi d psi star dt. So, now what we do is that we put it into this equation. So, we put 1 replacing h and let us call this as 4. So, put 1 in 4. And then we have so minus so we have this Psi star minus h squared by 2m del squared plus V of r, and you have psi, and that is minus psi. We have minus h squared over 2m del squared plus V of r. And we have a psi star, and then we have i h-cross d dt of—so you have psi star d psi dt and psi d psi star dt. So, this can be written as d dt of psi psi star, okay. As I mentioned, we are talking about a region where there is no source for charge or current. So, this would be psi star minus h squared over 2m del squared psi, plus V of r, which is a real quantity. So, V of r and psi star psi, and we have plus psi. And then you have a h cross square over 2m del square psi. And then we have V of r. So, there is a minus sign here and then that will be minus V of r psi star psi equal to i h cross to remind you this is called as a probability density.

Continuity equation.

No source of charge and current.

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(r) \quad (1)$$

$$H\psi = i\hbar \frac{d\psi}{dt} \quad (2)$$

$$H\psi^* = -i\hbar \frac{d\psi^*}{dt} \quad (3) \qquad H = H^*.$$

Multiply (2) by ψ^* and (3) by ψ & subtract.

$$\psi^* H\psi - \psi H\psi^* = i\hbar\, \psi^* \frac{d\psi}{dt} + i\hbar\, \psi \frac{d\psi^*}{dt} \quad (4)$$

Put (1) in (4)

So, which is equal to rho. So, rho is equal to psi star psi. okay, and this conveniently cancels out and what we have is that we can write this term as minus h cross square over 2m and divergence of psi star del psi minus psi del psi star and you have i h cross d rho dt. If you define, so 1 h cross goes away and this i can be brought downstairs so that this i reappears here. and if we define J which is a function of say r and t in general, this is equal to Psi star with h cross over 2 mi and Psi star del Psi minus Psi del Psi star. This gives you that divergence of J plus a del rho del t is equal to 0. So, this is the equation of continuity.

$$\psi^*\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\psi - \psi\left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\psi^* = i\hbar\frac{d}{dt}(\psi\psi^*).$$

$$\psi^*\left(-\frac{\hbar^2}{2m}\nabla^2\psi\right) + V(r)\psi^*\psi + \psi\frac{\hbar^2}{2m}\nabla^2\psi - V(r)\psi^*\psi = i\hbar\frac{d\rho}{dt} \qquad \rho = \psi^*\psi$$

$$\Rightarrow -\frac{\hbar^2}{2mi}\vec{\nabla}\cdot\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) = \hbar\frac{d\rho}{dt}.$$

$$j(\vec{r},t) = \frac{\hbar}{2mi}\left[\psi^*\nabla\psi - \psi\nabla\psi^*\right]$$

$$\vec{\nabla}\cdot\vec{J} + \frac{\partial\rho}{\partial t} = 0 \qquad : \text{Equation of Continuity.}$$

And if you have a source term there, then, of course, the right-hand side won't be equal to 0. That's equal to the source term that is present. So this is precisely the equation of continuity, which is expected to hold good for any system that we discuss or we have discussed or we are going to discuss. So let's go on to another topic, which is called the Pauli spin matrices. To remind you, spin does not arise naturally in non-relativistic quantum mechanics. So, we have to put that spinor. If a particle has a spin, the spinor wave function has to be put in by hand, which otherwise naturally occurs in the Dirac equation.

So, in any case, we will put simply by hand and then the spinner wave function has a dimension which is you know 2s plus 1. So, 2s plus 1 is a dimension and any spin operators they have the dimension of 2s plus 1 into 2s plus 1. So, for s equal to half, we have a 2 component wave function and we have an operator which is a 2 cross 2. For s equal to 1, we have a 3 component wave function and we have a 3 cross 3 operator whose eigenvalues are going to be the eigenvalues for the spin operator and the eigenfunctions are called as the spinner eigenfunctions. So, what are important in this context are S equal to half operators and these operators are known as Pauli spin matrices. they are in general denoted by this sigma. So, sigma it has three components sigma x, sigma y and sigma z, okay. And so, in general a spin and this spin half operators or this Pauli matrices have this relation that is equal to h cross by 2. So, h cross is the scale of the angular momentum including the spin angular momentum. So, it sort of absorbs that h cross there so that sigma can be written in terms of the just simply numbers because the dimension has already been taken into account. So, the three Pauli spin matrices are 0, 1, 1, 0, sigma y equal to that is for sigma x it is 0, minus i, i, 0, okay.

And the sigma z, which is diagonal 1 0 0 minus 1, okay. So whichever you take, they have certain properties that are common. And one of the property that is common is that the eigenvalue of each one of them are, you know, plus minus 1, okay. And so suppose you talk about sigma z. And then the wave function corresponding to that, so let us call it as chi up corresponding to the plus 1 eigenvalue. This has a form 1, 0, whereas the chi down corresponding to the minus 1 eigenvalue, this is equal to a 0, 1. And of course, there are other things such as a trace equal to 0 of each one of them.

Pauli Spin matrices.

Spinor wavefunction has a dimension $(2s+1)$.

Spin operator $(2s+1)\times(2s+1)$.

$S=\frac{1}{2}$ operators $\longrightarrow$ Pauli spin matrices $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma}$$

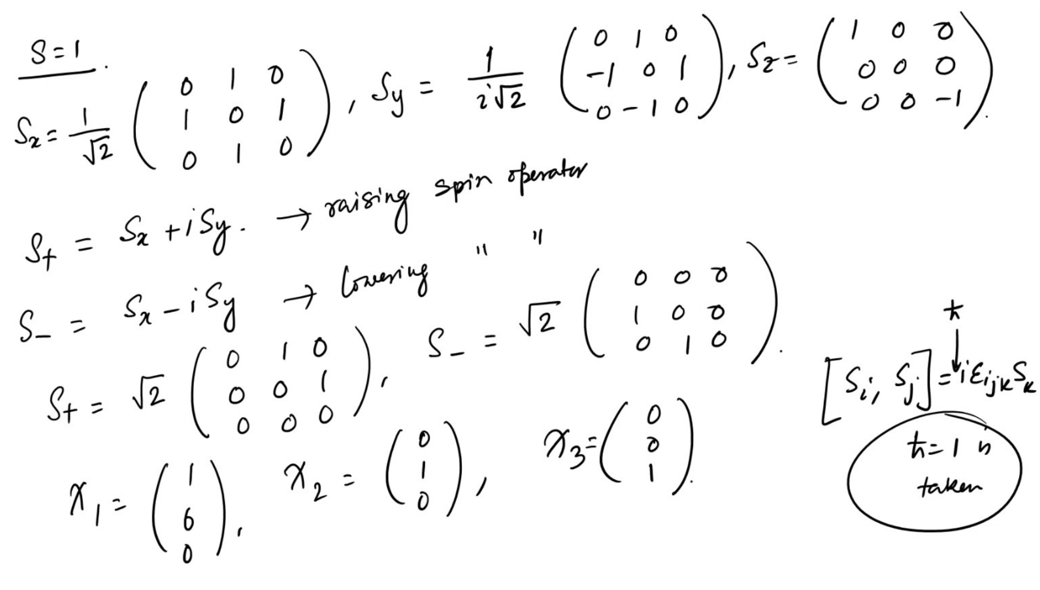
$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

1) Eigenvalues ± 1 $\quad \chi_{\uparrow}(+1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_{\downarrow}(-1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$

2) Trace $= 0$

You can see that there are some of the diagonal elements equal to 0 and so on. We are just going to come to that soon. But anything, you know, a spin 1 system, for example, which are not the Pauli matrices, but these are spin 1. And these, as I said, that there will be 3 by 3. So it's 1 by root 2, 0, 1, 0, 1, 0, 1. 0, 1, 0 and the Sx, sorry Sy, Sx I have already written down, it is 1 divided by i root 2 and then you have 0, 1, 0, minus 1, 0, 1, 0, minus 1, 0. And Sz, again, it is diagonal 1, 0, 0, 0, 0, 0, 0, 0, 0, minus 1. So, in each of these cases, one usually chooses Sz to be diagonal. And nevertheless, I mean, they all have these properties, some of which are quite universal for all the spin matrices. It is also useful at times for some problems to combine this.

Sx and Sy are written as Sx plus i Sy and S minus as Sx minus i Sy, and these are called the raising spin operator and the other one is called the lowering spin operator, okay. And they're called raising because they raise the eigenvalue by one. In any case, we do not go too much into details. You can read and get some knowledge about these things. In addition to that, there will be some assignment problems that will help you to get more familiar with these spin matrices. It is 1 0 0 and 0 1 0 and so on. The eigenvalues for this are chi 1 is equal to 1, 0, 0; chi 2 is equal to 0, 1, 0; and chi 3 is equal to 0, 0, 1. So, let us repeat some properties of this equation. Pauli matrices and so these are really some key properties which are often used. So, this is for an arbitrary vector rather. We have this sigma dot A; we can write this down as sigma x A x plus sigma y A y plus sigma z A z, and this can be written as A z, A x minus i A y, and A x plus i A y, and minus A z.



These are Hermitian matrices, and the way you should check Hermiticity is that you interchange the row and column, which is like taking a transpose, and then you should also take the complex conjugate. Now, if you take a transpose, then this A x plus i A y goes here and this one comes down, but then when you take a transpose, this i will change sign and you will have the same matrix. So, all these are, you know, Hermitian matrices. So, what are the commutation relations? And in fact, these are important. All of these are Sx, Sy, Sz that you see there; they do not commute with each other. And there is a general commutation relation that S i S j is equal to i epsilon i j k S k. You can put an h cross here. So h cross equal to 1 is taken. But you should not take it all the time. If you want to write the h cross here, so it's i h cross, which means that they follow a nice cyclic relation where S x S y commutation will give you i h cross S z. If you reverse the order, that is, if you take S y S x, then you will get minus i S z. That is because the epsilon i j k, if it is not cyclic, it is called a Levi-Civita symbol. When these i j k are clockwise or taken in a cyclic order, they give you one. If two of the indices are the same, then it's equal to zero. And if you interchange two indices, that is, i going to k, k going to i, and j going to i, i going to j, and so on. So, if you change two of them, then you again bring back a sign. So, each time you change these cyclic properties of these indices i, j, k, you bring up a negative sign. If two of these indices become the same, you have it 0, and that is what the property is.

So, Sx, Sx is, of course, equal to 0, because i and j are the same indices. So, we write down the commutation relations for these sigma matrices, and the sigma matrices have the same commutation relations: sigma x sigma y is equal to i sigma z, and sigma z sigma x is equal to i sigma y. And sigma y sigma z is equal to i sigma x. If you reverse any of the indices or rather take it in the other direction, that is not sigma x y, sigma x sigma y, but it is sigma y sigma x, you are bound to get a minus sign here. These are all understood. So, sigma i sigma j in general is equal to i epsilon i j k sigma k, very similar to what I have spoken for any general spin matrices or any general spin operators or any general angular momentum operators. This has some special relations, which are the anti-commutation relations. And these anti-commutation relations are written as the sigma x, sigma y with a plus sign, this is equal to sigma y, sigma z with a plus sign, equal to sigma z, sigma x with a plus sign. This is equal to 0.

Some Key properties of the Pauli Matrices

1) For arbitrary vector $\vec{A}$.

$$\vec{\sigma}\cdot\vec{A} = \sigma_x A_x + \sigma_y A_y + \sigma_z A_z = \begin{pmatrix} A_z & A_x - iA_y \\ A_x + iA_y & -A_z \end{pmatrix}$$

2) Commutation relations.

$$[\sigma_x, \sigma_y] = i\sigma_z, \quad [\sigma_z, \sigma_x] = i\sigma_y, \quad [\sigma_y, \sigma_z] = i\sigma_x$$

$$[\sigma_i, \sigma_j] = i\epsilon_{ijk}\sigma_k \Rightarrow \sigma_i\sigma_j - \sigma_j\sigma_i = i\epsilon_{ijk}\sigma_k$$

3) Anticommutation relations

$$[\sigma_x, \sigma_y]_+ = [\sigma_y, \sigma_z]_+ = [\sigma_z, \sigma_x]_+ = 0$$

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0$$

So, the anti-commutation of a pair of them is always equal to 0. Now, what it means is that you have a sigma x sigma y plus a sigma y sigma x, which is equal to 0. You see that the commutation relation actually has a minus sign: sigma i sigma j minus sigma j sigma i. That is the commutation relation, which is equal to i epsilon i j k sigma k. However, these anti-commutations are written with a plus sign in between. And if you combine these commutation and anti-commutation relations, then you get something nice. Let us see for one of them. So, sigma x sigma y minus sigma y sigma x is equal to i sigma z, that's for the commutation relation. For the anti-commutation relation, we have sigma y sigma x plus sigma x sigma y equal to 0. If you add both of them, then it gives you 2 sigma x sigma y, which is equal to i sigma z. So, that tells you that sigma x sigma y is equal to i over 2 sigma z. So, if you take the two matrices and multiply them in this order, of course, you have to maintain the order because matrix multiplication requires respecting the order—you have a sigma x and sigma y. If you multiply them in this order, you get i by 2 sigma z. That is an important property. We also have the square of them equal to 1.

So, sigma x squared equals sigma y squared equals sigma z squared, which is equal to 1. But 1 should be written carefully because now they are no longer just 1, but there is an identity matrix, a 2 by 2 identity matrix. So, to generalize this property, it is like sigma i sigma j equals i by 2 sigma k, and you should keep the cyclic order of x, y, and z or i, j, and k. So, let us take a simple example for these properties, the one we just saw: the commutation, anti-commutation, and the square of them, each equal to 1. Let us say you are given to simplify sigma x, sigma y, sigma z, sigma y, sigma z, sigma x. What is the product of all these matrices? One way, of course, is to simply multiply them—write the 2 by 2 matrices and keep multiplying. But that is not the point. Here, what you can do is note that sigma z sigma x equals i by 2, sigma y, okay? And then you have a sigma y here, so this sigma y and this sigma y will give you 1 because sigma y squared equals 1. So, till now, we have sigma y sigma z, and then we have an i by 2 coming in, and the rest will be equal to 1.

$$\sigma_x \sigma_y - \sigma_y \sigma_z = i\sigma_z .$$

$$\sigma_x \sigma_y + \sigma_y \sigma_x = 0 .$$

$$2\sigma_x\sigma_y = i\sigma_z \Rightarrow \sigma_x\sigma_y = \frac{i}{2}\sigma_z . \Rightarrow \sigma_i\sigma_j = \frac{i}{2}\sigma_k .$$

4) $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$

Example 1.

$$\sigma_x \sigma_y \sigma_z \sigma_y \underbrace{\sigma_z \sigma_x}_{\frac{i}{2}\sigma_y} .$$

$$= \frac{i}{2}\sigma_x \underbrace{\sigma_y\sigma_z}_{\sigma_x} = \left(\frac{i}{2}\right)\left(\frac{i}{2}\right)\sigma_x^2 = \frac{i^2}{4} = -\frac{1}{4}\mathbb{1}$$

Now again, the sigma y sigma z will get another i by 2 here, and we will write that as sigma x squared because sigma x comes from here. So, this again becomes equal to 1. It becomes i squared over 4, which is minus 1 over 4—of course, minus 1 over 4 means minus 1 over 4 times an identity matrix. So, that is the answer for this. You may want to calculate—so this is another property—suppose theta by 2 (there's no particular reason you can write it as theta, but what I'm trying to show is how to actually calculate the exponentiation of the Pauli matrices, because this is important when you talk about rotation in spin space, if you're talking about s equal to half, that is, spin-half objects, and consider the rotations. This is going to be the rotation operator, and this can be simplified because you can, you know, sort of expand this. So, this is i theta by 2 sigma y plus 1 by 2 factorial theta by 2 squared sigma y squared, and so on and so forth: 1 by 3 factorial theta by 2 whole cubed sigma y cubed, and so on.

Now, remember that sigma y square is equal to 1 or this identity matrix that tells you that sigma y cube or any odd power of sigma y is simply equal to sigma y. So, you have the odd powers of sigma y will give you just sigma y, and even powers of sigma y will be free from sigma y because sigma y square is equal to 1. So, this exponential i theta by 2 sigma y, the even powers will combine to give you the expansion of cosine. So, this is like a cosine theta by 2 multiplied by this one, this one is often called sigma 0, that is the identity matrix, and plus i sigma y sine theta by 2 because the odd powers will give you sigma y and will give you the expansion of the sine function. So, this is nothing but equal to sigma 0 cosine theta by 2 plus i sigma y sine theta by 2. Then another property is that any 2 by 2 matrix, any arbitrary 2 by 2 matrix can be written as in terms of sigma 0 and the components of sigma, or let us write it as sigma i with i equal to x, y, and z. This is easy to show.

Let us take a general matrix with its components as m11, m12, m21, m22. This can be written as M11 plus M22 divided by 2 sigma 0, that is identity, and then you have M11 minus M22 divided by 2 sigma Z. And then m12 plus m21 divided by 2 sigma x plus m12 minus m21 with the i here. and by 2 sigma y, and this is nothing but these a0 sigma 0 plus a dot sigma where a's are the coefficients of sigma x, sigma y, and sigma z. So, a0 is equal to half of the trace of M, and of course, we have trace of sigma i equal to 0 and sigma i squared equal to 1. Alright, so these are in brief some of the properties of the Pauli matrices, which are the spin angular momentum corresponding to spin half, and these have certain relations. These relations are important in solving problems or, rather, they help us to formulate the properties of these spin half objects. So, let us go on to a new topic known as Perturbation Theory. In general, they form a class of solutions known as the approximate methods.

5. $$e^{i\frac{\theta}{2}\sigma_y} = 1 + \frac{\theta}{2}\sigma_y + \frac{1}{2!}\left(\frac{\theta}{2}\right)^2\sigma_y^2 + \frac{1}{3!}\left(\frac{\theta}{2}\right)^3\sigma_y^3 + \cdots$$

$$\sigma_y^2 = \mathbb{1} \Rightarrow \sigma_y^3 = \sigma_y .$$

$$e^{i\frac{\theta}{2}\sigma_y} = \cos\frac{\theta}{2}\,\underset{\sigma_0}{\underset{\sim}{\mathbb{1}}} + i\sigma_y \sin\frac{\theta}{2} = \sigma_0 \cos\frac{\theta}{2} + i\sigma_y \sin\frac{\theta}{2} .$$

6. Any 2×2 matrix can be written in terms of σ_0 and σ_i $(i = x, y, z)$.

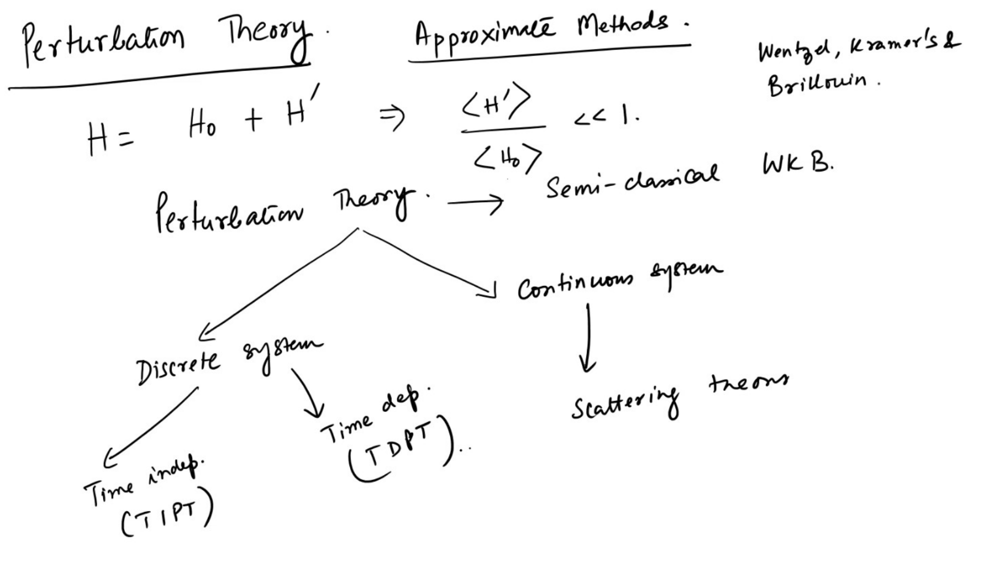
$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \frac{m_{11}+m_{22}}{2}\sigma_0 + \frac{m_{11}-m_{22}}{2}\sigma_z + \frac{m_{12}+m_{21}}{2}\sigma_x + i\,\frac{m_{12}-m_{21}}{2}\sigma_y = a_0\sigma_0 + \vec{a}\cdot\vec{\sigma}$$

$$a_0 = \frac{1}{2}\mathrm{Tr}[M]$$

As the name suggests, it is not an exact solution, but it is a solution which is approximately done. How good that approximation is will depend upon how controlled the approximation is. and how well we can explain these situations as a limiting case or when this perturbation goes to zero, whether we recover the results that we have obtained for the unperturbed problem. In general, very few problems in the real world in quantum mechanics or otherwise are exactly solvable. There are very few problems, and in fact, we have encountered this in classical mechanics where just barely two-particle problems are solvable. Even though that looks difficult because you need six coordinates for two particles and you have only three momentum conservation equations and one energy conservation equation. So you have four equations but six unknowns. Somehow we could solve this by going to the center of mass frame and establishing a relationship between the lab frame and the center of mass frame. We have seen such solutions. Going to three-body becomes very difficult and most of the time it is unsolvable. However, statistical mechanics routinely takes into account Avogadro's number of particles and presents a statistical description which seems to work very well.

Here we are not talking about that we are simply talking about a Hamiltonian which has got two parts let us say the H0 and some part which is H prime and of course we know that you know this H0 and H prime they do not commute and so H0 and H0 commute as well. But why these approximate method works is that when you have these H prime, the expectation value of that as compared to this H naught, that is much, much smaller than one, okay. Or you can say that the expectation value of H prime and the expectation value of H naught, it's much smaller than H naught. So there is an ambiguity that you may notice that we are talking about expectation values without knowing the eigenfunctions of these H prime because certainly we do not know the exact solution of H. Then how do we know the how to actually take this expectation value that is what are the eigenfunctions used for evaluating the numerator for this problem. And the answer is that we are still using the eigenfunctions of the unperturbed problem that is those of H0 and the reason that they work is that these perturbation does not take you far away from the unperturbed problem. If it does then of course this perturbation theory will not work and the approximation methods that we are talking about will not work. But it works in most of the cases where the strength of the perturbation is small. Say we talk about hydrogen atom and then you can use an electric field or a magnetic field

which is weak enough such that it doesn't cause, you know, the hydrogen atom to ionize, that is, it doesn't get converted into an ion so that the perturbation term or the energy associated with the electric field or the magnetic field, it's much smaller than the minus 13.6 electron volt. And in those cases, it will work and we'll see the problems which are called a Stark effect and Zeeman effect there. Okay, so we use a perturbation theory and one can actually build up a flow chart that this is a perturbation theory and so this has semi-classical thing which is called as WKB approximation and WKB, it is Wentzel-Kramers and Brillouin. So that's semi-classical approximation and then one can use it in continuous systems which gives you scattering theory and the scattering, these scattering matrices and this partial wave approximation to get you the scattering cross section. And we are mostly interested in this part where we talk about discrete system or quantized systems with discrete energy levels.



And for that, we have a time independent perturbation theory and we have a time dependent perturbation theory. Okay, so this will call as a TIPT, time independent perturbation theory and this will be called as a TDPT, time dependent perturbation theory. Okay, so it is going to be a little bit of mathematics, but that is because I am deriving everything from first principles. One should not be bothered about the mathematics. The final result is usually very simple and easy to apply to systems that you want to. So, let us do this TIPT or the time independent perturbation theory and build up the basis. So, you have a Hamiltonian which can be written as H0 plus H prime. But now, to explain the smallness of this H prime, let me put a lambda and so that I can actually do a perturbation theory in orders of lambda, okay. And so, lambda is basically a term which is much smaller than 1, then we do not have to talk about the perturbation. expectation values of H prime and H naught that we have talked about before. So, we have introduced by hand a parameter which is much smaller than 1 and you have to also acknowledge the fact that this problem is known that is this is the unperturbed problem and these 0 actually refers to the unperturbed problem or the zeroth order problem.

Or, as I said, 0th order, and so this is known to us. This solution is known, which means that En0 and Psi n0 are known, and n is the quantum number, whatever it may be in a given case. It could be the angular momentum. In some other case, it could be just a linear momentum k, or it could be just energy or a quantum number that explicitly expresses energy, like what we have seen in the harmonic oscillator or even in the hydrogen atom, and so on, okay. So, we want to solve for this En and Psi n. These are not known. And what are they? They are the eigen solutions of H, which means the Hamiltonian has these as the eigenvalue and the eigenvectors, and that is not known. So, what we do is the following: we write the one that we do not know as the one that we know, plus some unknown coefficients where k is not equal to n and cnk, which is a function of lambda, and then Psi k0. So, whenever you see a 0, you understand that 0 corresponds to the unperturbed problem or the 0th order problem. And the C nk is basically the coefficients that are unknown, and C nk is a polynomial in lambda.

So, because it is a polynomial in lambda, we can talk about 0th order, 1st order, 2nd order, and so on. And these corrections to the energy and to the wave function would only come from the 1st order. And for some reason, the selection rule excludes the 1st order correction from existing. It may happen at times, then we would resort to the 2nd order calculation. Now, this will give you, as I said, the different orders of the perturbation theory.

TIPT

$H = H_0 + \lambda H'$ $\quad \lambda << 1$

0: unperturbed problem (zeroth order).

$H_0 |\psi_n^0\rangle = E_n^0 |\psi_n^0\rangle$

$\rightarrow$ known. $(E_n^0, |\psi_n^0\rangle)$

$H \xrightarrow{\text{eigensolutions}} E_n, \; |\psi_n\rangle \rightarrow$ Not known.

$|\psi_n\rangle = |\psi_n^0\rangle + \sum_{k \neq n} C_{nk}(\lambda) |\psi_k^0\rangle.$

$C_{nk}(\lambda)$ is a polynomial in λ.

So, in the 0th order, let us see what happens. That is a known problem, but we still want to get it from the perturbation expansion. So, we have lambda equal to 0. So, H is simply equal to H0. That tells you that Cnk as 0 is equal to 0, and the total Psi n has to be equal to Psi n0, which is understandable because you have H equal to H0. The entire Hamiltonian has eigenstate Psi n, which should be nothing but the unperturbed eigenstates of the system. Now, Cnk lambda has a sort of solution which is Cnk 1 plus lambda square Cnk 2 plus lambda cube Cnk 3 and so on, okay. And the energy will have a form which is En 0 plus lambda En1 plus lambda square En2 plus lambda cube En3 and so on. This is, of course, known. So, if you substitute this into the Schrodinger equation, all these things that we have learned, then we have H0 plus lambda H prime.

That's the total Hamiltonian. And this will be on this Psi n0. That's the unperturbed plus you have these lambda Cnk 1. So, this corresponds to the zeroth order. lambda nk1 Psi k0 —well, that 0 I am writing it without a bracket—and so k is not equal to n, and you have k not equal to n again, and you have a lambda square Cnk2, and we have a Psi k0 and so on. Okay, so this is the left-hand side, and the right-hand side would be En 0 corresponding to Psi n 0 and plus lambda En 1 plus lambda square En 2 and so on. Psi n0 plus lambda Cnk 1. So, we are writing the 0 without a bracket in the superscript, but all other orders we are writing it with 1. Anyway, your Cnk 0 equal to 0, so that is not there, but these in the eigenfunction and eigenvalues will not put a bracket. So, we will simply write it like that to make this unperturbed problem appear distinct. So, this is Psi k0, and then we have a k not equal to n again, k not equal to n here, lambda square Cnk2, and we have a Psi k0 and so on, okay.

Zeroth order $\quad H = H_0$.

$\lambda = 0$

$C_{nk}(0) = 0$

$|\psi_n\rangle = |\psi_n^0\rangle$

$$C_{nk}(\lambda) = \lambda C_{nk}^{(1)} + \lambda^2 C_{nk}^{(2)} + \lambda^3 C_{nk}^{(3)} + \cdots$$

$$\& \quad E_n = \underset{\downarrow \text{ known}}{E_n^{(0)}} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \lambda^3 E_n^{(3)} + \cdots$$

$$(H_0 + \lambda H')\Big[\underset{\downarrow \text{ zeroth order}}{|\psi_n^0\rangle} + \sum_{k\neq n} \lambda C_{nk}^{(1)} |\psi_k^0\rangle + \sum_{k\neq n} \lambda^2 C_{nk}^{(2)} |\psi_k^0\rangle + \cdots\Big]$$

$$= \left(E_n^0 + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\Big[|\psi_n^0\rangle + \sum_{k\neq n} \lambda C_{nk}^{(1)} |\psi_k^0\rangle + \sum_{k\neq n} \lambda^2 C_{nk}^{(2)} |\psi_k^0\rangle + \cdots\Big]$$

Now, you have terms which are proportional to lambda, proportional to lambda squared, and so on, on both sides, and lambda is a small parameter. So, if this equation has to be satisfied, then we should equate equal powers of lambda on both sides of the equation. And that gives us, so match powers of lambda on both sides. So, we have H0, Cnk1, Psi k0. So, the first, the zeroth order is known. So, we are starting from the first order. So, this plus H prime psi n0 equals to en0 and k not equal to n, and we have a cnk1 and a psi k0 plus en1 psi n0 and so on. Okay, so you see that we have just equated the first power of lambda on both sides. So, if you recognize that H0 psi K0, that is the problem that you are familiar with, that is, this is known. And psi k0, then we have this En1, which is the first-order correction in energy, that acting on psi n0, this gives you k not equal to n.

Ek0 minus En0 and Cnk1 psi k0 plus h prime psi n0. Now, it is a standard trick to extricate this or rather to find this En1, you have to take an overlap with inner product with Psi n0 or take a conjugate of Psi n0, the bra Psi n0. And that will give you En1. This is equal to, so this is psi n0. Now, if you take an inner product on the left-hand side, the psi n0, psi n0 will give you 1. So, you are left with En1. That is the first-order correction to energy on the left-hand side. The first term on the right-hand side, when you take an overlap with psi n0, it gives you 0. This gives you 0 because n is not equal to k, and each of these psi n 0s and psi k 0s, this n and k are different labels of the wave function corresponding to the unperturbed problem, and each one of those psi n 0 is orthogonal to the psi k 0 when n is not equal to k. So, that eliminates the first term on the right-hand side, and you are left with this, and even though the calculation seems a little, you know, sort of messy, but what you get is that the first-order correction to energy is nothing but the expectation value of the perturbation term that you have started with. The perturbation term is here.

H0 plus lambda H prime. The lambda H prime is the perturbation term. The lambda is just a dimensionless number. And we are like when you have H prime, you know, H prime there. But it cannot solve the problem for H. That is, you can't find the eigenfunctions for H and the eigenvalues for H. But you can treat the problem perturbatively and can arrive at a solution for the first-order correction in energy, which is given by the psi n0 H prime psi n0. So, the matrix elements of H prime have to be calculated within the eigenstates of the unperturbed problem, and that is going to give you the first-order correction to energy. So, that tells you that at this level, if your perturbation is really weak, then you are doing fine if you just simply take the total En to be equal to En0, which is the unperturbed problem, and the En1 that you find out. So, En1 is here, okay. En1 can be found out like this. What about the wave function? Let us see the wave function. And so, to calculate the wave function, what we do is that we take an inner product with some, with a dummy psi m0, psi m0, okay. So, what we get is that we get an En0 minus Ek0 that's coming from the first term in, I mean, inner product of psi m 0 and m not equal to n.

Match powers of λ on both sides.

1st order

$$H_0 \sum c_{nk}^{(1)} |\psi_k^0\rangle + H'|\psi_n^0\rangle = E_n^0 \sum_{k\neq n} c_{nk}^{(1)} |\psi_k^0\rangle + E_n^{(1)} |\psi_n^0\rangle$$

$$H_0 |\psi_k^0\rangle = E_k^0 |\psi_k^0\rangle.$$

$$E_n^{(1)} |\psi_n^0\rangle = \sum_{k\neq n} (E_k^0 - E_n^0)\, c_{nk}^{(1)} |\psi_k^0\rangle + H'|\psi_n^0\rangle.$$

$\langle\psi_n^0| \cdots = 0$

$$\boxed{E_n^{(1)} = \langle\psi_n^0 | H' | \psi_n^0\rangle.} \Rightarrow E_n = E_n^0 + E_n^{(1)}.$$

If m is not equal to n, the left-hand side does not give you anything because if you use this equation and take an overlap with psi m 0 with m not equal to n, then this term gives you 0. And this term, of course, will give you a result because m can be equal to k and m may not be equal to k. But it turns out that if m is not equal to k, then this term does not give you anything. And now we get bra psi m0 and H prime psi n0, which means the matrix elements of H prime will have to be calculated between two different orthogonal eigenstates of H, and that's what we do there, and we get this result as En0 minus Ek0, and we have a Cnk. 1 delta k m, and that's equal to psi m 0 H prime psi n 0, and this is precisely what you wanted, the Cnk1, but now this will click only when k is equal to m. So, we have a Cnm1. This is equal to psi m0 H prime psi n0 and divided by En0 minus Em0. Okay, so this is the correction to the first eigenfunction.

Inner product $\langle \psi_m^0 |$. $m \neq n$

$$(E_n^0 - E_k^0) C_{nk}^{(1)} \delta_{km} = \langle \psi_m^0 | H' | \psi_n^0 \rangle$$

$$\boxed{C_{nm}^{(1)} = \frac{\langle \psi_m^0 | H' | \psi_n^0 \rangle}{E_n^0 - E_m^0}}$$

$$|\psi_n\rangle = |\psi_n^0\rangle + \sum_m C_{nm}^{(1)} |\psi_m^0\rangle$$

So what it means is that you have a psi n, which is equal to a psi n 0 and plus a Cnm1. And then, of course, you have these. The one that we have written down here is this psi k 0 or 0. So, Cnm and psi m0 and so on. So, of course, we have to sum over this m, where m can be any of these indices. Let us see what happens in second order because as you go higher up in the order, it becomes more and more complicated. Now you could ask me this question: if this perturbation is weak enough, do we really need to go into the second order? Yes, you do not. You are absolutely correct. But then for some reason, you may have the first-order correction to the energy to be equal to zero, in which case you have to go to second order. It is an obligation because otherwise, you do not get any correction in the energy due to the perturbation. So, that is an important thing. So, the second-order correction. And so this is H0, and then we have k not equal to n, and we have a Cnk 2, and we have a psi k0 plus H prime, and then we have k not equal to n, and we have Cnk1 psi k0 equal to En0.

And then k not equal to n, and you have C n. So, we have this. So, yes, that is correct. So, Cnk 2, and we have a psi k 0. And so, we are taking up to lambda square. So, only matching terms on left and right up to lambda square. And so, this En0, and then you have an En1. Then k not equal to n, En1 is of course known, Cnk1 Psi k0 that is known as well, and En2 Psi n0. So, we want to find En2 and Cnk2. So, again, in order to do that, you have to take an overlap with psi n 0. So, this is psi n 0. So, take an overlap and so on. So, that makes this En 2 free, and you have k not equal to n, a little sort of expression that you have to get to, but there is no option. So, we have a psi n 0 and psi k 0, which again gives you 0 because they are orthogonal, and then we have this plus.

psi n0 h prime and now we put in these values of Cnk1 which are psi k0 h prime psi n0 divided by En0 minus Ek0. We have already assumed that n is not equal to k so the denominator is never going to blow up. And we have a psi k0, okay? So, that's the overlap for this term that you see there, the second term on the right, and then you have minus En0, so that is k not equal to n, Cnk2. I mean, these look a little boring, but then they are instructive because the final result that you arrive at is quite simple. And again, this is equal to 0 because they are orthogonal. And we again have these psi n 0, H prime, psi n 0, and so on. And then you have a term which is k not equal to n. And there is a term there, but this is equal to 0 again. because H prime so psi n is not an eigen state of I mean so this will give you a term which is I mean not this one so this one is okay but then you have a term which is psi k 0 H prime psi n 0 and then you have a En 0 minus Ek 0 and then you have a psi n 0 and a psi k 0 which is equal to 0 again.

2nd order. upto λ^2

$$H_0 \sum_{k\neq n} C_{nk}^{(2)} |\psi_k^0\rangle + H' \sum_{k\neq n} C_{nk}^{(1)} |\psi_k^0\rangle = E_n^0 \sum_{k\neq n} C_{nk}^{(2)} |\psi_k^0\rangle + E_n^{(1)} \sum_{k\neq n} C_{nk}^{(1)} |\psi_k^0\rangle + E_n^{(2)} |\psi_n^0\rangle$$

Take an overlap. $\langle\psi_n^0|$ ·

$$E_n^{(2)} = \sum_{k\neq n} E_k^0 C_{nk}^{(2)} \underbrace{\langle\psi_n^0|\psi_k^0\rangle}_{=0} + \langle\psi_n^0| H' \sum_{n\neq k} \frac{\langle\psi_k^0|H'|\psi_n^0\rangle}{E_n^0-E_k^0}|\psi_k^0\rangle$$

$$- E_n^0 \sum_{k\neq n} C_{nk}^{(2)} \underbrace{\langle\psi_n|\psi_k^0\rangle}_{=0} + \langle\psi_n^0|H'|\psi_n^0\rangle \sum_{k\neq n} \frac{\langle\psi_k^0|H|\psi_n^0\rangle}{E_n^0-E_k^0} \underbrace{\langle\psi_n^0|\psi_k^0\rangle}$$

So, you are left with very little number of terms for the second order correction in energy. So, En 2 is simply equal to a k not equal to n and we have a psi n 0 H prime psi k0 and psi k0, H prime psi n0 and this is divided by En0 minus Ek0 and that's equal to mod square of, so k not equal to n. And we have these psi n0 H prime, psi k0 and a mod square of that divided by an energy denominator which does not diverge, okay? So, this is the first order energy correction and so on. I will sort of simply write down this final result for the wave function, which can be easily again determined by the same prescription that you have seen that is, you know, taking an overlap with psi m0 and all that. So, Cnm2 is H equal to 1 divided by Em0 minus En0.

$$E_n^{(2)} = \sum_{k\neq n} \frac{\langle\psi_n^0|H'|\psi_k^0\rangle\langle\psi_k^0|H'|\psi_n^0\rangle}{(E_n^0 - E_k^0)}$$

$$= \sum_{k\neq n} \frac{|\langle\psi_n^0|H'|\psi_k^0\rangle|^2}{E_n^0 - E_k^0}.$$

$$C_{nm}^{(2)} = \left(\frac{1}{E_m^0 - E_n^0}\right)\left[\sum_k \frac{\langle\psi_m^0|H'|\psi_k^0\rangle\langle\psi_k^0|H|\psi_n^0\rangle}{E_k^0 - E_n^0} - \frac{\langle\psi_n^0|H'|\psi_n^0\rangle\langle\psi_m^0|H'|\psi_n^0\rangle}{E_n^0 - E_m^0}\right]$$

So, this and then there is a sum over k psi m0 H prime psi k0. and psi k0 H prime psi n0 and minus, so this divided by another energy denominator, which is Ek0 minus En0. And minus Psi n0 H prime Psi n0 Psi m0 H prime Psi n0 and En0 minus Em0, okay? So, most of it we have derived, we have derived the second order completely up to second order, the energy correction and correction to the wave function. And these will be required for us to proceed farther because we'll apply it to different cases. And as I said that the first order energy correction being 0, you have no option but to go to the second order to see the correction. It cannot happen that just because the first order is 0, there is no correction that occurs. And in fact, we will see that that is the case. And this problem is a little more complicated for degenerate perturbation theory where once again the degeneracy means that you have different eigenstates corresponding to the same energy or the same eigenenergy.

And if that happens, you can have this Ek0 and En0 to have the same value for a given system. And in that case, this formalism doesn't work because it doesn't make sense that the second order correction to energy or second order correction to the wave function blows up. Then we are not applying the perturbation theory in a controlled fashion. So a new formalism has to be developed. We thought we'll be able to finish that, but nevertheless, we'll do that in the next class where we first talk about the degenerate perturbation theory and then hopefully apply it to these hydrogen atom in a weak electric field, which is called as a stark effect. So we'll stop here with this. We've started with the proximate methods or talking about the perturbation theory, specifically the time independent perturbation theory and have derived the first order correction to energy, second order correction to energy, first order correction to the wave function, second order correction to the wave function. And they have in general, they have complicated terms. But you see everywhere we have this psi k0, psi m0, psi n0, Em0, En0. These are known for a given problem. Suppose we give particle in a box, then these are the different eigenstates of the particle in a box, the ones that we have derived. But, so these have to be, you know, used in order to calculate the corrections. So, we build up our theory from all these quantities that we already know, that is, we are aware of these psi m0, psi n0, em0, en0 and all that.

So, these are simply the corrections to that in terms of all these eigen solutions of H0. We will stop here for today and we will carry on with degenerate perturbation theory on the next class.