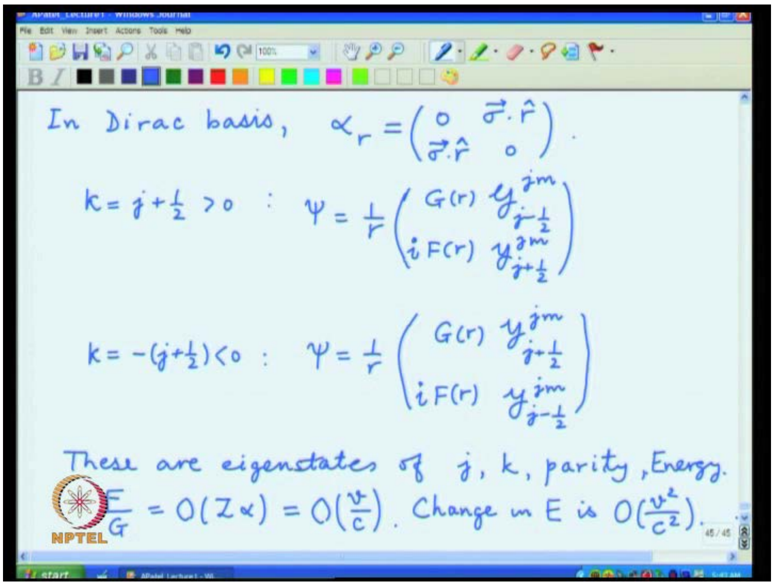
Relativistic Quantum Mechanics
Prof. Apoorva D. Patel
Department of Physics
Indian Institute of Science, Bangalore

Lecture - 9
Non-relativistic reduction, The Foldy-Wouthuysen transformation

In the last lecture we explicitly went through the solution to the hydrogen atom problem, finding an exact result in terms of hypergeometric functions appearing as the Eigen states and corresponding energy levels described by quantum numbers. Now, some features of these particular solutions relevant to further discussions are related to the contributions of various terms which appears in this equation.

(Refer Slide Time: 01:12)



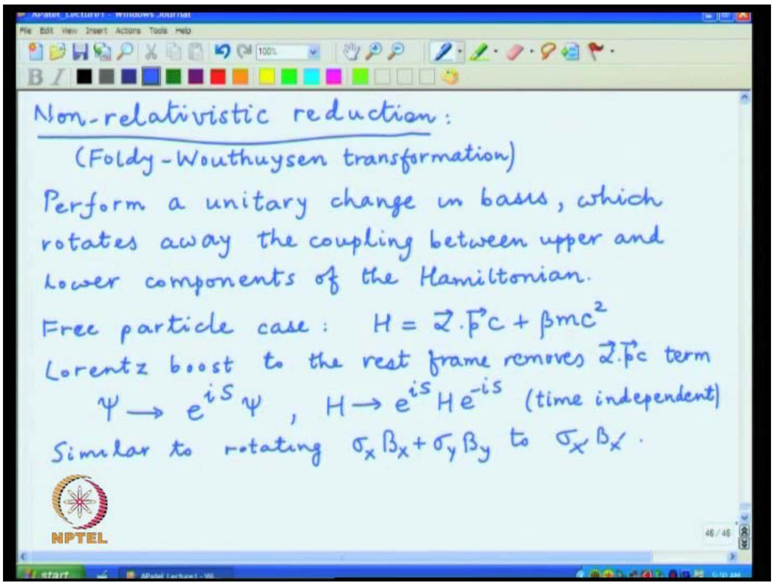
In particular, the ratio of lower components to upper components in these equations is the ratio of this functions F over G, and one can substitute the answers which were obtained through the power series solutions. And, it shows that this ratio is of the order of Z times alpha, essentially the fine structure constant; or, one can express it as the ratio of velocity to the speed of light. And, this is the magnitude generated by the alpha dot p term in the Hamiltonian which is the off diagonal term.

And, the diagonal term is m c square, and the ratio of off diagonal to diagonal terms is v over c, and that produces the mixing to the lower components starting from the upper components, which is exactly of this particular size. And, the effect on the energy coming from this correction, the correction to the over function is first order in v over c.

The correction to energy requires a second order perturbation theory, in the sense that the upper component mixes into the lower component, but it has to come back to the upper component to produce an Eigen state. And, that second order contribution gives a term which is order v square by c square compared to the rest mass, and that is the standard Rydberg formula which gives a corrections of the order of z square, alpha square on top of the rest mass energy.

So, change in energy is order v square by c square. And, this particular pattern has a long history in trying to connect non relativistic solutions to the relativistic problems as the power series in a, the, this same ratio v over c. And, today we will discuss about that formulation in detail.

(Refer Slide Time: 03:46)



So, this is the so called non relativistic reduction of the Dirac equation, is also goes under the name of the two people who gave a complete formulation. So, it is, goes as the name as Foldy-Wouthuysen transformation. And, what it does is, it performs a unitary change in basis, so that the coupling between the upper and lower components in the Hamiltonians gets rotated away.

And, once that is done the Hamiltonian becomes effectively block diagonal, and one can look at the one part or the other as the solution corresponding to the positive and the negative energy components. The upper and the lower components are essentially described by the sign of the energy in the Dirac basis, so to illustrate this, first just look at the free Hamiltonian. So, free particle case where H is nothing but alpha dot p c plus beta m c square; the beta m c square part is the diagonal, alpha dot p c is the off diagonal one and we want to rotate it away. It is a unitary change of basis.

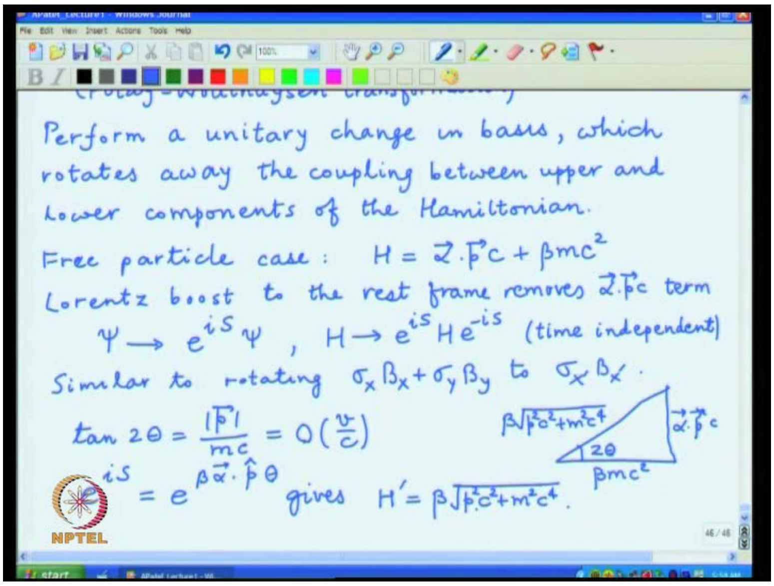
And, in this particular case, the unitary change of basis just turns out to be a Lorentz transformation. Because, as we can easily see, this is a particle moving with a particular momentum and you can go to the rest frame of the particle in which the momentum vanishes, and so the, whatever survives is the diagonal term. And, that is all one can get by just doing a simple Lorentz boost, described by whatever the momentum of the particle is. So, Lorentz boost to the rest frame removes alpha dot p term.

Now, just look at this thing in a mathematical sense. What we are going to do is, construct a transformation which takes psi to e raise to i S times psi. It is a unitary transformation that is why this e raise to i S factor is written explicitly. And, under this same transformation the Hamiltonian will become e raise to i s times, H times e raise to minus i s. And, under this scheme the equation will remain form invariant. And, in particular, this transformation is a time independent.

So, what we are constructing here is a very similar to taking a simple spin aligned in one particular direction and rotating it to another direction, the fact that it can be interpreted as Pauli matrices in involving spins, because the Pauli matrices over the same kind of anticommutation relations which alpha and beta also obey. So, the mathematical form turns out to be essentially the same things.

So, if I have sigma x, I can perform a rotation around the third component which is a z axis, and take this thing to a new axis where the term just becomes sigma x prime, times B x prime, in a sense you rotate to a basis where B y prime is equal to 0. And, this is just a simple trigonometric factor which takes the coordinates x, y, to x prime, y prime.

(Refer Slide Time: 10:53)



And, in case of the free particle situation, one can express as a simple triangle, right angle triangle form where one side is beta m c square, the other side is alpha dot p times c. There is a angle and we want rotation. And, the rotation obviously takes this thing to the total field which is the square root of the sum of this two terms, which in our case is p square c square plus m square c raise to 4. And, this rotation angle, this then determined by the properties of these 2 matrices.

The particle which we are dealing with is a spin half particle. So, the rotation angles actually appear are half angles. And, one can easily verify that the angle needed here is a described by the relation that tangent of 2 theta is a magnitude of p divided by m c, and in particular this object is order v by c, the typical non relativistic expansion factor. And, the transformation which will do this particular rotation is this e raise to i s term which is a now expressed in terms of this explicit angles.

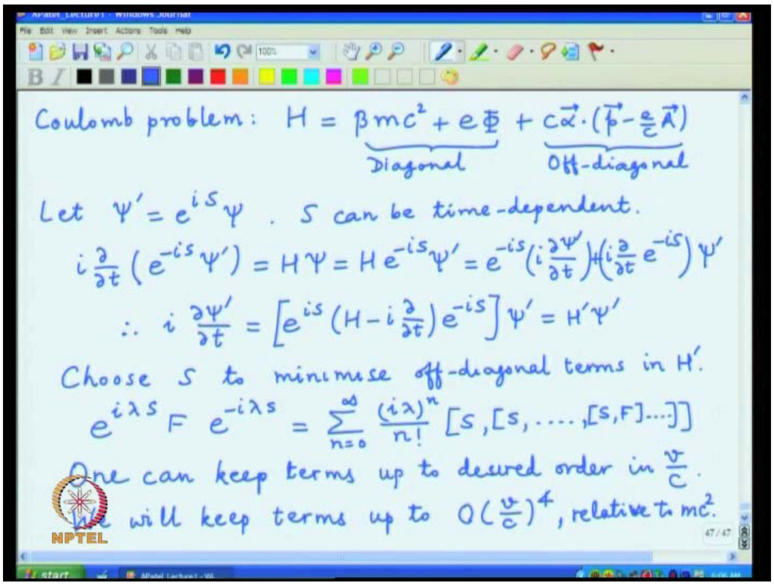
And, the formula is e raise to i s is exponential of beta alpha dotted with the unit vector in the momentum direction, and theta is a rotation angle. This is a unitary matrix because the product here, beta times alpha, is antiHermitian. Beta and alpha themselves are individually Hermitian, but the product when you take a Hermitian conjugate it produces alpha times beta. And, because of the anticommutation it produces a negative sign denoting anti Hermitian behavior.

And, that will gives now, the new Hamiltonian which I can just write H prime, and its structure is explicitly diagonal. There is only one matrix beta appearing inside there, and it corresponds to the rest frame of the particle, and the total energy is described by this energy momentum dispersion relation which is what it should be.

So, this is a simple situation of a free particle. We want to look at the same kind of analysis in the case of the hydrogen atom solution where this analysis cannot be done exactly as in the case of free particle problem. The rotation can be performed, but it does not diagonalize the whole matrix exactly. So, we will do it in steps, iteratively rotating away one term after the other.

And, that produces a description of a Hamiltonian as a series. The largest term is of course, m c square, but then all the sub leading terms will appear as different kind of operators involving the momentum as well as the electromagnetic field. And, the purpose of doing that is you can look at it as a systemic expansion in a powers of v by c. And, various terms will appear at different orders in v by c, and one can try looking at their physical interpretation. And, that helps in understanding the problem in a more detail because such a perturbative expansion is helpful in many other problems also. And, one can see the effects of relativistic corrections appearing term by term. So, for that we need some more machinery.

(Refer Slide Time: 15:53)



So, let us go to the coulomb problem. And, one can write the Hamiltonian also in the same fashion separating the diagonal and the off diagonal parts. So, the diagonal part is the rest mass energy plus the electromagnetic potential, and then there is a off diagonal term which is alpha dotted with the covariant momentum. So, this is diagonal, and this is off diagonal.

Our purpose is the same. We want to rotate away this off diagonal term, and see how it will fit back to the diagonal part giving a particular description, but exact diagonalization does not work in this way. So, we will construct a systematic expansion involving a, powers of this term, alpha dot p. For that we will now construct a transformation. So, which has the same form as we used earlier, but now we will allow a little more general case, where s can be time dependent. That is necessary because the electromagnetic fields which are introduced here can have some time dependence.

So, we need now the equation satisfied by psi prime, and then we will do a power series expansion of that particular equation. So, we will start with the usual equation which is the one for psi, and convert it into equation for psi prime. And, that is time derivative of e raise to minus i s of psi prime which happens to be equal to psi, and that now has the structure that it is H psi which is also H e raise to minus i s psi prime. And then, one can now evaluate this particular derivative explicitly which gives a e raise to minus i s times, i del psi prime by del t. And, the derivative acting on the first term will give, i del by del t e raise to minus i s, acting on psi prime.

And, the result of this jugglery is to obtain a equation just for, i del psi prime by del t, and then we can easily see by just simple rearrangement of terms that it has the structure. It is a, i del psi prime by del t, is a, e raise to i s. You have to multiply by this e inverse factor on to the other side, and put together all the other terms involving this H and i del by del t. So, it is, H minus i del by del t, which acts on, e raise to minus i s, and then this whole block acts on psi prime, and we can call this H prime psi prime.

So, H prime has this particular structure, e raise to i s and e raise to minus i s. The extra term is this time derivative which acts only on this e raise to minus i s factor and not on the psi prime, that is understood in this notation. And, that extra term is needed whenever the Hamiltonian is explicitly time dependent. And, now the problem is to find a suitable function s, so that H prime will have this off diagonal term minimized. And, so, choose S to minimize off diagonal terms in H prime, and we will do this iteratively.

We will take one term and try to rotate it away, then take another term whatever is left rotate it away, etcetera. But, to be able to do that we need a general way of constructing this particular product of, a bunch of operators, and for that it is a kind of convenient to develop a general formula, if there is some operator say F, here how it rotates, by multiplying by e raise to i s on one side and e raise to minus i s on the other side.

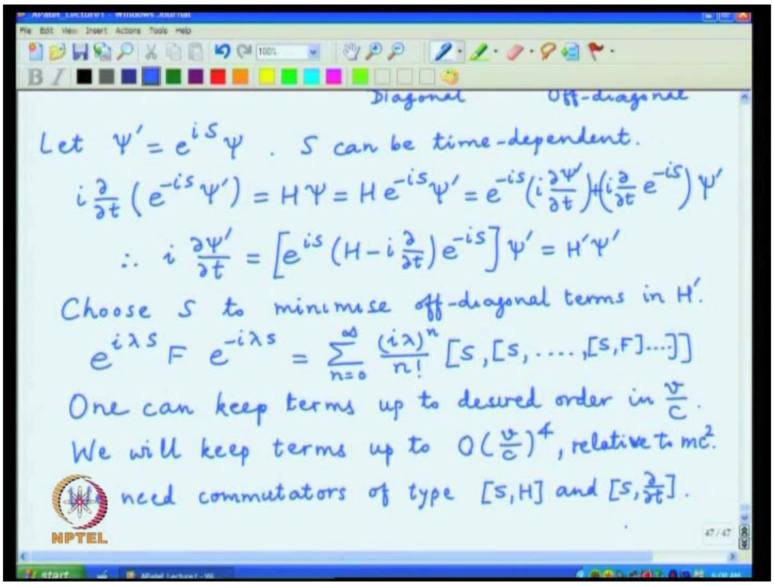
And, that can be expressed by a simple trick of doing a Taylor series expansion of a slightly different looking function which is e raise to i lambda s, F, e raise to minus i lambda s. And, we will express this object as a Taylor series in lambda. And, the result is then is a simple factor of i lambda raise to n by n factorial, and then a whole series; the leading term on the series just corresponds to lambda equal to 0 which just happens to be F. But, all the sub leading terms are derivatives of this thing evaluated at lambda equal to 0.

And, now one can see what happens when one takes a derivative of such a quantity. If the derivative acts on the first factor, you will get the same exponential, times s. If it acts on the second factor, you will get the same exponential, but now it is minus s. And so, when you combine the two terms together you will get, s f minus f s, which is nothing but the commutator of s and f. And, so, first term is f, the second term is s commutator of f. And, one can repeat this thing. Every time there is a derivative which is a taken as a extra commutator generated because of the this two factors of e raise to i lambda s and e raise to minus i lambda s.

And, the whole scheme can be written then as a set of nested commutators of s and f. It is understood that the 0th order term is just f itself. And, this is a compact enough formula to be able to rewrite this particular Hamiltonian to a desired order. We know that this s, as in the case of non relativistic expansion is of the order of v over c. So, we can go to a desired order of the same expansion parameter of v by c.

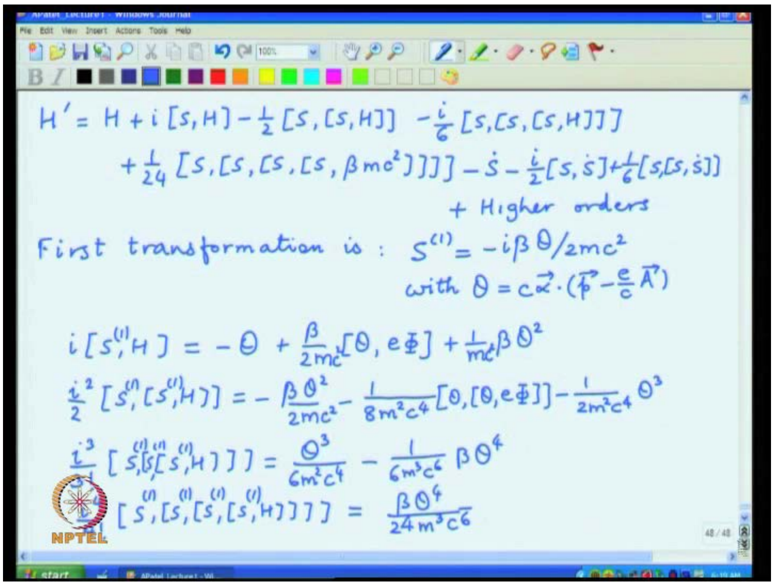
And, in this particular case of hydrogen atom we will keep terms upto order v by c raise to 4 compared to the rest mass term which is m c square. And, this is exactly the same order to which we worked out, the analogy between hydrogen atoms spectrum. It is the m rest mass is the leading term order v square by c square provided the Rydberg formula, and v by c raise to 4 will give the fine structure of the energy levels. And, in principle, one can go further down, but this much is enough to understand most of the different features which appear in a relativistic solution to the problem.

(Refer Slide Time: 27:23)



So, we need a general form which now can be written down as all this particular factors. And, we need basically commutators with both the different types of operators, S commutator with H, and S commutator with this time derivative operator. And, since the time derivative acts only on S, this is equivalent to taking a del s by del t with a negative sign. But these things are nested inside this whole construction, and we have to evaluate all these things term by term.

(Refer Slide Time: 28:37)



So, upto the order v by c raise to 4, let me just write down the expression which is needed. And, that involves the form where H prime is the original term plus i [s, H] minus half, second commutator, the third commutator, then in the fourth commutator we will keep only in the leading part of H which is beta m c square because the other terms kind of drop out.

And then, the same things now with the time derivative part- the first commutator is just s dot, the second commutator produces s commutator with s dot, and then the third term produces the triple commutator of s and s dot. And, keeping upto the fine structure level expression we do not need anything more than this.

And now the exercise is to evaluate all those things carefully, and then choose a value of s which cancels the unwanted terms. So, we will start in the same way, in case of free particle case where the first rotation is a just determined as minus i beta O by 2 m c square, where this O is the off diagonal part of the operator. And, we have taken a first order expression in the sense that we are not worried about taking the angle with its explicit form in terms of a tangent function, but just a linear expression assuming this rotation is small, so equivalently tan theta is roughly the same as theta.

So, this is exactly the same structure, and now one can see how the cancelations occur. If you perform this rotation we except this alpha dot p term to go away, but some other terms will survives. And, we will have a expression which will of diagonal term, but whose contribution will be suppressed by one more order of v by c compared to what was provided by alpha dot p. And, that can be just easily worked out by evaluating all the terms which are listed over here. Let us just do that.

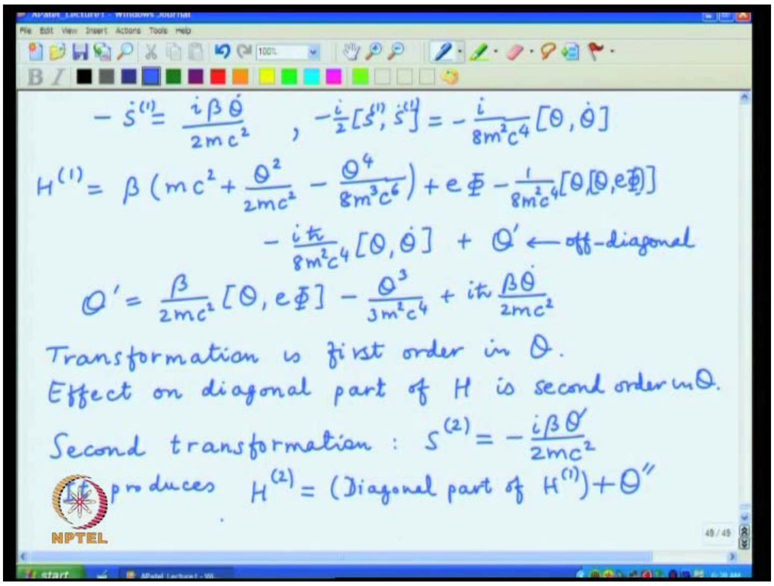
The first is a commutator of this object with the Hamiltonian. It is essentially the matrix structure alpha; it anti commutes with a beta, and also the different components of alpha itself. And, one case easily see what it produces when it anti commutes with beta, the two terms contributes the same thing. Coefficient of beta is m c square, and that just produces the operator minus O. And, this is exactly what we needed to cancel of the off diagonal part of the Hamiltonian.

And then, there is a contribution which comes for the electromagnetic field which produces a certain commutator of O with e phi, and then there is a commutator of alpha with itself. And, what it produces is a rather simple term which can be written as O square. So, this is a result of taking the first commutator. And, one can basically now keep on iterating it. Every stage there will be appearing commutators of O with either itself or O with beta, and they can be evaluated. I will only write down the explicit forms.

One more commutator of this whole object with O- the first term will produce exactly this term, s is beta o by 2 m c square combined with O produces this part, and then the second will coming from this particular commutator. I should keep the coefficient of m c square kind of explicit instead of dropping all this relative terms. So, then it is 8 m square c raise to 4. So, O, commutator with O, and this term, here is also c square. And then, the third term also gives m square c fourth O cube.

And, well, one can continue. I do not want to keep any terms which go beyond fourth power of O. So, then there is i cube by 3 factorial, where one can now evaluate again the various parts. O cube divided by 6 m square c raise to fourth, and then 1 by 6 m cube c raise to 6 beta times O fourth. And, the fourth commutator, keeping only the leading term is; I forgot one term here; and this amounts to beta O raise to 4 divided by 24 m cube c raise to 6. So, these are the all the terms which appear in by a commutator of s with H, and this where s is actually the first term of the expansions. So, one can actually put all the coefficients here to make it explicit.

(Refer Slide Time: 39:37)



One can also now work out what are the corresponding terms for the time derivatives. And, that is actually easier. The time derivative is quite straight forward. It is a derivative of this operator. And, the relevant terms are just this. So, now, one just has to add all these things together and see what happens, and the Hamiltonian after this first rotation.

So, these are again has a structure. The diagonal terms are O square by 2 m c square and O fourth divided by 8 m cube c 6. This particular piece is just the relativistic expression for kinetic energy expanded out to desired or, in particular square root of p square plus m square expanded out to second order. So, this is a rest mass, the linear term in p square by 2 m. And then, this is order p raise to 4.

In the other part which is diagonal is just the electromagnetic term, and then there are now new diagonal parts generated by the electromagnetic term. This can be written as various commutators. And then there is a term which is diagonal, but coming from the time derivatives; and then, the O prime which is a. So, all these first things which I have written explicitly are all diagonal, and the ones which are off diagonal are implicit inside this O prime.

The particular structure of the diagonal term is that operator O comes as even powers because it involves this matrix alpha. And, the odd powers of alpha will remain off diagonal, but the even powers of alpha turn diagonal. And, all of them are included inside here. And, O prime will have the remaining part of this whole calculation which is this combination of O, O cube, and O dot.

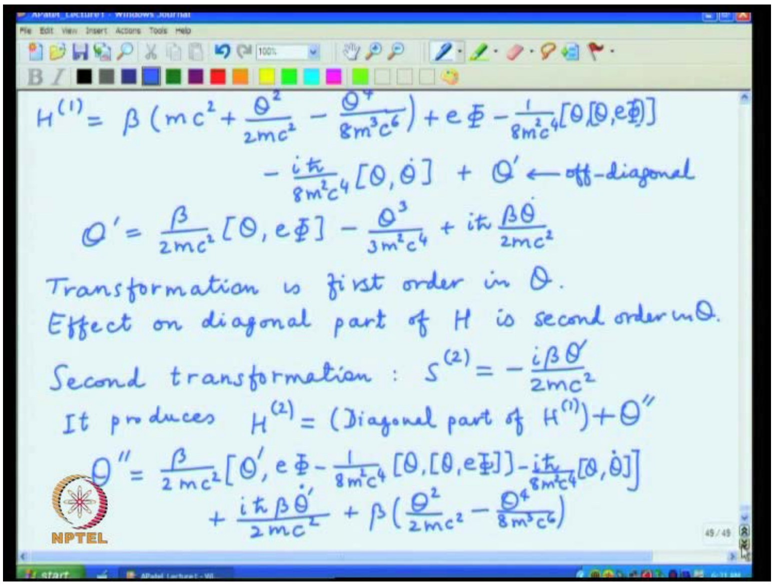
So, the next step now is to rotate away this O prime by a second transformation, and see how much the Hamiltonian changes. As it turns out, this second transformation corresponds to the off diagonal component. But, for it to feed back into the original Hamiltonian you need the second order perturbation theory. The same way that the upper component mixed into lower one, but the lower one had to feed back again into upper component which requires a second order perturbation theory.

So, this will generate a transformation which is first order in O prime, but it will affect the Hamiltonian only as a second order in a prime. So, transformation is first order in O, while effect on diagonal part of H is second order in O. We saw it explicitly in this H 1 construction, and the same pattern we will keep on holding. And, so, we will get some correction to H 1 which will be of order O prime square.

As it turns out, once you evaluate this objects, explicitly the terms which are second order in O prime square are sub leading compared to the terms which are already obtained. And, you do not change the diagonal part of the Hamiltonian much. The wave function does get recorded according to O prime, and that will be done by the same trick for the free particle case where the second transformation is again i beta O prime by 2 m c square.

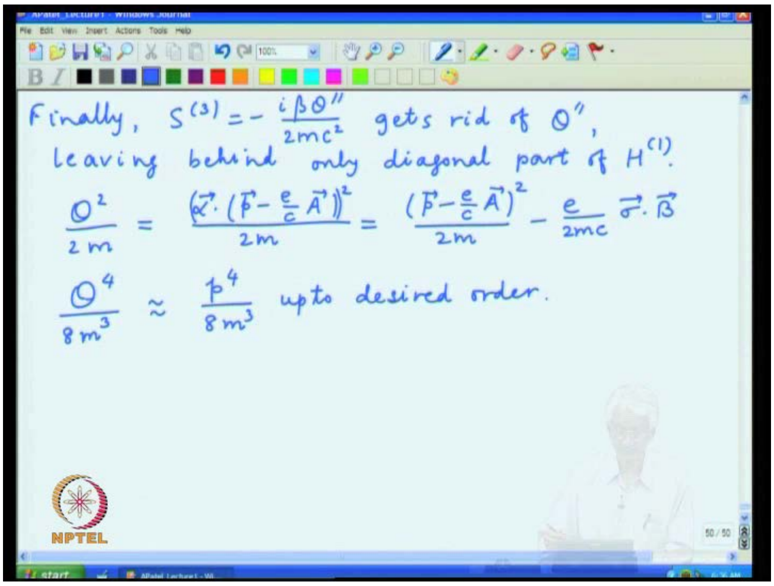
And, what it does is, it produces H 2 which is equal to diagonal part of H 1 written explicitly above, plus a modified of diagonal piece which I will call O double prime. And, that O double prime is now actual operator which has to be worked out by commutator of this object O prime with this big Hamiltonian. It is a matter of some detail, and I would omit it out the complete calculation, and only give the final result.

(Refer Slide Time: 48:24)



And, the expression is; is just the first order commutator of this whole O prime which is written over here with the corresponding terms in the Hamiltonian, plus terms which are coming from various derivatives with respect to time. And, one can keep track of it; of all these explicitly, but it turns out to be sub dominant, and not worth worrying about it. What one can do is, still this is some non 0 term.

(Refer Slide Time: 50:48)



And, in principle, extend this to a next step which is a third transformation which I will just denote as S 3. And, only the diagonal part of H 1 which I explicitly wrote down remains when I say that certain terms are eliminated. They are only eliminated upto the desired order which we are working which is v by c raise to 4; terms which are higher than v by c raise to 4 have been dropped. So, they are effectively there, but not of interest, and we will just treat them as 0 in all these equations which I wrote down.

And, so, now we have a expression, and we need to now evaluate it by constructing the various powers of O which appeared in that expression. And, it is again a straight forward exercise to do. I will write down some of the simple terms. So, the kinetic part is rather straight forward. It is just square of p minus A, and alpha matrix is actually inside the square.

And so, once squares it, alpha square will produce a delta function for the symmetric part, but also a sigma matrix for the anti symmetric part. So, the result for this thing is the usual e square by 2 m, but also the cross term which produces sigma dot B. This is the spin coupling explicitly appearing inside the kinetic energy which you have seen before.

One can now work out the higher order terms as well, and they can be reduced to give a little simpler form. So, in particular, we had this O raise to 4 by 8 m cube. And, this is roughly the same as p by 4 by 8 m cube upto desired order. The electromagnetic potential produces a correction to this, but the effect of A is much smaller than that of p, and so it can be just dropped. So, this is the contribution from the kinetic energy part of the H 1.

And then, there are the 2 new terms explicitly for the interactions which we need to work out. I will do that next time, and then give the interpretation of what those terms mean. They are the ones which are not yet seen in the explicit solution of the non relativistic case. So, they appear automatically in this expansion, and we know their physical interpretation of particular type of correction coming from relativity.