ELEMENTS OF MODERN PHYSICS

Prof. Saurabh Basu
Department of Physics
IIT Guwahati

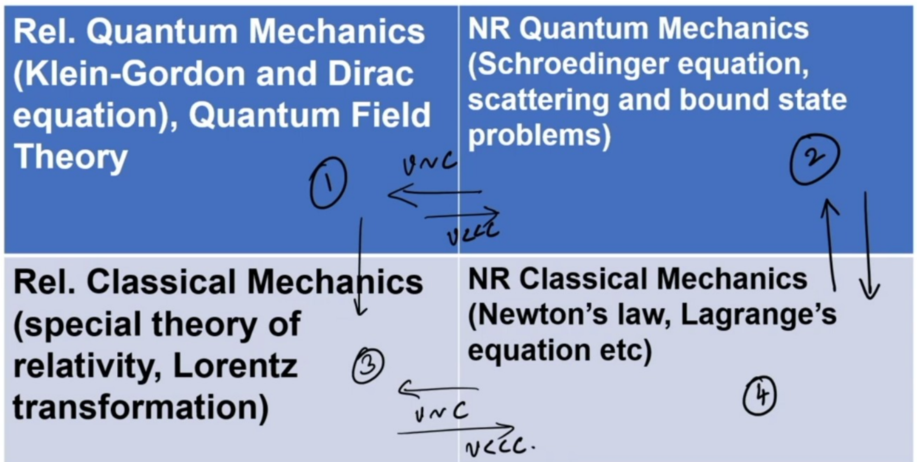
Lec 31: Klein-Gordon and Dirac Equations

Welcome to this last module of the course called Elements of Modern Physics. And in this last module, we will talk about relativistic quantum mechanics. In particular, we will discuss the Klein-Gordon and Dirac equations, though without any interactions such as Coulomb interaction, etc. We will just talk about free particle Klein-Gordon and Dirac equation. Then we will discuss the physics of elementary particles.

So that is the idea. And these later part, which is physics of elementary particles, will also be done in very brief. So this is the structure of mechanics, classical and quantum mechanics, as it has been presented to you so far. So we start with this fourth quadrant. And now if so, there's non relativistic classical mechanics in which we have studied Newton's laws, Lagrange's equation, et cetera, which tells us that under a certain force field, how

the particle evolves or the trajectory of the particle evolves as a function of time. And Lagrange's equation all farther gives that which exactly is the path taken by the particle because the Lagrangian density is extremum along that path only. And then when we went to relativistic classical mechanics, that is we actually came to this 3 in the limit that your V is comparable to C. So if the velocity of the particle is comparable to that of the speed of light, then these non-relativistic formalism, they do not work. And in the relativistic formalism, the main difference that comes in the form of space and time being coupled and we have looked at special theory of relativity and Lorentz transformation and we have talked about various phenomena such as which are paradoxes that arise because of the simultaneity problem.

We have also looked at you know this in this second sort of this quadrant in which we have talked about non relativistic quantum mechanics in fact we have spent the maximum time For this course in this non-relativistic quantum mechanics, where we have studied Schrodinger equation for various kind of potentials, including, you know, a constant potential and then a particle in a box and then harmonic oscillator, hydrogen atom and so on so forth. And this is true or rather, you know, in the limit that you have the sizes of the particle to be small or we are talking about the microscopic world, then these classical, these Newtonian mechanics or the non-relativistic classical mechanics that does not work. And the energy levels in the quantum mechanical systems are quantized in unit of h cross or these all these angular momentum, etc. This carry a scale of h cross and so on h or h cross.



I mean, basically the Planck's constant. Now again one can go from here to here when you consider V to be of the order of C and the relativistic quantum mechanics we have not looked at it and this is precisely what we are going to do here and this would correspond to Klein-Gordon equation, Dirac equation, quantum field theory, etc. We will simply stop at this Klein-Gordon and Dirac equation. So it looks like if you look at the four quadrants, the most complete description is given by these first quadrant, which talks about relativistic quantum mechanics, because this will go to the non relativistic quantum mechanics for V to be much smaller than C. Or it can go to relativistic classical mechanics when h cross can be neglected or, you know, the sizes of the particle or you talk about the macroscopic world and so on.

And similarly, the non-relativistic quantum mechanics can go to non-relativistic classical mechanics when, you know, so this way it is possible when the h cross is low. And this is possible again when v is much, much smaller than c. v is the speed of the particle and c is the speed of light, okay. So, we are really going to talk about these relativistic quantum mechanics, namely the Klein-Gordon and the Dirac equation. So, we start with Klein-Gordon equation.

And in short, we will write KG equation, okay. So, we will write it as a KG equation. So whenever you build up a formalism for studying systems, physical systems, these have to be inconsistent with the special theory of relativity or the Lorentz transformation. And this is precisely the reason that the non-relativistic Schrodinger equation fell short of that requirement because of you have a ih cross del psi lt or d psi dt equal to minus h cross square over 2m d square psi. We are writing just like a free particle wave equation or Schrodinger equation, but you can put a plus a v psi without a problem.

Now, this tells you that this is a second derivative in space. So, this is second derivative in space and this is a single derivative or a first derivative in time. Okay. But if you look at the low range transformation equations, which are like x prime equal to x minus, you know, vt divided by 1 minus that v square by c square and so on so forth, you see the x and t are at the same footing. So, if you write down or you try to apply these low-range transformation, this equation will not satisfy low-range transformation and hence it will not be consistent with the theories of special relativity and they have to be discarded when you are talking about V to be comparable to that of C, okay.

So, this is one of the main drawbacks of Schrodinger equation, which is only valid for V to be much, much smaller than C. In fact, there is no sort of special status of C here, C is just any velocity, but we know that from special theory of relativity that C is the ultimate velocity that a particle can attain. Okay, so this is the reason that one needs to actually go and find out new equations whose solution would give a description of the particle when it is moving in this with this speed which is that comparable to that of light. So, we have two things possible solutions, one is that upgrade the time derivative that is d dt to d2 dt2. that is upgrade the first derivative of time to second derivative of time so that it matches with the second derivative of space and hence will be low range invariant or the second option is that you downgrade the, you know, your d squared to just simply, you know, d or a delta, okay. So, not the Laplacian but the first derivative with respect to space.

This, when you do this, it is called a Klein-Gordon equation, and when you do this, it is called a Dirac equation, and we will see them separately for each of the cases, okay. Now, just a priority. You know, this KG equation or the Klein-Gordon equation has a difficulty in its interpretation because it gives a probability density for particles to be negative. However, if you interpret it properly, then it actually corresponds to probability charge density. which can be negative because of the charge being negative.

So initially, there was a problem with this Klein-Gordon equation, and it's thought to be a perfect description for classical fields such as an electromagnetic field. But then, with a proper interpretation of the charge density, or the probability density, one can actually relate it to the proper wave equation or the appropriate wave equation for integral spin particles. Now, Dirac equation, of course, does not have any of the anomaly and it is actually relevant to spin half particles such as electrons. And in fact, most of the time, we actually talk about electrons.

Klein-Gordon Eqn. (KG equation).

$$i\hbar \frac{d\psi}{dt} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi.$$

first derivative in time. Second derivative in space

$$x' = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}$$

$\boxed{v \sim c.}$

(1) upgrade $\frac{d}{dt} \to \frac{d^2}{dt^2} \to$ KG

(2) downgrade $\nabla^2 \to \nabla \to$ Dirac Equation.

So, the Dirac equation, in that sense, is very important. All right, so we will start with the Klein-Gordon equation, as we have said. So we do not have a dispersion which looks like E by E equal to P square by 2m which is non-relativistic, okay. This dispersion has been well studied; it is a free particle where E is proportional to P squared. And the relativistic dispersion, in fact, is equal to, you know, P square C square plus M square C4.

So, this is the relativistic dispersion. So, you know this is true at speeds which are of the order of speed of light. So if you replace E by I h cross del del t and if you replace P by minus I h cross del then this equation that is this equation the relativistic equation so one becomes minus h cross square del 2 psi del t2 equal to minus h cross square del square c square psi plus m square c4 psi and let us call this as equation 2.

So, this is precisely the Klein-Gordon equation. If you write it a little more neatly, one can write this as a del square minus 1 by c square d2 dt2 and psi, this is equal to m square c4 divided by h cross square psi. And this operator is often called the D'Alembertian operator. It is written with a square sign, but we can simply keep writing the double derivative with respect to space, which is del squared, and the double derivative with respect to time, which is 1 by c squared del 2 del t 2. Now, this is clearly Lorentz invariant because the space and the time derivatives are on the same footing; they are of the same order.

$$E = \frac{p^2}{2m} \quad \text{Non-relativistic.}$$

$$E^2 = p^2c^2 + m^2c^4 \quad \text{Relativistic} \qquad (1)$$

$$E \rightarrow i\hbar\frac{\partial}{\partial t}, \quad \vec{p} = -i\hbar\vec{\nabla}.$$

(1) becomes,

$$-\hbar^2\frac{\partial^2\psi}{\partial t^2} = -\hbar^2\nabla^2c^2\psi + m^2c^4\psi \qquad (2).$$

$$\text{or, } \underbrace{\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)}\psi = \frac{m^2c^4}{\hbar^2}\psi \qquad (3)$$

$$\rightarrow \text{KG equation.}$$

So, this is basically nothing but the KG equation. And so, what is the difficulty with this equation? The difficulty comes when we try to write down the continuity equation, which is nothing but the divergence of J plus del rho del t is equal to 0. To remind you, J is the probability current density and rho is called as the probability density. So, rho in non-relativistic physics is simply equal to mod psi square, where psi is the wave function.

So if in order to, you know, find this, let us call this as equation 4. Let me see whether all equations are numbered. Yes, they are numbered and so on. So to find rho and j, what we do is that we take this equation, equation number 3. and multiply this by psi star and then take a conjugate of this equation and multiply it by psi and subtract the two.

So, this technique has been applied in the non-relativistic case as well to prove the continuity equation. So, the continuity equation is valid here as well. I will just write it that you multiply psi star to equation 3 minus psi conjugate of star the equation 3 star of that equation or conjugate of that equation and take the you know the difference between the two. So that gives that a psi star del square minus 1 over c squared del 2 del t2. This and psi, this is equal to psi star mc squared.

We can write the psi star later. So we have these mc squared by h cross squared psi star psi. m squared c squared. Okay, so by h cross squared psi star psi and this can be called as equation 5. So this is the first term that this term is here and this term is we can write it as psi.

And del square minus 1 by c square del 2 del t 2 psi star is equal to m square c square h cross square psi star psi. And let us call this equation 6. I hope this is clear that we have taken equation 3 and have multiplied it by psi star. Then we have taken the conjugate of 3 and then multiplied by psi and take a difference between the two. And so if you take that, then you can write down this equation 4 and this equation 4, then we have these definitions of the rho and j as so putting in 3 in 4, putting in 4 or putting means after subtraction.

So let me write that after subtraction. putting in four yields, the continuity equations with, you know, the rho given by i h cross by 2 m c square psi star del psi del t minus psi del psi star del t. So this is can be written as h cross by mc square and imaginary part of psi del psi star t del t. So this that is the imaginary part of that and this is nothing but the rho which is a function of r and t and let us call this as equation 7. and this is the probability density and the probability current density J in a function of R and T can be written as minus ih cross by 2m and Psi star del Psi minus Psi del Psi star. That is the same definition that we have seen for the non-relativistic case and let us call this as equation 8, okay.

$$\vec{\nabla}\cdot\vec{J} + \frac{\partial \rho}{\partial t} = 0. \qquad (4)$$

$$\psi^* \times (3) - \psi \times (3)^*$$

$$\psi^*\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\psi = \frac{m^2c^2}{\hbar^2}\psi^*\psi. \qquad (5)$$

$$\psi\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\psi^* = \frac{m^2c^2}{\hbar^2}\psi^*\psi \qquad (6)$$

After subtraction Putting in (4) yields,

$$\rho = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) = \frac{\hbar}{mc^2}\,\mathrm{Im}\left(\psi\frac{\partial\psi^*}{\partial t}\right)$$

$$= \rho(\vec{r},t) \qquad (7)$$

If you assume a free particle solution, Now, for equation 3, so equation 3, so we can write this down as psi of RT, this is equal to exponential i k dot r minus omega t. Okay, you may write an amplitude, you may not write an amplitude, it really does not matter. So, let us call this as equation 9 and if you substitute 9 into 3, that is the unsolved solution, 9 into 3 and what you get is basically k square minus omega square by c square equal to minus m square c square divided by h cross square. And if you identify, you know, so let us call this as 10.

So, if you identify p is equal to h cross k and e equal to h cross omega, then your e square becomes equal to p square c square plus m square c 4 and which gives two values of e plus and minus. This is equal to plus minus C and we have P square plus, you know, M square C square. So, that is the energy and now there are of course, in addition to the plus energies, there are negative energies that are allowed as well. Okay. So, if you take this psi from 9 and then, you know, plug it into this expression, which is 7.

So, plugging 9 into 7, just to remind you that 7 is nothing but this probability density, which is rho RT. So, if you do that, then what you get is that your rho RT becomes equal to ih cross by 2mc square. I write it again here just for convenience, psi star minus psi star del psi del t minus psi del psi star del t and this with this definition or rather these answers written in 9, it gives you E divided by mc square and a mod psi square. Okay, and where, of course, E is equal to this thing.

$$\vec{J}(\vec{r},t) = -\frac{i\hbar}{2m}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right) \qquad (8).$$

Assume a free particle solution for Eq.(3),

$$\psi(\vec{r},t) = e^{i(\vec{k}\cdot\vec{r}-\omega t)}. \qquad (9).$$

Substitute (9) in (3)

$$k^2 - \frac{\omega^2}{c^2} = -\frac{m^2c^2}{\hbar^2}. \qquad (10)$$

$$p = \hbar k, \quad E = \hbar\omega.$$

$$E^2 = p^2c^2 + m^2c^4 \Rightarrow E_\pm = \pm c\sqrt{p^2 + m^2c^2} \qquad (11)$$

Plugging (9) into (7)

$$\rho(\vec{r},t) = \frac{i\hbar}{2mc^2}\left(\psi^*\frac{\partial\psi}{\partial t} - \psi\frac{\partial\psi^*}{\partial t}\right) = \frac{E^{\pm}}{mc^2}|\psi|^2.$$

Should be interpreted as probability current density which can be both positive and negative.

So, E can take positive and negative values and hence these probability density can be either positive or negative. But negative probability density defies traditional wisdom. It doesn't make sense to say that at a given region of space, the probability of finding the particle is negative. So since it doesn't make sense, it can be interpreted as a probability charge density. And then this plus minus sign associated with E, which is written in equation 11, that would be interpreted as since the charge can be both positive and negative.

So if it is a probability charge density, then of course, there is no problem. In fact, this has been told by a lot of people, including Pauli and Weisskopf, etc., In fact, they said that initially they said that this may not be a quantum mechanical equation, but these equations, the Klein-Gordon equation that we have talked about is a valid description of classical field equations such as electromagnetic fields. So, now going back to this, so it should be interpreted as as probability current density, which can be both positive and negative.

So, this is thing or rather this is the issue with Klein-Gordon equation which can be sorted out. So this rho RT can be you know written as just putting the values of that we have, so this is equation set 12 and this can be written as this rho RT is equal to plus minus C and P square plus M square C square divided by MC square and mod psi square, okay. So, this is equation 13 and these positive and negative values can be interpreted in the way that we have talked about. So, let us go back to equation 12 which you know says that rho RT is equal to E divided by MC square into mod psi square.

So, if you take the non-relativistic limit and the non-relativistic limit means that V is much much smaller than C. So, the only energy scale that remains is E equal to MC square because there is no kinetic energy is only the rest mass energy. So, this is called as the non-relativistic limit. So, this is, if you put that then rho RT becomes equal to, so MC square and MC square will cancel and will give you the non-relativistic form which you have used or which you know to be true. So, that equals mod psi squared. Okay, so that, so basically the Klein-Gordon equation is important.

There is an initial problem with, you know, interpretation of the probability density. But if you do that as the probability charge density, then it can take positive and negative values and it gives you the correct non-relativistic form which is known to be mod psi square, okay. But of course, it applies to integer spin. I mean, it basically applies to integer spin and spin 0, specifically spin 0 particles. Or let me write that does not apply to spin half particle, which is what we want to find.

$$\rho(\vec{r},t) = \pm \frac{c\sqrt{p^2 + m^2c^2}}{mc^2} |\psi|^2 . \qquad (13).$$

$$\rho(\vec{r},t) = \frac{E}{mc^2}|\psi|^2$$

Non-rel limit $(v \ll c) \rightarrow \boxed{E = mc^2}$

$$\rho(\vec{r},t) = |\psi|^2 \rightarrow \text{non-relativistic form.}$$

Does not Apply to $S=0$ particles.

and this is what the Dirac equation is going to give us. So, we will study Dirac equation, okay. So, we write down Dirac equation and what we have said is that in order to you know write down a relativistically invariant wave equation and initially there was no hint that it is for spin half particles but later on it emerged. and not only for spin-off particles, there are positive and negative energies as well. So, this is quite important and the way it started is that Terak linearized the equation, the Schrodinger equation.

In fact, as I said earlier that he downgraded the, you know, the space derivative from a double derivative to a single derivative, okay. So, the linearized Hamiltonian It looks like H is equal to C alpha dot P plus beta MC squared. This is the, you know, the kinetic energy, and this is the rest mass energy. Alpha and beta, these are matrices.

and beta they are matrices. And we do not know the rank of the matrix and neither we know the form of the matrix, but they are matrices for reasons that we are going to tell you right away. But one is that that alpha has three components, alpha x, alpha y and alpha z, sometimes called as alpha 1, alpha 2 and alpha 3 and beta is another and so on. And there are certain properties that are going to emerge of these alphas and betas, which necessitate us to take these forms that they are not scalar, you know, quantities. They are actually matrix quantities.

and if you write this as equation 1 and then you plug it into h psi is equal to ih cross d psi dt. So, it is a linearized equation. So, one obtains e minus c alpha dot p minus beta m c square. This psi will be equal to 0 and we can use e equal to minus ih cross, sorry, e is equal to

ih cross del del t and p is equal to minus ih cross del, then this thing looks like this equation, it is ih cross d dt or del del t, let me write it as del del t. minus i h cross c alpha dot del minus beta m c square psi equal to 0. So, this is called as a Dirac equation. And let us call this as equation 2 and this equation is the most important equation. Now, this equation is low range invariant because you have a single space derivative which is del and you have a first order.

What I mean to say by single is that first order space derivative and first order time derivative as well. So, what we do is that in order to you know proceed we will operate this equation by these operator because this operator acting on Psi gives you zero. So, we do that because you have a quadratic you know E square is equal to P square C square. So, we want to see that how this equation can be related to this relativistic dispersion which is E square equal to P square C square plus M square C 4. So, what we do is that we operate this equation by these whatever is under braced here.

And that gives us E plus C alpha dot P. I'm writing this equation, which is one step above. So there is the P is negative. So there's a plus sign here. So E plus C alpha dot P plus a beta MC squared operating on E minus, you know, this C alpha. So, if you write it in terms of P, then there is a minus sign.

And if you write it in terms of P equal to minus IH cross del, then there is a plus sign. So, we should not make the sign mistake. So, E minus C alpha dot P minus beta MC squared. And then acting on psi will give you 0. And this can be, you know, simplified.

Dirac Equation.

Linearized Hamiltonian,

$$H = c\vec{\alpha}\cdot\vec{p} + \beta mc^2 \qquad (1). \quad \vec{\alpha} = (\alpha_x, \alpha_y, \alpha_z), \ \beta$$

Plug into $H\psi = i\hbar\frac{d\psi}{dt}$.

$$\left(E - c\vec{\alpha}\cdot\vec{p} - \beta mc^2\right)\psi = 0$$

$$E = +i\hbar\frac{d}{dt}, \quad \vec{p} = -i\hbar\vec{\nabla}$$

$$\left(i\hbar\frac{\partial}{\partial t} + i\hbar c\,\vec{\alpha}\cdot\vec{\nabla} - \beta mc^2\right)\psi = 0 \quad (2) \rightarrow \text{Dirac equation.}$$

$$\left(E - c\vec{\alpha}\cdot\vec{p} + \beta mc^2\right)\cdot\left(E - c\vec{\alpha}\cdot\vec{p} - \beta mc^2\right)\psi = 0.$$

It will give you a big expression, but there is nothing much to worry about. So you simply, you know, operate it or rather multiply it. And we have these E square minus C square, then you have alpha X square P X square plus alpha Y square P Y square plus alpha Z square P Z square. So, this is one term, then there are cross terms which come as, you know, commutator or anti-commutator. alpha x plus alpha y plus alpha y alpha x and px py, then we have alpha y alpha z plus alpha z alpha y and py pz.

and then we have these plus alpha z alpha x plus alpha x alpha z and then we have this p x p z p x p z okay. So, this is one term and I mean, these are so this alpha squared term and this all these cross terms, etc. And then there is a minus m square c4 beta square and then minus m c square and then we have a alpha x beta plus beta alpha x. This is just simple multiplication. So you have to do it carefully.

There are three terms and this is a three by three. So there will be nine terms and so on. So, plus alpha y beta plus beta alpha y, this is P y and then there is the last term is alpha z beta and beta alpha z P z and psi, this is equal to 0, okay. Now what you do is that you compare it with E2 equal to P2C2 plus M2C4. You see that there is a P2 term and which means that these are like terms.

So, there is like if I open it will be px2 plus py2 plus pz2 and all that. So, all these cross terms you know these terms will all be 0 because there is no term that connects different components of P in this equation. So, that tells us that alpha and beta so they satisfy these equations which these equations are let us call them as so let us call this equation as 3 and these equations are like alpha x square equal to alpha y square equal to alpha z square equal to beta square is equal to 1.

So, that gives that all these operators or rather these quantities the alpha x, alpha y, alpha z, they square up to unity each one of them individually. And, further these, so let us call this as 4a and then alpha x alpha y plus alpha y alpha x equal to alpha y alpha z plus alpha z alpha y equal to 4a. alpha x alpha z plus alpha z alpha x, this is all equal to 0 because there is no cross term. So, this let us say 4B and not only these alphas they, you know, anti commute with themselves, alphas also anti commute with beta. So, alpha x beta plus beta alpha x that is equal to 0 as well.

Because if you see these things, there is no linear term in P in the dispersion in the E square form. So, they would be 0 as well. So, this is equal to alpha y beta plus beta alpha y equal to alpha z beta plus beta alpha z and all this will be equal to 0. So, that tells that these squares are unity. And all of them anti-commute with each other.

$$E^2 - c^2\Big[(\alpha_x^2 p_x^2 + \alpha_y^2 p_y^2 + \alpha_z^2 p_z^2) + \underbrace{(\alpha_x\alpha_y + \alpha_y\alpha_x)}\, p_x p_y + (\alpha_y\alpha_z + \alpha_z\alpha_y) p_y p_z$$
$$\underbrace{(\alpha_z\alpha_x + \alpha_x\alpha_z)}\, p_x p_z\Big] - m^2c^4\beta^2 - mc^2\Big[\underbrace{(\alpha_x\beta + \beta\alpha_x)}\, p_x + (\alpha_y\beta + \beta\alpha_y) p_y$$
$$+ (\alpha_z\beta + \beta\alpha_z) p_z\Big]\psi = 0. \qquad (3).$$

$$E^2 = p^2c^2 + m^2c^4$$

$\vec{\alpha}$ and β satisfy

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1. \qquad (4a) \rightarrow \text{Squares are unity}$$
$$\alpha_x\alpha_y + \alpha_y\alpha_x = \alpha_y\alpha_z + \alpha_z\alpha_y = \alpha_x\alpha_z + \alpha_z\alpha_x = 0 \qquad (4b)$$
$$\alpha_x\beta + \beta\alpha_x = 0 = \alpha_y\beta + \beta\alpha_y = \alpha_z\beta + \beta\alpha_z \qquad (4c)$$

(4b), (4c): all of them anticommute with each others.

Now that tells us that they cannot be simple numbers because of this anti-competition relations and all of this squaring up to 1, that is only possible for matrices. So, alpha and beta are matrices, but we do not know what is the rank of the matrix and so on. But at least we know that they are not a single component object. They are, you know, they have basically a column vector which has, or rather a matrix, not a column vector, sorry, it is a matrix. Psi is a column vector.

H is Hermitian $\longrightarrow$ $\vec{\alpha}$ & β are Hermitian.

Assume β to be diagonal, α_i off diagonal.

$\beta^2 = 1 \Rightarrow \beta$ has eigenvalues ± 1.

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \left(\begin{array}{c|c} \begin{matrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \end{matrix} & 0 \\ \hline 0 & \begin{matrix} -1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & \cdots \\ \cdots & & & \end{matrix} \end{array} \right) \quad m \times n.$$

And because of the hermeticity of H, H is of course Hermitian. And because of that, alpha and beta are Hermitian as well. So, we got a few very important inputs or rather very useful inputs. One is that while linearizing this Schrodinger equation or making it relativistically invariant, we wrote down this equation which is the Dirac equation. which is H minus or H equal to C alpha dot P plus beta MC square operating it on the wave function and writing down the equation, the corresponding wave equation gives us E minus, you know, C alpha dot P minus beta MC square equal to operating on psi equal to 0.

And in order to get further, you know, insights into the problem, what we did is that we have multiplied these operators. So, from the left and on this Dirac equation, the same operator that appears in the Dirac equation, we have multiplied it there. Once that is done, we get a long equation including this Px square, Py square, Pz square term and then the cross term Px, Py and then the term that are linear in X and there is also one term that is completely independent of X and depends on the rest mass energy. If you compare these with this equation, the equation that is the relativistic equation, then we understand that these certain relations of alpha and beta that emerge, which are written in 4A, 4B and 4C. And because of these relations, we are forced to think them not as single, you know, numbers, but they are matrices.

We are yet to ascertain the rank of the matrices or the shape of these matrices. Now these are if we accept that they are matrices then we can take arbitrarily we can take one to be diagonal and others to be you know off diagonal and let us take beta this purely kind of you know convenience beta to be diagonal and all other alpha i these are off diagonal. Once again, it is not clear at this moment that what is the size of the matrix or what is, you know, the rank of the matrix, so to say. Okay. So, why is this taken?

Because one is diagonal and the others are off diagonal and that is the only way that they can commute or rather they can anti-commute because if more than one is diagonal, then they would not, you know, they would not all amongst themselves, they would not anti-commute. So, this is an essential requirement. Now, if you ask me that why beta has been taken as diagonal and the other three are off diagonal, you can, you know, there is no embargo on that. You can take one of the four to be diagonal and the other three to be off diagonal. But this is the standard procedure that is, you know, kind of followed.

Okay, so the squares of the matrix that become unity, so beta square equal to 1 and beta is diagonal, that tells us that beta has eigenvalues which are equal to plus and minus 1. Okay. So, if these are the eigenvalues, then beta and beta being diagonal, so beta will have, you know, in its diagonal elements, we will have this. And just to make things a little general, we can write this not as 1, but like this, which would stand for, you know, a sort of big matrix, which is like this. And it is like a 1, 0, 0, and so on, and then 0, 1, 0, so on, and then 0, 0, 1, so on, and then so on here.

And similarly, all these are 0. And then we have these as, you know, so basically, we can choose this as not as 1, but as minus 1. And then we have a minus 1, 0, 0. and so on. It is still diagonal, of course, but with 1, 1, it becomes an identity matrix which will commute with everybody.

So, that is not a possibility. I wrongly wrote it as, you know, plus 1, but it has to be minus 1. So, this is 0, minus 1, 0 and so on so forth, okay. So, that is the size of the matrix, but we really do not know or the shape of the matrix rather. And we have said that alpha and beta are

you know they are Hermitian matrices. Now, if beta has a structure like this and each of them, you know, it is some kind of an m cross n matrix. So, each of these is like a m cross n matrix, okay. So, each of these diagonal blocks is like m cross n matrix. So, this is like an m cross n matrix and this is like an m cross n matrix as well and we are yet to ascertain that what are the values of m and n.

Now, very important input comes when we look at the anticommutation relation between the alpha x beta plus or any of these not only alpha x but alpha y but just we are looking at this alpha x. So, beta alpha x this is equal to 0. So, if you take the jk-th element then what happens is that the alpha x jk and these beta j plus beta k, this is equal to 0. Now, this beta j, beta k are the diagonal elements and because it has eigenvalues plus 1 and minus 1, so beta j can be equal to beta k can be equal to say either plus 1 or minus 1. In both the cases, your alpha x is equal to 0.

if beta j is not equal to beta k which means one of them is plus 1 and the other is minus 1 in that case alpha X j k may not be equal to 0 I am not saying not equal to 0 so may not be equal to 0. That is because of this in order to satisfy this equation. So, once again, I will just go over it. If beta j and beta k both are same, then they would add up either, you know, both plus 1 or both minus 1. In that case, necessarily alpha x j k will have to be equal to 0.

And if it is not equal to they are not equal that is 1 is plus 1 and 1 is minus 1 then that could be 0 then alpha x may not be equal to 0. And in fact it is clear that it is coming from that structure that these alpha X and similarly alpha Y because alpha Y obeys the similar relation as we have written here this equation. So, and alpha z as well, that satisfies the same relation as well. So, all of alpha x, alpha y and alpha z are off diagonal. And this is what we have, you know, kind of started off with.

And when we mean off-diagonal, which means that, you know, it has a structure where these are zeros and there will be some non-zero elements, you know, here. These will be some non-zero elements, okay. And, in fact, this alpha matrix can be written as, we will just show that as a 0 and sigma and sigma and 0, where sigma are the Pauli matrices. So, sigma x, sigma y, and sigma z are And then automatically it tells you that it becomes equal to, you know, a 4 by 4 matrix.

So, beta is then equal to, it is like 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, minus 1, 0, 0, 0, 0, minus 1. And these alphas will be formed by this matrix. You know, these 2 by 2 matrices, which are sigma x, sigma y, and sigma z, just to remind you that sigma x is equal to 0, 1, 1, 0, sigma y is equal to 0, minus 1, 0, sorry, minus i, i, 0, and sigma z is the diagonal one, which has 1, 0, 0, minus 1. Okay, so now why is it 4 by 4 and not larger than that? There are two, you know, sorts of rationales or two reasons behind that.

We can say that, you know, any matrix, suppose first we say why it is not a 2 by 2 matrix and it is a 4 by 4 matrix. So, if you for any matrix, any 2 by 2 matrix, which can be written as linear combinations of the Pauli matrices and the identity matrix. And if it can be written as in terms of the Pauli matrices, then they will commute. And clearly these matrices anti-commute with each other. So they can't be two by two.

And so basically also it's not difficult to show that if they are two by two matrices, then a fourth matrix really, you can't find a fourth matrix that anti-commutes with all three. And this again reiterating the solution or rather the statement that for any 2 by 2 matrix like a matrix which is A, B, C, D that can be written in terms of something sigma X plus something sigma Y plus something sigma Z that is some coefficient sigma Z which are formed of A, B, C, D and so on and sigma Z and plus something which is sigma 0. All right. So if that is true, then, of course, these matrices won't anti-commute with each other because each one is like a sigma X and or sigma Y or sigma Z. Okay, so let me give you another proof, which is probably even, you know, more mathematically sound.

$$\Rightarrow \alpha_x\beta + \beta\alpha_x = 0.$$

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

$$(\alpha_x)_{jk}\,(\beta_j + \beta_k) = 0.$$

$$\beta_j = \beta_k = \pm 1 \Rightarrow (\alpha_x)_{jk} = 0.$$

$$\beta_j \neq \beta_k. \Rightarrow (\alpha_x)_{jk} \text{ may not be zero.}$$

$$\left[\begin{array}{c|c} 0 & \text{NZ.} \\ \hline \text{Z.} & 0 \end{array}\right]$$

α_x, α_y & α_z are off-diagonal.

$$\vec{\alpha} = \begin{bmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{bmatrix}$$

$$\vec{\sigma} : (\sigma_x, \sigma_y, \sigma_z)$$

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

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (\ )\sigma_x + (\ )\sigma_y + (\ )\sigma_z + (\ )\sigma_0$$

And so we start with this anticommutation relation, which is alpha i beta equal to minus beta alpha i, okay, where i is x, y and z. So, if I take a determinant of both sides alpha i beta which is equal to determinant of minus 1 and determinant of beta alpha i. So, this becomes equal to determinant of alpha i into determinant of beta and equal to determinant of minus 1 and determinant of beta and determinant of alpha i. Now, you see that determinant becomes, you know, kind of a number. So, determinant of alpha i, determinant of beta will cancel from both sides.

And that gives us the determinant of minus 1 is equal to 1. Now, for general n cross n matrices. this determinant of minus 1, you know, it is equal to minus 1 whole to the power n and a minus 1 whole to the power n can only occur which becomes equal to 1, can only occur for n equal to 2, n equal to 4, n equal to 6 and so on. So, the simplest one or the smallest one is n equal to 2 that tells us that these alpha matrices are really like these matrices that we have written earlier. So, these are 2 by 2 matrices, these are 2 by 2 matrices and 2 by 2 and similarly for beta as well they should definitely have the same rank in order to you know follow all these anti-commutation relations.

$$\alpha_i\beta = -\beta\alpha_i \qquad i \in x, y, z$$

$$\det(\alpha_i\beta) = \det(-1)\det(\beta\alpha_i).$$

$$\Rightarrow \det(\alpha_i)\det(\beta) = \det(-1)\det(\beta)\det(\alpha_i)$$

$$\Rightarrow \det(-1) = 1.$$

for $N \times N$ matrices $\det(-1) = (-1)^N$.

$$(-1)^N = 1$$

$$N = \underset{\substack{\uparrow \\ \text{Smallest.}}}{2}, 4, 6 \ldots\ldots$$

$$\vec{\alpha} = \begin{pmatrix} 0 & \overset{2\times 2}{\vec{\sigma}} \\ \underset{2\times 2}{\vec{\sigma}} & 0 \end{pmatrix}$$

So, now it is clear that all four of them are really, you know, four-by-four matrices, and it is important to note that, you know, these two of them actually correspond to positive and negative energy solutions, and the other is basically because of the spin, which is like the Hilbert space is 2s plus 1. It's also important to note that we did not have an analog of spin in non-relativistic quantum mechanics. It only came naturally, you know, with an effort to linearize the Schrödinger equation and make it Lorentz-invariant; it came as a dividend. So, let me finally write these equations, which are: 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, minus 1, 0, 0, 0, 0, minus 1. That's beta.

Alpha x is that sigma x there. So, it's 0, sigma x, sigma x, 0. So, that's 0, 0, 0, 1. 0, 0, 1, 0, 0, 1, 0, 0, and 1, 0, 0, 0. You can check all of them to be, you know, Hermitian.

So, even though alpha y will have an i there, it is still Hermitian. So, 0, 0, 0, minus i, so now it has a structure which is like a 0, sigma y, sigma y, 0, okay, so this structure is reflected here, and this structure is 0, sigma x, sigma x, 0, that is reflected here. So, 0, 0, i, 0, then it is 0, minus i, 0, 0, then i, 0, 0, 0, okay. And finally, alpha z, which is equal to this sigma z. So, sigma 0, sigma z, sigma z, 0. So, that has a distinct form from beta, although it looks like.

So it's a 0, 0, 1, 0 and 0, 0, 0, minus 1. Then it's a 1, 0, 0, 0. Then it's a 0, minus 1, 0, 0. And so these are the four matrices. Now, assume a solution, free particle solution.

$$\beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \alpha_x = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \leftarrow \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}$$

$$\alpha_y = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}, \quad \alpha_z = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}$$

psi i RT, so this is equal to ui exponential i k dot r minus omega t. So if you substitute this in the Dirac equation, we have kept, of course, amplitude, and you get a 4 by 4 equation. So now the Dirac equation, unlike Schrodinger equation, which was just a single equation, now it's a matrix equation. So there'll be four coupled equations. So, it is E minus MC square u1 and minus C pz U3 minus C px minus i py u4 equal to 0, there is no u2 here.

The second one has E mc2. This is obtained by putting all this alpha beta which are given there and then putting this unsolved solution. u2, this is C px plus I py sigma U3 and plus C pz u4, that is equal to 0. Again, there is no U1 here. Then we have a E plus beta mc, sorry mc square, beta has been put, mc square u3 minus cpz u1 minus c px minus ipy u2, this is equal to 0.

and e plus mc squared u4 minus c px plus ipy u1 and plus cpz u2, that is equal to 0. So, these are 4 by 4 equations. Now, they have a unique solution only when the determinant of the coefficient matrix is 0. And this gives us these unique solutions and these solutions for, so there are these two energies E plus and E minus. So, corresponding to E plus, we have E equal to E plus, that is the plus sign in front of this.

Assume a free particle solution:

$$\psi_i(\vec{r},t) = u_i \exp\left[i(\vec{k}\cdot\vec{r} - \omega t)\right].$$

$$(E - mc^2)u_1 - c p_z u_3 - c(p_x - i p_y) u_4 = 0.$$
$$(E - mc^2)u_2 - c(p_x + i p_y)u_3 + c p_z u_4 = 0.$$
$$(E + mc^2)u_3 - c p_z u_1 - c(p_x - i p_y)u_2 = 0.$$
$$(E + mc^2)u_4 - c(p_x + i p_y)u_1 + c p_z u_2 = 0.$$

$\boxed{\text{Det of Coefficient Matrix is zero.}}$

So, if you write this E to be equal to, you know, plus minus C, p squared, c squared plus m, sorry, not c squared, c already has been taken out. So it's p squared plus m squared, c squared. So this plus sign, e equal to e plus, so you have these u1, u2, u3, u4. And these are, you know, so there are these two linearly independent solutions.

And these linearly independent solutions are C pz divided by E plus plus MC squared and C px plus i py divided by E plus plus MC square. Let me write this neatly. So, this is one and the other is the two linearly independent solutions are, it is a 4 by 4, the two eigenvalues are E plus and the two eigenvalues are E minus. So, it is a C px minus i py.

divided by E plus plus MC squared and minus C pz divided by E plus plus MC squared. Yeah. So these are the two linearly independent solutions for E plus. And for E minus, we have these u1, u2, u3, and u4 as these C pz divided by E minus, minus MC squared, C px minus i py divided by E minus MC squared. and 1 and 0 and the other solution is C px minus ipy divided by e minus minus mc square and Cpz divided by e minus minus mc square.

$$E = \pm c\sqrt{p^2 + m^2c^2}$$

$$E = E_+$$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \dfrac{cp_z}{E_+ + mc^2} \\ \dfrac{c(p_x + ip_y)}{E_+ + mc^2} \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \dfrac{c(p_x - ip_y)}{E_+ + mc^2} \\ -\dfrac{cp_z}{E_+ + mc^2} \end{pmatrix}$$

$$E = E_-$$

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} \dfrac{cp_z}{E_- - mc^2} \\ \dfrac{c(p_x - ip_y)}{E_- - mc^2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \dfrac{c(p_x - ip_y)}{E_- - mc^2} \\ \dfrac{cp_z}{E_- - mc^2} \\ 0 \\ 1 \end{pmatrix}.$$

and 0 and 1. So this is corresponding to E minus and that corresponds to E plus. So these are the two linearly independent solutions. So we got positive and negative energy solutions and then we have as an important dividend we got a sort of a matrix structure of this Hamiltonian which corresponds to you know up and down spins which are relevant for the spin half particles like electrons. So electron

So, it is basically the spin has this two-fold structure as well as the energy, the two energies plus and minus that makes it four. So, it is basically a four by four equation or rather the four by four Hamiltonian. whose solution gives us all these things for the free particle. Of course, we stop at free particle, but the problem can also be solved. Suppose you want to solve hydrogen atom, so that solution would have to be done in presence of a, you know, the 1 over r kind of a Coulomb potential.

We will not do that. From here on, we will discuss very briefly the physics of elementary particles. So, we will stop here. Amen.