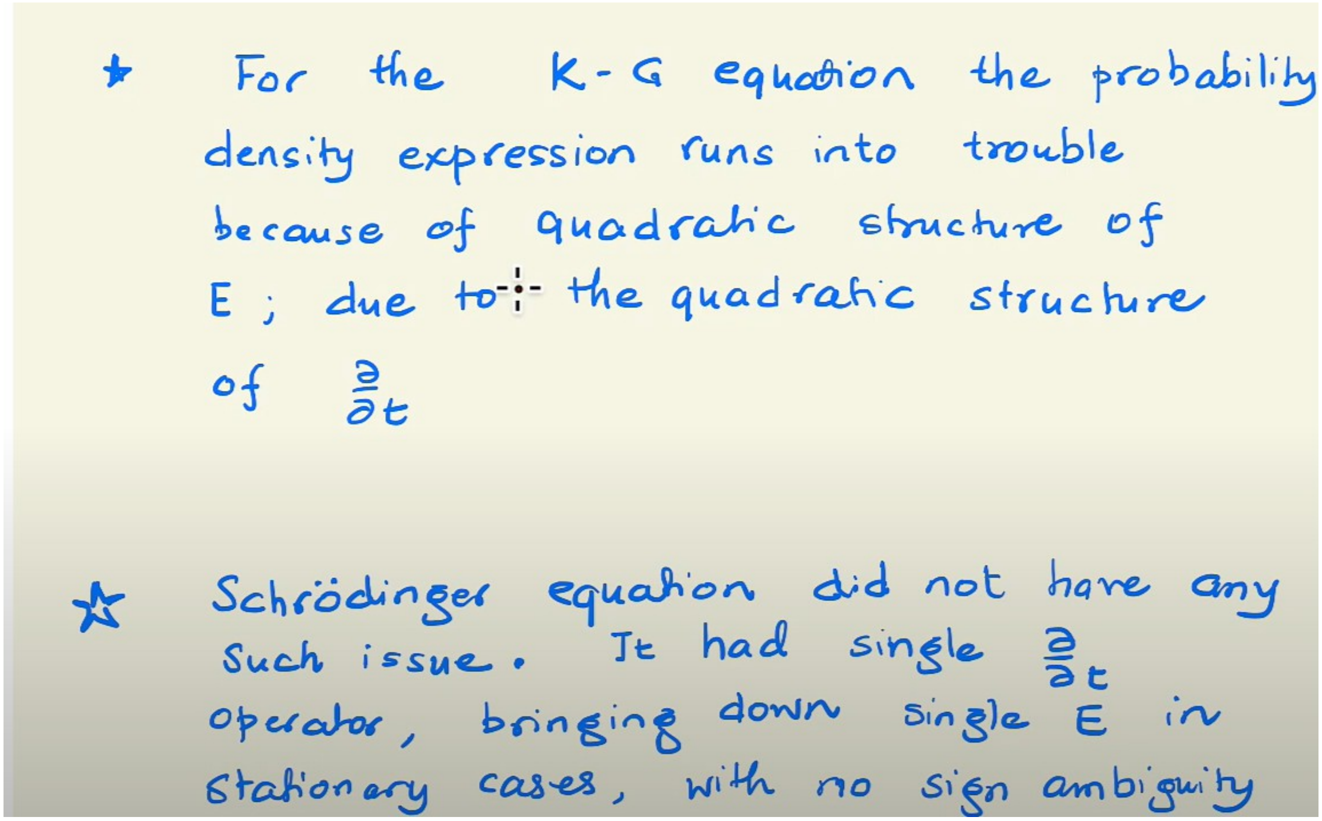
Foundation of Quantum Theory: Relativistic Approach
Relativistic Quantum Mechanics 1.3

Prof . Kinjalk Lochan
Department of Physical Sciences
IISER Mohali

Dirac Equation I
Lecture- 13

So, in this discussion session we will move beyond the Klein–Gordon discussion which we have till now and then we will try to see the onset of discussions on Dirac equation. So, remember when we were discussing the Klein–Gordon equation we ran into problems of consistency or conceptual clarity where the Klein–Gordon equation it turned out that the probability density expression becomes difficult to handle because of its quadratic structure, it admits a negative solutions for energy as well. And therefore, the probability density, the quantity which we were tempted to call probability density runs the risk of becoming a negative and therefore, an inconsistent way of describing the probability physics which quantum theory is typically built/. So, therefore, it was not a good idea that energy should be allowed a negative solution, otherwise the probability density is in jeopardy. And this happened because of its quadratic structure in double derivative with respect to time. that means the Klein–Gordon equation had a structure like $\partial t\ 1/c^2$ This gave rise to a quantity like E^2/c^2, if you remember in the stationary case and therefore, the solution of this equation came up with both + or –sign, where the –sign created problem for us. On the other hand, when we were doing non relativistic quantum mechanics, the Schrodinger equation that does not have any quadratic structure. It has a single $\partial/\partial t$ operator which controls the dynamics. So, $i\hbar\partial\psi/\partial t$ and in the stationary case, it gives just the one energy E which is positive semi–definite. So, therefore, there is no sign ambiguity in the non–relativistic regime because of single derivative structure while in the Klein–Gordon scheme of things due to its double derivative structure a quadratic form of the energy came about coming up with subsequently a negative sign as well. So, the question which was naturally asked by Dirac first that is there a special relativity compatible first order equation. So, we had to give up Schrodinger equation because it was not consistent with Schrodinger with special relativity because of its structure that it is first derivative in time, but double derivatives in space. So, it was not compatible with Lorentz transformation. However, making that also a double derivative, the temporal derivative also has a double derivative quantity, then negative energy solutions creeped up. Now, the question is instead of making the temporal derivative of second order quantity. Can we make first order Hamiltonian where even the spatial derivative is made up from the first derivative and that would be a first derivative spatial relativity compatible equation and that is what the search of Dirac which ultimately led him to the Dirac equation. And then we will see whether doing this kind of business do we really elevate the problem which was faced by the Klein–Gordon equation namely the negative energy solutions and the negative probability density or not. So, that is the discussion which we will have in this session today and it would be the Dirac equation with which we will understand things.



★For the K-G equation the probability density expression runs into trouble because of quadratic structure of E_{jf} due to quadratic structure of $\partial/\partial t$.

★Schrodinger rquation did not have any such issue. It has single $\partial/\partial t$ operator, bringing down single E in stationary cases, with no single ambiguity.

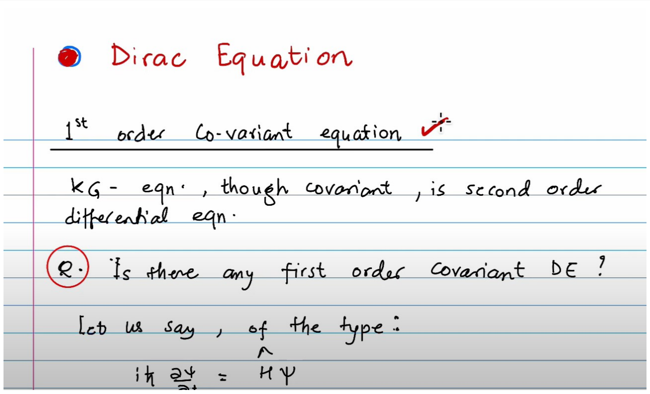
So, the idea is to obtain a first order covariant equation just like we had a invariant equation or a low range compatible equation at the second order, we are trying to find a first order equation. So, we are looking for a question, is there any Hamiltonian and of any kind made up of first derivatives only, because we do not want inconsistency between second derivative on the Hamiltonian side and first derivative on the temporal derivative side. The question is, is there a first order covariant Hamiltonian such that it is made up entirely of the first derivative and left hand side is a one single derivative of time. So, let us propose that it is made up of quantities like $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, which here right now I am writing as $\partial/\partial x^1$, $\partial/\partial x^2$, $\partial/\partial x^3$. And better they should come up with some coefficients α_1, α_2, α_3. And maybe some special relativity invariant quantity m0 c^2 also we can throw in. And that will also come up with some coefficient. So, the idea is we want the whole equation to be either invariant or covariant. So, you can readily see that this cannot have an invariant structure if α_1, α_2 or α_3 are supposed to be constant numbers. If they are just numbers, this thing will not work out because left hand side of the equation is the temporal derivative. Right hand side we have something, a partial derivative, partial derivatives coming with some coefficient. Let us think of a case where I am doing spatial rotations. If I do spatial rotations, I know that the coordinate differentials *dx, dy, dz or* dx^1, dx^2, dx^3, these quantities transform, they do not remain the same under rotation. They transform with a rotation matrix given by this. If these transforms like this rotation matrix, we can figure out how the derivatives will transform, what will happen to $\partial/\partial x^1$, what will happen to $\partial/\partial x^2$ and what will happen to $\partial/\partial x^3$. This is an easy exercise which you should have done in some classical mechanics transformations as well as ordinary quantum mechanics, where you will realize that these quantities transform as with inverse of this matrix *R*, *R* inverse. But since this matrix *R* is an orthogonal, so this matrix R is orthogonal matrix. So, therefore, you would have the inverse of this as just a transpose of that matrix. So, therefore, the transpose of R transforms these quantities, partial derivative with respect to *x, y, z or* x^1, x^2, x^3. If α_1, α_2, α_3 are constants, then it so happens that under rotations, the constants α_1, α_2, α_3 would not change while the derivatives with respect to x^1, x^2, x^3 would change with *R* transpose. that means the Hamiltonian would change. It will not remain the same Hamiltonian under rotation. Whereas on the left hand side, I have a temporal derivative. under spatial rotations temporal derivative are not supposed to change. So, therefore left hand side would not change at all. So, we have an inconsistency at the hand that under rotation a right hand side has changed left hand side has not changed. So, this equation cannot be true and since rotations also are a part of general special relativistic transformation or Lorentz transformation. Recall, low range transformations are made up of rotations and boosts. So, if under rotations things have changed, that means under low range transformation things change. So, this is not an invariant equation. So, therefore, if I try to write down a first order equation, the clear realization we should have is that such a first order equation cannot be obtained from constants α_1, α_2, α_3. And better there should be something which also transformed in a particular way that exactly cancels the changes brought about by $\partial/\partial x^1$, $\partial/\partial x^2$, $\partial/\partial x^3$ such that overall this quantity remains unchanged. So, first realization is this, this α_1, α_2, α_3 *s*hould be.

If we have single derivative equation probably then we will be safe and we will be getting energy, energy density and other things that is the domain of Dirac equation which we will do in the next class.

So, remember when we were discussing the Klein–Gordon equation we ran into problems of consistency or conceptual clarity where the Klein–Gordon equation it turned out that the probability density expression becomes difficult to handle because of its quadratic structure, it admits a negative solutions for energy as well. And therefore, the probability density, the quantity which we were tempted to call probability density runs the risk of becoming a negative and therefore, an inconsistent way of describing the probability physics which quantum theory is typically built/. So, therefore, it was not a good idea that energy should be allowed a negative solution, otherwise the probability density is in jeopardy. And this happened because of its quadratic structure in double derivative with respect to time. that means the Klein–Gordon equation had a structure like $\partial t^2 1/c^2$ This gave rise to a quantity like

E^2/c^2, if you remember in the stationary case and therefore, the solution of this equation came up with both + or − sign, where the − sign created problem for us. On the other hand, when we were doing non relativistic quantum mechanics, the Schrodinger equation that does not have any quadratic structure.

From segment 51
So, first realization is this, this $\alpha_1, \alpha_2, \alpha_3$ should be probably components of some vectors. Then this is a vector, this is a covector. This is a component of a vector, this is a component of a covector. And it has a structure like a vector undergoing the gradient dot product with the gradient. This quantity presumably will not change under a spatial rotation. So, now we would have a structure that the coefficients which are coming in $\alpha_1, \alpha_2, \alpha_3$ are components of some vector $\dot{\boldsymbol{r}}$ gradient operator presumably should not change under rotation. Then it looks like that we have a case for a first order differential equation. Left hand side also does not change under rotation, right hand side also does not change under rotation. However, we will quickly see that this cannot be an ordinary Position vector it has to have more feature than that Okay, so let us go and see what kind of other features we can expect this to have Okay, so now that we have realized the coefficients α and the β which were appearing over here and Here they are not supposed to be ordinary c numbers. They have better be components of some vector and We will soon see that this cannot be ordinary vectors components, but they have to have certain more structure than ordinary vectors. More precisely we will see that α's and β's should be matrix based vectors. that means they should be α should be α should be the component of certain matrices R. which is $\alpha_{1i} + \alpha_{2j} + \alpha_{3k}$ such that these are just not numbers apart from being components of the vectors, they should be matrices.



★Dirac Equation

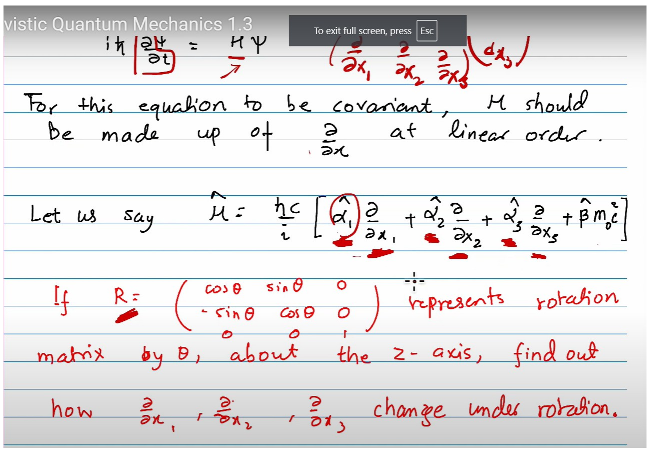
1st order Co-variant equation

KG eqn', though covariant is a second order differential equation.

Q. Is there any first order covariant De?

Let us say, of the type:

$$i\hbar\frac{\partial\psi}{\partial t}=\hat{H}\,\psi$$



For this equation to be covariant, $\mathscr{H}$ should be made up of ∂/∂x at linear order.

Let us say

$$\hat{H}=\frac{hc}{i}[\hat{\alpha}_1\frac{\partial}{\partial x_1}+\hat{\alpha}_2\frac{\partial}{\partial x_2}+\hat{\alpha}_3\frac{\partial}{\partial x_3}+\hat{\beta}\,m_0c^2]$$

If R = $\begin{pmatrix} cos\theta & sin\theta & 0 \\ -sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ represents rotation matrix by θ_1 about the z aixis, find out *how* $\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2}$ *and* $\frac{\partial}{\partial x_3}$ will change under rotation.

So, r is not a vector, it is a matrix valued vector that we will soon realize as we will see in upcoming slides. So, ultimately these α's and β's which are appearing on the Hamiltonian side, turn out to be matrices. So, they are not ordinary numbers they are matrices apart from being components of the vector. So, overall on the right hand side this becomes a matrix hitting/the wave function ψ. So, therefore if the Hamiltonian becomes a matrix Then the wave function should also be a column vector. A matrix cannot act on a function. A matrix has to act on a column vector. What dimensional column vector? It should have the same dimension as that of the matrix. Matrix has a *n* cross *n* dimension, then the column vector will be n cross 1 dimension. What decides the *n*? What should be the size of the matrix? All these things will become clear in the discussions coming below. So, ultimately this is an upshot which I am giving. We have not yet seen how α and β's are going to be matrices, but this is an upshot of upcoming discussion. However, we have seen that they should be component of some vectors. Now, I will motivate why they cannot be ordinary vectors, but matrix valued vectors. Okay the picture that will emerge as again to re–emphasize it would be left hand side there will since it has right hand side Hamiltonian has become a matrix left hand side also better should be a matrix so what better matrix than the identity to put on the left hand side therefore. The consistency structure will demand that I have a i$\hbar$ times some identity n cross n dimensional matrix is equal to $\hbar$c/I times some matrix component α, α_1, α_2, α_3 coming with the partial derivative ∂x_1, ∂x_2 , ∂x_3 and the β which are also n cross n dimensional matrix. And both sides are acting on wave functions. They are written as a n dimensional column vectors. So, the wave function itself does not remain just one function, but it is a collection of n functions. So, it is n–tablet equation. So, demand of first derivative covariant equation, low edge compatible equation would ultimately lead us to a collection of many functions rather than a singlet function. And that is why a more features will be recovered in the Dirac equations structure. And that is why more features will be recovered in the Dirac equations structure. So, do $\dot{\psi}_1$, $\dot{\psi}_2$, $\dot{\psi}_m$, $\dot{\psi}_n$. And ultimately one particular ψ m dot would be equal to on the right hand side the Hamiltonian acting/the wave function vector and ultimately mth component of that. So, that in the matrix notation will be written as this box equation where $i\hbar\partial\psi_m/\partial t$ is equal to mth row and n^{th} column element of that nth row the element of that row are summing over the column elements of the vector ψ_n that means this row goes with the column vector and just sums it over over all the elements of the column vector this element this element this element so this is $\dot{\psi}_m$ and this was the hit has many row,s, I will select the mth row because I am looking at mth component in the wave function and that will just undergo multiplication with the column vector that is the usual matrix multiplication structure. Now, we are left out to see why I am claiming that α and β's are supposed to be matrices and if they are matrices what are their dimensions. So, that discussion we will take and see where do we get the matrix structure from. So, the question is how to fix down what should be α's and β's which we have put/. So, let us see we know that the structure which we are after is the first derivative structure which we have repeatedly written. It is something like that. I have written as a proposed equation is equation number 1.

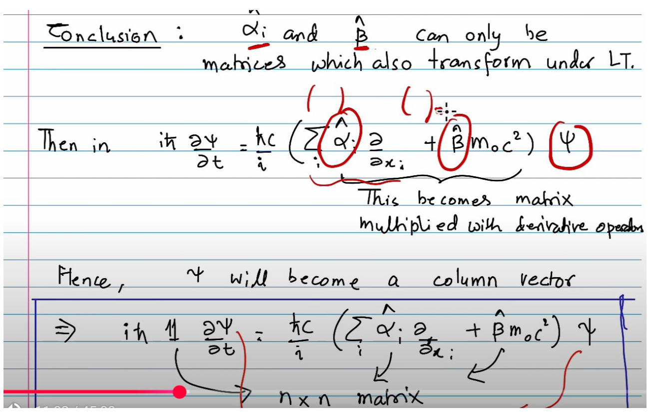
Now, what I do, if I take one more derivative of this equation from the left hand side, so I will have a double derivative with respect to time, $\partial^2\psi/\partial t^2$. And then this $\partial/\partial t$ will go across. And hit the ψ because α's or $\partial/\partial x$, they are all time independent. So, if I introduce one more $\partial/\partial t$ from the left hand side, this will be hitting the right hand side term as well and in right hand side only the wave function is supposed to be time dependent. All other quantities are time independent. So, the derivative operator which we have introduced will go and hit this ψ and I will have a $\partial_t\psi$ term.

Let me multiply additional i$\hbar$ on both sides so I will have i$\hbar$ multiplied both sides now you can see what emerges out in the left hand side I will have a $-\hbar^2\partial^2\psi/\partial t^2$ just from this on the right hand side I can combine this i$\hbar$ to this $\partial\psi/\partial t$ and equation use the first equation a back that means I already know what is $i\hbar\partial\psi/\partial t$ and that information is fed into equation number B. $i\hbar\partial\psi/\partial t$ is equal to right hand side of equation number a. So, that means $-\hbar^2\partial^2\psi/\partial t^2$ which has appeared on the left hand side would be equal to $-\hbar^2c^2$. This will come from Since the ih $\partial\psi/\partial t$ coming here would be replaced by another factor like

this. So, this whole thing will go and replace this and this $i\hbar$. So, I will have a already there was a $\hbar$ c/I present. Another $\hbar c/i$ is coming from the second replacement. So, I will have a $-\hbar^2c^2$ and then the two times of the same bracket this and this. Remember since these are series I will use different indices I here and j here. j runs from 0, 1, 2, 3 and I also runs from 0, 1, 2, 3. So, this would be the structure which I will get. So, I can open it up. So, this whole series here is $\partial/\partial x$, $\partial/\partial x^2$, $\partial/\partial x^3$ and whole series here also is $\partial/\partial x^1$, $\partial/\partial x^2$, $\partial/\partial x^3$. So, when I open the bracket, sometimes I will get two derivatives with $\partial/\partial x$ 1, sometimes two derivatives with $\partial/\partial x^2$, sometimes two derivatives with $\partial/\partial x^3$ and there will be cross terms as well. So, what we can do effectively is just open up the bracket and find out what kind of derivative terms on the ψ which is here do we end up earning. So, $-\hbar^2c^2$ is a common to all it will remain there the two sums I and j they can directly. So, this term this term combined will give rise to α_i α_j and the double derivative hitting the ψ, $\partial^2\psi/\partial x_i\partial x_j$. that $\alpha_i\alpha_j$, if I have, I am just writing for the consistency of this, I will get a $\alpha_i\partial/\partial x$ I and then j $\alpha_j\partial/\partial x$ j hitting the ψ. that term I am writing as α_{ij} and $\alpha_i\alpha_j + \alpha_j\alpha_i$ by 2. This is just a symmetrization I have written and ∂^2x ψ ∂^2x^i and ∂x^j. So, this is the first term I have written. This is just the symmetrization of the quantities which, sorry here, this is the symmetrization of the quantities, the double derivative which has appeared. Then there will be a next order term, which will be this term combining with this βm_0c^2. And similarly, this term combining with this derivative operator. As a consequence of that, you will get a $\hbar^2c^2$ and m_0c^2 further. Okay, so together they will generate for you $\hbar m_0c$ cube.
One $\hbar$c is coming from the first bracket and the second one is coming from this m_0 c^2. So you will have this structure where one α and one β combine. So you will have a $\alpha_i\beta$ from here and here and then a β times α which is here and here. And lastly, those two β's combine and you will have β^2 term with $m_0^2c^4$ and a ψ. So, ultimately the right hand side if I look at opening up the bracket, I will get a double derivative terms, I will get a single derivative terms and I will get no derivative terms. All these terms are double derivative, all the terms are single derivative and this is no derivative. So, we have got a double derivative structure. Now we know that since we started from a special relativity compatible theory, its double derivative structure should also be special relativity compatible. And we know at the level of double derivative there is one operator which is special relativity respecting covariant that was the D'Alembertian, the Klein–Gordon operator. So, therefore, the Dirac equation once more derivative is acting on that, it should better reduce down to the Klein–Gordon operator because no other operator is special relativity compatible. So, right hand side I should get $-\hbar_0^2c^2$ times $\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$. that is all I should not get any cross term I should not have ∂/∂single derivative term and I should not have a mixed derivative in two different variables are there for example I should not have a $\partial^2$of ∂x ∂y $\partial^2$of ∂x ∂x ∂z or $\partial^2$of ∂y ∂z those terms should not be there so therefore that will put restrictions on the coefficients first if I want this.

Double derivative terms should be equal to either ∂_x^2 *or* ∂_y^2 *or* ∂_z^2. that means this coefficient should better sum up to Kronecker delta, twice of Kronecker delta. Twice will cancel the half and Kronecker delta will make sure that these two quantities appearing here, x_l and x_a are the same thing. They should not be mixed. So, the first demand of a Klein–Gordon structures of this kind tells me that the coefficients $\alpha_i\alpha_j$ has to satisfy this summation rule.Secondly, we do not want in a Klein–Gordon equation there is no single derivative term. So, therefore this single derivative term should vanish that means coefficient should also vanish and we will get The second demand that $\alpha_i\beta$ and β times α_i should be 0.

So, that is the second demand that there should not be a single derivative term. And lastly, whatever was over here, in the Klein–Gordon equation, right hand side is just $m_0^2c^2$ or $m_0^2c^4$ if you have multiplied c^2 in all the terms.



Conslusion: $\hat{\alpha_i}$ + β ˆcan only be matrices which also transform under LT.

Then in

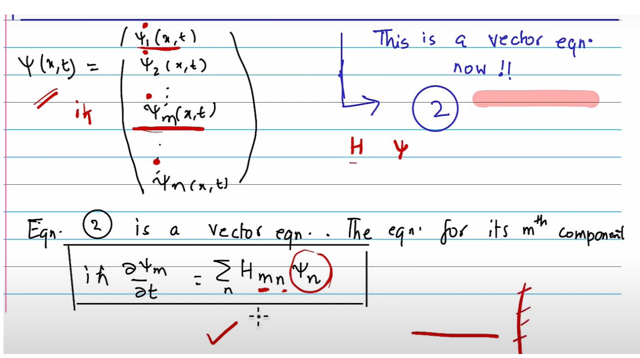
$$i\hbar\frac{\partial\psi}{\partial t}=\frac{\hbar c}{i}\Big(\sum_i \hat{\alpha}_i\frac{\partial}{\partial x_i}+\hat{\beta}\, m_0 c^2\Big)\psi$$

This becomes matrix mulplied with derivative operators.

Hence, ψ will become a clulmn vector.

$$\rightarrow \quad i\hbar\frac{\partial\psi}{\partial t}=\frac{\hbar c}{i}\Big(\sum_i \hat{\alpha}_i\frac{\partial}{\partial x_i}+\hat{\beta}\, m_0 c^2\Big)\psi$$

→ n×n matrix

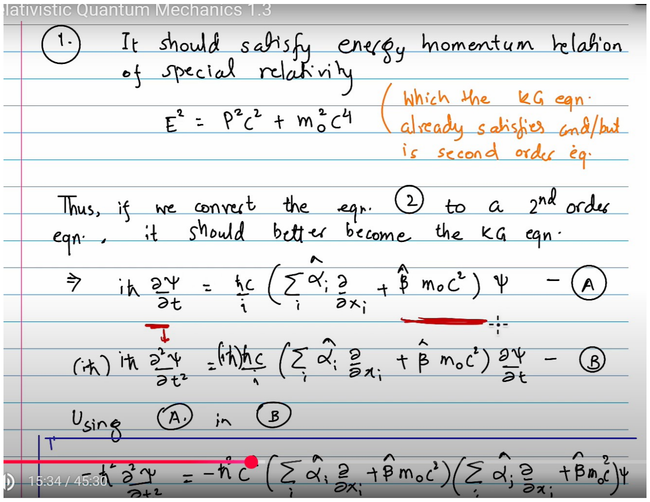


$$\begin{pmatrix} \dot{\psi}_1(x,t) \\ \dot{\psi}_2(x,t) \\ \cdot \\ \cdot \\ \cdot \\ \dot{\psi}_m(x,t) \\ \cdot \\ \cdot \\ \cdot \\ \dot{\psi}_n(x,t) \end{pmatrix}$$

This is the vector equation now!! (2)

Eqn. (2) is a vector eqn. . The eqn for it's m^th^ component

$$i\hbar \frac{\partial \psi_m}{\partial t} = \sum_n H_{mn}\psi_n$$



1. It should satisfy energy momentum relation of specila relativity.

$$E^2 = p^2 c^2 + m_0^2 c^4$$

(Which the KG eqn. already satisfies and/but is second order eq.)

Thus, if we convert the eqn (2) to a second order eqn., it should better become the KG eqn.

$$\rightarrow \quad i\hbar \frac{\partial \psi}{\partial t} = \frac{\hbar c}{i}\left(\sum_i \hat{\alpha}_i \frac{\partial}{\partial x_i} + \hat{\beta}\, m_0 c^2\right)\psi \; \ldots . \quad (A)$$

$$(i\hbar)\, i\hbar \frac{\partial^2 \psi}{\partial t^2} = (i\hbar)\frac{\hbar c}{i}\left(\sum_i \hat{\alpha}_i \frac{\partial}{\partial x_i} + \hat{\beta}\, m_0 c^2\right)\psi \; \ldots . \quad (B)$$

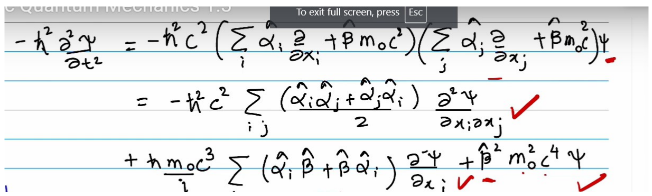
Using (A) in (B) have,

$$-\hbar^2 \frac{\partial^{2\psi}}{\partial t^2} = \hbar^{2c}\left(\sum_i \hat{\alpha}_i \frac{\partial}{\partial x_i} + \hat{\beta}\, m_0 c^2\right)\psi .$$

And lastly, whatever was over here, in the Klein–Gordon equation, right hand side is just $m_0^2c^2$ *or* $m_0^2c^4$ if you have multiplied c^2 in all the terms. So, therefore, the coefficient of that should be 1. that means identity. So β^2 should be identity. Again if I look at this first equation, when j is equal to i, right hand side will become twice of Kronecker delta iI and therefore Kronecker delta of same argument is 1. So twice of identity.
And left hand side will become α_i times α_i which is $\alpha_i^2 + \alpha_i{\cdot}\alpha_i$ which is again α_i^2. So, left hand side will become twice of α_i^2 that is equal to twice of identity that means α_i^2 should be identity.

α_1^2 should be identity, α_2^2should be identity, α_3^2 should be identity. So, therefore, α_i^2 α should all be equal to identity or 1 if you want and β_i^2 should all be identity or one from here you can see the structure that they these $\alpha_i\beta_i$'s cannot be real numbers only numbers because if or for a real number if α_i^2 is one if α_i^2 is one and β_i^2 is also one. Then this equation which is $\alpha_I\beta$ and $\beta\alpha_I$ which is just order. So two numbers which are not 0, they should not, their product will not be 0 after flipping them for real numbers. If suppose it was 1 not identity, then α_I has a choice of either +1 or –1 and similarly β_i has a choice of +1 or –1. The second equality can never be received for any of the combination. So, that means it can never be ordinary numbers. It has to be matrices. Matrices have this possibility that their squares being identity, they can still anti–commute to become 0. This is anti–commutative. So, we are looking for them two matrices, set of matrices, four matrices, α_i's, β's such that their squares are identity while they anti–commute to 0, that they are anti–commutative matrices. So, this cannot be achieved by real numbers.
If I want to have a second derivative structure which is also special relativity compatible, $\alpha_I\,\beta$'s cannot be c numbers, cannot be real or complex number. Therefore, the vector r times the gradient which we

wrote initially cannot be ordinary vector. It is matrix valued vector. So, we have figured out it should be matrix. What kind of matrix? What is the dimension of a matrix?
Should also be very clear to us from this demand. First of all, if I want my Hamiltonian which is made from α and β is Hermitian. So, remember the Hamiltonian which we have written is simply this. If I want the whole combination to be Hermitian, there should be some demands on Hermeticity on α and β.



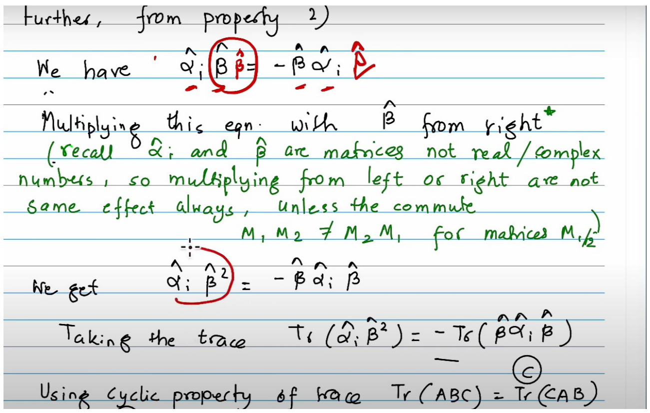
$$-\hbar^2 \frac{\partial^2 \psi}{\partial t^2} = -\hbar^2 c^2 \left(\sum_i \hat{\alpha}_i \frac{\partial}{\partial x_i} + \hat{\beta} m_0 c^2\right)\left(\sum_j \hat{\alpha}_j \frac{\partial}{\partial x_j} + \hat{\beta} m_0 c^2\right)\psi$$

$$= \quad -\hbar^2 c^2 \left(\frac{\sum_{ij} (\hat{\alpha}_i \hat{\alpha}_j + \hat{\alpha}_j \hat{\alpha}_i)}{2}\right) \frac{\partial^2 \psi}{\partial x_i \partial x_j} + \frac{\hbar m_0 c^3}{i} \sum_{ij} (\hat{\alpha}_i \hat{\beta} + \hat{\beta} \hat{\alpha}_i) \frac{\partial \psi}{\partial x_i} + \hat{\beta}^2 m_0^2 c^4 \psi$$

You remember $i\partial/\partial x$ was a momentum operator in some sense, $-i\hbar\partial/\partial x$ was a momentum operator. Momentum operators are Hermitian. So, if momentum operators are Hermitian which are appearing in this and you want the whole combination also to be Hermitian, the whole Hamiltonian to be Hermitian, you will be led up to the demand that α's and β's should be Hermitian operators themselves and Hermitian matrices themselves. This can be achieved just by looking at the hermiticity property whether or not $\partial/\partial x^\dagger$, $\partial/\partial x^\dagger$ is equal to $\partial/\partial x$ or not. You know it is not because momentum is p dagger is equal to p. So, $\partial/\partial x^\dagger$should be $-\partial/\partial x$. So, therefore the whole Hamiltonian remaining Hermitian will force you to look for Hermitian matrices not just some arbitrary matrices, Hermitian matrices which satisfy these demand. Again this demands cannot be met by real numbers they have to be met by matrices and not just any matrices but Hermitian matrices. So, let us see now.
Coming from these structures, what kind of matrices apart from being Hermitian we are landing up to? So, since we have realized that α's and β's are supposed to be Hermitian matrices, we know that Hermitian matrices have real eigenvalues, so their eigenvalues are real. that means in the diagonal form, I can just write α_i's in their diagonal forms as just diagonal entry elements which are their eigenvalues and similarly β's in their diagonal form will be just their eigenvalues lying across their diagonal elements.

So, let us call the α_{ik}, is the kth eigenvalue. So, you see Since we do not know what is the dimension of matrix, we are assuming it is some n cross n dimensional matrix. What is the size of n, we do not know. So, if it is n cross n dimensional matrix, then it has n eigenvalues in principle it can have. They may be same, some of them may be same, some of them may be different, all of them may be same. But right now, we do not make any specialization, we just say that in principle if it is n cross n dimensional matrix, it could have in principle n eigenvalues. So, let us call a_{ik} as the kth eigenvalue of α_i. Remember, there are three α_i's, α_1, α_2, α_3.
And all of them have n different eigenvalues. So, a_{ik} will tell you which matrix I am talking about. The I index will tell me which α's I am talking about. And k index will tell me which is the eigenvalue I am talking about. Second eigenvalue, third eigenvalue or fourth eigenvalue. So α_{ik} is the k^{th}

eigenvalue of ith α. β does not have such a discrepancy because there is one single β.
So β_k will just be the kth eigenvalue of b. So remember in this α_{ik} upper index will tell you which α we are talking about, lower index will tell me which eigenvalue of that α_i's and β_k only which eigenvalue of β we are talking about. So, Hermitian matrices, real eigenvalues, three α's and one β. So, this is what the structure we have.
Now, let us try to fix what should be the dimension of that. Now, we need to have four different matrices, α_1, α_2, α_3 and β.
So, this is what the structure we have.



$$i\hbar\frac{\partial\psi}{\partial t}=\frac{\hbar c}{i}\left(\sum_i \hat{\alpha}_i\frac{\partial}{\partial x_i}+\hat{\beta}\, m_0 c^2\right)\psi$$

Multiplying this eqn. $\hat{\beta}$ from right★
(Recall $\hat{\alpha}_i$ and $\hat{\beta}$ are matrices not real/complex numbers, so multiplying from left or right are not same effect always, unless they commute.

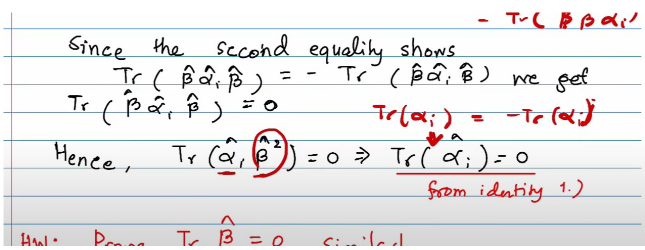
$M_1M_2 \neq M_2M_1$ *for* matrices $M_{1/2}$

We get $\hat{\alpha}_i\hat{\beta}^2 = -\hat{\beta}\hat{\alpha}_i\hat{\beta}$

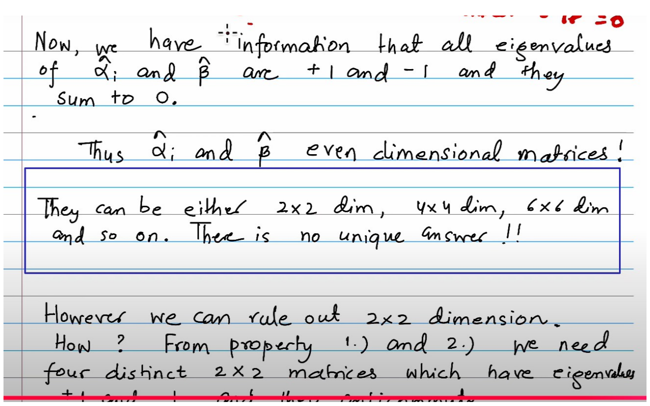
Taking the trace $Tr(\hat{\alpha}\hat{\beta}^2) = -Tr(\hat{\beta}\hat{\alpha}_i\hat{\beta})$

C

Using cyclic property of trace $Tr(ABC) = Tr(CAB)$



Since the second equality shows
$Tr(\hat{\beta}\hat{\alpha}_i\hat{\beta}) = -Tr(\hat{\beta}\hat{\alpha}_i\hat{\beta})$, we get
$Tr(\hat{\beta}\hat{\alpha}_i\hat{\beta}) = 0$
Henec, $Tr(\hat{\alpha}\hat{\beta}^2) = 0 \rightarrow Tr(\hat{\alpha}) = 0$
from identity 1)



Now we have information tha t alll eigenvaluesw of $\hat{\alpha}_i$ and $\hat{\beta}$ are +1 and –1 and they sum to 0.
Thus $\hat{\alpha}_i$ and $\hat{\beta}_i$ are even dimensional matrices!
Tehy can be either 2×2 dimension. How? From property 1) and 2) we need four distinct 2×2 matrices which have eigen values.

Now, let us try to fix what should be the dimension of that. Now, we need to have four different matrices, α_1, α_2, α_3 and β. We need four different matrices. And all four of them as we have seen need to satisfy these set of demands that their anti–commutator should be twice of delta ij times identity. Their anti–commutator of α and β's should be 0 and α^2 and β^2 should be identity. This is what we had realized realizing that they are matrices and not real numbers. And also we do not want all these α_1, α_2, α_3 to be the same We need distinct eigenvalue, distinct matrices. α_1 should be different from α_2 and α_3. We need to have four different matrices, not just α_1 and this is twice of α_1, this is thrice of α_1 and this is four times of α_1.

I do not want one matrix and multiplication by constant into other matrices.

So, four linearly independent matrices we want. So, that will happen only if they are linearly independent in some way. First of all from this structure which we have property number 3. We know that α_i^2 and β_i^2 should be β^2. There is no β_i. β^2 should be identity. that means if I write them in their diagonal form, then also they should be true. The diagonal form square should be identity. Diagonal form square is very simple. All their eigenvalues squared should be equal to identity. Identity is what? 1, 1, 1, 1 across the diagonal element.

that means all the eigenvalues squares are one α_{ik}^2 is equal to one for all k and all I doesn't matter which α.I am talking about doesn't matter which eigenvalue I am talking about for every α and each eigenvalue of every α should have this property that their square should be one because demand number three is forcing us to have α^2 even in their diagonal form equal to identity so that means all the eigenvalues of all the α's can either be +1 or could be –1 only. There is no other possibility. Since eigenvalues are supposed to be real not complex and those real numbers square should be equal to 1, then I am left with only possibility that it could be either +1 or could be –1. It is not forcing that all of α in case should be the same. Some of them could be +1 and some of them could be –1, but no other choice. Similarly, for β_k since β_k also has to equal to its square equaling to identity even in their diagonal form it would be true and therefore, eigenvalues of β should all be squaring up to 1 that means it could be either +1 or –1. So, now we have more information that α, β apart from being matrices, hermitian matrices their eigenvalues are +1 or –1 and nothing else. And one more step, look at property number 2. Property number 2 was telling me that α's and β, all the α's individually, α_1 should anti–commute with β, α_2 should anti–commute with β and α_3 should also anti–commute with β. So, that means I have this structure, that α matrix multiplying β matrix should be equal to the other way, β matrix and α matrix multiplied together with a – sign. Now, since this is a matrix equation, what I can do, I can multiply this equation with β from the right hand side. I put a β here and I put a β here as well.

This side will become β^2 and β^2 is identity. So, identity would be converting into 1.

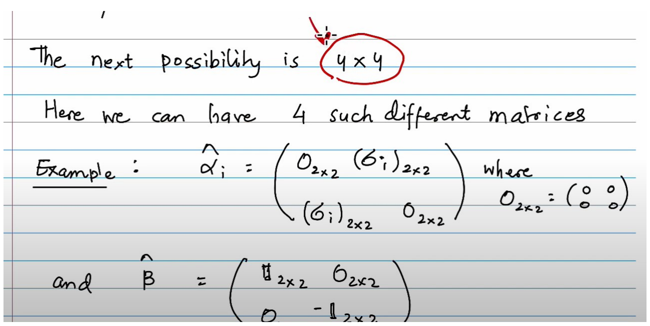
So, this I will have a identity like this, $\alpha_i\beta^2$ and right hand side will be $\beta\ \alpha_i\beta$, $\beta\ \alpha_i\beta$

Since these are matrices, I cannot flip them. For matrices *M1, M2* as we know is not always equal to M2, M1 unless they are commuting matrices. I do not know if they are commuting Hermitian matrices or not. I just know that they are Hermitian matrices. So, I cannot flip them.

So, left hand side β^2 will be obtained which is manifestly there, right hand side I will have a 2 β, one on the left of α and one on the right of α and I cannot do anything about it. So, I have a $\alpha_i\beta^2$ in the left hand side is equal to $-\beta\alpha_i\beta$. So, this equation will be obtained. Then what I can do, this is a matrix after all, a matrix multiplying another matrix is and right hand side also three matrices are there.

So, what I can do, I can take trace of this matrix. This is a resultant n cross n dimensional matrix and this is also a trace of resultant n cross n dimensional matrix. Since, so that will give me this equation trace of $\alpha_i\beta^2$ should be negative of the trace of this triple product $\beta\alpha_i\beta$. So, I could not bring β close to

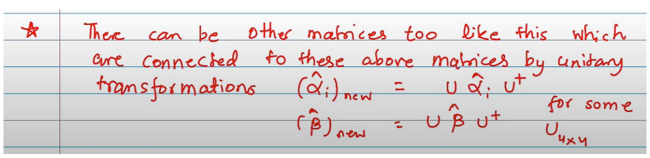
this before the taking trace. But in trace I know, trace of matrices satisfy an interesting property that trace of A, B, C, three matrices multiplied together is equal to the trace of their cyclic rotation. that means if I rotate it by one bit, C goes to location of A, A goes to location of B and B goes to the location of C, then the trace remains the same. So, trace of A, B, C is also equal to trace of C, A, B. So, that means if I call this as a A, if I call this as a B, if I call this as a C, that means under the trace I cannot rotate them. So, then CAB that will become the trace of, CAB will become trace of $\beta\beta\alpha_i$, again a β^2 which has appeared here. So that means trace of $\alpha_i\beta^2$ should be equal to –of the trace of β^2 α_i. And since β^2s are all identity, this will go to identity, this will go to identity. So that means trace of α_i is equal to –of trace of α_i. So trace of α_i is equal to $-Trace$ α_i. This is also could be written as trace of α_i times β^2. So, this equation tells me trace of α twice of Trace $\{\alpha_i s\}$ 0 that means trace of α should be 0. So, therefore, what we have realized from property number 2 and cyclic rotation of matrices under trace that α_i s have a property that their trace is 0.
So four things now they are matrices they are hermitian matrices their eigenvalues are + one or –one and their trace is zero the same thing you can prove this is a homework exercise the same thing you can prove that β also satisfy the same property the trace of β is zero so that means this is true for all α's and β's that they are matrices they are hermitian matrices. Their eigenvalues are +1 and –1 and their trace is 0. that means if a matrix whose eigenvalues can only be +1 or –1, its trace can only be 0 if equal numbers of + eigenvalues and equal numbers of negative eigenvalues are present. So, that means it has to be even dimensional matrix and odd dimensional matrix whose eigenvalues are either +1 or –1 would not have a trace 0. It would be either trace 1 or trace –1 or any other odd number, but not 0.



The next possibility is 4×4.
Hence we can have 4 such different matrices
Example :

$$\begin{pmatrix} O_{2\times2} & (\sigma_i)_{2\times2} \\ (\sigma_i)_{2\times2} & O_{2\times2} \end{pmatrix} where\, O2\times2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

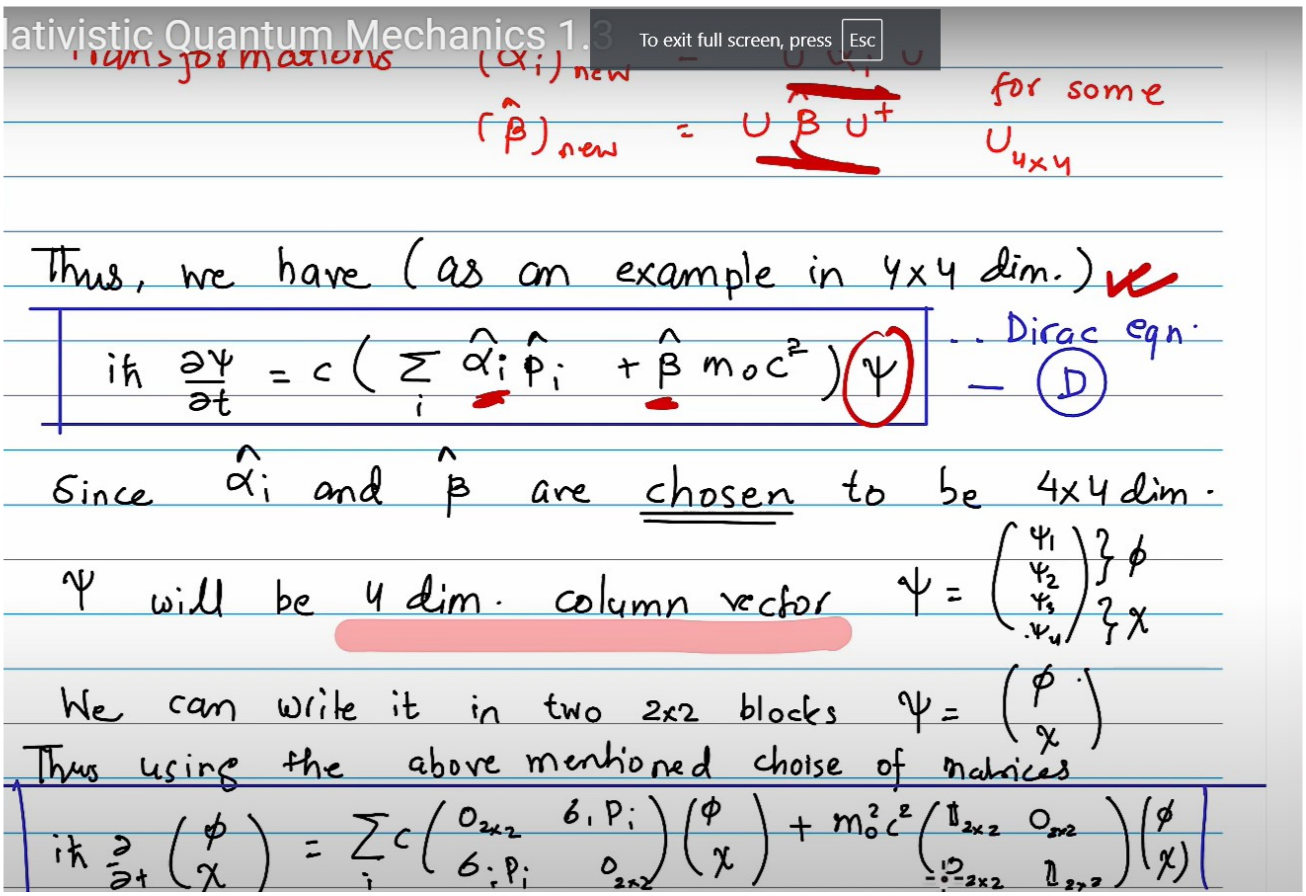
and $\beta^{\wedge} = \begin{pmatrix} \square_{2\times2} & O_{2\times2} \\ O_{2\times2} & \square_{2\times2} \end{pmatrix} where\, O2\times2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$



★There can be other matrices too like this which are connected to these above matrices by unitary transformations $(\hat{\alpha_i})_{new} = U\alpha^{\wedge}U^{\dagger}$
$(\beta) = U\beta^{\wedge}U^{\dagger}$ *for some* $U_{4\times4}$

It would be either trace 1 or trace –1 or any other odd number, but not 0. In order to have a trace 0, you should have an even dimensional matrix. Whether it can be 2x2 dimensional matrix, 4x4 dimensional matrix, 6x6 dimensional matrix or 8x8 dimensional matrix, so on.v Till now, we have no unique answer that what set of matrices will satisfy these three properties, the three properties which we have listed. Actually, these three properties are not sufficient to give you unique matrices. They will give you only possible candidate which can do that and there can be many many candidates which satisfy this bill there can be two cross two dimensional matrices which satisfy that I am giving just an example that in two dimensions there are σ_1 σ_2 σ_3 you can verify that they all satisfy the

property they anti–commute amongst each other to twice delta ij. Their diagonal portions are 1 and –1, their eigenvalues are 1 and –1, their trace is 0, they are Hermitian.
So, this looks like in 2×2 dimensional matrix is possible. However, in this case, I did not have a luxury to get a fourth matrix which was β. β should also be linearly independent. One can prove in 2 cross 2 dimensional matrix only these 3 are sufficient to generate any Hermitian matrix whose square is identity. So that means I would not get 4 such matrices, instead we will have only 3 such matrices and I do not want that. We should have at least 4 such independent matrices. And therefore, the next possibility is that it should be 4×4 dimensional matrices. α's and β's cannot have 4 such independent matrices in 2 cross 2 dimension. They cannot belong to 2×2 dimensional matrices space. They have to belong at least 4 cross 4 dimensional matrices.



Thus, we have (as an example of 4×4 dimension)

$$i\hbar\frac{\partial\psi}{\partial t}=\frac{\hbar c}{i}\left(\sum_i \hat{\alpha}_i\frac{\partial}{\partial x_i}+\hat{\beta}\,m_0c^2\right)\psi \quad \ldots(D)$$

Since $\hat{\alpha}$ and $\hat{\beta}$ are chosen to be 4×4 dim.

ψ will be 4 dim. Column vector

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

We can write it in two 2×2 blocks $\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix}$

Thus using the above mentioned choice of matrices

$$i\hbar \frac{\partial}{\partial t}\begin{pmatrix} \phi \\ \chi \end{pmatrix} = c\sum_i \begin{pmatrix} O_{2\times 2} & \sigma_i P_i \\ \sigma_i P_i & O_{2\times 2} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix} + m_0^2 c^2 \begin{pmatrix} \Box_{2\times 2} & O_{2\times 2} \\ O_{2\times 2} & \Box_{2\times 2} \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix}$$

There also we can list down what is the possibilities the three α's can be obtained again from the poly matrices poly matrices along the off diagonal this is 2 cross 2 block of poly matrix this is a 2 + 2 block of the poly matrix this is a 2 + 2 0 matrix which is this and this is 2×2 0 matrix so you see off diagonals are poly matrices diagonals are 0 matrices so this is a 4 cross 4 dimensional matrix and β is identity in the first block diagonal and -identity in the second block diagonal. Now, you can verify that these four matrices are linearly independent, they properly do the anticommutation with respect to α and β and their anticommutation among α and α_is twice of delta ij times identity. in four cross four this is one possibility I have given actually this is the only possibility in four + four dimensional space because there is a theorem which will not prove in this course but you should know that any other choice which you can generate satisfying all these three demands in four four + four dimensions are unitary equivalent to these choices which we have just written.
So, any other set of four matrices in 4×4 dimensional space which satisfy all the three requirements which we will have are just the unitary transformed versions of these matrices. So, therefore, we have learnt that our story can start only with 4×4 dimensional matrices. We will have a Dirac equation which is first order equation compatible with the requirement of special relativity whose second derivative is the Klein–Gordon equation. This can only be generated at least by 4×4 dimensional matrices.

Therefore, this ψ cannot be wave function. It is a collection of 4 wave functions, which is a 4 cross 1 column vector. And that structure is there of the Dirac equation. It is not ordinary wave equation, it is a matrix equation and the quantities appearing here are called Dirac matrices and the wave function which has now become 4 cross 1 dimensional wave function is called a Spiner. So, what kind of properties this spiners and what new features the spiners come up with that we will see in the next class. But here we should realize that the requirement of first derivative equation, requirement of Klein–Gordon at the second level forces us to start discussing α's and β's to be matrices at least of 4 cross 4 dimension because 2 cross 2 dimension they cannot have 4 such α's and β's. So, from here we realize that the ψ which we are talking about has to be replaced by a quarterlet, 4 tuplet. It has 4 entries, ψ_1, ψ_2, ψ_3, ψ_4 . Typically, we have learned to do quantum mechanics with one wave function ψ. But now there is a more broader structure that the wave function is a collection of four functions that together generate one wave function. that means it has more features. So, Dirac equation is not talking about the ordinary quantum mechanics which we have learned. This is talking about a theory of a wave function which is more richer than what we have learned to deal with. What these new functions are ψ_1, ψ_2 , ψ_3, ψ_4 what properties of the full wave function they describe that we will learn in the next class. Okay, so I stop here for this.