

Foundation of Quantum Theory: Relativistic Approach

Dirac Equation I
Prof . Kinjalk Lochan
Department of Physical Sciences
IISER Mohali

Dirac Equation I Lecture- 15

In today's class we will discuss the quantum field theory structure and we will have our first attempt of dealing with fields and their quantization and we will do so by studying the quantization of a most elementary field which is possible which is scalar fields and from with here we will see how various insights developed with scalar fields can be translated to other fields in coming lectures as well okay to remind you. We have till now learned that if I try to do special relativistic treatment to quantum mechanics, we end up on the propositions that wave functions probably or probability densities which we were dealing with start to lose their significance as they were in the Schrodinger picture or Schrodinger equation level.



In quantum field theory the wavefunctions of relativistic quantum theory become space-time dependent operators!

$$\Psi(x, t) \longrightarrow \hat{\Psi}(x, t)$$

- Position and time are demoted to ordinary c-number variables

$$\hat{x} \longrightarrow x$$

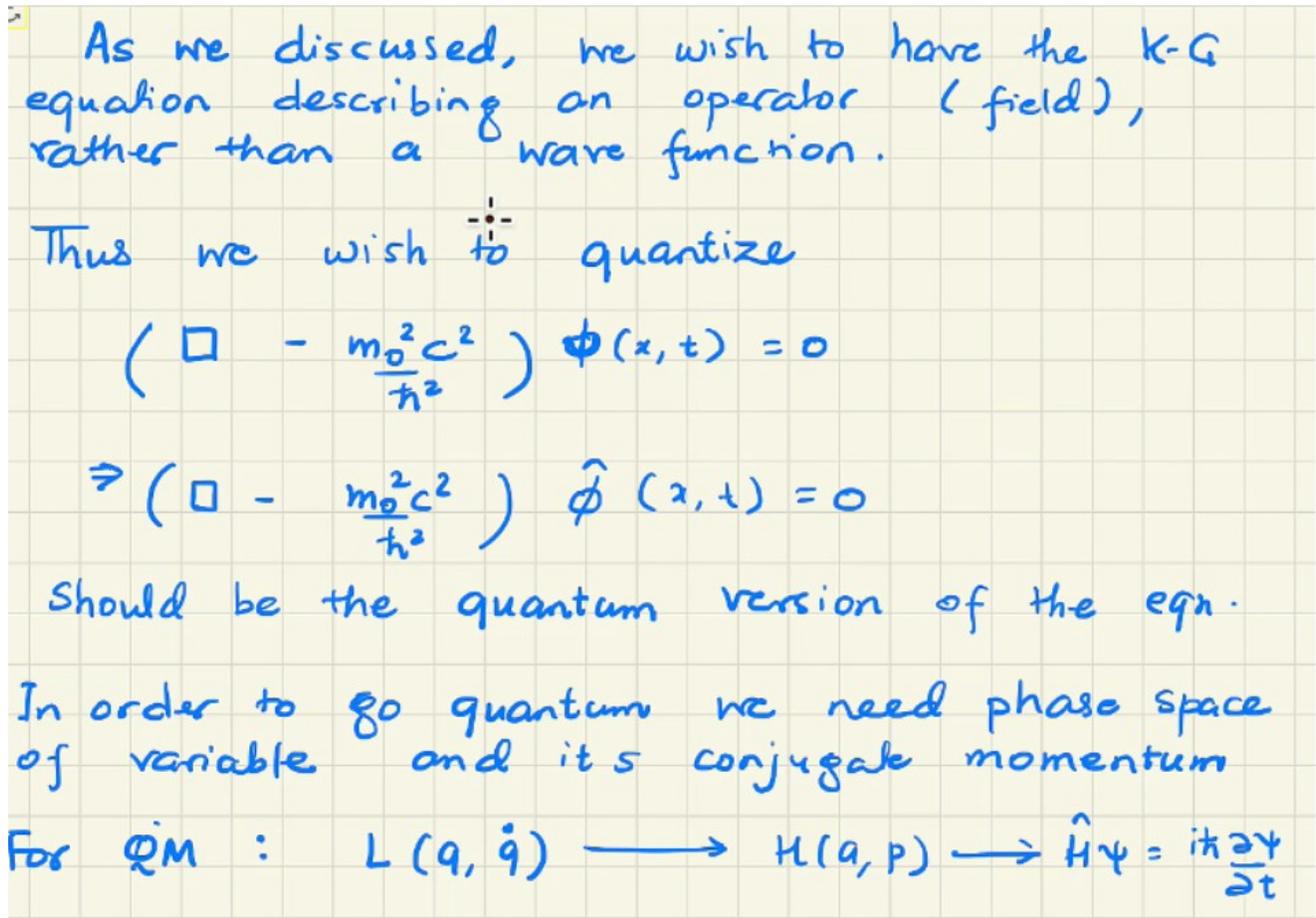
- A field is a space-time dependent variable

- Quantum mechanics is a theory of time dependent variables promoted to operators

$$q(t) \rightarrow \hat{q}(t)$$

At relativistic level, we seldomly have any consistent probability density expression or the wave functions and ontology in terms of providing the wave probability density is at the geopathy. So therefore, in order to get rid of those structure where we have negative probability density which makes the probability theory meaningless. We come up with the idea of a second quantization where the wave function themselves are supposed to be not a functions but promoted to operators so ultimately in order to have still one go at a quantum theory of wave functions we will try to promote the wave function itself into operators. Previously at any location x, t in space and time this wave function was giving me the probability amplitude in some sense of finding the particle around the window of small $\delta(x, t)$. Now, this will not give me a number anymore. It will provide me an operator at location x and t . And in that sense, this position x and position t , therefore, cannot be themselves operators because operator at operators is a meaningless statement. So, we should rather demote the position and time back into c number variables. They should be just variables with continuous real values they contain. So we will talk about positions of space and time not as operators but the fields there are operators and those fields in some limit are supposed to develop the structure of position operator or momentum operator as we will see along the way. So as long as we are Oblivious with the quantum structure of the fields, we are not doing justice to quantum relativity and quantum theory together. And one attempt as we saw was to quantize the wave function or get a field which is a set of operators at all points. Previously $\psi(x,t)$ was a set of complex number at all point and time. Now we have operators at all points and time. What kind of operators? Hermitian operators, non-Hermitian operators, what is their theory, how to find out their structure, expectation values, all those things we will deal with. So, therefore, a field as we discussed is a space time dependent variable, just like $\psi(x,t)$ was a field as well, that means it was giving you complex numbers at location x,t . Now a field could be a different variables as well for example if instead of getting you complex numbers it can give you matrices at all location x and t think of a matrix field each location you go you will get a matrix okay for instance we have seen that happening for $f_{\mu\nu}$ the electromagnetic $f_{\mu\nu}$ tensor that's supposed to give you quantity six quantities at each location which can be written in forms of anti-symmetric matrix. So, similarly there could be other matrices as well which define various parameters at a particular point of space and time. Now, this is just analogy if you want to draw with quantum mechanics in your mind. Quantum mechanics was a theory of time dependent variables Q of t . For example, initially we started with Lagrangian. which was just a theory of time dependent variable not space time dependent Q was just a function of t it was not supposed to be a function of x and later on we promoted $q(t)$ into operators $\hat{q}(t)$ that means each time we will talk of operator position operator at time t momentum operator at time t and so on the same game we are going to play just in this case in this time we will make it a field would become a operator not only at allbut all space points as well so at this location there will be some operator here let us see some operator O_1 at location x_1 at time t_1 and at same time let us say t_1 at some other location x_2 there is some other operator O_{x_2} at time t_1 . So, similarly the collection of all points and all operators together generate for you a field operator. So, thus quantum field theory would be a theory of space time dependent variables and then they are promoted to operators. So, we will see the artifact of turning on one more dependent variable. Previously, it was just a function of t . Now, we want it to be changing from point to point in space as well. And that leads to a field theory that we will try to see where do we land up with. So, now let us start with the problem statement we had. We discussed that

we wish to have a Klein-Gordon kind of equation, but this time describing an operator or a field. Our attempt of Klein-Gordon equation describing a wave function has failed in terms of various inconsistencies with probability theory. So therefore, we want to salvage the situation by making them operators. So previously we had this kind of thing. In this notations, I am trying to write all the wave functions, Klein-Gordon wave function with symbol ϕ rather than a ψ . This is done to just match with the notations of textbooks which you will find that most of those fields are discussed with symbol ϕ rather than a ψ . So, we wanted a Klein-Gordon equation which was doing this to wave functions.



As we discussed, we wish to have the K-G equation describing an operator (field), rather than a wave function.

Thus we wish to quantize

$$\left(\square - \frac{m_0^2 c^2}{\hbar^2} \right) \phi(x, t) = 0$$

$$\rightarrow \left(\square - \frac{m_0^2 c^2}{\hbar^2} \right) \hat{\phi}(x, t) = 0$$

should be quantum version of the eqn.

In order to go equation we need phase space of variable and its conjugate momentum.

$$\text{For QM : } L(q, \dot{q}) \rightarrow H(q, p) \rightarrow \hat{H}\psi = i\hbar \frac{\partial \psi}{\partial t}$$

Thus we need to identify momentum conjugate
to $\phi(x, t)$: $\pi(x, t)$

Find operators made up of (ϕ, π)

- their expectations
- their fluctuations

Thus we need to identify momentum conjugate to $\phi(x, t)$: $\Pi(x, t)$

Find operators made up of ϕ, π .

- Their expectations

- Their fluctuations

-

Now, since we have promoted the wave function to fields, now the same equation will be talking about the dynamics space time derivatives of a field. So, this is the quantum version of the Klein-Gordon equation. Now in order to go quantum we wanted to know the phase space of a theory this has happened even at the quantum mechanics level initially we had some lagrangian let us say of a free particle which was $1/2\dot{q}^2$ if it has it is not a free particle but it lives in a potential then lagrangian is given as a $1/2\dot{q}^2 - V(Q)$ that is the traditional thing which we know but the quantization does not happen in the space of Q and \dot{q} which is identified as the configuration space variable. But quantization happens in the space of Q and p , its conjugate momenta Q and its conjugate momenta p . This space is called the phase space. And there something like a Hamiltonian is defined. And that Hamiltonian ultimately tells you the rate of change of the state you are looking at. So two things are needed. Whatever your variables is, You have to write down first the Lagrangian in terms of the variable itself and its derivative. And then you have to jump to phase space where you will trade off \dot{q} will be replaced by p by something called a Legendre transformation as you might have known. And the Q will become an element of the phase space. The same thing we have to do for field now, previously we had done for a single field dependent variable object q , Q was a function of t only, now we have a ϕ which is a function of x and t . So, therefore we have two jobs at hand, we have to find out what is the momentum conjugate to ϕ . I am giving you some symbol π of x of t , just like p was conjugate momenta corresponding to q , Π is conjugate field corresponding to ϕ .

Previously, p was to be obtainable from the Lagrangian by taking the derivative of the Lagrangian with respect to \dot{q} . This time also, we want to know about the dynamics which is driven by the Hamiltonian. Therefore, I would want to know the momentum corresponding to a field by taking the derivative of a Lagrangian with respect to $\dot{\phi}$. However, we have to be slightly careful in taking the derivatives with respect to derivatives of a multivariable function. This previously it was just a derivative with respect to a single variable function and the similar kind of logics we have to follow with slight care at hand in dealing with those functions. So, once we have identified what are the phase space variable ϕ and its conjugate field Π .

We would try to operatorize them just like we had operatorized Q and p in quantum domain we had

operatorized all the functions which are functions of Q and p for example momentum angular momentum hamiltonian they all become operators similarly we would write down operators made up of ϕ pi or their mixtures and therefore all these things expectations their fluctuations and other things will come about just like all the operators in quantum mechanics made up of Q and p had some fluctuations and expectations. So this is what we are supposed to do. So let us get going with this. So in the first attempt of doing that, we need to know the Lagrangian corresponding to which the theory of field would be described and then we have to take the derivative of the Lagrangian with respect to $\dot{\phi}$ in order to get the momentum just like we used to do in quantum mechanics. So, again to remind Lagrangian for quantum mechanics was Q made up of Q and its first derivative with respect to time. Now since we are doing dealing evolution not only in time but in space as well the \dot{q} which is just first derivative d or zero thought or derivative dq/dx_0 now in special relativity compatible theory derivative with respect to time is not special because I know under logic transformation time and space mix up. So, better I should have a derivative with respect to all the position coordinates as well $dq/dx_1, dq/dx_2, dq/dx_3$ and in order to have that to happen Q should also depend on not only on t but also on locations x_1, x_2, x_3, xyz and therefore we will keep calling it ϕ rather than a q . So, therefore the Lagrangian now would be a function of the field and its all derivatives not only time derivative but all the spatial derivatives as well. So, the Lagrangian which we had previously obtained for quantum mechanical system or even classical particle system would be integration dt of something called a Lagrangian which is function of Q and \dot{q} and this integration will happen from some initial time t in to some final time t_f or t final and we know what are the boundary conditions existing at initial time and final time okay now the same game we have to play for a particle not for a particle but for a field so previously we had variation in time only because only thing the dynamical variable was changing was in time. This time the dynamical variable is changing not only in time but in space as well so we will have a additional spatial integral spatial integral of the lagrangian made up of field and its derivative by the virtue that I have to integrate spatial variations as well spatial integrations as well that makes this object which is which is curly l not as a lagrangian but a lagrangian density because integrating over the space completely will give me a function which is legitimately called Lagrangian. Remember, action is just $L dt$. Now, if I call, if I generalize to double integral of dt and d^3x , the quantity d^3x with a single integral in x , sorry, the triple integral in x is performed over this Lagrangian density and whatever quantity is left out is yet to be integrated with respect to T , that quantity will be called Lagrangian. So, therefore, the integrand of d^3x will be called Lagrangian density. You have to keep in mind that typically it is called its name is Lagrangian density but many of thein the discussions of quantum fields this Lagrangian density is loosely called as Lagrangian but keep in mind this is Lagrangian density it has to undergo integration with respect to all space to give you Lagrangian and that quantity will further undergo integration with respect to all time to give you an action all right so with this identification at hand. Now we will try to do principle of least action or variation of action in order to locate the extremal positions of the fields.



Now

$$\frac{\delta L}{\delta(\partial_\mu \phi)} \partial_\mu (\delta\phi) = \partial_\mu \left[\frac{\delta L}{\delta(\partial_\mu \phi)} (\delta\phi) \right] - \partial_\mu \left(\frac{\delta L}{\delta(\partial_\mu \phi)} \right) (\delta\phi)$$

$$\delta S = \int \int dt d^4 x \left[\left[\frac{\delta L}{\delta \phi} - \partial_\mu \frac{\partial L}{\partial(\partial_\mu \phi)} \right] \delta\phi - \partial_\mu \left[\frac{\delta L}{\delta(\partial_\mu \phi)} \delta\phi \right] \right]$$

Since $\delta\phi \rightarrow 0$

$$\frac{\delta L(\phi, \partial_\mu \phi)}{\delta\phi} = \frac{\partial L}{\partial \phi}$$

$$\frac{\partial L}{\delta(\partial_\mu \phi)} = \frac{\partial L}{\partial_\mu \phi}$$

That means what values of the field, what configurations of the field generate maximum action or extremal action let us say. So for that to happen I have to do the infinitesimal variation in the action itself that will be obtainable by infinitesimal variation in the Lagrangian density itself. Because action is made up of only the Lagrangian density, only way the action can change is through the change in the integrand. So, that would be change in the Lagrangian density. Fine, but I know the Lagrangian density is a function of ϕ and all its derivatives. So the only way the Lagrangian density can change is through either ϕ changes by some amount which we are going to call $\delta\phi$ and therefore the total change the partial change rather in the Lagrangian density due to change in ϕ alone will be $\partial L/\partial\phi$. This is called a functional derivative. This is what you might have seen even in a classical action extremization process. This time it is with respect to space time dependent quantities. But philosophy is same. We will just touch that functional differentiations as well. So this is functional derivative that depicts the change in the Lagrangian due to only change in ϕ without caring for other changes. So that would be partial derivative something like a partial derivative but it is a functional derivative of lagrangian density with respect to change in ϕ and then the amount of change in ϕ itself similarly change in lagrangian density when the derivatives are changed and then the amount of change in the derivatives think of it as a function of two variables f of x y and the total change in function f df is partial derivative of f with respect to xdx + partial derivative of ydy. Only thing is that here x and y are supposed to be real numbers. This time we are talking about functionals changing through changes in functions. So these things are not just real numbers. These are just fields. We have to further supply x positions to give you a real number. $\partial\phi$ themselves is just a function. Okay all right so we move further and we write this thing that this delta of $\partial_\mu\phi$ that is talking about how much change is there brought about infinitesimal change that is brought about in the derivative of ϕ that suppose there was a some function before it had some derivative at all points at this point it has a derivative value. So suppose this is a function $f(x)$. And if I plot its derivatives, f primed of x, I will get a function which is something like that. At this location, function's derivative is 0. It will start with 0.

Here it is negative. Here it is positive. It becomes 0 again here and then it starts becoming positive. So you can see the derivative of f , not the value of f . You calculate derivative of f . This will change along with X . And what this delta is doing is taking this derivative function. It has some plot. Suppose, let me give a nice example rather in order to make things very clear. So let us think of, let us make some space. So let us think of a function which is, let us say f of x is x^2 . Its plot would be something like this.

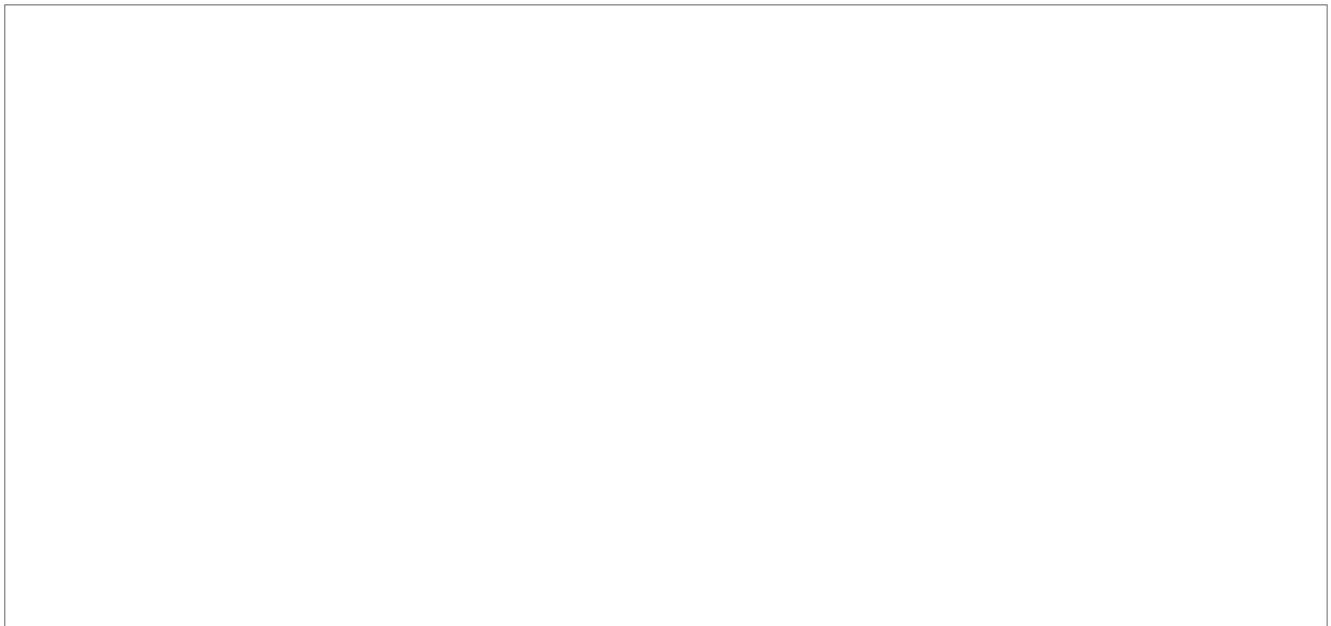


Start from zero and rises in the quadratic fashion. So this is f of x and this is x . Now take its derivative f prime of x that will be twice of x and if I try to plot f prime of x I will get a linear function. It would be a linear function. Now when I try to do infinitesimal variation of derivative of f $\partial f / \partial x$ that means I will try to bring some amount of change infinitesimal amount of change some some different function I will try to add which is very small that will introduce slight wiggleness about this curve so I am going to change it infinitesimally so that is the meaning of this delta of the derivatives all the derivatives have to be marginally changed Now the same thing can be done by first changing the function slightly and then taking the derivative. So this curve here comes from derivative of not x^2 but some other function will be which will be just first integral of this. So let us say something like $\int_0^x g(y) dy$ such that $g(0)$ is 0. Then if I take the derivative. So this is the infinitesimal change in the function and its derivative if I calculate you will get exactly this curve. So, it is first integral antiderivative. Therefore, when you take the derivative you will get this function. So, the statement is infinitesimal variation in the derivative of a function can also be obtained from first changing the function infinitesimally and then taking the derivative. So, that is what we can do. I was supposed to take the derivative of the function and then change it infinitesimally, but I can reverse the order and I can first change the function and then take its derivative. So, therefore, the second term will become partial derivative or space time derivative of the change function $\delta\phi$ is the amount of change in the function ϕ you have done. So, you see the first term is the functional derivative of L . with respect to ϕ itself and then infinitesimal change of ϕ and then the second term is the functional derivative of the Lagrangian density with respect to the gradient of the ϕ that is the derivatives of the ϕ and then ∂_μ the derivative of the infinitesimal change into the function. So, this is what the flip has done. Now as we have seen even in classical mechanics the same *philosophy* will hold here as well. The second term which is just written on the left side here can also

be written as the total partial derivative of this whole quantity – partial derivative hitting the first quantity and leaving the second quantity alone. If you open it up, the first square bracket on the right hand side, you will get two terms.

First the partial derivative hits the function in the red boundary, the red circle and $\delta\phi$ will be left alone and the second term will be the partial derivative hitting the $\delta\phi$ and the first functional derivative of L with respect to $\partial_\mu\phi$ will remain intact. The first term will be exactly cancelled by this and the second term will just give you this function which was already there so left hand side can be obtained from the space time derivative of a product quantity – partial derivative of the first function and second function left alone this you have seen in classical mechanics as well so instead of The second function I write down the both the terms one is the divergence space time divergence it is this time of the whole product quantity which is over here and – from here actually it this this should have been a + over here mine so – this second term over here okay so now you see now you see I have a divergence this is space time divergence this is not just spatial derivatives all the four derivatives mu is equal to zero is time one two three of a product quantity which depends on $\delta\phi$. $\delta\phi$ is infinitesimal variation in the function. And then there is another term which is $\partial L/\partial\phi$ and – the second term. The second term was also proportional to $\delta\phi$. The first term which is surviving here is also proportional to $\delta\phi$. So these two terms can be collected together and overall $\delta\phi$ can be put out. Now you know the Gauss's divergence theorem. That in any space dv of a gradient of certain function will take you to the surface of the volume and dot product with the divergence of a vector rather. Divergence of a vector will take you to the surface with a vector dotted with the area element. So, same thing will happen with this supposedly space-time vector here. It would be taken to the boundary of this integral t tending to t_1 to t_2 , the boundary of t , t initial to t final and boundary of a space which is presumably at infinity. But one thing I know that at the boundary of time t initial and t final I exactly know what field configuration was there I do not know what is the field configuration in between I am trying to find out a field configuration which extremizes the action but I do know what is the field configuration at the boundary so therefore $\delta\phi$ at the boundary of time will be 0 because I do not want to bring any variation which changes my boundary conditions. So, $\delta\phi$ is supposed to be 0 at t initial and t final and that is what this integral will project u2. So, therefore, the extra divergence term which we have added is actually harmless because under integration it contributes nothing. So therefore, you are left with only, so let me clean it up once more. So that means once we had written the whole variation in the action that became made up of two pieces, one term proportional to delta pi and another divergence. The divergence term is useless because under integration it gives you 0. That means delta s is just proportional to this term, the first term and that we want to be 0 because extremal action will want you to have delta s is equal to 0. So that means this whole integral should be 0 and we want it to be true for all possible $\delta\phi$ that means whatever variation I do infinitesimal variation I do I should get the extremal action that means I am not looking for a solution of $\delta\phi$ I am rather looking for any arbitrary $\delta\phi$ variation remember what was $\delta\phi$ $\delta\phi$ was a variation in the already some known function trajectory so for any arbitrary variation I want this integral to vanish this can only happen if the integrand vanishes so see what is the integrand it is $\partial L/\partial\phi$ – the ∂_μ term hitting the first function. The first function was functional derivative of Lagrangian with respect to $\partial_\mu\phi$. So, therefore this should be equal to 0 and that is what the equation of your equation of motion is. So, ultimately you have to be just put you have to just put this square bracket is equal to 0. In this case, you have a functional derivative of L with respect to ϕ and functional derivative of L with respect to $\partial_\mu\phi$. We will see soon that for some simple n of Lagrangian densities, this could be replaced by partial derivative of Lagrangian density with ϕ and partial derivative of Lagrangian density with $\partial_\mu\phi$. So, ∂ of $\partial_\mu\phi$. So, we will soon see, but once this happens that means we are going to get this quantity, this quantity over here square bracket over here is equal to 0 will be my equation of motion. The same thing was there for quantum mechanical classical mechanics case where mu was only one index time so ∂t /so you if you recall the discussion you will realize that we are if mu is just restricted to time you are just obtaining the euler lagrangian equation of classical particles this time we are getting

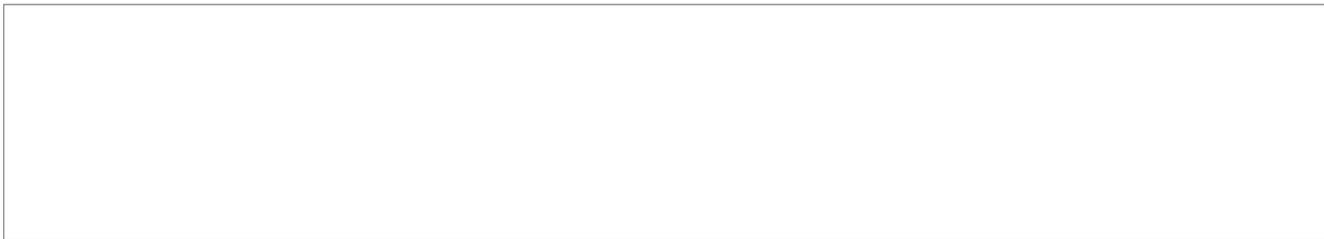
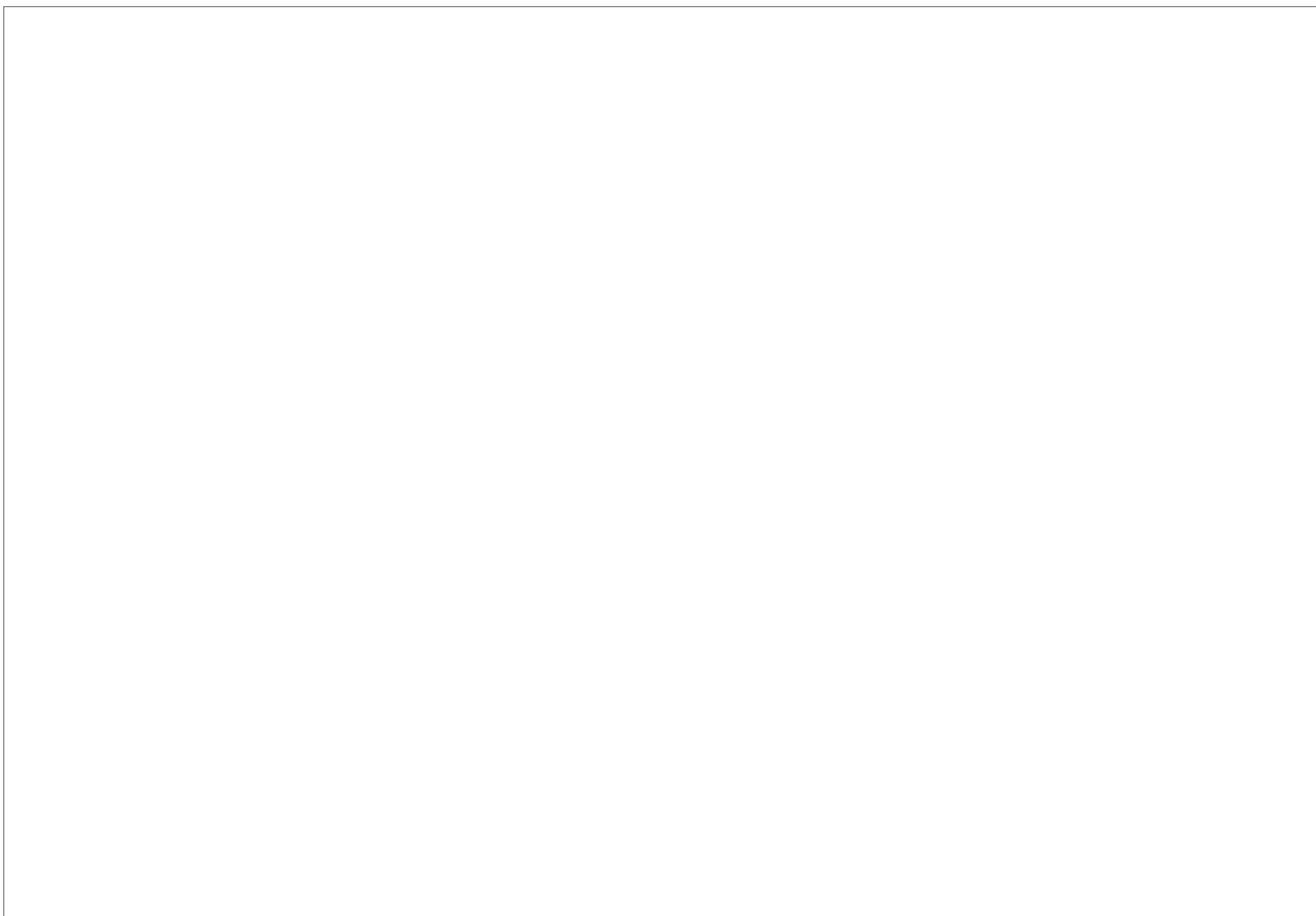
euler lagrangian equations of motion for fields we are not yet true to the quantumness of that but we will just see from equation of motion what we have to do to quantize things as we had discussed the same thing for classical particles as well. Let us conceptualize and visualize the variation in the field with respect to extremization of the action. So I just would remind you when we were discussing, when you had discussed in classical mechanics the extremization of an action for a classical particle, what was being done was suppose it was given to you that the particle is at the boundaries at time t is equal to t_1 , it is at some location Q initial and time t is equal to t_2 , it is at some location Q final.



The extremization process tries to find out the path along which the action becomes extremum. So, let us say there is a blue curve along which from Q initial to Q final during time t_1 to t_2 , you can propagate or move from this point to that point along blue curve, red curve or the green curve or any other curves

as well. And we will find out which of the curves extremizes the action for you and that is the equation of motion or the classical solution along which the particle propagates. And we will find out which of the curves extremizes the action for you and that is the equation of motion or the classical solution along which the particle propagates. Quantum mechanics we know, we have some weightage of all the paths possible and in the limit \hbar tending to 0, the classical path C merges out. So, similarly when we are talking about fields rather than particles, what is known to me is the field configuration at time t is equal to t initial and time t is equal to t final. For example, let us say this axis is the time axis here and this axis is the space axis x . Now at time t is equal to t_1 , so this will be time t is equal to t_1 , the constant line. The field configuration at different-different locations would be this curve, this bold curve which we are plotting. The third axis is the value of the field. For example, at this point, this is the location, some location x_1 , let us say. This in between is some location x_2 . This is some location x_3 . So you see the height of the curve tells you what is the field value at time t_1 . So this is time t_1 at different, different positions. So this is like a snapshot of the field. At time, time is frozen, only space you are moving along and you are seeing what kind of field configuration is generated. Here the field is of this height, this much value, that much value, somewhere else it can go negative as well. So this kind of at time t is equal to t_1 . This is how the field configuration is. This is known to me. Similarly, at t is equal to t final. So this is t is equal to t final. If I move along this dashed axis, I'm moving along the space. Time remains the same. Again, I know what is the final field configuration that if I go to location x_1, x_2, x_3 , what happens to the field? It becomes different from its initial configuration. So what is known to me is $\phi(x), t_1$ which is here and $\phi(x), t_2$. That means I know what is the field existing in all positions at time t_1 and I also know what is the field existing at all location at time t_2 . But I do not know what is the field configurations in between. For example, some intermediate time let us say some t here. I do not know what is the field. It could be this curve, for example, at time t . Okay, so you see time t , this might be this height here at this location. At some other location, this is this height. At some other location, this is this height and so on. It could exist like this and at a future time, it again goes like that. So, this blue kind of surface between these two fixed Fixed bold curves. So these bold curves are final configuration. I know what is there. Here and here I know what is the field configuration. So these are like this kind of shaped rods. And I do not know what has happened to field in between. So it can take any configuration in between. So this is like this blue sheet you can think is all possible positions. This is one not all possible. This is one possible realization. So this blue sheet effectively is a realization of one particular class meaning from initial field configuration to final field configuration. The field can propagate along this blue sheet or it can propagate along this green sheet green sheet is like this that initially you start with this then you go down at a future time and then you come up again at a future time so look at one particular location so at this location what is happening initially field was this value here and I move into future meaning along this axis so let's draw a parallel so at the same location the field becomes negative. Then you wait further at the same location field becomes less negative sometimes it becomes zero then it becomes negative again then becomes negative then it becomes positive so this is like one variation at one point similarly at some other point you can draw axis parallel to time and find out what is the field configuration a different field configuration like is like a different sheet between this initial and final thing see in these two curves here and here the boundaries are the same the field configurations at the boundary are the same only thing changing is in between the configuration so that is like shape of the sheet in between is changing this is like this sheet is this shape the green sheet is this shape similarly there could be a red sheet which could be a different shape so in principle there are infinitely many shapes possible with these two boundaries. You can more easily visualize this as between two fixed rod you have put up again a sheet and just make it vibrate meaning whatever shape you could generate those are all possible field configurations at differentat various locations axis. So the analog of quantum a classical particle along one path is fields different fields configuration between settling into a final field configuration into from the initial field configuration. Again a visualization you can think of that initially I know at time t is equal to 0, temperature is let us

say $1/\text{some } t_0/1 + \alpha \text{ square } x^2$. This is how temperature changes at time t is equal to 0 along the distance x . And t is equal to final time, let us say t_f . The temperature, t initial was this, t final was this. is some $T_0 \text{ tilde}/1 + |\alpha x^2|$, let us say. So, this is the final temperature at different, different locations. But you do not know what has happened in between. It could have come from here to there along one particular T_x and T_x of t or it could have been a different functions in between. So those different functions are these visualizations of these curtains or sheets between the fixed boundaries. So all possible boundaries or with the same boundary all possible variation that means all possible shapes of this sheet are analogous to this different curves here. So we want to know what is the optimal sheet shape which extremizes the action. And that is what the action for extremization process does just to realize we have landed up with this condition when I demand the extremization of δS while sending these things to zero the boundary terms are sent to zero because at the boundary I exactly know what the field configuration are only in between again in the diagram as we saw the field variations are not changing things on the boundaries the bold curves remains the same only thing changes is in between so $\delta\phi$ I am not going to have anything at time t_1 and t_2 which are initial time and final time So, I am left with only the first term which is this square bracket here. So, let me just clean it up so that we can see it more clearly. So, the boundary term is thrown out only the first square term which is remaining is the action extremization equation which is $\partial L/\partial\phi - \partial_\mu L/\partial\mu$ of file this is supposed to be zero four this is supposed to be zero four equations of motion for a field here also μ and μ here are coming repeated and they are along the diagonal that means there is a soft sum implied that μ runs from 0 1 2 3 okay so I have not written it in the consistent notations of Einstein's summation convention which we had previously discussed as well that if same index appears in the upper and lower both that means there is a sum implied. Why there was a sum? Because there are four derivatives $\partial_0\phi, \partial_1\phi, \partial_2\phi, \partial_3\phi$. So, the variation of the action through the Lagrangian density should come not only with field variation, it is temporal derivative variation, spatial derivative variations all these things. So, here the soft sum in μ is implied. When I had written like this, all the partial derivatives of fields were appearing and variation with respect to that was taken care of. So, this is the equation of motion for a variation of a field. We will now try to understand what is the functional derivative doing and we will just visualize how this claim works that this functional derivative for simple enough Lagrangians can be used as partial derivative for elementary calculation purpose. So that we will see in the following discussion. So let us look at the properties of functional derivative through the discussions of functional so you see g over here as we discussed is a functional a functional is a map which takes a function in and gives a number out for example this $f(x)$ is a function function takes a number x and gives you a real number or a complex number after feeding it a number for example x^2 is a function because if I put x is equal to 5 I will get a number 25 however this functional is a map from space of function to real number that means it eats up a function and then gives a number you do not have to feed it a point point x you do not have to give information of x you have to give information of the function for instance let us take one example a functional g of f which is obtainable from integrating the function $f(x)$ to the power n from let us say a to b .



$$G[f] = \int dx (f(x))^n$$
$$f(x) \rightarrow f(x) + \epsilon \delta(x - x_0)$$

$$\delta G = \int dx ([f(x) + \epsilon \delta(x - x_0)]^n - [f(x)]^n)$$
$$= \int dx [(f(x))^n + n \epsilon \delta(x - x_0) (f(x))^{n-1} - (f(x))^n]$$

$$n(f(x))^{n-1} = \frac{\delta}{\delta f} (f(x))^n = \lim_{\delta f \rightarrow 0} \left[\frac{(f(x) + \delta f)^n - f^n}{\delta f} \right]$$
$$= \frac{\partial (f(x))^n}{\partial f(x)}$$

So whatever you feed in here, I will just take n th power of that and integrate that function from a to b . So for example, if you give me x^2 , $f(x)$ is equal to x^2 , then I will do x to the power $2n$ from a to b and I will get answer $b^{(2n+1)/2n+1} - a^{(2n+1)/2n+1}$. Otherwise, if you give me another different function, $f(x)$ is equal to 1 by x , then I will integrate 1 by x to the power n from a to b and get a different answer out. That means with different functions being fed up, I will get different answers. So, this functional is a map from function to real number, not from real number to real number. Okay and then I want to know how does this functional change if I change the function slightly remember even in the action I had a functional S was a functional of ϕ and $\partial_\mu \phi$ you have to tell me what function and what derivative I should supply in $\int d^4x L(\phi, \partial_\mu \phi)$ then you have to supply a function. You do not have to supply a point. You have to supply a function and its derivative. So therefore, then you will get a real number. So S is a functional of f and $\partial_\mu \phi$. And you change the functions, you will change the action. So one example as we were discussing is this. So suppose I want to change the function slightly. As we discussed when we were looking at the diagrammatic representation, and not changing the functions at positions. For example, I can be at location x_2 here. So x_2 here in this map is a visualization that intermediate configuration of the field could be this value, this much value here, this point. This height determines what is the field value at location x at time t . Now, if at the same location and the same time, a different realization can be obtained by this green sheet. So at the same location x_2 at an intermediate time t , one configuration could be that the function has taken this much small value, the green thing. Or it could take the red thing. You can see different, different sheets if you visualize between the solid curves. Those are the different field configuration which exist between the boundaries. And therefore at any point fixed at t and x , you could have different, different realization depending/the curves and the bends in the sheets. So therefore, what we are changing is not the location, but the value of the function at that location. That means we are adding something to it. So let us think of an infinitesimal change in the function. By infinitesimal, I mean that most of the time the function remains the same. Only at a particular location, let us say x_0 , the functions become slightly different. And that is realized by a delta function. So look at this curve. The blue curve over here is the function $f(x)$. I want to do infinitesimal variation about it. So what I would do, I would add something very tiny around x_0 . So, there will be a kink developed at x_0 due to this delta function and the height of the kink will be controlled by ϵ . For visualization purpose I have scaled the height very high, but in principle it could be as tiny as possible. So, that would be infinitesimal change. Now, under this infinitesimal change I want to know how much the functional has changed. So as per the job, I would supply the new function, which is the previous function + a small variation to the power n . And then I will subtract out what was its earlier version. That means without any fluctuations or without any variation, what was the functional value. And the difference of that will be the difference in the functional, δG . Now, again we can do the Taylor expansion or let us say binomial expansion to the power n of this quantity $(x + y)^n$ for small y , y is $\epsilon(\delta x - x_0)$ and we are repeatedly saying that it is infinitesimal chain that means very small. So therefore I would just retain the leading order term in ϵ . I would forget about all ϵ higher powers, ϵ^2 , ϵ^3 because at the end of the day I am going to take ϵ tending to 0 . So let us find out the first order change. So the first order change would be obtainable. This $(x + y)^n$ is $x^n + ny$ and x^{n-1} . So, this is the first item and this is the second item. So, n th the second item and first item to the power $n - 1$ that would be the leading order term. There will be other higher order terms as well which I am not writing because at the end of the day I am interested in taking ϵ tending to 0 .

That is the new functional value and then I have to subtract out what was the previous functional value that was $f(x)^n$ integrated from let us say a to b here also a to b . So, therefore this $f(x)^n$ and that $f(x)^n$ will cancel each other and I will be just left with the deviation term which is this $nf(x)$ to the power $n - 1$ and then $\epsilon \delta x - x_0$. And this $\epsilon \delta x - x_0$ remember is the infinitesimal variation δf which we had brought about. So, the answer I am getting after doing this integration since I have done a delta function change at a

particular location x_0 , I will get answer ϵn and f^{n-1} at the location x_0 at the location where the change was brought about. So, therefore, this is the change in the functional and suppose I want to know what is the change per unit deviation in the function that is to say the quantity which is appearing over here. is nothing but the change in the functional with respect to the change in the function. So we are going to define this quantity as the function + infinitesimal change – the function itself functional itself rather and divided by the change in the limit δf tending to 0. This is very reminiscent of the derivative of a function. The function x + small amount δx – $f(x)/\delta x$ and then you are supposed to take the limit δx tending to 0. That gives you the df/dx . Similarly, the process over here gives you functional derivative of this $f(x)^n$ taking the derivative with respect to function f gives you this result. So, this quantity which is appearing over here apart from ϵ . So, what you have to do? You have to take the ϵ out. So, that is the definition that you do with infinitesimal change per unit infinitesimal change. This is the definition of the functional derivative. And you can see for practical purpose for polynomial powers. It behaves as if you have just done the partial derivative of this with respect to f . So, for polynomial functionals, you can just take the functional derivative as similar to partial derivative for all practical purpose. So, therefore, this equation of motion which was there, I can convert it into partial derivatives – $\partial_\mu \delta L / \partial_\mu \phi$. This is just an operational step, otherwise the function derivative is the most fundamental quantity which we should use. But for just for computational purpose, we can write it like that. So, therefore, this should be the equation of motion for the field given a Lagrangian. And we want a Lagrangian which gives rise to the equation of motion which is nothing but the Klein-Gordon equation because we want to know quantum theory of a Klein-Gordon equation. So, we want to know the phase space for a Klein-Gordon field. In order to go to the phase space, I want a momenta. A momenta would be obtained from the Lagrangian itself. So, therefore, I want to obtain a Lagrangian which under extremization of action gives me the Klein-Gordon equation. And what kind of action does that? That we will see in the next class.