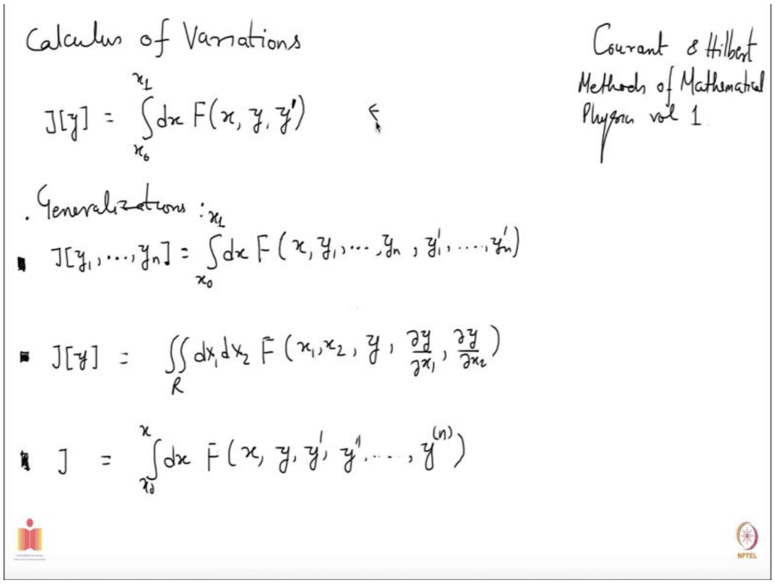
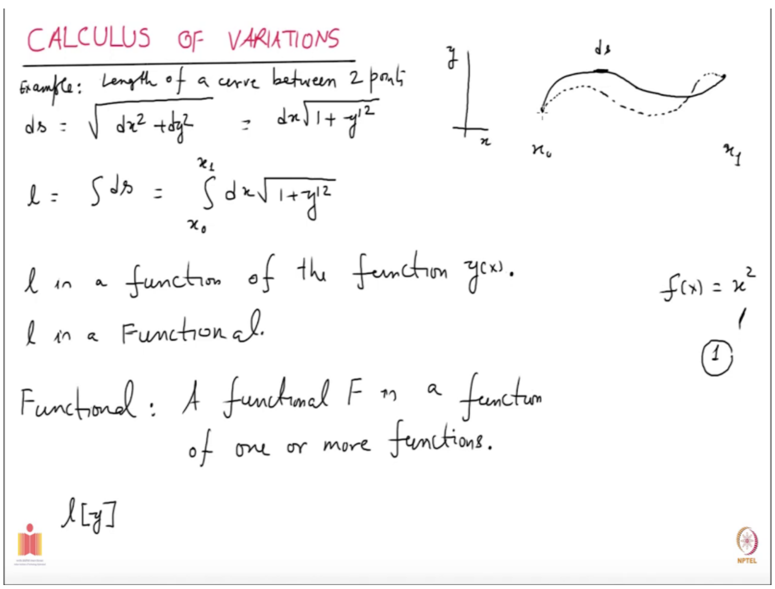
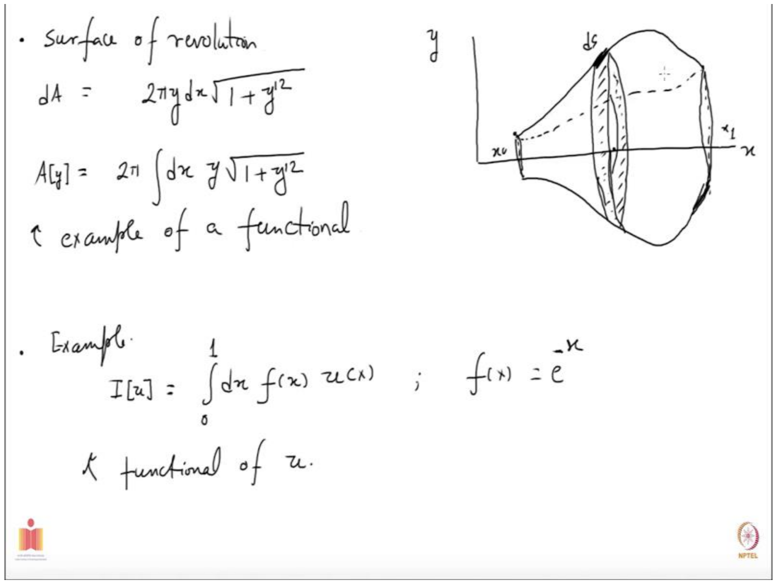
Introduction to Classical Mechanics
Assistant Professor Doctor Anurag Tripathi
Indian Institute of Technology, Hyderabad
Lecture 55
Calculus of Variations:
Condition for Extremum

(Refer Slide Time: 00:17)







Let us, continue with our calculus of variations, very useful discussion for a very useful discussion of whatever we are going to cover today can be found in this classic book by Courant and Hilbert methods of mathematical physics volume 1, let me write it down here, Courant and Hilbert methods of mathematical physics volume 1 methods physics volume 1, it is a very nice classic book and I would encourage you to have a look at it. So, I gave you a few examples where we need calculus of variations.

For example, I was asking about what is the I mean if I want to know that curve which minimizes the distance between these two points, the length, the length is minimum between these two points, what is that curves? That is one question you can ask and also you can ask about the, what should be the shape of this curve so that you get the minimum area of revolution around the x axis? So, you again need to minimize an integral of this kind.

So, in general we are going to be looking for minimizing or extremizing integrals which are of the following form. So, we will have some mint functional J does not look like J, J of y where y is a function and the two examples I showed you I have integrals of this form. So, they are all integrals over one variable x and then you have a function F which depends on the curves y and y prime, the function is y and y prime.

Now, you may immediately realize that this question of looking at the minimum or minimizing these integrals can be asking more general situations, for example you could let me write it generalizations, so you could have a situation where instead of having just one function y, this y prime is the derivative of y with respect to x, so our generation would involve for example instead of having just one function y you have several functions, y1, y2 and so forth to yn, so let us say n of them.

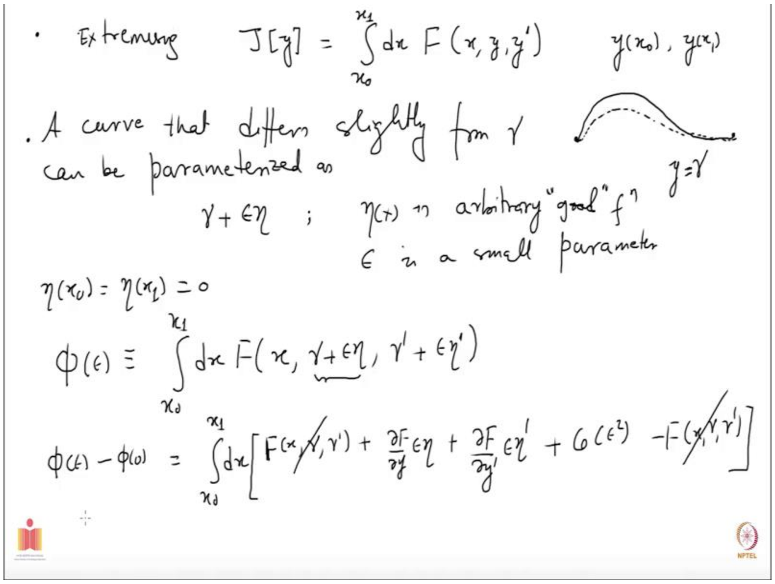
So, J is a functional of all these functions and you have again the endpoints I have fixed here and then you have dx F and then you have x y1, so forth to yn and then you will have all the derivatives of y1 and all the remaining yn prime. So, that is one generalization possible where instead of just one function y you have several of them, you can still think of more cases, more generalizations where you may have not only one variable x here, but a multiple of them.

So, let us say you have two x1 and x2 and let me denote that generalization by putting only one variable one function y here, so you could have like x naught to x1, sorry not x naught to x1, now you will be integrating over some region R in the space span by x1 and x2 and then you have your function F as I said there are two now x1 and x2 and I am still keeping one variable which you can have you can of course have more than one.

And then the derivatives of these with respect to the two variables dx1 dx2, that is one generalization possible and you can have still more, so instead of having only the first order derivatives and the function here you may have higher order derivatives also present. So, let us say that I have let us say from x naught to x, I am looking at only one variable right now, so you have some function F, which not only depends on the function y but also its first derivative, second derivative and so forth, let us say up to nth derivative.

So, you may ask like I want to minimize such an integral, so there are many generalizations that are possible, let us begin with our first one the simplest one and try to see what y the curve y will give the extremum for J. There is the question you want to ask. Let us, see how to proceed. So, let me write it down again.

(Refer Slide Time: 07:12)



We are looking at extremizing our J of J is the functional and first thing is my limits of integration are fixed dx F x, y, y prime, also the values of the function y at the end points are specified. So, I am going to search for all those functions y of x, I mean search among all those functions whose endpoints the values at the end points are fixed.

So, I am saying when I am searching for the function which is going to extremize this integral, I should keep the values of y x at x naught and y at x1 to be fixed, these values cannot change. So, let us say here is x naught and there is x1 and I am looking at some y of x. so, the value here and the value there is not allowed to change. So, that is the thing.

Now, let us see, let us our goal is to find out the y which will extremize this, which is equivalent to I mean in the parallel for functions if it was a function not a functional is here, you know the answer you just have to look at the first derivative to be vanishing. So, all those points where the first derivative vanishes is going to provide you the answer and we want to find the equivalent thing here for the functionals.

So, let us say this is the solution. So, let us say this curve which I have drawn here, let me call it gamma, so let us say y is equal to gamma of x, y of x is equal to gamma of x. So, this is the solution which is going to extremize this functional. And let us imagine a small deviation or variation of this curve. So, of course as I said the value of the function y should be same at these two endpoints.

So, it has to be of something of this sort, so I am imagining a small variation of away from gamma, let me write this, so curve or a function, let me write curve you understand that by a curve I mean really y of x, a curve that differs slightly from gamma, gamma can be written as can we parameterize as gamma plus epsilon eta, this is a trivial statement all I am saying is this curve which is given by some another function y of x will be to, I mean it will be almost gamma, but they will be small deviations which will be controlled by a small parameter epsilon.

So, epsilon is the small parameter which is going to govern or control how much this curve dotted curve deviates away from the curve gamma. And eta is any function with all the good properties that we may need. So, eta is arbitrary function, so eta of x is arbitrary, good function, by good I mean all the properties that one may require are available to us and eta is a small parameter.

And of course because this is some different y, but I have already said that the y should not change or is fixed at the end point, so this also should I mean this is just y, so it should not change its value at the end points, which means the epsilon, sorry the eta at the end points should be 0, that is required because I have already assumed gamma to be the correct solution which means the gamma itself has the correct values of correct values that it should take at these two endpoints.

So, any deviation should not change the values of the endpoint. So that is why you have this condition that is good, that is good. Now, if I take this y which is gamma plus epsilon eta and substitute is in this functional, then this functional will not remain a functional but will become a function because now I am not anymore changing a function because eta is going to be chosen to be something, gamma is fixed anyway, the only thing that you are changing is epsilon.

And epsilon is the number, so let me define a function phi which is a function of epsilon to be this, x naught to x1, dx and here you have F of x and instead of y I am going to put up gamma plus epsilon eta and y prime will become gamma prime plus epsilon eta prime. So, there is the definition of the function phi. Note that this is not a functional for not a functional it is a function because now you are whatever you change in here is just epsilon. That is good.

Now, let us see what happens as we change epsilon. So, phi 0 is, I mean when you are putting phi 0, epsilon is 0, then you are at you are using the gamma here, that is what it becomes, it becomes x gamma and gamma prime. And when you are on this curve, than you have this phi of epsilon. Now, let us look at phi epsilon minus phi naught.

So, we are asking how much the function phi changes as you go from this curve to a neighbouring curves. And the difference is clearly x naught to x1 dx, so I am going to write this for expand it around epsilon equal to 0, so let me write it down as so the first term in the expansion is F of x gamma gamma prime, then you have del F over delta y epsilon, so here you know this is the place where y sits.

So, I am looking at the rate of change of F with respect to what you have in here and then what is the diffusion deviation, so it is epsilon gamma plus del F over del y prime, that is the place what you have here is y prime and the deviation is now epsilon eta prime, that is the change in the function, so this gives your change in F up to order epsilon and then of course there will be order epsilon square higher order terms. That is good.

Now, I should also subtract phi of 0, so which I can include in here minus F x gamma gamma prime. And these two of course cancel, now this term is there, then these two have epsilon in common which I will pull out and take to the left, so I will have phi epsilon minus phi 0 over epsilon and I want to take the limit epsilon going to 0.

(Refer Slide Time: 16:55)

$$\phi'(\epsilon) = \lim_{\epsilon \to 0} \frac{\phi(\epsilon) - \phi(0)}{\epsilon} = \int_{x_0}^{x_1} dx \left(\frac{\partial F}{\partial y}\eta + \frac{\partial F}{\partial y'}\eta' \right) = 0$$

Integration by parts:

$$\frac{\partial F}{\partial y'}\frac{d\eta}{dx}$$

$$\phi'(\epsilon) = \lim_{\epsilon \to 0} \frac{\phi(\epsilon) - \phi(0)}{\epsilon} = \int_{x_0}^{x_1} dx \left(\frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right) = 0$$

Integration by parts:

$$\frac{d}{dx}(f \cdot g) = \frac{df}{dy}$$

$$\phi'(\epsilon) = \lim_{\epsilon \to 0} \frac{\phi(\epsilon) - \phi(0)}{\epsilon} = \int_{x_0}^{x_1} dx \left(\frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right) = 0$$

Integration by parts:

$$\boxed{\phi'(0) = 0}$$

$$\frac{d}{dx}(f \cdot g) = \frac{df}{dx} g + f \cdot \frac{dg}{dx} \Rightarrow \int_a^b dx \, \frac{df}{dx} g = \int_a^b dx \left(-f \frac{dg}{dx}\right) + \underbrace{\int_a^b \frac{d}{dx}(f \cdot g)}$$

$(f \cdot g)(b) - f \cdot g(a)$

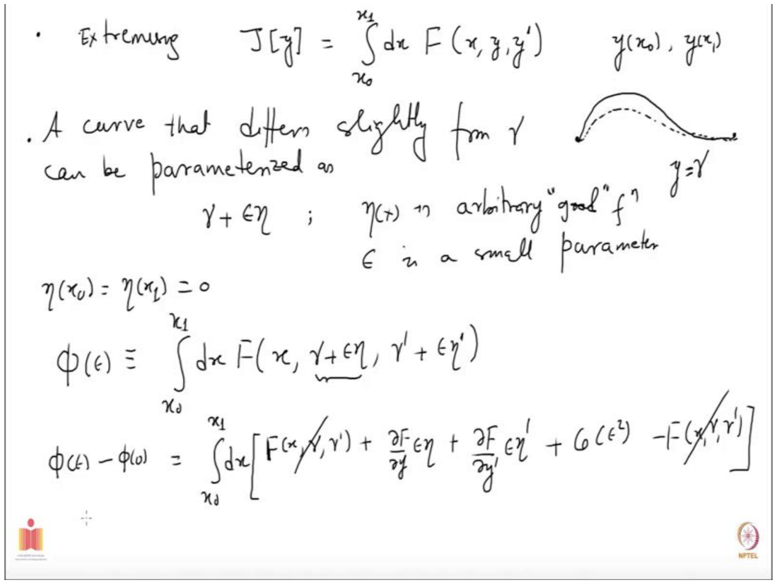
→ boundary terms.

$$\phi'(\epsilon) = \int_{x_0}^{x_1} dx \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] \eta(x) + \left(\frac{\partial F}{\partial y'} \cdot \eta(x) \right) \Bigg|_{x_0}^{x_1}$$

→ boundary term vanish

For γ to be extremum,

$$\int_{x_0}^{x_1} dx \left[\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] \eta(x) = 0$$



And in that limit I have I get phi prime of epsilon, which is phi of epsilon minus phi of 0, one let me put it here epsilon and limit epsilon going to 0, which is phi prime of epsilon, what is that? It is so epsilon has already been divided, so you get only these two terms, let me write that down, you get integral x naught to x1 dx and in the you have del F over del y eta plus del F over del y prime eta prime, let us see whether that is correct, that is correct, very good.

Now, remember what we are looking for, we are looking for a way to say that gamma is the one which extremizes this integral, which is equivalent to saying that the function phi of epsilon should have an extremum at epsilon equal to 0 at because epsilon equal to 0 corresponds to the curve gamma. So, I have changed the problem of finding the extremum of a functional to finding the extremum of a function phi.

And that is easy that I know, I know that if I ask that this quantity be 0 that is phi prime of epsilon is 0, then that is the location where the function could be will have an extremum or maybe a saddle point, so at least a stationary point. So, this is the condition that we have but we can make further progress by using a very useful trick which you will encounter several times in in physics and mathematics that is called integration by parts. So, let me tell you about that, I will use some colour so that it appears to be a separate part, a trick, integration, looks horrible, integration by parts.

Let me say in words what it is going to do for us. You see what is this? It is df over dy prime some function, then you have an eta prime, eta prime is d del eta over, sorry d eta over dx, now what I can do by using this trick integration by parts is let me write down this thing is delta F over delta y prime d eta our dx.

So, what I can do is what integration by parts does is it helps you to take away the derivatives from eta and put it on this one and that you may want to do because there is an eta here and then you will be able to combine the two terms. So, that is what integration by parts does and which you have learned in your school.

So, let me show you again in case you have not used it before, so all it does is let us say you have some function f times g, so they are two function f and g and you are looking at a derivative of this product, than this is as you remember df over dg, so df over dx times g plus f times dg over dx, which means that if I am looking at df over dx times g, so right now the derivative of derivative is on f and not on g and you get d sorry f dg over dx plus d over dx of f times g, is that correct? That is correct.

Now, if you are looking an integral of this, left hand side then it becomes let me write it, dx and put the minus sign here plus integral of this, let us put some limits a to b, so you see here the derivative is sitting on f, but in this one the derivative is sitting on g and it has been pulled away from f, but you get a minus sign, so you always remember that when you do this, you get a minus sign and this term is called boundary term, because this is equal to what? It is a total derivative.

So, it will give you f times g evaluated at b minus f times g evaluated at a and many times you will not have a boundary term because maybe one of the functions is fixed and it is not allowed to change, so this will be 0, you will see, but anyway there is a boundary term, so you should remember these are called boundary terms.

So, let us use that in our analysis, so what I can do is I can put the derivative from eta to this one, so there will be a d over dx sitting on this and eta will be freed up and as you saw there is a minus sign, so there will be minus sign and then they will be boundary terms. So, let us see what we get. So, your phi prime of x becomes x naught to x1 dx delta F over delta y and there is an eta which I will keep outside and then you get a minus sign and the derivative is sitting now on delta F over delta y prime.

And there is an eta again which now I will take common and just make eta affects to be more explicit and then this will generate a term of this kind right d over dx acting on del F over del y prime and eta, so you will get let me write down plus, so this d over dx will when you are integrating it will go away and it will lead you this, lead you to this, so it is del F over del y prime times eta of x and you have to evaluate at the boundary.

So, it will be evaluated at x1 and subtracted at x naught, so that is what you will get, let us see whether I miss something or not, fine, looks good, very good, nothing has been missed. Now, if you look at this one, I have already said that eta should remain I mean it should vanish at x naught and x1. So, that your whatever function this is it should not change at the end points.

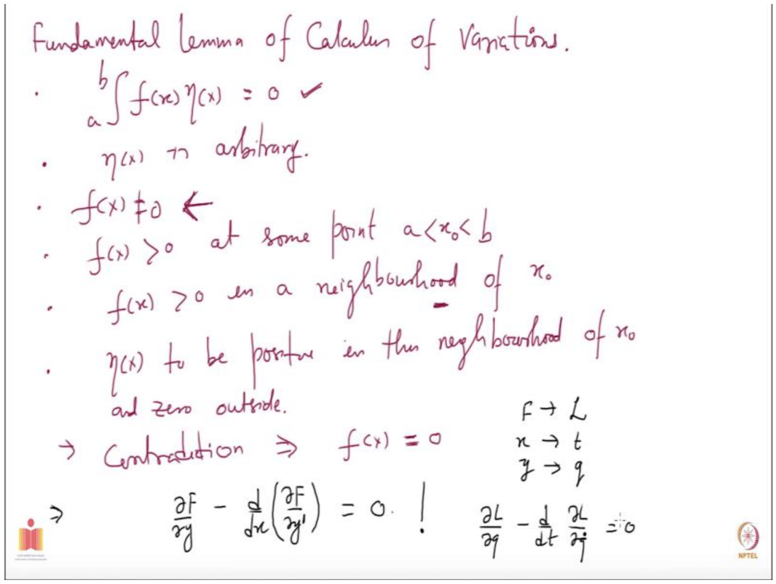
Which means when you calculate this boundary terms you will put an eta x1 here and of course this will be also evaluated at x1, but then this will vanish, similarly at this boundary also it will vanish, so the boundary term vanishes, boundary term vanish, boundary terms vanish, so at the both ends it vanishes.

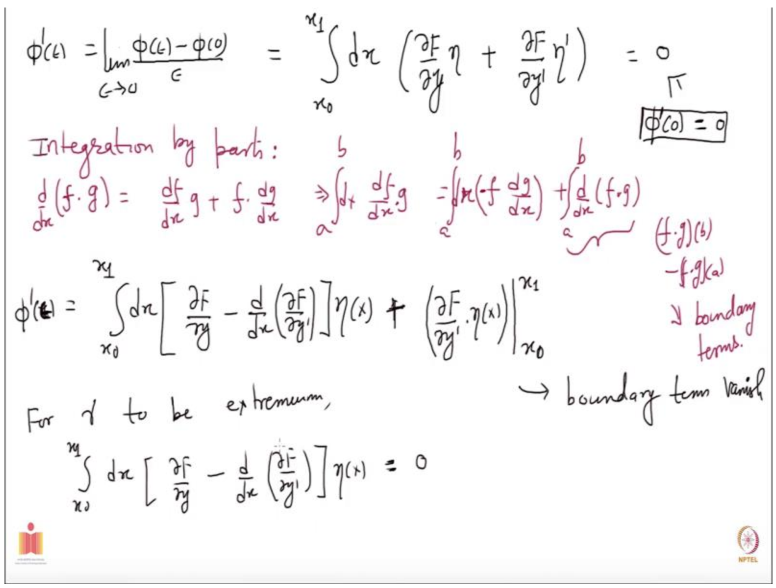
So, now what you are left with is this, which means for an extremum, for gamma to be extremum or a stationary point phi prime of not psi epsilon psi prime of epsilon should be 0, so psi prime should be 0 at epsilon equal to 0. So, here this is this is I am saying phi prime at 0 should be 0, I should have been more careful in writing this, so that is the condition we are looking at.

So, what does that mean? It means, so for this to be 0, so what I want is this to be 0, let me write again prime times eta this this is what I want to be 0, so if this is 0 then my functional is an extremum at gamma. That is good. Now, it is just few steps away or few are one argument away from the final result, now because your eta is an arbitrary function I have never specified what that function is, you could choose it freely, it should have all the good properties that we want for it, but rather than that it is completely arbitrary.

Now, I can conclude from here, that this whatever you have in square brackets is 0, let me tell you how I we can conclude this what happened, I need to add a sheet here, let us go. So, let us see how I can make that conclusion. And this conclusion is based on what is called, what is called fundamental Lemma of calculus of variations. Let me put it again in red so that it appears not to be a mean part.

(Refer Slide Time: 28:47)





So, fundamental Lemma of calculus of variations and what is that? So, imagine you are given the following integral f of x times eta of x and you are told that this integral is 0 and eta is arbitrary, so let us say it is told that this is true for arbitrary eta. And let us say that f of x is not identically 0. If that is the case, let us say at some point so I am saying f of x is not 0, let us assume this, if I assume this then it means that the function is going to take some positive or negative value at some point, so let us take a point where it takes a positive value.

So, let us say your f of x is greater than 0 at some point, the point let me call it x naught and which lies between a and b, it has some value non-zero value and I will assume that the function f of x is given to be continuous. Now, if it is continuous then if it is positive at some point then it will be positive in a neighbourhood of this point, because it cannot immediately become 0, so it has to be positive in some neighbourhood.

And in the same neighbourhood, so it is positive in some neighbourhood, so let us say f of x would then be positive in the neighbourhood of x naught, so that is clear. Now, because eta is arbitrary I can choose eta to be positive in this neighbourhood in the same neighbourhood I choose eta to be positive, it may be it may take a maximum value somewhere in the neighbourhood and it may be falling down and then it goes to 0 outside the neighbourhood.

So, when it is outside the neighbourhood the function is 0 and in the neighbourhood within the neighbourhood it takes a positive value, so let me say we take eta x to be positive in this neighbourhood, this neighbourhood of x naught and 0 outside, you can make this choice for eta but then it is evident that this integral f times eta integrated over this x from a to b will not vanish then.

Because it is 0 everywhere outside the neighbourhood, so it does not care about what f is doing, so that will be 0, unless these are singular but we have assume there are no similarities. And within the neighbourhood where there is a support for this function an eta, sorry support for eta this is positive that that is positive so it will give you a answer which will be positive.

Which is in contradiction to what we have set out to say that the integral should be 0, which means that this assumption is not correct, so we get a contradiction, the diction which implies that my f has to be 0 identically. So, if that is the case then I should first change the colour and get black and conclude immediately that whatever you have in the square brackets here should be 0.

So, this implies that del F over del y minus d over dx, del F over del y prime is equal to 0, meaning the function y should satisfy this equation for it to be an extremum to provide an extremum of the integral, so that gamma will be obtained from this equation, that is very nice, we have got the solution to this and note that if I replaced my F by Lagrangian x by time and y by a generalized coordinate q, then this equation becomes del l by del q minus d over dt delta l over delta q dot is equal to 0 and that is what is your Euler Lagrange equation which you have encountered in the very beginning of this course.