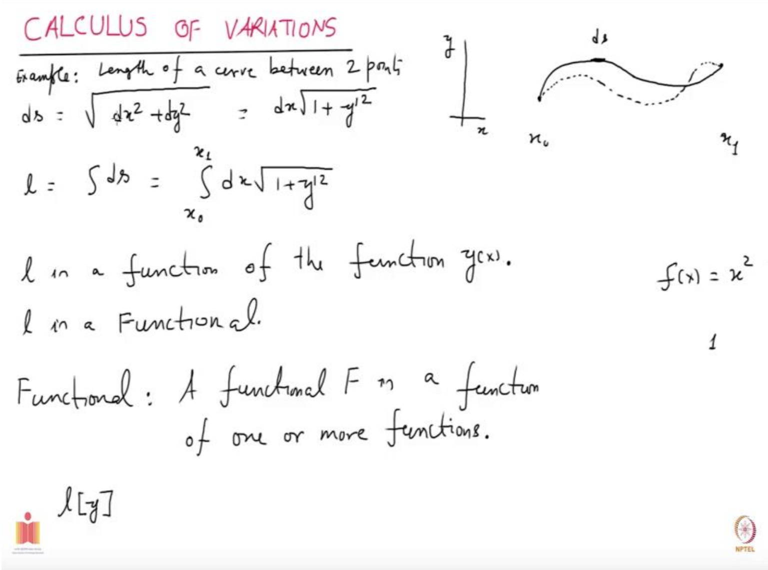
Introduction to Classical Mechanics
Professor. Anurag Tripathi
Indian Institute of Technology, Hyderabad
Lecture 53
Calculus of Variations: Functionals

(Refer Slide Time: 00:14)



For next couple of videos, I want to talk about the use of variational principles in mechanics. So, that is the topic which we will be taking up. And I will begin with the Calculus of Variations which some of you may have encounter in some of your courses and if you have not, then there is no, no issue because I am going to not assume anything that about this. So, let us begin with calculus of variations, I will give you a typical example of a problem that you may want to study in calculus of variations.

Imagine that you are given two points, which are located at x naught and x1 in the xy plane, so this starts at some point x naught, this starts at some point x1, the x and x coordinates are x naught and x1, there it is also specified here (())(01:32). Now, you want to, let us say you are given a curve like this, some curve is specified and we ask what is the length of this curve.

So, how am I going to find the length of the curve, I will add up small pieces of it. Let us say this has a length ds, where ds square is given by dx square plus dy square, so ds becomes square root of dx square plus dy square, so the length of the entire curve would be given by the following, let me first finish this so I take the dx out, 1 plus dy over dx which I call as y prime and you have a square of it.

So, that is the differential element of the length along this arc and the total length of this curve is the integral of ds which becomes dx 1 plus y prime square and of course you have to integrate from x naught to x1. Now, let us have a look at the quantity we have in front of us, the length l is a function of the function y of x, let me write it down first, l is a function of the function y of x.

Meaning, if you specify a different y of x, if you give a different curve then the length will change, of course that is an obvious statement. So what I am saying is, if instead of this one if I give you another curve, let us say I give you this curve, then you evaluate this quantity again on the right hand side the integral and you will have to substitute y prime for this curve, take a square of it, do the square root, do the integral and you get another value of l.

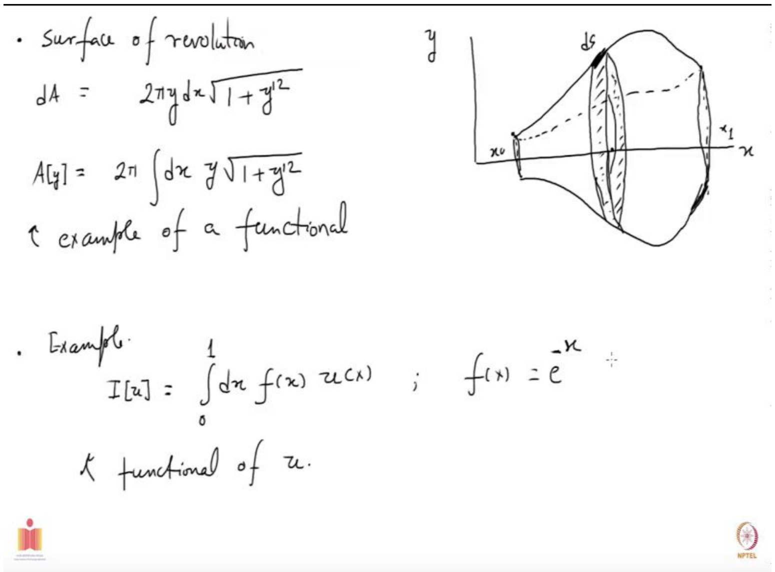
So, l is a function of the function of y of x, it is just like for ordinary functions when you give a number the function produces a number, so let us look at this. What is f of x, let us say f of x is x square, so you give a value of x, so you give a number let us say x is 1, then what you get is again a number back, you get x square is 1 you get f of x is 1, so you input a number you get a number.

But here l is, l has a different nature you input a function and you get a number, so this is not a function but something different, these quantities are called functionals. So, for me that quantity above l is a functional. And what is the functional, functional is a function of function. So, I will make this thing as a prototype of prototype definition, prototype of a functional so I will, I will define functional to be a function of a function.

So, let me just write it down functional, a function, a functional f is a function of one or more functions. So, that is the definition of functional and just pay attention how we have arrived at such a definition it was because of the prototype example that we took. And it is standard to notate or denote by a functional, for example in this case the l as the following.

So, you write l with a square bracket y, so l is the functional of in this case y where y denotes any curve between these two points. I will give you another example, so that was example number one, let me put it down here, this was length of curve, length of a curve between two points.

(Refer Slide Time: 07:20)



You can think of several such problems in which you will have a functional and one such I will give you now. So, it is surface of revolution, so let us see here, so the thing is this, this is your x axis, this is your y axis and between two points you are given some curve, some curve is given to you, let us say like this. Now, what we want to do is, so let us say before I tell you what we want to do, let us say this is located at x naught, this is located at some value x1.

Now, what I want to do is take this curve and rotate it about the x axis, so when I rotate it about the x axis I get, it will be difficult for me to draw on the slate let me try anyway, not bad, it is fairly good, so when you rotate it this is what you will get. So, there will be a flat end here, there will be a flat end there, which this bad hand drawings it is difficult, but anyway let us now you can ask what is the area of this revolution of this curve?

So, whatever shape what kind of thing you are looking at, what is the area of surface of this pot. And let us find out what it will be, again I take a small element ds here, then I make a differential area element here, it just look bad, so that is the small area element that I am looking at which is, if you look from here it is this ds is located a distance x away from this line.

So, if you look at the dA it will be 2 pi x, it is not x it is y, 2 pi y so this, this is y direction. So, the radius of this trip is y, so the area element is 2 pi y times ds and we already know what ds is, ds we found to be dx 1 plus y prime square, that is what we saw in the previous example. So, if I want to know the area of this surface of revolution it will be 2 pi times integral dx y square root of 1 plus y prime square and to emphasize that this is a functional I right here y.

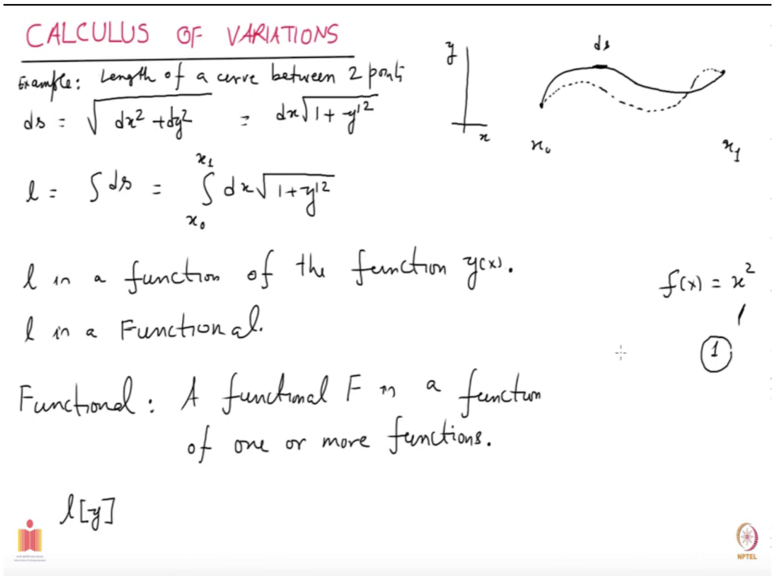
So, again as before you will get area will be some number, it will be some number, it will have some value, but that number you are going to get by specifying a particular curve, so if you specify this curve you get some value, if you specify a different curve, let us say you specify a different curve like this, then you will have a different area of revolution and you will get a different number of, number for A of y.

So, this is another example of a functional. I will give you one more example. Suppose I give you an integral over one variable dx f of u, f of x times some u of x, where I will specify f of x for you. So, let us say I fix the f of x to be say e to the minus x or e to the minus let us say e to the minus x.

Now, this integral now is a functional of u, if you choose u to be something and you do the integral, let us have some limits, let us say 0 to 1, these limits are fixed. So, if you take some function, I mean this function is specified for u, as you change the function u of x, you get different values and that is why your integral on the right hand side is a functional.

So, this is a functional of u, functional of the function u, with different choices of u you get different values. Now, it will happen that we will be interested in the maxima or minima of functionals, for example you could be interested in asking, what is that curve that will give you the minimum surface area of revolution.

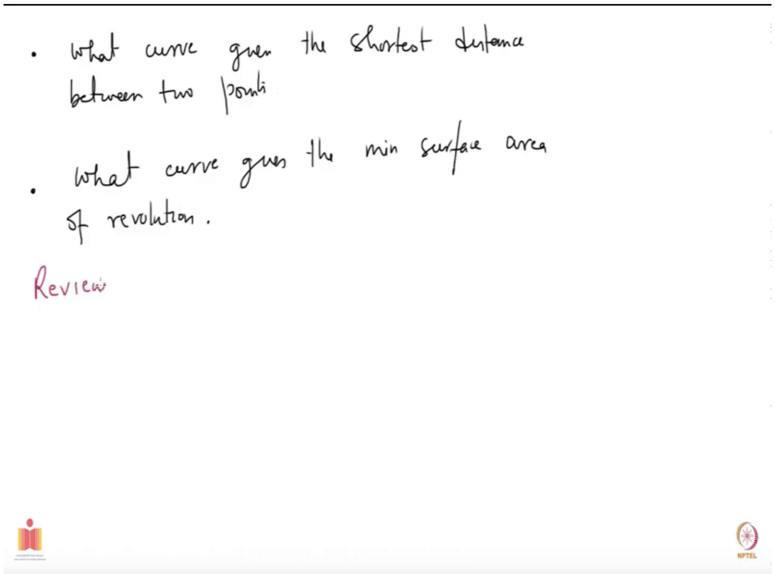
(Refer Slide Time: 14:02)



Or in this example, given these two points x naught and x1, what curve gives you the minimum length of the curve. So, for what function y the length gets minimized, that could be a question that you can ask. And these are similar to the questions that you ask when you are studying maxima and minima of ordinary functions, you ask at what point you will get a maximum of this function.

Now, remember the point, for example here in this example x was 1, that point is replaced by the function, so you just asked in the case of functional for what function you get minimum of the functional. That is about functions that what I wanted to say, maybe I should write it down, I have, I think I should write it down.

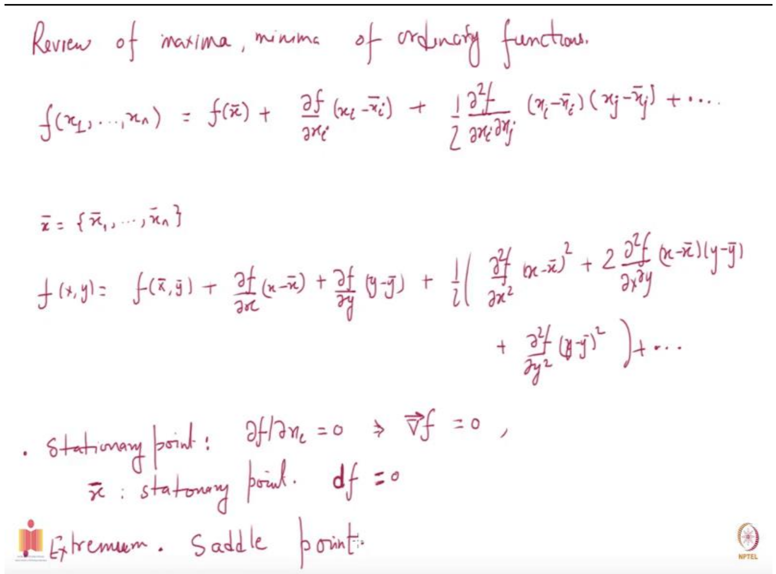
(Refer Slide Time: 15:08)



So, the questions we will be generally interested in asking in this case of functional is, for example, what curve gives the shortest distance between two points, between two points, that is in the xy plane for example or xyz. Or what curve gives the minimum surface area of revolution, what curve gives the minimum surface area of revolution?

Now, before I go into looking for these questions for functionals, let us review what we know about the maxima, minima and all these things for ordinary functions of n variables. So, that is what I am going to do, then it will be simpler or easier to move two functionals. So, let us look at now, maybe some colour will be nice, let us go to next page.

(Refer Slide Time: 16:45)



Review of maxima of ordinary functions. So, imagine we are given a function f which depends on n variables x1 to xn, then about a regular point I can always do a Taylor expansion, if there are no similarities there, then let us say I want to expand it about some point x bar which is basically this set x1 bar xn bar, so that is point x bar is really all these things evaluated at that point.

Now, you know that I can expand it around x bar as follows, so I get f x bar, then I have terms with the first derivative del f over del xi, then xi minus xi bar, where you have a summation over i implied by this repeated indices, then you get second order terms delta 2 f over delta xi delta xj xi minus xj, xi minus xi bar and then you have xj minus xj bar plus higher order terms.

And you can easily check that if you expand this for case of two variables you get f of x, y as let me put as x bar, y bar plus del f over del x x minus x bar plus del f over del y, y minus y bar and here you have three terms when i is x, I mean let us say xi is x, so it gives you delta 2 f over delta x square.

So, there should have been a factor of half here which I miss, so we get half delta 2 f over delta x square, this gives x minus x bar square, then you get the second term which has delta 2 f over delta y square y minus y bar squared. And the third term is where you have one of them is x and other one is y, but they are going to come twice because there is a summation over i and j, so you get a factor of 2 x minus, delta 2 f over delta x delta y x minus x bar and y minus y bar and plus all the higher order terms.

Now, if all the partial derivatives del f over del x del f over del y and in case of multiple variables all the partial derivatives they all are 0, then you call that point as a stationary point. So, let me right, if all the del f over del xi are 0 which is equivalent to saying that your gradient of f is 0, then that point is called stationary.

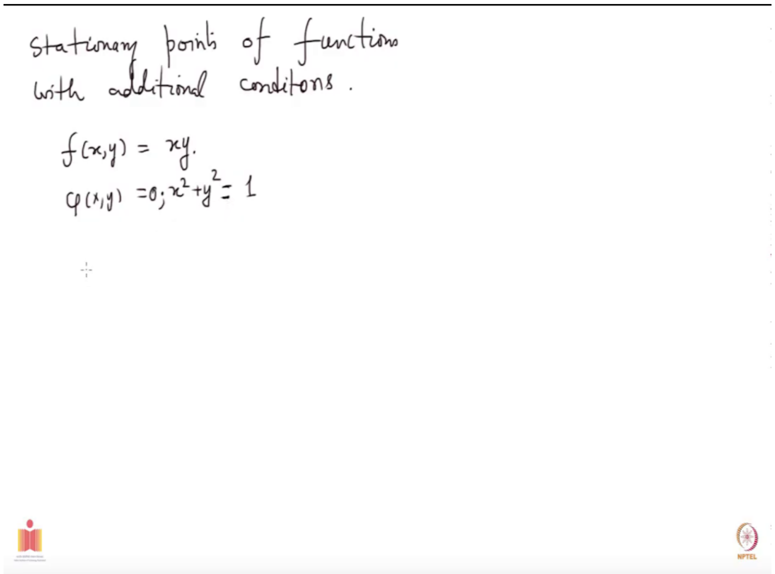
Let us, say this happens at some point x bar where x bar is x is x1, bar x2 bar and so forth. Then this point is called a stationary point. It is stationary because if you do a, if you move away from point x bar the value of the function f does not change to the first orders it, because these are all 0 your df is 0. So, there are no first order changes, so point has not the value of the function has not moved from its original value, so it is stationary there, so that is why it is a stationary point.

Now, depending on what these terms are doing or these terms are doing these coefficients are doing, you may or may not have a extremum there you may or may not have a maximum at that point. So, if the function, if at this point the function is maximum or minimum then you say the point is extremum point, otherwise it may happen that these derivatives are all 0 and then there is no maximum or minimum and it is a saddle point or point of inflection, that is good, that is fine.

Now, I would be interested in asking, for example, what if I am given a function and I am looking for its maxima and minima or the stationary points let us say, let us say I am not interested in knowing whether it is a maximum or a minimum or a saddle, I am happy with knowing whether it is a stationary point or not.

So, in this case, this is fine, you just take the derivatives and if it is, if the derivatives all the derivatives are 0 then you say that it is a stationary point. But what if all the variables x1, x2 and so forth xn they are not independent of each other, but there are conditions which constrain them, then how do I find out where the maximum or minimum or let us say, more generally stationary points are located. So, the question I am asking is the following.

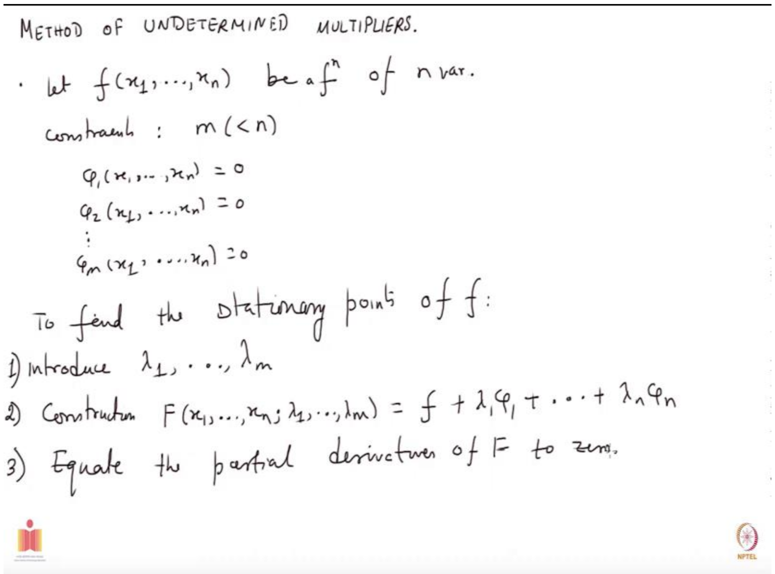
(Refer Slide Time: 23:55)

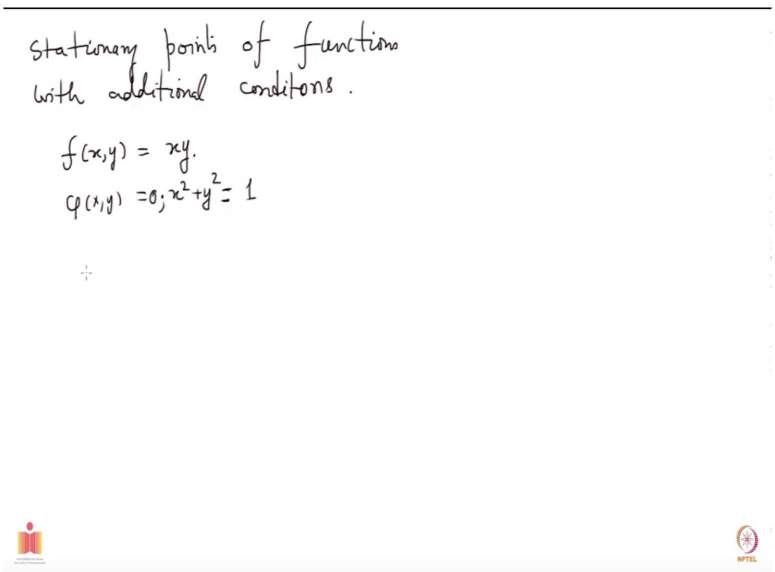


I am interested in finding the stationary points, points of functions with additional conditions, see that is what I would like to ask. For example I may, let us say you are given some function f of, let us say two variables is xy and you are given a constraint that x square plus y square should always be 1. And now I ask where are the stationary points located for the function f given phi, let me write it more neatly, let us say phi specified to be this, x square plus y square equal to 1.

So, this is these are the kind of things which we want to ask and I will give you the answer to this and actually it is, it is easy to why that is the answer and you can also look at the book by Courant and John, Introduction to Calculus and Analysis Volume 2, I think it is volume 2, there you can find a good discussion on this subject, so let me state this first, maybe I should, probably it will be nicer if I move to the next page.

(Refer Slide Time: 25:57)





So, method of Undetermined Multipliers. I will give you the answer and then we will see why that is an answer. So, question is the same, stationary point of a function with some additional constraints. So, let us say I have a function f which depends on several arguments and let us say they are total n of them, which are x1 to xn. And the constraints are the following, you have in total let us say m constraints and m has to be less than m.

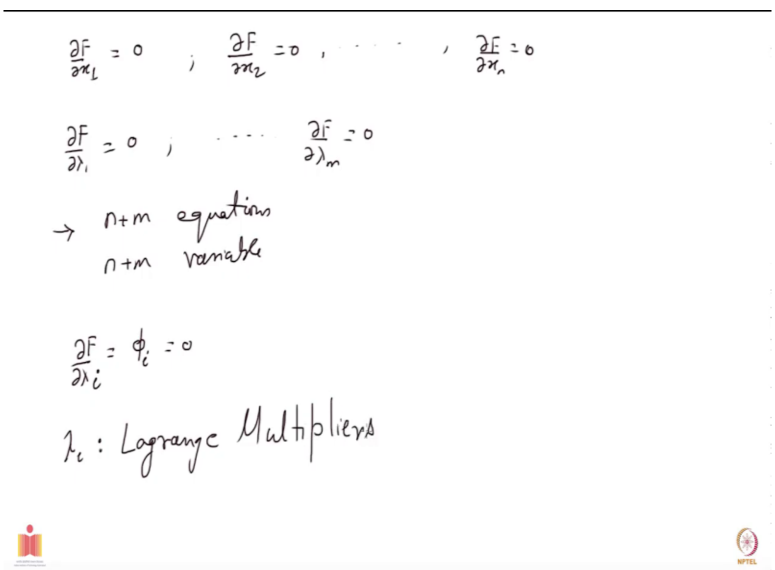
And let us say we denote them by f1 of x1 to xn being 0 than you have some other constraint phi 2 which constrains all these variables and you have phi m, now we are asked to find the stationary point of f. So, to find the stationary points, where is it, to find the stationary points of f we introduced the, introduce m multipliers, so one multiplier corresponding to each of these.

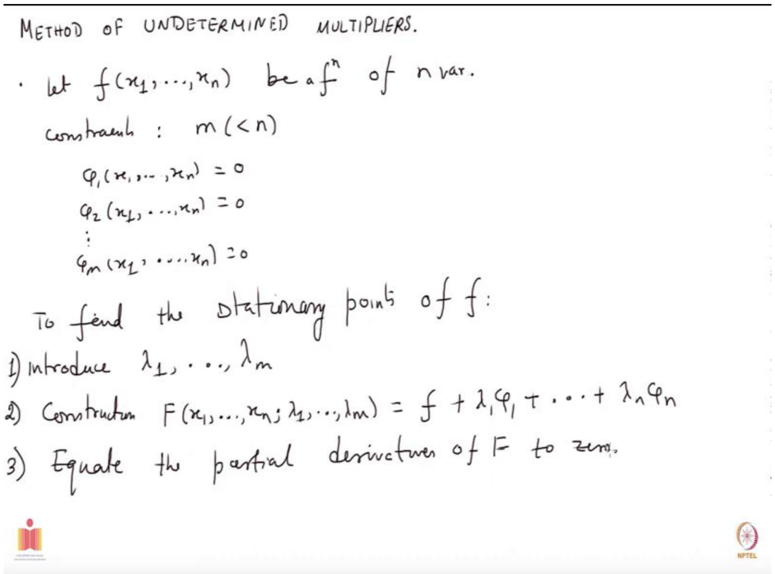
So, let me denote by lambda 1, lambda 2 and so forth up to lambda m, a total of m multipliers this is step one, so you introduce these multipliers, then you construct, a function a new function actually, F which is a function of all your variables which you had originally and then it also is a function of all these multipliers and the definition of the function is the following.

f is your original function plus lambda 1 phi 1 plus lambda 2 phi 2 so forth lambda n phi n, so that is the new function that we have constructed. Now what we do the, what we do is the following remember if all the phi's were 0 meaning there were no constraints, so let us say all these were all, I mean there were no constraints at all, so let us say these were all not there, then your function is just f and to find the stationary points you have to take the partial derivatives of f with respect to all the variables and equate them to 0.

But now in this case what you have to do is again take the partial derivatives of capital f with respect to all the variables and equate them to 0. So, it is roughly, I mean it is exactly like the original thing but instead of acting the partial derivatives on small f you act them on capital F, so let me write it down. Equate the partial derivatives of F, of F to 0 and these derivatives have to be evaluated for all the variables in the problem, meaning let me go to the next page.

(Refer Slide Time: 30:40)





So, you have to calculate delta F over delta x1, then you have to calculate and equate it to 0, then you get delta F over delta x2 put it 0, delta F over delta xn put it to 0, calculate again delta F over delta lambda 1 put that equal to 0, delta F over delta lambda m and put that equal to 0. So, how many equations we have now, n of these equations I have, and m of these equations I have, so I have total of n plus m equations.

So, there are total of n plus m equations coming from here. And how many variables do I have, I have again n plus m variables, n for the axis and m for the lambda, so they are total n plus m variables. Now, n plus m equations, n plus m variables the system can be solved. So, I can find out all these things.

But now the question is, why this is the solution of our problem. So, let us go back, let us see what delta F over delta x1 gives, delta, delta F over delta lambda 1 gives, let us put generically lambda i, this gives you phi i and when I am equating to 0 this is the condition I had originally, so these are m equations and I have phi i equal to 0 coming from here. So, this anyway I needed to solve and then you have the partial derivatives with respect to x1, x2 and so forth.

And together with these conditions you get back your delta F over delta x1 equal to 0 or delta F small f over delta x2 equal to 0. So, that is why this, this system of equations provides you the stationary point of the small f together with the conditions, that is why this constitutes the solution. And these lambda i's are called Lagrange multipliers.