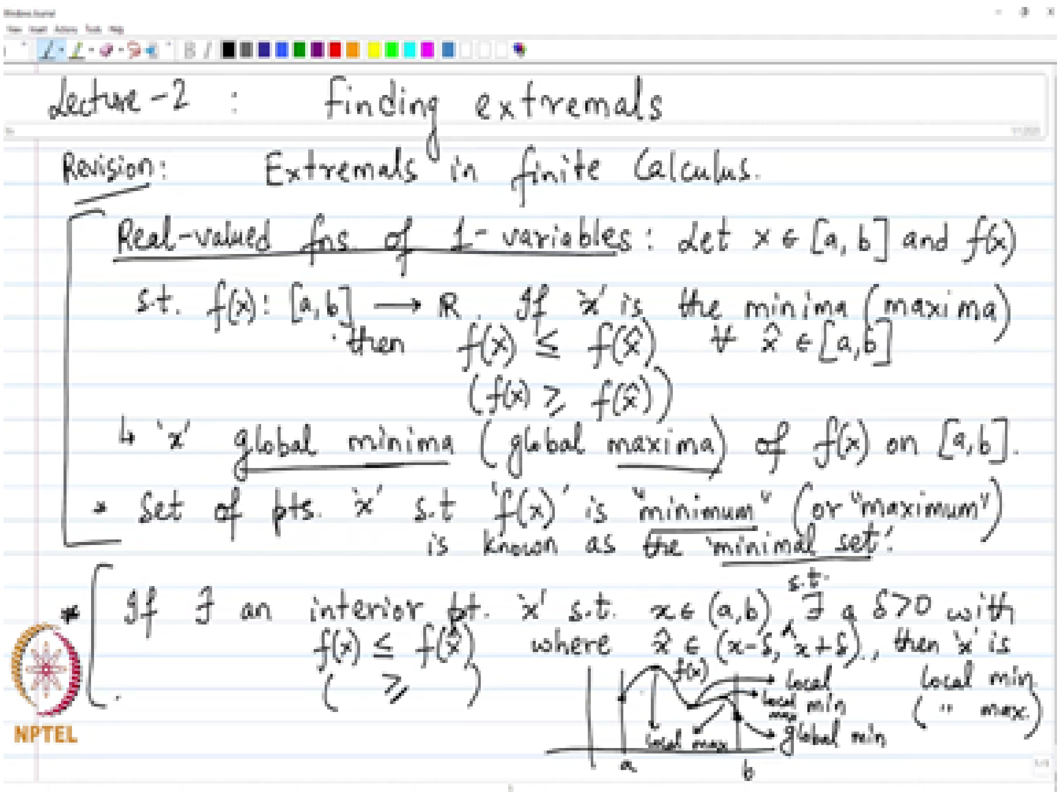
Variational Calculus and its Applications in Control Theory and Nano mechanics

Professor Sarthok Sircar
Department of Mathematics
Indraprastha Institute of Information Technology, Delhi

Lecture – 04

Introduction –Euler Lagrange Equations Part-4

(Refer Slide Time: 00:19)



So, in today's lecture, I am going to talk all about finding extremals. So now, before I dive into finding extremals of functionals or talk about real problems of finding extremals using calculus of variations, let us do a brief revision. So, I am going to start my topic today by describing Extremals in Finite Dimensional Calculus. So, students who have done primary courses in multivariate calculus, vector calculus, ordinary calculus, they must be familiar with most of the results, that I am going to talk over the next 25 to 30 minutes.

So, let me start with, you know Real-valued functions of 1-variable, and how to find extremals for such functions. Now, let me just define what do we mean by the extremal. Now, extremals in real-valued functions of 1-variable could either be maxima or minima.

So, let us say we talk about minima. So, let us say I have a point x and $x \in [a, b]$, which is part of the real axis. And let us also define a function f(x), s.t $f(x) : [a, b] \to R$.

If 'x' is the minima then $f(x) \leqslant f(\hat{x}) \;\; \forall \quad \hat{x} \in [a, b]$

If 'x' is the maxima then $f(x) \geqslant f(\hat{x}) \;\; \forall \quad \hat{x} \in [a, b]$

so, x is known as the global minima, respectively the global maxima. x is known as the global minima, respectively the global maxima of f(x) on [a,b]. Suppose we can always describe that, we can collect all such points, where f attains the minimal value respectively the maximal value and denoted by the minimal set.

So, what I just said is the set, the set of points x s.t function f(x) is minimum. When I say 'minimum', I have described what do I mean by minimum. So, it is the set of points x for which, f(x) is minimum or respectively maximum, maximum is known as the minimal set.

So, I have described the global minima or global maxima. So, in terms of the definition, if I have to write a proper math definition of a minima or maxima,then we need to look at points, which are interior to the interval.

So, if I have an interior point, if I have an interior point $x \in (a, b)$, then there $\exists$ a $\delta > 0$ with $f(x) \leqslant f(\hat{x})\ \ \forall\ \ \hat{x} \in (x-\delta, x+\delta)$ then x is local min.
and if $x \in (a, b)$, then there $\exists$ a $\delta > 0$ with $f(x) \geqslant f(\hat{x})\ \ \forall\ \ \hat{x} \in (x-\delta, x+\delta)$ then x is local max.

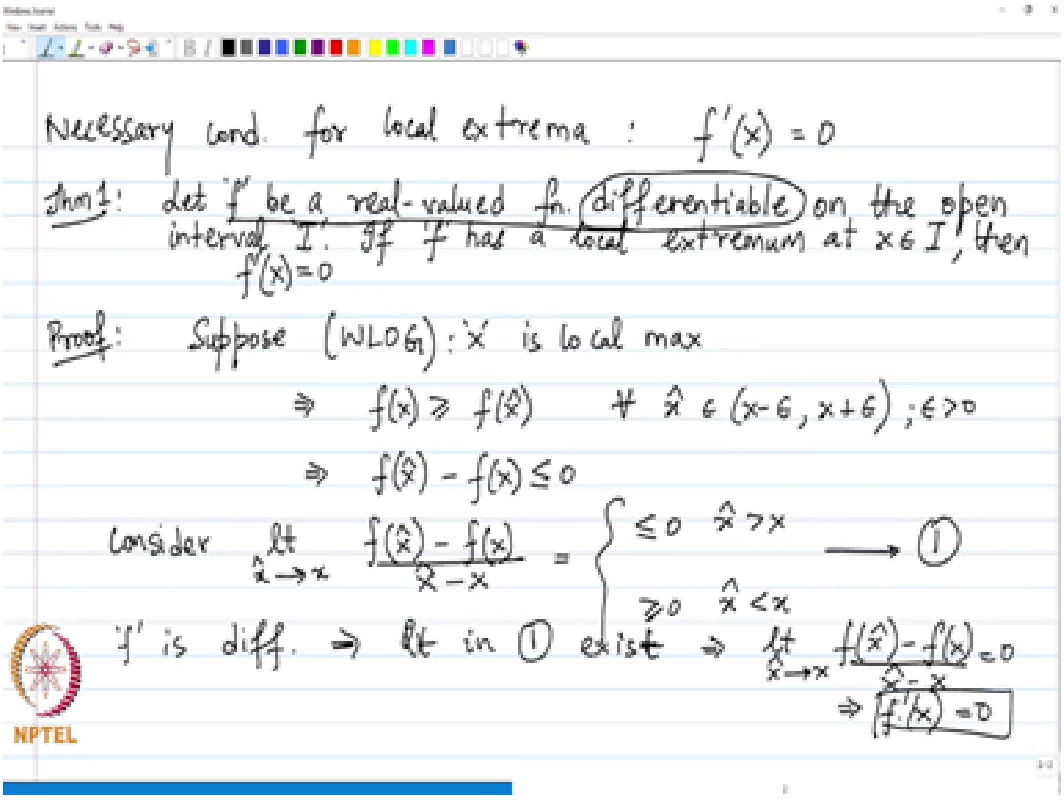
So, we see that the definition of the global minima follows the comparison of the function with all such values, possible over the entire interval. While the definition of the local minima follows the comparison of the values of the function inside a small interval around the point under consideration.

And similarly, we can describe the same definition using local maxima. So, if I were to draw a figure explaining all these concepts quickly, let us say I am describing a function over an interval (a, b) as shown above. And let us say, this is my function f(x), a function of 1-variable.

We see that, these are my points which are local maxima. And this is a point, which is local minima. While we see that, this point is above this, but this point has the lowest value. So, this will be a point on the boundary, which is also the global minima. Because the value of the function attained at this point, is the lowest among all such possible points. While finally, this particular point is local maxima.

So, in short finding local extremals, we can always look around the points. We can always look around the neighborhood of those points. While for finding the global extrema, we have to compare the value of the functional at all the points inside the interval, over which the function is defined.

(Refer Slide Time: 8:47)



So, then let us now start describing the tests for finding the extremals. So, in particular let us describe the necessary condition for local extrema. So, we are only going to describe the necessary condition for local extrema. And then when we compare the boundary values with these local extrema, we get the global extrema, when we include the boundary values.

So, the necessary condition for students who have done classes in multivariate calculus, they all know that the necessary condition has to be $f^{'}(x) = 0$ And that can also be shown using a very simple result in the form, which I state in the form of a theorem.

So, the theorem says, let f be a real-valued function differentiable function on the open interval 'I'. And if f has a local extremum at $x \in I$, then $f^{'}(x) = 0$ And the proof of this result also follows, in a very straightforward manner.let us assume without loss of generality, assume that x is the local max, we can prove similarly for local min.

$$\Rightarrow \quad f(x) \geqslant f(\hat{x}), \quad \forall \ \hat{x} \in (x-\varepsilon, x+\varepsilon), \varepsilon > 0$$

$$\Rightarrow \quad f(\hat{x}) - f(x) \leqslant 0$$

consider

$$\lim_{\hat{x} \to x} \frac{f(\hat{x}) - f(x)}{\hat{x} - x} = \begin{cases} \leq 0 & if \hat{x} > x \\ \geq 0 & if \hat{x} < x \end{cases} \qquad \mathbf{1}$$

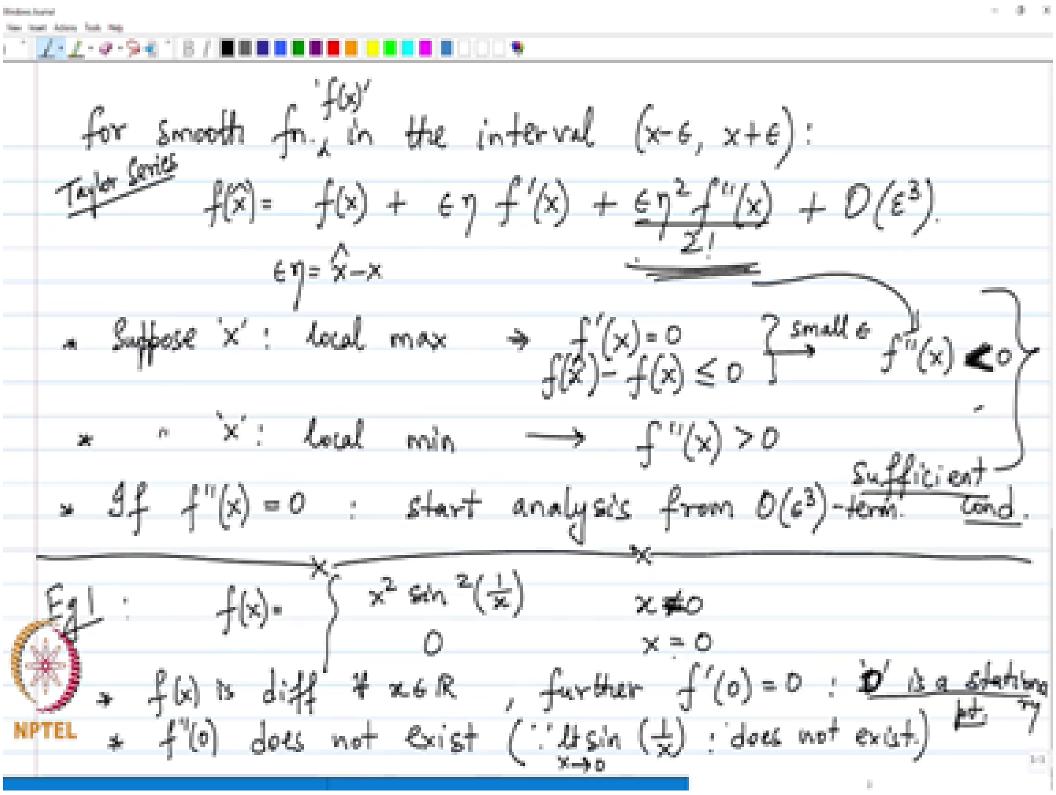
So, since f is differentiable, I limit exist in **1**, Or in summary, both the limit whether going from the left or coming from the right or must be equal to each other. And that can only happen if this limit is equal to 0 i.e

$$\Rightarrow \qquad \lim_{\hat{x} \to x} \frac{f(\hat{x}) - f(x)}{\hat{x} - x} = 0$$

$$\Rightarrow \qquad f^{'}(x) = 0$$

Because that is the only common value from the left or from the right. And we all know that this particular limit is the derivative of the function f with respect to x. And hence, the result follows. So, what is the moral of the story here?

(Refer Slide Time: 13:49)



So, the moral of the story is that, for a smooth function in the interval $(x-\epsilon, x+\epsilon)$, I have that $f(\hat{x})$, what I am trying to do is that, for any smooth function f(x), I am trying to rewrite the function around the neighborhood of the point x, using Taylor Series Expansion.

So, I am using the result by Taylor's expansion. So, I know that if , $\hat{x}$ is in the neighborhood of x, I can always represent $f(\hat{x})$ in terms of f(x), as follows.

$$f(\hat{x}) = f(x) + \varepsilon\eta f^{'}(x) + \frac{\varepsilon\eta^2 f^{''}(x)}{2!} + O(\varepsilon^2), \qquad \varepsilon\eta = \hat{x} - x$$

So, suppose x is the local max. Then I know that there are two pieces of information. I know that $f^{'}(x) = 0$ because x is an extremum. And also, $f(\hat{x}) - f(x) \leqslant 0$ because x is the local max. And that, from these two information, it follows that for sufficiently small epsilon, the leading order term in this Taylor's expansion is the second order term and from here, we conclude that $f^{''}(x) \leqslant 0$. In fact, for local max $f^{''}(x) < 0$.

Now, similar result for x being the local min. We can conclude that in this case, $f^{''}(x) \geqslant 0$. So, from here, we get the so-called sufficient condition for the existence of the extremals. So, the previous slides I showed the necessary condition. And in this slide, I am highlighting the sufficient condition.

And of-course, I could always have that the second derivative of f with respect to x could be 0. And then, Again, we have to start our analysis like the one done above, starting from the third order terms onward; from order ϵ cube terms onwards.

So, you know recapitulate our revision using some examples that I have compiled. So, the first example is, let us look at a function

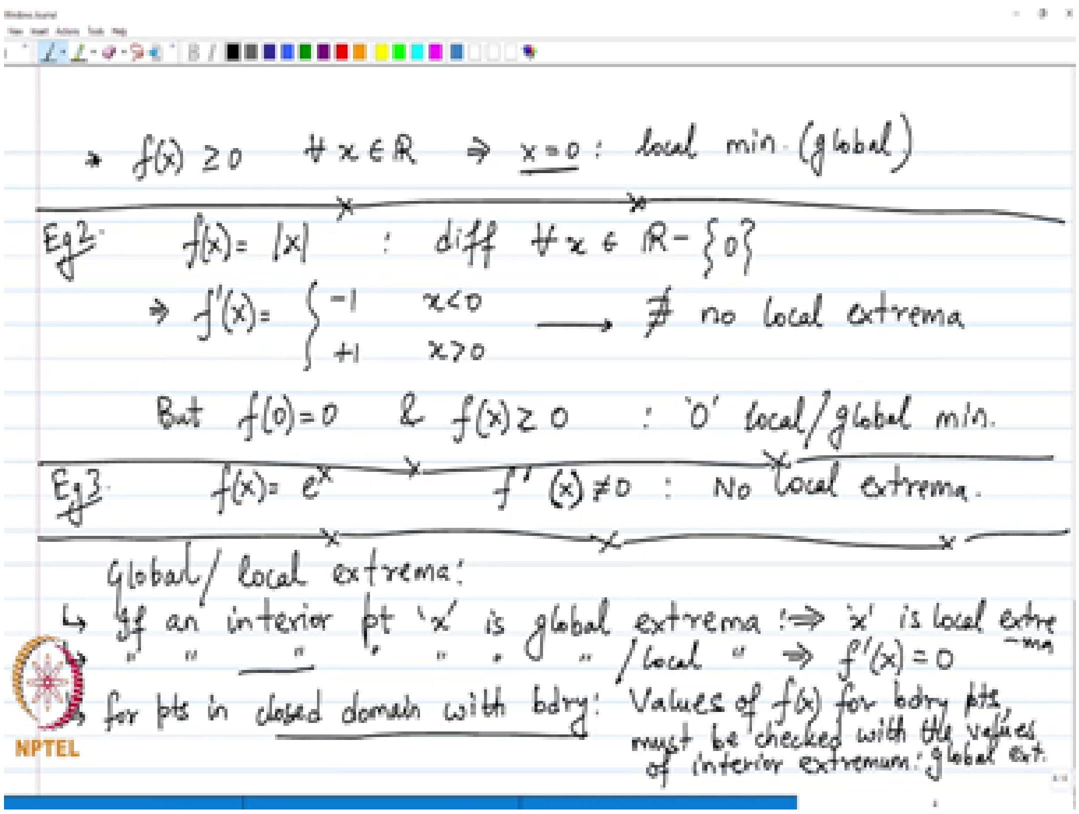
$$f(x) = \begin{cases} x^2 \sin^2 \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

So, we can show that f is differentiable for all x real-valued. And in fact, we can further, we can further show that f is also differentiable at 0, using our definition via the limit. And the derivative at 0 is 0 i.e $f^{'}(0) = 0$ So, it turns out that 0 is a stationary point.

So, then what I have is, let us say that $f^{''}(x)$, when we try to calculate the second derivative of the function at 0, we see that it does not exist. All the students who are going through this exercise should do this assignment, to see that the second derivative of the function at 0 does not exist. People can check that. That is because while calculating the second derivative, we have to evaluate this $\lim_{x\to 0} \sin \frac{1}{x}$ we all know that this does not exist. So, that creates the problem. Well, that does not prevent us to figure out what is the nature of this stationary point.

We know that 0 is a stationary point, but the second derivative test will not work, because we can see that the second derivative does not exist. However, we see that f is a square of a real-valued function. So, whatever be the value of x, we know that f(x) will be non-negative, which means that the lowest value of f will be 0. And that is attained at x = 0. So, 0 is my minima. So, what I said is the following.

(Refer Slide Time: 20:33)



We note that $f(x) \geqslant 0 \ \forall \ x \, \varepsilon \, R$ and this means that x = 0 is a local minima, because only at 0, the function attains the lowest possible value, which is 0. And in fact, it is also the global minimum, because that is the minimum value that the function can take.

So, here is, let us look at quickly look at another example. So, let us consider $f(x) = |x|$, we know that this function is differentiable ∀ x, except at x = 0. So, it is differentiable for all x, which are all real values of x, except at 0.

And further, we can see that

$$f^{'}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

which means that there is no local extrema because the first derivative is never 0.

However, we know that, f(0) = 0 and we also know that f is always a non-negative real-valued function. So, which means that, from these two observations we can conclude that 0 is a local as well as a global minima. So, this is just some basic problems, which we are revising at the moment.

Let us quickly again look at another example, a simple case could be, let us say $f(x) = e^x$ Now, this exponential function has derivatives for all orders. However, it never vanishes. None of the derivative vanishes, which means that there would not be any local extrema.

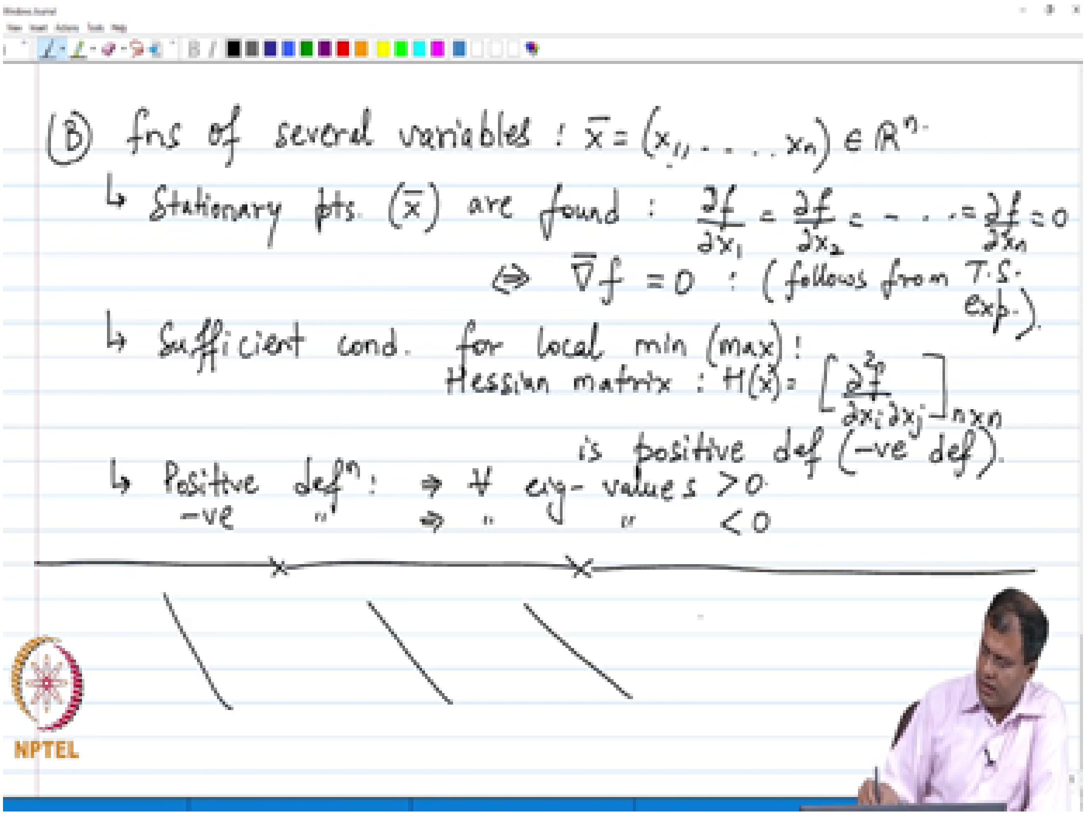
So, this; so, f^nth derivative is never 0. Well, all I need is the first derivative. So, f, the first $f^{'}(x) \neq 0$ and from here, we can conclude that there is no local extrema. So, our derivative test would not work on this function. So, to summarize what we have so far seen in our revision topics, what we have seen is, are the following basic points.

In order to find a global or local extrema, what we see that if I have an interior point x, which is a global extremum, then this interior point will also be a local extremum. if I have an interior point x, which is also global or even local extrema, then for an interior point it is always necessary that the first derivative will vanish.

So, whether it is local or whether it is global, for an interior extrema the first derivative of the function will always vanish. Finally, for points in closed domain, for points in closed domain with boundary points included. We see that if the values of f at the boundary points must be checked with the values at the interior points to evaluate the extrema. So, what I just said is the following.

For points in a closed domain, with boundary points included, the values of f(x) for boundary points must be checked with the values of the interior extremum to figure out the global extremum. So, this is the basic of our finite dimensional calculus. And we can, similarly extend this 1-variable calculus to several variable calculus as follows. So, let me term our first case to be case A.

(Refer Slide Time: 26:43)



And we can go to look at the second case as case B. That is the case with functions of several variables, $\bar{x} = (x_1, ...x_n) \in R^n$. So, like in the function, similar to the observations in functions of 1-variable, to figure out the extremum for functions of several variables, we must have that the first derivative of these functions must vanish at each, at the first derivatives taken with respect to each component.

So, what I just said is, to find the critical points or the stationary points, let us say $\bar{x}$, which is now vector in this case are found by setting up the following constraints.

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = ... = \frac{\partial f}{\partial x_n} = 0 \quad \Leftrightarrow \quad \bar{\bigtriangledown} f = 0$$

The first derivative of the function at with respect to each of the individual component variables, must vanish. So, this result is consistent with functions of 1-variable. Or in short, we say that the gradient of the function is 0 and this result again follows from this one and the one that I am going to just describe, follows from our Taylor Series expansion.

So, then the next result is about the sufficient condition. So, the sufficient condition what we saw was, for functions of 1-variable. The sufficient conditions led to finding the second derivative of the functional, function and checking the sign. If the sign is negative, then local max. Sign is positive, then local min.

So, similarly if the sufficient condition for local min or local max simultaneously, it must be that the Hessian matrix, the Hessian matrix which are the matrices involved using, let us say the second derivatives with respect to the variables, and so on. Let me let me denote this matrix in the short hand notation $\frac{\partial^2 f}{\partial x_i \partial y_j}$ which is in the short hand notation, a matrix of order n x n.

And if this Hessian matrix is positive definite, then I have found the condition for local minimum. Simultaneously for local maximum, it must be negative definite. So, what do I mean by positive definite? So, a matrix is positive definite, if all eigen values of the matrix are positive.

And similarly, if I have a negative definite matrix, it implies that all eigen values must be negative. So, all we need to check to figure out the sufficient condition, i.e to check the eigen values of the associated matrices, that are found by evaluating this Hessian at the critical points. So, I think at this point, the revision for the Finite Dimensional Calculus is over.