Indian Institute of Science

Variational Methods in Mechanics and Design

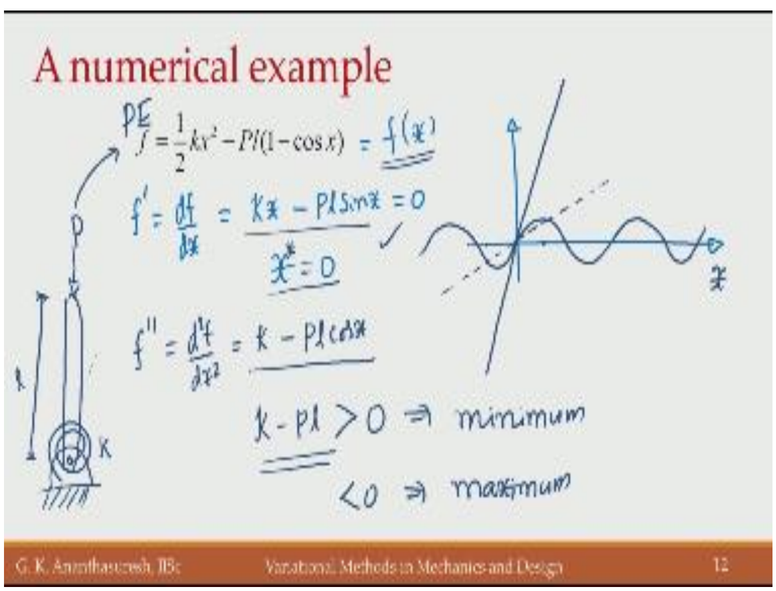
Prof. G. K. Ananthasuresh

Department of Mechanical Engineering
Indian Institute of Science Bangalore

NPTEL Online certification Course

Hello again so let us continue our discussion of this detour of finite variable optimization necessary sufficient conditions without constraint unconstrained minimization we discussed in essence sufficient conditions now let us take this one variable problem and do a numerical example that is shown here.

(Refer Slide Time: 00:38)



F = ½ Kx^2 - pl – cos x, this is our f of X here we have taken an example okay now if I do F' which is in our notation DF / DX okay so what do we get here 1/2 KX^2 ½ and this too in X^2 will cancel so I will have k x okay and then we have PL which are constants X is the variable here so we have minus and there is minus when I take derivative of cosine as I get another minus so minus minus minus eventual have my necks p l sine x that should be equal to 0 that is our condition for X to be a minimum right necessary condition.

If you solve this thing so KX minus PL sine X equal to 0 if you just observe this equation X = 0 is definitely a solution right because K X = 0 pls and X sine x is 0 and X equal to 0 that is a solution is the only solution clearly not because if I imagine that for some values of P and L let us say I draw a sine curve so we got the necessary condition that KX - PL sine X = 0 X equal to 0 is a solution let us call that X star okay is that the only solution but if you look at K X and P L sine X let us draw for some value of P&L we can draw the sine curve and in fact it will go both sides right.

And then now if I take a value of K that is very large I can I will have a straight line like this now these two intersecting only at the origin X = 0 but if K were to be different let us say then I have one more and third right where these two things are equal that is what we are saying red KX should be equal to P L times sin X taking a multiple solution so for a simple one variable problem there can be multiple solutions some of them will be minima other will be Maxima or in between for that we have to take the second derivative that is d^2 f / d x^2 square okay.

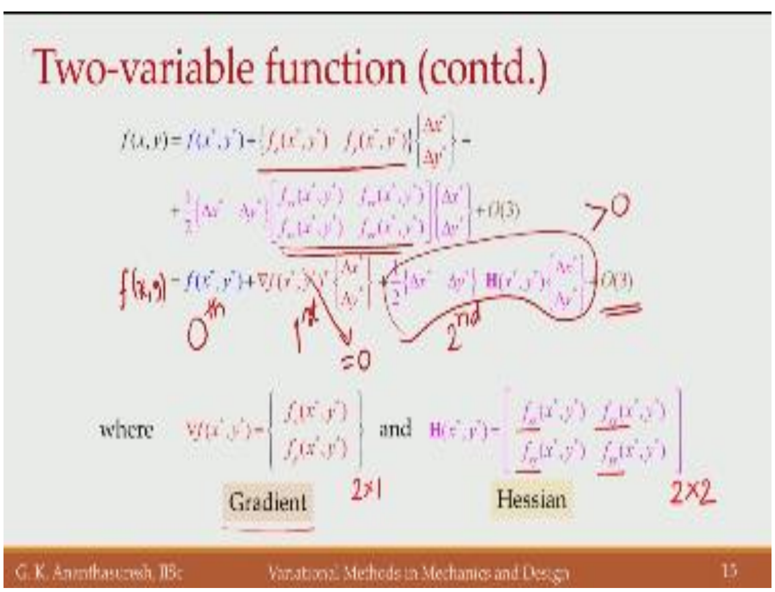
That in this problem if I take derivative of this one more time I get this k and then minus P L cosine X okay, so here at x* = 0 cosine X* will be cosine 0 is 1 and then PL so if I have K -PL should be greater than 0 for this point to be a minimum of this function okay so as I increase p at some point k is let us say chosen such that that is larger than p times L for some value of P as you increase p at some point this will cease to be positive in which case this point which is a minimum until then will cease to be a minimum and then when it becomes negative it is going to become a maximum account of a sufficient condition right.

We use necessary condition to solve for the problem to solve the problem to get all these values of x star okay here I did graphically but you can do it analytically or numerically and then every one of those you have to check against the sufficient condition we have written and then verify that it is greater than 0 then it implies that it is a minimum otherwise if it is less than zero it would imply it is a maximum okay, so you should work out this numerical example yourself to see as you vary of value very the value of P how X*which is zero goes from a minimum to maximum.

It is in fact it is a wonderful simple example that illustrates buckling of if I have a rod which is hinged over here and let us say there is a torsion spring here of constant K or κ then if this length is l this is a rigid rod here there is a compressive force p for that this is the potential energy as we considered minimum potential energy principle gives you static equilibrium we are actually minimizing this F which is actually the potential energy for this problem okay.

So you should work it out and understand how mineral maximum change as you increase p as increased p beyond a certain threshold value which is equal to K / L then it is actually a buckling then what is straight will suddenly shift that side or this side okay you should really try this problem out it is a one variable optimization problem okay.

(Refer Slide Time: 06:17)



Now let us move on to two variables x and y okay since two variables let me change the color of the ink okay we have two variable problem for that leads to the Taylor series expansion okay and that is how it looks right so we have the 0^{th} order term which is in blue color and then the first order term there are two of them there is one and there is two when I write FX here this FX what I mean is that it is ∂f / ∂ X F sub X this our notation similarly if I write F why I mean this to be ∂ f / ∂ Y partial derivative.

So here I can do the perturbation in λ X *λ Y*right we remember that we had a disk in two variables I can from this point that I have so we had a disc around a point if I move somewhere I will have a λ X* there is λ Y* which is what is indicated here this oh this is λ X *this distance and this distance will be λ Y* okay that is a 0^{th} order term and a partial derivative perspective X partially with respect to Y.

Now in the case of second order term we have three terms there is one there is two there are three right so this is λ X* square perturbation the x then y this is X this is y second order term and this is mixture λ X * and λ Y*okay so we have three terms in the second order term now again if you say the first order necessary condition that these two terms should be equal to 0because if λ X * and λ Y*X are very small we can neglect the second order term and higher order terms then the first order term because you do not know λ X * and λ Y*can be positive or negative depending on the sign of ∂f / ∂ X and ∂x / ∂ y.

We do not know whether the function will be larger or smaller require first order condition so we say that both f of X and F of Y should both be equal to 0 that is the necessary condition here okay all the notation FXFY FXX is second order second derivative with respect to X F X Y mixed derivative secondary way to ∂ square f by ∂ X ∂ Y then do YY is ∂ square f by ∂ y square okay which is how we have written now we can write in the matrix form like this so we have arranged this in the form of a row vector and this as a column vector.

So that when you multiply I get these two terms right so these terms I get and likewise if I right this second order term which has three terms in it in the form of a matrix with a row vector and a column vector and this is a matrix here at two by two matrix in the case of two variables right so let us look at that so we have written this in this short hand notation again this is zero order term and this is first order term and this is second order term and these are third and higher-order and when we have this we write it in the form of this row vector here right.

So which I have written as transpose so what we have written here this thing is called the gradient okay likewise we have indicated this matrix here in the latter edge which is called Hessian okay, we have a gradient and a hessian gradient determines the first order term here senator my second order term now when we want first order term to vanish for X* Y*r to be a local minimizing point then the gradient should be zero just like in the one variable case we said DF by DX should be equal to zero here we say graded.

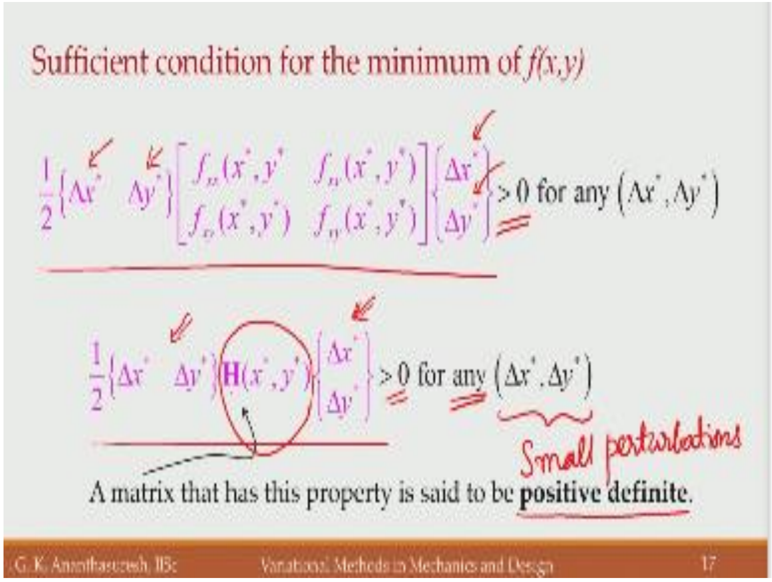
So that the necessary condition outer necessary condition goes away we want our f of X Y here that should be larger than f of X* Y* now that this has gone to zero because of necessary condition then the second order condition the term that we have that is this whole thing here that has to be greater than zero because only then X star Y star will be minimizing point or minimum value right every other word should be larger than the sufficient condition as we had seen and that is the meaning.

So of the first order term goes away second order term should be positive that makes X* Y*a minimize but understand the gradient which is the partial derivative of F with respect to x and y and the Hessian which is second order partial derivatives which we will see what we have written gradient and hessian which I have overwritten let me erase it so that you can actually see yeah, so now so you have that as a gradient and this is the hessian okay.

We get the pen back yeah so this is gradient which is a vector you see it is a arranged in this case of size to buy one okay and the hair cell is a matrix in this case it is 2/2 write it is a gradient which is actually the direction in which this function at that point changes the most that is a direction we would say the 2 x1 array column vector or it is a vector if you say x and y are two directions that is a gradient in the direction which the function changes its value or the fastest rate okay.

Let us see in a second order term has this thing called Hessian which is 2 / 2 matrix that includes all these second derivatives X square Y square and then XY.

(Refer Slide Time: 13:17)



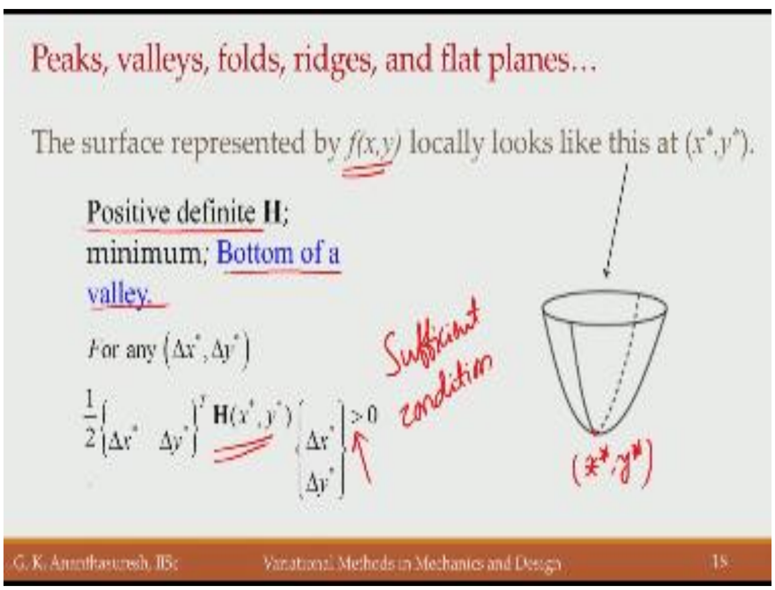
Necessary condition is that the gradient should be equal to0 for the argument that we have described now right well so this is a vector so it should be equal to a zero vector meaning that there are actually two scalar equations in it means that FX is 0 and then f sub y is equal to 0again do not forget the notation F sub X is 2 f / ∂ X and F sub y is ∂ f / ∂ Y so that is a necessary condition okay that is what we have written here right these are the necessary condition in two variables when you are solving two variable optimization problem you have two unknowns you do not know X* you do not know x*.

So you need two equations which is what we have we have two equations to solve for two variables to solve for x* and y* okay two unknowns two equations that is what these conditions are useful for not only you can tell whether a given point is minimum or not but also you can use these conditions necessary condition to solve for those minimizing point now when you go to sufficient condition then the second order term which is what we have shown here right when we expand it out it will have three terms that we get.

Now this one for any value of these perturbations λ X* and λ Y* which are over here and here is a row vector and a column vector when you expand it all out then you get the three terms that some of that which is the second order term should be greater than zero the sufficient condition which in shorthand notation looks like this for any is important any small perturbations these are small perturbations for any small perturbations this should be true and all this if it is for any values of these right then if it is have to be greater than zero the property lies within this matrix hessian okay.

That property is called positive definiteness okay definitely positive that is no matter what values you take for this F of λ X* and λ Y* okay this whole quantity that we are taking a row vector and then a matrix and column vector that will be greater than 0 if it is true then we say such a matrix is positive definite at that point right we are doing X* and Y* evaluate at that point will be called positive definite. So if I if I have.

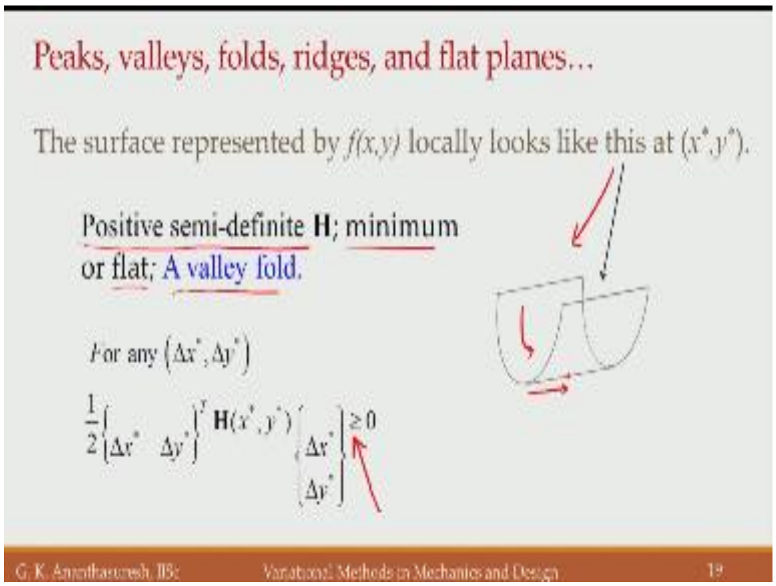
(Refer Slide Time: 16:11)



Let us say problem where function f is a function of two variables X and Y then if locally if it looks like a bow like this okay let us say this is my X* Y*coordinates at this point or whatever surface that f of X Y can be repeated as a surface in2d x and y and then third dimension f of XY locally it will look like a bowel which is the bottom of a valley that is a minimum condition for that is this should be greater than zero that means that H is positive definite that is corresponds to a minimum and that is the sufficient condition for two variables.

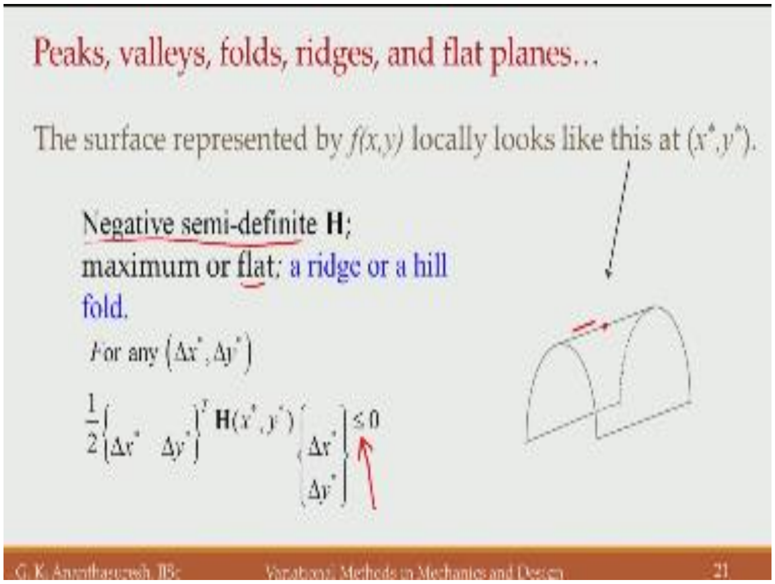
What was second derivative is simply greater than zero over several condition that is d square f by DX square greater than zero here this matrix should be positive definite meaning that this thing should hold okay, that is positive definite.

(Refer Slide Time: 17:22)



But what if it is not greater strictly but greater than or equal to 0 such a thing will look like this okay so it looks like a fold right like a valley fold as we call it then we call the matrix H positive semi-definite it is greater than or equal to 0 okay then you can customable let is now considered strictly greater when this is a minimum greater than or equal to it is a valley fold okay it is either minimum or a flat that is if I look if I come in this way right whatever point I take it looks like a minimum but if I come like this it is flat it does not change because in second order term is zero in that direction if I perturb right that is the value fold.

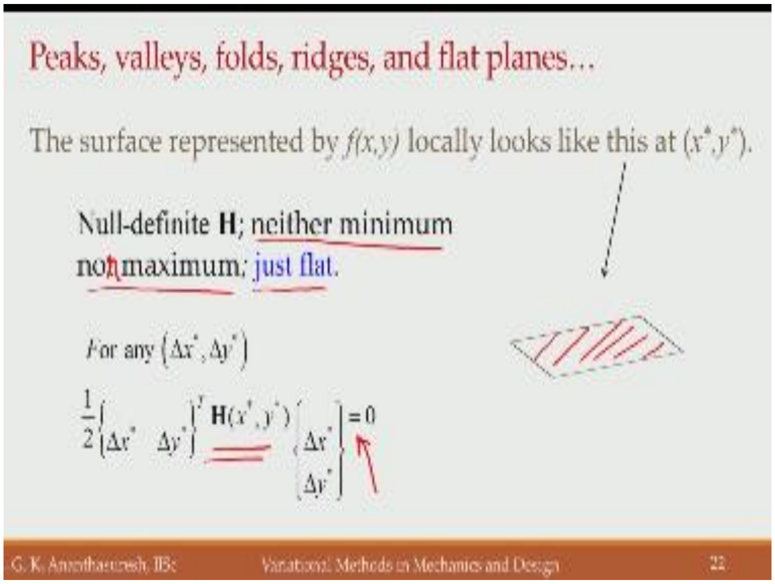
(Refer Slide Time: 18:14)



Similarly if it is strictly less then it corresponds to maximum right then you are the top of a hill peak of a hill you are the king now you are the top right in that case actually you are a king but it is negative it has a negative connotation negative definite when a matrix satisfies this property where it is strictly less than 0 when I definite that corresponds to a maximum at the peak of a here now you can talk about negative semi definite right then it is not strictly less than 0 that is second order term it is less than or equal to 0 that is a Ridge or a hill fold.

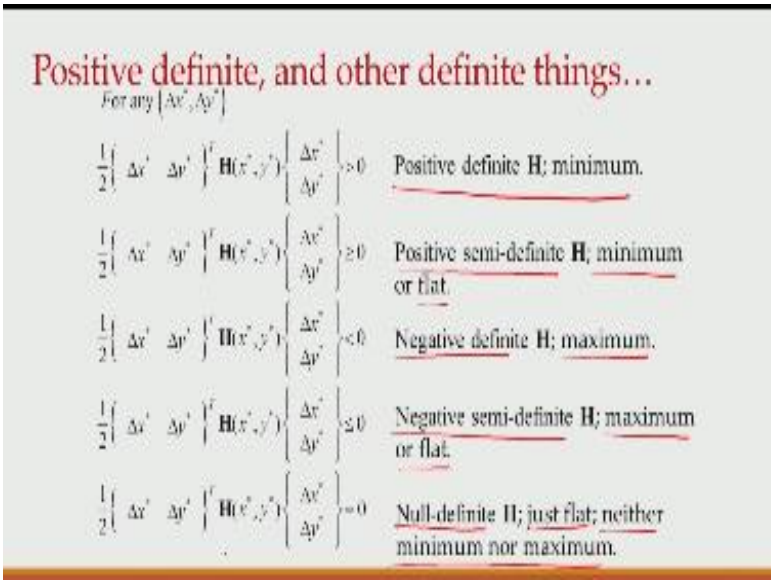
So any point here going this way it looks like it is flat okay but if I go in the other direction that is if I go like this then actually you have maximum Maxima a flat okay.

(Refer Slide Time: 19:14)



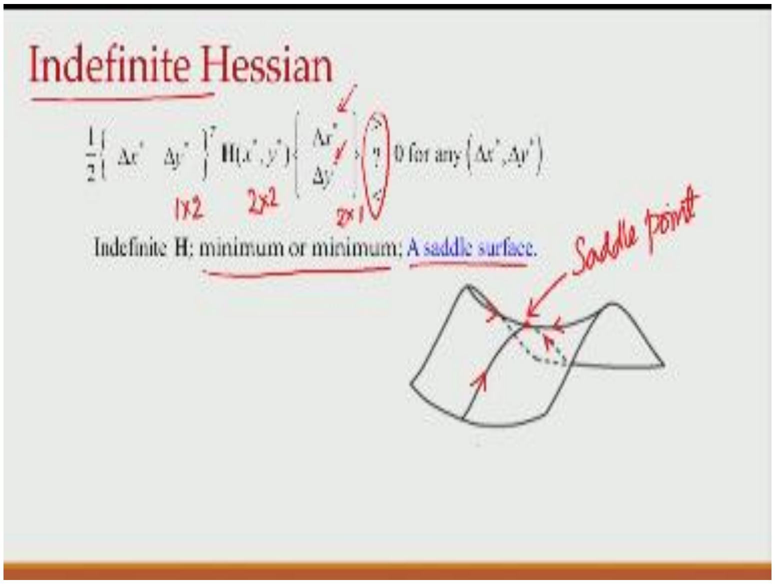
Now there is another one right where this is just equal to zero so no matter what λ X* and λ Y* star you take this matrix H at that point has a property that it just gives you zero in such a case locally the surface look like this plane okay plain old plane so that is there just flat then it is neither a minimum nor a maximum either a minimum nor a maximum right so such a thing can also happen there is one more case you can think of.

(Refer Slide Time: 19:53)



Which is before that let us first consider the five cases that we have considered passionate corresponds to a minimum positive semi-definite corresponds to a minimum or a flat region negative definite corresponds to a maximum negative semi definite very good sponsor maximum or flat null definite okay there it is just flat neither minimum nor maximum look at the condition greater than 0 greater than or equal to 0 less than 0 less than equal to 0 then equal to 0 okay, so all that is contained in this Hessian sufficient condition.

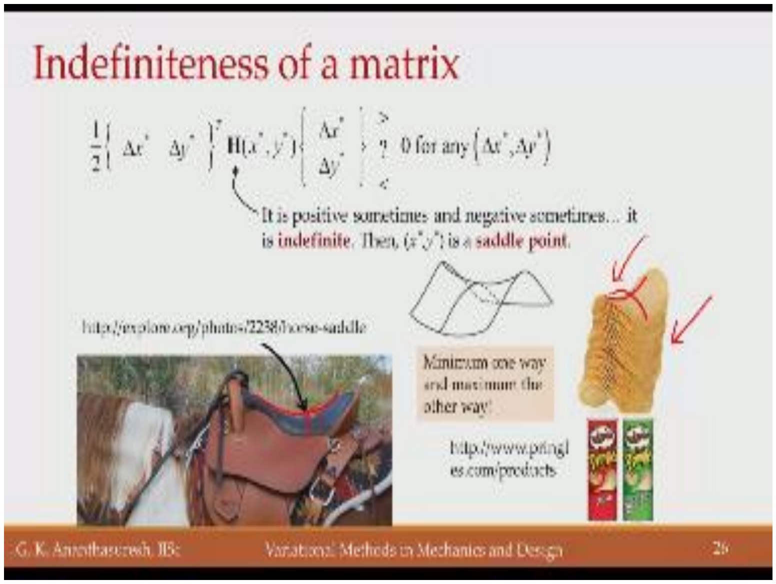
(Refer Slide Time: 20:34)



Now there is one more right so we have all of these again above a valley fold a maximum like a here hill fold and flat that is what we had seen pictorially there is one more something like this that which we call a saddle point if we are here going this way you see that it is a maximum but coming this way it is a minimum so the point is the Hessian now is indefinite it is not able to decide that is if you take arbitrary values of λ X* and λ Y* and do this multiplication which leads to a scalar because this is a 2 by 2 matrix this is a 2 by 1 vector is 1 by 2 row vector.

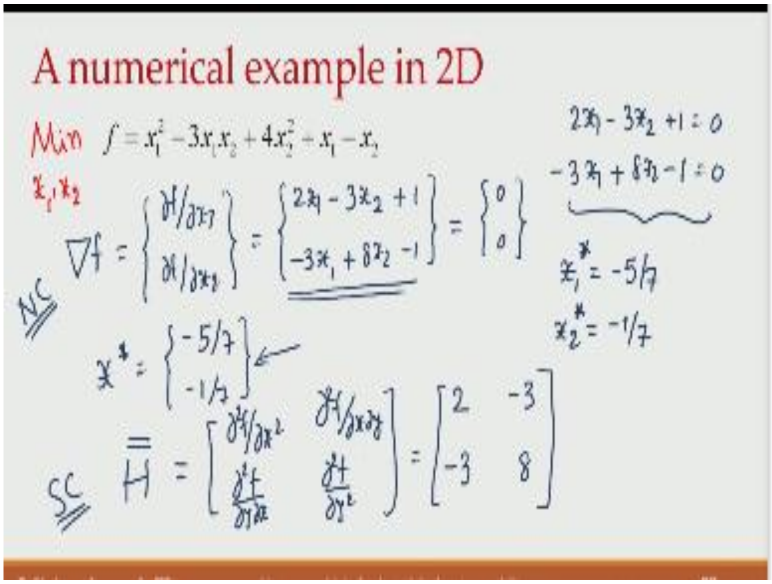
So overall you will get one by one write a scalar that we do not know whether it is greater than 0 less than 0or equal to 0 so that is kind of an indefinite matrix Hessian in which case we call it a saddle point or this point is called a saddle point that is how the horse saddle looks like okay there our f is neither a minimum nor maximum or rather it is both minimum and a maximum right coming one way it is minimum other way it is maximum right it is not absolutely minimum absolutely maximum it is both such a thing is called a saddle point problem saddle point example okay.

(Refer Slide Time: 22:10)



So for you if you have not seen a saddle that is how it looks like here right so it is a minimum that way going that way it is a maximum okay, that is that is a horse saddle looks like but you do not have to go and look at a horse saddle if the horses are not close to you but you can always see the chips right the Pringles chips but most of the chips have this profile most likely they pack them so that you know they do not break before we bite and break them so they do this these are all the saddle think that you can clearly see okay minimum this way maximum that way okay.

(Refer Slide Time: 22:55)



Now let us be a simple example in this two variable case now instead of putting as x1 x2 I am writing a reading as x1 x2 in writing X Y that we had taken does not matter two symbols X Y or X 1 X 2 if I want to minimize this function with respect to x1 and x2 okay, first thing is necessary condition that says that we have a gradient right so not a one variable problem so let me raise it okay first a tour at the gradient of F that means that I have to have ∂f / ∂ x 1 ∂ f / ∂ X 2.

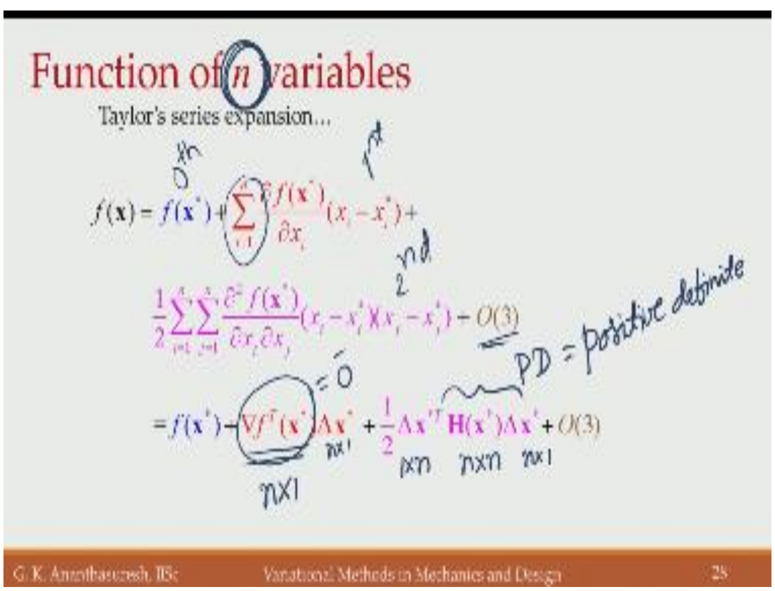
What is it now for this problem we just do a partial derivative so that is 2 x 1- 3x to wear take derivative respect to X 1 + 1 and then with respect to X to V artier derivative of F ∂ f / ∂X 2 there will be 3 x 1 + 8x2 – 1, right now this should be equal to 0 0 there is our necessary condition so we get two equations that we can solve so I have to solve these two equations 2 x 1 – 3 x 2 + 1= 0 -3x1 + 8x1 – 1= 0, if I solve these two this is a linear equations I will get the answer X 1 and X 2 so X 1turns out to be minus 5 x 7 x 2 turns out to be minus 1 by 7 you can plug it in and see.

Now if I want to know if this x* now right so what is our X *our X* is minus 5 / 7 - 1 / 7if I want to know if it is a minimum or a maximum because this necessary condition is what we have used this is NC right necessary condition that right sufficient condition to be sure so for that we need to write this Hessian matrix that means that how to do ∂^2 f b/ ∂ X^2 ∂^2 f / ∂ x ∂ y and then ∂^2 f /∂Y ∂X ∂ / ∂Y is the same mixed derivative and then those square f by ∂ Y square if I do that okay.

So I need to take ∂ f / 2 X 11 mode with respect to X 1 that will give me to here with respect to X to give me minus 3 here again to square f by ∂ / ∂ X looking at this thing right so it will give you minus 3 and then ate okay now according to our thing we have to check whether this H is positive definite we know the definition but operationally useful thing there are several one of the easiest is to check the Eigen values of this matrix H you both of them are positive you say it is positive definite okay.

You can look at the principal minors and check the signs of it they are all positive it will be positive n matrix and so forth but most easiest is to check the Eigen values in this case if H if you look at determinants if it is positive and you look at the Eigen values both of them were positive then I think we can conclude it is a minimum in which case in fact this point X star turns out a minimum you can plug it in and check okay that both Eigen values of this age are going to be positive and hence this point X star that we found is actually a minimize for this problem okay.

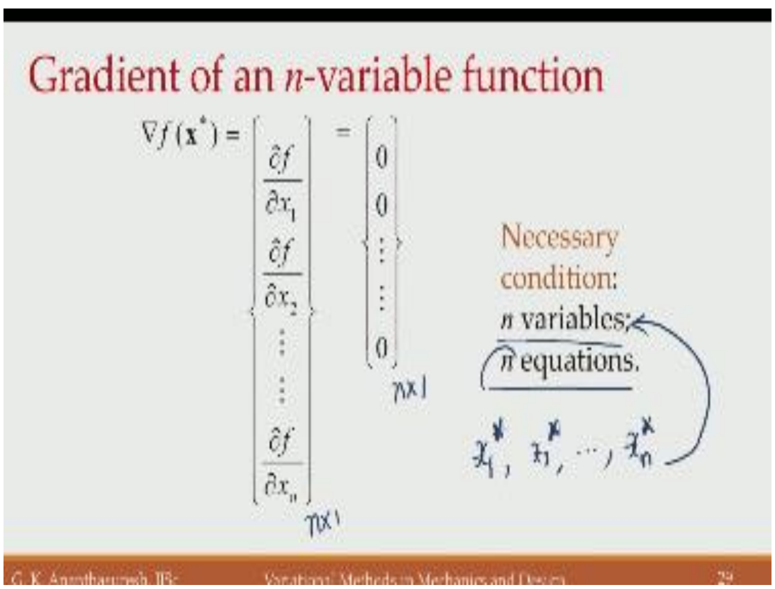
(Refer Slide Time: 27:09)



Now let us jump to n variables right then I can write the Taylor's expansion we just have the summation of if you see this is the first order term again our 0^{th} order term and our first order term and then I have our second order term and then hire our terms right in the same thing now the gradient we have a size n / 1 and the Hessian will have a size n / n okay, and we have perturbations also n / 1right so this is n /1 this is 1 / n okay.

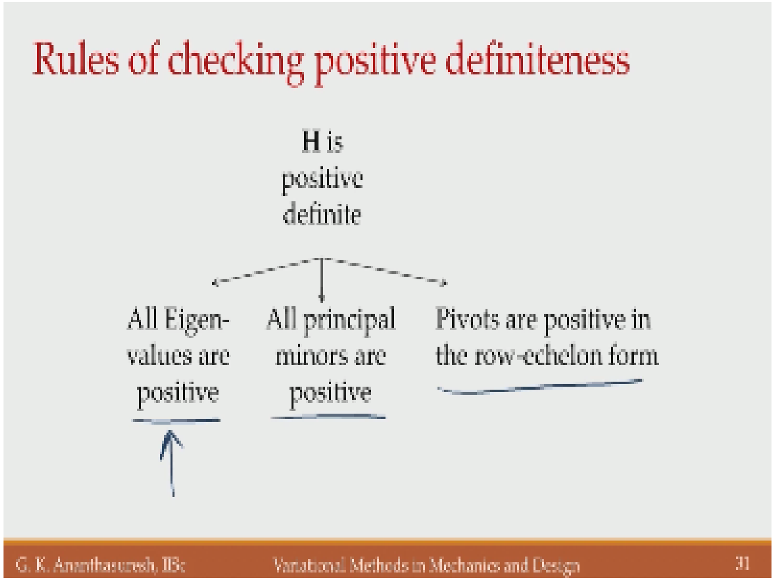
this is the N / 1 and we have transpose here so it is 1 / n this is n / 1 so we get a scalar eventually f of X so again if you think about our argument for necessary condition we want this gradient to be equal to zero vector and we want this to be pd meaning positive definite for it to be a minimum positive definite except that we cannot imagine this bowl and a hill and a valley fold and the ridge fold and so forth we just have to imagine that in the end dimensions our n dimension problem they will be bowels and hills and so forth valleys and hills and so forth right.

(Refer Slide Time: 28:34)



So necessary condition variable problem is that this gradient in n variable should be equal to n / 1right that that vector these n / 1these n / 1 so we have n equations herein n variables n variables meaning that we do not know x1 * x 2 *so fourth up to X and star these are all the unknowns that we have but we have enough equations n equations to solve for all of them.

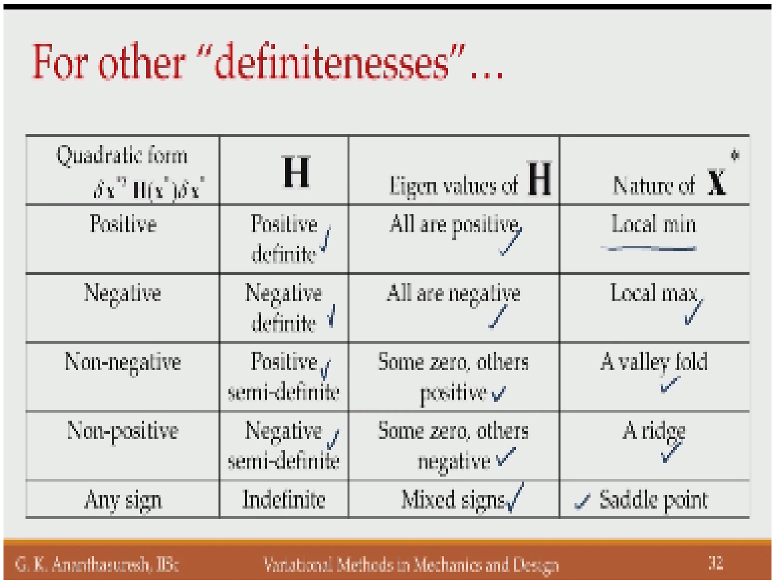
(Refer Slide Time: 29:08)



And then you look at the Hessian matrix which is now in n by n matrix a big one you look at the Eigen values if they are all positive then you conclude that that X star that you found is actually a minimize for the problem that is f of X that you have okay you should be first definite one way to check is that all Eigen values are posited for it so all Eigen are positive are all principal minors are positive all provides if you reduced row echelon form that you are learnt in linear algebra that is also there.

But as I said looking at Eigen values is the best way is easiest to a because in many ways new motor analysis you can easily find the musing software not yourself but you can find them.

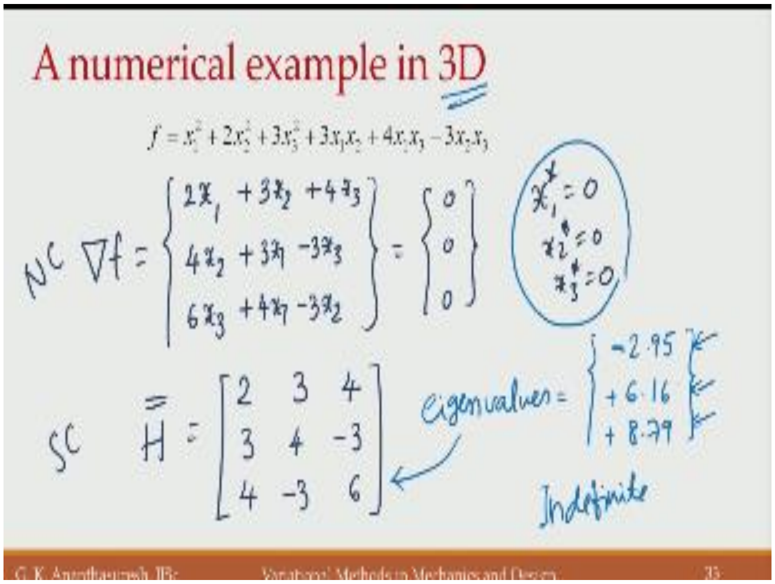
(Refer Slide Time: 29:58)



Quadratic form $\delta x^{*T} H(x^*) \delta x^*$	$\mathbf{H}$	Eigen values of $\mathbf{H}$	Nature of $\mathbf{x}^*$
Positive	Positive definite	All are positive	Local min
Negative	Negative definite	All are negative	Local max
Non-negative	Positive semi-definite	Some zero, others positive	A valley fold
Non-positive	Negative semi-definite	Some zero, others negative	A ridge
Any sign	Indefinite	Mixed signs	Saddle point

Again to summarize if it is a local minimum then you will have quadratic form is positive and the matrix is positive definite and the negative definite for local maximum semi definite for a value for then Eigen values are some are zero others are positive here all of them are positive here all of them are negative here some of them 0 or positive here summer 0 negative then we have negative semi definite okay then you get a ridge for and a saddle point will be mixed signs for this okay.

(Refer Slide Time: 30:33)



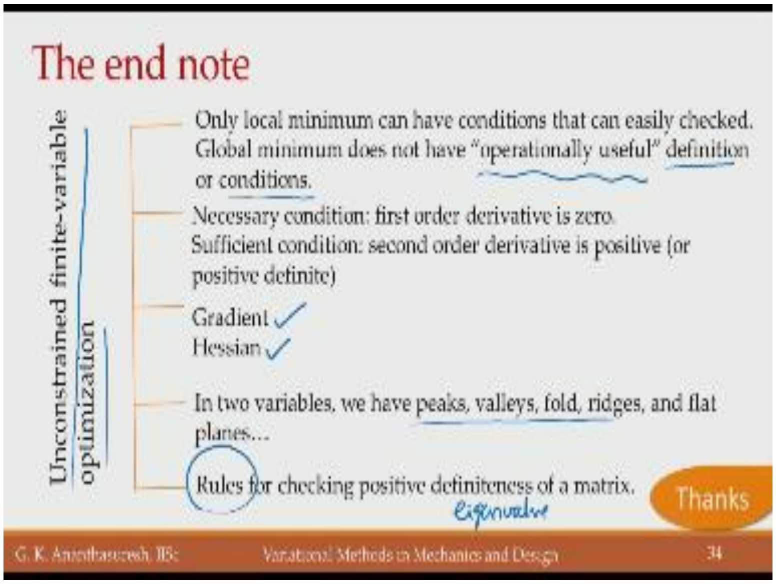
So if you take a numerical example in 3d now ND will be a little too much to solve in few minutes so I took a problem there are x1 x2 x3 so let us write the gradient for this so this will be partial derivative of F with respect to x1 that will give you 2X 1 and then 3 x2 and then 4 X 3 ok that is ∂ f / ∂ X 1 let us to do f / ∂ X 2 that is for x 2 + 3 x 1 - 3x3 derivative of F with respect to X 2now derivative F respect to X 3 will get 6x 3 + 4 x 1 - 3x to right.

So we have all of these should be equal to 000 each of them is equal to 0 get three equations and three unknowns right here there is a trivial solution actually if x 1 = 0 x 2 = 0 x3 = 0 actually satisfies this so you can actually call this okay, it is one solution the linear system we have one solution now whether the solution is necessary condition right we need to look at the sufficient condition also so let us write the sc here okay, that will be a 3/3.

So I had to write ∂^2 f /∂X 1^2 already have with respect to X 1 1 derivative that I will have 2 here and then 3here and for with respect to X 2 this will be 4 this 3 this will be - 3 and this will be 6 that is ∂^2 f /∂X 3^2 and then this will be for this will be - 3 this our SC okay, the Eigen values for this which I have so we have the hessian which is 2 3 4 3 4- 3 4 - 36 now we have to look at the Eigen values so Eigen values of this matrix which I have computed ahead of time are these here.

So 1 is -2.95 6.16 and 8. 79 these are the three Eigen values for this matrix they have this is negative these two are positive and positive mixed signs so if this matrix is indefinite so in this three variable problem we have this point which is like a saddle point in three dimensions it is not your horse saddle where you imagine a higher dimension so such a problem is not a minimize not a Maximize in one direction if I say x1 direction it looks like a maximum because I conveyor is negative other direction is positive so it is a minimum. So it is a saddle point okay.

(Refer Slide Time: 33:47)



Just to summarize today we talked about unconstrained finite variable minimization or optimization and we first talked about definitions and then we went to conditions definitions are not operationally useful but they tell you the condition of a local minimum and a global minimum and operational useful definition or condition we got it for the necessary condition and salvation condition established from it for a local minimum.

And then we talked about the concepts of gradient and Hessian and in two variables we have graphical interpretation of peaks valleys hill folds and ridges and floodplains and saddle points and mainly what we remember is to remember is that rules were checking past definiteness which is based on Eigen values okay which is the easiest in the next lecture we consider the constraints also in n variables and then after the G tour move on to calculate variations again. Thank you.