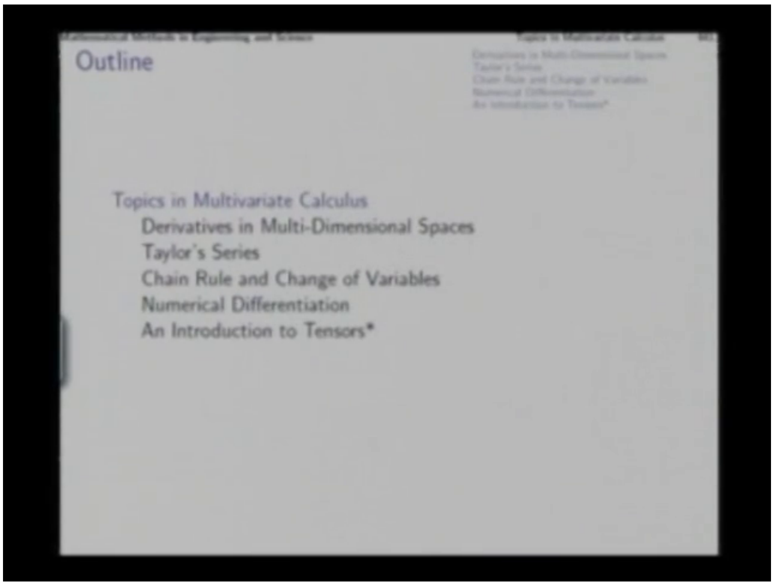
Mathematical Methods in Engineering and Science
Prof. Bhaskar Dasgupta
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur

Module – III
Selected Topics in Linear Algebra and Calculus
Lecture – 03
Multivariate Calculus

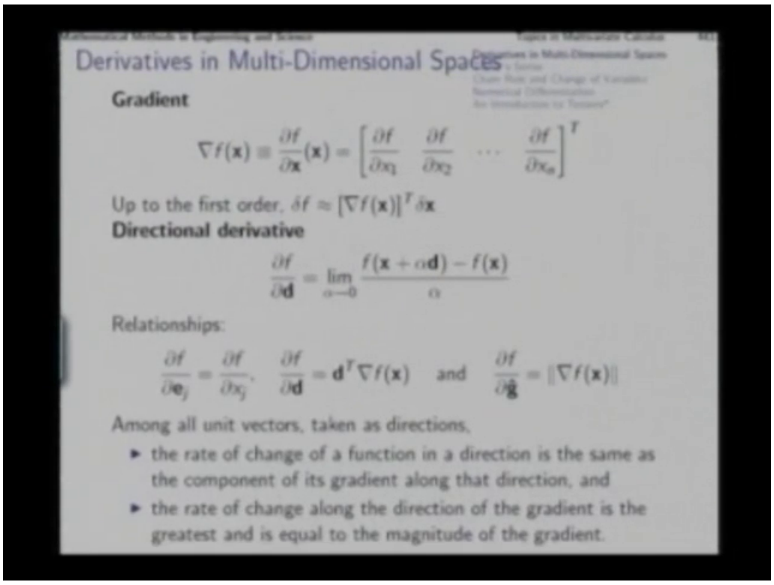
Good Morning, in the last lecture we completed our module of linear algebra. In the present lecture we start the small module on Calculus this will have 3 lessons topic in Multivariate Calculus and then 2 lessons on Vector Calculus.

(Refer Slide Time: 00:44)



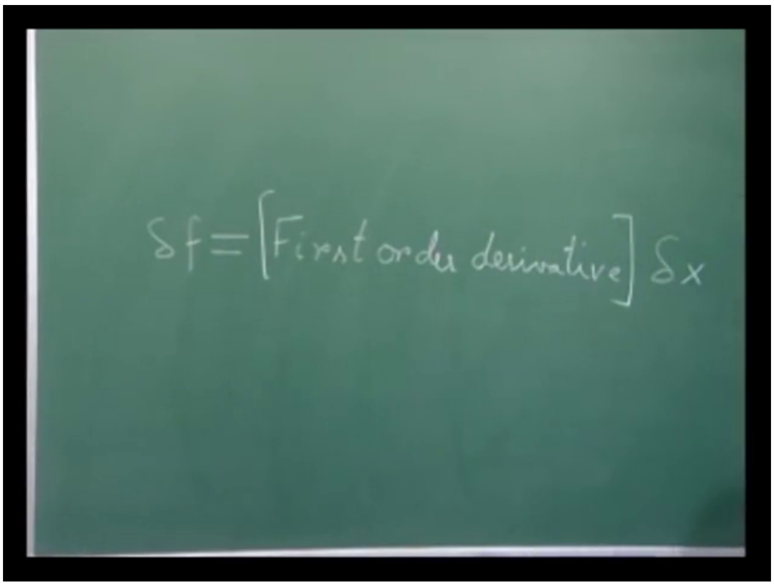
So, in this lecture in this lesson on topics in Multivariate Calculus, we will briefly summarize those topics of Multivariate Calculus which are likely to be confused or miss used inadvertently..

(Refer Slide Time: 01:01)



First issue is the derivatives in Multi-Dimensional Spaces first of all if we have a scalar function f of a vector variable for that the first order derivative is the gradient which has these following as the components del f by the del x 1 del x by del x by 2 and so on. So, the n partial derivative with respect to the n variable x 1, x 2 up to x n. These form a vector which is a column vector that is a transpose of this rho vector that is called the gradient. And in what sense it is the first order derivative of the function f in the sense is this that is the differential change the first order differential change in the function value is the product of the gradient and a first order change in the value of x.

(Refer Slide Time: 02:12)



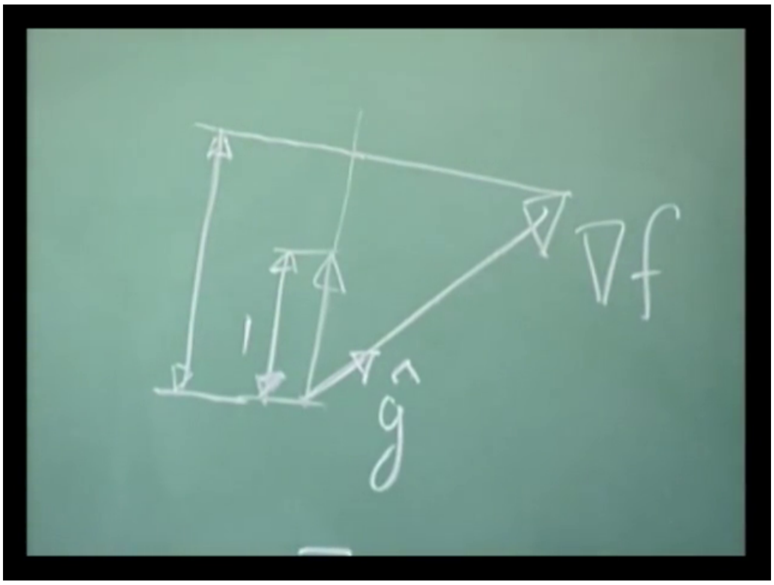
So this will be considered always the notion of first order derivative that is the first order derivative is something which multiplied with a differential change in the independent variable produces the corresponding differential change in the function or the dependent variable. So in this sense the gradient vector forms the first order derivate of the function f scalar function of a vector variable so this is the meaning of it.

Now, if further multi dimensional Multivariate function in a particular direction you want to find out the rate of change in that direction then the corresponding rate of change the scalar value of it is called the directional derivative and the definition of it is like this that is from the current point x in the direction d if we move in little step alpha and then we consider the change of the function value between the original point and the changed point that is this numerator and then divide that by the little step alpha that we took and then if we take the limit of this quotient as alpha tends to 0, then what we get is the rate of change of the function along the direction that is called the directional derivative.

Keep in mind that this vector d need not be a unit vector though quite often if we use a unit vector there then the meaning of the step size alpha is becomes more appropriate, however, it is not necessary that the vector d is unit vector it can be any vector for that method.

In particular you should sometime verify that these relationships always hold that is if we try to take the directional derivative of the function in the coordinate direction e j the j-th coordinate direction then that turns out to be the same as the ordinary partial the partial derivative with respect to x j that is a j-th variable and then you can also verify that the directional derivative with respect to direction d turns out to be equal to the inner product of the vector d and the gradient vector gradient.

(Refer Slide Time: 05:13)

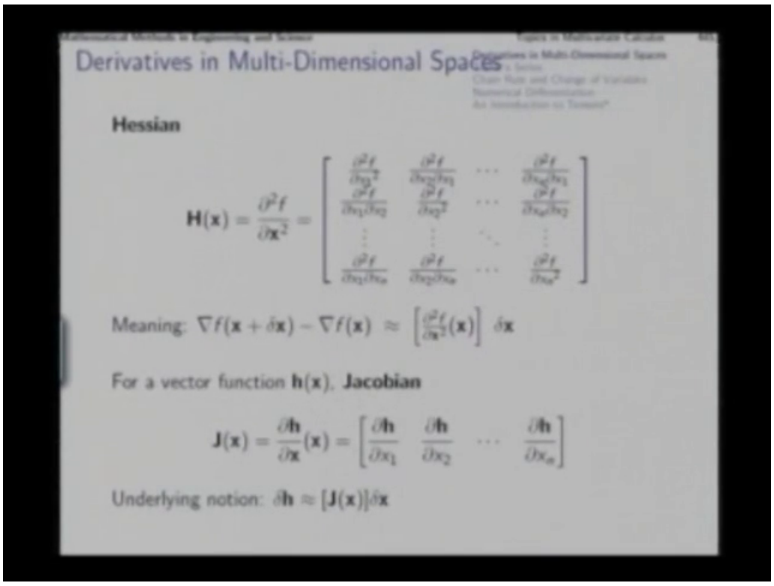


In particular this is another important relationship that is if you take the vector g hat as the unit vector along the gradient vector itself suppose grad f that vector is in this direction and if you take g, g hat as the unit vector in this direction. Then if you try to find out the directional derivative with respect to this unit vector then the magnitude that you get is the same as the magnitude of the gradient of f these relationships you should workout and verify.

The points to note are the following among all unit vectors taken as directions the rate of change of a function in a direction is the same as the component of it is gradient along the direction if you take this vector as the gradient and you want to find out it is component along this direction and suppose this is a unit vector then if you work out it is component along this direction like this then this component turns out to be the directional derivative of the function f in this direction this is one important point to note.

And the second point to note is that the rate of change along the direction of the gradient is the gradient that is among all directional derivatives the directional derivative in the direction of the gradient is the maximum among all unit vectors taken as directions now this is for the first order derivative. Now when you go to find out the second order derivative, then what do we get? We again should have the notion of the definition of a derivative that is the second order derivative should be a quantity which when multiplied with delta x should give us a small change in the first order derivative that is the gradient.

(Refer Slide Time: 07:31)

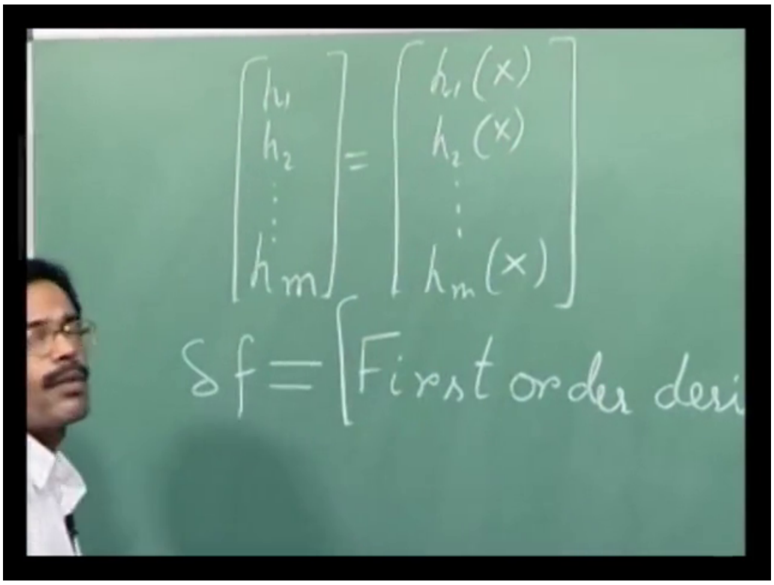


With that understanding the second order derivative the role of the second order derivative is played by this matrix known as the Hessian which is the n by n matrix formed by the second derivative second partial second order partial derivative of function f which respect to x 1, x 2, etcetera.

So, the diagonal entries are the direct second derivatives with respect to individual variables say this is del 2 f by del x 1 square this is del 2 f by del x 2 square and so on and the off diagonal elements will be in the form of del 2 f by del x I del x j right and this is a symmetric matrix this is a symmetric matrix.

Now in what sense this matrix is the second derivative of the function the sense is this a small change in the gradient at x and x plus delta x that small change in the first derivative is roughly equal to this matrix into delta x so this is the role of the second derivative that this matrix plays you can work out you can multiply this complete matrix with the vector del x delta x and see that what you get we will turn out to be the small change in the gradient vector that is a column vector. Now so far we have considered the function f to be a scalar function of a vector variable x, you can also consider a vector function of a vector variable.

(Refer Slide Time: 09:12)

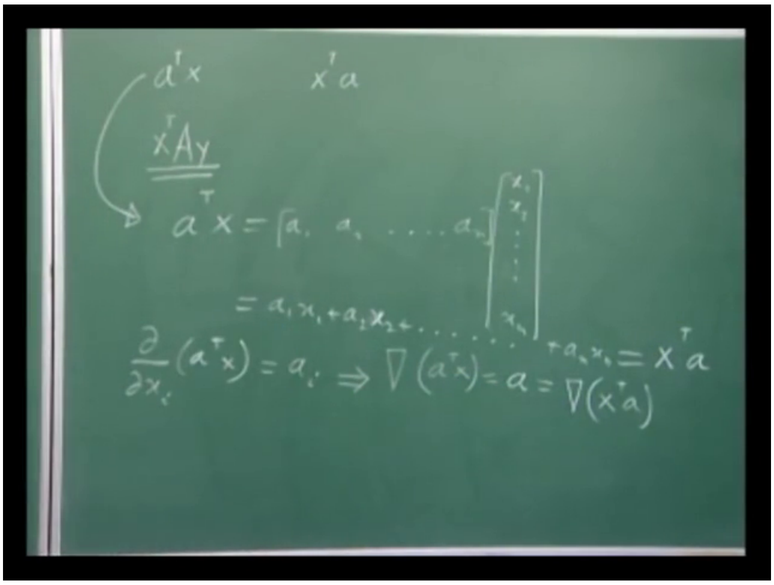


So, now you can say that suppose we have a vector function of m components like this, so here the function itself is also vector and the independent variable is a vector variable that means, the x the variable x is a vector which has so many elements that is x 1, x 2, x 3 up to suppose x n.

So, then if you try to work out the first order derivative of this vector function with respect to the vector variable then again you get a matrix and that matrix is called Jacobian, then you find that the Jacobian is given by this expression so each member of this del h by del x 1 then del h by del x 2 and so on each of them is a column vector because the function h itself is a column vector.

So, now in this matrix you will find that there are n columns and sorry yeah n columns and m rows h is a m component vector so there will be m rows so this will be a column vector similarly this will be another column vector and so on so such n column will be there and this when multiplied with delta x having members delta x 1, delta x 2, delta x 3 etcetera that will produce delta h that is in it is rows will have delta h 1, delta h 2 and so on. So this is called the Jacobian of the vector function h. So, that way you can say that the Hessian turns out to be the Jacobian of the gradient because gradient is a vector function of x it is derivative the n by n matrix is the Hessian the second order derivatives.

(Refer Slide Time: 11:46)

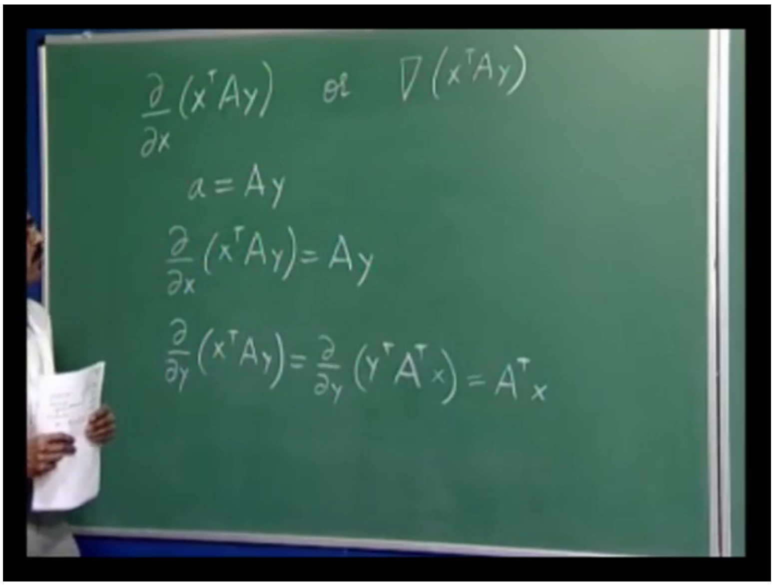


Now in order to see how we get the gradient and Hessian of sum very simple functions, let us see some examples we will try to find out the derivative of a few simple functions this is 1 this is another now this is a scalar function of x, x is a vector variable this is a scalar function of x and then we will consider derivatives of this with respect to x and with respect to y and in particular we will consider that situation in which y is the same as x that is another special case that you consider first this, this is a scalar function. So, we can we can find out it is gradient with respect to x so first of all let us verify that these 2 are actually same.

How do we do that we open it and say a transpose x is equal to a is a column vector so we will have a 1, a 2, a n and x is a column vector right so a itself is a column vector that is why a transpose becomes this row vector and as we open this we will get a 1, x 1 plus a 2 x 2 and so on. If we try to find out x transpose a then here we will have x 1, x 2, x 3, x 4 etcetera and there we will have a 1, a 2, a 3, a 4 etcetera right so the product will be the same thing so that is why this a transpose x and x transpose a are actually the same thing right.

Now if we try to find out it is gradient then in particular let us try to find out it is partial derivative with respect to the i-th variable so with respect to i-th variable all these partial derivative so all these we will go to 0 except the term a i x i so the derivative of that will be a i right so then as we try to find out del by del x 1 we will get a 1 del by del x 2 we will get a 2 and so on. So, the when we frame the complete gradient we will get a 1, a 2, a 3, a 4 etcetera so we will get a which will be the same thing as a gradient of this right. So gradient of a transpose a or x transpose a will be simply the vector a.

(Refer Slide Time: 14:59)

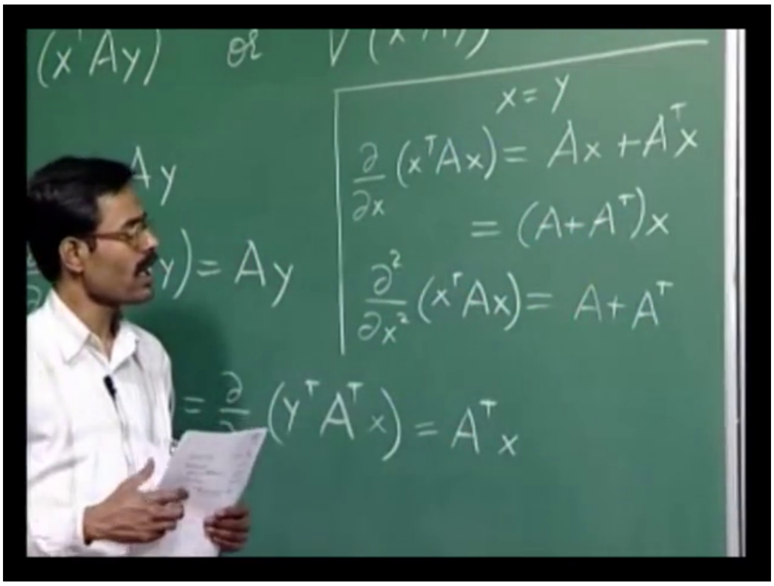


Now, consider this we want to find out the gradient of x transpose a y we can denote it as this or you can denote it like this we are looking for gradient of this functions. Now consider x is an n dimensional vector, a y will also be an n dimensional vector otherwise this multiplication x transpose a y will not make sense. Now in place of a y, if we can use this a itself in place of this a if we use a y then what we will find in that same original expression in place of a if we write capital A y then directly from that expression we get the gradient this will be A y right.

Now note that this function x transpose A y is actually a function of 2 vector variables if we consider y also as variable this derivative is it is gradient with respect to variable x with respect to variable y also we can find out the derivative when we want to do that we note that this x transpose A y is a scalar function now a scalar is a 1 by 1 matrix it is transpose is that scalar itself so we can replace this with it is transpose.

If we do so right we get y transpose A transpose x and now we have got a similar situation here we were trying to find out derivative with respect to x and x appeared here we had x transpose something and the derivative turned out to be that something here we are trying to find out the derivate with respect to variable vector y and the function is y transpose something. So the derivative will be that something so; that means we have A transpose x.

(Refer Slide Time: 17:50)



Now, consider this that in the special case where x is equal to y. Then what will happen then when we want to differentiate this we will find that this derivative will have 2 components; one considering this x as variable treating this as constant and the other in which we will be differentiating this x treating this as constant. So, the first one in which this is the second this x has been considered as constant. We will find that the derivative is A x.

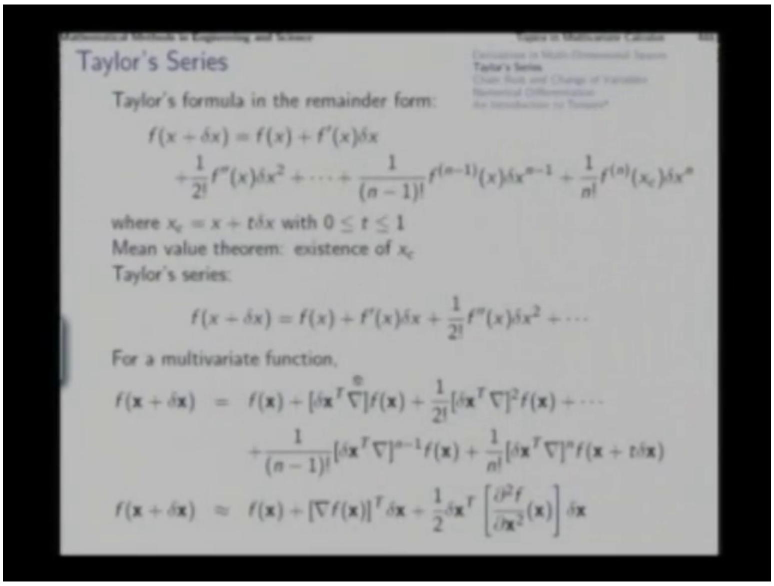
On the other hand in the second case where this is differentiated and keeping this as constant we will have the derivative as a transpose x from here. And that means, we have the derivative gradient as A pus A transpose x. Note that this matrix this is systemic with respect to what is A this matrix is symmetric.

Typically in this kind of a function which is called a quadratic form which we have encountered earlier also, A is taken as symmetric but even if originally A is not taken as symmetric. Finally this will be symmetric anyway. Now note that this is a vector function this is a gradient. If we differentiate this then we get the Hessian of the original function that is the second order derivatives.

And that will turn out to be A plus A transpose, because a small change delta x here will produce this into delta x that much change in the gradient in this function so this will be the Hessian. In the case of symmetric A this will be twice A because A and A transpose will be same.

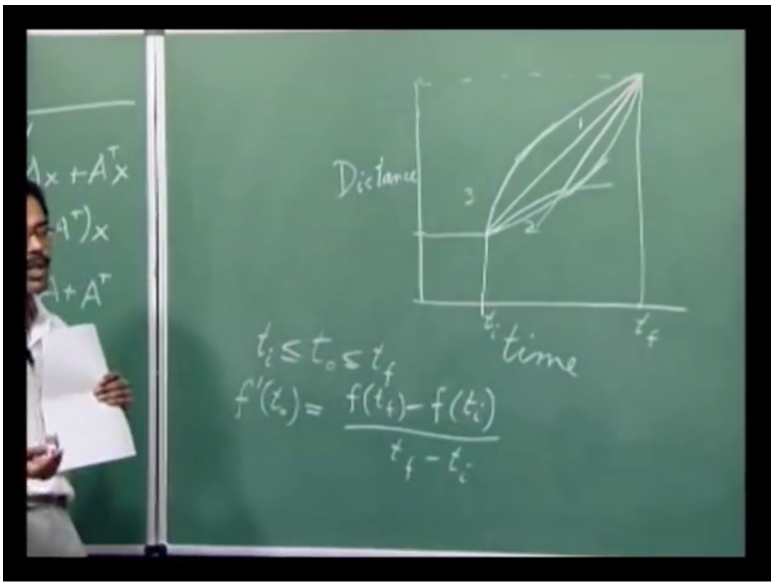
Now, the second important issue in this lesson that we explored is the Taylor's formula and Taylor's theorem and Taylor series.

(Refer Slide Time: 20:14)



So, let us try to motivate the discussion through a very practical issue.

(Refer Slide Time: 20:31)



We consider 3 trains all of them at a particular time start from one station and at another time reach another station. So from the first station to the second station the distance that is covered the same distance is covered by all these trains in the same duration. So one of the train go that a constant speed. So, this will be it is time versus distance curve graph. Now this is constant speed and speed is the rate of change of the displacement.

So, constant speed means here we get constant slope for the train. This is the first train. Another train, which initially goes a little slower than the first train so this is train 1 let us say this is train one. Train 2 initially goes little slower so that means, it is initial speed is less than this means little lower slope like this it starts. But then at the same time it reaches here.

If initially it was a little slow, but at the same time it reaches the final destination as train 1 then that means, that somewhere else it must have made up. So, if initially it was slow, but it could makeup finally, that means that somewhere else train 2 must have moved faster than train 1 so that means, with higher slop compared to this like this. Now if initially it was slow and in between; it must have become faster than the first train somewhere that means, that from slow to fast at some point of time it must have equaled it is speed with the speed of the first train. For the first train the speed is constant, so that means that between this initial time and the final time there must be some sometime where when the speed of the second train is exactly the same as this constant speed of the first train.

So, wherever that time happens to be there must be some point of time. Where that happened? where the slope of this curve this graph is parallel to this the same slope so, the tangent is parallel. Similarly if there is a train 3 which initially was moving very fast. But then finally, it reach the destination station at the same time that means, somewhere it must have turned from faster to slower.

So, there must be some point sometime when it is slope was the same as this one. So, the tangent somewhere must have been parallel. So from faster to slower there must have been a point of time when if speed same as this right. Now this is necessary because the speed cannot change suddenly that is because the speed is a continuous function of time.

If that were not so, if the speed could suddenly change then this was not necessary. If speed could suddenly change that means, the graph could have suddenly turn it is direction then it would be possible to have it like this. At this point suddenly there is a change and there is no tangent there is one tangent like this another tangent like this. But in this kind of situation if we say that the first order derivative is continuous that is speed is continues, then it becomes necessary that at some point of time between this and this there must be the slope which is parallel to this which is same as this ok.

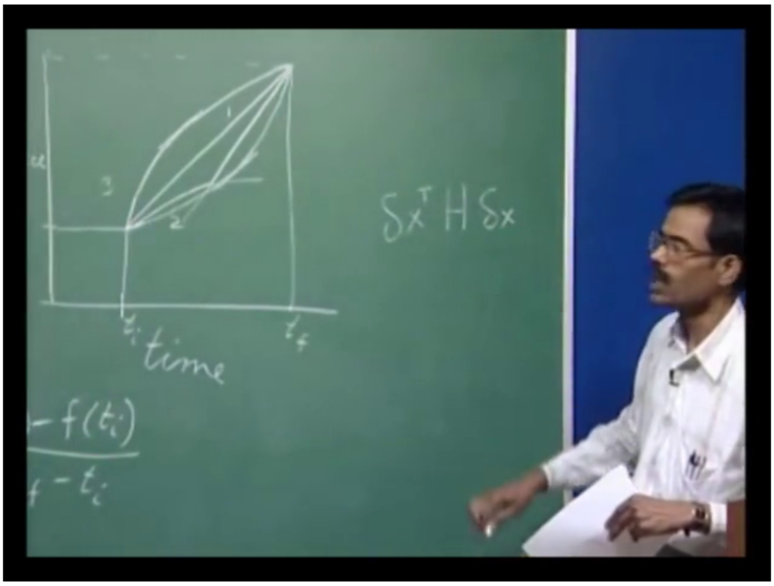
So, this in mathematical terms we will mean that if the function is continuous and if the first order derivative is continuous. Then between the t initial and t final there must be some time say t 0, such that at t 0 the slope is equal to the average slope that is this is Lagrange's mean value theorem right. So, the Lagrange's mean value theorem says that, if the function between these 2 points is continuous and it is derivative is also continuous then there must be some point in this interval, where the first order derivative is the same as the rate of average rate of change. So, that is the statement here if we consider only up to first order that is f of x plus delta x is equal to f x plus here in the place of x, we will have some x x c here which I have represented as t 0 in that place we have x c.

So, that is Lagrange's mean value theorem first order derivative. If the function is n times differentiable then we can go on extending that and we can include first order change, second order change, third order change up to n minus one-th order change. And that final mean value we can write in this form the n-th derivative in the Taylor's formula is then evaluated at x c some point in between the interval.

So, here the way we have written here that is t 0 some value between t i and t f we could have said that where t 0 is equal to t I plus some parameter into t f minus t i it would be same thing and that parameter can be between 0 and 1. So, we can say in that sense also, so if we say that in that sense with up to n-th order derivative included. Then this term is known as the remainder term and this is Taylors mean value theorem. Taylors mean value theorem basically assures us of such n x c such a value x c.

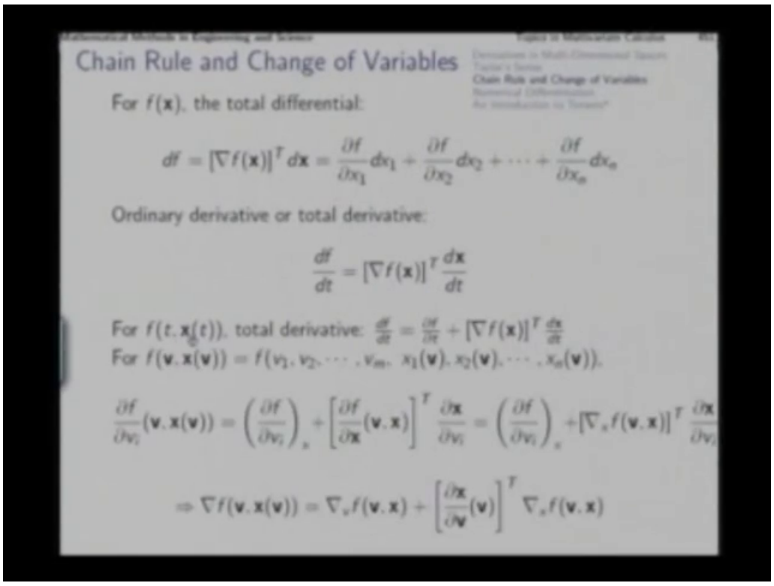
Now, if we say that there is a function in which is infinite times differentiable. Then this reminder term we can go on postponing and we can have an infinite series and that is this Taylor series it goes on right this is for a scalar variable. Now what it is a analog for vector variable or for Multivariate function. For a Multivariate function you will find that this f x time is same here. This f prime x into delta x that term will be included like this delta x transpose the gradient plus this second order term will be taken as delta x transpose the Hessian into delta x and so on. So, you can say that this delta x transpose grad del square when we write it we get when you open it out we get this ok.

(Refer Slide Time: 29:19)



So, the third term third order term onwards terms becomes very complicated and you cannot write it in the form of matrix multiplicity and that is why most of the sensible analysis goes only up to second order. So, this is what is written here is the second order truncated Taylor series that is we have truncated it up to the second order. This is an expression which is going to be very useful for many of our analysis later.

(Refer Slide Time: 29:54)



Another important issue that we will need quite often is that change of rule is Chain Rule and change of variables.

So, we know that for a function for a scalar function of a vector variable we can change x 1, x 2, x 3, x 4 etcetera independently. And if we make several such changes then this quantity is called the total differential in which the components d x 1 the smaller small changes d x 1, d x 2 etcetera effect these individual changes in the function value. Now if we divide this whole thing with dt and then take the limit when dt tends to 0 in the sense that x the variable vector x itself is a function of p another parameter.

Then what we get? As the total derivative d f by d t which is actually the ordinary differential coefficient of f with respect to t turn out to be grad f transpose d x by d t. So, this is the way we differentiate a scalar function f of a scalar variable t when the description of the function is available not directly with p but through a vector variable x whereas, f is a function of x which is a vector and x itself is a vector function of a scalar variable t. In that sense in that case the application of Chain rule d f by d t in the ordinary Calculus would be d f by d x d x by d t.

So the Multivariate analog in which x is Multivariate turns out to be like this gradient of f transpose d f by d t where, x is vector, t as well as f are scalar. Many situations arise in which the function f is expressed as a function of p and x f of p and x in which x itself turns out to be f function of p.

In that case the total derivative of f with respect to t should include the contribution in the derivative to direct dependence on t and also the contribution through the dependence by the dependence over t through the vector variable x. Then what we will have we will have this term as well as the ordinary derivative the partial derivative with respect to t. So, the total derivative d f by d t turns out to be the partial derivative with respect to t.

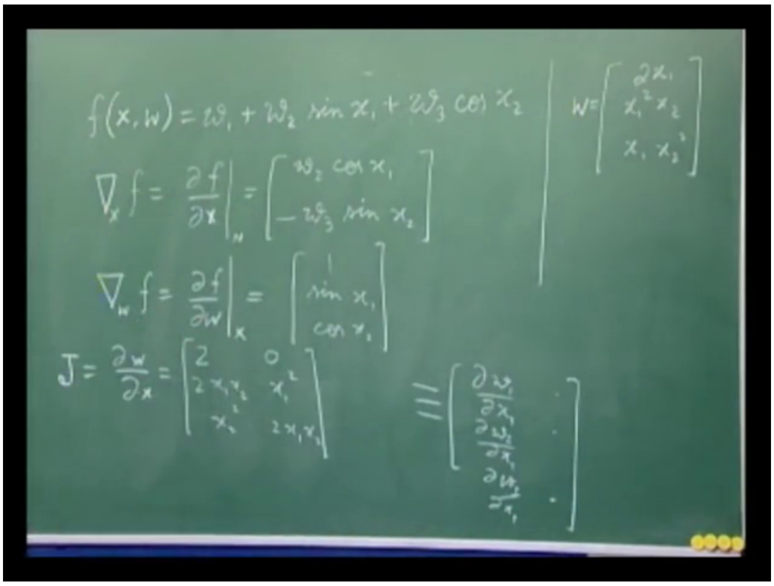
Considering this entire x as constant and then a separate component added to that which contributes the derivative part which is due to it is dependence over t through this so that is coming from this expression. Now it may happen that f is a vector function of v and x in which v is a vector variable and x is another vector variable which is again function of v.

So, in a similar model the way in which we worked out this. We find that the derivative of f with respect to v i turns out to be 2 parts: one by considering v i only as the variable and v 1, v 2, v 3 other v's as well as these x's this v is kept constant and it is differentiated with respect to x and then multiplied with the derivative of x with respect to v i like this so this is in the same sense as this.

And when such partial derivatives we assemble together then we get it is derivative with respect to v ok that is the gradient that is the full gradient of f with respect to v. While using this kind of expressions you should exercise question to note this transpose and then the order in which these matrices and vectors are multiplied and so on.

So, quite often the sizes of the matrices and vectors would give you a quick check. But sometimes they may not, but always if you try to see what each quantity means and what you will get if you write the components of those matrices clearly and what is the meaning of the Chain rule applied in the Multivariate context you will find that the confusions get removed. In order to see how it is used, Let us consider this small example.

(Refer Slide Time: 35:01)

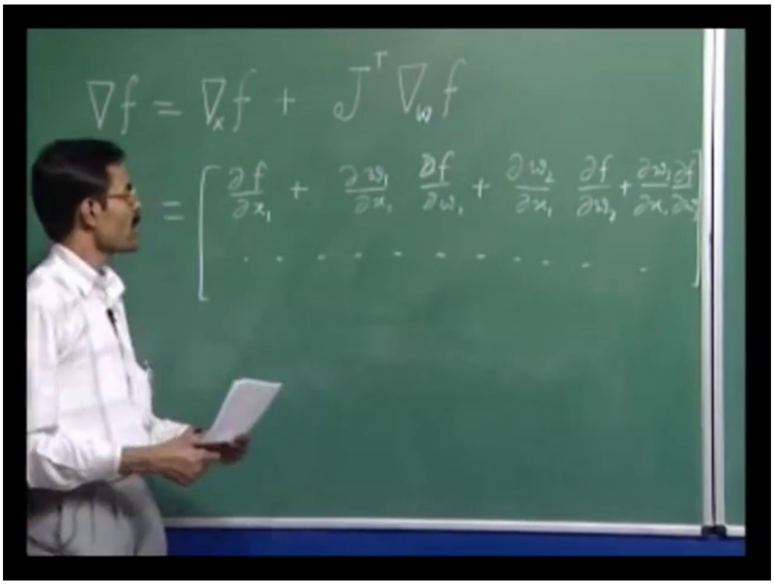


We have got a function of 2 vector variables x and w and this is the function in which we have got a 2 dimensional vector x, x 1, x 2 and x 3 dimensional vector w 1, w 2, w 3 and the function is like this. Now in this the vector w itself is a function of x which is given like this. Now from here if we want to find out it is gradient with respect to x then we can talk of 2 such gradients one is the partial gradient with respect to x that is evaluated keeping w as constant. So, if we do that then the gradient of this, the partial gradient keeping w's as constant will be w 2 cos x 1 and then minus w 3 sin x 2 and that is it right.

If we similarly construct the gradient with respect to the variable vector w keeping x constant then we will find the derivative with respect to w 1 is 1, derivative with respect to w 2 is sin x 1 and derivative with respect to w 3 is cos x 2. Now this relationship which gives w as a vector function of x if we differentiate this with respect to x then we get the Jacobian that is del w by del x and that will be this c by 2 matrix. This is the derivative of w 1, this is the derivative with respect to x 1 that is derivative of w 1, w 2 w 3 with respect to x 1 derivative of w 1, w 2, w 3 with respect to x 2.

Now that formula there this formula apply to this particular problem will give us grad f, the total gradient of f which accounts for the derivative the rate of change because of this direct change as well as the change through a change in w because of change in x.

(Refer Slide Time: 37:19)



So, that will be this direct change in f because of change in x plus Jacobian transpose this Jacobian transpose multiplied to this 1. So, when we construct this now see what it means this Jacobian what does it mean it means del w 1 by del x 1 ok.

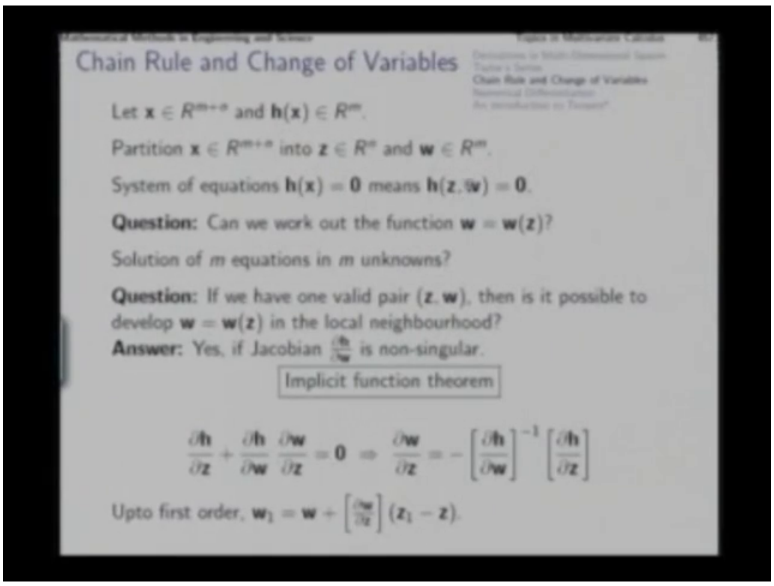
Then del w 2 by del x 2 del x 1 and so on. Then del w 3 by del x 1. Similarly these are with respect to x 2 now it is transpose will have these 3 fellows in the first row. So, here we will have in the first row we will have del f by del x 1 plus here we will have the 3 elements that I have written there they will be in the first row of j transpose multiplied with this.

So, then we will get del w 1 by del x 1 into del x by del w 1 del f by del w 1 plus the second one del w 2 by del x 1 del w 2 by del x 1 into from here the second entry that will be del f by del w 2 plus the third entry from there del w 3 by del x 1 into the third entry from here that will be del f by del w 3 that is the first element of this.

Similarly, there will be a second element of it in which in place of x 1 we will have x 2. Now note that this is a direct variation for x 1 this is a variation with respect to x 1 through variation of w 1 del f by del w 1 into del w 1 into del x 1 and so on. So, the variations in w 1, w 2, w 3 due to a change in x 1 account for small changes in f through this, through this and through this. And the direct dependence over x 1 is accounted for here you can expand this and find out the gradient that you get from this and then try to do the same thing all over again by first putting those w values here and getting it in terms x 1 and x 2 only and then finding the derivatives directly.

So, through both methods you should find out the final derivatives and see that they match. And we proceed to another important issue that concerns such quite often.

(Refer Slide Time: 40:43)



For that, let us consider a vector of m plus n dimensions and a function of it h of x which is an n dimensional function that is a vector function h of a vector variable. The variable x is of m plus n dimension and the function is of m dimensions only. If we partition this vector into 2 parts n variables in z and the rest of the n variables in w like this. Then this relationship becomes h of x and w, right.

And now if we equate this to 0 vectors that will mean h of z and w equal to 0. Now this gives us m equations. The function h of x is an m dimensional function m components, so, this equal to 0 gives us m equations in all these unknowns all these variables z and w. Now note that z has n variables and w has m variables and there are m equations.

Now if we give the value of z that is the n variables listed in z if we describe their values then what it becomes, it becomes m equations in m variables in w. Now m equations in m variable same number of variables can we solve it and then say that for every set of values for z can we work out w? If we could then we would basically have a straight forward function w of z this is a question can we work out the function w of z that is by prescribing z1, z 2 up to z n by prescribing those n variables can we determine the rest of the variables that is w 1 to w m from these m equations.

So, in general for the non-linear problems all over the domain we cannot do this through a single close form expression. But then if we ask for something less that is if we say that we have 1 valid pair z and w which satisfy this requirement. Then in the immediate neighborhood can we frame can we form a first ordered approximation this is possible under a certain condition.

How? So, for that what we do we consider the derivative of this is a vector function of a vector variable of 2 vector variables. So, if we try to differentiate this then we get del h by del z plus we are considering the derivative with respect to z because z is prescribed z is going to play the role of the independent variable and w is going to play the role of the dependent variable.

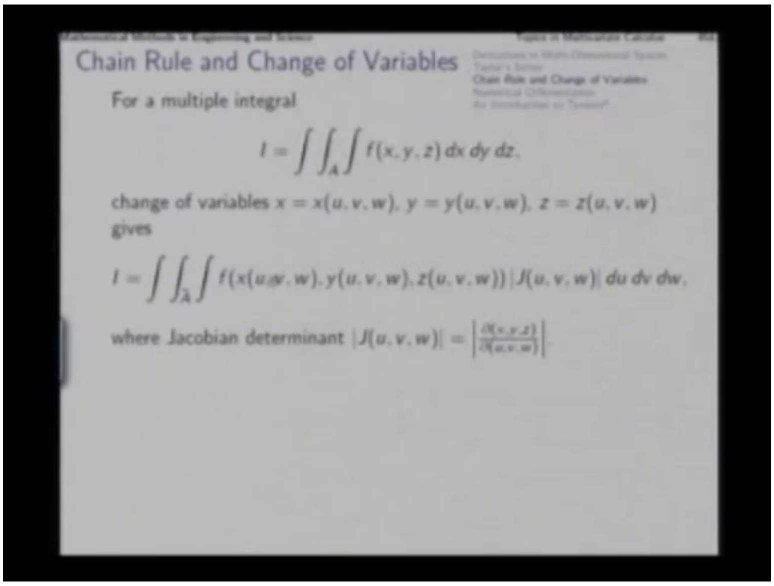
So, we try to find out it is derivative with respect to z, so it is derivative with respect to z we will have 2 parts one direct and the second 2 w. So, the derivative of this is del h by del z direct derivative and plus the other component will be del h by del w del w be del z, so we have this.

Now note that h is of size m, w is also of m this will be a square matrix. So, we can talk of the problem of finding it is inverse so for that we take del h by del z on the other side try to pre multiply with it is transpose and then we get del w by del z as this that is if this matrix is invertible. If we can do this then we can say the first order change in w can then be found out by this that is through a change in z by this amount.

The corresponding change in w that is w 1 minus w will be given like this del w by del z into this. So, this is a first order approximation of the function w of z around a point pair z and w. Now what is why this local neighborhood description in this manner if possible if the Jacobian del h by del w is non singular if this is invertible. So, that is the condition and this result is known as the implicit function theorem which is going to be very useful in many of our applications where at one point we try to find out a locally valid first order approximation and continue with that particular approximation as the local description of the function. So, this will be possible when the Jacobian matrix is invertible, non singular.

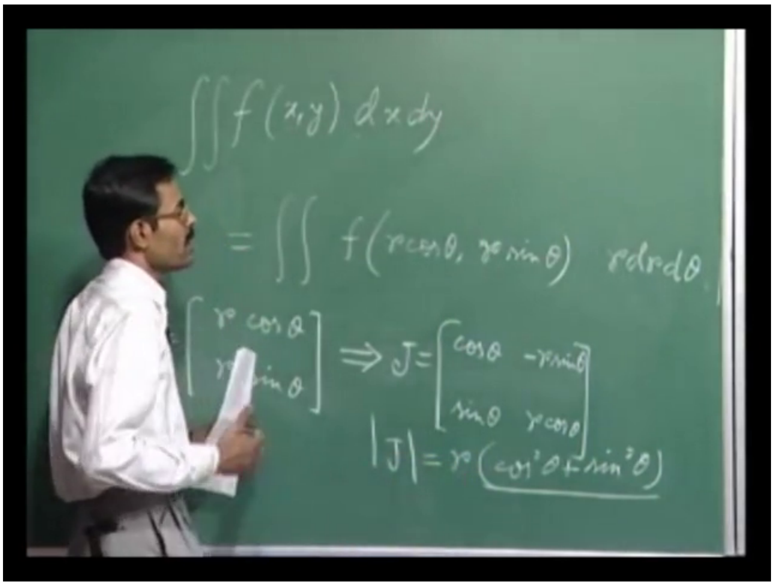
After this we have another few small issues that we need to quickly have a look into:-

(Refer Slide Time: 46:09)



For a multiple integral, when we conduct a change of a variable say the original integral is in terms of x y z. Now if you conduct a change of variables in this manner x, y, z all 3 of them are taken as functions of 3 new variable u v w. Then first of all we need to transform the domain from A to A prime A bar where A bar is the corresponding domain in the u v w space. And then here in place of x y and z we put x y and z in terms of u, v, w and then in place of d x d y d z we have this determinant of the Jacobian into du dv d w. So, this Jacobian determinant is the element that transforms a volume in the u, v, w space to a corresponding volume in the x, y, z space. As an example consider this small case in the 2 by 2 situation. So, quite often for evaluating a double integral we transform from rectangular to polar.

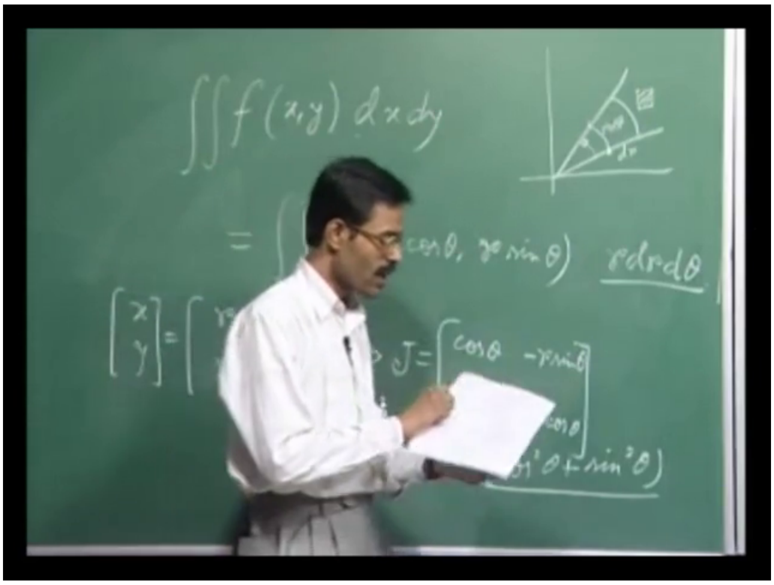
(Refer Slide Time: 47:30).



And then apart from transforming the limits of integral, when we do the transformation from rectangular to polar we keep here r cos theta r sin theta and in place of d x d y we use r d r d theta. What is this r doing here for that you can see that if you take x y as this vector function of r sin theta. Then if we try to work out it is Jacobian, The Jacobian will be here 2 by 2 matrix in which we will have del x by del r here and then del x by del theta here del y by del r here from here and del y by del theta that is.

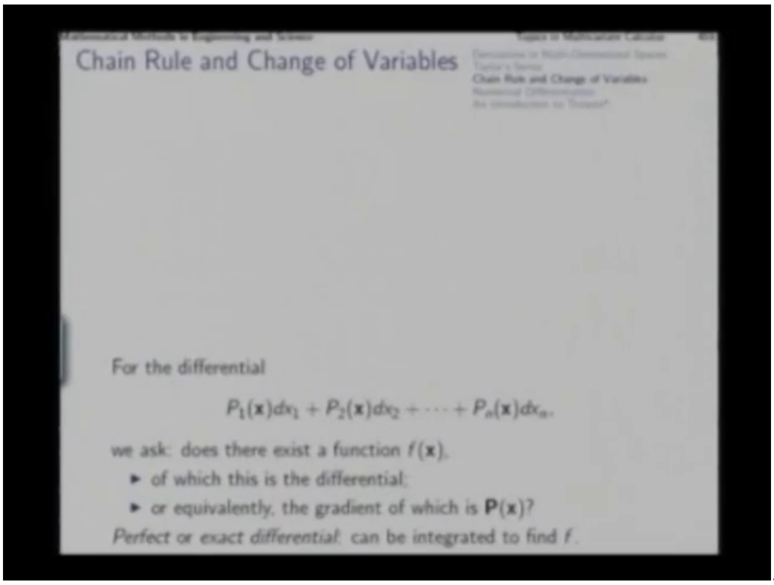
So, if you try to take the Jacobian determinant you will find that this cos theta into r cos theta minus sin theta into minus r sin theta. So, you will find r cos square theta minus minus plus r sin square theta that is like this, this is 1. So, the Jacobian determinant is r that is why while transforming the unit volume or unit area in this case d x d y transformed to r d r d theta.

(Refer Slide Time: 49:49)



The same thing you can find out geometrically also, In the rectangular coordinate system from this point a small area element turns out to be d x into d y. In the case of polar coordinate the typical area element is like this. Now this is point r theta, this is angle has changed to d theta. So, this length is r d theta, on the other hand this is radial change this is d r. So, you find that the elemental area is d r into r d theta. So, that gives you this as the elemental area ok.

(Refer Slide Time: 50:41)

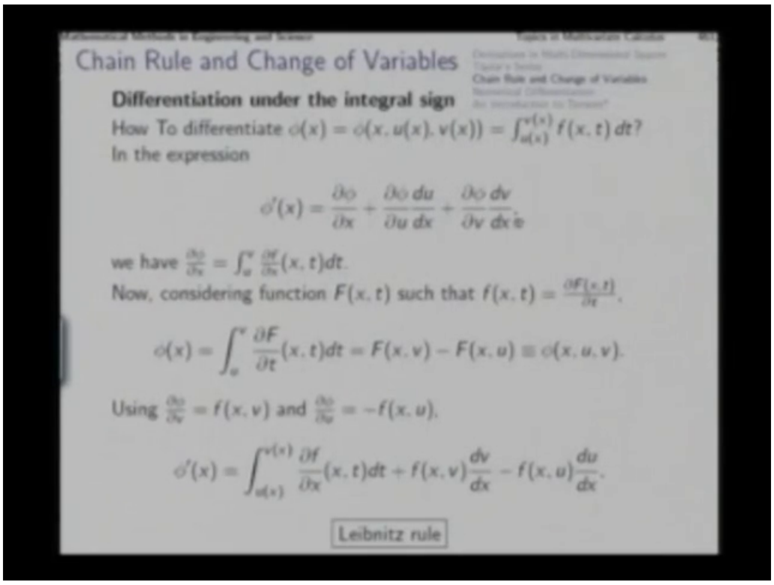


.

Now, if we continue in our next important issue here this is going to have a lot of use in the vector (Refer Time: 50:45) segment. So, here if we have a differential quantity and then if we ask this question that does there exist a function f x for which this is the differential that is we can talk about a function f for which d f turns out to be this or equivalently we can say that for which the gradient is the vector function t x. The components of which are t 1 x, t 2 x sitting here.

So, if the answer is yes then we say that this particular differential is a perfect differential or an exact differential and it can be integrated to find f for every differential this may not be valid.

(Refer Slide Time: 51:42)



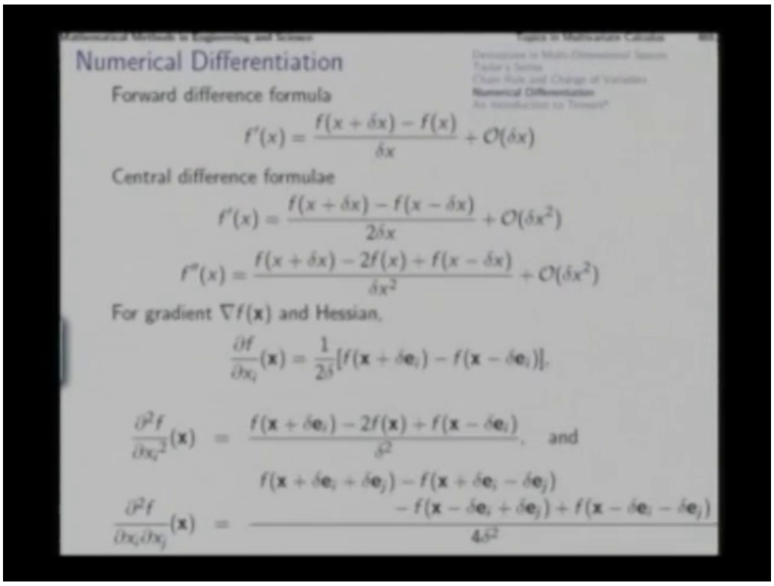
The last point that we consider in this lecture is the formula for differentiation under the integral sign so that is useful in differentiating a function which is available in the form of an integral. The integrand here is a function of x and t integral is with respect to t and the limits of the integral are functions of x.

So, if we try to differentiate this considering this as a function of x u and v u and v themselves being functions of x again then the straight forward Chain rule will give us this expression for the derivatives. This first term is the direct derivative with respect to x which does not consider the dependence of x and t which is here through the limits of integral so that is simply this in which the derivative is taken into the integral time. But that is only the first term, in these 2 terms d u by d x and d v by d x are derivatives of u and v which are known functions of x, but we need to find out these 2 partial derivatives.

So, for that what we do we considered a function capital F whose derivative with respect to t is this small f our small f here and then phi x turns out to be here itself in the place of f if we put del f by del t then we get this. And the integral of that is capital f of x t evaluated at v minus evaluated at u. Now here you see this whole thing is the corresponding function of x u and v. Now if we differentiate this here with respect to u partially and with respect to v partially then we get these 2 partial derivatives. So, we do that so derivative of this partial derivative with respect to u will be coming from here that is minus del f by del u that is minus del f by del t evaluated at t equal to u and similarly for this.

So, from here we get these 2 expressions and insert these 2 expressions for the partial derivatives here and then we get the complete expression like this, this is called the Leibnitz rule. The special case is of course, that one in which u and v, the limits of integrals are constants or independent of x, in that case we will have only this first term these 2 terms will not be there.

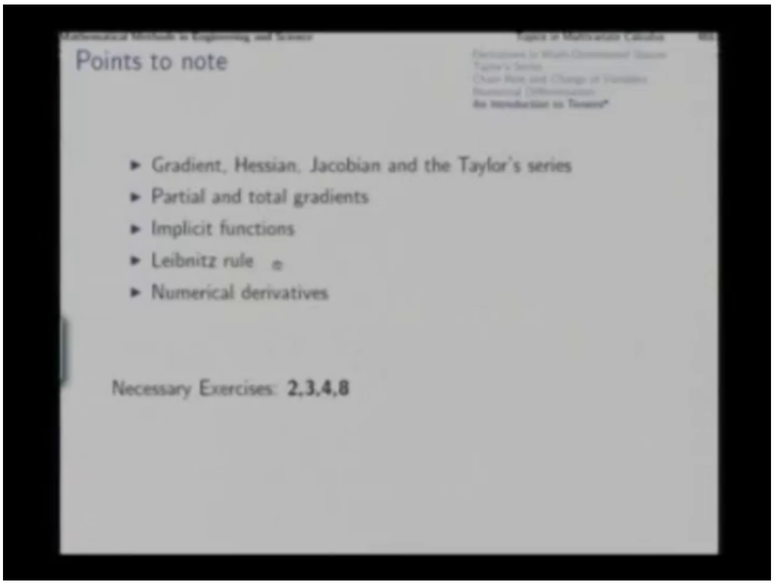
(Refer Slide Time: 54:37)



So, this is called a Leibnitz rule and it is used for differentiating under the integral sign. There is another small topic here in the lesson which we will omit because it is quite straight forward, but I advise you to go through it in the textbook and be conversant with it because it will be useful in many of the applications that we consider later in the course.

So, these are the derivative formulas for scalar functions and for variance and Hessians the vector extensions of them can be worked out like this.

(Refer Slide Time: 55:14)



So, in this lesson these are the important issues that we have considered, the Multivariate functions derivative the sense of it:- The Partial and total gradients, Implicit functions Leibnitz rule. So, these are the important topics that we will be using again and again in the rest of the course. In this lesson the necessary exercises that you must complete to develop the essential amount of understanding are these.

Thank you.