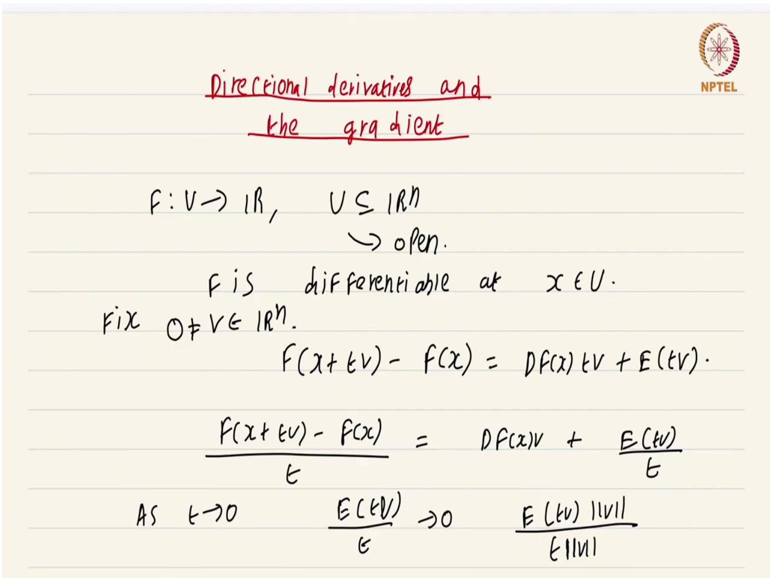
Real Analysis II
Prof. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

Lecture - 10.3
Directional Derivatives and the Gradient

(Refer Slide Time: 00:22)



In this video, we are going to learn how to compute the derivative directly by our knowledge of one variable calculus. This will be facilitated by the notion of directional derivatives and partial derivatives. So, as usual the setting is F is from U to R, U is subset of R n open. We are in this setting.

Suppose, we know that F is differentiable, F is differentiable, at this particular point x in U, ok. Now, fix a vector, fix V in R n and fix this vector to be a nonzero vector, fix a nonzero vector. Then, we know that we can do F of x plus tv minus F of x and this will be equal to DF x tv plus E of tv.

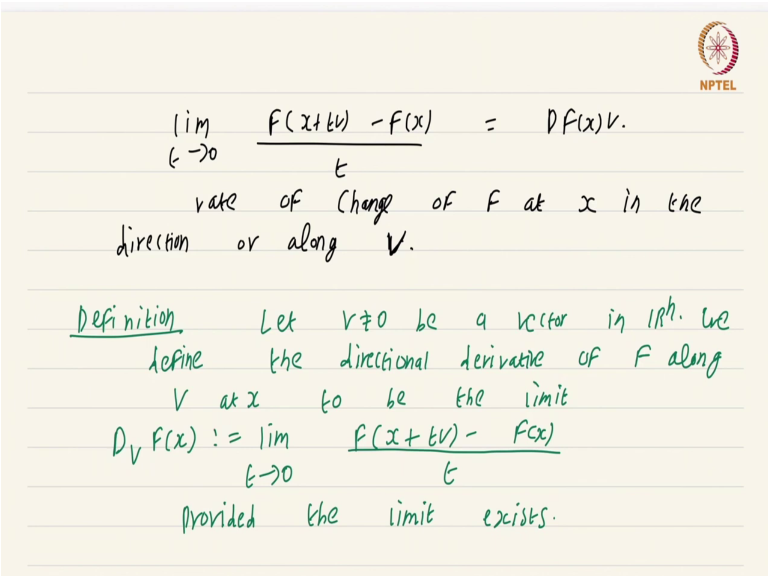
This equation will be valid whenever t greater than 0 is suitably small, ok. In fact, this will also work when t is less than 0 and is suitably small, ok. All you are saying is that the vector tv is sufficiently small that x plus tv is in the domain of both the x plus tv is in the domain of U, ok.

Now, we divide by t on both sides. So, what you get is F of x plus t v minus F of x by t is equal to DF x V plus E of tv by t. Note, I have gotten rid of that t this t I have gotten rid of that is because DF x tv is just D t times DF x v because DF x is a linear transformation, it is a linear functional.

Now, observe that as t goes to 0, as t goes to 0 E of t v, E of t v by t goes to 0. Well, to see this all you have to do is you will just have to multiply and divide by norm v, by norm v, ok. So, as t goes to 0, this will go to 0, of course, you might object that there should actually be a modulus of t here, but that really does not matter, that is just going to change the sign.

And as we are claiming that this term goes to 0, it really does not matter whether there is a plus t or a minus t in the denominator, ok. So, writing E of t v by t as E of t v norm v divided by t norm v, we see that this term E of t v by t goes to 0 as t goes to 0. What is the net upshot? Well, the net upshot is the following.

(Refer Slide Time: 03:28)



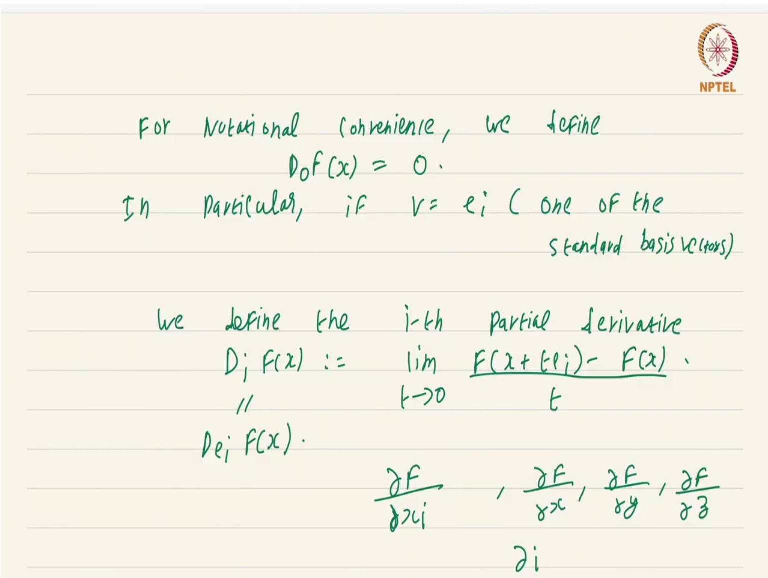
It is that limit, t going to 0, F of x plus p v minus F of x by t, this is equal to DF x v, ok. Now, intuitively what this expression F of x plus tv minus F of x by t limit t going to 0, what this set gives? This gives the rate of change, so this is the rate of change, rate of change of F at x in the direction in the direction or along V, ok. So, essentially what we are doing is we are just considering a line passing through this point x along the direction V that is what this x plus tv is supposed to denote. And we are restricting F to that line.

So, F in some sense becomes a one variable function and we are differentiating F in the usual sense that is what this is doing. And the net thing that you get limit t going to 0, we will measure the rate of change in that particular direction.

The new thing that we have got is the fact that if F is differentiable at the point x, this is in fact, equal to DF x v. So, the derivative linear functional DF x allows you to compute the rate of change in each direction, ok. So, we formally make this into a definition now, definition.

So, let V not equal to 0, be a vector in R n, vector in R n. We define the directional derivative of F along V at x to be the limit t going to 0 F of x plus t V minus F of x divided by t. Of course, provided the limit exists, provided the limit exists, ok. So, as I have said we will consider V not equal to 0, ok.

(Refer Slide Time: 06:29)



Now, for what do you say; for notational convenience; by the way I did not introduce the notation this directional derivative is denoted by D V of F of x, ok. Directional derivative of F at x along V is denoted dv F of x, ok. So, for notational convenience we define D 0 F of x to be just 0. So, the directional derivative along 0 is defined to be 0 just by convention, just for it is a formal definition, ok.

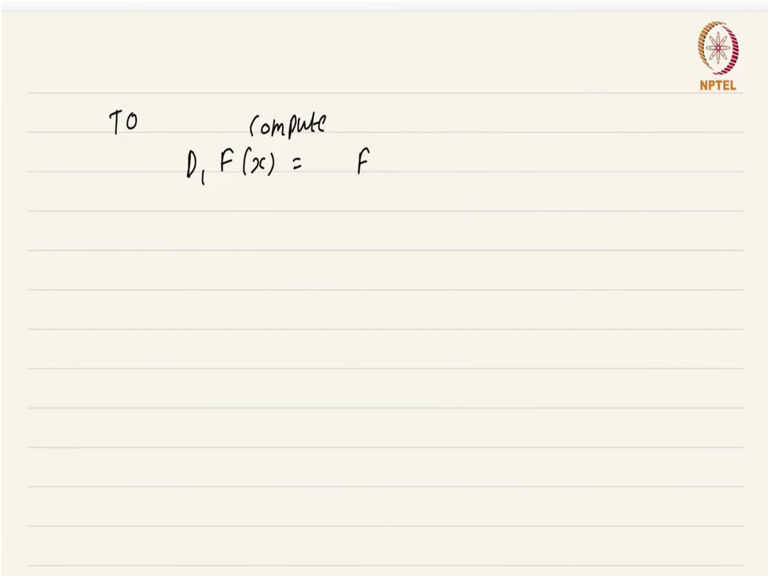
In particular, if V is e i that is one of the standard basis vectors, one of the standard basis vectors, we define the i-th partial derivative D i, we do not put the e i for convenience D i of F of x is just as you can guess, F of limit at t going to 0, F of x plus t e i minus F of x by t, ok. So, the i-th partial derivative is actually nothing, but D e i of F of x, ok. So, this is a specific case of the directional derivative, ok.

Now, I am going to assume that you have already studied basic multivariable calculus in your first year of BSc or whichever course BTech or whichever course you are taking. I am going to assume that you have already studied taking partial derivatives and directional derivatives.

Now, other commonly used notations for the partial derivatives include things like del F by del x i, ok or in the case of just 3 variables del F by del x or del F by del y, del F by del z, I am sure you have seen such notations and also del i. These are various notations that are used by various authors to denote the partial derivatives.

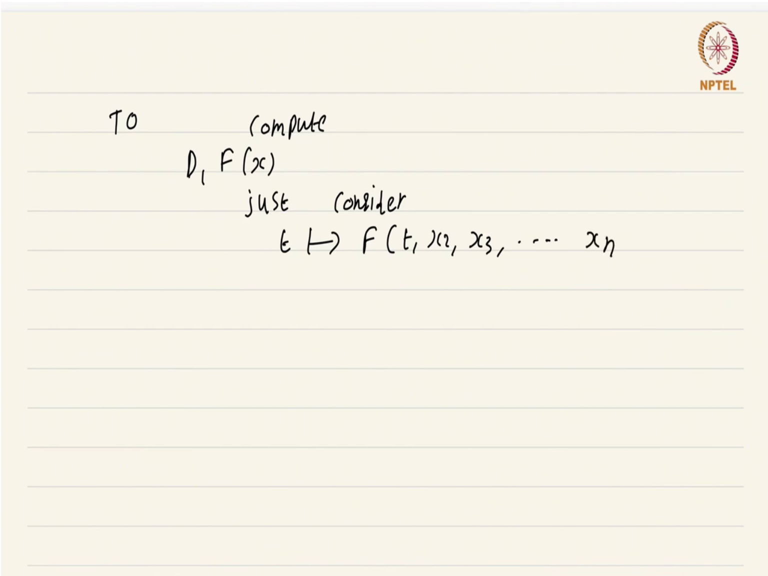
Now, the partial derivatives are almost exactly the same concept as one variable differentiation. To compute the partial derivative at the point x 1 to x n, you just fix all the variables.

(Refer Slide Time: 09:40)



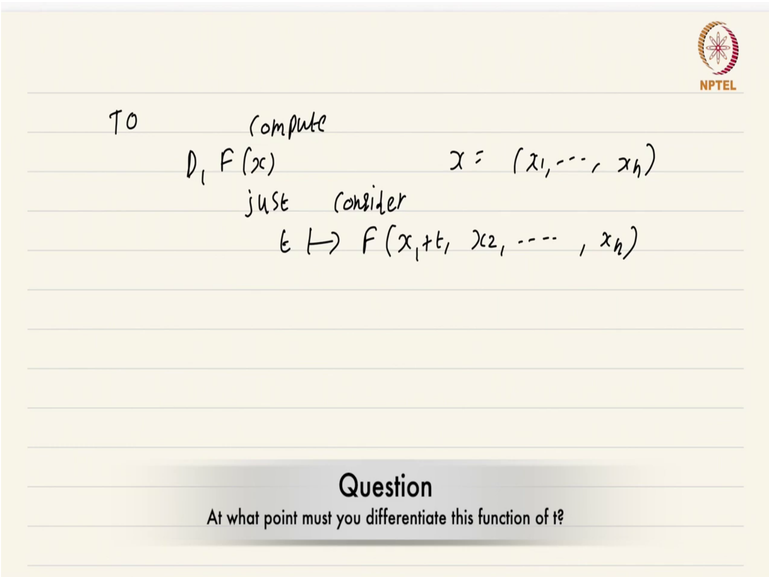
So, suppose you want to compute let us say D 1 of F at the point x, what you do is you just consider the map, you just consider F of. So, let me just erase this.

(Refer Slide Time: 10:07)



To compute this just consider t going to F of t, x 2, x 3, dot dot dot x n.

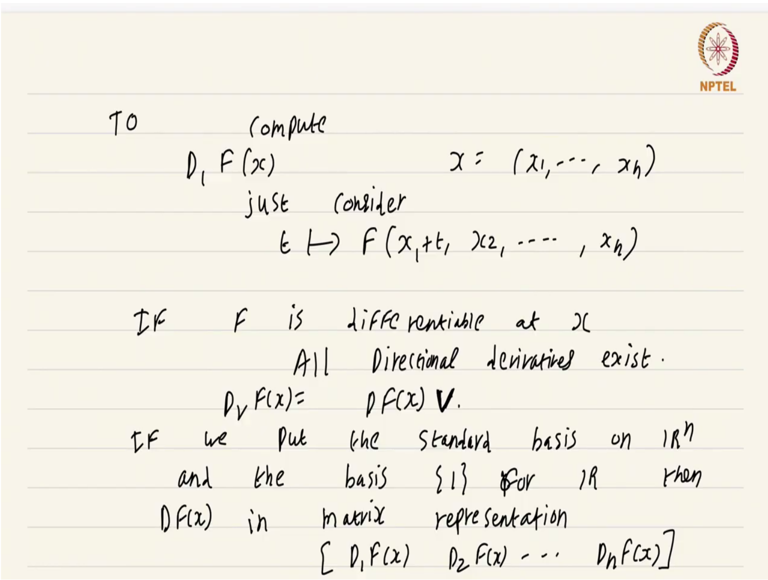
(Refer Slide Time: 10:31)



Actually, it should not be considered F of x 1 plus t consider the map F of x 1 plus t x 2 dot dot x n, where this x is equal to x 1 to x n. So, fix the values x 1 to x n and substitute it substitute x 1 plus t dot dot dot x n into the expression, you will get a function of t, right, you will get a function of t. So, this function of t you just differentiate it with respect to t, just the way you would do if you are given a one variable function, ok.

So, the derivative in this scenario is exactly the same as one variable differentiation. You just treat all the other variables as constant. Just the variable that you are interested in just plug in x 1 plus t get an expression of t and differentiate as usual, ok.

(Refer Slide Time: 11:37)



Now, what we have concluded is that if F is differentiable at x; is differentiable at x all directional derivatives exist, all directional derivatives exist that was the motivation before I defined the directional derivative. And in fact, we have D V of F x is just DF x acting on the vector V, ok.

Now, the LHS as you notice is a number and the RHS is a linear functional acting on V that will also give a number. So, this actually makes sense. Now, suppose we know that the RHS always exists that is directional derivatives we are not assuming sorry about that.

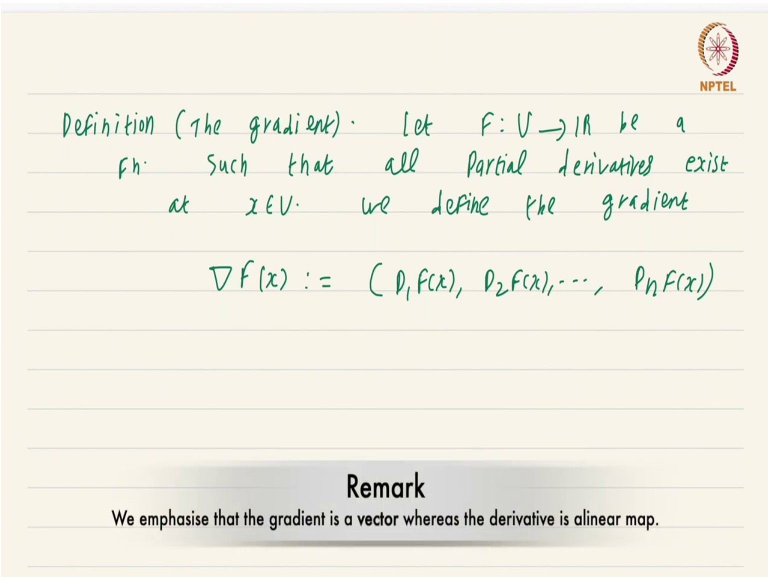
Suppose, we assume that the LHS, suppose we assume that the LHS always exists that is directional derivatives along all directions, suppose we assume that it exists, then does it mean that the RHS DF x which is what we are interested in does it mean that it exists? Well, let us think about this question for a moment.

Notice that the right hand side allows you to give an expression for DF x. If we put the standard basis, the standard basis on R n and the basis and the basis just set with 1, this is just one element.

Just set with one as a basis for R, then DF x in matrix representation in matrix representation, how do you find out the matrix representation of a linear map? Well, you act the linear map on the basis vectors and then just put the elements the what the output vectors write down in the basis for the codomain and put them as the columns. So, immediately you will see that the matrix representation is just D 1 F x, D 2 F x dot dot dot D n F x. This is a row matrix consisting of the partial derivatives.

So, what this is saying is if F is differentiable we have a nice expression D 1 F x, D 2 F x, D n F x, ok. So, this prompts the definition of the gradient which is one of the central topics of this video.

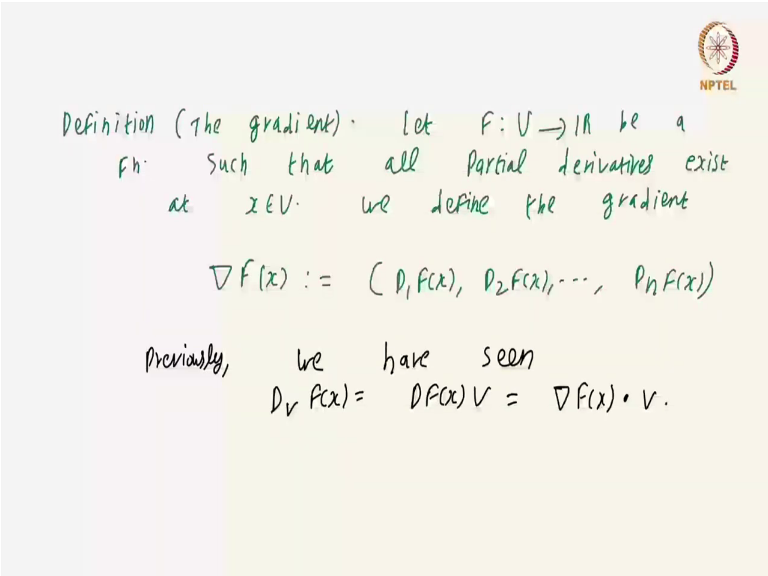
(Refer Slide Time: 14:34)



So, definition; this is the definition of the gradient. Let F from U to R be a function such that all partial derivatives, all partial derivatives exist at x in U. We define the gradient. Nabla, this is read nabla or just grad F x is by definition equal to this vector D 1 F x comma D 2 F x comma dot dot dot D n F x, ok.

Now, rephrasing this expression in this slide that D V F x is DF x V, we can write that in terms of the gradient as well.

(Refer Slide Time: 15:58)

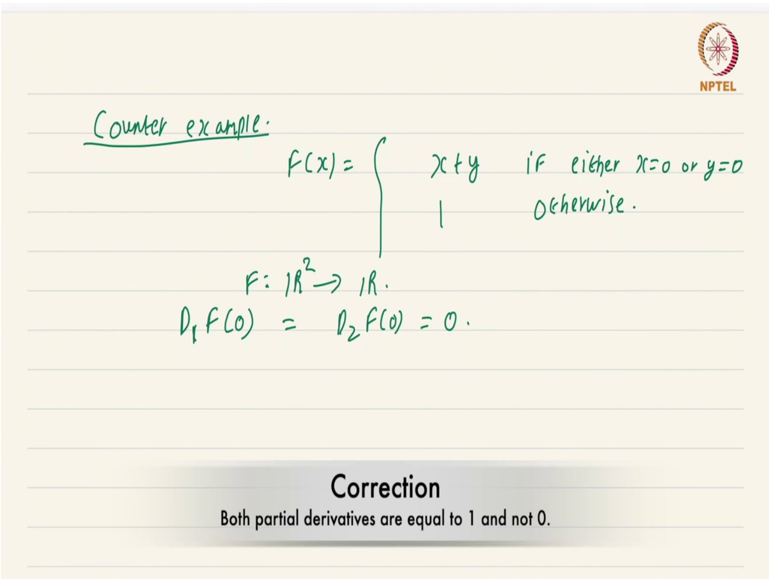


So, previously we have seen previously we have seen; we have seen that D V F x is nothing but DF x v, but this expression that we have seen DF x in its matrix form is just going to be the row matrix D 1 F x dot, dot, dot D n F x.

So, this is same as the gradient of F x dot product, the standard dot product, the standard inner product on R n dot V, ok. So, we have several ways of writing down the directional derivative along V of the function F at x provided F is differentiable.

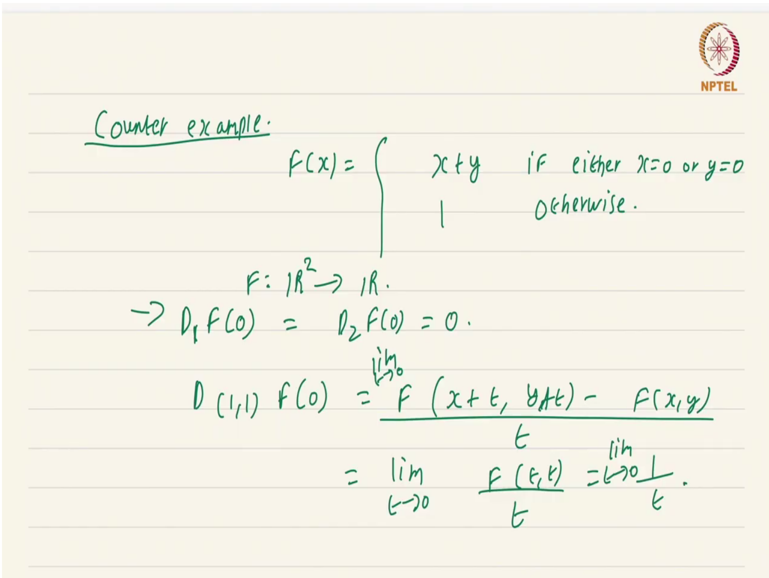
Now, the natural question arises. Suppose, I know suppose I know the partial derivatives exist at the point x, then does that mean that this D 1 F x, D 2 F x dot dot dot D n F x, that row matrix, does that mean that this row matrix is nothing, but the derivative of F at x? No, we have a counter example, ok.

(Refer Slide Time: 17:17)



Consider, F of x to be equal to x plus y if either x equal to 0 or y equal to 0, ok and 1 otherwise. So, this is a rather simple function of two variables. So, this is a function F from R 2 to R, ok.

(Refer Slide Time: 18:21)



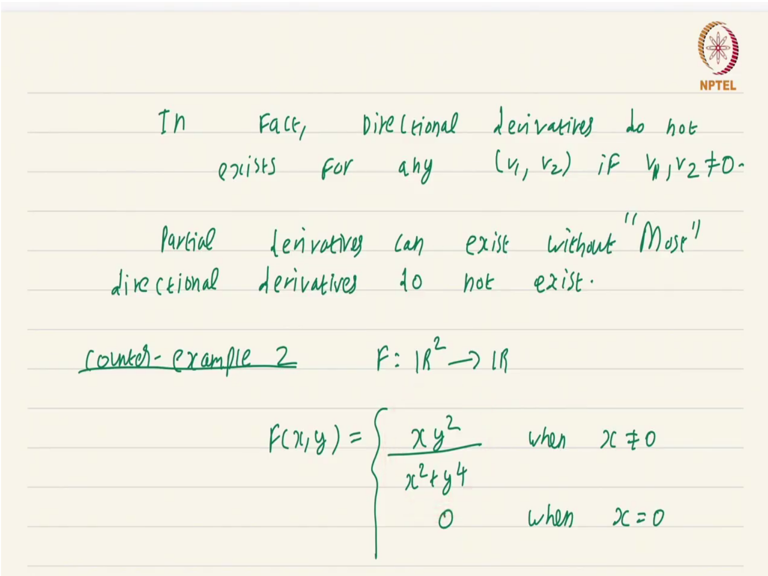
Now, you can easily compute that D 1 F of 0 equal to D 2 F of 0 is equal to 0. Why is this the case? Because you will fix the second variable in the first case, second variable to be 0, so y is going to be 0, so the function F is 0. Now, when you differentiate it you are just going to differentiate the constant function 0, so you will get 0.

Now, let us compute D 1 comma 1 of F at 0. This just means taking the directional derivative along the vector 1 comma 1, ok. So, by definition this is just going to be limit F of x plus t 1 comma 1 or rather F of x, y sorry this is going to be F of x plus t comma y plus t minus F of x, y by t, right. Why x plus t comma y plus t? Because the vector v we have taken is 1 comma 1.

Of course, the point we have taken is 0, so this just is nothing, but I have to write limit t going to 0 of course. So, this is limit t going to 0 of F of t, t by t because F of 0, 0, when you substitute 0 comma 0 for x, y is 0. This is just 1 by t right. So, the limit does not even exist, of course, limit p going to 0.

So, when you take the directional derivative along the direction 1 comma 1, then the directional derivative does not exist in this direction at x comma y equal to 0 comma 0, ok.

(Refer Slide Time: 20:16)



In fact, the same argument shows, in fact, directional derivatives do not exist for any v 1 comma v 2 if both v 1 and v 2 are nonzero, if v 1 comma v 2 are both not 0. Then, that directional derivative along that vector is non-existent the same argument will work.

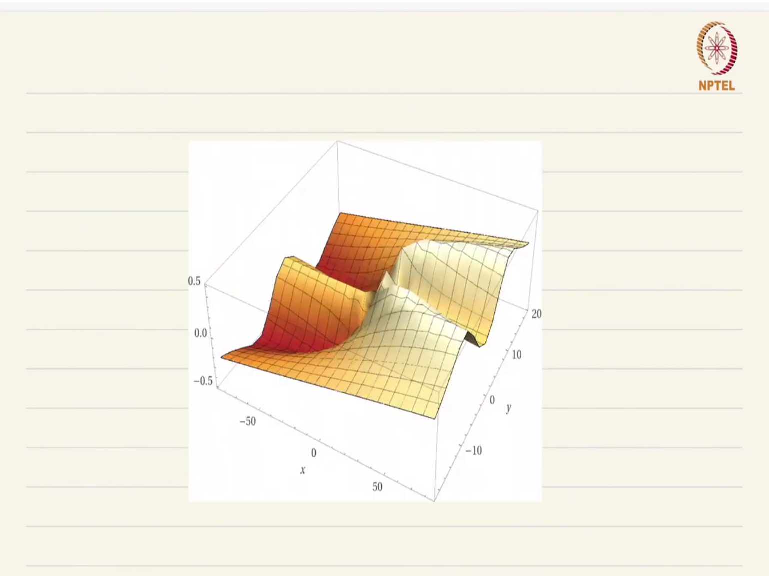
So, what this shows is that partial derivatives can exist, derivatives can exist without most directional derivatives existed, ok. But from what we have seen if the function F is differentiable then all directional derivatives exist, right.

So, just the mere existence of partial derivatives do not guarantee that the gradient is going to give you the derivative by taking a dot product, ok. You might think, ok this is sad, but why do you assume only partial derivatives exist.

So, counter example 2 is going to say even if you assume all partial derivatives exist, it need not be the case that F is differentiable. So, again consider F from R 2 to R given by F of x, y is equal to x, y squared divided by x squared plus y power 4, this is when x is not equal to 0, when x is not equal to 0, and 0, when x equal to 0, ok.

Now, this function looks rather complicated. So, I have just plotted this on wolfram alpha just the part x, y squared plus x squared plus y power 4, I asked wolfram alpha to plot it and it gave me something like this.

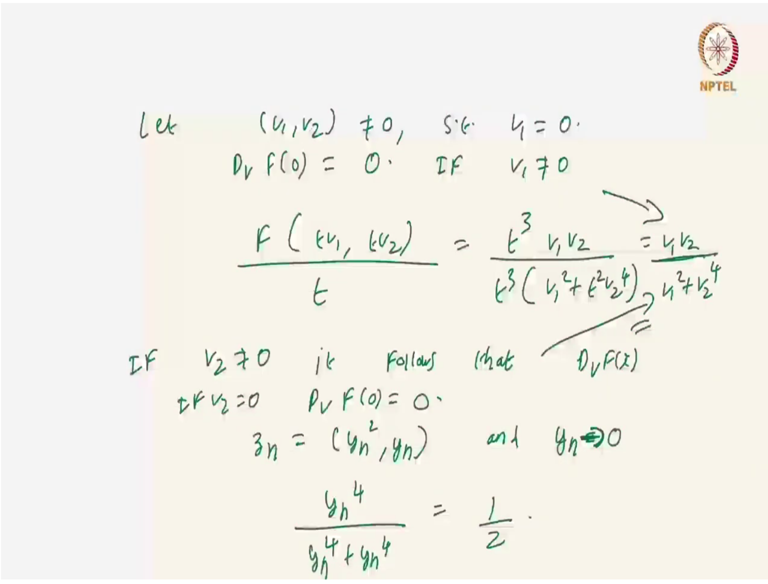
(Refer Slide Time: 22:51)



As you can see from this picture, there seems to be some weirdness happening at the origin. It seems to be taking different values at the origin depending on which direction you approach from.

So, this picture sort of suggests that F is not even continuous at 0, ok. So, that is what we are going to be showing now. We are going to show that F is not continuous at 0, but nevertheless all directional derivatives exist. So, F cannot be differentiable because a differentiable function is always continuous, ok. So, let us just first prove that all directional derivatives exist at the origin.

(Refer Slide Time: 23:31)



So, let v 1 comma v 2 be a vector such that. So, let this not be 0 such that v 1 equal to 0, such that v 1 equal to 0, ok. Now, you can immediately see that in this scenario D V F of 0 is nothing but 0, ok. Now, if v 1 is not 0, if v 1 is not 0, we can compute F of tv 1 comma tv 2 by t, this is what we want, this is the this is going to give you the expression for the derivative, directional derivative at the origin in the direction v 1, v 2.

When you substitute you will get t cube v 1 v 2 divided by t cube into v 1 squared plus t squared v 2 power 4 which is nothing but v 1, v 2 divided by v 1 squared plus v 2 power 4, ok. So, this computation is valid whenever v 1 is not 0. In fact, it is valid throughout, but the interesting case is when v 1 is not equal to 0.

If v 2 is not equal to 0, if v 2 it is not equal to 0, it follows that this expression above if this expression above gives you a D V F x, follows that dv F x is just this, right. And of course, if v 2 is 0, just like the case before D V F at 0 is just 0. So. In fact, I could have just not complicated by taking v 1 not equal to 0, v 2 not equal to 0 and all that first wrote down this expression and say that this expression is valid if both v 1 and v 2 are nonzero or if at least one of v 1 or v 2 is not 0, ok.

So, the net upshot is irrespective of my clumsiness. The net upshot is the directional derivatives at the origin exist. But F is not even continuous at 0. How do you see that? Well, consider the sequence z n to be equal to y n squared comma y n and y n going to 0.

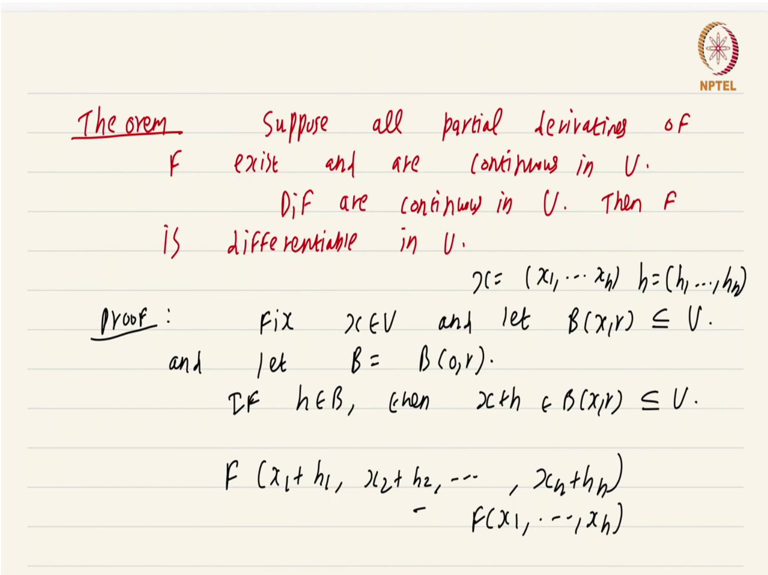
So, you are approaching, you are approaching the origin via this parabola y n squared comma y n, ok. Substituting in that expression for F we will get y n power 4 divided by y n power 4 plus y n power 4 which is just half, ok. So, when you approach along this parabola y n squared comma y n you get half, but F of 0 comma 0 is equal to 0 because we had defined F to be 0 whenever x is 0, ok.

So, this shows that F is not continuous at the origin and therefore, F cannot be differentiable at the point x, at the point 0, ok. So, this shows that the directional derivatives existing, all directional derivatives existing does not guarantee the existence of the derivative, ok. Now, the reason for this is the following. All we know, from the existence of the directional derivative is that along each direction the function behaves reasonably well because the derivative exists.

However, the definition of differentiability requires the existence of a good linear approximation irrespective of the direction, ok. So, this linear approximation in some sense does not depend on the particular direction of approach. You have sort of a uniform expression that allows you to approximate the function F. So, intuitively the notion of the derivative existing is much stronger than good behavior in each direction. And this is captured in this example.

Nevertheless, a minor twist can make the existence of the derivative guaranteed. What is that minor twist? You just have to assume that the partial derivatives not only exist, but they are continuous. Then, it will turn out that F is differentiable, that is the content of the next theorem, ok.

(Refer Slide Time: 28:24)



Suppose, all partial derivatives, all partial derivatives of F exist and are continuous in U, ok. So, we are assuming all partial derivatives of F x exist at all points of U and the partial derivatives that is these D i F they are all continuous, continuous in U. Note, D i F would also be a function from U to R, taking partial derivatives gives you back a function from U to R, then F is differentiable in U, F is differentiable in U, ok.

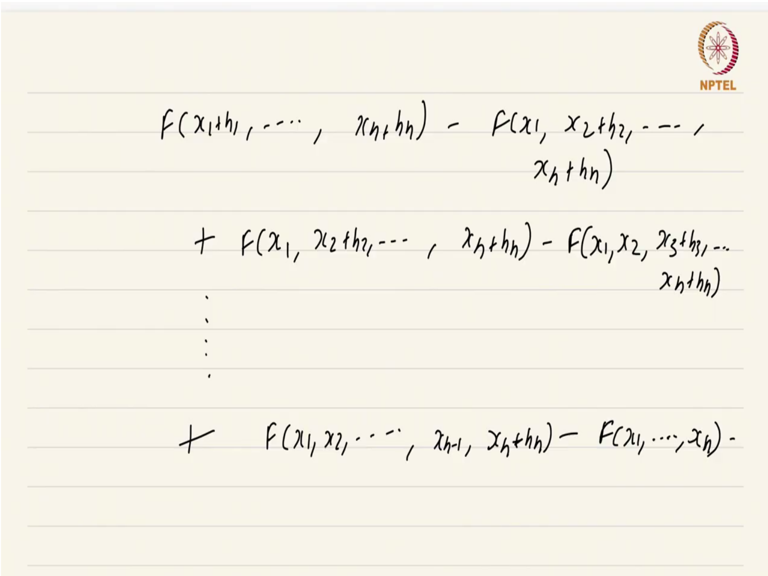
The proof is rather nice it involves a clever telescoping type thing proof. So, fix x in U and let B x, r; B fully contained in U, choose an open ball that is fully contained in U, ok. Now, this ball I am just going to call it B, ok. Just a moment. I made a slight error. So, let B x, r be in U and let B be equal to B 0 r not B x, r.

You will understand in a moment why I am making this change, ok. Now, if h is in B then x plus h is an element of B x, r which is contained in U that is the reason why I defined this B this way, so that I get I am guaranteed that x plus h is an element of B x, r which is in fact, a subset of U. So, x plus h is an element of u.

So, what do we want to investigate? We want to investigate F of x 1 plus h 1, x 2 plus h 2 comma dot dot x n plus h n and approximate it using some linear map and of course, I have to subtract F of x 1 to x n, I have to approximate it by a linear map. Of course, x is x 1 to x n, and h is h 1 to h n, ok. Our goal is to get a nice approximation for this difference, ok.

Now, here is the trick. I am going to write this difference as a telescoping sum. So, that is why you would need actually that F is partial derivatives of F exists throughout U. In fact, you could have just assumed exist in some open set that contains x and get a local conclusion that F is differentiable at x, but I prefer to make this more what do you say simpler statement, but less general, ok.

(Refer Slide Time: 32:01)

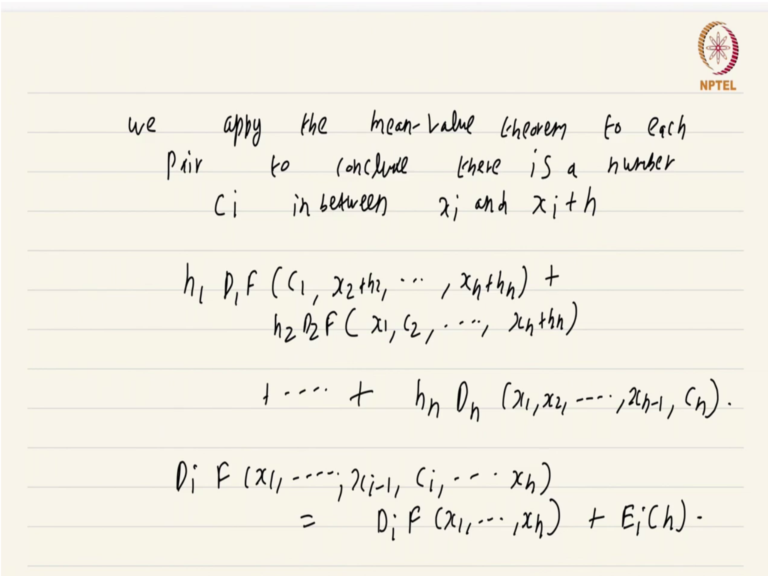


So, what you do is you write this as F of x 1 plus h 1 comma x n plus h n minus F of x 1 comma x 2 plus h 2 comma dot dot dot x n plus h n, ok. You have subtracted something, so you will have to add back something. So, you add back the same thing, F of x 1 comma F of x 1 comma x 2 plus h 2 comma dot dot dot x n plus h n, ok. Now, subtract F of x 1 x 2 comma x 3 plus h 3 comma dot dot dot x n plus h n.

So, you have subtracted one term, now you have to add it back again, so you keep going like this. Finally, you will get F of x 1 x 2 dot dot dot x n minus 1 comma x n plus h n minus F of x 1 to x n, ok. So, we have just, there is just going to be cancellation. So, this term gets cancelled with this term, this term will get cancelled with whatever term was supposed to be there within these dots. So, finally, you will be left only with this minus this, ok.

Now, what we are going to do is we apply the mean value theorem to each such pair, ok. We apply the mean value theorem to each such pair, observing that in this difference only one variable is changing. So, we can apply the mean value theorem.

(Refer Slide Time: 33:55)



We apply the mean value theorem, we apply the mean value theorem, value theorem to each pair to each pair to conclude there is a point there is a point or rather there is a number, there is a number C i in between x i and x i plus h.

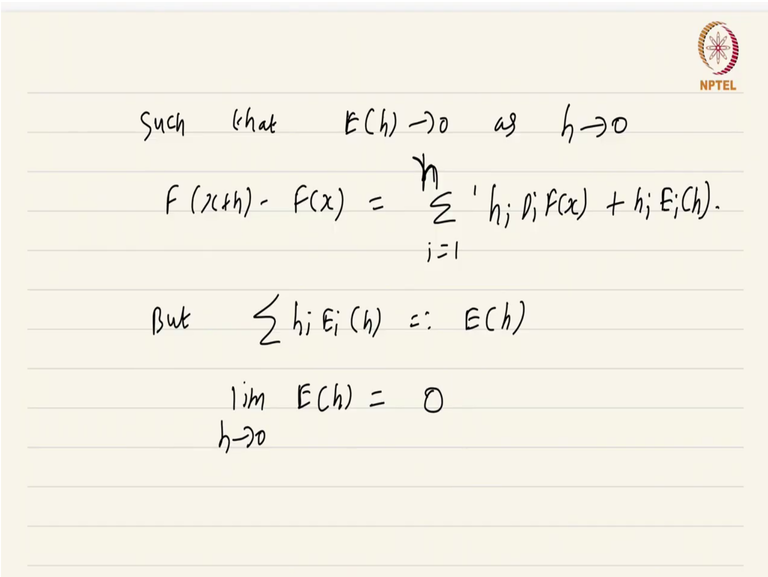
There is a number C i in between x i and x i plus h, such that this telescoping sum just becomes h 1 D 1 F of C 1 comma x 2 plus h 2 comma dot dot dot x n plus h n, plus h 2 D 2 F x 1 comma C 2 comma dot dot dot x n plus h n, plus dot dot dot finally, h n D n of x 1 x 2 dot dot dot x n minus 1 comma C n, ok.

I have applied the mean value theorem to each pair of equations here, each pair here and found out a C i in between x i and x i plus h i, such that I am applying the mean value theorem to that particular variable. So, you get h 1 D 1 of F, C 1 comma x 2 plus h 2 dot dot dot x n plus h n, and finally, the last term h n D n x 1 comma x 2 comma dot dot x n minus 1 C n, ok.

Now, what we want to do is we want to show that the derivative of F at x exists, so we need to construct an error term. So, what we do is because each one of these partial derivatives exists, each one of these partial derivatives exists, we know that D i of F of x 1 comma dot dot dot x i minus 1 comma C i comma dot dot dot x n each one of these partial derivatives I can just write this as D 1 of F sorry D i of F of x 1 dot dot dot x n plus E i h some error term, ok.

Now, why does this follow? Well, this you can always write down an error term like this.

(Refer Slide Time: 37:25)



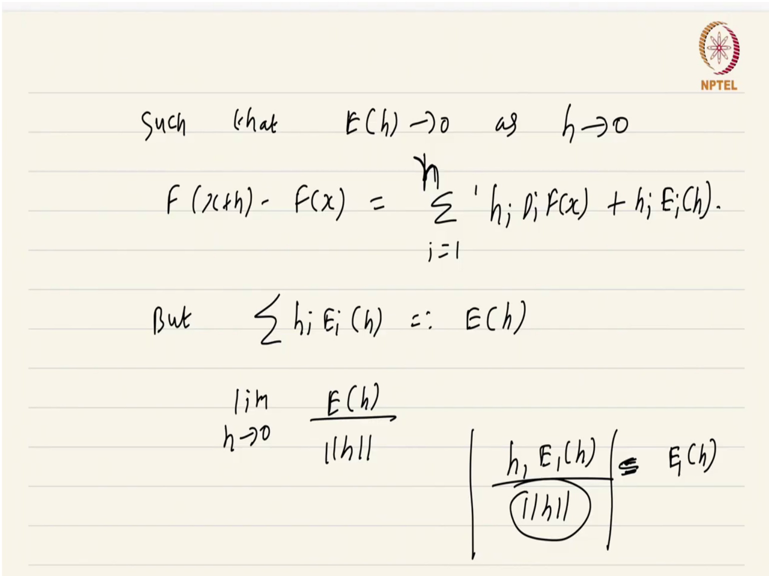
But what is key is that, if such that E of h goes to 0 as h goes to 0, ok. So, let us look at this expression once more. What I am saying is each such partial derivative D i F of x 1 comma dot dot dot x i minus 1 comma C i comma dot dot dot x n is equal to the partial derivative at the point x 1 to x n plus this error term E i of h and this error term goes to 0 as h goes to 0.

Well, this just follows because partial derivatives are assumed to be continuous in the hypothesis. We have assumed that each partial derivative is continuous. So, there will be this error term and this error term will go to 0 as h goes to 0. So, net upshot of all this is F of x plus h minus F of x is nothing but summation i equals 1 to n, let me write the n properly 1 to n, h i D i of F x plus plus h i E i of h, ok.

This just follows from this expression here each one of these partial derivatives I have this further expression, I just substitute. So, I get F of x plus h minus F of x is equal to summation i equals 1 to n, h i D i F x plus h i E i h, ok. But, call this summation h i E i h to be just E of h, define this to be just E of h, limit h going to 0 E of h is equal to 0, right.

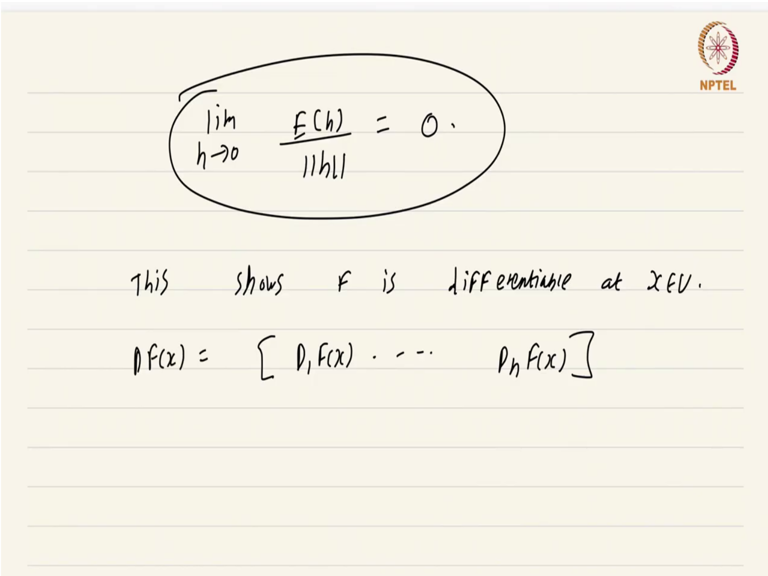
In fact, limit, in fact, this is not enough to guarantee, is not enough to guarantee. What we want? We want limit we want to consider limit h going to 0, E i E of h by norm h, right.

(Refer Slide Time: 39:37)



So, let us just take one term, let us just take E 1 of h 1 E 1 of h by norm h, ok. But we know that norm h would be greater than h 1, so if you want to take the modulus or rather the norm of this whole thing, no need to take the norm, this is just a real valued thing. Just take the modulus of this, this is going to be equal to, this is going to be equal to or rather less than or equal to E 1 of h because modulus of h 1 is less than or equal to norm h, ok.

(Refer Slide Time: 40:29)



So, this just shows that a limit h going to 0 of E of h by norm h is equal to 0 because each term E 1 h by nor h 1 E 1 h by norm h this one goes to 0, ok. So, we conclude that the error term that we have defined E of h has the property the limit h going to 0 E of h by norm h is equal to 0, ok. This shows F is differentiable at x in U.

Not only does this show that F is differentiable at x in U, we get the matrix representation of the map. It is just D 1 F x comma dot dot dot D n F x. This is the expression for the derivative in terms of the standard ordered basis on R n and the single vector 1 on R. With respect to this pair of basis this is going to be the expression for the derivative.

Why is this going to be the expression for the derivative? Well, because we have written F of x plus h minus F of x as summation i equals 1 to n, h i D i F x plus this error term within this error term is sub linear, ok. So, this concludes the proof that whenever the directional derivatives, sorry, whenever the partial derivatives exist and are continuous throughout an open set U then the function F is in fact, differentiable at U, ok.

So, this allows us to compute derivatives explicitly using one variable calculus, everything gets reduced to doing classical differentiation which we are all experts at this point of time. So, this concludes this video on directional derivatives and the gradient.

In the next video, we will study some basic properties of the gradient and also interpret the gradient as giving the direction of maximum increase of the function. This is a course on Real Analysis, and you have just watched the video on Directional Derivatives and the Gradient.