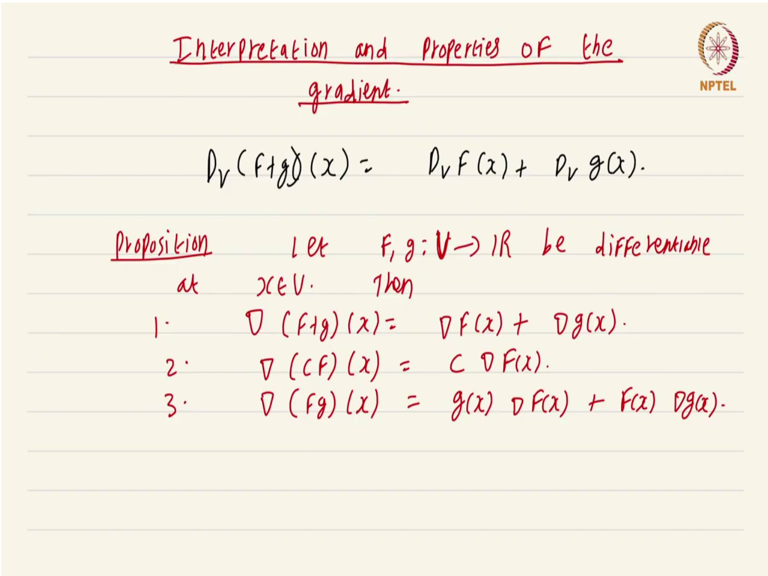
Real Analysis II
Prof. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

Lecture - 11.1
Interpretation and Properties of the Gradient

(Refer Slide Time: 00:23)

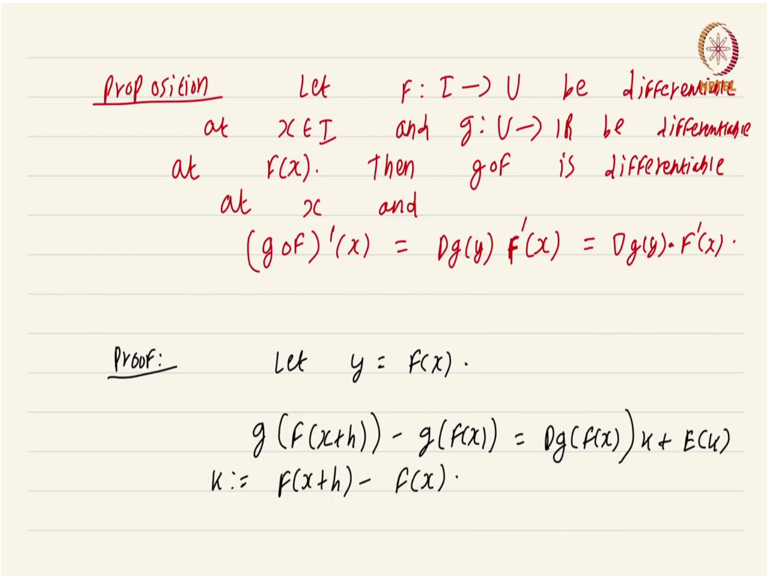


In this video we will discuss the basic properties of the gradient and also interpret the gradient in terms of normal's to hypersurfaces and directions of greatest increase and all that. So, first some basic properties of directional derivatives are rather obvious I am not even going to bother stating them. These are things like D V of F plus g x is equal to D V F x plus D V g x. So, such things are very easy and are obvious, I am going to leave it to you.

So, let me begin by stating a proposition and not proving it because it this is also very easy proposition this involves the gradient. So, let F comma g from U to R; U to R be differentiable at x in U ok. Then we have gradient of F plus g x is nothing but gradient of F at x plus gradient of g at x.

We also have gradient of C F at x is equal to C times gradient of F x and finally, we have gradient of F g at x is equal to g of x gradient of F of x plus F of x gradient of g of x. These are all straightforward properties to prove and I am going to leave them to you to work out the proof ok. We also have a chain rule and this chain rule is somewhat involved. So, I am going to prove the chain rule now.

(Refer Slide Time: 02:35)

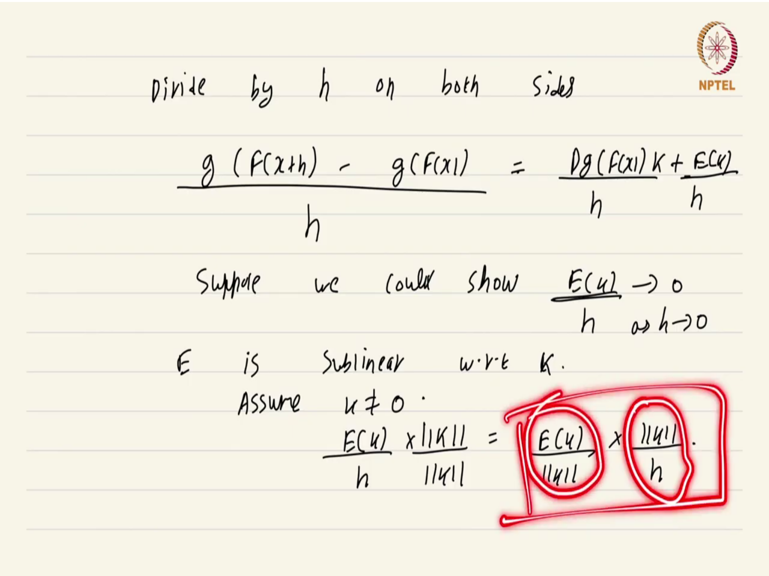


Proposition: let F from I to U be differentiable at x in I. Recall I is an interval in R and obviously, g and g from U to R be differentiable at F of x ok. So, here I am going against the standard convention of having F from U to R. Here we have g from U to R. U subset of R is of course, open.

Then, g composed with F is differentiable at x and g composed with F prime at x is nothing but D g of y, the derivative of the function g at the point y multiplied by this quantity F prime x ok; not multiplied by I am sorry about that acting on the vector F prime at x ok. D g y is a linear functional. So, it eats up a vector and spits out a scalar.

This same expression can be written as gradient of g of y dot product F prime x. So, let me bold the dot ok. So, this is one version of the chain rule ok Let us see a proof. Proof: so, we can of course, write let y be equal to F of x. We can of course, write g of F of x plus h minus g of F of x is just D g F of x times K plus E of K, where K is by definition F of x plus h minus F of x ok.

(Refer Slide Time: 05:35)



Now, what we are going to do is divide by h on both sides on both sides ok. So, what we will get is g of F of x plus h minus g of F of x by h is equal to D g at F of x by h in into acting on K plus K by h not K by h sorry E of K by h E of K divided by h ok.

Suppose we could show; suppose we could show; suppose we could show; suppose we could show that E of K by h by norm h to be more precise actually norm is not needed because h is just a single variable it S just a real number.

Suppose we could show that E of K by h goes to 0 as h goes to 0 then we are done. Because this particular quantity K by h this quantity K by h as h goes to 0 is nothing but F prime at x right because this quantity is nothing but F prime at x. So, this whole thing will just become D g F of x acting on F prime of x and if this goes to 0 then we would be done.

Now, observe that this function E was coming from the definition of the derivative of g. So, E is sublinear ok. So, E is sublinear with respect to K remember that. E is sublinear with respect to K ok.

So, now, so, assume K is not 0 for the time being. You will understand why I am making this assumption. So, we want to do E of K by h. So, what you do is you multiply and divide by norm K multiply and divide by norm K. So, you will get E of K by norm K into norm K by h ok.

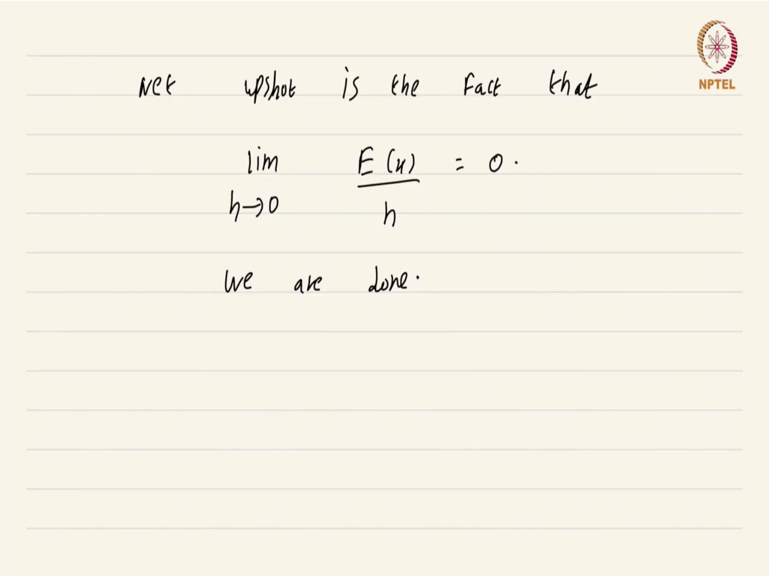
Now, observe I mean I am this is valid because I am assuming K is not 0. This term goes to 0 as h goes to 0. Why is that the case? Well, because as h goes to 0 as h goes to 0 K must also go to 0 because recall that K was F of x plus h minus F of x and F is going to be continuous at x because it is differentiable at x. So, K will go to 0 as h goes to 0. So, E of K by norm K will go to 0 because E is sub linear.

Now, this quantity norm K by h will be bounded ok that is because this quantity K F of x plus h minus x by h that is ultimately going to give you the derivative at x which we know exists. So, this norm h by K will remain bounded in the neighborhood of h equal to 0. So, at least when K is not equal to 0 and if h goes to 0 and K is not equal to 0 then this quantity can be made really small.

Of course if K is equal to 0, if K is equal to 0 we already know that E of 0 is just 0 right. E of 0 is just 0 that is just because E is coming from this definition of the derivative of g at x. Always the error function will take the value 0 at 0 ok.

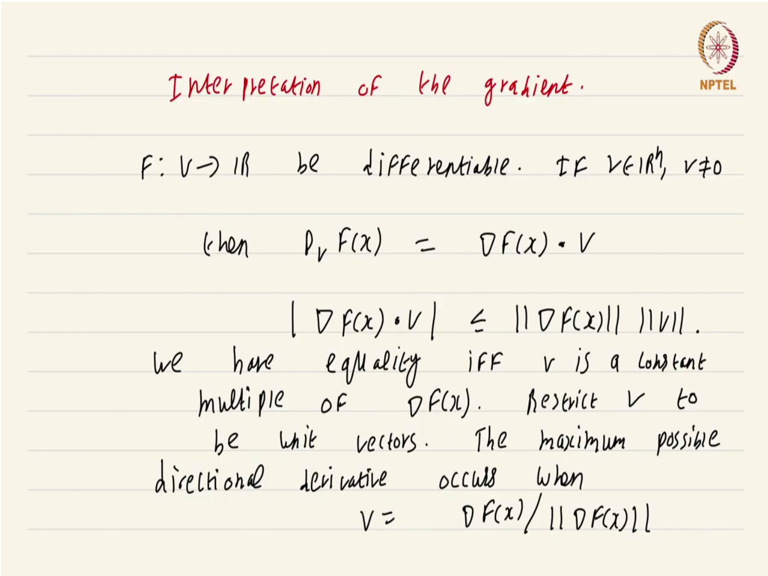
So, since E of 0 is going to be 0, this quantity E of K by h is already 0 when K is equal to 0. So, irrespective of what happens as h goes to 0 even if K is 0 we get 0. If K is not 0 we can make this quantity really small because this quantity is bounded and this quantity is going to 0.

(Refer Slide Time: 10:19)



Net up shot is net up shot is upshot is the fact is the fact that limit h going to 0 E of k by h is equal to 0. In other words we are done. In other words, we are done. So, what happens is as h goes to 0 here as h goes to 0 here this expression here goes to D g F x acting on F prime of x and this quantity goes to 0. So, we are done ok. So, this concludes the proof of the chain rule for the directional not directional the derivative in for scalar valid functions of a vector variable.

(Refer Slide Time: 11:21)



Now, we are going on to interpreting the gradient interpretation of the gradient. So, all of you are familiar with this no doubt from a basic course on multivariable calculus. You would have definitely learnt that you get the gradient measures the direction of maximum increase and so on, let us make this precise.

So, let F from U to R be differentiable ok. Now, we know that if V is in R n V is not 0 then the directional derivative along V at the point x is nothing but the gradient of F at the point x dot product the vector V ok. Now, we already know that the Cauchy Schwarz inequality is going to tell you which applies for more general inner products. So, in specifically it applies to the dot product.

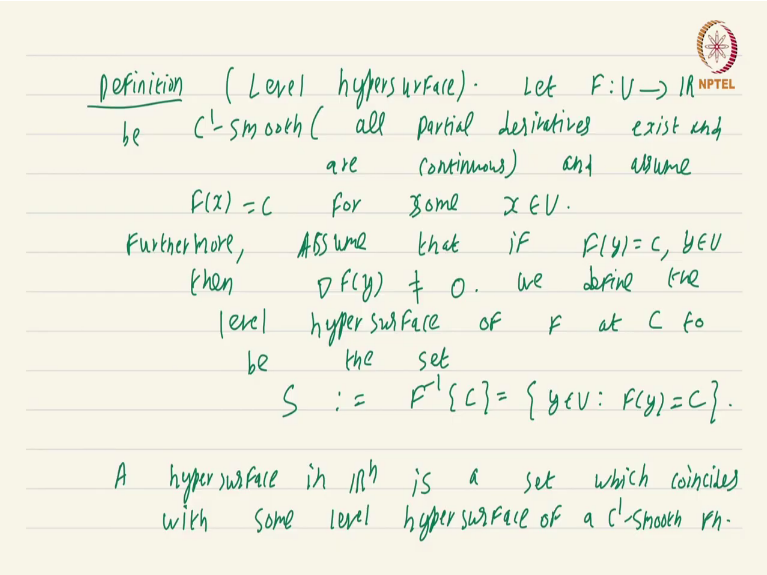
So, modulus of this is less than or equal to the norm of the vector F of x times the norm of the vector V ok. And in fact, we have equality we have equality recall from our study of norm vector spaces. We have equality if and only if and only if V is a constant multiple of this gradient ok. Now, restrict V to be unit vectors to be unit vectors ok.

Once you restrict V to be unit vectors then it is clear that the maximum possible directional derivative maximum possible directional derivative occurs when V is nothing but nabla F x gradient divided by the norm of the gradient. This is going to be a unit vector. I am assuming the gradient is non 0 ok for this to make sense. So, in other words the gradient gives you the direction of the maximum rate of change of the function ok.

So, this is the standard interpretation of the gradient for that actually you require the Cauchy Schwarz inequality to show this. Now, another interpretation of the gradient is that it gives you normals.

So, I am going to make that precise now. For that we need a definition of the hypersurface or the level hypersurface of a function ok. So, we will explore a hypersurfaces more when we come to manifolds in a later chapter, but for the time being I am going to give an ad hoc definition.

(Refer Slide Time: 14:51)



So, definition, this is the definition of a level hypersurface ok. So, let F from U to R be C 1 smooth. What do I mean by C 1 smooth? This just means that all partial derivatives exist and are continuous and are continuous. In particular F will be differentiable in U because of the theorem that we have established earlier that if partial derivatives exist and are continuous then the function is differentiable ok.

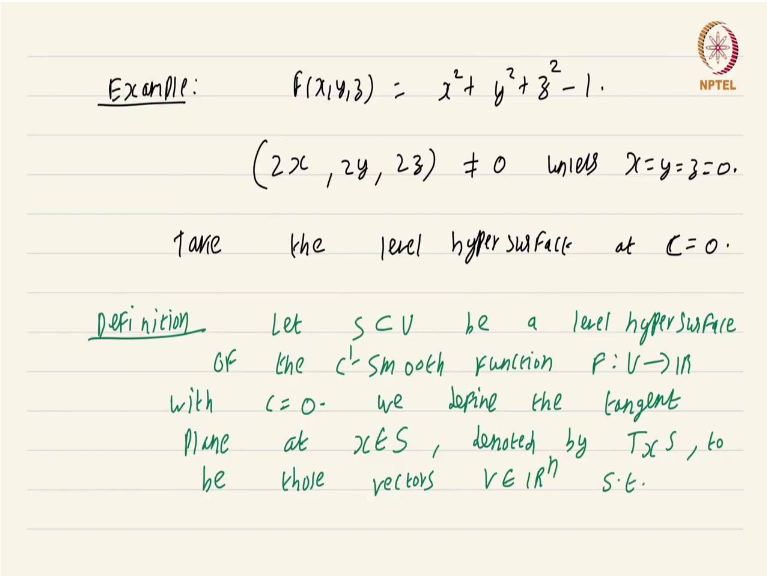
So, let F from U to R be a C 1 smooth function and assume F of x equal to some C, ok. So, for some I mean for x for some x in U. So, you are fixing a point x in U and the value is at the point x is just C, ok. Now, furthermore if assume that, if F of y is equal to C, where y is coming from U then the gradient of F of y the gradient of F at y to be more precise is not 0 ok.

What we are assuming is that for all points in y in U we are calling those points y all points y in U such that F of y equal to C, the gradient should not vanish at that point. So, we are going to assume that the gradient does not vanish at any point at which F takes the value C ok.

In this scenario we define the level hypersurface the level hypersurface the level hypersurface of F at C to be the set; to be the set S by definition equal to the pre image of the point at C. This is just y in U such that F of y equal to C ok. So, a level hypersurface of a function is just the inverse image of this points this value C ok.

A hypersurface in R n is a set which coincides with some level hypersurface. So, it is just a set which is going to be the pre image of some function that is C 1 smooth and such that the gradient does not vanish and so on and so forth. So, some level surface of a C 1 smooth function ok. So, this definition might seem a bit complicated. So, it is good to see at least one example. We will see many more.

(Refer Slide Time: 18:49)



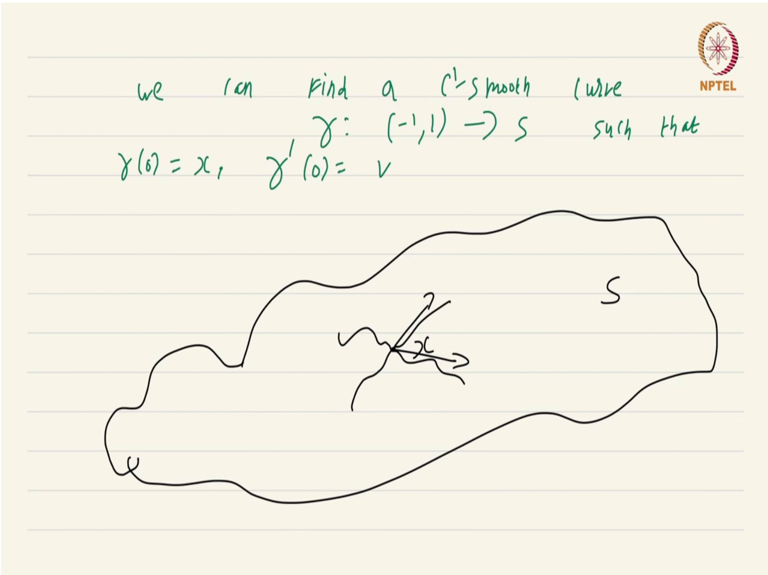
So, in the future when we study manifolds, so, let us just see one example to just get an idea of what is going on. Let us just take F of x comma y comma z to be x squared plus y squared plus z squared minus 1 ok. Now, note that this gradient of this is 2x, 2y, 2 z and the gradient is not 0 unless x equal to y equal to z equal to 0, right.

So, take the level hypersurface at C equal to 0 that will give you exactly the unit sphere that will give you the unit sphere in R 3 and this is going to be a level hypersurface because the gradient does not vanish at any point on the unit sphere ok. Now, one remark I am going to make. The reason why we are assuming that this normal this gradient grad F of y is not 0 is because that is going to give you the normal.

So, a level hypersurface is sort of a figure for which there is a well defined normal at each point. All this will make more sense when we define manifolds, but for the time being we can make this precise in an ad hoc way. So, I am going to define what the tangent plane to a hypersurface is. So, you will have to bear with me for a few more weeks before we precisely define what a manifold is and this will make sense.

Definition: so, let S subset of U be a level hypersurface ok of the C 1 smooth function F from U to R with C equal to 0 ok. I am going to of course, assume that level hypersurfaces are always non empty ok. I am going to assume that there is some point that F takes the value 0. We define the tangent plane the tangent plane at x in S denoted by T x at S to be those vectors V in R n.

(Refer Slide Time: 22:05)



Such that we can find; we can find; we can find a C 1 smooth curve C 1 smooth curve gamma from minus 1, 1 to S such that gamma prime at 0 is equal to V and of course, gamma of 0 is equal to x ok. So, what is this definition saying? You want to define what the tangent plane at the point x on this level hypersurface is. So, what you are doing is you are considering. So, let us draw a crude picture of what is happening to get an idea.

So, we have something like this in space. This is supposed to be our level hypersurface. We are focusing on this point x. We are focusing on this point x and I am saying that to define the tangent plane at this point x just consider curves that pass through this point. There could be many many curves, but the key factors these curves the image of these curves should all lie within the hypersurface they must not move outside the hypersurface ok.

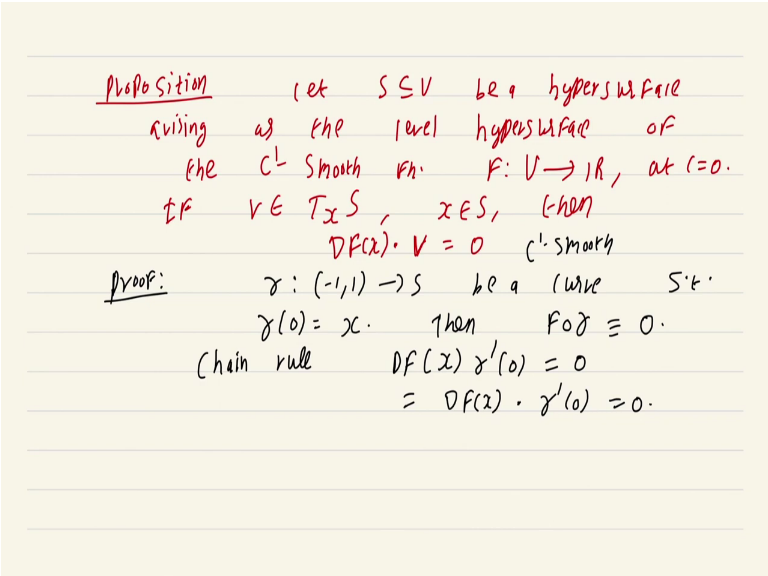
Now, what you do is you look at the various derivatives that is the tangent vectors of these curves ok. Put them all together and call them the tangent plane ok. Now of course, the picture I have drawn is again inaccurate because these vectors that I have drawn here these vectors they are all going to be situated at the origin, but I have made a remark when we discussed these vector valued functions of a scalar variable that is curves that we will usually translate the vector and make it emanate from the point x.

So, we are just considering all such vectors that arise this way. Look at all the possible curves passing through x, look at all the possible tangents of these curves at x put them all together call them the tangent plane. Now, we do not yet have the tools to prove that this tangent plane is a vector subspace of R n.

One reason why we took the vectors gamma prime at 0 and not just x plus gamma prime at 0 is because we want the tangent plane to be a subspace. So, 0 should in particular be an element of this tangent plane.

We do not have the technology yet to prove that this is going to be a subspace of R n what we have defined, but all that will come later. So, for the time being I have done this just to give interpret the gradient as the normal to the hypersurface. All this will make more sense once we study manifolds ok. So, we have done some abstract definitions at least we must gain something out of it.

(Refer Slide Time: 25:01)



We gain this proposition. Proposition: Let S subset of U be a hypersurface be a hypersurface arising as the level hypersurface of the C 1 smooth function; C 1 smooth function F from U to R at C equals 0 at C equals 0. If V is an element of T x at S is of course in S, then gradient of F at x dot product V is 0.

So, in other words the gradient is normal to the tangent plane. So, proof this is looks very complicated, but the proof is even shorter than the statement. Let gamma from minus 1 to 1 to S be a curve C 1 smooth curve of course, C 1 smooth curve such that gamma of 0 is equal to x ok.

Then F composed with gamma is identically 0 because F is going to be 0 on S. S is nothing but the level hypersurface of F at 0 ok. Now, chain rule tells us that D F at gamma of 0 which is just x acting on gamma prime at 0 is just 0. This is just same as gradient of F of x dot product gamma prime of 0 ok.

So, this was a rather easy proof. So, the gradient gives you the normal to the level hypersurface or rather that normal to the tangent plane of the level hypersurface. All this will be more made more precise when we study manifolds.

This is a course on Real Analysis and you have just watched the video on properties of the gradient and interpretation.