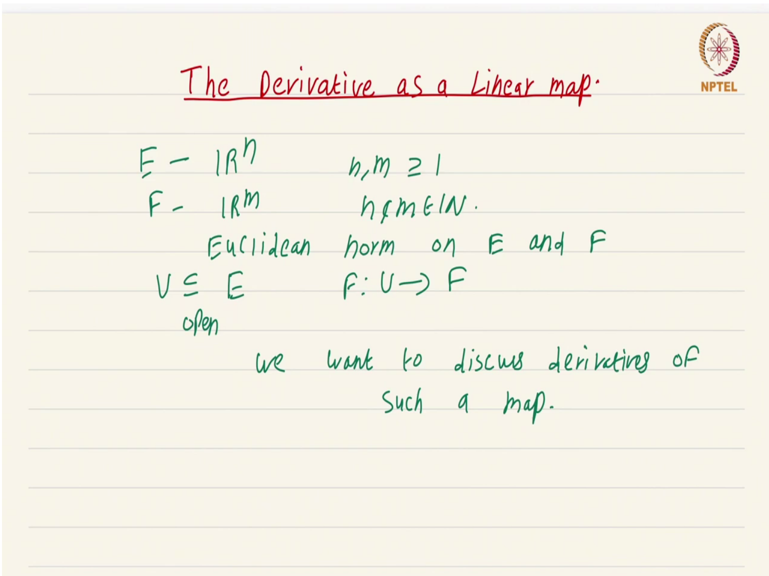
Real Analysis II
Prof. Jaikrishnan J
Department of Mathematics
Indian Institute of Technology, Palakkad

Lecture - 12.1
The Derivative as a Linear Map

(Refer Slide Time: 00:24)

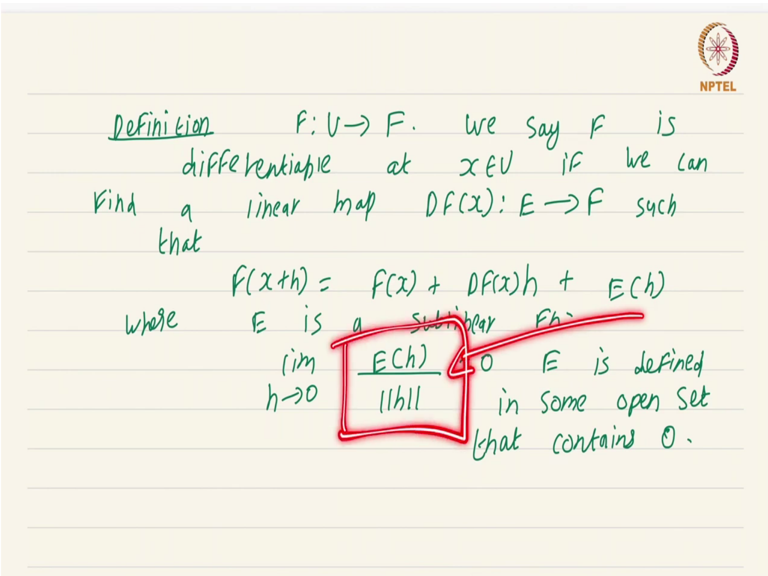


Now, that we have discussed both the cases of a vector valued function of a scalar variable and a scalar valued functions of a vector variable. We are now in good shape to proceed to the most general case where we will consider a mapping between Euclidean spaces R n and R m.

So, we will fix notation once and for all throughout the rest of this part of the course, E will always denote R n, F will denote R m and n comma m are greater than or equal to 1, n comma m are natural numbers ok and we will always put the Euclidean norm; we will always put the Euclidean norm on E and F.

So, our objective is to consider an open set U in E and a map F from U to F and we want to discuss; we want to discuss derivatives of such a map. So, what should the derivative be of a mapping from an open set in Euclidean space taking values in another Euclidean space? We have already done the hard work of treating the derivative of a scalar valued function of a vector variable as a linear functional. So, the definition now that I am about to give will become utterly straightforward.

(Refer Slide Time: 02:07)

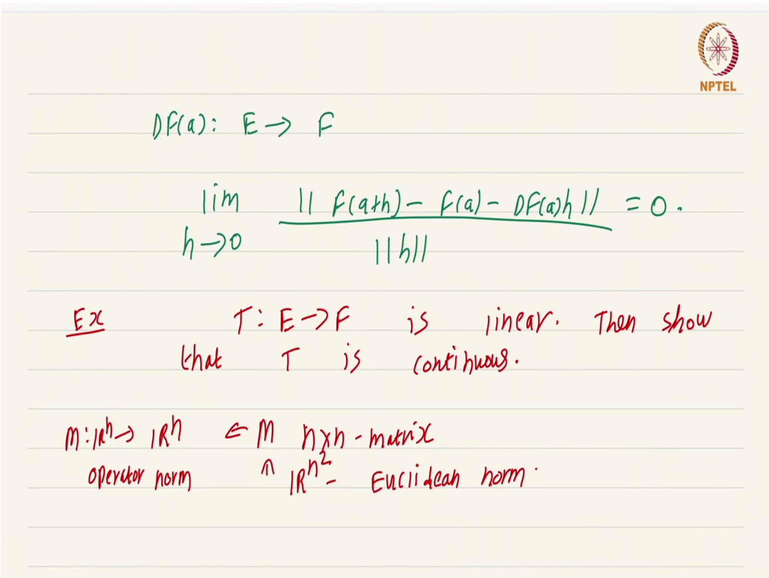


So, this is the definition. So, as usual F is from U to F, we say F is differentiable at the point x in U; at the point x in U if we can find; if we can find a linear map; a linear map, I am going to denote this linear map by DF x, this is a linear map from E to F such that you can approximate F in a nice way using this linear map that is F of x plus h is just F of x plus DF x acting on the vector h plus E of h where; where E is a sub linear function; is a sublinear function.

To be more precise, limit h going to 0 of E of h by norm h is equal to 0, E is defined; E is defined in some open; set some open set that contains 0. So, this definition, the way we have been formulating the definition of the derivative, this definition is exactly the same as what we saw some twenty videos ago when we first introduced the derivative in the context of one variable functions taking values in r.

The exact same definition works, this part is the affine linear good approximation, this E is the error term, the fact that the error term is small is quantified by the fact that limit h going to 0 E of h by norm h equal to 0 ok and similar remarks that we made after the gradient applies, this h has to be chosen so small that x plus h is still in U, we can always do that by shrinking the domain of h to be a suitably small neighborhood of 0.

(Refer Slide Time: 04:45)



So, we can also formulate this definition without involving the error term explicitly by saying that F is differentiable if you can find DF a from E to F such that limit h going to 0; h going to 0 norm of F of a plus h minus F of a minus DF a h divided by norm h is equal to 0. So, you can formulate the same thing without involving the error function explicitly. In fact, you can even delete the norm, this norm in the numerator is actually not needed we even without that the definition is true you can formulate this in alternate ways ok.

So, immediately let us see an exercise. In the next video, I am going to explore certain complicated examples of the derivative but let us see a simple example. Suppose T from E to F is linear; suppose T from E to F is linear. Then show that T is continuous ok. Now, note E and F are not general vector spaces, they are just shortcuts for R n and R m respectively. Any linear map between R n and R m is automatically continuous and the proof is really easy, I want you to show that ok.

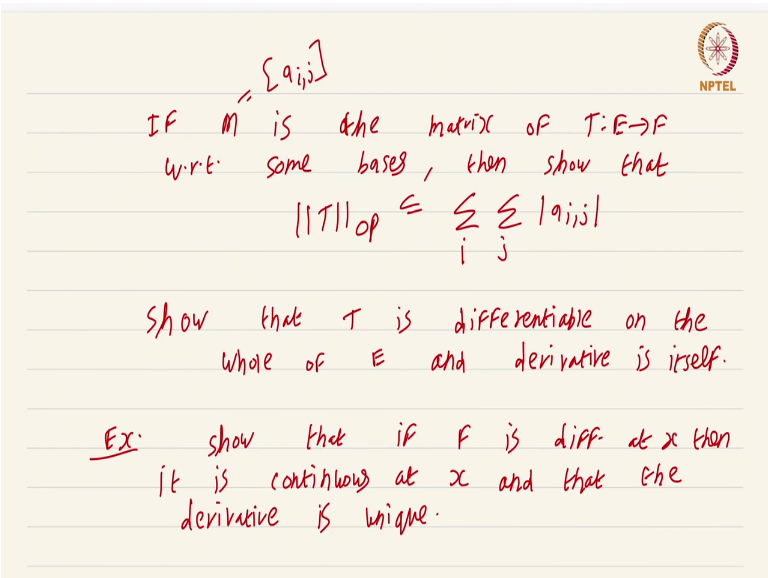
Now, we will always put the operator norm on linear mappings. So, but we will be treating matrices themselves as elements or vectors. So, we will be considering functions defined on matrices and differentiating such functions and I said that whenever we consider E and F vector spaces which are R n and R m, we always put the Euclidean norm.

So, if you have a matrix M; if you have a matrix M, then it is let us say it is an n cross n matrix, if you have an n cross n matrix, then this is actually you can treat it as an element of R n squared in a natural way. Just put the entries of the matrix consecutively as a single vector, you will get an element of R n square and on R n squared, you have the Euclidean norm; you have the Euclidean norm.

But on the matrix M, you can treat M as a map from R n to R n because it is an n cross n matrix and on the matrix M, you automatically have; you automatically have an operator norm in fact, you have this even when you have a matrix which is not a square matrix, a matrix which is a rectangular matrix an n cross m or an m cross n matrix. So, the question arises are these two related?

Well, both norms are going to be equal and that is an exercise that we have already seen when we studied nonlinear spaces. However, you can say something even better or worse depending on your perspective.

(Refer Slide Time: 07:46)

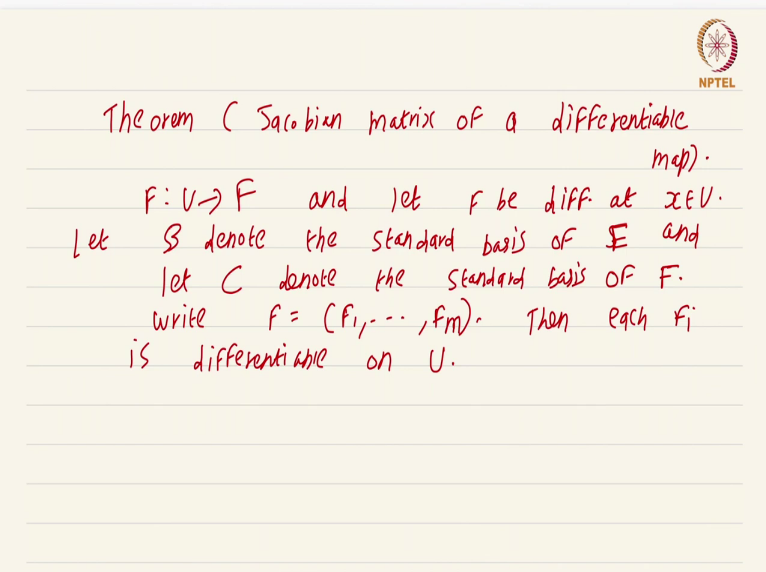


So, this is a part 2 of this exercise. If M is the matrix; is the matrix of T from E to F with respect to some bases; with respect to some bases, then show that the operator norm of T is less than or equal to the summation over i summation over j mod a i, j ok so, where this m is nothing but a i, j. So, suppose you have a linear transformation from E to F and you represent it as a matrix M with respect to some pair of bases, then the operator norm of T is bounded by this quantity ok.

So, this is going to be very used useful in various scenarios ok. Now, finally, show that T is; show that t is differentiable on the whole of E; on the whole of E and derivative is itself. This last part is rather trivial is itself, this will just follow from the way we have formulated the notion of derivative as a best linear approximation ok.

Now, I am going to leave some more facts for you to check, these are easy. Show that if; show that if F is differentiable at x, then it is continuous at x; it is continuous at x and the derivative is unique; and that the derivative is unique. You cannot have two linear mappings that simultaneously satisfy the definition of the derivative of F at the point x ok. So, these proofs go exactly the same as what we did for the derivative of a scalar valued function of a vector variable, exact same proofs will work ok.

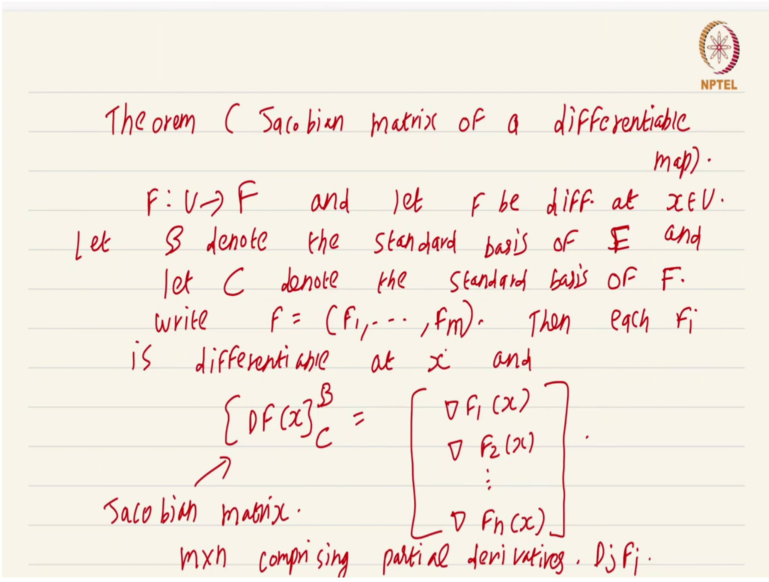
(Refer Slide Time: 10:28)



Now, in the scalar valued function of a vector variable case, we found out that we can represent the derivative using partial derivatives, a similar thing is true even in this general scenario. Theorem I am going to call this the Jacobian matrix of a differentiable map; of a differentiable map ok.

Theorem is as follows as, always F is from U to F ok and let F be differentiable at x in U ok. Let B denote the standard basis; standard basis of E; of E and let C denote the standard basis; standard basis of F ok. Now, write F as F 1 comma dot dot dot F n. You can write F in terms of its coordinates. Then each F i is differentiable on U. What are these F i's? These F i's are nothing but scalar valued functions of a vector variable. These F i's are the various components of the function F each of which would be scalar valued.

(Refer Slide Time: 12:37)



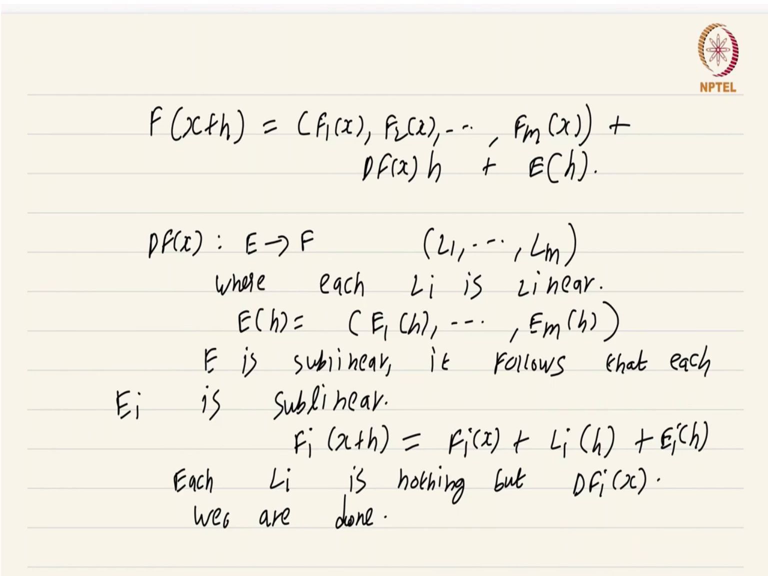
So, there will be m of them not n of them because E is R n and F is R m. Each F i is differentiable not on U at x because we assumed differentiability of F only at the point x.

So, then each F i is differentiable at x and the matrix DF x the matrix of the linear transformation with the basis B on the domain side and the basis C on the codomain so, the standard basis is nothing but; is nothing, but DF 1 x; DF 1 x, then let me just write it in a better way DF 2 x dot dot dot DF n x; DF n x rather I should not put D because that will be inaccurate, this is gradient of F 1 x, gradient of F 2 x dot dot dot gradient of F n x.

So, the various rows of this matrix are going to be gradient of F 1, gradient of F 2, dot dot dot gradient of F n x ok. So, this is called, this matrix is called the Jacobian matrix; is called the Jacobian matrix ok and this Jacobian matrix which is an m cross n matrix comprising partial derivatives.

So, when the map F is differentiable, you can write down the matrix quickly with respect to the standard basis by computing the various partial derivatives and putting them as a matrix. So, this is the jth partial derivative of the ith function, this is what the ijth entry is going to be. Let us see a proof of this, let us see a proof of this theorem.

(Refer Slide Time: 15:00)



All the hard work has actually been done when we prove the corresponding result for the gradient that the gradient is actually the representation matrix representation of the derivative map of a scalar valued function of a vector variable. So, let us write F of x plus h is equal to F 1 x, F 2 x comma dot dot dot F m x plus DF x h plus the sublinear error term, this is just the definition of F being differentiable at the point x.

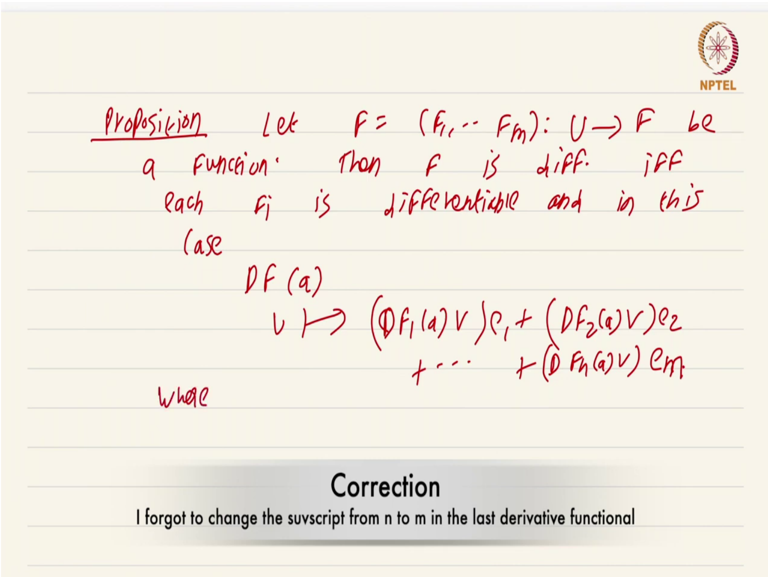
Now, DF x is itself a linear map from E to F which I can write in components as L 1 dot dot dot L m simply because F is an m dimensional space so, you can just write it as L 1 to L m where each L i; L i is linear because the map DF x is linear, each component must also be linear ok and we can also write this E of h likewise in components as E 1 h comma dot dot dot E m h. So, essentially, we have just taken the definition and written everything in components.

Now, because E is sublinear it is trivial that it follows that each; that each it follows that each E i is sublinear, this is utterly straightforward. Now, putting all this together, what we get is F i of x plus h is nothing, but F i of x plus L i of h plus E i of h, this is just equating all the components ok.

Now, it immediately follows that each L i; each L i is nothing but d F i at x which shows that each F i is differentiable and the fact that the matrix; fact that the matrix of DF x with respect to the standard basis is going to be just the various gradients of F i's follows from the fact that we have proved earlier that if you have a scalar valued function of a vector variable, then the matrix representation of the derivative map is nothing, but the gradient putting all this together, we are done; we are done; we are done.

So, this was fairly easy because we spent some time proving the easier case of scalar valued functions of a vector variable.

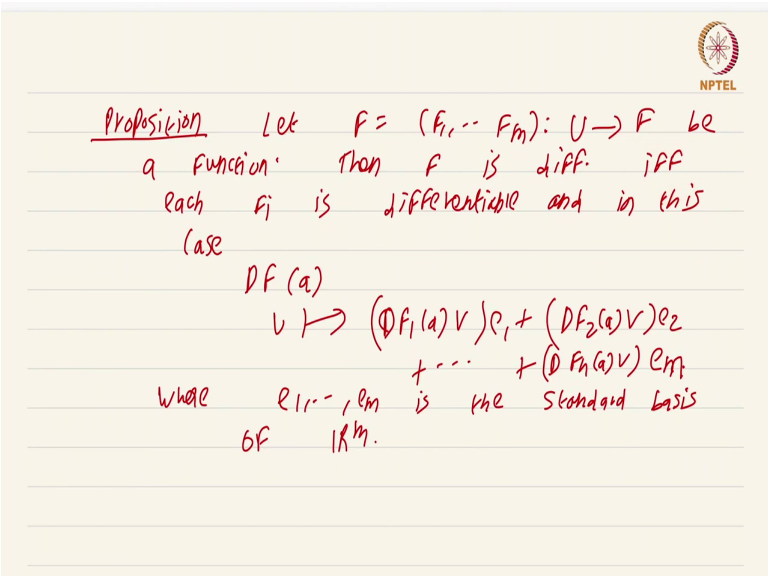
(Refer Slide Time: 18:06)



Now, what about the converse? Suppose I know that each F i is differentiable, then can I get the derivative, can I show that F is differentiable and what is its derivative? Well, we can do that by formulating the same proposition sort of the converse, but this time I am going to do it independent of these components ok.

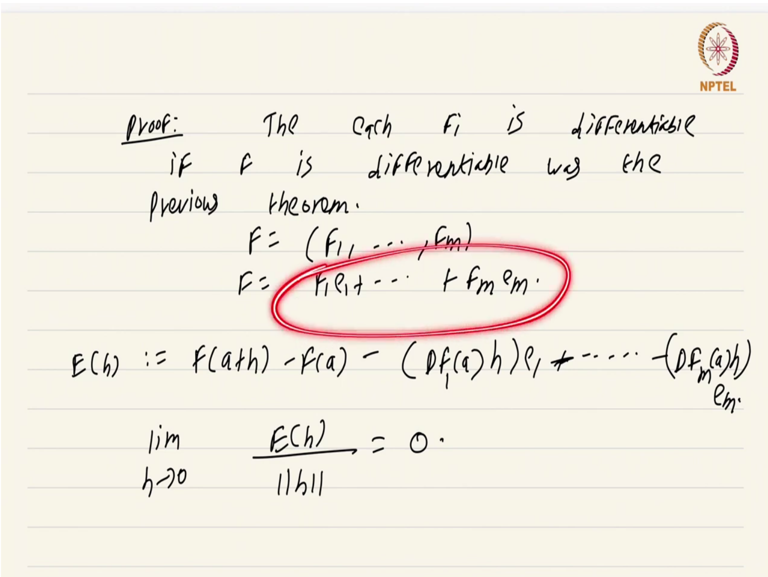
So, let F equal to F 1 to F m from U to capital F be a function, then F is differentiable; F is differentiable if and only if each F i is differentiable; if and only if each F i is differentiable and in this case; in this case, the derivative at a given point a is given by the formula v maps to D; DF 1 a e 1 so, just a second DF 1 a acting on the vector v e 1 plus DF 2 a acting on the vector v e 2 plus dot dot dot D F n a acting on the vector v e n ok.

(Refer Slide Time: 19:33)



So, wait a second, this should be e m right because I want a linear map into the space F which is an m dimensional space where of course, where e 1 to e m is the standard basis; is the standard basis; basis of R m. So, this is a formulation that does not involve matrices, it is exactly the same result including the convex, there is an if and only if condition now ok.

(Refer Slide Time: 20:02)

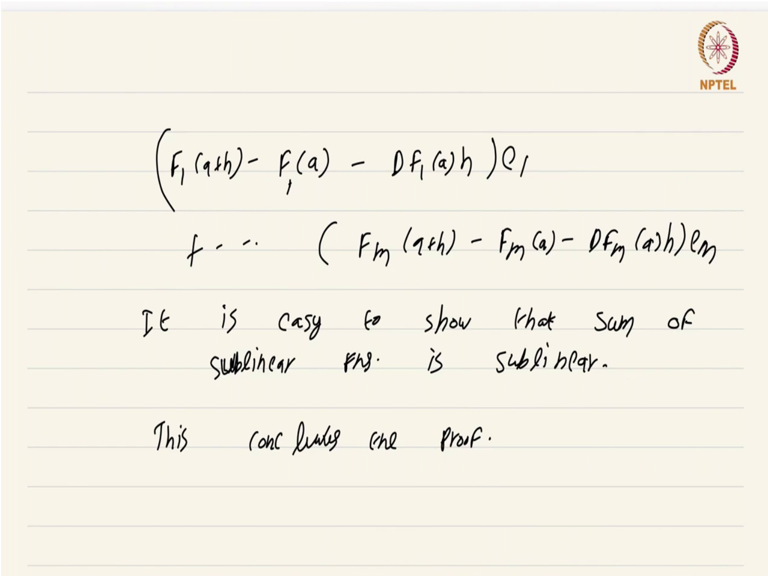


Now, one part has already been established as part of the last theorem so, proof that each F i is differentiable; that each F i is differentiable if F is differentiable; if F is differentiable was the last theorem, was the previous theorem ok. So, now, we have to prove the converse, we have to assume that each F i is differentiable and prove that the map F itself is differentiable. So, we have written F in terms of its components as F 1 to F m ok, this is same as writing F is F 1 e 1 plus dot dot dot F n; F m e m, this is exactly the same thing in a different notation.

Now, as we are going to assume that each F i is differentiable and conclude that F is differentiable. So, let us consider the difference F of a plus h minus F of a minus the map that we claim is the derivative that is DF a h e 1 plus dot dot dot actually its minus dot dot dot minus DF so, this is DF 1 so, this is DF m a h e m ok.

So, we have just written down the map we are interested in and what we are going to show is that this thing, this what remains is I am just going to call it E of h and show that limit h going to 0 E of h by norm h is equal to 0. This will show by the remark that I made just after I defined the derivative that the function F is differentiable and the candidate derivative that we have is indeed the derivative map ok.

(Refer Slide Time: 22:10)



Now, how do we know that this is going to be I mean a sublinear? Well, observe that by this remark that I made that F is nothing, but this, what we get is we can group the terms together and we will get F 1 of a plus h minus F a sorry minus F 1 of a minus DF 1 a h e 1 plus dot dot dot F m of a plus h minus F m of a minus DF m a h e m ok.

Now, by the fact that each F i is differentiable and; obviously, this is the derivative of each F i by definition, each one of these is sublinear and it is an easy fact, it is easy to show; it is easy to show that; show that sum of sublinear functions; sublinear functions is sub linear. So, I am being a bit vague here because there are these vectors e 1 to e m so, what exactly do I mean by sum and all that I am going to leave it to you, these are easy checks ok. So, this concludes the proof; this concludes the proof ok.

What is the net upshot of all this? Well, the net upshot is if you want to compute the derivative map, there are only two real ways, one is to somehow guess the derivative map and prove that in indeed it is the case that this is the derivative map alternatively, you can just use these results proved in this video and compute this which is just essentially one variable calculus, just compute the various partial derivatives of each component on the map F and put them all in a matrix that will give you the matrix representation of the derivative max derivative map.

In many treatments, we do not even introduce linear maps, we directly defined the derivative in terms of this Jacobian matrix ok. Now, but I prefer to do this coordinate independent definition of the derivative map as a linear map because that fits in with our philosophy that the derivative is supposed to be the best linear approximation of a function.

Remember, linear phenomenon are easy. We have spent several generations studying linear phenomenon, we have a huge set of tools that allows us to tackle linear phenomenon. So, whenever possible, whenever you are confronted in with a nonlinear phenomenon, try to reduce it to a linear case and the derivative map is the appropriate way to reduce whatever you are studying to the linear case.

This concludes this video, this is a course on real analysis, and you have just watched the video on the derivative as a linear map.