

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

Lecture - 18.2
The Implicit Function Theorem

(Refer Slide Time: 00:22)

The implicit Function theorem.

Definition Let $F: U \rightarrow \mathbb{R}^m$ be a function. The graph

$$\Gamma(F) := \{ (x, F(x)) : x \in U \}.$$

The graph of F is a subset of \mathbb{R}^{m+n} .

$x^2 + y^2 + z^2 = 1$ S² - unit circle

$$S^2 := \{ (x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1 \}.$$

In this video we shall prove the all important Implicit Function Theorem. The implicit function theorem is probably the best application of the inverse function theorem and it is also the starting point of manifolds, the theory of manifolds. We will also use the implicit function theorem to study optimization problem Lagrange multipliers these are the various applications of this theorem.

Since this theorem is very important let us motivate what this theorem is trying to say by a concrete example. For that I need a basic definition that of a graph. All of us are familiar from middle school where we took a graph paper and plotted various functions like x squared by

hand. So, the graph the notion of the graph it is just the set of points in the graph paper that we actually draw, so we make this precise in the following manner.

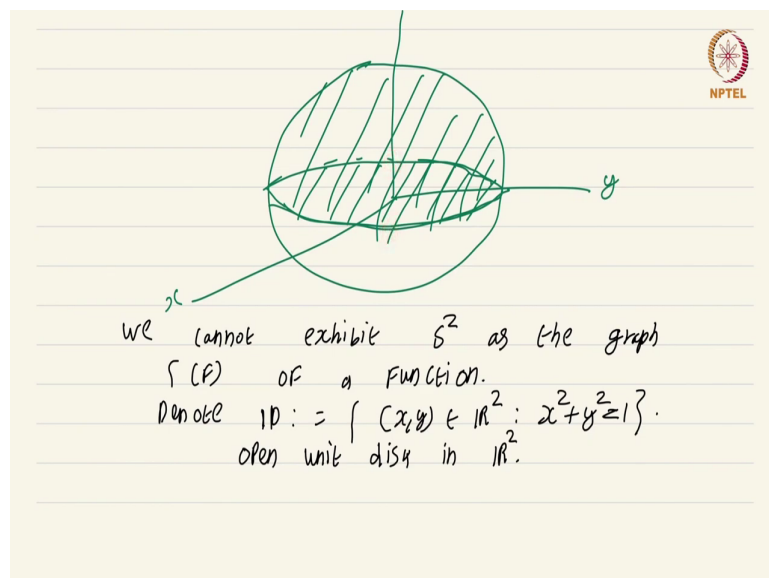
Let F from U to \mathbb{R}^m be a function. The graph which is usually denoted by capital gamma of F this is just defined to be the set of all points x comma F of x , if you recall this is what we actually plot in the graph paper such that x comes from U . So, you look at all the pairs x comma F of x and put them together into the set; obviously, this we can view as an element of \mathbb{R}^m plus n ok.

So, the graph of a function from open subset of \mathbb{R}^n to \mathbb{R}^m is just the collection of points x comma F of x such that x comes from U . So, we just note down for specificity that the graph of F is a subset of \mathbb{R}^m plus n ok. Now, we are going to illustrate a concrete version of the implicit function theorem with the all familiar unit sphere in \mathbb{R}^3 all of us know that the unit sphere in \mathbb{R}^3 is given by the equation $x^2 + y^2 + z^2 = 1$ ok.

Now, note this is the sphere that is essentially just the boundary of the unit ball the interior part is not there in this set. So, we are looking at the set which is classically called S^2 sphere of dimension 2 this will start to make sense once we define manifolds.

So, what this dimension is so, but the classical notation is S^2 for this object and of course, S^1 denotes the unit circle and S^3 will denote the unit sphere in higher dimensions in \mathbb{R}^4 ok. So, S^2 is defined to be the set of all points x, y, z in \mathbb{R}^3 , such that $x^2 + y^2 + z^2 = 1$, fine. So, pictorially this object is nothing but the unit sphere which looks like this.

(Refer Slide Time: 03:57)

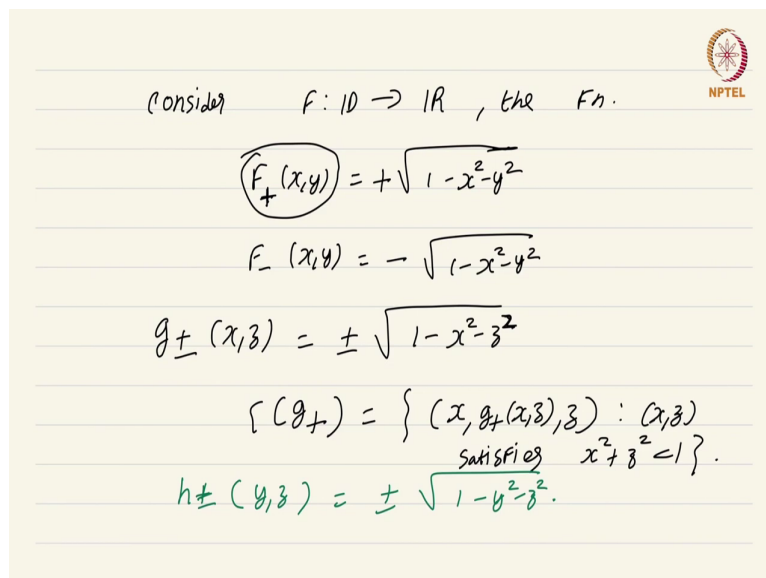


Of course what I am about to draw is badly drawn, but you get the idea. Now, you should think about why we cannot exhibit S^2 as the graph γ_F of a function. You cannot find a single function such that S^2 is the graph think about this carefully and try to come up with the solid reasoning as to why this is impossible.

However, pieces of S^2 you can write down as a graph, how does one do that? Well observe that this upper hemisphere can easily be written down as the graph of a function which is defined on this open set in \mathbb{R}^2 this bit which is essentially the unit disk ok. So, denote D to be x, y in \mathbb{R}^2 such that $x^2 + y^2 < 1$, this is the open unit disk, this notation this D with another line is borrowed from complex analysis open unit disk in \mathbb{R}^2 .

In complex analysis this object D is probably the most important domain on which complex functions are defined. So, we have a specific notation for that and you can consider the upper hemisphere of this unit sphere in \mathbb{R}^3 as the graph of a very simple function what is that function well consider F from B to \mathbb{R} .

(Refer Slide Time: 06:06)



consider $F: D \rightarrow \mathbb{R}$, the F_n .

$$F_+(x, y) = +\sqrt{1-x^2-y^2}$$

$$F_-(x, y) = -\sqrt{1-x^2-y^2}$$

$$g_{\pm}(x, y) = \pm \sqrt{1-x^2-y^2}$$

$$\{G_+\} = \{(x, g_+(x, y), y) : (x, y) \text{ satisfies } x^2 + y^2 \leq 1\}$$

$$h_{\pm}(y, z) = \pm \sqrt{1-y^2-z^2}$$

The function F of x, y is nothing but under root 1 minus x squared minus y square ok. So, here by to emphasize that I want the upper hemisphere let me put a plus sign to emphasize that this is the positive square root ok. So, this function again I will put a plus here. So, that it is it will become convenient in a moment. So, this functions graph is clearly the upper hemisphere ok.

In an analogous way I can get the lower hemisphere as the graph of the function F minus x, y is equal to the negative square root of 1 minus x squared minus y square. So, every positive real number has 2 square roots, so I take the positive square root for the function F plus and

the negative square root for the function F minus ok. Now fairly easy computations will tell you that this is the graph of this is nothing but the upper hemisphere and the graph of this is nothing but the lower hemisphere.

So, it looks like we have not I mean it is impossible to write the entire sphere as the graph of a function. But we have managed to write the unit sphere S^2 as the union of 2 graphs. Wait a second think carefully if you look at this picture this F plus the graph would be this upper part minus this equator the equator will be missing.

And the same thing is true for the lower the graph of the lower hemisphere that function will take I mean the graph will consist of all the points here, but not the equator. So, we have missed out all the points of the equator, but fear not we can write those points also as a graph except we change the domain.

What we do is so if we draw the 3 axis essentially what we have done is the upper hemisphere we have represented as the graph of the domain D which is lying in the $x y$ plane right. Well there is nothing stopping us from considering the unit disk on the $y z$ plane or the $x z$ plane and considering graphs over that. I mean it is just essentially we are just changing the base axis and what we are plotting.

So, this is really nothing happening in a substantial way it is just the same concept. So, what we can do is we can define more functions we can define more functions, we can look at the functions g plus or minus y sorry. You can consider the functions g plus or minus $x z$ which is just again plus or minus under root $1 - x^2 - z^2$ ok and consider the graph of this function.

So, the graph of this function will look slightly peculiar. So, this for concreteness let me write gamma of g plus this is just going to be the collection of all points $x g$ plus $x z$ comma z ok; such that $x z$ satisfies $x^2 + z^2 < 1$. So, what we have done is we have not just considered the graph as a function of $x y$, we are now considering the graph I mean

the function is a function of x and z and it is a graph in this sense that the middle variable y has the function in it.

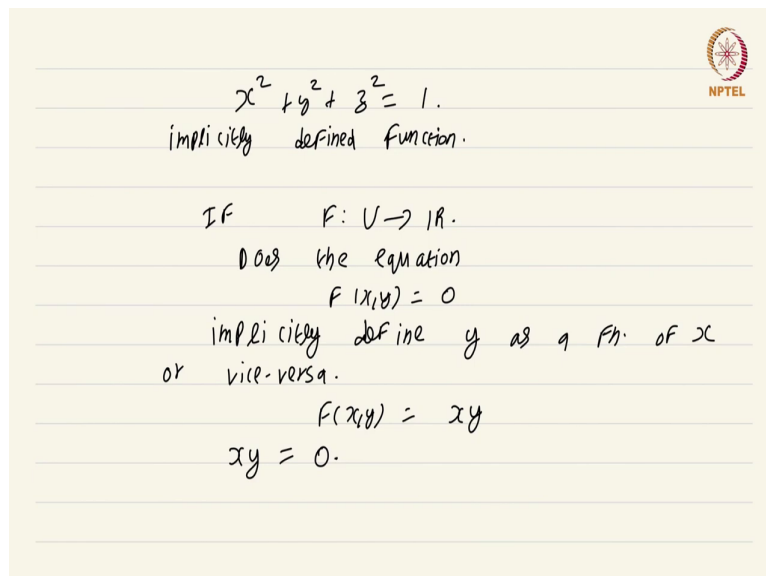
So, there is nothing really sacred about z always being a function of x and y . So, here we can treat y also as a function of x and z ok. So now, we have 2 more graphs one coming from g plus and one coming from g minus and if you think about it looks like we have covered all of the sphere, but no. So, if this in this picture so if this is the x this is the y and this is the z , we have managed to write down the left and the right hemisphere.

So, the right hemisphere would still be a graph and the left hemisphere will also be a graph. But again this what do you say this equator the equator along the x and z plane or whatever that will be missing from both these graphs, but they are there in the graphs that we have already done not quite two points will be missing; two points will be missing those are the two points that are there in the intersection of both these great circles. So, essentially this point and this point will be missing.

But again, if we are not we can remedy this situation quite easily by considering two more functions h plus minus h plus minus, this time I am going to consider y and z which is equal to plus or minus root of 1 minus y squared minus z squared. So, I want you to mull over this example for some time and convince yourself together the graph of these 6 functions will the union of this will be exactly the unit sphere in \mathbb{R}^3 .

So, we have not managed to write the unit sphere as the graph of a single function that is simply impossible. But what we have done is we have managed to write it as the union of the graph of 6 functions. So, locally every point on this sphere is the graph of a function and this observation will lead on to the definition of a manifold quite soon. But before that let us see what is it that we have managed to do using this discussion.

(Refer Slide Time: 12:42)



$x^2 + y^2 + z^2 = 1.$
implicitly defined function.

If $F: U \rightarrow \mathbb{R}.$
Does the equation
 $F(x, y) = 0$
implicitly define y as a fn. of x
or vice-versa.
 $F(x, y) = xy$
 $xy = 0.$

Well if you look at this x squared plus y squared plus z squared equal to 1 what we have essentially said is this equation implicitly defines several functions and we have explicated these implicit functions ok. So, you must have come across the term implicitly defined function, when you studied multivariable calculus at a basic level.

So, what we have done is from an implicitly defined function via this equation via this equation what we have managed to do is we have managed to construct several functions out of this. So, the question now follows if you have a function F from let us say U to \mathbb{R} for simplicity sake, suppose you have a function from U to \mathbb{R} we will be considering a more general scenario when we study the implicit function theorem.

But for the time being just consider the function F from U to \mathbb{R} , does it does the equation does the equation F of x, y equal to 0 implicitly define y as a function of x or vice versa. Suppose

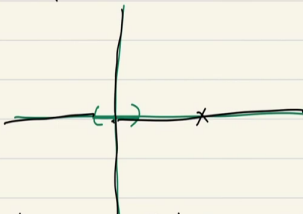
you have a function F from U to \mathbb{R} and you just consider the equation $F(x, y) = 0$, does it follow that you can write y as a function of x in a more geometric way.

What this is asking is suppose you are given $F(x, y) = 0$ and you consider the 0 set of this function that is the collection of all points (x, y) such that $F(x, y) = 0$, then is that 0 set a graph that is a geometric way of looking at this. Well let us look at a simple function for which this will not be true. Consider $F(x, y) = xy$ and look at the equation $xy = 0$ look at this equation.


Now no simple manipulation like what we did for the unit sphere is going to obtain y as a function of x or vice versa that is clear, we can actually make this precise ok.

(Refer Slide Time: 15:13)

Suppose $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ is a fn.
whose graph is precisely

$$\Gamma(g) := \{ (x, y) : xy = 0 \}.$$


This is not possible because $g(x)$ is going to be multi-valued.

 NPTEL

So, suppose g from \mathbb{R} to \mathbb{R} is a function whose graph is precisely x, y such that xy is equal to 0 oh sorry what I wrote down is nonsense such that x times y is 0. So, suppose you could write the collection of all points x, y such that the product is 0 has a graph Γ of g , then just think about what this graph is going to look like it is just going to be the union of the x axis and the y axis ok.

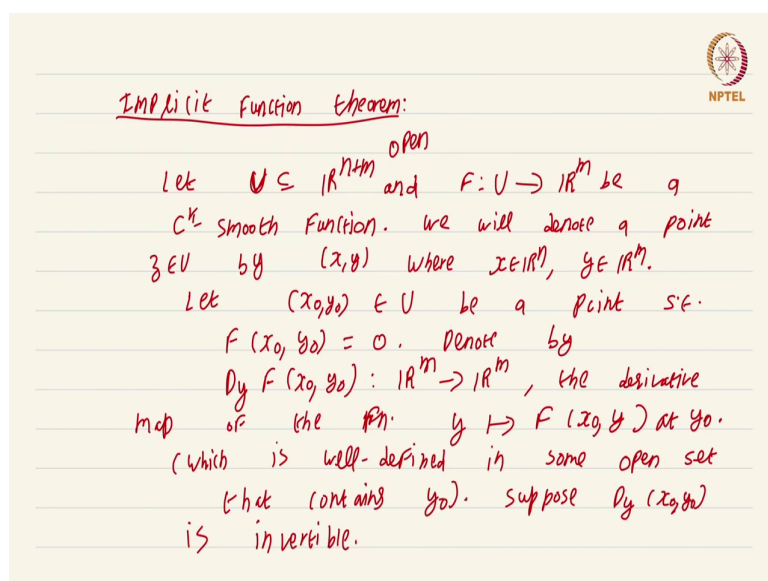
So, which means that if this is the graph this is the graph it near any interval of 0 at this point at least at this point 0, the graph will have to be multi valued. So, g this is not possible this is not possible because g of 0 is going to be multi valued is going to be multi valued ok.

So, there is no clear way it is very easy to define this function g away from 0 you just set a g of x equal to 0, so that will give you this portion and this portion. But at the origin you need to obtain this entire thing as the value of g of 0 which is nonsense such no such function exists, I mean the term multi valued function itself makes very little sense functions are by definition single value. So, the multi valued function is what is called an oxymoron such a thing is simply not possible.

So, we are not able to solve xy equal to 0 and get a graph. So, the implicit function theorem gives you a precise condition on F of x, y which allows us to extract a graph from this function at least locally. What goes wrong with the function xy at the origin is the fact that the gradient is 0 as you can check, the moment the gradient is not 0 then the implicit function theorem will allow you to extract a graph out of it locally.

So, we are going to now state and prove the implicit function theorem, it is going to be a straightforward consequence of the inverse function theorem, but there is a trick involved in the proof. So, let us go on to the statement of the inverse function theorem and prove it.

(Refer Slide Time: 18:11)



Implicit Function Theorem:

Let $U \subseteq \mathbb{R}^{n+m}$ be an open set and $F: U \rightarrow \mathbb{R}^m$ be a C^k smooth function. We will denote a point $z \in U$ by (x, y) where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.

Let $(x_0, y_0) \in U$ be a point s.t. $F(x_0, y_0) = 0$. Denote by $D_y F(x_0, y_0): \mathbb{R}^m \rightarrow \mathbb{R}^m$, the derivative map of the fn. $y \mapsto F(x_0, y)$ at y_0 . (which is well-defined in some open set that contains y_0). Suppose $D_y F(x_0, y_0)$ is invertible.

So, implicit function theorem, so the statement is as follows. Let U subset of \mathbb{R}^n plus m and F from U to \mathbb{R}^m be a C^k smooth function. So, the reason why we wrote down U as a subset of \mathbb{R}^n plus m instead of in a weird way is to make notation simpler, some people just write U as a subset of \mathbb{R}^n and F from U to \mathbb{R}^m then you will have n minus m and all that.

The way I have written it down the notation will become clean. Now, we are going to denote a point we will denote a point z in U by x comma y , where x comes from \mathbb{R}^n and y comes from \mathbb{R}^m . So, we are going to split the variables on which this function F depends into the variables coming from \mathbb{R}^n and the variables coming from \mathbb{R}^m . So, the first n we have club together and called it x and the remaining we have called it y .

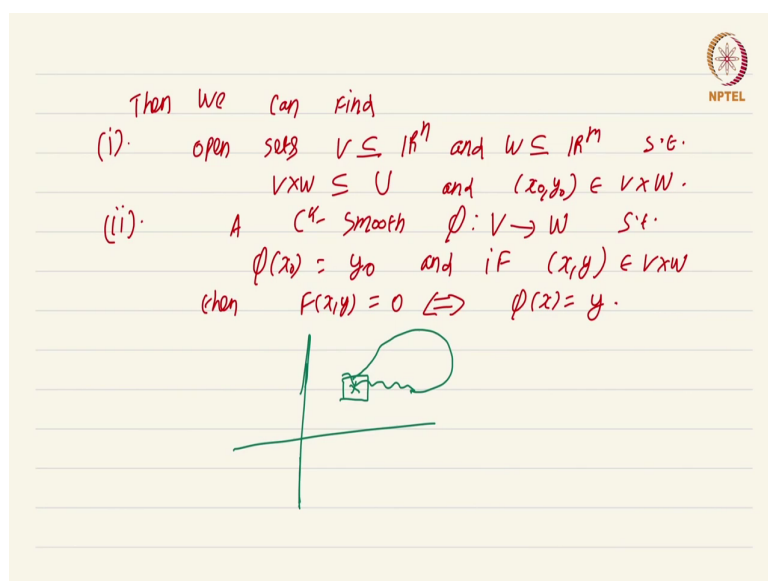
Let x_0, y_0 in U be a point such that $F(x_0, y_0) = 0$. So, consider a 0 of this function denote by dy . So, this is a new notation partial derivative with respect to the variables y , dy of F of x_0, y_0 . Well what is this? Well this is just a linear map from \mathbb{R}^n sorry it is a linear map from \mathbb{R}^m sorry about that \mathbb{R}^m to \mathbb{R}^m this is a linear map from \mathbb{R}^m to \mathbb{R}^m , this is the derivative map the derivative map of the function; of the function y maps to F of x_0, y_0 ok.

Now, observe that this map will be well defined in some ball in \mathbb{R}^m that contains the point y_0 because U is an open set because U is an open set. So, let me just emphasize that always there is a global remark that U is always open, but let me make that precise. So, because of the openness of U this function y maps to F of x_0, y is well defined in some open set that contains the point y_0 . So, which is well defined well defined in some open set some open set that contains y_0 ok.

So, this map the we are assuming denote by dy the derivative, of course this derivative will exist because F is a C^k smooth function so the derivative will exist ok. So, derivative of the function at y_0 of course that is crucial I have not specified what the point is ok.

Suppose dy_{x_0, y_0} is invertible, so this is a linear map from \mathbb{R}^m to \mathbb{R}^m I am going to assume that this map is invertible. So, this is the precise condition that allows you to write the 0 set of F as a graph. So, the conclusion unfortunately has to come in a different page, but the conclusion is long it is unavoidable.

(Refer Slide Time: 22:31)



Then we can find

(i). open sets $V \subseteq \mathbb{R}^n$ and $W \subseteq \mathbb{R}^m$ s.t.
 $V \times W \subseteq U$ and $(x_0, y_0) \in V \times W$.

(ii). A C^k smooth $\phi: V \rightarrow W$ s.t.
 $\phi(x_0) = y_0$ and if $(x, y) \in V \times W$
 then $F(x, y) = 0 \Leftrightarrow \phi(x) = y$.

The diagram shows a coordinate system with a horizontal and vertical axis. A curve starts from the left, goes up and right, then loops back down and left to a point marked with a small square and an asterisk, representing the point (x_0, y_0) .

Then number 1 or rather then we can find then we can find number 1 open sets V subset of \mathbb{R}^n and W subset of \mathbb{R}^m such that this $V \times W$ is contained in a U and x naught y naught is there in this product neighborhood $V \times W$. Number 2 a C^k smooth map so the first part is trivial that just the openness will give those neighborhoods V and W these open sets V and W .

But the 2nd part tells you there is something special about this V and W we can find V and W with some special properties. We can also find A C^k smooth mapping which I am going to call ϕ from this V to W notes this is A C^k smooth mapping, such that ϕ of x naught is equal to y naught and I am going to quantify that the 0 set can be written as a graph of the variable x y can be written as a function of y sorry y can be written as a function of x and that is made precise by saying.

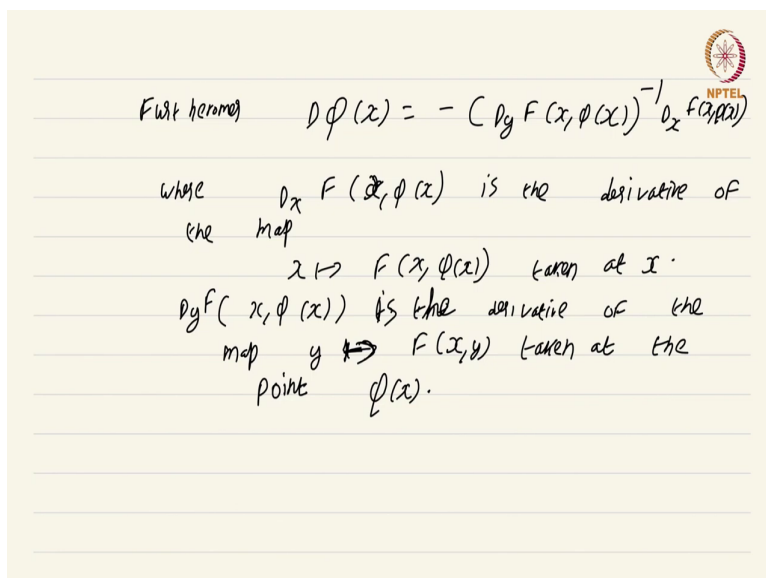
And if $x, y \in V \times W$ then $F(x, y) = 0$; that means, if you look at the 0 set of F which is also within the product open set $V \times W$; this can happen if and only if $\phi(x) = y$ ok. So, at this point it is appropriate to draw pictures to illustrate what is happening, well essentially what is happening is suppose let us just for argument sake say that we have only two variables.

And let us say this is the 0 set this is the 0 set of this given function F what it is saying is that at a given point you can find a product neighborhood you can find a product neighborhood in which you can establish the 0 set you can establish the 0 set as the graph, as the graph of a C^k smooth function. So, this may not be globally true, so what can happen for instance is that you it could happen that this function 0 set turns back and comes back inside ok.

So, you may not be able to globally do this there is no way to now write the 0 set as a global function, but by shrinking this $V \times W$ we can still see that it is going to be a graph of C^k smooth functions locally at least. So, the conclusion of the implicit function theorem is a local construction this is not surprising because the conclusion of the inverse function theorem was also local.

So, you cannot determine the global property of a function just by knowing that one partial derivative with respect to the variables y is invertible ok. So, there is one more part of this conclusion.

(Refer Slide Time: 26:13)



First hence $D\phi(x) = - (D_y F(x, \phi(x)))^{-1} D_x F(x, \phi(x))$

where $D_x F(x, \phi(x))$ is the derivative of the map $x \mapsto F(x, \phi(x))$ taken at x .

$D_y F(x, \phi(x))$ is the derivative of the map $y \mapsto F(x, y)$ taken at the point $\phi(x)$.

Furthermore $D\phi(x)$ you can write down the derivative of $D\phi(x)$ and you must have done this in the one variable case when you studied implicit differentiation informally, this is the negative of D_y of F of x comma ϕ of x . So, that is you take I mean there is an inverse.

So, you take the derivative with respect to the variables y at the point x comma ϕ of x and then invert the derivative and this has to get multiplied by the derivatives with respect to the x variables also taken at the point x comma ϕ of x no inverse here, because we do not even know that this is invertible. So, where $d_x F \in F(x, \phi(x))$ is the derivative is the derivative of the map of the map x maps to ϕ of x comma ϕ of x ok.

And D_y of x comma ϕ of x is the derivative $D_y F$ ok is the derivative of the map of the map y goes to $D_x y$ taken. So, here I made a mistake this is x goes to F of x comma ϕ of x

ok taken at x the derivative taken at x and here it is the map derivative of the map y goes to y goes to F of x y taken at the point x comma ϕ of x or rather taken at the point ϕ of x .

So, x is fixed here x is fixed here and you consider the map y going to F of x y and take the derivative at the point ϕ of x , not very different from what we did to define D_y of F of x naught y naught ok. So, this final conclusion is a bit subtle you have to understand what exactly is happening well come to that in the proof ok, now let us go to the proof.

The proof is not hard so before that let me make one more remark the statement of this theorem is rather complicated, the way I have written it is actually 3 slides. But what it is essentially saying is fully illustrated in this picture. So, it says that roughly if you have a C^k smooth function such that derivatives with respect to some variables is nonsingular, that just means that the derivative with respect to some variables is invertible.

Then the 0 set can be locally expressed as the graph of a C^k smooth function that is essentially what the implicit function theorem is saying in mathematical English. The variables with respect to which the derivative is non singular they are the dependent variables.

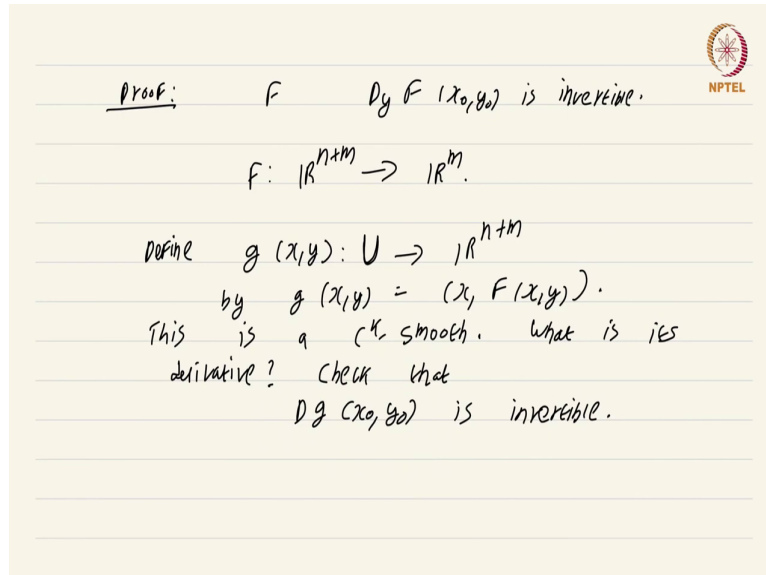
So, here we were writing y as a function of x and the rest of the variables are the independent variables. Now one remark the way I have stated this theorem you notice that this will not if you look at the previous example we wrote the middle variable as a function of the first and the last.

So, it will not happen in general that the final set of variables will be the dependent variables and the first set of variables will be the independent variable that was usually not the case. What will essentially happen is that some set of variables will be dependent and some set of variables will be independent.

This is just essentially the same scenario as what we are considering we are just renumbering the coordinates. So, it can happen that the second the third and the fifth variables are the dependent variables and the other variables are the independent variables. So, this is just a

minor variant of the implicit function theorem, I am not going to dwell too much on that because it is just a slightly notationally complicated thing, but the idea is exactly the same ok.

(Refer Slide Time: 31:08)



Proof: F $D_y F(x_0, y_0)$ is invertible.

$F: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m.$

Define $g(x, y): U \rightarrow \mathbb{R}^{n+m}$
 by $g(x, y) = (x, F(x, y)).$
 This is a C^k smooth. What is its
 derivative? Check that
 $Dg(x_0, y_0)$ is invertible.

So, we are going to move on to the proof of the implicit function theorem and the proof is somehow manufacture a function from \mathbb{R}^{n+m} to \mathbb{R}^{n+m} whose derivative at some point is invertible. So, we are given this function F and we know that $D_y F$ at x naught y naught is invertible is invertible. But we cannot really apply the inverse function theorem because F happens to be a function from \mathbb{R}^{n+m} to \mathbb{R}^m ok.

Now, this $d_y F$ x naught y naught the way it is defined this is going to be a linear map from \mathbb{R}^m to \mathbb{R}^m . So, there are n variables missing in the co domain of the function F somehow we have to increase the co domain by m variables. So, what is the simplest and most natural


thing to do we will fill in the missing variables exactly as they are. What do I mean by that define g of x, y g of x, y from U to \mathbb{R}^n plus m by g of x, y is equal to x comma F of x, y .

So, the variables that are missing in order to apply to make both the domain and co domain equidimensional you just tag it along in the first coordinates of the co domain just tag it along. So, this is the simplest thing we can do this is a C^k smooth map that is very easy to check because F is already a C^k smooth map and we have just tagged along essentially just the identity function in x . So, this is going to be a C^k smooth map.

What is it is derivative what is it is derivative? That is the key right we want to somehow by hook or by crook apply the inverse function theorem. So, for that we need to find out the derivative and I am going to make that your headache by simply saying that check that Dg_x naught y naught is invertible ok. So, I am going to I mean I will give a proof I am going to leave one part to you can compute the derivative map as a matrix and see that it is invertible.

What I mean by that is you already know that the sub matrix which is going to be contributed by F that is going to be an m cross n invertible matrix and this part is essentially going to contribute the identity and you can use the properties of determinant to actually check that the matrix of Dg_x naught y naught is going to be invertible. I will give a proof that does not depend on matrix x , it a matrices it is a direct proof. So, let us directly prove that the derivative of $g(x, y)$ is going to be invertible.

(Refer Slide Time: 34:25)


$$\begin{aligned} \text{Let } (h, k) &\in \mathbb{R}^n \times \mathbb{R}^m. \\ g(x+h, y+k) - g(x, y) &= (h, F(x+h, y+k) - F(x, y)) \\ &= (h, DF(x, y)(h, k) + E(h, k)). \end{aligned}$$

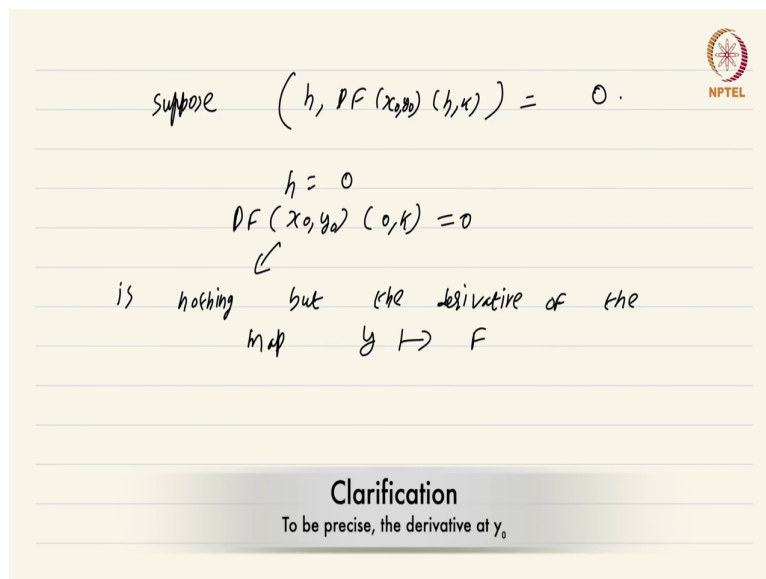
where E is sublinear. It follows that g is diff. and derivative of g is $(h, k) \mapsto (h, DF(x, y)(h, k))$.

So, what you do is let h comma k be vectors in \mathbb{R}^n cross \mathbb{R}^m and choose them to be really small. So, that this g of x plus h comma y plus k minus g of x y . So, this first part should be well defined g of x plus h comma y plus k choose h and k . So, small now if you just look at the definitions and notice what will happen, because x plus the first few coordinates of g is just going to be x plus h minus h this is going to be just h and the second coordinates is going to be F of x plus h comma y plus k minus F of x y ok.

Now, because this F is differentiable we can write this as DF x y acting on the vector h comma k plus a small error term e h k this is just the differentiability of the function F ok. So, E is sub linear where E is sub linear and comes from the definition of the differentiability of F ok. Now it follows easily that it follows that it follows that g is differentiable that g is differentiable and the derivative of g is h k maps to h comma DF x y acting on h k .

So, this is exactly the remark I was making in the matrix proof, you will get a sub matrix which is going to be contributed by the derivative of F of x, y and another sub matrix which is essentially going to be the identity mapping. So, identity matrix this is just the coordinate independent way of saying the same thing, I am just saying the same thing in term without using matrices not really coordinate independent without using matrices ok. So, we want to show that this map is actually invertible.

(Refer Slide Time: 36:53)



Suppose $(h, DF(x_0, y_0)(h, k)) = 0$.

$h = 0$
 $DF(x_0, y_0)(0, k) = 0$

↙
 is nothing but the derivative of the
 map $y \mapsto F$

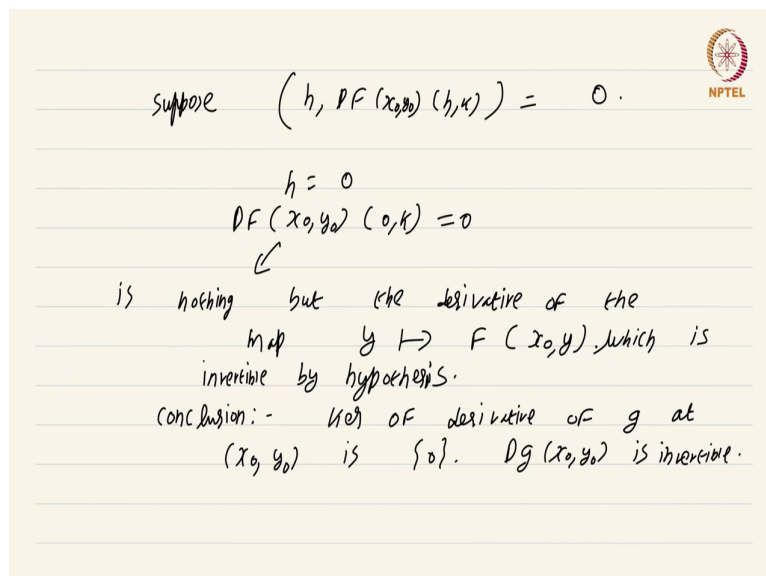
Clarification
 To be precise, the derivative at y_0

So, what I am going to say is suppose this derivative map h comma $DF(x, y)(h, k)$ suppose this is 0. So, essentially I am trying to find out the kernel of the map and show that it goes to 0. Immediately we get that h is 0 and we also get that $DF(x, y)(0, k) = 0$ just a minor modification I am interested only in the point x_0, y_0 right. Because I want to show

that it is invertible there in any case I have data about the sub matrix or sub the portion of the derivative $dF_{(x,y)}$.

So, it follows that $D F_{(x,y)}$ acting on h is also 0, but h is 0. So, this is essentially just saying that $D F_{(x,y)}(0,k) = 0$. Now I leave it to you to observe that this map $D F_{(x,y)}(0,k)$ this is nothing but the derivative of the map of the map $y \mapsto F(x,y)$.

(Refer Slide Time: 38:24)



Suppose $(h, DF(x_0, y_0)(h, k)) = 0$.

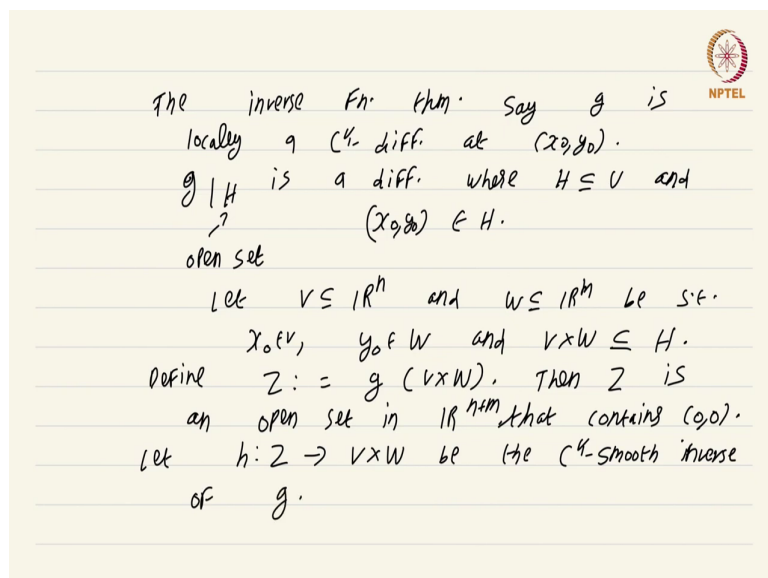
$h = 0$
 $DF(x_0, y_0)(0, k) = 0$

↙
 is nothing but the derivative of the map $y \mapsto F(x_0, y)$, which is invertible by hypothesis.

Conclusion: - Ker of derivative of g at (x_0, y_0) is $\{0\}$. $Dg(x_0, y_0)$ is invertible.

I want you to check that this is a rather trivial, this is just the map derivative of the map $y \mapsto F(x,y)$ which is invertible which is invertible by hypothesis ok. I want you to check this precisely what this is saying. Now, conclusion is that kernel of derivative of g at (x,y) is set with 0 and by basic linear algebra this just means that $Dg_{(x,y)}$ is invertible it is nonsingular.

(Refer Slide Time: 39:33)



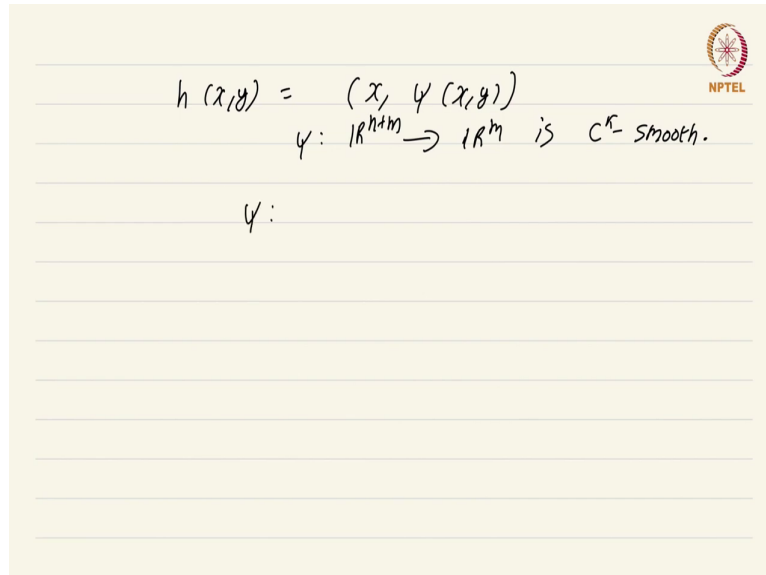
Now, the inverse function theorem the inverse function theorem inverse function theorem says g is locally a C^k diffeomorphism at this point x naught comma y naught ok. Let us say g restricted to h is a diffeomorphism is a diffeomorphism where h is a subset of U and x naught y naught is contained in h .

So, we are just choosing an open set we are just choosing an open set in which g is going to be at C^k smooth diffeomorphism ok. Now what we do is let V subset of \mathbb{R}^n and W subset of \mathbb{R}^m be such that x naught is in V y naught is in W and V cross W is a subset of h ok.

And define z to be the image of V cross W under g ok. Now, because this is a diffeomorphism because g is a diffeomorphism then z is an open set in \mathbb{R}^n plus m think

about why this is true? This just follows because it is a diffeomorphism and this set contains the origin that contains 0 ok. Let h from z to V cross W be the C^k smooth inverse of g ok.

(Refer Slide Time: 42:01)



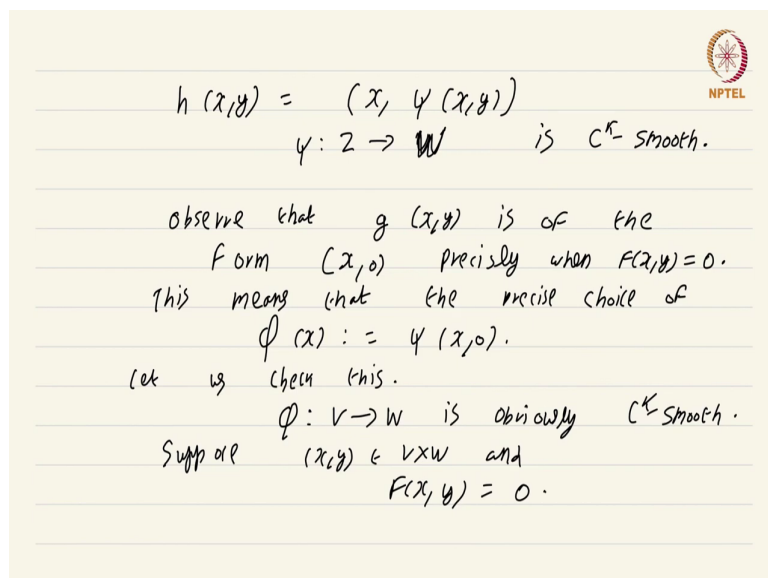
$$h(x, y) = (x, \psi(x, y))$$

$$\psi: \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m \text{ is } C^k \text{ smooth.}$$

$$\psi:$$

Now here is the crux h of x, y is of the form x of x comma ψ of x, y , because g is of the form x comma F of x, y h of x, y the inverse must be of the form x comma ψ of x, y where ψ from \mathbb{R}^{n+m} to \mathbb{R}^m is C^k smooth ok. Now this ψ is actually sorry this is not from \mathbb{R}^{n+m} to \mathbb{R}^m this is from z to W .

(Refer Slide Time: 42:42)



$h(x, y) = (x, \psi(x, y))$
 $\psi: Z \rightarrow W$ is C^k smooth.

observe that $g(x, y)$ is of the form $(x, 0)$ precisely when $F(x, y) = 0$.
This means that the precise choice of $\phi(x) := \psi(x, 0)$.

let us check this.
 $\phi: V \rightarrow W$ is obviously C^k smooth.

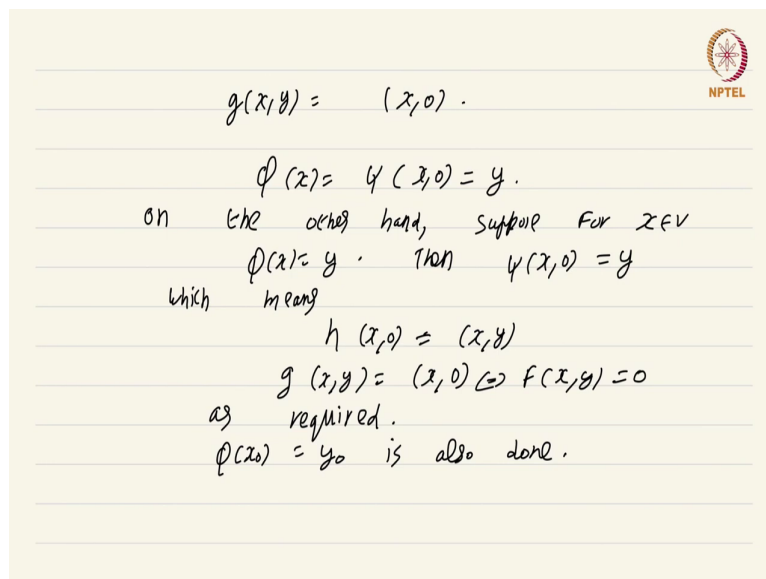
Suppose $(x, y) \in V \times W$ and $F(x, y) = 0$.

Because we know the domains of g and W g and h this is from z to W sorry about that. So, this ψ from z to W is a C^k smooth map ok. Now here is the crux of the proof observe that g of x, y is of the form $(x, 0)$ precisely when F of x, y is 0 right, I am writing down a bunch of triviality so do not get overwhelmed by this. This means that the choice the precise choice of ϕ of x is to just define it to be ψ of $x, 0$.

So, this is the crux of the proof and really there is nothing I can do to explain it, because it is actually a triviality when you realize why this is the only choice for ϕ . So, I want you to sit under a tree and look at the nice blue sky and figure out why, why is it that ϕ is the nice and appropriate choice that we need to finish this theorem. But I am not just going to leave it dangling in the air and let you finish the proof let us just confirm this. So, let us check this.

Let us check this mathematically I want you to think about why this is clearly true when you think about it intuitively. Now of course, ϕ from this function ϕ from V to W is obviously C^k smooth, it is obviously C^k smooth no questions about that the issue is the condition that $F(x, y) = 0$ and $x, y \in V \times W$ if and only if $y = \phi(x)$. So, to confirm that suppose (x, y) is in $V \times W$ and $F(x, y) = 0$ suppose you take a 0 of this given function F we want to confirm that y is equal to $F(x)$ ok.

(Refer Slide Time: 45:13)



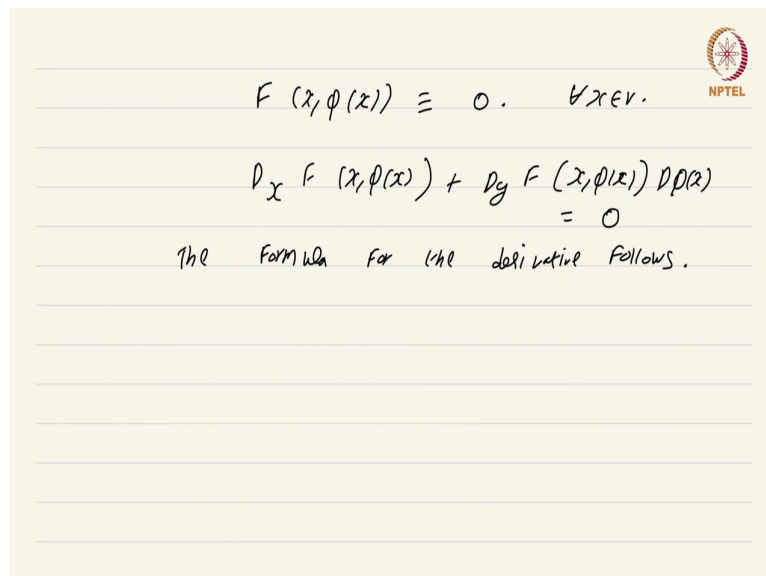
$$g(x, y) = (x, 0).$$

$$\phi(x) = \psi(x, 0) = y.$$
 on the other hand, suppose for $x \in V$
 $\phi(x) = y$. Then $\psi(x, 0) = y$
 which means
 $h(x, 0) = (x, y)$
 $g(x, y) = (x, 0) \Leftrightarrow F(x, y) = 0$
 as required.
 $\phi(x_0) = y_0$ is also done.

Then this just means this just means $g(x, y)$ is equal to x comma 0 right that is just by definition and that just means $\phi(x)$ which is just defined to be $\psi(x, 0)$ has to be y ok and this is all just by definition. So, really as I said this is nothing but a triviality, on the other hand on the other hand suppose for x and V we have $\phi(x)$ equal to y .

Then that just means then that just means that the ψ of x comma 0 is y ψ of x comma 0 is y which means h of x comma 0 is x comma y ; which means g of x comma y is x comma 0 ok which means F of x y is 0 as required. So, this entire proof is just unraveling the definitions of what g h and ϕ are and that will give it. Of course, this ϕ of x naught equal to y naught is also done ok. So, only thing that remains to be done is to show the formula for the derivative.

(Refer Slide Time: 46:52)



$$F(x, \phi(x)) \equiv 0. \quad \forall x \in V.$$

$$D_x F(x, \phi(x)) + D_y F(x, \phi(x)) D\phi(x) = 0$$

The formula for the derivative follows.

But from what we have established we know that ϕ of x comma ϕ of x is the identically 0 function ok for all x and V , that just means the derivative of this map is identically 0 . But what is the derivative of this map by chain rule the derivative of this map is nothing but $D_x F$ of x comma ϕ of x plus $D_y F$ of x comma ϕ of x $D\phi$ of x this is just this just follows from the chain rule and which I want you to check.

And we know that this has got to be the 0 map and the formula follows the formula for the derivative follows ok. So, this concludes the proof of the implicit function theorem it is a consequence of the inverse function theorem. So, let me also remark that in some treatments you first establish the implicit function theorem and obtain the inverse function theorem as a corollary.

But I my opinion that is not the correct way to go about the proof simply because the way we have done it this will generalize to other situations for instance to infinite dimensional non vector spaces, which we briefly talked about when we talked about higher derivatives. So, this approach is sort of the natural one to first establish the inverse function theorem either using Newton's method or by the contraction mapping principle and obtain the implicit function theorem as a corollary.

Now there are new approaches to the implicit function theorem in recent times that give a completely elementary proof that does not rely on any sophisticated theorems like the Banach fixed point theorem the contraction mapping principle Newton's method and so on these proofs. However, are still long there are no simple proofs of this theorem that are short as well as elementary.

Anyway the rest of this course is essentially a big application of the inverse and implicit function theorems. In the immediate application in the next video we will fix that part about the tangent space of a hyper surface, we will prove that the tangent space of a hyper surface is going to be an $n - 1$ dimensional vector space that is the first application of the inverse function of the implicit function theorem. This is a course on real analysis and you have just watched the video on the implicit function theorem.