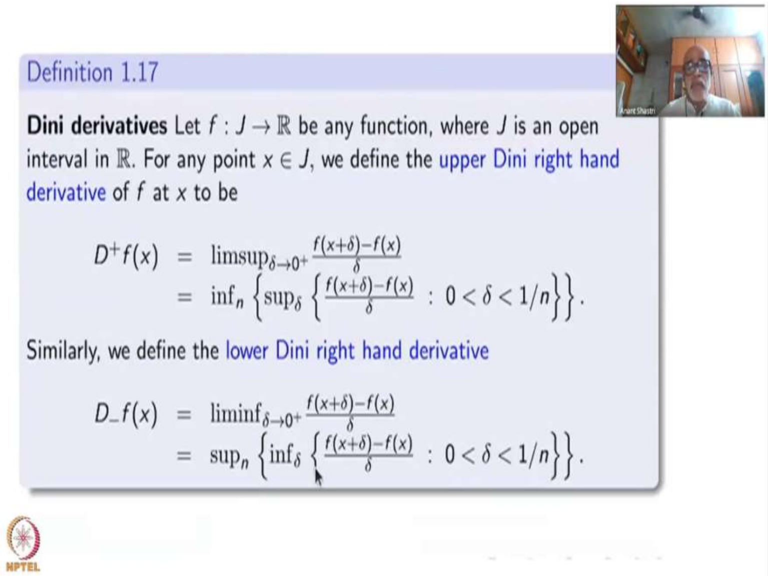
**An Introduction to Point-Set-Topology (Part II)**
**Professor Anant R. Shastri**
**Department of Mathematics**
**Indian Institute of Technology, Bombay**
**Lecture 3**
**Preliminaries from One-Variable Real Analysis**

(Refer Slide Time: 00:16)




Hello. Welcome to NPTEL NOC, an introductory course on Point Set Topology, Part 2, Module 3. So, today, we shall begin with some preliminaries required from one variable real analysis. I can say that it is one variable calculus, but in calculus courses, we do not go this deeper. So, I take this opportunity to do this one. I will explain why this is needed a little later.

Start with any function $f$ defined on an open interval $J$ to $\mathbb{R}$. For any point $x \in J$, we will define the upper Dini right hand derivative of $f$. They are attributed to Dini. He was one of the French mathematicians, I believe. So, let us just call them upper right-hand derivatives. They are not right-hand derivatives. You are already familiar with the right-hand derivatives and left-hand derivatives.

Now, we are only taking upper right-hand derivatives, notation is $D^+(f)(x)$, which are nothing but the limsup of these divided differences, $f(x+\delta) - f(x)/\delta$, where the only thing which makes it the right derivative is that $\delta$ tends to $0^+$. So, only from positive part, all is always taken positive here. $\delta$ tends to $0$ but from positive side. Note that you are not taking the limit which may not exist but you are taking the limsup which always exists.

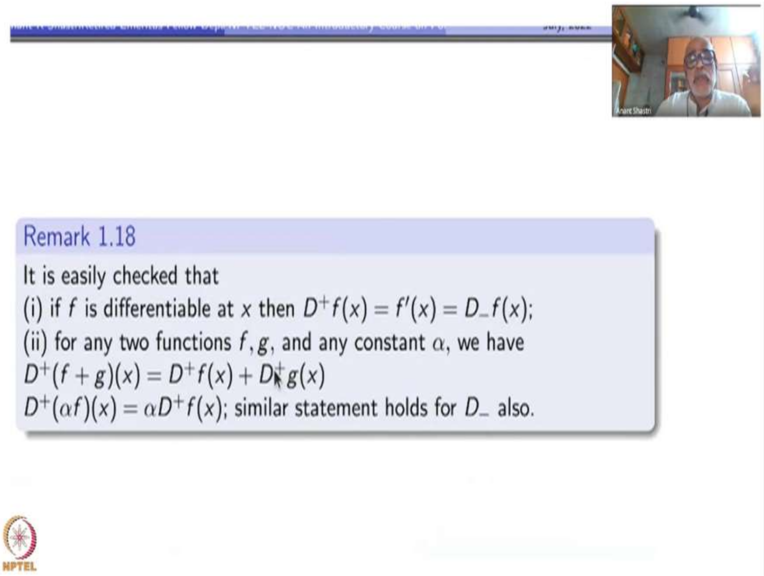
The whole idea is that limsup always exists, no matter what the function is, if you allow infinity also. If $f$ is a bounded function, then the simsup will be finite. In any case, it will always exist. So, I am just recalling the definition of limsup which is nothing but infimum of the following sequence,--- for all $n$ here, where the inside thing is supremum of $f(x+\delta) - f(x)/\delta$, over all the $\delta$'s between $0$ and $1/n$.

Note that as $n$ becomes larger, the range interval for $\delta$ becomes smaller and smaller and so does the supremum. So, this sequence, as a sequence in $n$, is a monotonically decreasing sequence, therefore its infimum exists, which actually the limit of the sequence. So, $D^{+}(f)(x)$ always exists. This is called upper Dini right hand derivative. Similarly, I can define the lowered Dini right hand derivative.

The only thing I change is instead of limsup, I take liminf --- first you take infimum over deltas ranging in these intervals to get a monotonically increasing sequence and then you take supremum of that sequence. It can be shown that I mean, it is very clear from this that liminf is always smaller than limsup and so on.

Both of them always exist but if they are equal then it will be actually the limit namely the right hand derivative of $f$ at $x$. So, these things you must be knowing, but now, we have these symbols here $D^{+}$ and $D_{-}$ for upper right hand derivative and lower right hand derivatives.

(Refer Slide Time: 05:14)

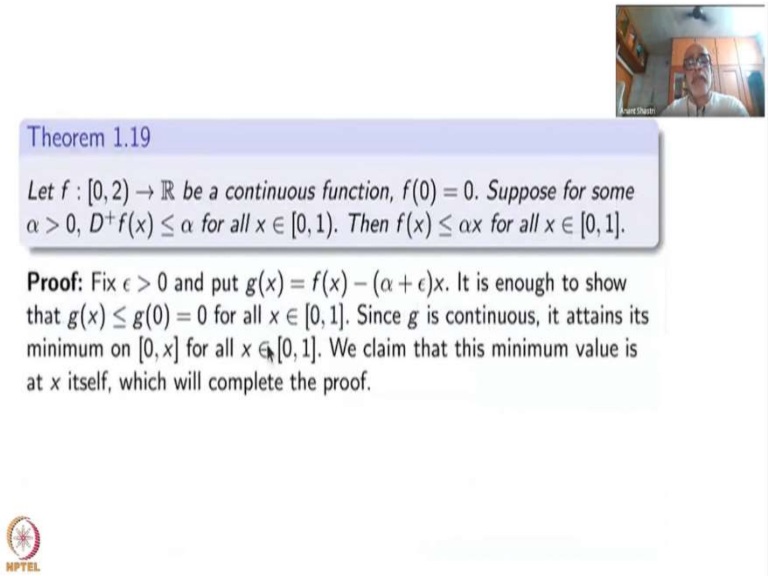



So, it is easily checked that if $f$ is differentiable at $x$ then $D^+$ is equal to $D_-$ equal to the derivative. For any two functions $f$ and $g$ any constant $\alpha$, we have $D^+(f+g)$ is $D^+(f) + D^+(g)$ and $D^+(\alpha f) = \alpha D^+(f)$. So, these are just linearity of $D^+$ which follows by the corresponding property for limsup. The same thing for $D_-$ also.

By the way, I have only talked about the right-hand derivative. It is exactly same way you can introduce two more $D$'s here from the left hand also. So, there will be four such quantities. If $f$ is differentiable then all these four quantities are equal to the derivative itself. Indeed this is if and only if. So, that is the easiest way to see that.

But I am not interested, right now, in the left-hand derivatives here at all. So, neither I am going to do a lot of things with these things. These things are very very helpful in analysis in general. So, I take this opportunity just to introduce them. My main aim is to use them to prove the so-called weak mean value theorem in the case of Banach spaces. So, let us go ahead with that.

(Refer Slide Time: 07:14)




So, this is a theorem which will help us to do that job. This is very simple thing. It depends only on $D^+$ which always exists. So, all that I do is to start with some open interval say containing the closed interval $[0,1]$, and a continuous real valued function $f$ on it. Assume $f(0) = 0$. These are harmless assumptions. You can always assume $f(0)$ equals something and then subtract that. That is not a problem. So, this is just a technical assumption. The basic assumption here is that this is a continuous function.

Suppose, now for some $\alpha$ positive, $D^+(f)(0) \leq \alpha$ for all $x \in [01)$. Suppose we have found out a bound, upper bound for $D^+(f)$. So, pay attention, I am taking it only in the interval $[0,1)$. All that I want is this condition should hold in $[0,1)$, but I want the function to be defined at $1$ also. So, the conclusion is that $f(x) \leq \alpha x$ for all points in $[0,1]$.

So, just for a continuous function, if the Dini derivative $D^+$ is bounded, it will give you a bound for the function a very specific bound in terms of the given bound for the Dini derivative. So, this bound for the Dini derivative must hold on the half open interval $[0,1)$. But the conclusion is for the entire closed interval.

So, how do I go ahead? If I show that $f(x) \leq (\alpha + \epsilon)x$ for every $\epsilon$ positive, then it must be true for $\alpha$ as well.

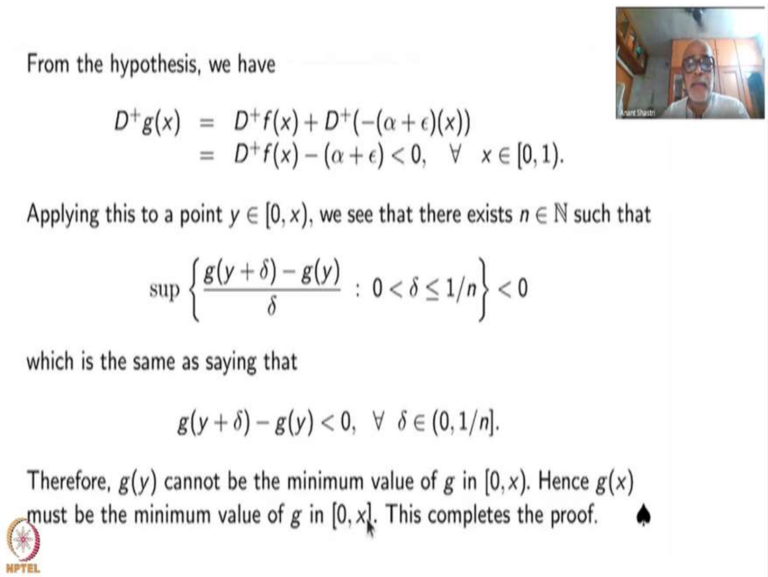
So, I consider the function $g(x) = f(x) - (\alpha + \epsilon)x$ and show that this function $g(x) \leq g(0) = 0$. $g(0)$ here is what? Here, $x$ equal to $0$, if $x$ equal to $0$ here is $f(0)$ which is $0$.

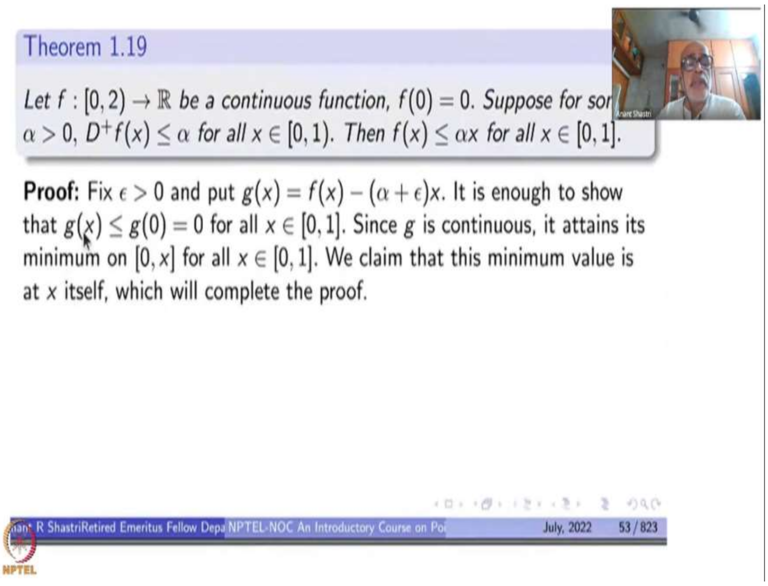
So, we have show that $g(x)$ is non-positive for all $x$. The same thing as saying that $f(x) \leq (\alpha + \epsilon)x$, for every $\epsilon$. So, how do I prove that $g(x) \leq g(0)$ for all $x \in [0, 1]$.

Now, the only hypothesis that I have is that $g$ is continuous because $f$ is continuous. This $g$ is just a difference of two continuous functions here. So, $g$ is continuous. Continuous function on a closed interval attains its minimum on each closed interval $[0, x]$ for all $x \in [0, 1]$.

We claim this minimum value is at $x$ itself. But then in particular we get $g(x) \leq g(0) = 0$, because $0 = g(0)$ is also one of the values in the interval. So, we want to prove $g(x) \leq g(0)$ but we are proving a very strong result namely, that the function itself takes minimum value at $x$. So, there is a lot more we are proving here (monotonically decreasing!) but we are not using all that. We are just using that $g(x)$ is non-negative.

So, how do you prove that the minimum value of this $g$ is at $x$. So, that is where the Dini derivative $D^+$ will help us. (Refer Slide Time: 12:57)





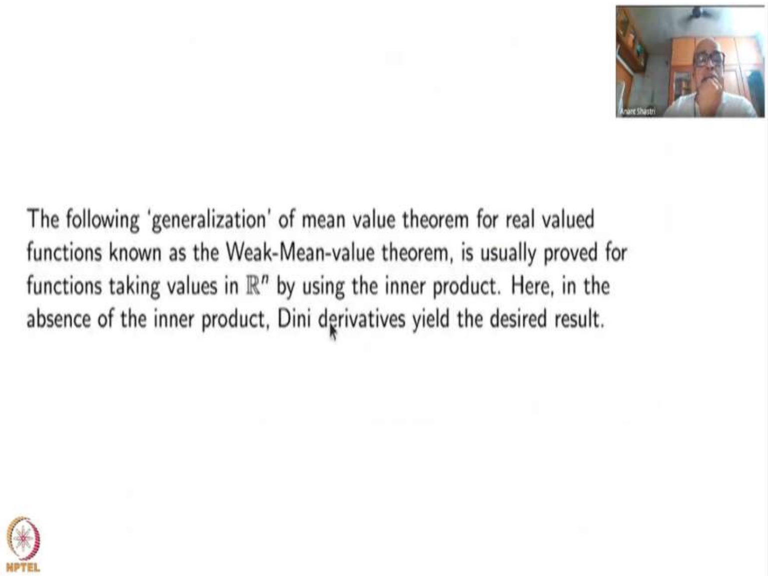

By the way, every time you have to use limsup, you have to do this epsilon business. Given an $\epsilon$ subtract or add, whatever, then something happens. That is the only way to catch this limsup and limpinfs, Not only that even while handling limits also, this is the way. So, $D^{+}(g)$ , (remember $g$ was difference of two functions), equal to $D^{+}(f)$ plus $D^{+}$ of this function, viz, $-(\alpha+\epsilon)x$. But second function is differentiable and what is the derivative? This is linear map. So, the derivative is $-(\alpha+\epsilon)$. So, the $D^{+}$ will be equal to this derivative, which is $-(\alpha+\epsilon)$. So, this $D^{+}(g)$ is nothing but $D^{+}(f)-(\alpha+\epsilon)$. We started with our assumption that this is negative.

$D^{+}$ this is less than $\alpha$ itself is what we have assumed. So, apply this one to a point $y$ inside $[0,x]$. This is true for all points inside this open interval $[0,1]$. Fix an $x$ and then take $y$ inside $[0,x)$. Now I am going to use limsup correctly. If all these supremums were greater than equal to $0$, then the infimum of all of these supremums will be also greater than equal to $0$. So, for some $n$ this supremum must be less than $0$.

Because, $D^{+}$ is infimum over all these supremums, for some $n$, we have the supremum of $(f(y+\delta)-f(y))/d$ must be negative for all $\delta\in(0,1/n)$. This is the same thing as saying, that $f(y+\delta)<f(y)$ for all positive $\delta\in(0,1/n)$. Therefore, $g(y)$ is not the minimum of $g$ in the entire of $[0,x)$. Therefore, the minimum of $g$ must be at $x$. You understand.

The minimum may not be attained in an open interval but we have a continuous function on a closed interval $[0,x]$. Therefore the minimum is attained. But this minimum cannot be inside $[0,x)$. That is what we have seen. Therefore, the minimum must be at $x$ and $g(x)$ must be the minimum value of $g$ in $[0, x]$. And that is precisely what we wanted to say. Therefore, $g(x) \leq g(0)$ in the entire of $[0, 1]$.

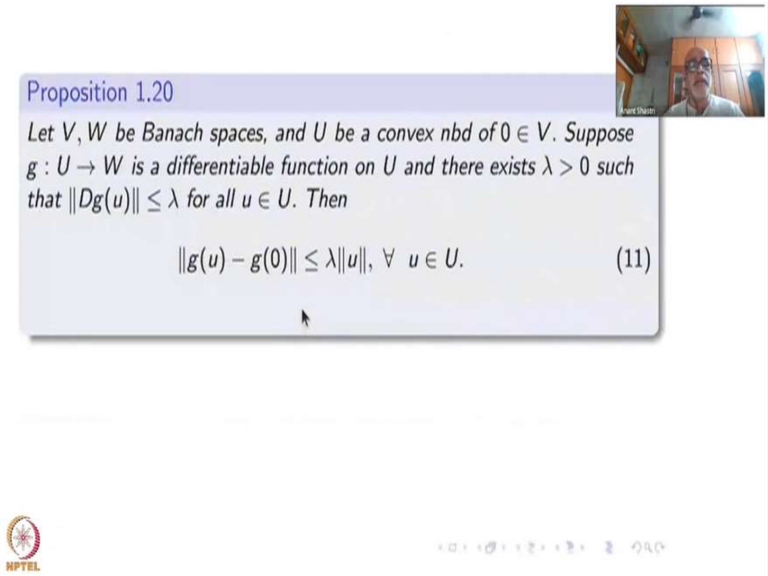
(Refer Slide Time: 17:23)




So, now we are ready to do the generalization of Weak-Mean-value theorem. Why I am calling generalization? Weak-mean-value theorem is true for all differentiable functions on a convex domain in $\mathbb{R}^n$ into $\mathbb{R}^m$, all vector valued functions. Now, we are going to do it on Banach spaces. Same thing, same statement for Banach spaces. This is what we are going to do. The usual proof for this in the case of $\mathbb{R}^n$ and $\mathbb{R}^m$ becomes easier because the Norm-Square function is differentiable on $\mathbb{R}^n$. What is the Norm-Square function?

Namely the Euclidean norm. Euclidean norm-square is just summation $x_i^2$. So, we can use that to do our job. Because all other norms are equivalent to Euclidean norm. But in the general case, we do not have any such theorem. And we do not know, in fact, it is not even true perhaps that given a Banach space with a norm, that norm in our definition may not be differentiable, even the square of the norm may not be differentiable. Putting extra condition that norm being differentiable makes the result too much restrictive. Almost you are begging that it must be a Hilbert space. So, that is why the Dini derivatives are brought in to help us to prove the Weak-Mean-Value theorem.

(Refer Slide Time: 19:17)

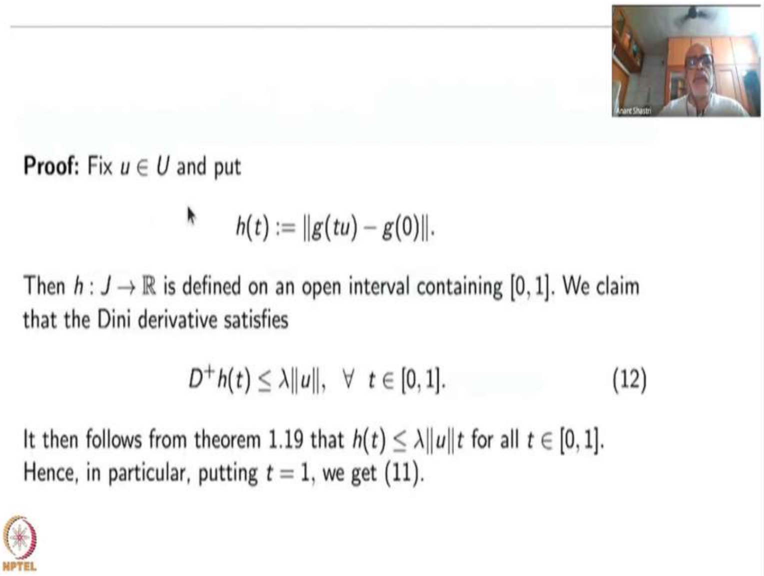


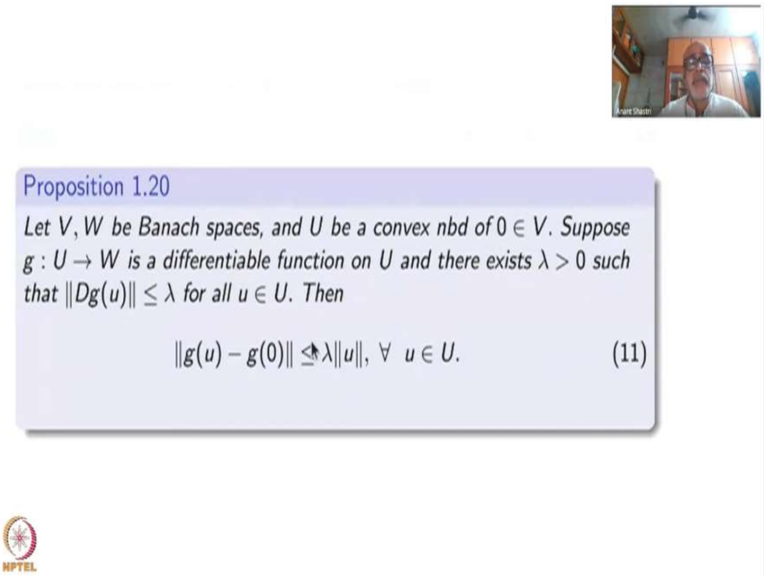

So, the proof itself is not at all difficult now. Let us go through this one. Start with $V$ and $W$, Banach spaces, $U$ is a convex neighbourhood of $0$ belonging to $V$. By the way, I have already told you that this assumption that $0$ belongs to $V$ is just a technical thing, you can do it for any other point.

Suppose $g$ from $U$ to $W$ is a differentiable function and there exists a $\lambda$ positive such that all the derivatives are bounded by $\lambda$: norm of $D(g)$ at all the points $U$ is less than or $\lambda$. Then the $\|g(u) - g(0)\|$ itself is less than or equal to $\lambda\|u\|$ for every $u \in U$. So, this is the weak-mean-value inequality.

In the case of one variable calculus, this was deduced by using the mean-value theorem wherein there is equality, namely, $g(u) - g(0)$ is equal to $g'$ of something in the interval and that $g'$ is bounded by $\lambda$ so, you get this one. But we do not have the mean-value theorem itself, in the case of vector valued functions, but what we have is an inequality, directly. So, this is what we are going to prove in the case of Banach spaces directly. Of course, it will work for any $\mathbb{R}^n$ also because $\mathbb{R}^n$'s are also Banach spaces anyway.

(Refer Slide Time: 21:08)



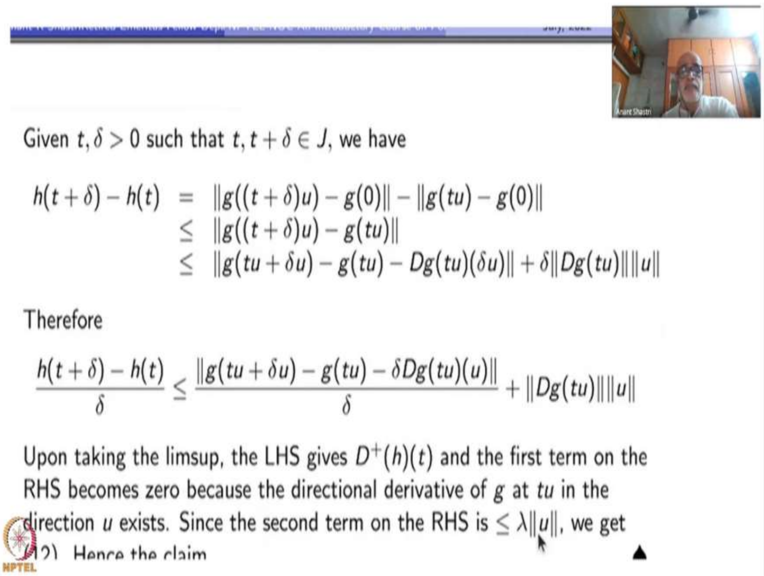



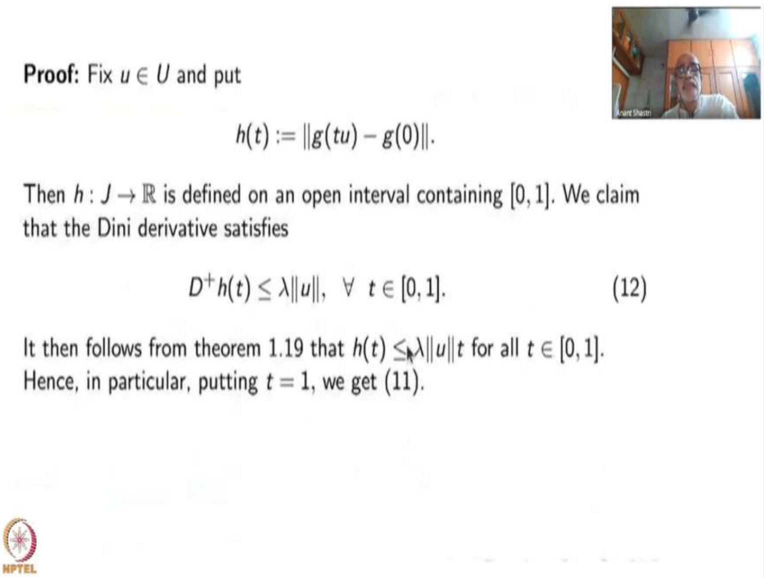

So, fix one point $u \in U$ and define $h(t) = \|g(tu) - g(0)\|$. See if you just take $g(tu) - g(0)$, this is precisely one variable calculus. Only trouble is that this would be a $W$-valued function, not a real valued function. How to get a real valued function? We are left with only taking norm or norm square. Taking norm square does not make it easier. Just take the norm function. This is a continuous function. You do not know whether it is differentiable.

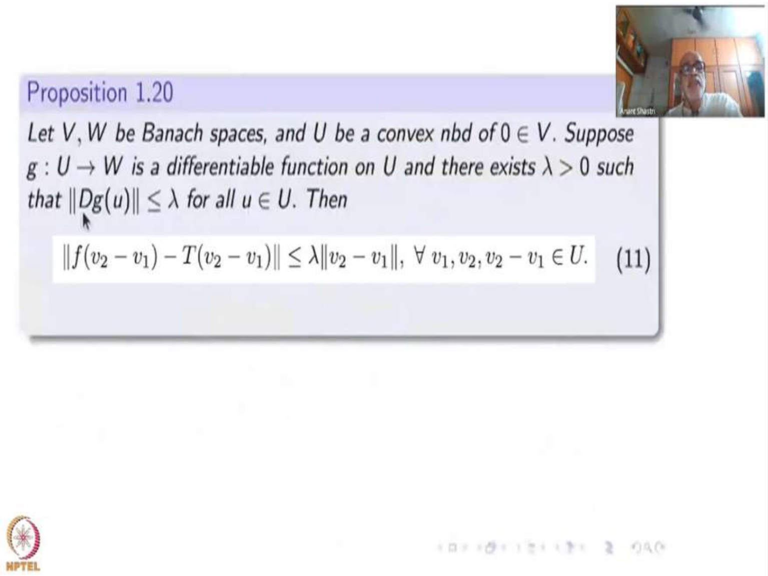
Now, we have got into one variable calculus. $t$ ranges over $[0, 1]$. Because $U$ is convex, when you take $tu$, it will be still inside you. Therefore, $h$ from $J$ to $\mathbb{R}$ is defined on an interval $J$ containing $[0, 1]$. Because, in the open subset $U$ the line segment can be extended just a little bit, on both sides.

We claim that the Dini derivative satisfies the inequality: $D^{+}(h)(t) \leq \lambda u$. So, this was the condition that we needed. So, this $\lambda$ is the same thing as the $\lambda$ in the hypothesis of this proposition. Once we prove this one, you can use our earlier theorem to see that this function $h$ itself is less than to $\lambda\|tu\|$ (the $t$ factor is there after all), for all $t \in [0,1]$. But then you can put $t = 1$, we get is $\|g(u) - g(0)\|$. That will be less than or equal to $\lambda\|u\|$. That is what we wanted to prove. So, we have to prove this inequality, this formula (12). So, this comes very easily now.

(Refer Slide Time: 24:02)






Take $t$ and $\delta$ positive such that both $t$ and $t + \delta$ are inside $J$. Because, $J$ is the interval on which the function is defined, so, you have choose $\delta$ as small as necessary, so that $t + \delta$ is also inside $J$. So, once that is satisfied, I can look at $h(t + \delta)$ as well as $h(t)$ and take the difference, divide by $\delta$ and take the supremum and estimate the $D^+(h)(t)$.

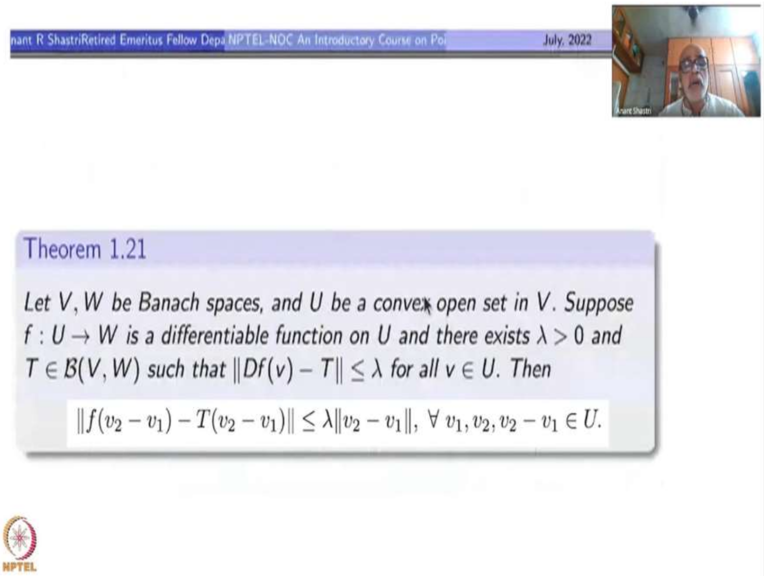
So, start with $h(t + \delta) - h(t)$. That is by definition difference of the norms of $g((t + \delta)u) - g(0)$ and $g(tu) - g(0)$. You know, by triangle inequality, add and subtract) this will be less than or equal to the $\|g((t + \delta)u) - g(tu)\|$. (If you take this term on the other side, it becomes easy to see the triangle inequality).

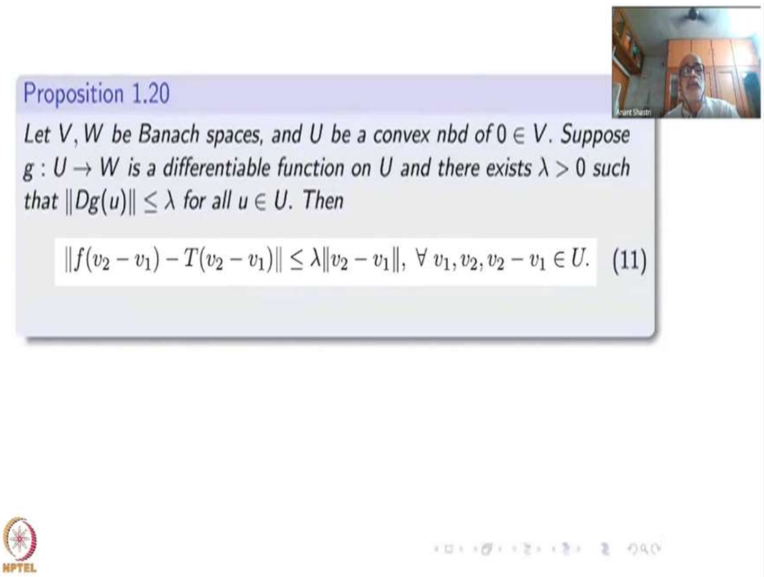
Now, once again you add and subtract, $D(g)(tu)(\delta u)$, so that the norm will be less than equal to norm of this plus norm of this second factor which I have subtracted here. So, there this delta comes out, $\delta$ is positive. It is nothing but norm of $D(g)(tu)$ into norm of $u$. So, norm of this is less than or equal to that. I have pulled out this $\delta$, because, now, I can divide out by this $\delta$. When you divide out by this delta on the left-hand side what you have is the corresponding term for our $D^+$ definition. $h(t + \delta) - h(t)/\delta$ which is less than or equal to some of two terms.

Now you take the limit limsup of the LHS. That is the Dini dervative. On the RHS, the first term is the norm of term occuring in the definition of the directional derivative of $g(tu)$ in the direction $u$. Since the derivative exists, the directional derivative also exists and hence the limsup of this term becomes zero. The second term is independent of $\delta$.

So, what you can conclude is that $D^{+}(h)$ is less than or equal to this second term which by hypothesis, less than or equal to $\lambda\|u\|$. So, what we have proved is this mean value inequality for Banach spaces for all differentiable functions on a convex open set.

(Refer Slide Time: 28:49)



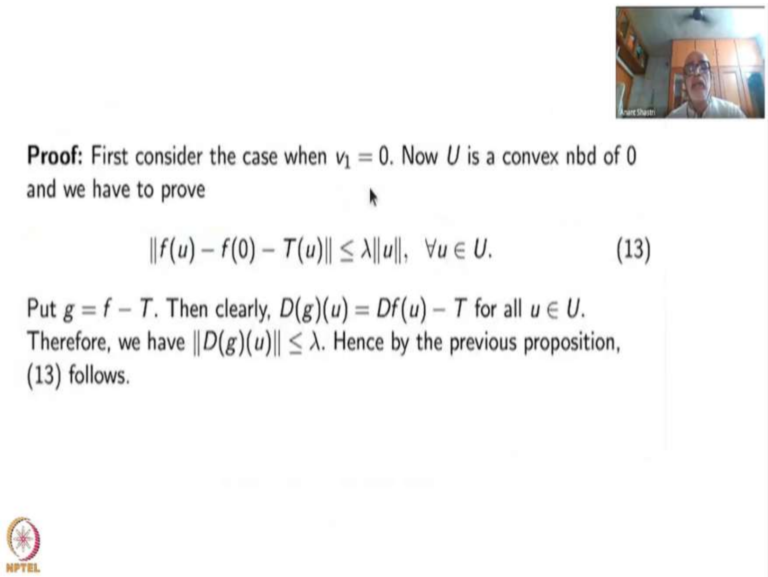

So, let us now, convert this into the following theorem, a ready-to-use result for our implicit function theorem etc.

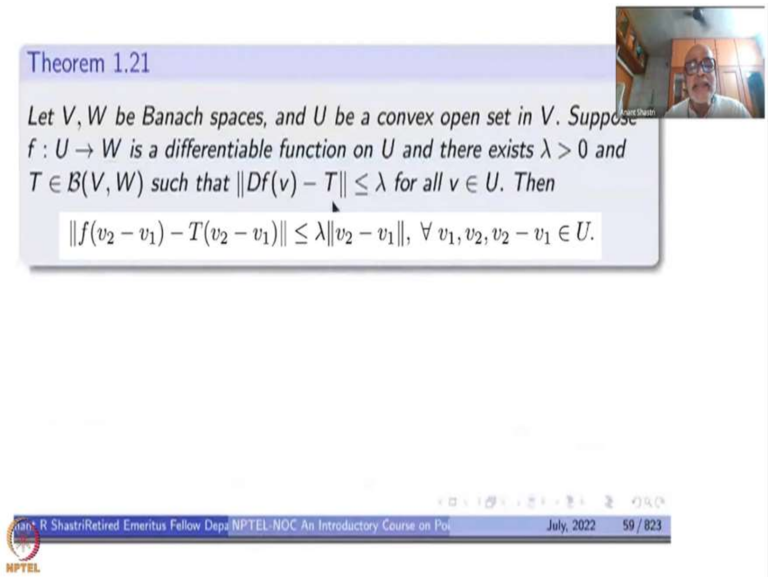
Let $V$ and $W$ be Banach spaces, $U$ is a convex open subset of $V$, $f$ is a differentiable function on $U$ taking values in $W$, $\lambda$ is a positive constant and $T$ is a some bounded linear function, $T$ in $\mathcal{B}(V, W)$, (see, so far everything is same as the previous proposition, but here now, I am bringing an arbitrary bounded linear function $T$ from $V$ to $W$) such that $\|Df(v) - T\| \leq \lambda$

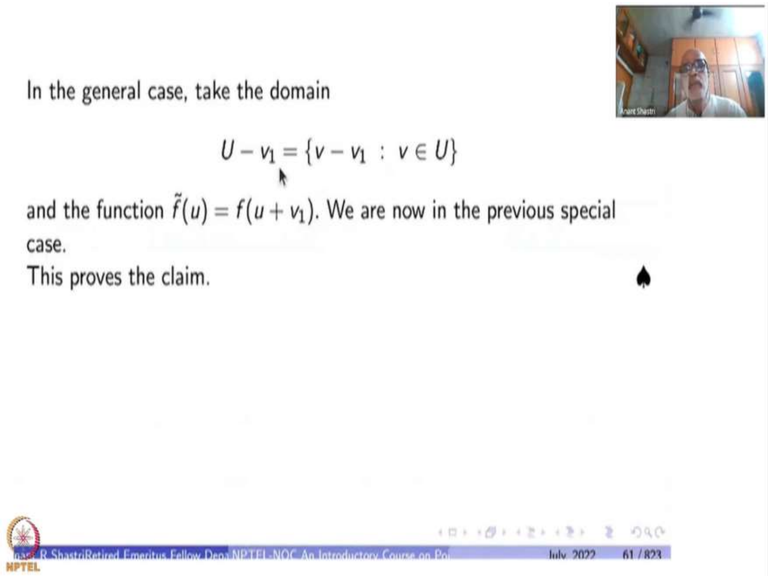
for all $v \in U$. (If you put $T$ equal to $0$ then you get the earlier case. So, this is an extension of the earlier case.)

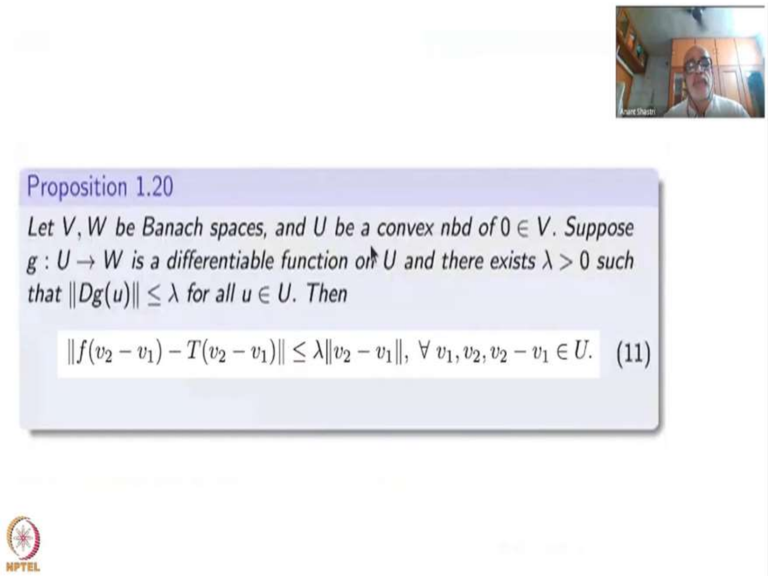
Then the conclusion is (also slightly more general namely,) that the $\|f(v_2) - f(v_1) - T(v_2 - v_1)\| \leq \lambda \|v_2 - v_1\|$ for all $v_1, v_2$ inside this convex neighbourhood. This $U$ itself is a subset of the Banach space $V$. So, this is very easy now. Of course, using the previous result. We do not need to go back to any more Dini derivatives.

(Refer Slide Time: 30:40)





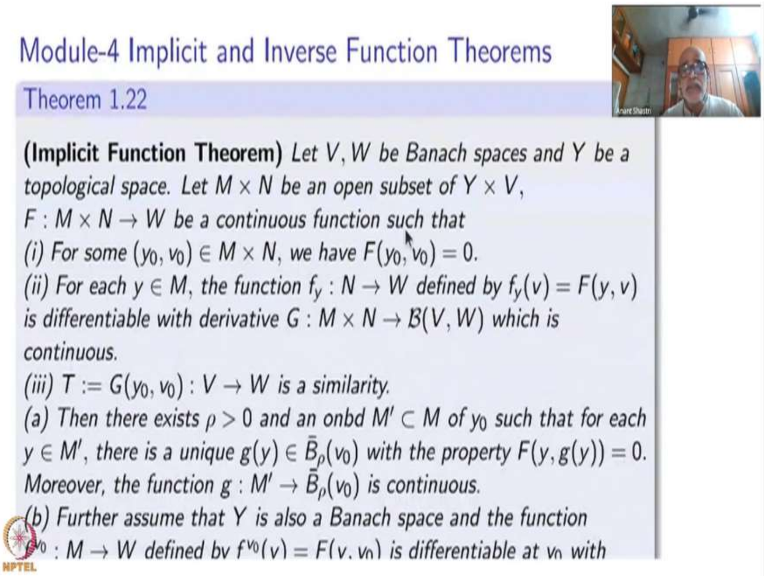



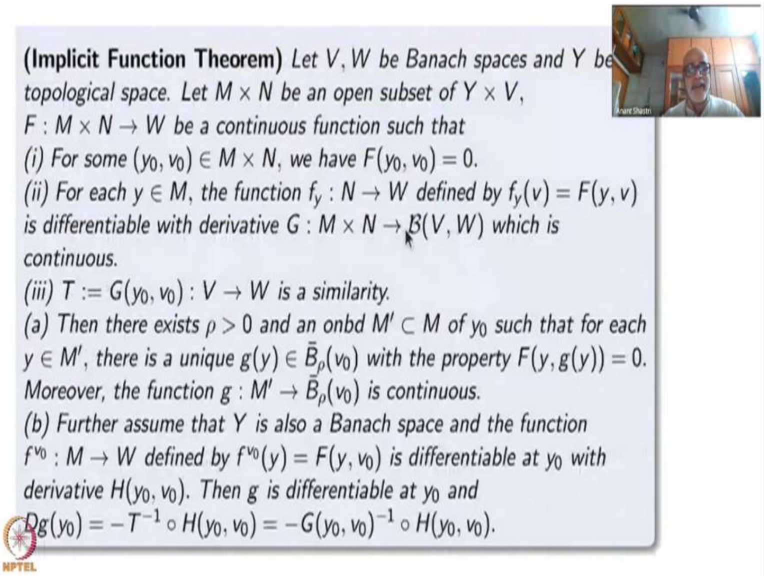

So, first consider the case when $v_1$ itself is 0. This is my assumption. This may not be the case but I am taking a special case. Now, $U$ is a convex neighbourhood of 0 because $v_1$ is after all an element of $U$. So, we have to prove that $\|f(u) - f(0) - T(u)\| \leq \lambda \|u\|$ for every $u \in U$, by replacing $v_2$ with $u$.

For this, what I do is to put $g = f - T$. Because in the hypothesis here $Df(u) - T$ is bounded. So, I take $g = f - T$. What is the derivative of $g$? It is $Df(u) - T$. Therefore, we can apply the previous proposition. Now one more step. I have to remove the simplification we have made, namely, I have assumed $v_1$ is 0. So, how to do this in the general case?

You first take the domain itself to be $U - v_1$; translate $U$ by $v_1$. (That is shift the origin to $v_1$). That means we now consider the open convex set consisting of all points $v - v_1$ where $v$ is inside $U$. On this one, you change the function also now, namely, instead of $f$, $\tilde{f}(u) = f(u + v_1)$. $v_1$ is a point of $U$. Therefore $0$ is a point of $U - v_1$. $U$ is convex therefore $U - v_1$ is convex. So, you can apply the previous conclusion for $\tilde{f}$. What you get is precisely the required statement now.

(Refer Slide Time: 34:40)




So, next time we shall prove the implicit function theorem just for the sake of what we are up to, I will just begin just statement and then we will prove this one next time. So, implicit function theorem, the statement itself is somewhat long here. The proof is not as long as that but, you know, you do not have to be threatened by the big statement.

What are the hypothesis here? $V$ and $W$ are Banach spaces. $Y$ is any topological space. So, that is part of the generality that we have achieved here. Take $M \times N$ to be any open subset of $Y \times V$. In other words, $M$ is an open subset of $Y$ and $N$ is an open subset of $V$; $F$ from $M \times N$ to $W$ is a continuous function such that :

(i) for some point $(y_0, v_0)$ belonging to $M \times N$, we have $F(y_0, v_0)$ is 0;

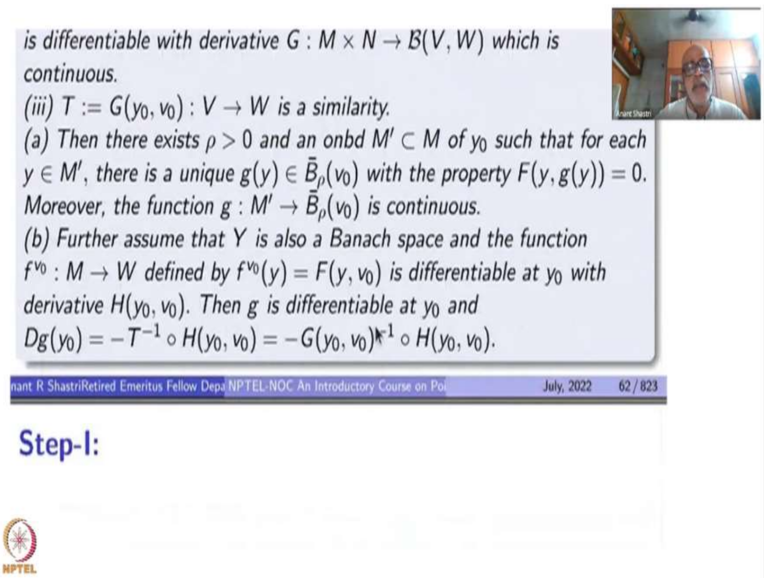
(ii) for each $y \in M$, the function $f_y$ from $N$ to $W$ given by $f_y(v) = F(y, v)$, namely, $y$ is fixed now to get a function of just the variable $v$, $f_y$ is from $N$ to $W$. That is differentiable. Its derivative will be a bounded linear map $D(f_y)$ from $V$ to $W$ and the function $G$ from $M \times N$ to $W$ given by $(y, v) \mapsto D(f_y)(v)$ is continuous.

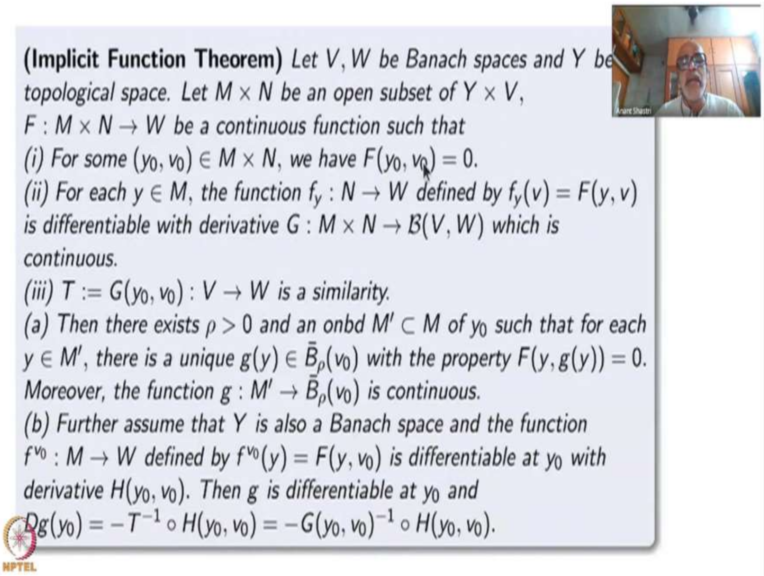
So, for each fixed $y$, $D(f_y)$ it is continuous. That is not enough. $G$ as function of two different variables $y$ and $v$ must be jointly continuous. Finally,

(iii) the derivative at $(y_0, v_0)$, $T = G(y_0, v_0)$ is a similarity from $V$ to $W$. This hypothesis is very important.

In particular, you know that $V$ and $W$ are similar. Other hypothesis is $F$ is $t$ times differentiable with respect to $v$ and this partial derivatives of $F$ are jointly continuous. The hypothesis $F(y_0, v_0) = 0$ is only a technical one.

(Refer Slide Time: 38:01)






With these hypothesis, now, the conclusions. There are two conclusions here. For the second conclusion we will need little more hypothesis that is why it is separated out in the statement.

(a) The first conclusion is that there is a neighbourhood of $(y_0, v_0)$ in $M \times N$, more specifically, a smaller neighbourhood $M'$ of $y_0 \in M$, some smaller and a positive real number $\rho$ (that would give a neighbourhood of $v_0 \in V$) such that for all $y \in M'$ there exists a $g(y)$ belonging to $\bar{B}_\rho(v_0)$, the closed ball of radius rho around $v_0$, with the property that $F(y, g(y)) = 0$. So, you see $F(y_0, v_0) = 0$. This of $f(y, v) = 0$ as an equation which we want to solve. One solution has been given. Then you want a continuous solution here on a small neighbourhood. So, that is precisely what is achieved here. In fact, when you are trying to get a solution in a neighbourhood, you are already ensured of a unique solution. In the solution of all existence theorems, the uniqueness part helps you a lot.

So somehow, we have got this uniqueness here also. You see the only thing to catch is that in the conclusion, we have to cut down the domain properly on which we have no control. Some open subset we do not know, some row we do not know. On the closed ball for each point inside $M'$, $g(y)$ will be unique. One is not satisfied with just continuity of $g$. We would like to have differentiability also. For that we have to put a little more hypothesis.

Since $M'$ is a subset of a topological space $Y$, it does not make sense to demand that $g$ must be differentiable. So, in order to make that sense we have to say that $M$ is an open subset of a some normed linear space. So, more safely, we assume that $Y$ is a Banach space in the next part

(b) So, further assume that $Y$ is also a Banach space and the function $f^{v_0}$ from $M$ to $W$ defined by $f^{v_0}(y) = F(y, v_0)$ is differentiable at $y_0$.

You see I have used now upper index here to indicate that now, I am fixing the right-hand slot; $v = v_0$ is fixed and $y$ is the variable. But the function is same a function $F$ is restricted to $v = v_0$ and $y$ is varying that must be differentiable. We are not demanding that the function $F$ itself from $M \times N$ to $W$ is differentiable from the product space. This is a much stronger condition. But we are assuming that it is continuous. Whereas differentiability is partial with respect the two variables $y$ and $v$ separately. Fix $y_0$, you get a differentiable function, fix $v_0$ you get another differential function. That is all we are demanding. So, this part is that $f(v_0)$ where $v_0$ is fixed that must be differentiable. Denote its derivative by $H(y_0, v_0)$.

Then what happens? $g$ will become differentiable at $y_0$. What is this $g$? The unique solution given by the part (a). That will be differentiable at $y_0$ and its derivative is given by $-T^{-1}H$. See $H$ is bounded linear map, $T$ is a bounded linear map. This $T$ is invertible, it is a similarity. So, I can talk about $T^{-1}$. So, $D(g)(y_0) = -T^{-1}H(y_0, v_0) = -G(y_0, v_0)H(y_0, v_0)$.

So, let us prove this statement next time. Thank you.