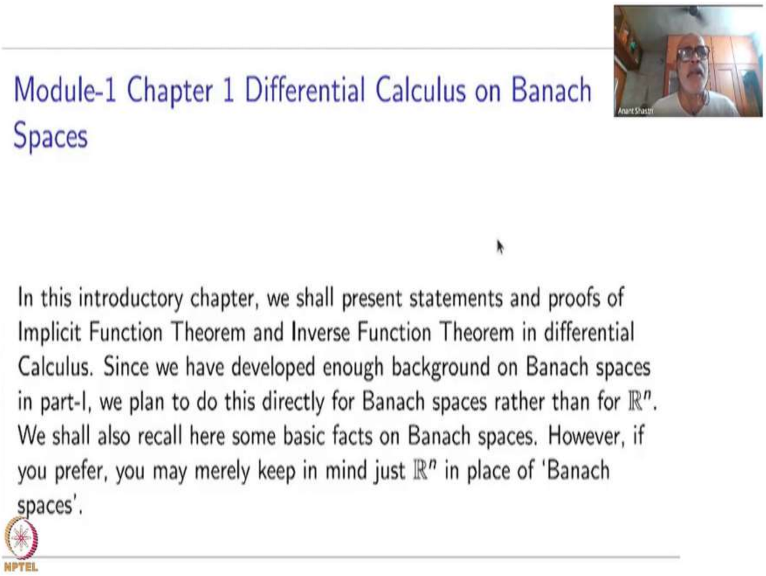
An Introduction to Point Set Topology Part 2
Professor Anant R Shastri
Department of Mathematics
Indian Institute of Technology, Bombay
Lecture 01
Preliminaries from Banach Spaces

(Refer Slide Time: 0:17)



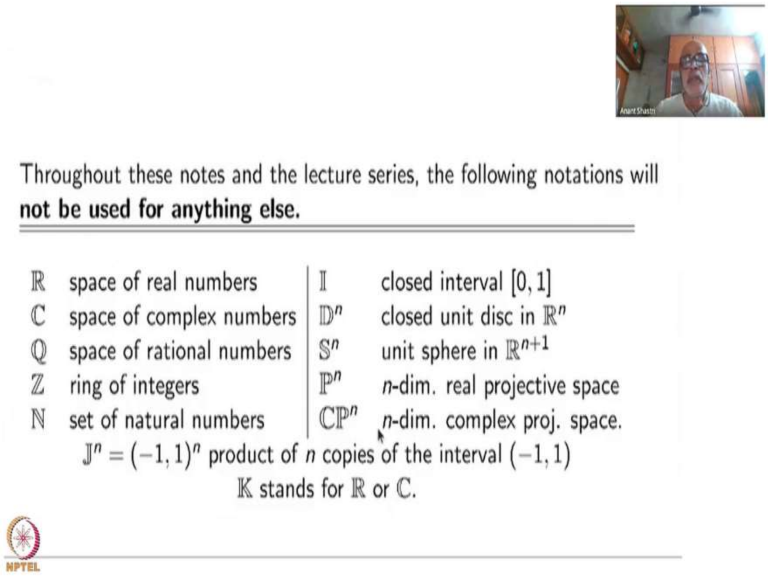
Welcome to NPTEL NOC course on Point-Set-Topology part II. This is module 1/ chapter 1 on differential calculus on Banach spaces. In this introductory chapter we shall present statements and proofs of implicit function theorem and inverse function theorem in differential Calculus.

Since we have developed enough background on Banach spaces in part-I, we plan to do this directly for Banach spaces rather than for $\mathbb{R}^n$. Usually for $\mathbb{R}^n$, one can prove inverse function theorem first, and then prove implicit function theorem which becomes a little more transparent. In the case of general Banach spaces such a method is not possible. We have to first prove the implicit function theorem and then deduce inverse function theorem.

One of the key factors which needs to be sharpened in the case of Banach spaces is the so-called weak mean value theorem. In the case of $\mathbb{R}^n$, because it has a rich structure namely Hilbert's structure the proof is much simpler. Here the proof a uses little deeper result in analysis viz., what are called Dini derivatives.

If you have some difficulty in understanding Banach spaces you can just replace all occurrence of Banach spaces by $\mathbb{R}^n$ and just think of $\mathbb{R}^n$ and try to follow the material. Afterwards you can fill-in your Banach spaces once you learn Banach spaces thoroughly.

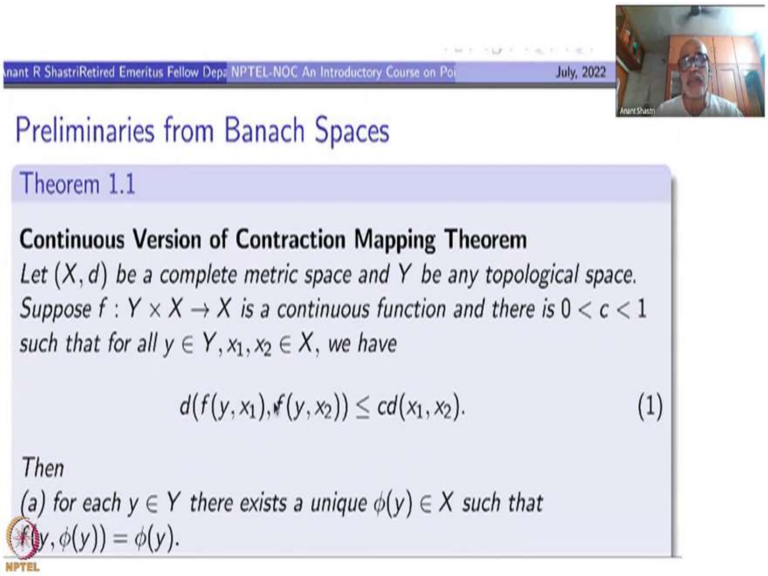
(Refer Slide Time: 2:50)



Let me recall some notations which I have already introduced to you. These notations will not be used elsewhere during this course. So, they are kind of frozen especially for this chapter. These are all standard ones of course, the Euler fonts $\mathbb{C}, \mathbb{Q}, \mathbb{Z}, \mathbb{N}$ and in addition, I was also using this Euler font $\mathbb{I}$, for close interval $[0, 1]$, and $\mathbb{D}^n$ for the closed unit disc, $\mathbb{S}^n$ for the unit sphere. And finally, $\mathbb{P}^n$ and $\mathbb{CP}^n$ come very rarely but are standard notations. I do not use them for any other thing.

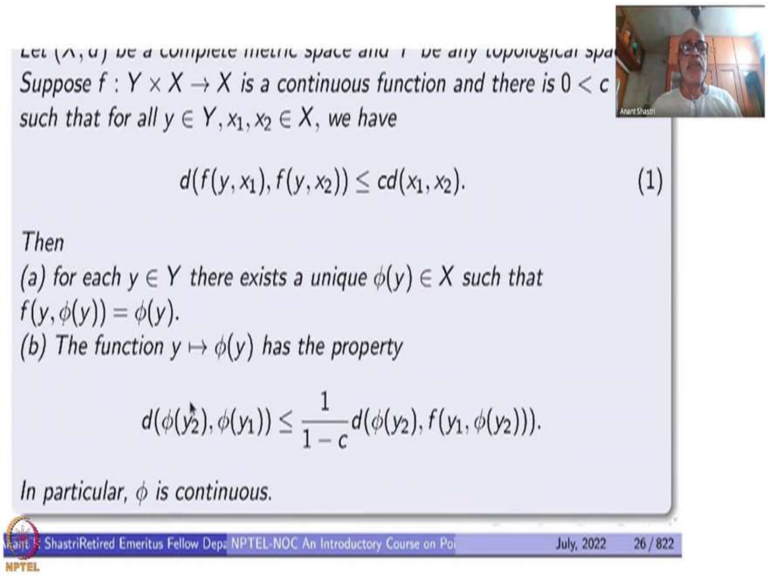
Sometimes I need the open interval $(-1, 1)^n$ for which I will use this (Euler font) $\mathbb{J}^n$. Most important one is that often I have to deal with both real and complex fields simultaneously. So, in that case I will use this notation (Euler font) $\mathbb{K}$. In the context of a special case, we will mention which one is being used, otherwise this $\mathbb{K}$ will be either the field of real numbers or complex numbers.

(Refer Slide Time: 4:18)



So, let me begin with a modification of the contraction mapping theorem that we have proved in part-I. A modification which is an extension actually. There, we had proved it for one map. Here we will prove it for a family of maps that is the important difference. So, start with a complete metric space (X, d) and let Y be any topological space. Take a function on $Y \times X$ to X which is continuous, so that there is a real number between 0 and 1, strictly between 0 and 1 such that this property holds: the distance between $f(y, x_1)$ and $f(y, x_2)$ is less than or equal to c times the distance between x_1 and x_2 for all $y \in Y$ and for all $x_1, x_2 \in X$. So, this is some kind of uniform condition uniform-continuity condition.

(Refer Slide Time: 5:32)



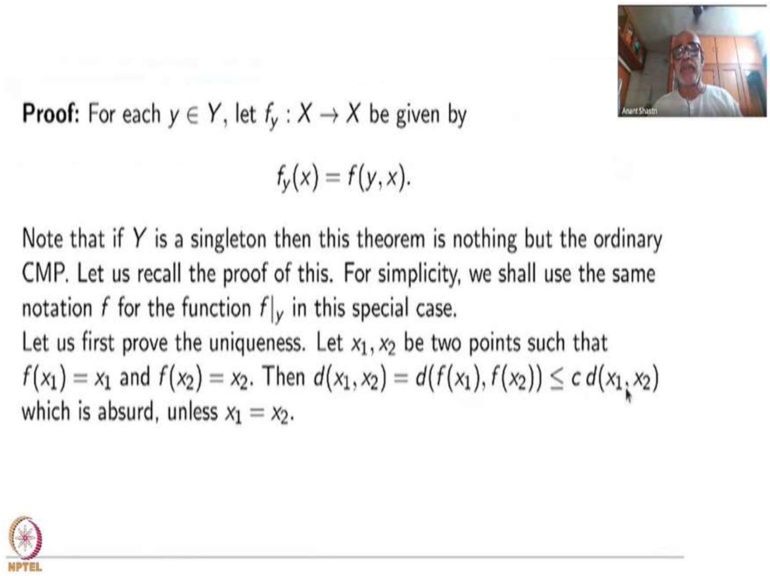
If this condition holds then we have a conclusion. What is the conclusion? For each $y \in Y$, there exists the unique $\phi(y) \in X$ such that $f(y, \phi(y)) = y$. You can think of this as

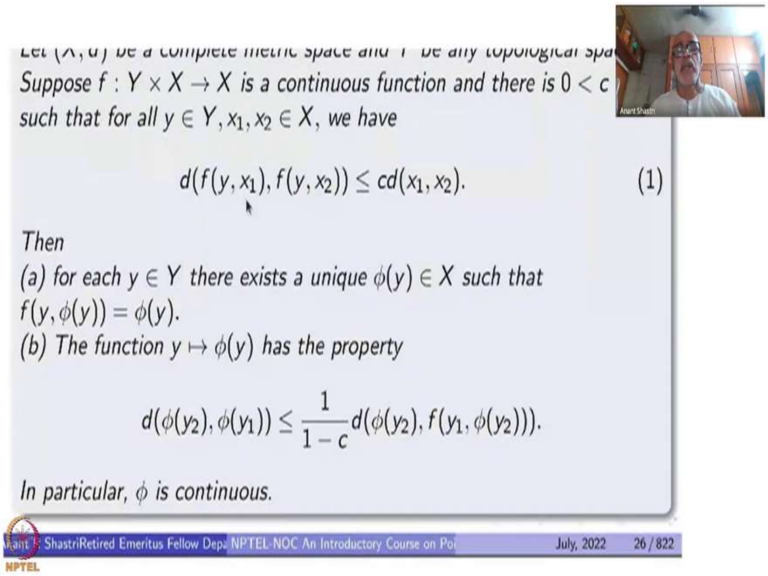
$f(y,x) = x$, by replacing $\phi(y)$ by x which is a unique point of X. This x will be called as $\phi(y)$ because it is a unique one and depends on y.

So, thus we get an assignment from Y to X, a function ϕ. This ϕ has the property that the distance between $\phi(y_2)$ and $\phi(y_1)$ is less than or equal to $1/(1-c)$ times $d(\phi(y_2), f(y_1, \phi(y_2)))$. That same c here. This is a technical result that will be very helpful. In particular you can immediately see that phi is continuous.

This distance between $\phi(y_1)$ and $\phi(y_2)$ is dominated by this one on the right hand side. So, if y_1 and y_2 can be controlled then this can be controlled therefore this left-hand side can be controlled is the conclusion here. If you forget about this capital Y here (by taking it to be a single point) and look at just a single function from X to X, one single function then we have proved this statement viz., there is only part (a) namely there is a unique point which is the fixed point of the contraction mapping. So, the additional statement (b) here is there when you have a family of functions indexed by the space Y, which satisfies this `uniform continuity condition. We get a continuous function out of it. It is stronger than saying just continuous So, that property we have put here and we will use this one once again. So, let us go through the proof of this one all right?

(Refer Slide Time: 8:07)



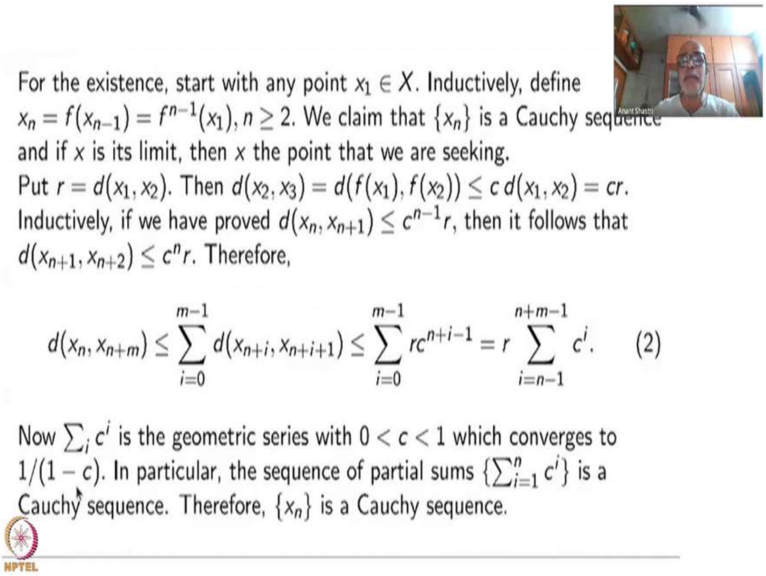


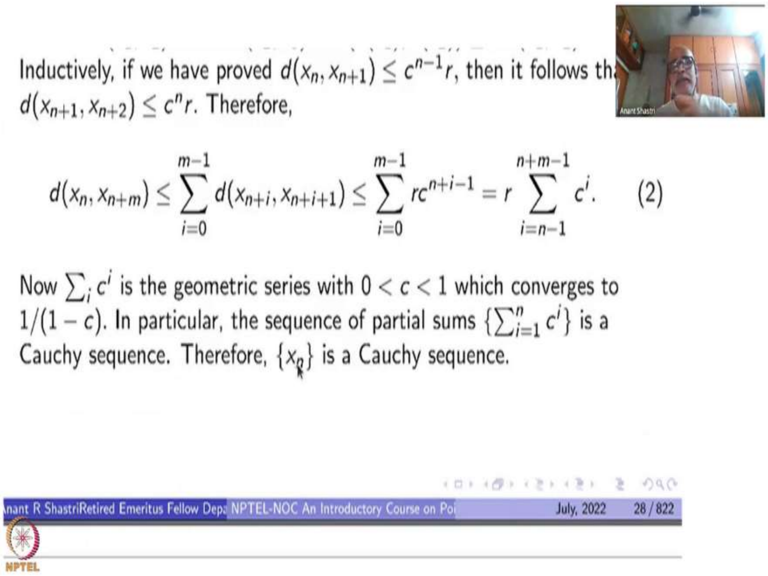
First of all, just recall the proof in the case of when there is a single function only. For each $y \in Y$, let us look at f_y, one single function X to X given by $f_y(x) = f(y, x)$. The y coordinate is fixed so you get one single function. For this let us say what is the proof of this theorem which we have done already. That is if Y is a singleton space, then this theorem is nothing but the ordinary contraction mapping principle.

So, let us recall the proof. For simplicity we shall use the same notation f for the function f restricted to y namely f_y in this special case. This singleton space $Y = \{y\}$, so in this special case let us first prove the uniqueness. Let x_1 and x_2 be two points such that $f(x_1) = x_1$ and $f(x_2) = x_2$.

Then distance between x_1 and x_2 will be distance between $f(x_1)$ and $f(x_2)$ because x_1 is $f(x_1)$ and x_2 is $f(x_2)$. But this is less than c times the distance between x_1 and x_2. Now y is suppressed here, that is all. Distance between $f(x_1)$ and is less than c times the distance between x_1 and x_2. This is condition 1. So, condition 1 will tell you that we have distance between x_1 and x_2 is less than or equal to a fraction of the same distance and this is possible only if this this real number is 0. That means that x_1 is x_2. This is the way the uniqueness part was proved and we have just recalled it here.

(Refer Slide Time: 10:03)





For the existence part what we do? We follow the iteration method. Start with any point x. Inductively define $x_2 = f(x_1), x_3 = f(x_2)$ and so on $x_n = f(x_{n-1})$, which is nothing but f operating n-times on x_1, i.e., $f^{n-1}(x_1)$.

We claim that the sequence got by iterating the powers of f on x_1 namely, this $\{x_n\}$ is a Cauchy sequence. Then we will appeal to the fact that the metric space is complete, to get a limit of the this sequence. That limit, if we denote by x, then we will show that this x is the fixed-point of f, i.e., $f(x) = x$. So, how does one prove that this is a Cauchy sequence? Let us for the sake of simplicity put r equal to the distance between x_1 and x_2.

What is x_2? x_2 is $f(x_1)$. If $f(x_1)$ is already x_1, then this distance will be 0. But then we have already solved this problem we do not have to go any further. Never mind. So r may be 0 never mind. But whatever we have if r is not 0 namely $f(x_1)$ is not equal to x_1, then only we have to iterate. So, we keep iterating. So do not worry about it right now whether it is equal to the old one or not. If it is equal to the old one by chance, then you can stop there. That is no problem. Anyway, $d(x_2 x_1) = r$ and $d(x_2, x_3) = d(f(x_1), f(x_2))$ which is less than or equal to $cd(x_1, x_2) = cr$.

Now, you repeat this one: $d(x_3, x_4)$ less than or equal to $c^2 r$ and so on. Distance between x_n and x_{n+1} will be less than or equal to $c^{n-1}r$. Therefore, assuming this inequality, you repeat once more to get distance from x_{n+1} and x_{n+2} is less than or equal to goes to $c^n r$. So same formula is there now.

Therefore, to compute the distance between any x_n and x_{n+m} now, I use the triangle inequality m times. So, I go for x_n to x_{n+1}, then x_{n+1} to x_{n+2}; x_{n+2} to x_{n+3} etc. take all these distances, add them up put a less than equal to sign, summation from 0 to $m-1$ of distance between x_{n+i} and x_{n+i+1}. But just now we have proved this formula this distance is rc^{n+i-1} is equal to r times the summation from 0 to $m-1$ of c^{n+i-1}.

What are these? These are nothing partial sums of the series $1 + c + c^2 + c^3 + \ldots$ which is a geometric series, where c is between 0 and 1. So now summation c^i is a geometric series which converges to $1/(1-c)$. So, that explains why we have got this $1/(1-c)$ here in this inequality.

So, continuing with the proof of this Cauchy sequence; in particular the partial sums summation from 0 to n of c^i is a Cauchy sequence. Therefore $\{x_n\}$ is also a Cauchy sequence. It converges because X is a complete metric space.

After that if you apply limit of the sequence $\{x_n\}$ is the same thing as limit of the sequence $\{x_{n+1}\}$ which is the same as the sequence $\{f(x_n)\}$. f is a continuous function. So, I can take out f from the limit and hence I get $f(x) = x$.

(Refer Slide Time: 15:42)

Let $\lim_n x_n = x$. (This is where we use the completeness of X.) Since f is continuous

$$f(x) = \lim_n f(x_n) = \lim_n x_{n+1} = x.$$

By taking the limit as $m \to \infty$ in (2) it follows that

$$d(x_n, x) \le rc^{n-1}/(1-c), \quad \forall n \ge 1. \qquad (3)$$

Returning to the general case, since the fixed point $x \in X$ depends upon the function $f_y : X \to X$, let us denote it by $\phi(y)$. Also, since r depends

NPTEL

Inductively, if we have proved $d(x_n, x_{n+1}) \le c^{n-1}r$, then it follows th $d(x_{n+1}, x_{n+2}) \le c^n r$. Therefore,

$$d(x_n, x_{n+m}) \le \sum_{i=0}^{m-1} d(x_{n+i}, x_{n+i+1}) \le \sum_{i=0}^{m-1} rc^{n+i-1} = r \sum_{i=n-1}^{n+m-1} c^i. \qquad (2)$$

Now $\sum_i c^i$ is the geometric series with $0 < c < 1$ which converges to $1/(1-c)$. In particular, the sequence of partial sums $\{\sum_{i=1}^n c^i\}$ is a Cauchy sequence. Therefore, $\{x_n\}$ is a Cauchy sequence.

NPTEL

$$d(x_n, x_{n+m}) \le \sum_{i=0}^{m-1} d(x_{n+i}, x_{n+i+1}) \le \sum_{i=0}^{m-1} rc^{n+i-1} = r \sum_{i=n-1}^{n+m-1} c^i. \qquad (2)$$

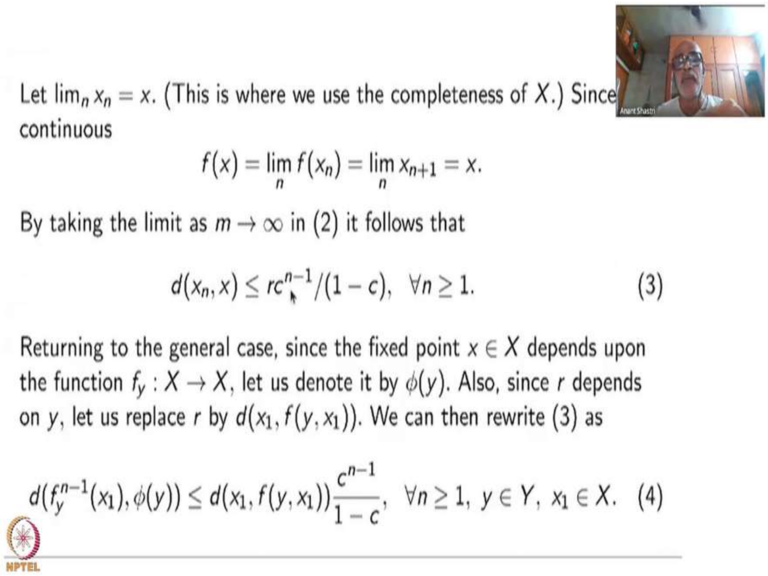
Now $\sum_i c^i$ is the geometric series with $0 < c < 1$ which converges to $1/(1-c)$. In particular, the sequence of partial sums $\{\sum_{i=1}^n c^i\}$ is a Cauchy sequence. Therefore, $\{x_n\}$ is a Cauchy sequence.

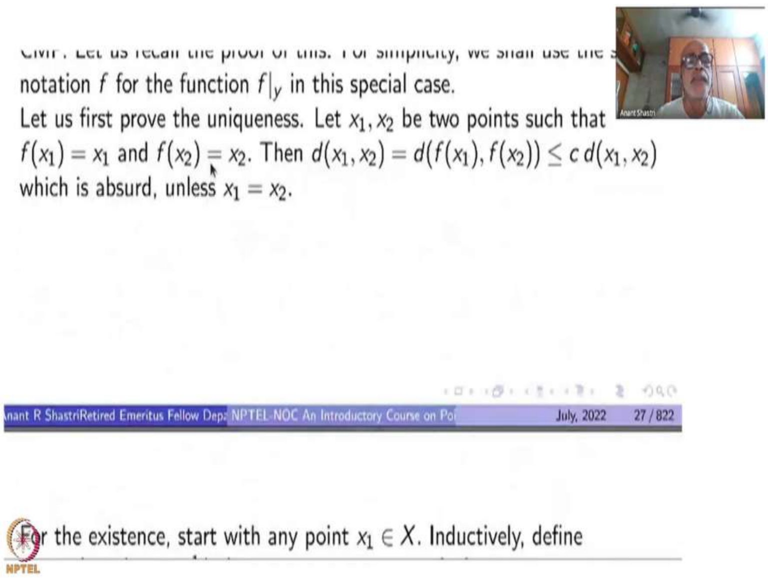
Let $\lim_n x_n = x$. (This is where we use the completeness of X.) Since f is

NPTEL

So, now if you take he limit as m tending to infinity in the inequality (2) on the LHS you get $d(x_n, x)$. What is the limit of this summation on the right? What you get is distance of x_n and x is less than equal to rc^{n-1} divided by $(1-c)$, the partial sums from the this is the remainder after in terms.

(Refer Slide Time: 16:26)





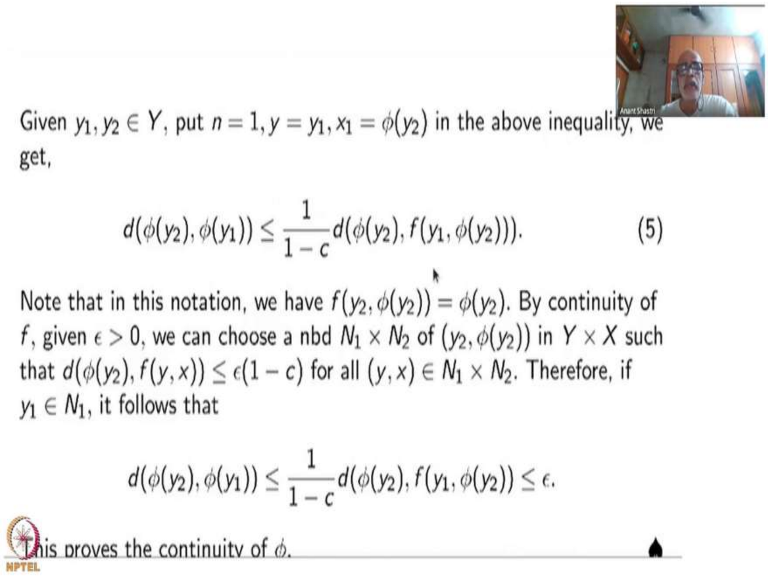
So, for all n greater than equal to 1 this is true. So, now we copy the same thing by taking the variable y also. The idea of the proof is exactly the same, there is no change at all.

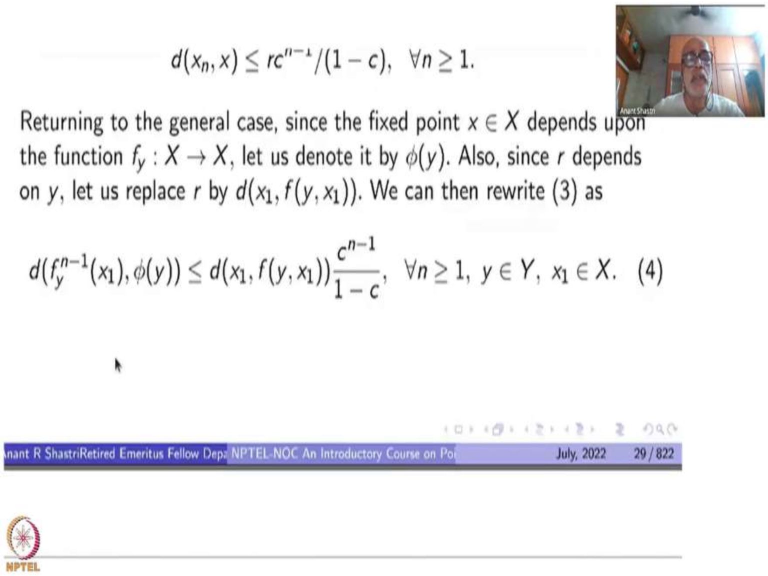
Returning to general case, since the fixed-point x of f_y depends upon y, we are changing its notation: for each y, I am getting a fixed point which I will denote by $\phi(y)$.

Also, while proving the existence, this r which we have fixed as the first distance between x_1 and x_2 earlier, this will now depend upon y because this is now r is the distance between x_1 and $f(y, x_1)$. So, let us denote this r_y, let us replace r by r_y in the formula (3).

We can then rewrite this (3) as distance between... you see this now x_n is $f_y^{n-1}(x_1)$ and so the distance between $f^{n-1}(x_1)$ and $\phi(y)$(which is the fixed point of f_y) is less than or equal to $r_y c^{n-1}/(1-c)$. This is independent of y, the constant c is independent of y remember that. This holds for every $n \geq 1$, for all $y \in Y$ and $x_1 \in X$.

(Refer Slide Time: 18:29)





So, now given y_1 and y_2 in Y, put $n = 1, y = y_1$ and $x_1 = \phi(y_2)$ in the above inequality to get this this one. Here, $n = 1$ so this will be 1, $r_y = d(x_1, f(y, x_1)) = d(\phi(y_2), f(y_1, \phi(y_2)))$. That is we have $d(\phi(y_2), \phi(y_1))$ is less than or equal to $d(\phi(y_2), f(y_1, \phi(y_2))/(1-c)$.

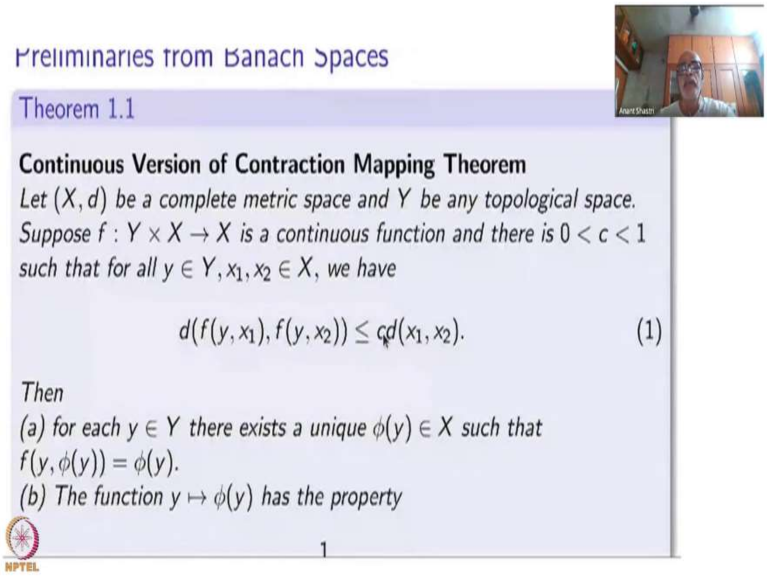
So, this is what we wanted to prove this is a conclusion of the part of theorem. Note that in this notation, we have $f(y_2, \phi(y_2)) = \phi(y_2)$ and similarly, $f(y_1, \phi(y_1)) = \phi(y_1)$, and so on, $f(y, \phi(y) = \phi(y)$ for all y.

By continuity of f, given $\epsilon > 0$, we can choose a neighborhood $N_1 \times N_2$ of this point $(y_2, \phi(y_2))$ in $Y \times X$ such that the distance between $\phi(y_2)$ and $f(y, x)$ is less than $\epsilon(1 - c)$ for all $(y, x) \in N_1 \times N_2$. (Formula (5).) This is just by continuity of f from $Y \times X$ to X. The point $(y_2, \phi(y_2))$ is taken to $\phi(y_2)$ by f. So a neighbourhood will this point will be taken to a neighbourhood of $\phi(y_2)$.

Because, in the domain you have Y which is an arbitrary space cross x which is of course a metric space, so I am writing the neighborhood as $N_1 \times N_2$ instead of choosing ball neighborhoods and so on of the point $(y_2, \phi(y_2))$. This is inside $Y \times X$. This entire neighborhood is taken f inside this ϵ' neighborhood of $\phi(y_2)$ by f, and I am making ϵ' to be $\epsilon(1 - c)$.

So, this is you can choose ϵ' here and then write ϵ' is $\epsilon(1 - c)$ no problem. The distance between $\phi(y_2)$ and any $f(y, x)$ as soon as y and x are inside this neighborhood is less than this one. So, this is continuity of f therefore now using the continuity of f here for $y \in N_1$, it follows that distance between $\phi(y_2)$ and $\phi(y_1)$ is $1/(1 - c)$ times this one but then $1 - c$ cancels out you are left with ϵ. Just to cancel out $1 - c$ factor, I had chosen ϵ' that way here above. That is all. So, this is last part (b). This helps to derive the continuity of the function phi that we have obtained as a solution function for the fixed point.

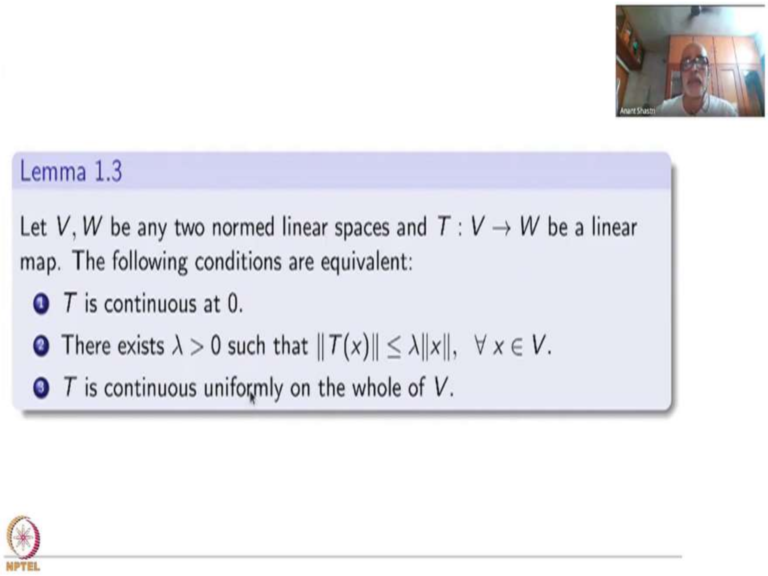
(Refer Slide Time: 22:15)

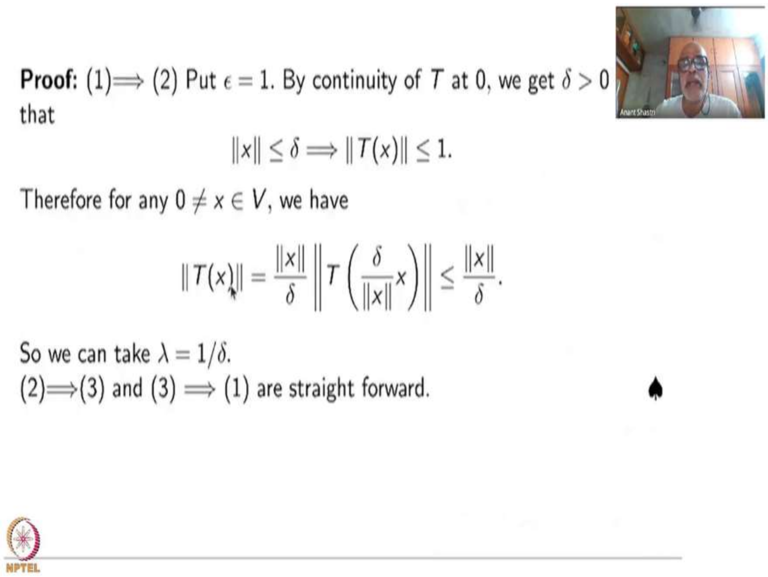


So, I am repeating condition (1) may be referred to as uniform contraction mapping condition. Why? Because in the right-hand side here, the choice of c does not depend upon y at all, for all $y \in Y$, you have this the same c. That is why I told you that this is like uniform Lipchitz. So, the theorem itself can be called as uniform contraction mapping.

Any questions?

(Refer Slide Time: 22:55)





Let us now recall a few basic facts from normed-linear spaces and Banach spaces and so on. Suppose you have two normed linear spaces V and W and T from V to W is a linear map. Then the following three conditions are equivalent. (Remember on an infinite dimensional vector space, a linear map may not be continuous. So, therefore these things become non-vacuous non trivial statements):

(1) T is continuous at 0.

(2) There exists λ positive such that norm of $T(x)$ is less than $\lambda\|x\|$ for all $x \in V$.

(3) T is continuous uniformly on the whole of V.

So, all these three are equivalent conditions. None of them may be true in general. When V is finite dimensional this will be automatically true for all linear functions.

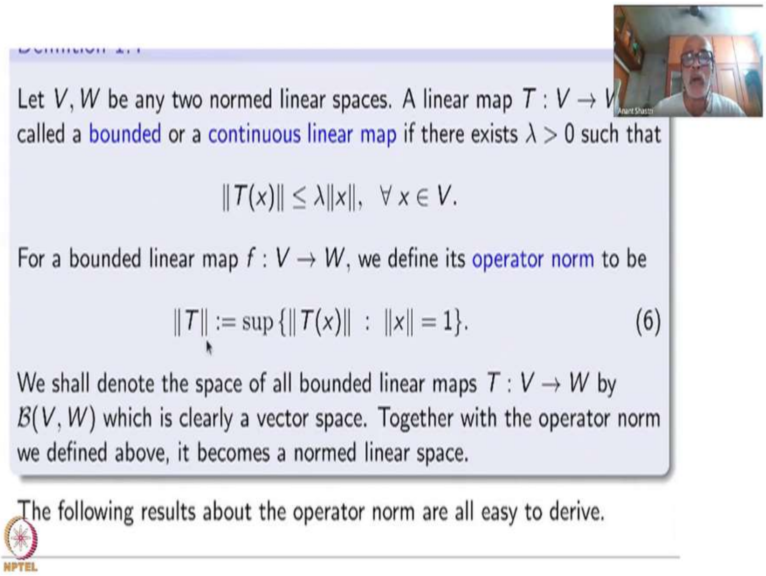
So, let me just recall this because this is so fundamental: Proof of (1) implies (2). That means first you assume that T is continuous at one single point 0. (By the way you can assume continuity at any other point also it is equivalent to condition (1). I could have added that condition also here in the list. Think about it take it as an exercise.) So, put $\epsilon = 1$. By continuity of T at 0, we get a δ positive such that $\|x\| \leq \delta$ implies $\|T(x)\|\ eq 1$. Excellent. This for just taking $\epsilon = 1$.

Now you take x to be any non-zero vector then by linearlity of T, we have norm of $T(x)$ is equal to $\|x\|/\delta$, I am writing $\delta/\|x\|$ times x here, T of that and then I have to compensate for this factor which comes out and cancels out with this one. So, this all this entire thing is just norm of $T(x)$. Now the multiplication factor $\delta/\|x\|$ is taken inside here, the inverse of the same factor is multiplied outside here.

What is the idea of this one? Now, what is the norm of the term inside the bracket here? It is less than or equal to δ and so, I can apply this inequality (1). It means that this part is less than or equal to 1. So, for all $x \neq 0$, we have $\|T(x)\| \leq \|x\|/\delta$. If $x = 0$, then the left hand side 0 and hence the inequality is still valid. So, for all x this is true actually but for writing down this proof I have to assume x is different from 0, because I have to divide by norm x here.

So, we can take $\lambda = 1/\delta$. Then what do I get? I get $\|T(x)\| \leq \lambda\|x\|$, so that is the conclusion of (2). As soon as you have such a uniform lambda uniform continuity follows on the whole of V: $\|T(x_1 - x_2) \leq \lambda\|x_1 - x_2\|$ for all $x_1, x_2 \in V$. So, if $\|x_1 - x_2\| \leq \delta$, which you have to choose this appropriately viz., ϵ/λ, that is all, so that (2) implies (3) is follows. Finally, (3) implies (1) is obvious because this is now actually continuous on the whole of V, and it contains 0 also.

(Refer Slide Time: 27:59)



Now, we make a definition here. Take V and W to be any two norm linear spaces, take a linear map T from V to W. We will call it bounded (or a continuous) linear map if there exists λ positive such that $\|T(x)\| \leq \lambda\|x\|$ for every $x \in V$.

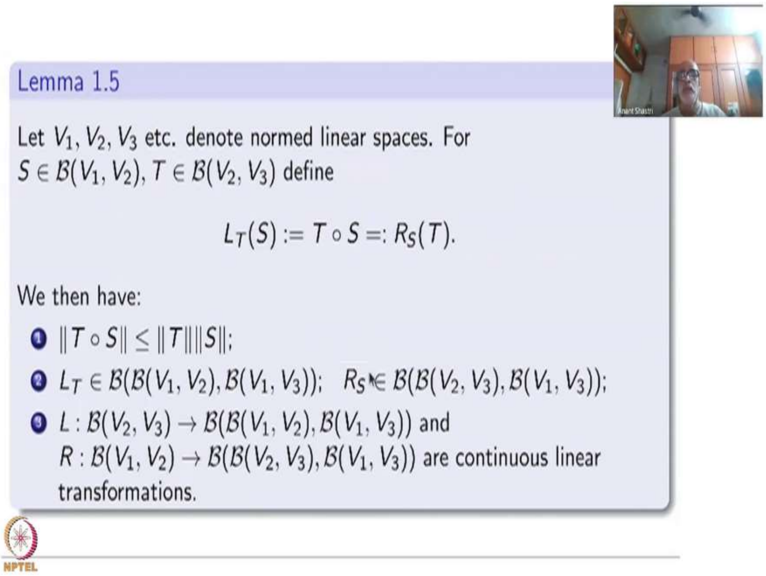
So, such a T is called a bounded linear map. This is a standard terminology in functional analysis. I cannot help it, it is not the standard meaning of bounded functions. For functions which take values in a metric space, boundedness has a different meaning altogether. This terminology is a bit unfortunate but you will get used to it when you are doing function analysis. Or instead of that you say continuous linear map.

For a bounded linear map f from V to W, there is something called the operator norm. We have introduced this one in part I itself. We have in fact studied the Banach space of continuous functions on a metric space and so on. So, I am just recalling this operator norm here. norm T is defined to be the supremum of all $\|T(x)\|$ where $\|x\| = 1$. That means that x is varying on the unit sphere in the domain. Domain is V, x must be inside V of course here.

We shall denote the space of all bounded linear maps T from V to W by this $\mathcal{B}(V, W)$. We will just read it $\mathcal{B}(V, W)$, which is clearly a vector space. What you have to do now? You have to show that if f and g are bounded linear, then $f + g$ is also bounded linear. Linearity is clear. Similarly, you have to show that αf is also bounded. That will show that $\mathcal{B}(V, W)$ is a vector space.

So, together with the operator norm as we have defined here it becomes a norm linear space. This norm has the standard properties: $\|T\|$ is always, you know, is bigger than equal to 0; it is 0 if and only if T is identically 0. And $\|\alpha T\|$ is $|\alpha|\|T\|$. And the triangle inequality in terms of addition, viz., $\|T+S\| \leq \|T\| + \|S\|$. So, these are the conditions which make a function into a norm. The following results about operator norm is a special thing. I want to tell you that these are not a general properties of a norm. As such they all very easy to derive.

(Refer Slide Time: 31:47)



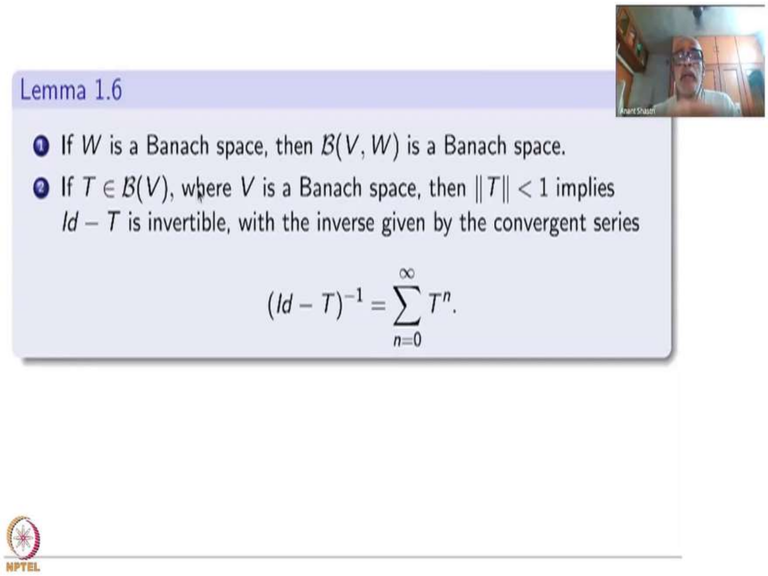
What are these? I have summed it up in this lemma. We will keep using this again and again. So, let us go through it carefully that you understand this one completely. Take vector spaces, V_1, V_2, V_3 they are all norm linear spaces. I am not going to mention various norms specifically. This is standard practice, just like we keep saying X is a topological space without mentioning the topology there. So, V_1, V_2, V_3 etc. are normed linear spaces.

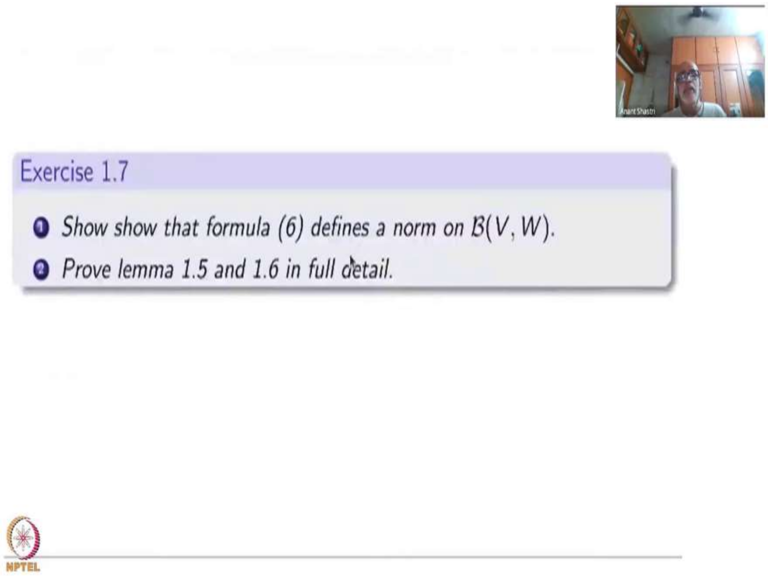
Take a bounded linear map S from V_1 to V_2 and another T from V_2 to V_3. Now I am defining an operation L_T on $\mathcal{B}(V_1, V_2)$ to $\mathcal{B}(V_1, V_3)$, viz., $L_T(S) = T \circ S$. The same thing, you can view it as an operation from $\mathcal{B}(V_2, V_3)$ to $\mathcal{B}(V_1, V_3)$, namely operating via S on T, so you can think of this as $R_S(T) = T \circ S$. $L_T(S)$ is composing T on the left, whereas $R_S(T)$ is composing with S on the right. So, if you vary S this will be a map from $\mathcal{B}(V_1, V_2)$ to $\mathcal{B}(V_1, V_3)$ and if you vary T and keep S fixed then it will be map from $\mathcal{B}(V_2, V_3)$ to $\mathcal{B}(V_1, V_2)$.

So, this is an interesting thing. Obviously, these functional compositions are non-commutative so that is why you have to worry about this whether right composition or left composition separately. In any case the first property (1) is very simple: $\|T \circ S\| \leq \|T\|\|S\|$. This is a very fundamental property which makes the norm linear spaces into what are called as normal algebras. Now, L_T is again a bounded linear map from $\mathcal{B}(V_1, V_2)$ to $\mathcal{B}(V_1, V_3)$. Similarly, R_T is the bounded linear map from $\mathcal{B}(V_2, V_3)$ to $\mathcal{B}(V_1, V_3)$.

They are linear maps is clear. They are bounded linear maps. To see this, you keep using this property (1). If you vary T, it will tell you that the operator norm of R_S is less than or equal to $\|S\|$. Similarly, the operator norm of L_T is less than or equal to $\|T\|$. Thus both L_T and R_S are continuous. That is what I wanted to emphasize. All of them follow from this one single property (1). Linearity is obvious for all of them. If you add S_1 and S_2 here and then compose with T, it is the same as the sum of $L_T(S_1)$ and $L_T(S_2)$. Similarly, R_S operating upon the sum of T_1 and T_2 is the same as $R_S(T_1) + R_S(T_2)$. So linearity etcetera is not a problem and continuity follows from (1).

(Refer Slide Time: 36:48)





So, I will have one more important lemma here namely now I am assuming that W is a Banach space. Banach space is nothing but a norm linear space in which the induced metric is complete, every Cauchy sequence is convergent. So, take W as this codomain as a Banach space. Then the set of all bounded linear functions from V to W with the operator norm becomes a Banach space itself. This is what we have studied last time in the part I.

Let T belong to $\mathcal{B}(V)$. Now $\mathcal{B}(V)$ is short notation for $\mathcal{B}(V, V)$. So, these are self-maps from V to V, bounded linear operators. Let T is an element of $\mathcal{B}(V, V)$, where V itself is a Banach space. Suppose that $\|T\| \leq 1$. Then $I - T$(or you can take $I + T$ also which is the same thing because I can change T to $-T$) is invertible with its inverse given by the convergent series summation from 0 to infinity of T^n; the geometric series viz., $I + T + T^2 + T^3$ and so on.

Why this is convergent? Look at the partial sums, they are bounded by the partial sums of the geometric series summation c^n, where $c = \|T\|$. That is the proof. So, if you look at identity map Id from V to V. That is invertible. You can now take the ball of radius 1 around it in $\mathcal{B}(V)$. All the elements in the open ball are inevitable. So, this is the hypothesis this is the conclusion here. In any Banach space the unit ball centered around the identity consists of only invertible elements. There may be more invertible elements of course. For instance non zero scalar multiples of invertible elements and so on.

The open ball if you take all the elements in that are inevitable. So, this is the hypothesis this is the conclusion here. In any Banach space the unit ball centered around the identity represents all invertible elements. There may be more and multiple elements of course but this is definitely a all these are invertible elements this is the meaning of this.

So, here are some elementary exercises for you to work out. It contains few other statements than whatever I told you already, without proof and so on. So, that is the first lesson you have to do. So, let me just read out these.

(1) Show that formula (6) defines a norm on $\mathcal{B}(V, W)$. I have just given the definition, you have to show, you have to verify those three conditions for the norm.

(2) Prove lemmas 1.5 and 1.6 and this one this is what you have to do. So, that is the exercise for you. Thank you we will meet next time.