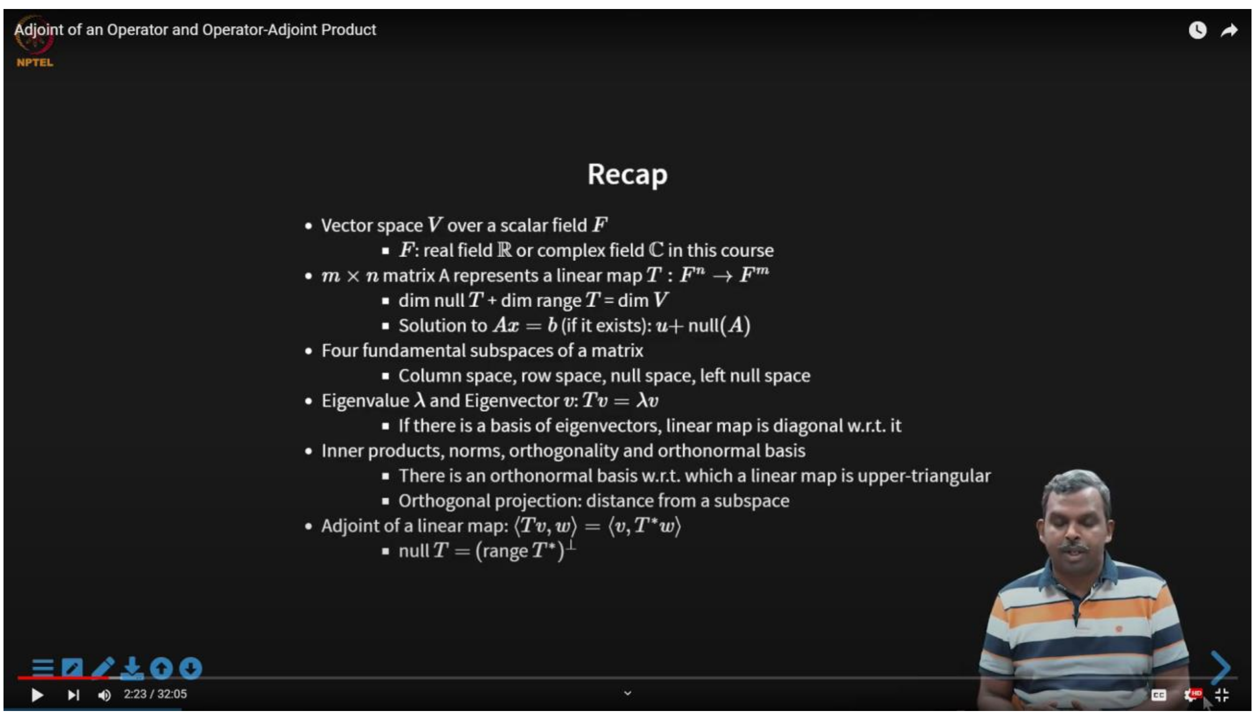
Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Week 9
Adjoint of an Operator and Operator-Adjoint Product

Hello and welcome to this lecture on further properties of adjoint and in particular the product of the operator and its adjoint. What happens when we do that, what kind of properties can we talk about in that context. So this will be a relatively short lecture but it will talk about some things which are very important. In the previous lecture we saw how when you compose two operators, take the prod… Two linear maps, when you take a product of two linear maps, there is this nice little thing you have to look at. Null space of one and the range of the other. And depending on how they intersect, some interesting things happen as far as the composition is concerned. The composition sometimes sort of retains the properties of one operator and in other cases it doesn't, okay? So that's, that was interesting. So one can also look at further properties which come from compositions, and that will play a key role later as well. So we will start looking at those kinds of things.

(Refer Slide Time: 02:23)



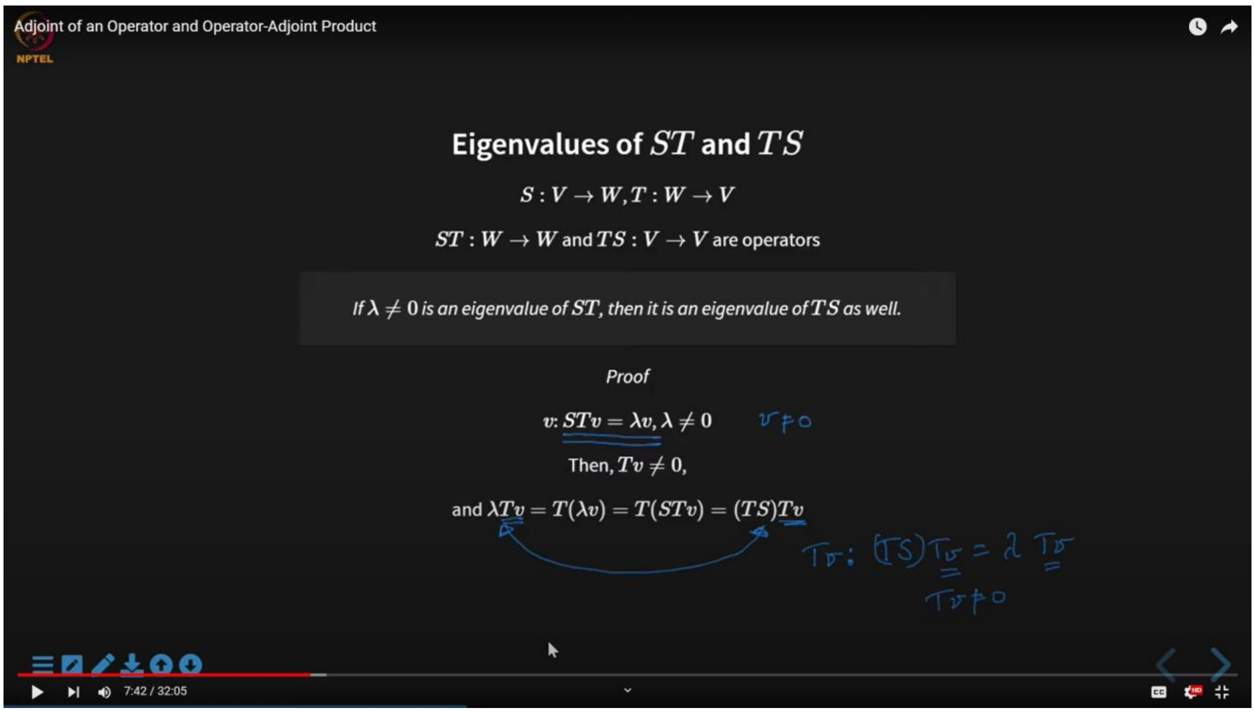
And another aspect of this lecture is that we will start looking at operators and their adjoints. So far we've been looking at linear maps and their adjoints, when the map goes from V to W. Now if the map itself is going from V to V, it is an operator, then that adjoint also is an operator and something interesting seems to be there between the adjoint and the operator itself, okay? So let's start looking at these things more closely, okay?

A quick recap. All the older things are there. The latest thing we are looking at is adjoint of a linear map which sort of gives you the mapping between inner products in some sense, right? So what happens when inner products are involved with linear maps? How do you associate, how do you relate inner products to linear maps? And there is a very nice relationship. Every map has an adjoint so that, you know, $< Tv, w >$ when you take inner product, after you've gone to the output side is also reflected on the input side through the adjoint, okay? So it's a very nice little relationship. And we saw that there are these nice connections between the null space of the map and the range space of its adjoint, okay? And all these interesting relationships come in handy when we try to derive more advanced results about operators, okay? So let us move on.

We will start, in this lecture… One of the ideas is trying to relate the eigenvalues of ST and TS in the case where both ST and TS are operators, okay? So there is a special case. It is sort of a special situation. Supposing you have a map S, linear map $S: V \to W$. And you have a linear map T, some other linear map $T: W \to V$, okay? So that's the situation where both ST and TS are well defined and they are both operators, okay? So you can see that this is a very interesting situation and it can happen quite often. And with an operator and its adjoint, with a linear map and its adjoint, this is exactly what happens, isn't it? So let's look at this situation more closely, particularly from the view of eigenvalues, okay? Anytime you have an operator, now you can start thinking of eigenvalues for it. It simplifies the description of the operator. So otherwise when you have a linear map $S: V \to W$, and V and W are different, you are not able to associate an eigenvalue with it, right? So we sort of went to operators for that. Now it seems like if there is a map, some sort of a map from W to V and maybe it is connected to S, then for the product of those two, both ways we can associate eigenvalues. And is there a connection between these eigenvalues and the original linear map? We will explore these things later. But it seems like an interesting way to study linear maps in the case where V and W are different, okay? And still try and associate some invariancy for subspaces and etc., okay? So this is something interesting from that point of view. So clearly you see that both ST and TS are well defined. And what about eigenvalues for these things?

So here is a very very interesting and simple to prove result. But very interesting. So if you have a non-zero eigenvalue for ST, it should also be an eigenvalue for TS, okay? So this is a crucial result with these kinds of products. If you look at ST and TS and both are operators, then they share the same set of non-zero eigenvalues. It cannot be that there is a non-zero eigenvalue for one but it is not an eigenvalue for the other, okay? It's a very interesting little result and the proof is not very complex as well. So this is interesting. So it seems like when both ST and TS are defined, it's enough if you worry about the eigenvalues of one of those, the other is the same, right? So proof is actually quite simple. So supposing you take an eigenvalue λ, there will be an eigenvector v which is non-zero, right? So v is also non-zero. I think I should maybe write that down somewhere, okay. So you have this eigenvector v which satisfies this property. So this is what makes λ an eigenvalue for ST, isn't it? STv should be equal to λv and $\lambda \neq 0$, okay? So if $\lambda \neq 0$, clearly this Tv will also be not equal to 0, right? If $Tv = 0$, then the left hand side becomes 0. Clearly right hand side is not 0. So this Tv is also not 0. This Tv not being 0 will be crucial, okay? It will be crucial in the next thing. So here is what I'm going to do. I'm going to look at λTv, okay? Notice this little bit of trickery here. So λTv is the same as $T(\lambda v)$. And what is λv? The same as STv, okay? And what is $TSTv$? There is commutativity for me for operators. So you can do TS first and then Tv, okay? Linear maps when you multiply, they compose, they commute like this, this is the associativity property of this thing. So it is okay. So TS you can do first, Tv.

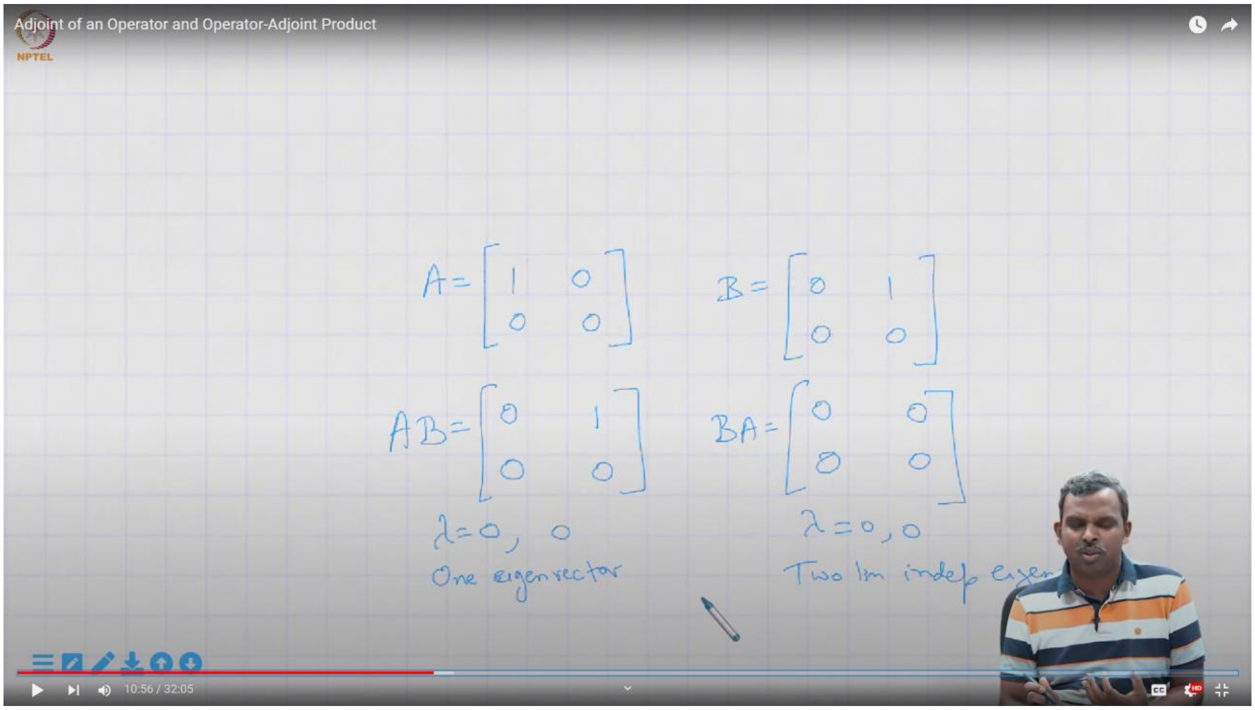
(Refer Slide Time: 07:42)



And notice this equation. This is an interesting little equation. It says, right, you started with this guy, right, v was an eigenvector for ST with eigenvalue λ. Now you see Tv, right, Tv is an eigenvector for TS with eigenvalue λ, okay? So Tv is such that TS times Tv equals λTv, the same λ, and Tv is non-zero, right? So Tv is non-zero and it satisfies this equation. So λ becomes an eigenvalue for TS and Tv is the eigenvector also. So that is an interesting thing, right? So if v is an eigenvector for ST, then Tv is an eigenvector for TS with the same eigenvalue λ. But λ has to be non-zero here. If λ is 0, things are a bit more murky and let's not, we will look at it in the next slide. But this is the main point here, okay? So hopefully this was illuminating. So it's very useful to use this result and we will use this result later on also. But it's important to know that the set of non-zero eigenvalues of ST will be equal to the set of non-zero eigenvalues of TS, okay?

So now the next question is: what about multiplicities, right? I have shown one non-zero eigenvalue is also present here, can the multiplicities be different? Can the algebraic multiplicities be different? Can the geometric multiplicities be different? It seems, it seems possible, right? So here is a, I think let me show you one example here. So here's an example. Supposing you take A and B as 2×2 matrices, okay? So here is an example. So maybe I should… So let us say A is $[0\ 1;\ 0\ 0]$ and B is also, I think I want $0, 0$ here and let me put 1 here… No I shouldn't put 1 here, so let me just think of this a little bit. The danger of cooking up examples on the fly. Let me just see this once. 1 here, so maybe I need a 1 here. Okay, maybe, maybe I will get this right, I think, yeah, this is right, okay? So I think this is okay. Yeah, okay, this okay. So let me see. I mean, I want a particular situation for AB and BA. If A and B were like this, I think I will get it. So let me just do this. $[0, 1;\ 0, 0]$. BA is going to be $[0, 0;\ 0, 0]$, right? Yeah, I think this is what I wanted.

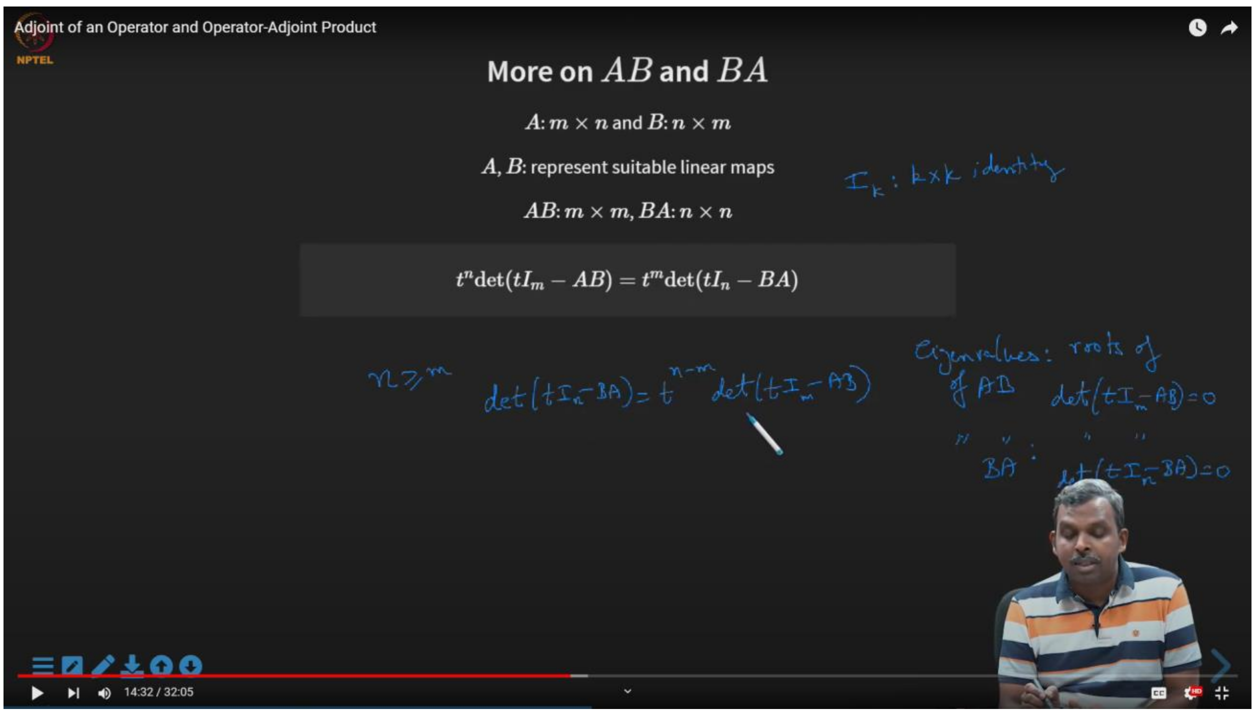
(Refer Slide Time: 10:56)



Okay. So so here is a… See when, this is ST and TS, right, I mean this is a simple 2×2 case where everything is well defined. But still this is valid, this is not wrong. This is ST and TS. And in one case you have AB being $[0, 1;\ 0, 0]$ and BA being $[0, 0;\ 0, 0]$. So what happened here for AB? You have eigenvalues 0 occurring twice and there is only one eigenvector, isn't it? So here, when I say one, you mean really, you know, with the one linearly independent eigenvector.

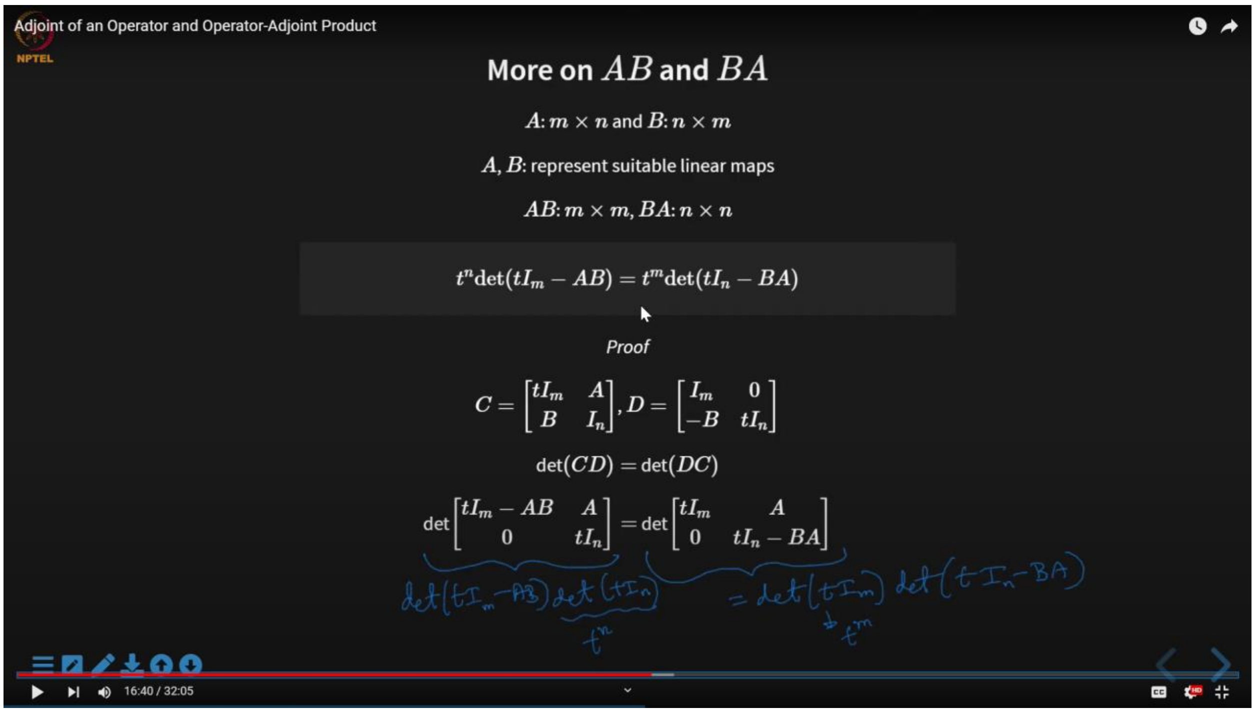
That's what I mean here. This this, you do not have two linearly independent eigenvectors, that's what I mean when I say one eigenvector. So here also you have λ = 0, 0. But you have two linearly independent eigenvectors. So what happened here is: the geometric multiplicity is different, but the algebraic multiplicity ended up being the same, okay? But we, I mean this is zero eigenvalue, it wasn't really covered in our previous result, right? Only non-zero eigenvalues, if they appear, they have to appear on both sides. But even then it is not clear how many times that non-zero multiplicity eigenvector should appear algebraically, how many times it should appear geometrically, okay? So geometrically at least it seems like the eigenvectors can, I mean the eigenvalues can, the geometric multiplicity may be different with AB and BA, that seems likely because geometric multiplicity depends on so many other properties of the matrix. But algebraic multiplicity, what about that? Can we say something about that is the question, okay? So this is nice. So let's look at it.

(Refer Slide Time: 14:33)



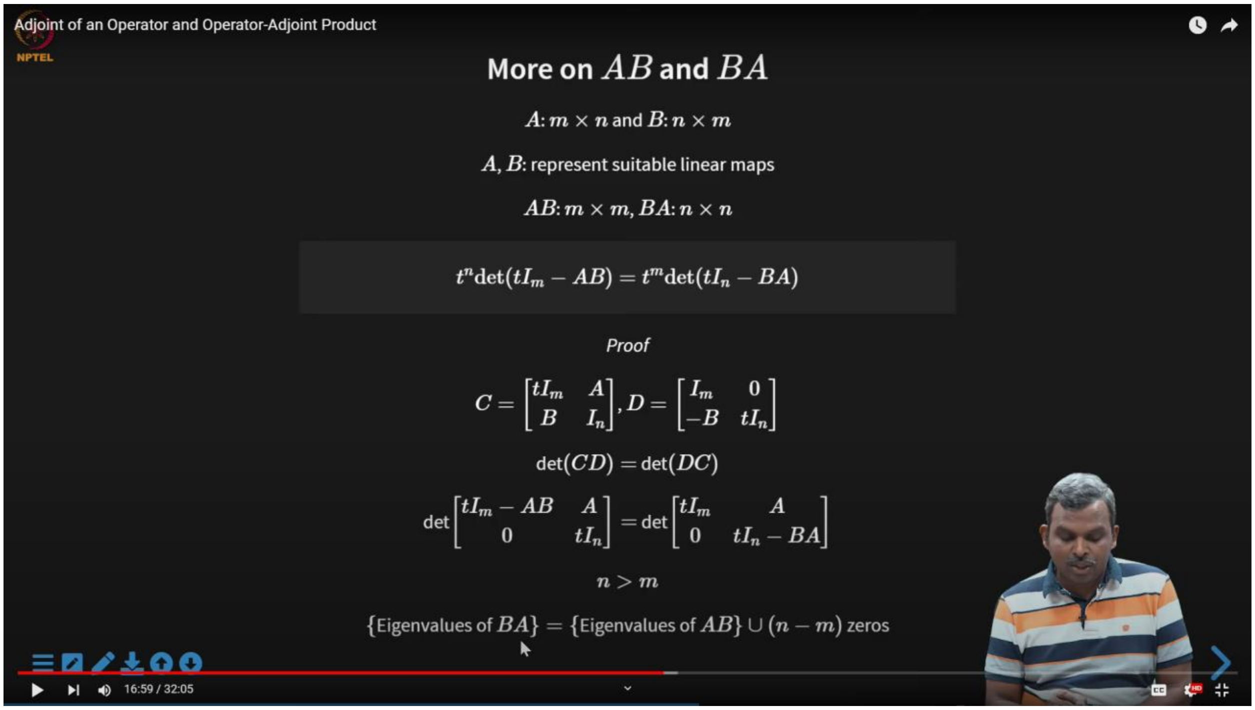
But it turns out, at least to the extent to which I have seen the proofs of these things, when you want to talk about algebraic multiplicities, it's easiest to deal with matrices and determinants, right? So let's do that. So in the matrix determinant world, it is easier to prove this result because I am thinking of algebraic multiplicity ultimately, you know, so that is better. It is better in this world. So let us look at AB and BA. A is an $m \times n$ matrix and B is an $n \times m$ matrix, okay? So you see, let us say A, B represent suitable linear maps S and T. They are $m \times n$ and $n \times m$. Now AB will be $m \times m$ and BA will be $n \times n$. So here is a similar situation to before, except that previously we thought of, you know, linear maps in an abstract way. Here I am talking about matrices specifically, okay? So here is a very, very interesting result. The result says, see, remember, $|tI_m - AB|$, right? So solution of this gives you eigenvalues, right? So eigenvalues, if you are worried about algebraic multiplicity of AB, are roots of $|tI_m - AB| = 0$, isn't it? AB is $m \times m$, determinant… I_m is basically, I_m denotes $m \times m$ identity matrix. In denotes an $n \times n$ identity matrix, okay? So that is the notation, $k \times k$ identity, okay? So similarly eigenvalues of BA are given by, so same thing for BA is given by roots of $|tI_n - BA| = 0$, okay? So what this tells you is another way of writing this. Supposing let's say $n \geq m$. Supposing you say n is bigger, then this simply says $|tI_n - BA| = t^{n-m}|tI_m - AB|$ okay? So this is what this result says.

(Refer Slide Time: 16:44)



And this is a very nice result, isn't it? See, $|tI_m - AB|$, it's a polynomial. The roots of this polynomial are the eigenvalues of AB. If you take that polynomial and multiply by t^{n-m}, what is t^{n-m}? It is simply, you know, just all the powers of t increase by $(n - m)$, just simple multiplication, okay? That gives you the polynomial whose roots are the eigenvalues of BA, okay? BA is $n \times n$ and that's the relationship between these two polynomials. So all the eigenvalues of AB will occur as eigenvalues of BA algebraically with the same multiplicities except for zero. The zeros' multiplicity will go up by $(n - m)$, right? So that is a very nice observation you can make when you have a result like this, okay? If you are concerned about algebraic multiplicity for eigenvalues, this relationship is very, very useful. Of course, geometric multiplicity is not mentioned here. So all the algebraic multiplicities of non-zero eigenvalues remains the same for AB and BA, the algebraic multiplicity of 0 alone increases by $n - m$ if n is bigger, okay? Other way round it will increase by $m - n$, okay? So that is the power of this result. It is a very nice result. Using determinants, we are able to get them, okay? So let's see a quick proof of it.

(Refer Slide Time: 16:59)

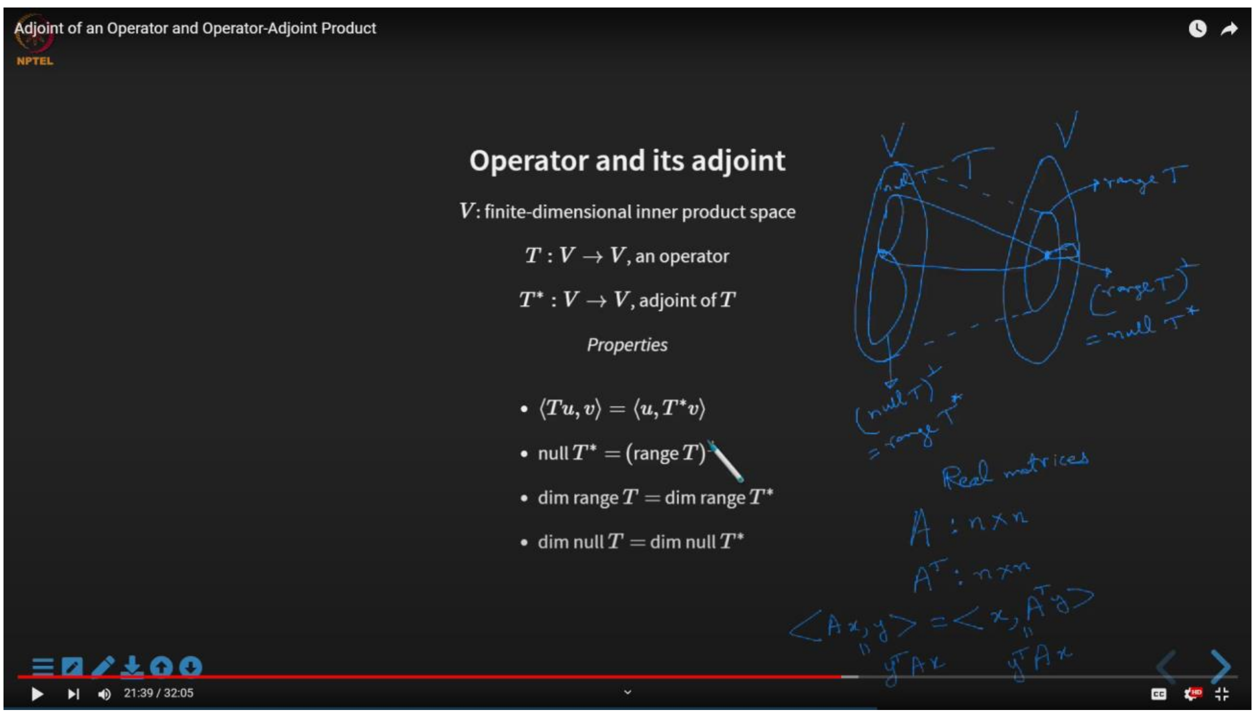


A proof is not very hard. It's sort of like one of those clever proofs by construction which is difficult to provide intuition for. But this is what you do for the proof. You construct two matrices C and D, okay? Which are block matrices involving I_m, I_n, A and B, okay? And this T also. Construct things like this. We know always the determinant of CD and determinant of DC are the same, right? So this is same. So if you multiply CD, you get this. You multiply DC, you get this. So what is determinant of this? Determinant of this guy into determinant of $(tI)^n$, that is t^n. What is determinant on this side? Determinant of this times the determinant of T… So maybe I should write that down. This guy is equal to determinant of $tI_m|tI_n - BA|$ and this is nothing but t^n, isn't it? Same thing here. This is $|tI_m - AB||tI_n|$ and this is t^n, okay? So there is nothing difficult about this proof except that this construction of the C and D seems very non-trivial. And how it comes about is a bit tricky. It's because it's got all these tie-ups here and you can see that it's quite nice, okay? It comes out very cleanly as a block triangular matrix and it gives you what you want. But the C and D are not very obvious, okay? But the result is quite nice. So now we know a lot more about BA and AB, particularly through this relationship for eigenvalues, right? So the eigenvalues, non-zero eigenvalue set is exactly the same. Zero eigenvalue just gets repeated more when you go from one to the other, okay? So this is a nice relationship to know.

Okay. So now let's move on to operators and adjoint. So far we haven't come to adjoint yet. We've just looked at general properties of this AB and BA. Later on we will use these results where we need them, okay? Yeah. So I've summarized it here. If $n > m$, then the eigenvalues of BA is basically the eigenvalues of AB union with $(n - m)$ zeros. So this is the relationship.

Okay. So now let us look at operator and its adjoint, okay? So now… Let me just go on for one minute. Yeah, okay, all right. So, so far we have looked at the adjoint of linear maps. Now we will specialize to the case of operators and look at what's interesting here, okay? So if you have a finite dimensional vector space, inner product space, you can have an operator from V to V and then that operator can also have an adjoint. I mean, nothing from the definition of adjoint stops this, right? We generally had a more general adjoint where we went to, I mean $T : V \to W$. So if it were to take V to V also the same adjoint definition holds. Nothing is wrong in that. Here are some interesting properties. The property for the adjoint, the basic definition, right, this is more or less the definition. $< Tu, v >$ will be equal to $< u, T^*v >$ and null of $T*$ is (range of T)$^{\perp}$ and dimension of range of T equals dimension of range of T^*, all these were true in general for linear maps, isn't it? Now because it's an operator, the dimension of null will also be the same, okay? So that ends up being the same.

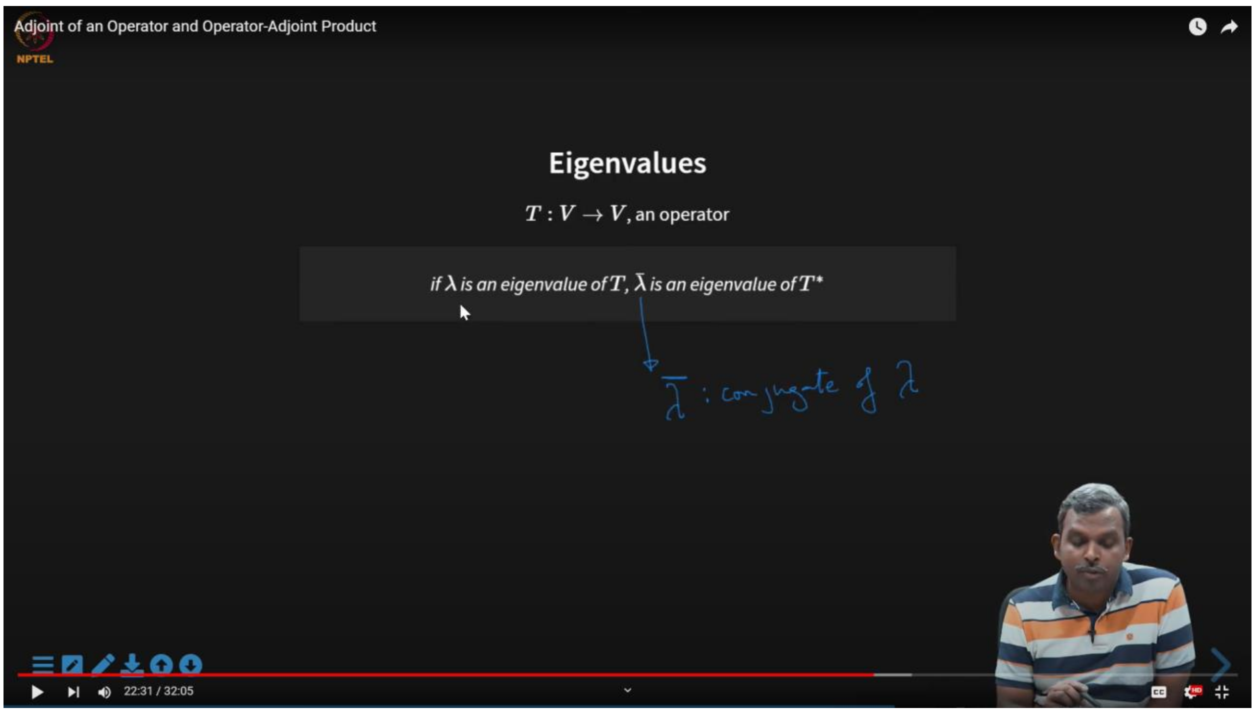
(Refer Slide Time: 21:39)



So let us draw a couple of pictures and then relate this to matrices. So you, if you have V, right? $T: V \to V$. Both of these are the same, okay? T is going to take you from V to V. There will be a range of T, right? This range of T which is a general subspace and then there is null of T, right?

So this is, let us say null of T. Null of T takes you to zero here, okay? And then the entire V goes to range, right? So we say that. Now what about T^*, right? So T^* is going to take you from V to V again. But let us just show it in the other direction. This way, the dual of null T, right? So if you will have the dual of null T, it will be something like this, right? It will intersect only at zero. This will be $(\text{null } T)^{\perp}$. What will be $(\text{null } T)^{\perp}$? And that is equal to range of T^*, isn't it? Okay, same thing here. The dual of this guy, okay? So it's sort of difficult to draw this here. So let me draw this at the end so you will have the dual here, okay? It does not have too much intersection, but this is $(\text{range } T)^{\perp}$. And that is equal to null T^*, okay? So this is sort of the picture that you would have. Any operator does this, okay? So what happens in the matrix world? If you have a being an $n \times n$ matrix representing this operator T, okay, and then you know A^T, okay? Let us say this is real matrices. A^T is also $n \times n$, it represents T^* and you can see all these things are true, right?

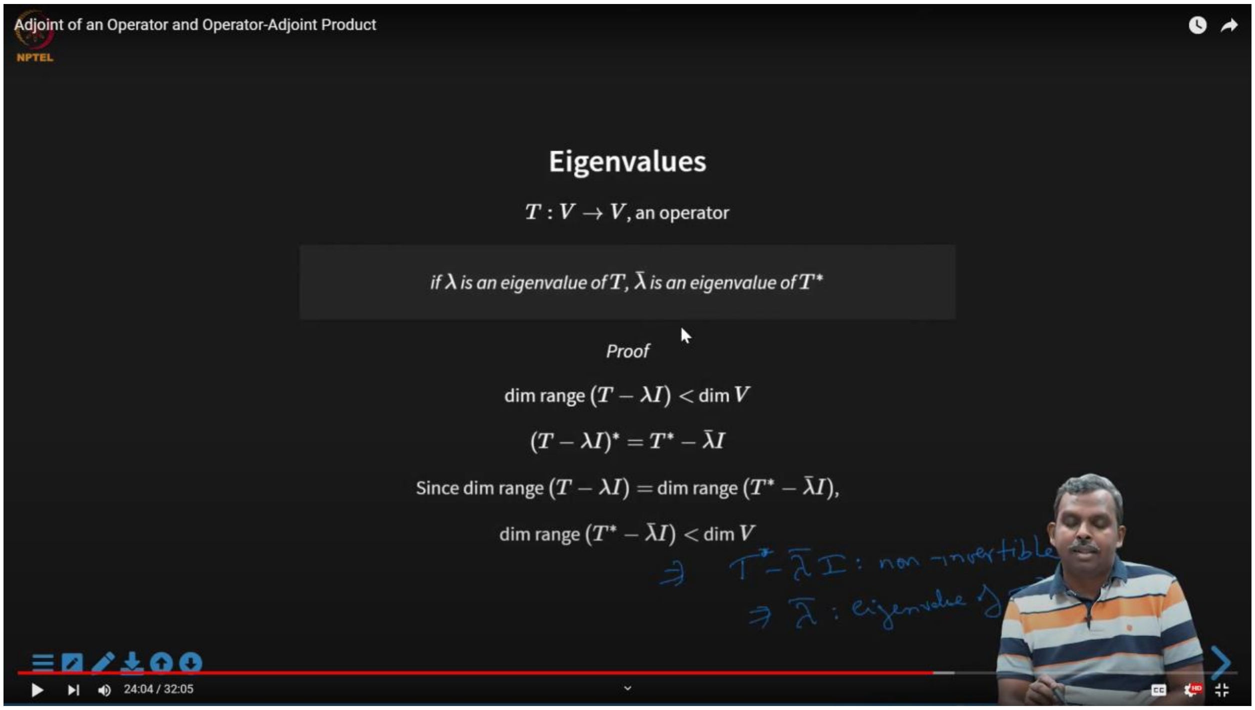
(Refer Slide Time: 22:31)



So what is $< Tu, v >$? It is basically inner product $< Ax, y >$ okay? And that is the same as the inner product $< x, A^T y >$, isn't it? So what is the inner product of these two? It's basically, you can write it in, see it's $< Ax, y >$, you can write it as… Let me just do it correctly. $y^T Ax$, right? And you can also transpose this and then you will get… This also is actually equal to $y^T Ax$, okay? All right. So both of these are exactly the same. You can work it out. So you get the same relationship. So you see the transpose is sort of well behaved with respect to that definition of the adjoint and you can see null T, the null T^* is nothing but the left null and that's going to be perpendicular to the range of T, all of this works out fine. And $\dim \text{range } T = \dim \text{range } T^*$ is nothing but the row rank equals column rank. And null T = null T^* is also a similar definition. So in the matrix world also all these definitions work out. So for operators and adjoint, whatever we saw before also holds okay? So this is interesting.

What about eigenvalues for the adjoint, okay? So when T itself is an operator, T will have eigenvalues. And T^* also is an operator, so T^* will also have eigenvalues. Is there a connection between these eigenvalues? It turns out there is a very nice connection. So if T is an operator in an inner product space, if λ is an eigenvalue of T, then $\bar{\lambda}$, okay… So this is $\bar{\lambda}$, okay? So conjugate of λ, complex conjugate of λ is an eigenvalue of T^*, okay? So that is what this means, okay? So this is a very simple and elegant result relating the eigenvalues of the operator and its adjoint. If λ is an eigenvalue, $\bar{\lambda}$ is an eigenvalue of T^*, okay? Proof is not really complicated. So if λ is an eigenvalue, then you know that the $(T - \lambda I)$ is invertible, it is not invertible, right? So $(T - \lambda I)$ ends up being not invertible, which means the range of $(T - \lambda I)$, the dimension of that is less than dimension of V, okay? So this we know for sure, the moment λ is an eigenvalue. Now what is the adjoint of $(T - \lambda I)$? We know that adjoint of sum of two operators is the adjoint of each of them, so you have $(T^* - (\lambda I)^*)$. Now if λI, then you have to do $\bar{\lambda}$, okay? So that's what happens here. So the adjoint goes like this.

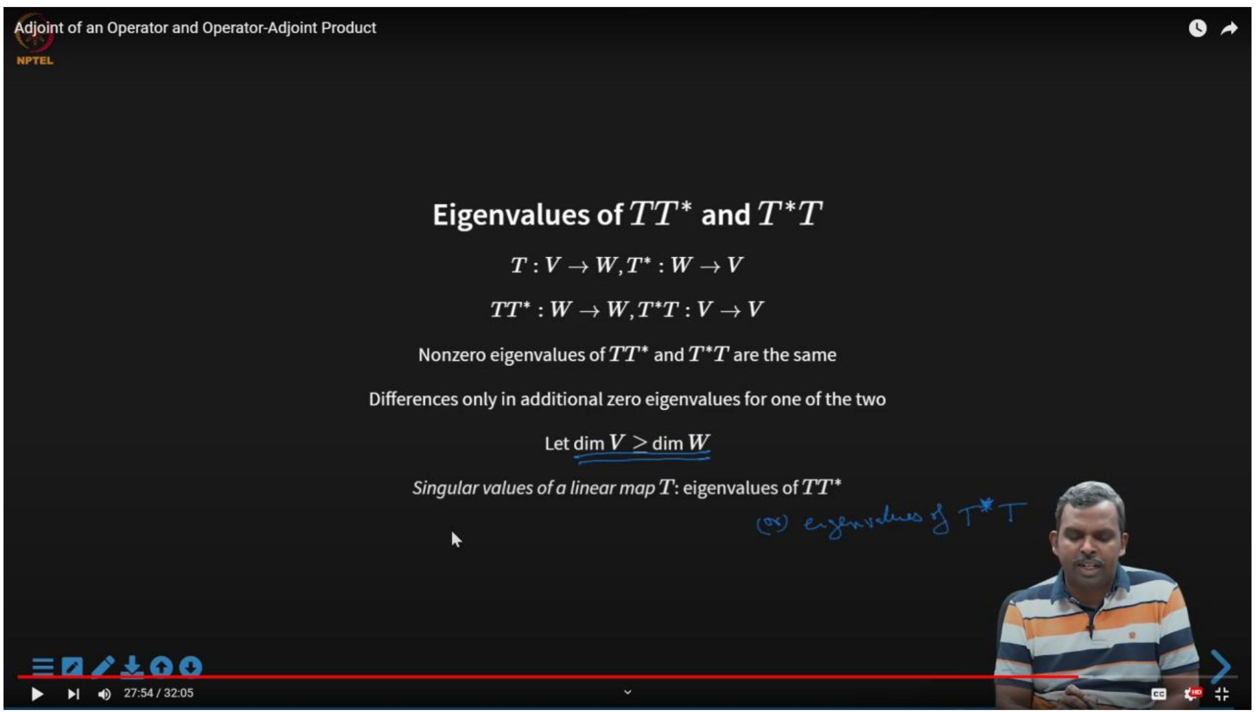
(Refer Slide Time: 24:06)



Now we know an operator and its adjoint have the same range dimension, right? So that is the property that we saw. So dim range $(T - \lambda I) = \dim range\ (T^* - \bar{\lambda} I)$, okay? So now this guy, if the left hand side is less than dimension of V, clearly the right hand side is also less than dimension of V. So this means $\dim \text{range}\left(T^* - \bar{\lambda}I\right) < \dim V$. That implies $(T^* - \bar{\lambda}I)$ is also non-invertible. So that implies $\bar{\lambda}$ is an eigenvalue of T^*, okay? Simple enough proof to show that this relationship is true, okay? So it's very nice. So if λ is an eigenvalue, $\bar{\lambda}$ is an eigenvalue of T^*. So if you are in the real world, if λ is real, then you know $\bar{\lambda}$ is also going to be real, So A and A^T will have, you know, if everything is real, then they will have the same eigenvalues. But if you are in complex, then you have to do conjugate transpose for T^*. Assuming, you know, the matrix is with respect to an orthonormal basis, you can do all that, okay? So this is a good relationship to have. So that's done.

Okay. So what about these eigen, the operator-adjoint products? We saw before that this operator-adjoint product is something really interesting. It sort of represents the operator in some way because it does not nullify the range in either direction. So let us look at the case where $T: V \to W$ now, is a linear map and $T^*: W \to V$. And then we look at TT^* which is an operator. T^*T which is also an operator. Now we can think of eigenvalues of these two and we know from our previous result, right, the non-zero eigenvalues of T^*T and non-zero eigenvalues of TT^* are the same, okay? This is the AB, BA result, right? So difference is only in additional zero eigenvalues of one of the two. So for instance, if you were to take dimension of V greater than or equal to this, so one defines something called singular values of a linear map T, okay? See, so far we couldn't associate eigenvalues with a general linear map T, right? If you had a map from V to W, and V and W were not the same, we don't have eigenvalues to assign.

(Refer Slide Time: 27:51)



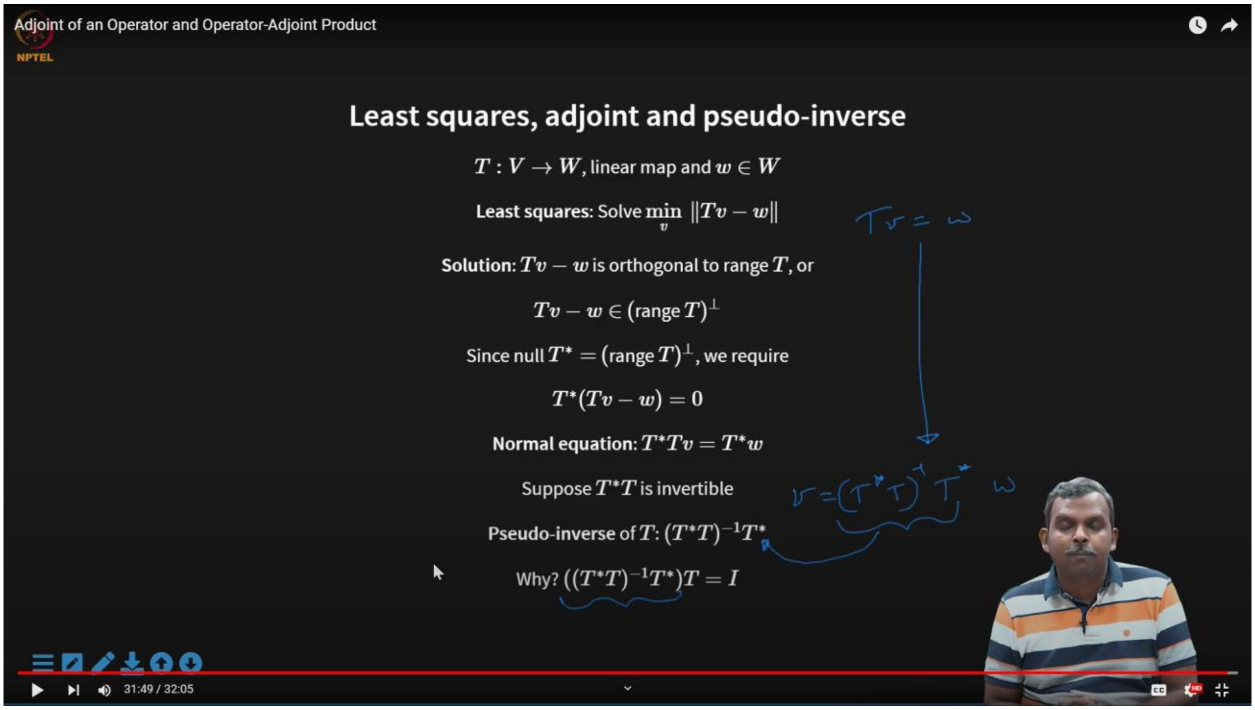
So in lieu of that, people usually talk about what are called singular values of a linear map, okay? What are singular values of linear map T? These are the eigenvalues of TT^*. I have just taken $\dim V \geq \dim W$, that's just to define this in terms of W, in terms of the smaller ones, right? So see, you can also think of it as, see you can also say, or eigenvalues of T^*T, isn't it? Okay? See, because these two have the same eigenvalues more or less except for some zero eigenvalues that are additional, so one can define it like this. But I am, I'm just being a little bit more careful here and I am taking the smaller dimension, I am defining it as the eigenvalues of the smaller dimensions. A bunch of zeros I am ignoring, but that's okay. I mean, you can do it either way, but the singular values are eigenvalues of TT^* or T^*T. I think it's okay, as long as you understand that zeros are the only things that differ here, you can define it either way. So we can take it as the smaller one so that you do not have to keep track of a lot of values, okay? So singular values of a linear map T is eigenvalues of TT^*, I mean, T^*T also is the same, okay?

So these singular values will come to represent T in some very fundamental, nice way, okay? So if you think about it, the eigenvalues sort of capture the essence of the operator, right? When you go to the eigenvector basis particularly for diagonalizable matrices, it becomes diagonal. So likewise these singular values will capture a lot of fundamental properties of the linear map, okay? Even when V and W are not the same, okay? So singular values are very, very important. Later on we will see something called a singular value decomposition for a linear map. And at that point you will see how important they are and they have huge applications today. They have come to dominate applications in linear algebra and matrices and machine learning. And in so many other areas singular values dominate applications, so it's very important to get a handle of that. And it's not surprising, we'll come back and later see more properties of this. And it's not surprising at all why this TT^* represents T also in some way, isn't it? So because, you know, it's not really changing the range in any way, there is no additional null space which creates the problem, okay? So singular values are important. I am just defining it here, we will redefine it later when we see singular values in more detail, okay?

Okay. So this final slide in this lecture captures a few things which I mentioned before in a slightly different language. I want to just bring it down to another language and nail it down, okay? So this is this least squares problem we talked about and its connection to adjoint and something called normal equation, and this pseudo-inverse. I'll just quickly summarize this. Supposing you have a linear map from V to W and you have a vector in W, okay? What is this least squares problem? It is like solving $Tv = w$, except that sometimes it may not have a solution. If it doesn't have a solution, you simply find the least distance between Tv and w, right? So that is this $\min_v ||Pv - w||$. If you solve this, you get a least squares solution, okay? We know by orthogonal projection that there is a solution and the solution is such that $(Tv - w)$ is orthogonal to range T, isn't it? So that's the solution. So what does that mean? That this $(Tv - w)$ is in $(\text{range } T)^\perp$, okay? Now what do I know? Null T^* is the same as $(\text{range } T)^\perp$, isn't it? So when you say $(Tv - w)$ belongs to $(\text{range } T)^\perp$, all you are saying is $(Tv - w)$ belongs to null

T^*, okay? So what is the meaning of a vector belonging to null of T^*? T^* times that has to be equal to zero, okay? So all these things work out in a very clean way and you get what is called a normal equation, okay? So this is $T^*Tv \ = \ T^*w$. So this is the normal equation. In simple ways, the orthogonal projection is sort of hidden here. But you can see I have just put it down. We have seen all this before, I am just putting it out in a language which is clear enough. So this gives you the normal equation.

(Refer Slide Time: 31:49)



Now if this T^*T is invertible… Now remember, T^*T is now an operator, right, from V to V, okay? So if this were to be invertible, you can do, you know this will have a unique solution. So we will have a unique solution here and even otherwise it will have a solution, but otherwise this is an invertible, you can see that there is a unique way to easily get the solution. And so that will look like this. So how will the solution look? When T^*v is invertible, it will look like $v \ = \ ((T^*T)^{-1} \ T^*w)$, okay? So what am I trying to solve here? I am trying to solve $Tv \ = \ w$, and then you get this $((T^*T)^{-1} \ T^*w)$, okay? So this is the solution. It's called the, so this, so this guy here is called the pseudo-inverse, okay? So you can see why. Because if you, if you hit it with T, you get I, okay? So this is, when T is invertible, this gives you, this gives you something, right? So that's nice to see, yeah. So, see, invertibility is a bit tricky here for T. See, if T were to be a square matrix, then some special cases will appear here. But if T is rectangular also, there is no notion of inverse for T as such, right? But this has that property, this $((T^*T)^{-1} \ T^*T)$ ends up being I, because you see this T^*T will come here and this $(T^*T)^{-1}$ will come here, you get I, okay? So you can think of what is left here as some sort of an inverse for T. It's not a proper inverse if T is not a square matrix but it's called a pseudo-inverse, okay? When T is not a square matrix. If T is a square matrix, it becomes the inverse when this is well defined, okay? So this is just some language which we did not use before. I just thought I will summarize all of this, okay?

So this concludes our study of adjoint. And beginning next week onwards we will start looking at very interesting operators such as self-adjoint operators, normal operators and these two operators, these two types of operators dominate applications. Quite a few applications use these type of special operators and we'll see some fantastic properties that they have in the next week, okay? Thank you very much.