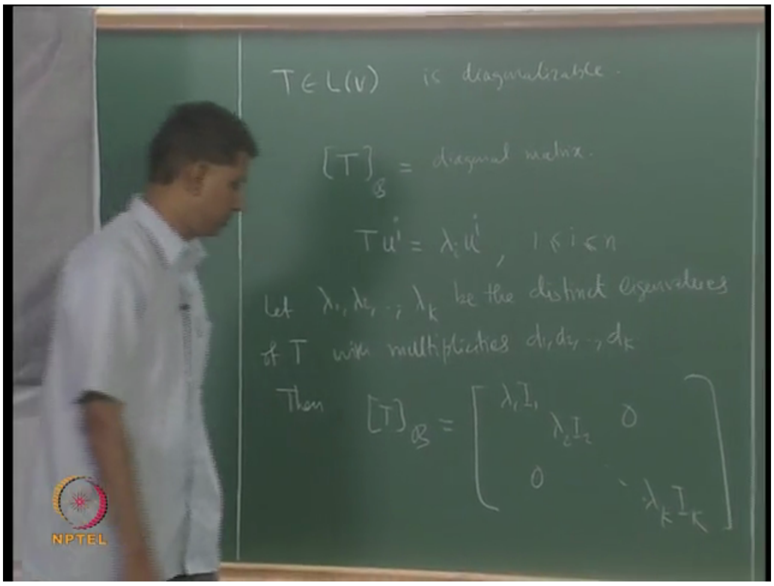
Linear Algebra
Professor K.C Sivakumar
Department of Mathematics
Indian Institute of Technology, Madras
Module 7 Eigenvalues and Eigenvectors
Lecture 27
Diagonalization of Linear Operators. A Characterization

See we are discussing diagonalizability. In today’s lecture we will derive one necessary un sufficient for diagonalizability of a linear operator.

(Refer Slide Time: 0:32)



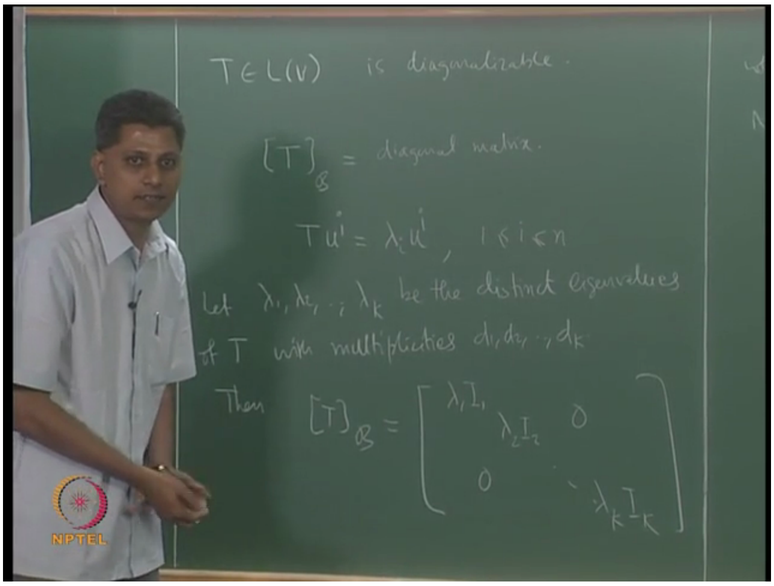
So let us go back and look at this problem diagonalizability of a linear operator T this means that I can write down the matrix of T relative to some basis as a diagonal matrix, okay. I am writing down this diagonal matrix let me assume that by the way is it clear that the eigenvalues of this diagonal matrix must be the eigenvalues of T the eigenvalues of the diagonal matrix.

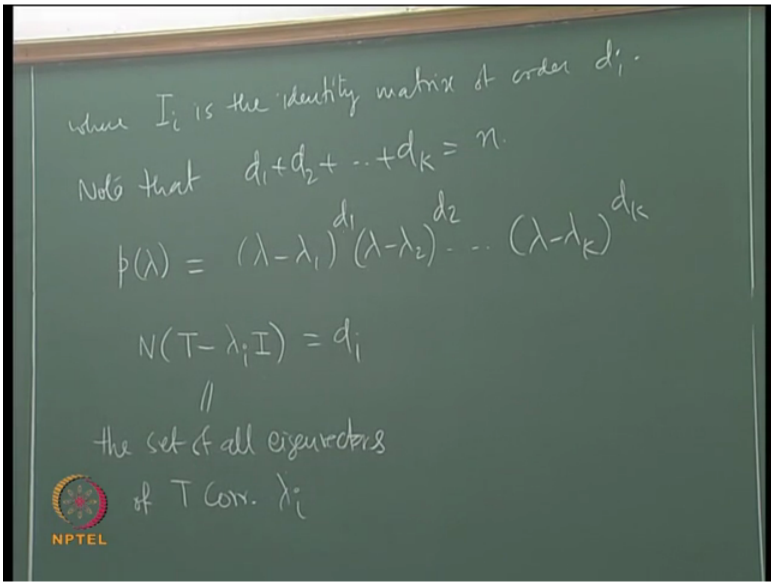
So what is the meaning of saying that T is diagonal? This means T u i I am just writing down the equation that I gave yesterday instead of alpha let me write lambda i u i, okay these are the numbers that will figure here these lambda i's are the numbers that will figure here, okay. Now some of we have seen example the second example where the eigenvalue repeats some of these could repeat.

So what I will do is write down the distinct eigenvalues and take care of multiplicity when I write down the diagonal matrix. In other words let me say lambda 1, lambda 2, etc lambda k be the distinct eigenvalues be the distinct eigenvalues of T with multiplicities with multiplicities I will use this notation different from what I used yesterday in the last lecture with multiplicities let me say d 1, d 2, etc d k that means lambda 1 comes as an eigenvalue of the operator T d 1 times in other words the characteristic equation that characteristic polynomial is determinant of (A min) determinant of T minus lambda i is 0, characteristic polynomial is determinant of T minus lambda i, characteristic equation is determinant of T minus lambda i equal to 0 this is the polynomial of degree n.

When I write down this what I mean is that lambda 1 appears d 1 times as a root of that characteristic equation, lambda 2 appears d 2 times, etc lambda k appears d k times, okay. So what is clear is that d 1 plus d 2 etc plus d k is equal to n the degree of the polynomial, okay. If this is the case then this can be rewritten then I can write this matrix diagonal matrix as lambda 1 now appears d 1 times so the first block will have lambda 1 appearing d 1 times is it clear that I can write it as lambda 1 I 1, lambda 2 I 2, etc lambda k I k all other entries are of course 0 with the convention that I 1 is the identity matrix of order d 1 cross d 1, I 2 is identity matrix of order d 2 cross d 2, etc I k is identity matrix of order d k, okay.

(Refer Slide Time: 4:32)





So let me confirm where I i is the identity matrix of order d i so I can first this is the first observation the diagonal matrix d can be written in this manner after listing the distinct eigenvalues. So let us observe that the statement that I made just now d 1 plus d 2 plus etc d k must be the dimension of the space that I started with I will always assume the dimension to be n in the rest of the discussion, okay.

If this is the case what is the characteristic polynomial of T? The explicit formula for the matrix of T has been written down we knows the characteristic polynomial does not change when I change the basis, okay because it is basically the determinant of some matrix which does not change under change of basis, okay. So the characteristic polynomial so I am using p for the characteristic polynomial I will use p here after for the characteristic polynomial.

So the characteristic polynomial of this operator T can be written as it will be lambda will be a variable lambda 1, lambda 2, etc are the 0's so I have lambda minus lambda 1 to the d 1 into lambda minus lambda 2 d 2, etc lambda minus lambda k to the d k is it clear that this is the characteristic polynomial of T, okay. See in general this cannot be done I am assuming if T, I am assuming the T is diagonalizable in that case I can do this.

For example look at the operator first operator that we considered yesterday the rotation matrix when the angle is 90, the characteristic polynomial there is lambda square plus 1, okay and lambda square plus 1 cannot be factorized as lambda minus lambda 1 into lambda minus lambda 2 for lambda 1, lambda 2 real because this is an irreducible polynomial over the polynomial ring F x, R x, okay irreducible polynomial over R lambda square plus 1 is one example we cannot factorize there that is an example of a operator which is not diagonalizable.

So this cannot be done always if it is diagonalizable then this can be done, if the operator is diagonalizable then the characteristic polynomial can be factorized into products of powers of linear factors let me mention this is the product of powers of linear factors, okay okay. So this is one information that the characteristic polynomial is a product of powers of linear factors is one information. I also have the other information look at look at the eigens look at the dimension of the subspace null space of T minus lambda i I, I want to calculate the dimension of the subspace null space of T minus lambda i I, I want to calculate the nullity of T minus lambda i, okay can you tell me what it is?

What is what is this subspace first? See I want T minus lambda i I, I have T here I can treat like matrix for being specific let us take lambda i to be lambda 1 i is 1 lambda 1 T minus lambda 1 I this first block is 0, all the other entries will remain why because they are distinct this will be lambda 2 minus lambda 1 times I 2, this will be lambda k minus lambda 1 times I k. I am doing T minus lambda 1 I so these entries will not be 0 distinct eigenvalues so lambda 1 will not be equal to lambda 2 etc lambda k.

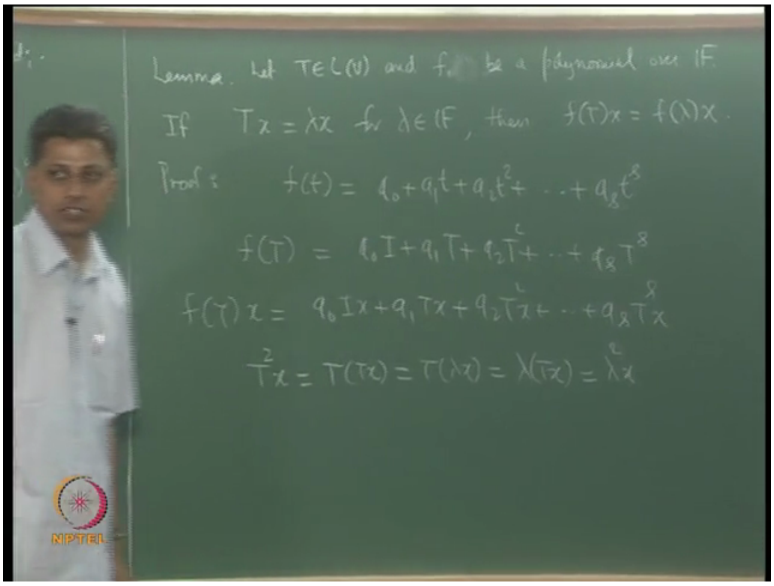
But this block is 0, I want the set of all solutions of the matrix equation let us say some some A x equal to 0 where A is a matrix whose first block is 0 all other diagonal entries are not 0 so what is the dimension of the solution space is it clear it is d i the rank of T minus lambda 1 I will be the rank coming out of these nonzero entries so nullity corresponds to just this so nullity of T minus lambda I is d i please check this, so what is that clear first? The null space has this dimension.

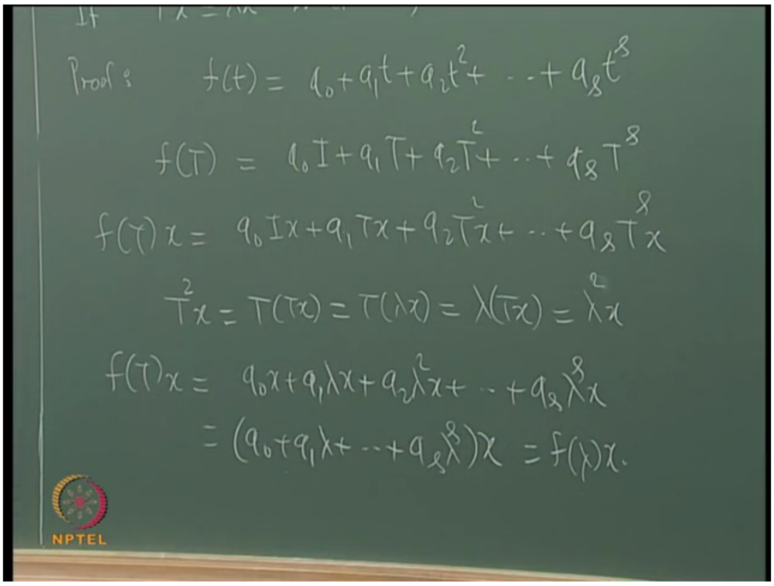
By the way what is the null space of T minus lambda 1 I? Can you see that it is the it is eigenspace corresponding to the eigenvalue lambda I this I did not mentioned in the last lecture. This is the set of actually the space the set of all eigenvectors of T corresponding to the eigenvalue lambda i this is the set of all eigenvectors of T corresponding to the eigenvalue lambda i first observation this is a subspace I did not make this point yesterday the set of all eigenvectors corresponding to an eigenvalue forms a subspace because it is a null space of a certain linear transformation null space of T minus lambda i.

So this subspace has dimension d i remember there are there are two objects coming here one object is a polynomial the other one is the dimension of some subspace, if T is diagonalizable these two numbers are the same this is an important observation because we will see that you can go back to example 2, example 3, example 2 we discussed only two examples we can go back to example 2.

Look at the dimension of the subspace corresponding to the eigenvalue 1 the dimension is 1 there whereas the multiplicity of 1 as an eigenvalue was 2 in general these are different, okay. For diagonalizability it is crucial that these two numbers are the same, okay this is the second important point first point is that the characteristic polynomial is can be written as a product of powers of linear factors, second fact is that the dimension the number of times lambda i appears as an eigenvalue of A that is d i that number is the same as the dimension of the eigenspace corresponding to eigenvalue lambda i, okay okay.

(Refer Slide Time: 12:22)





Let us proceed and we will this is important as you will see this is important for diagonalizability, okay. As I told you I want to look at a characterization let me give you one or two results before I proof that result. First I want to demonstrate the following suppose T is an operator on a finite dimensional vector space V and f of lambda okay f of t I use small t okay let me just say f be a polynomial be a polynomial over F, V is vector space over F, f is a polynomial over F by which I mean that this polynomial has its coefficients coming from F, okay so if F is R this is a real polynomial.

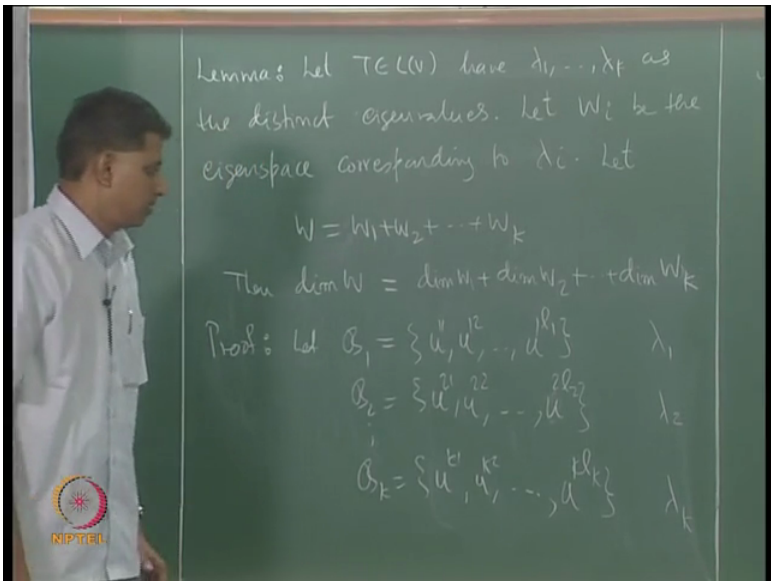
If Tx equals lambda x for lambda the underlying field, then f T of x is f lambda x you will need this little result when f T of x is f lambda x what this means is that if lambda is an eigenvalue for the operator T and if x is eigenvector then look at f of T f of T is also a linear operator see little f is a polynomial f of T is another linear operator for this linear operator we want to show that f of lambda is an eigenvalue with a same eigenvector x the same eigenvector x, okay.

Proof straight forward you just write down f of t let say it is a not plus a a 1 t plus a 2 t square, etc let us say a s t to the s where the coefficients come from the underlying field, then what is f of capital T wherever little t comes I must replace it by capital T the first term it is T power 0 that is identity operator, okay so my f of T is this polynomial a not I plus a 1 capital T plus a 2 capital T square etc a s T to the s you remember T is an operator on V it is a linear transformation from V to itself.

So it is T square T cube etc make sense composition this is f of T, what do I want to verify verify this. So let me write f T x this will be a not Ix plus a 1 Tx plus a 2 T square x plus etc I need to calculate each term, okay but it is easy to see that T is okay let us calculate T square x for example T square x is T of T of x T of x is lambda x, so this is T of lambda x this is lambda T of x again T of x equals lambda x. So this is lambda square x, so T square x is lambda square x by induction T power R x is equal to lambda power R x so this can be rewritten so it is a simple result.

So if you look at f T x it is a not x plus a 1 lambda x plus a 2 lambda square x plus etc plus a s lambda to the s x, x is the vector take that outside all the others are numbers coefficients a not plus a 1 lambda plus etc plus a s lambda to the s into x this is a this is a number coming from the field but this is precisely f of lambda instead of t I have lambda. So this is f of lambda x, okay so remember that f T is a polynomial in T if lambdas and eigenvalue and x is the corresponding eigenvector then we have found out an eigenvalue for f of T that is f lambda for the same eigenvector x, okay okay that is a simple computation.

(Refer Slide Time: 17:25)



I want one more result before I proof the main theorem so the framework is the same as before lambda 1, lambda 2, etc lambda k are the distinct eigenvalues of a linear operator T on a finite dimensional vector space V. Let me say that W i let W i be the eigenspace corresponding to to the eigenvalue lambda i so I have these k eigenspaces corresponding to eigenvalue lambda 1, etc lambda k these eigenspaces are subspaces I can talk about the sum of these subspaces.

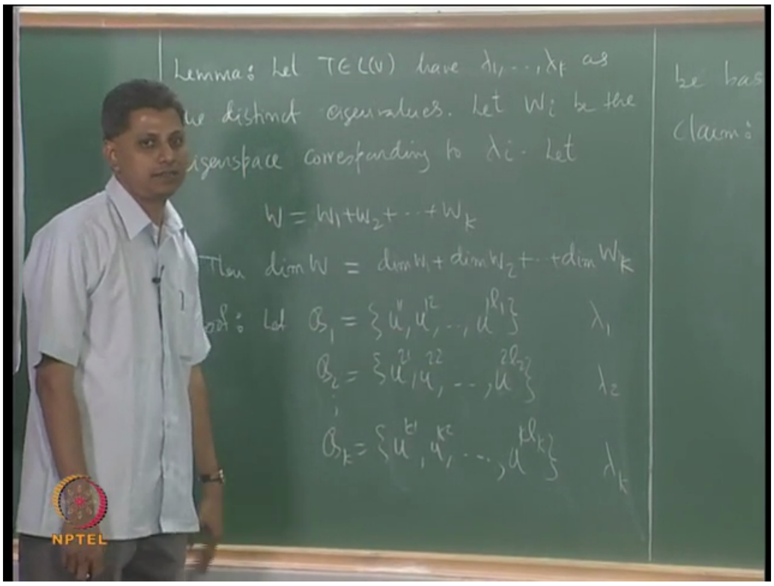
Let W equal to W 1 plus W 2 etc plus W k, I take the sum of these k eigenspaces. Now remember that dimension of W 1 plus W 2 is general is not equal to dimension W 1 plus dimension W 2 you need to remove the dimension of W 1 intersection W, okay but in this case the dimensions add up in this case the dimension will add up that happens if the subspaces are eigenspaces that is what we are trying to do here.

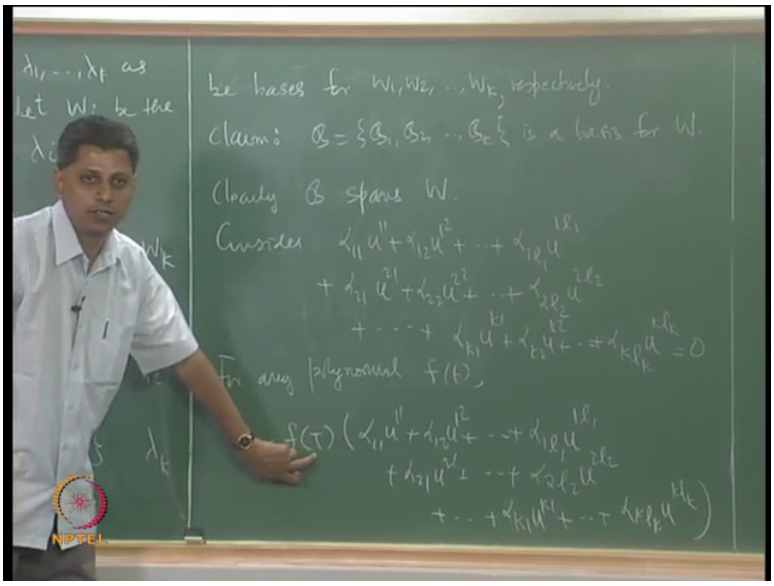
So what is the conclusion then dimension W is dimension W 1 plus dimension W 2 etc. In particular what this means is that eigenvectors correspond to distinct eigenvalues are linearly independent, okay this is this statement is encoded in this eigenvectors corresponding to distinct eigenvalues are linearly independent, okay okay.

So let us proof this then we will use this in characterizing diagonalizability. All that I will do is take a basis for W 1, take a basis for W 2, etc basis for W k show that the union is a basis for the sum W, is that okay? We will take a basis for W 1, basis for W 2, etc W k take the union I will show that that is the basis for W then it follows that dimension of W is the number of elements in B 1 plus the number of elements in B 2 etc number of elements in B k that is precisely this, okay.

So let us write down now basis explicitly let B 1 equal for the subspace W 1 corresponding to lambda 1. So B 1 I will call it u 11, u 12, etc u 1l 1 the first subscript denotes that it corresponds to that eigenvalue first subscript denotes it corresponds to eigenvalue lambda 1, okay so this is corresponds to eigenvalue lambda 1, B 2 is u 21, u 22, etc u 2l 2 this corresponds to lambda 2 etc B k u k1, u k2, etc u kl k this corresponds to the corresponds to eigenvalue lambda k.

(Refer Slide Time: 21:37)





Let these be basis ordered basis for W 1, W 2, etc W k. I must show that I must show that this is a basis for W, okay. Spanning set we will (())(22:12) of quickly linear independence is what is a little difficult in this problem little more involved than the other one spanning set. I want to show that this union this is the spanning set for W okay take anything in W little w that is that is W 1 plus W 2 etc W k each W 1 in term can be written in terms of these so it is clear that this is a spanning set.

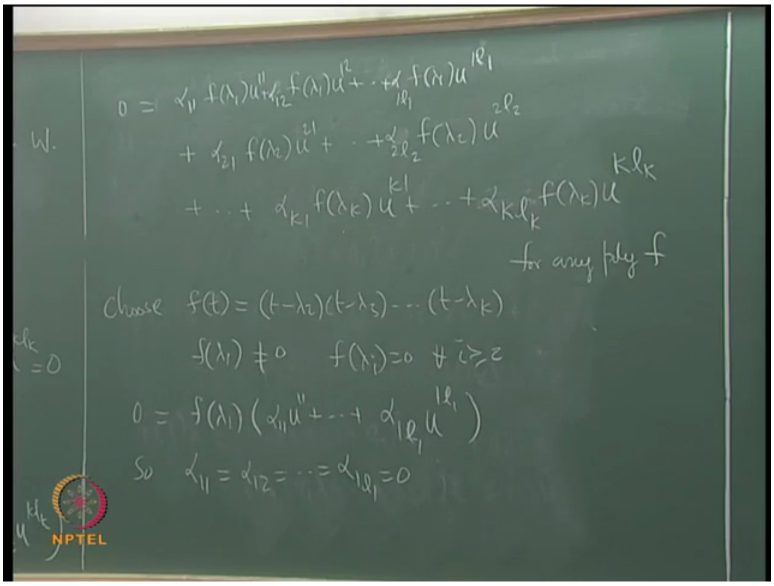
Clearly this is this spans W linear independence is see if you show that this is the basis for W, I am assure that it is a spanning set and it is a linearly independent set. The fact that this is a spanning set is straight forward for the following reason take any little w in W then by definition that W is sum W 1 plus W 2 etc W k but look at W 1 the first term that is the linear combination of these etc so this W is a linear combination of these vectors that is those vectors in script B and so this spans W linear independence.

Consider so we need to solve linear independence so we need to consider a combination. So consider alpha 11 u 11 plus alpha 12 u 12 I choose the scalars also according to the superscripts etc alpha 1l 1 u 1l 1 plus alpha 21 u 21 alpha 22 u 22 plus etc plus alpha 2l 2 u 2l 2 plus etc last term alpha k1 u k1 alpha k2 u k2 etc alpha k what is that last l k u kl k suppose this is 0.

I must show that each of the scalars is 0, okay consider this combination I must show that each scalar is 0 it would then follow that these vectors u 11 u 12 etc u 1l 1 u 21 u 22 etc u 2l 2 u k1 u k2 etc u kl k they are linearly independent, okay. Let us this is this vector is 0 vector f of t is take any polynomial for any polynomial f of t I will consider f of capital T f of capital T is also linear I will apply f of capital T on this of this whole thing, okay.

Alpha 11 u 11 plus alpha 12 u 12 etc alpha 1l 1 u 1l 1 this time I will just write down okay does not matter alpha 21 u 21 etc alpha 2l 2 u 2l 2 plus etc the last one last one is alpha k1 u k1 plus etc plus alpha kl k u kl k, f of T is also a linear operator so any linear operator has a property that its action on the 0 vector this is 0 vector so 0 is this, f of T is linear because T is linear I will apply this to each term, okay.

(Refer Slide Time: 26:23)



So can you see that the first term will be alpha 11 f T u 11 use the previous lemma it is f of remember u 11 is an eigenvector corresponding to eigenvalue lambda 1 all these u 11 u 12 etc u 11 1 they comes from they come from the eigenspace corresponding to the eigenvalue lambda 1 so each of this is an eigenvector corresponding to eigenvalue lambda 1 apply the previous lemma f T of x equals f of lambda x if lambda is the eigenvalue.

So the first set of terms is it clear? That it is alpha 11 f lambda 1 u 11 plus f lambda 1 u 12 etc u 11 1 first terms will go with f lambda 1 plus the second ones will go with the f lambda 2 the last one will go with f lambda k do you agree with this? I forgot the constants here, here also, okay is it clear the first set of terms coming from this bracket the first one here the first set of terms here they will go with f lambda 1 because each of those vectors u 11 u 12 etc u 11 1 each of those vectors is an eigenvector corresponding to the eigenvalue lambda 1 and I am appealing to the previous lemma, okay.

This is true for any polynomial f this is true for any polynomial f. I will make now particular choices of f and then show that this is 0, another choice I will show this is 0, etc okay. Suppose I have one choice for which the second set of terms, third set of terms etc the last set of terms vanish this only remains does it follow that those scalars those scalars will be 0, okay I will show such polynomial in such a way that this is not 0.

So it will follow that alpha 11 alpha 12 etc alpha 11 1 is 0 apply the next polynomial I will choose it in such a way that f lambda 2 is not 0 it will follow that these coefficients are 0, etc

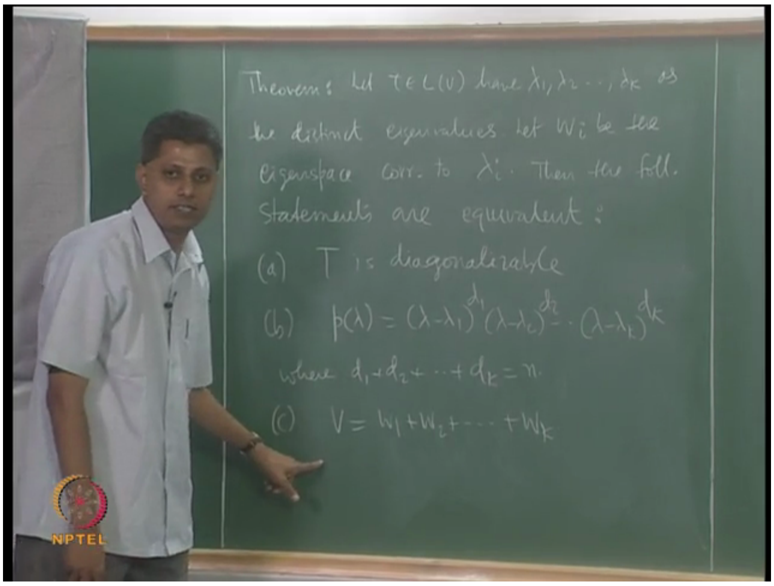
I will go back and substitute into this equation I am sorry I want to conclude each scalar is 0 so it will follow that these are independent, okay.

Now what is that polynomial? For the first one I will do it for the first one the rest is similar. For the first one I will choose f of t to be see I want these to be 0 I want a polynomial I want the polynomial to have lambda 2 etc lambda k to be 0 give me one choice t minus lambda 2, t minus lambda 3, etc if you choose this then f of lambda 1 f of lambda 1 is lambda 1 minus lambda 2 lambda 1 minus lambda 3 etc product lambda minus lambda k these are distinct so that is not 0, okay.

So f of lambda 1 is not 0 f of lambda 1 is not 0 but f of lambda i equal to 0 for all i greater than or equal to 2. So I will take this polynomial apply it to this equation then I get okay I take this polynomial apply this polynomial to this equation then I get the second set of terms are 0, etc I have only these terms remaining, I can write it as f lambda 1 outside into the rest of them alpha 11 u 11 etc alpha 1(k) l 1 u 1l 1 only the first set of terms will remain f lambda 1 is not 0 so this can be cancelled, look at the rest that is 0 but these are linear independent, right they form a basis for W 1 these are linear independent from this it follows that the first set of coefficients alpha 11 alpha 12 etc alpha 1l 1 first set of coefficients must all be 0.

So you apply the second polynomial which is T minus lambda 1 into T minus lambda 3, T minus lambda 4, etc T minus lambda k then you can show the second set of coefficient 0 etc. So it follows that each of these scalars started with this equation follows that each of the scalars is 0 and so these vectors are linearly independent so this is the basis for space W, okay.

(Refer Slide Time: 32:42)

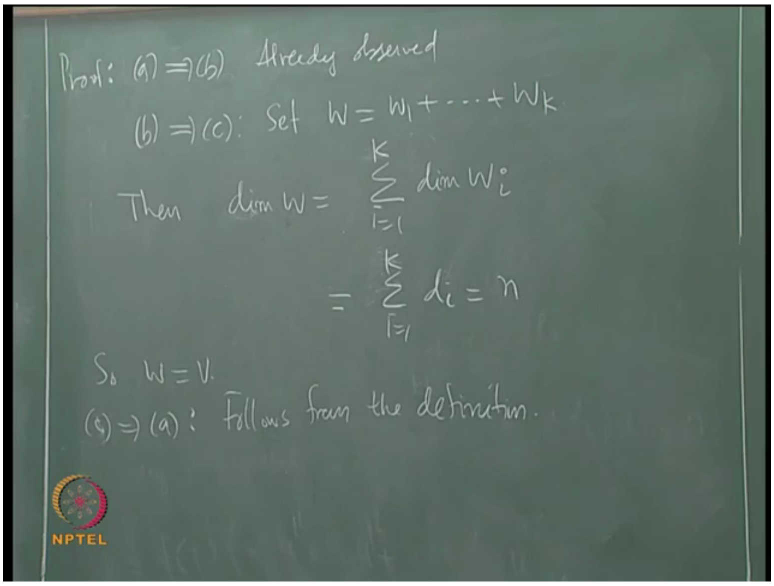


As I mentioned earlier in particular this means that eigenvectors corresponding to distinct eigenvalues are linearly independent we have proofed something more, okay let me proof this theorem then theorem that characterizes diagonalizability the framework is as before T is the linear operator in a finite dimensional vector space with dimension n over f V is defined over f I have lambda 1, lambda 2, etc lambda k as the distinct eigenvalues let W i be the eigenspace corresponding to the eigenvalue lambda i then the following statements are equivalent.

I will give two conditions that are necessary and sufficient for T to be diagonalizable first statement is T is diagonalizable. Second is the condition involved in the representation of the characteristic polynomial characteristic polynomial I am denoting by p I will write p of lambda it is lambda minus lambda 1 to the d 1 lambda minus lambda 2 to the d 2 etc lambda minus lambda k to the d k.

Where d 1 plus d 2 plus etc plus d k equals n I am assuming that n is a dimension of the space V. The last condition is in terms of sums of the eigenspaces, if you look at the sums of the eigenspaces call it W then this W the subspace is the whole of space V this is the last condition we have already observed that a implies b, T is diagonalizable then we have seen that the characteristic polynomial has this form from b implies c, okay.

(Refer Slide Time: 35:41)



Proof a implies b, already observed b implies c follows because okay you can set W to be the sum of these subspaces then by the previous lemma it follows that dimension of W is summation i equals 1 to n dimension W i this comes from the previous lemma because these are eigenspaces corresponding to distinct eigenvalues but what is the condition on look at condition V the condition on these numbers d 1, d 2, etc d k is that their sum is n.

So dimension W is summation i equals 1 to n d i that is n. So I have a 1 to k there are k subspaces ya so this sum is n so but we know that the sum of two subspace is again subspace, W is a subspace of V having the same dimension so W must be the whole of V that statement c the sum is equals to W, W is equal to V we have shown so b implies c also holds c implies a c implies a follows from the definition of what we mean by diagonalizability is that clear.

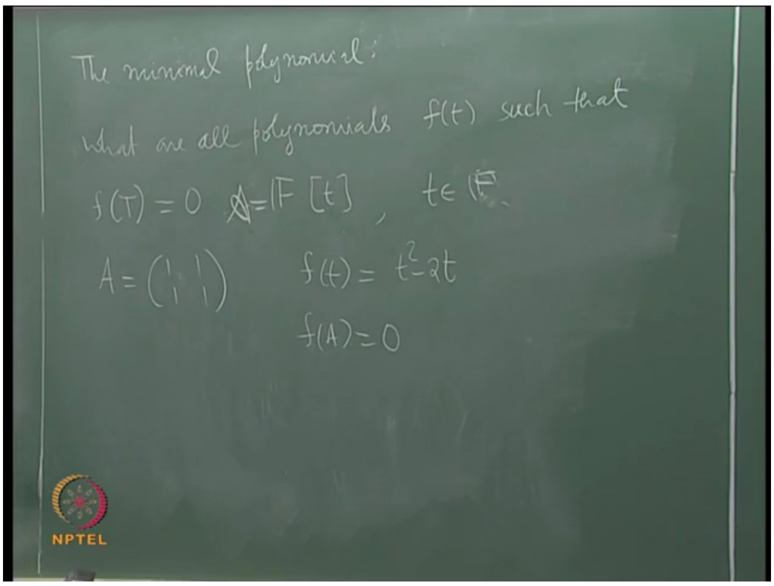
I want to show T is diagonalizable that is I want to show that there is a basis script b of V such that each vector from b is an eigenvector of T that is diagonalizability I saw this yesterday T is diagonalizable if and only if there is a basis b for b for V each of whose vector is an eigenvector for T. I know that V is the direct sum of is a sum of these subspaces look at the construction that we did earlier take a basis b 1 for W 1, b 2 for W 2, etc b k for W k each of the basis has the property that each of this basis for the subspaces b 1, b 2, etc b k have the property that their elements their vectors in b 1 for instance is an eigenvector corresponding to lambda 1, the vectors in b 2 are eigenvectors corresponding to the eigenvalue lambda 2, etc.

The combination is a basis for V the union of these basis for the subspaces is a basis for V and so each vector of this basis is an eigenvector of V is an eigenvector of T so T is diagonalizable. So can I just say c implies a follows from really the definition which we saw yesterday really the definition that there is a basis b for V each of whose vectors is an eigenvector for T, okay.

So this is one characterization that is necessary sufficient condition for T to be diagonalizable one is that the characteristic polynomial can be written as a product of powers of linear factors, second is that the whole space V can be written as the sum of these subspaces the subspaces being the eigenspaces, okay this is one characterization, we will also look at another characterization involving the so called minimal polynomials, okay that I will do in the next one or two lectures.

But before I conclude I want to atleast mention what is the minimal polynomial we want to know when precisely a linear transformation is diagonalizable, okay one answer has been given here two answers really look at the eigenspaces, take the sum, verify if that is the whole space V, the other thing is look at the characteristic polynomial verify if it is a product of powers of linear factors then you know that T is diagonalizable.

(Refer Slide Time: 40:23)



There is another answer as I told you which comes in terms of the minimal polynomial, okay so let me atleast give the definition of the minimal polynomial. Remember that we are looking at an operator a single linear operator T and we are analysing the operator T. In this study what is important is to identify classes of polynomials which have the property identify classes of polynomials let us say f of t such that f of capital T is 0.

What are all polynomials f of t such that f of capital T is 0? Now why why is this statement true that will be clear only a little later you will get the connection between minimal polynomial and the characteristic polynomial then it will be clear as to why these polynomials are important, okay but let me atleast give the concept of the minimal polynomial coming from this, okay remember f is a polynomial its coefficients come from the underlying field, okay.

Now if you look at if you look at F t call that script A look at F t call that script A then so what is this F t F t is the set of all it is it has an algebra structure, it is a set of all polynomials in a single variable t in a single real variable t, I should actually write F but for the sake of convenience I am writing it R single real variable t so I should actually write R of t, okay may be let me go back and change this to F.

So it is either for our discussion let us say it is a set of all real polynomials over a single variable t single real variable t, the coefficients are real, the variable t is also real I am using script A for that it is what is called as an algebra you know that it is a euclidean domain set of all polynomials is a euclidean domain it is a commutative of ring where you can do euclidean algorithm it is a commutative ring where euclidean algorithm can be applied.

It is it is an euclidean domain which has a property that you know the concept of an ideal concept of an ideal in subring, okay does not matter if you do not know you will learn it now this semester sometime. What can be shown is that an ideal is a subring of see this is what is called as an algebra an algebra is an algebraic structure where you can do multiplication of vectors, okay.

So an algebra is something more than a vector space where there is also a possibility of multiplying vectors. Now multiplication here is multiplication of polynomials multiplication of polynomials you know term by term multiplication of polynomials is term by terms. So one could do multiplication of polynomials it also has one or two little (())(43:55) but let us not worry about that this is an algebra this is a euclidean domain where I can do where product of vectors make sense, in such a euclidean domain the notion of an ideal comes, an ideal is a sub bring an ideal is a sub bring which has a property that if I take an element from the ideal and an element from outside the ideal that is f of T then the product will belong to the ideal, okay the product will belong to the ideal.

Ideals in a euclidean domain have the property that they are generated by a single unique polynomial ideals in a euclidean domain are characterized with the property that they are generated by a single unique element which means anything in the ideal is a multiple of a specific polynomial anything in an ideal given an ideal anything in that ideal is a multiple of a unique element coming from that ideal, okay.

I need this property one can also do without this property for the minimal polynomial. I wanted to find the minimal polynomial first of all the question is given a linear transformation T, does there exist a polynomial f such that f of T is 0, okay. I will give two answers one for a specific example linear transformation matrices they are equal actually so I will take this matrix 1 1 1 1 I want you to consider this polynomial f of t equals t square minus 2t, f of t is t square minus 2t how do I get this?

At the end of the next two lectures you will also be able to write down such polynomials, t square minus 2t then you can verify that f of A is a 0 matrix A square is equal to 2A, A square is equal to 2A for this matrix that is the reason why I choose t square minus 2t then f of A equal to 0. So given a linear transformation this make sense the question does there exist a polynomial f such that f t equal to 0 make sense I have given one example.

I will actually prove it I will actually prove that given a linear transformation on a finite dimensional vector space there is a polynomial which has the property that f of capital T equal to 0 tomorrow and then may be define the minimal polynomial and how it is related to the notion of diagonalizability, okay let me stop here.