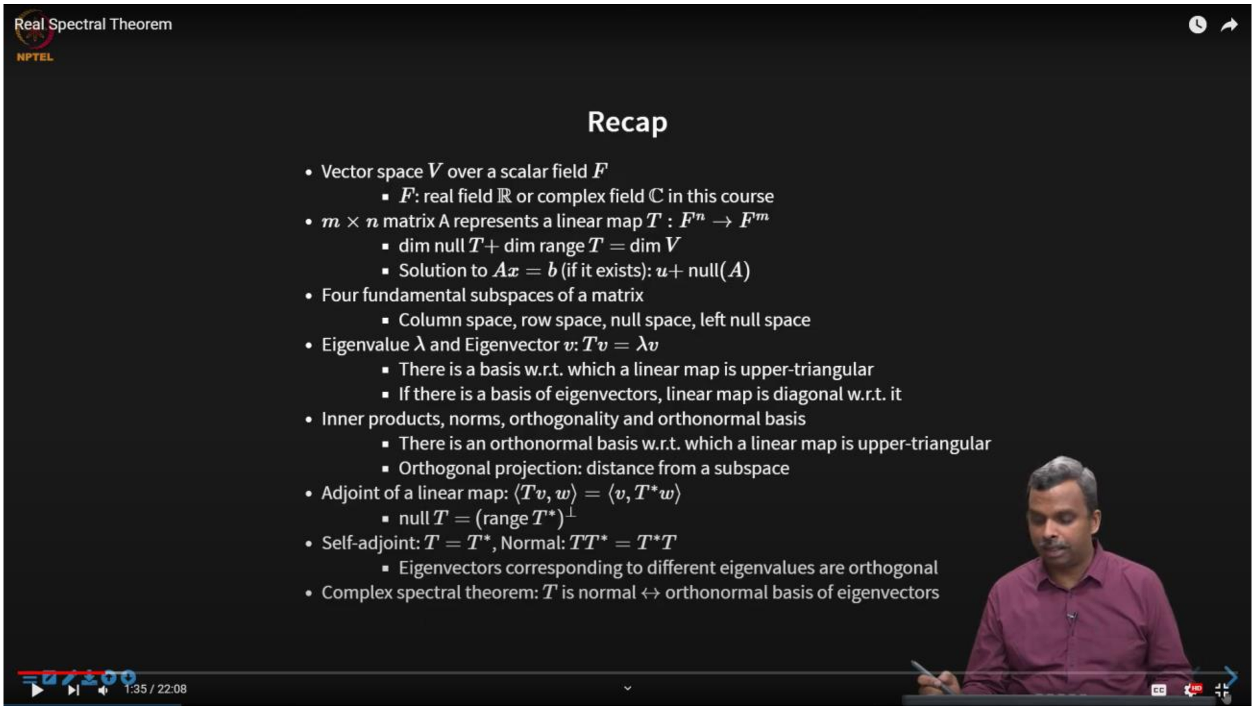
Applied Linear Algebra
Prof. Andrew Thangaraj
Department of Electrical Engineering
Indian Institute of Technology, Madras

Week 10
Real Spectral Theorem

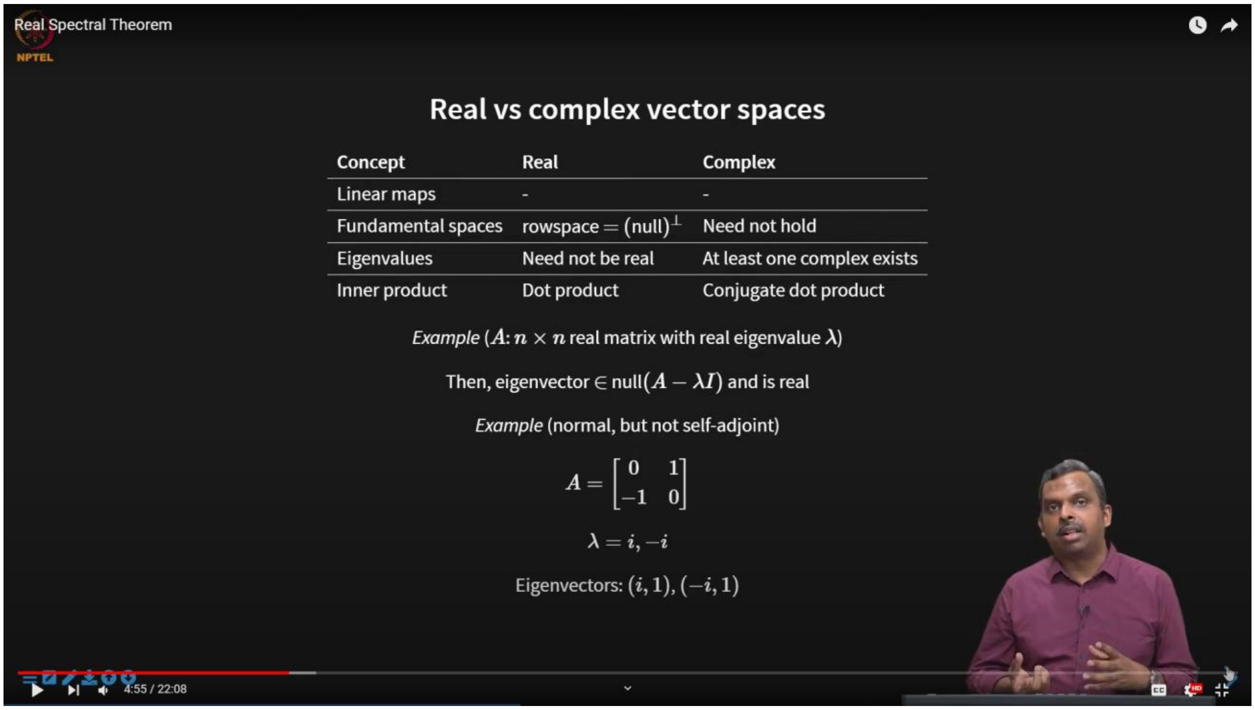
Welcome to this lecture. We are going to talk about the real spectral theorem. In the previous lecture, we saw the complex spectral theorem which was a very simple characterization of what a normal matrix is, what a normal operator is going to look like, okay? Any normal operator is going to have that form, it is going to be diagonal over an orthonormal basis. So if you want to express its matrix for instance, you can write it in terms of that very simple expression, which is $(\lambda_1 e_1 \overline{e_1^T} + \lambda_2 e_2 \overline{e_2^T} \ldots)$, okay? So we saw that very easy and simple result. A couple of illustrations of that as well. So now when you think of a normal matrix, its eigenvalues could be real or complex. And self-adjoint is only a subset of the set of normal matrices. And the matrix, I mean it's usually over the complex vector space, right? So there is an interest now to looking at: how do you exactly characterize self-adjoint operators, right? So that's one interest. And the other interest is: what about real vector spaces? Supposing you have a real matrix, is there something you can do within the real field without going to the complex field? So what is the kind of result that one can say? And that's sort of like the self-adjoint operator. What is normal in the complex vector space is the self-adjoint operator in the real vector space. So that is this theorem which we will see in this lecture, okay?

(Refer Slide Time: 01:35)



A quick recap. I think I mentioned the recap earlier. So let us look at what is it that is different between a real and a complex vector space. So in general, most of the properties we looked at in this class are going to hold whether the vector space is real, over the real field or the complex field, right? So that much is more or less true. You do not have to worry so much about it. But there are these small minor differences here and there which makes things different. One of the crucial differences is the fundamental theorem of algebra, which is: any polynomial is guaranteed to have, a polynomial with complex coefficients is guaranteed to have a complex root. You cannot say that about real numbers. If it's a polynomial with real coefficients, it need not have a real root, right? So we know that there are polynomials like that. So that makes some changes in, introduces some differences when you talk about eigenvalues and all that. But there are also a few other subtle differences. I thought I will make a small list of what is out there. So if you… When we studied linear maps, and you know fundamental theorem of linear maps and all that, there's absolutely no difference between… Everything held, I mean, every result we derived holds. It held for both, you know, whether the vector space was the real field or complex field there's no problem. But when we went to fundamental subspaces, some fundamental spaces of a matrix and all that, there was this subtle little result how, you know, in a real, for a real matrix, the row space became orthogonal to the null space. For the complex, it's not exactly that, it's the conjugate row space which becomes orthogonal to the null space, right? So that's some minor differences there. That's primarily because the inner product is different. The inner product in real space is just the dot product, inner product in the complex space is the conjugate dot product, so which is, which makes a big difference there.

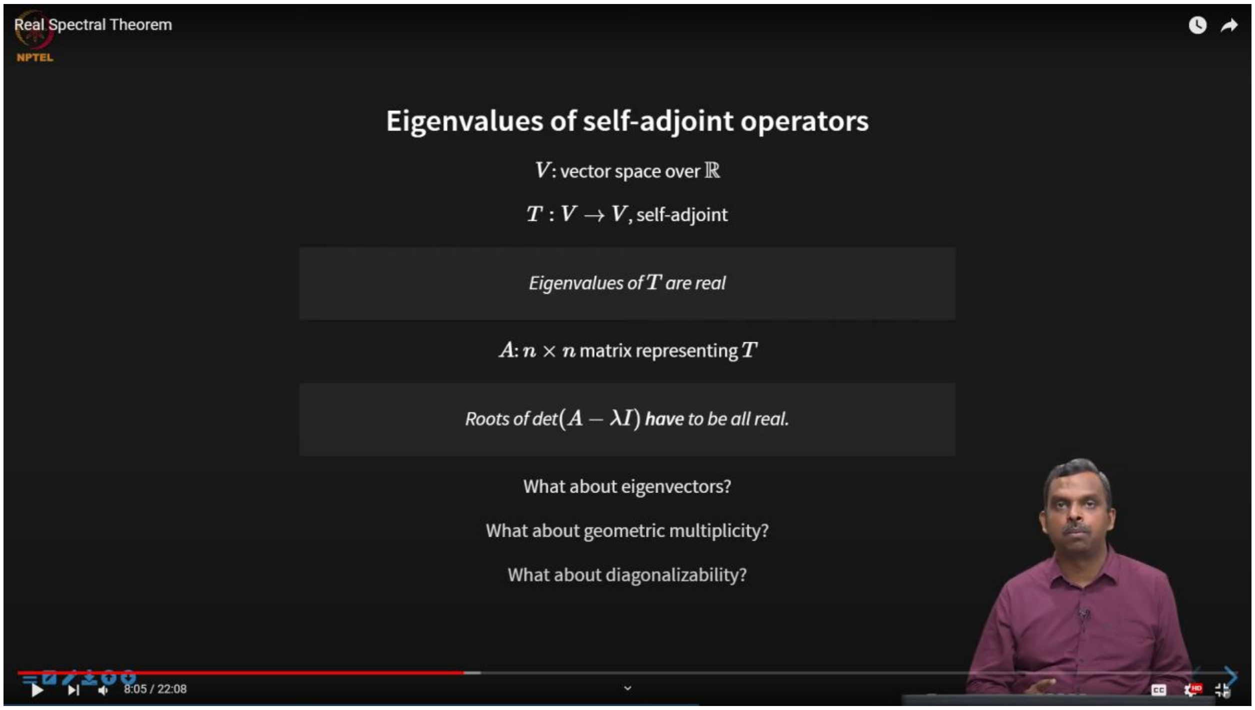
(Refer Slide Time: 04:55)



And the other clear difference is eigenvalues, right? So you need not have real eigenvalues for a real matrix, like, for what, for instance for a complex matrix or even the real matrix, there's always a complex eigenvalue. So let me just give a couple, one example to illustrate what happens with these eigenvalues and eigenvectors and all that. Particularly, see, all these are eigenvalues, what about eigenvectors? Should they be real or complex? What happens? So supposing you start with a nxn real matrix. And let's say it has a real eigenvalue lambda. Once the eigenvalue becomes lambda, real lambda, then the eigenvector is actually a null space vector in $(A - \lambda I)$. Now A is real, λ is real, I is real. So $(A - \lambda I)$ is real. You evaluate null space of that, you will only get real vectors, right? So the eigenvectors are real, okay? So when you have a real matrix and a real eigenvalue, eigenvectors are also going to be real, there is no problem with that, right? So there are maybe other ways to prove these things, but it's clear that that is true, right? So that's interesting. Now what… But that need not be true, right, so you can have, even, for normal matrices which are not self-adjoint. Like for instance this simple example. $[0\ 1;\ -1\ 0]$. You can have only complex eigenvalues, there is no real eigenvalue for this matrix. There are two complex eigenvalues: $i, -i$. And when, once the eigenvalue itself goes complex, even for a real matrix, the eigenvectors are going to be complex, right? So otherwise it won't make sense, right? So you will have eigenvectors being complex. So these are the two types of situations you can have typically when you have a real matrix. If it has a real eigenvalue, the eigenvector will be real. If it has a complex eigenvalue, then the eigenvector will be complex, okay? So this is something to keep in mind when we think of eigenvalues, eigenvectors for real versus complex, okay? So this is a simple, simple fact.

Okay. So what about self-adjoint operators? Like I said, I mean a normal operator is very easy to characterize. And I was telling you that, you know, self-adjoint operators similarly in the real space are very easy to characterize as well. So the crucial result there which we have seen already is that: for self-adjoint operators, whether they have complex entries or real entries, once they become self-adjoint, the eigenvalues are real, okay? So that makes a significant, that's an important stepping stone into looking at, you know, symmetric or self-adjoint operators in vector spaces over the real field, okay? So this is a nice guarantee to have. When you have a self-adjoint operator, whether it's real or complex, the eigenvalues are real, okay? So in particular, if you are going to look at vector space over real and self-adjoint operators, this is very very useful, okay? All right. So that is something good to know. So now what happens is, okay, so you have an operator T which is self-adjoint and then you go to a basis and find a matrix for T, an $n \times n$ matrix, okay? So let us say vector space V is dimension n, as usual. A is an $n \times n$ matrix. We know how to find all the eigenvalues of A, right? How do we do that? We take $|A - \lambda I|$, it's a polynomial in λ, you find the roots, okay? So now because T is self-adjoint, because A is symmetric, this polynomial can only have real roots, it cannot have complex roots, okay? So if it has complex roots, then that root will become an eigenvalue and that's wrong, right? It's a self-adjoint operator, it cannot have a complex eigenvalue, so this clearly tells you that this polynomial will have all real roots, okay?

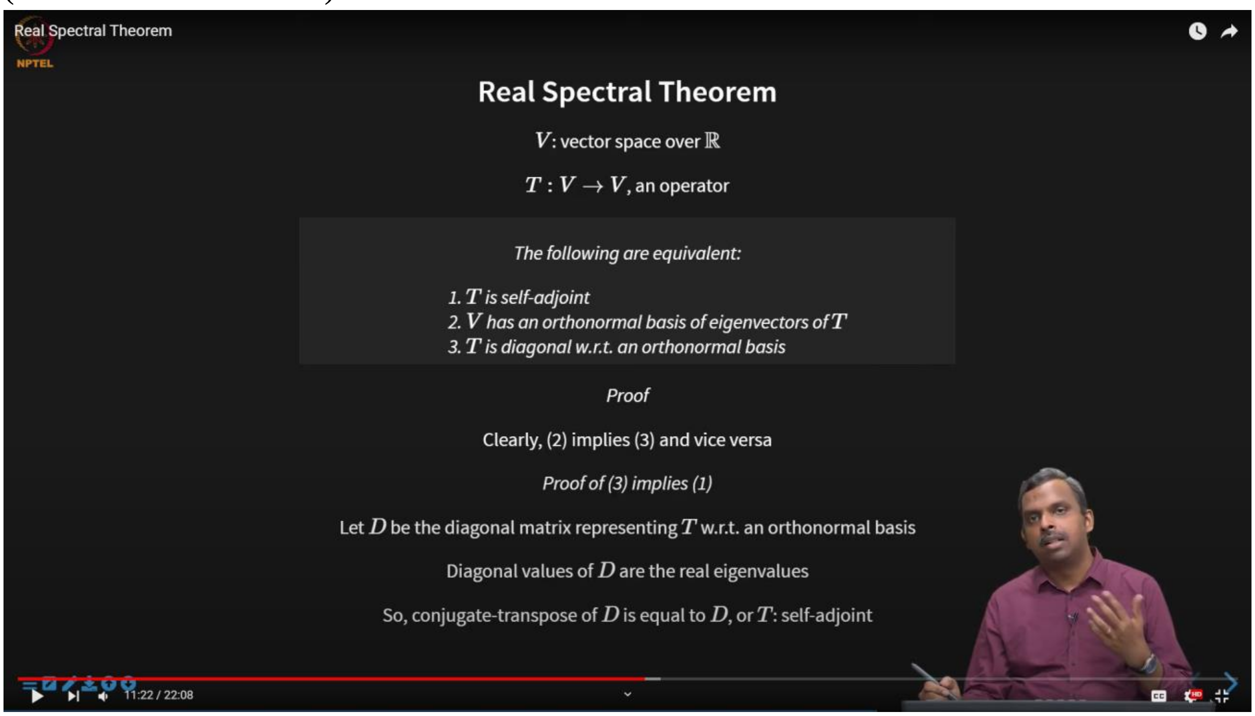
(Refer Slide Time: 08:05)



Now your book has a different proof for a similar result without using determinants, so determinants are not so crucial here. I just put it out here for getting us a simple little argument and getting over the proof. Otherwise, if you do not use determinants here, you have to do a long argument involving quadratic expressions with operators and all that, which is okay, it is, I will leave you to read it in the book. But for our purposes, since we are anyway okay with looking at determinants, we are simply going to say this will have all real roots, okay? So that's an easy result to see because eigenvalues of T are real, okay? Now what about eigenvectors, okay? See, so now this… We know already that there is a good start we have made for a self-adjoint operator over a real field. All its eigenvalues are going to be real. Now what about eigenvectors? For every real eigenvalue, I am guaranteed to have at least one real eigenvector, okay? But what about geometric multiplicity? Will algebraic multiplicity be equal to geometric multiplicity or can there be cases where you have a self-adjoint operator, a symmetric matrix, but its geometric multiplicity is strictly less than the algebraic multiplicity? Then what will happen? It will not be diagonalizable, okay? So can you have a self-adjoint operator which is not diagonalizable, okay? Even over the real field, okay? So that's an interesting question that one can answer. And all these questions are beautifully settled in the real spectral theorem which we will hopefully see in the next slide, okay?

So this is the real spectral theorem, okay? So you have a vector space over ℝ, okay? Over the real field. And T is an operator, okay? The following are equivalent. When you say equivalent, again, any one implies the other, they are all equivalent, okay? If T is self-adjoint, then there is an orthonormal basis of eigenvectors of T in V, in the real vector space, okay? And of course T is diagonal with respect to this orthonormal basis. And this is implied, I mean this is if and only if, okay? If T is self-adjoint, then T is diagonal with respect to an orthonormal basis. If T is diagonal with respect to an orthonormal basis, then T is self-adjoint. So it's a complete characterization of what a self-adjoint operator will be, okay? What a symmetric matrix will be, it will be something which is diagonal with respect to an orthonormal basis. Nothing else, okay? So this simplifies the description of self-adjoint operators. Not only that, it also tells you that any self-adjoint operator is diagonalizable. Its geometric multiplicity of any eigenvalue will be equal to the algebraic multiplicity. All of that is guaranteed by this real spectral theorem.

(Refer Slide Time: 11:22)



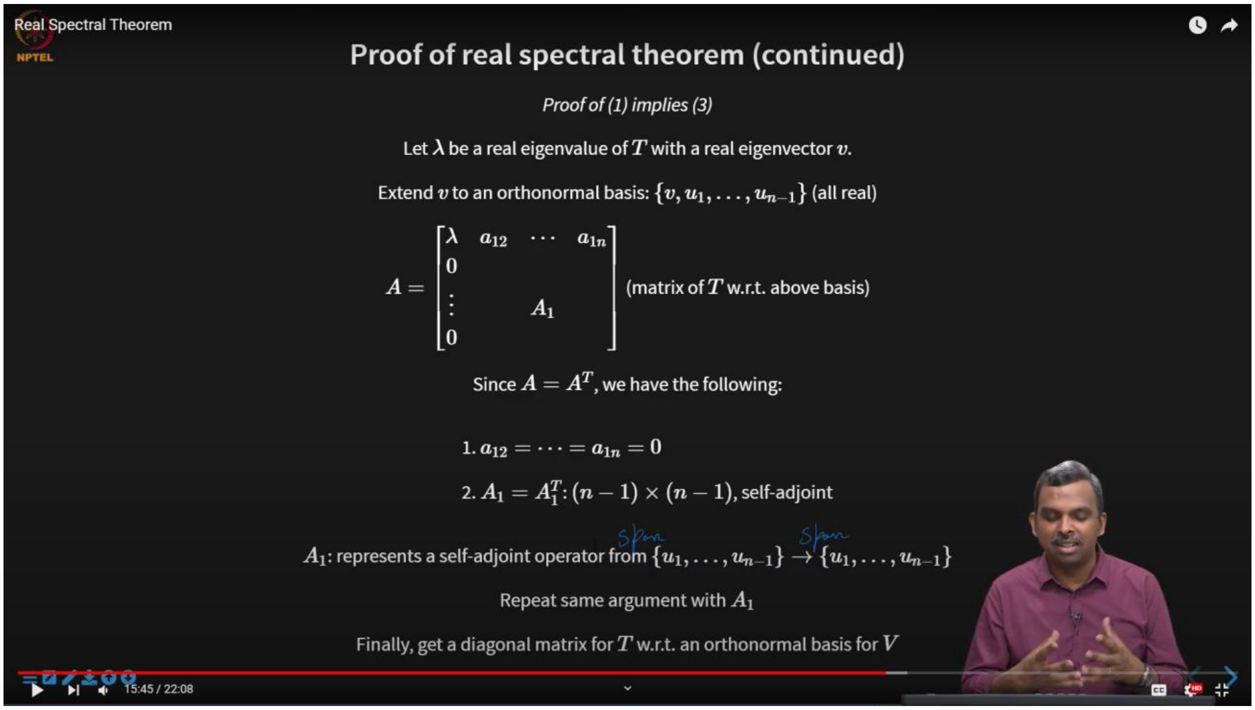
So it’s a very powerful result and you can see the analogy between this and the complex spectral theorem, right? What is normal operators in vector spaces over complexes is sort of captured, that spot is captured by self-adjoint operators in vector spaces over the reals, okay? So it’s a very simple and elegant result. It tells you that a symmetric matrix is always diagonalizable. You cannot have a symmetric matrix which is not diagonalizable, okay? So that's a nice result to have. It's, a couple of points, I mean, a lot of people get confused by this, so I'll re-emphasize that it's not necessary that a symmetric matrix should have distinct eigenvalues, right? So we saw before, if an operator has distinct eigenvalues, then it is diagonalizable. A symmetric matrix, whether it has distinct eigenvalues or not, it's always diagonalizable. So that's the power of this result, right? So any symmetric matrix is diagonalizable. Any, you know, self-adjoint operator is diagonalizable. So that's the power of this result. So simple examples are there, but anyway, that's what I mean. So keep that in mind, okay?

So we will do a quick proof. And proof is very similar to what we did in the normal case. So first of all, two and three sort of imply each other. We can forget about them. And three implies one is also very easy to show, okay? So let us say if you have T being diagonal with respect to an orthonormal basis, you will have a diagonal matrix representing T. Now all the diagonal values are the real eigenvalues of T. So they are all going to be real. So if you take conjugate transpose of D, you will get D itself, okay? Which is again real. So T becomes self adjoint, okay? So to show that T is self-adjoint if it has a diagonal matrix with respect to an orthonormal basis is easy, okay? So this part is easy, it's very similar to what we did in the normal case, okay? So in this case, because D is diagonal and real, the conjugate simply becomes equal to itself, okay? So in the previous case, the normal case, we only had commuting because it could be complex, right? So normal we do not know eigenvalues are real. Here the eigenvalues are real, so it becomes equal, okay? When you take conjugate transpose, okay? So that is good.

So a slightly non-trivial result is one implies three, okay? So same thing in the complex spectral theorem as well. So here we will follow an approach similar to, I mean sort of similar to what we did for the complex spectral theorem. The complex spectral theorem we started with the orthonormal basis which gave an upper triangular matrix, right? And then we showed that that is diagonal, okay? So here also we'll do something very similar. But instead of just invoking Schur’s lemma or Schur’s theorem which we can't because that really works for complex spaces. So here we are in real space. So we will just sort of rederive it slowly step-by-step. If you remember the proof of Schur’s theorem, the orthonormal basis, you start with one eigenvalue and then, you know, start going step-by-step from that. So something like that we will do here, okay?

So you start with one eigenvalue. I know it will be a real eigenvalue for T because it’s, T is self adjoint and it has a real eigenvector. Once you have a real eigenvalue, it will have a real eigenvector, it’s easy to see, okay? So we can extend v to an orthonormal basis, okay? The orthonormal is important. We extend v to an orthonormal basis and all of these are real, right? This is over the real field, okay? $\{v, u_1, u_2, \dots, u_{n-1}\}$. So in this basis, if you find a matrix for T, okay, the first column is going to be λ followed by all 0s. And the next columns will be various things. Now remember, this is an orthonormal basis with respect to which you found a matrix. Now this is real, everything is real here. So T has, T is self adjoint. So A has to be equal to A^T, okay? So $A = A^T$ then clearly the top row, a_{12} to a_{1n} have to be all 0, right? So it has to be equal to its transpose, it has to be symmetric, so that has to be true. And and then A_1 has to be equal to A_1^T, right? So this sub-matrix here, the $(n-1) \times (n-1)$ sub-matrix here has to become self-adjoint or symmetric in some sense, right? So these are the simple facts you get just by going to this. So previously we couldn't conclude, you know… So when we did not have an orthonormal basis and we did not have a self-adjoint operator, we couldn't conclude these things were 0. We had to live with them, you know… I had to define some other operator when I moved to the $(n-1)\times(n-1)$. In this case, when I go to $(n-1)\times(n-1)$, I simply get another symmetric matrix, okay? So that's a nice thing to have here. So the top row is all zero, okay?

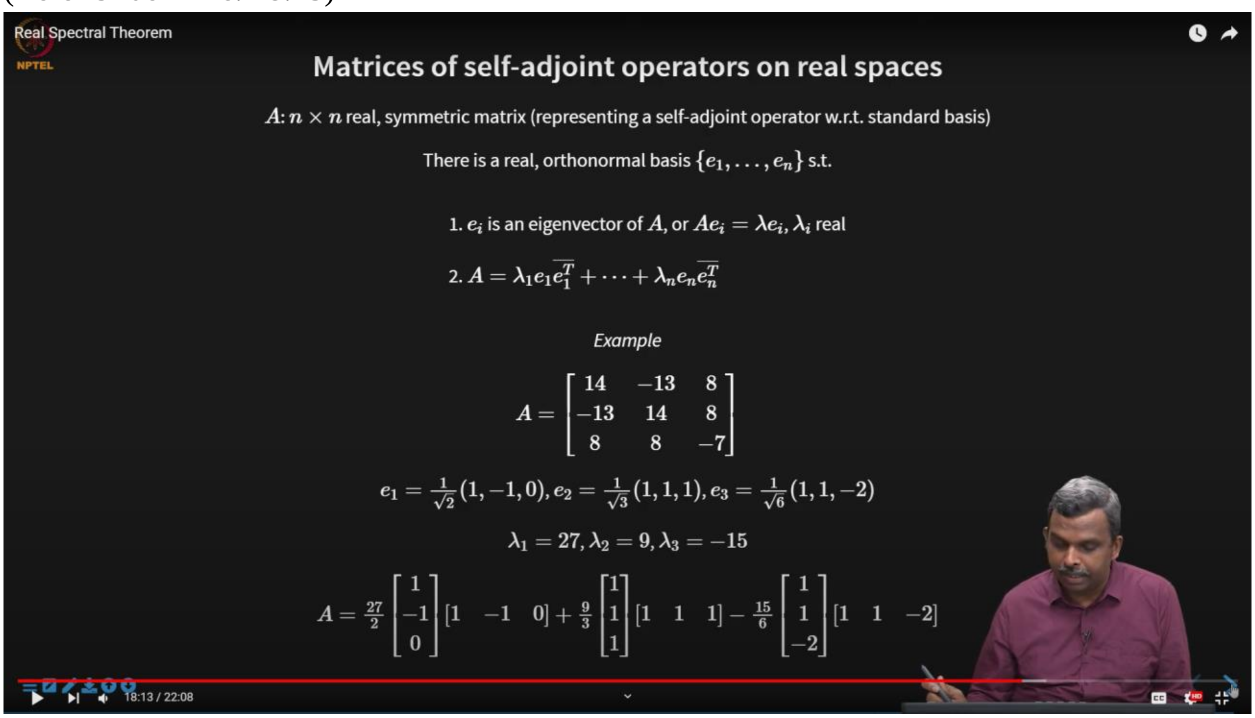
(Refer Slide Time: 15:45)



So notice how the upper triangular is slowly becoming diagonal, okay? So once I come to A_1, you notice, right, this A_1 represents the self adjoint operator from u_1 to u_{n-1}. So what is this? When I say from this, I mean really from span of this to span of this, okay? The span is sort of left out there, okay? So you can see that, see, because notice. See, A_1 the top row is all zero, right? For A_1. So when you use, when you apply u_1 on it, you are never going to get anything from v, okay? So v will not play a role when A_1 alone is operating. So $\{u_1, \ldots, u_{n-1}\} \to$ $|u_1, \ldots, u_{n-1}\}$, if you look at the span of those two, this A_1 will represent a self-adjoint operator, okay? So you have come from dimension n to dimension $n-1$ with some diagonal λ extension, okay? Then you repeat whatever you did with A for A_1, okay? So you can even write a program for doing this, it is easy to go step-by-step and step. And in every stage you will have an orthonormal basis. And A will start being diagonal diagonal diagonal, you will get smaller and smaller, this A_1 will become A_2 which is $(n-2)\times(n-2)$ then A_3, A_4 etc. And finally you will end up with an orthonormal basis with respect to which the operator T itself is diagonal, okay? And that's the result, okay? So this argument you can make in so many different ways. You can make it in a slightly more abstract way with, you know, self… I mean invariant U and $U^{\perp}$ and all that which is what your book does, and some induction sort of argument which can make it a little bit more clean. This one maybe is a little bit more dirty, involves matrices and all that but I think it gives you the crux of the idea of how this orthonormal basis works. So then look at how self-adjoint is so important, otherwise you can't get rid of the upper triangular terms, they go away because of the self-adjoint property, okay? So that's the end of this proof.

So this real spectral theorem is very powerful, particularly when you want to look at matrices of self-adjoint operators on real spaces or basically symmetric matrices. What happens when you have an nxn real symmetric matrix which of course represents a self-adjoint operator, say with respect to the standard basis, okay? Any other basis is fine. You can pick the standard basis. We know from the real spectral theorem that there is an orthonormal basis $\{e_1, \dots, e_n\}$, right, such that e_i is an eigenvector of A. Ae_i is λe_i. λ_i is real so… I forgot to put the i here, $\lambda_i e_i$, okay? And A itself becomes $\lambda_1 e_1 e_1^T + \ \dots$). See, because it becomes diagonal, right? A becomes diagonal with respect to this orthonormal basis. And once it becomes diagonal, you know A can be written in this form $\lambda_1 e_1 \overline{e_1^T} \ + \ \dots + \lambda_n e_n \overline{e_n^T}$. This is similar to the complex spectral theorem.

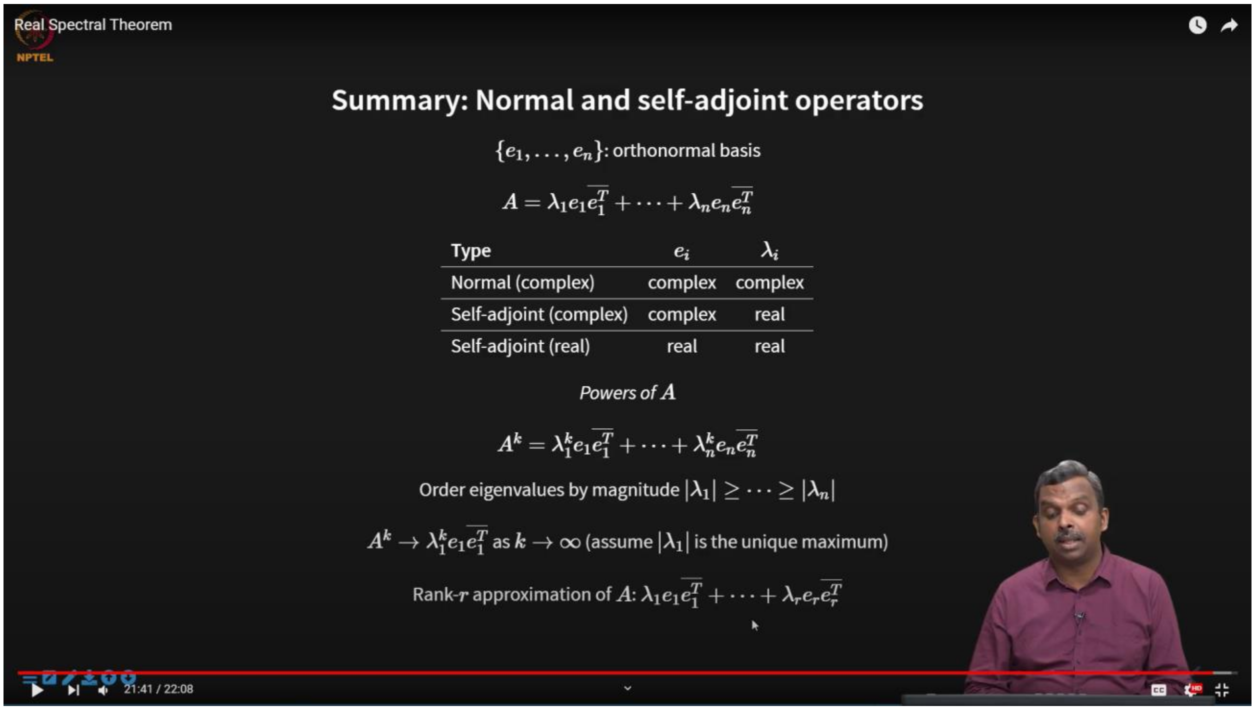
(Refer Slide Time: 18:13)



So we did, so this is a form for any $n \times n$ real symmetric matrix and that really simplifies things quite a bit here. You take any orthonormal basis, pick any λ of your choice, you multiply, you know, do a linear combination like this, each one is a rank one matrix. You add up n of them, you know, multiplied by λs, whatever they are, zero, non zero… They all have to be real though, right? Then you get a symmetric matrix, any symmetric matrix can be decomposed like this. Anything you form like this is a symmetric matrix. So that's what is, I mean, it's sort of easy to see one way. The other way maybe is a little bit tricky. So that is the symmetric matrix and the real spectral theorem here, okay?

You can see an example, this example is from your book. It's a very clean example. You have a matrix here and it has three distinct eigenvalues. It need not be distinct, this is an example of three distinct… It could be that, you know, two of the eigenvalues are the same but still they are diagonalizable, you will get an orthonormal basis with respect to which A is diagonal and you can write it like this. I mean, this form is something that I like. You will see it has a powerful use, okay? All right. So that is good. That takes care of the example, okay? So that's the real spectral theorem. Hopefully you got a feel for how it works, okay? And let's summarize all that we have looked at in terms of spectral theorems for normal and adjoint, self-adjoint operators, okay? So let, everything starts with sort of an orthonormal basis. So now you could be, you know, the vector space could be over complex or real, you know, the normal and self-adjoint operators have a special role there. You can start with an orthonormal basis and you can form a matrix A, let's say, which is $\lambda_1 e_1 \overline{e_1^T} +$ ... in this form, okay? So the e_i's are, you know, expressed in coordinate terms over a standard basis, let's say, and you can form these, some linear combination of rank 1 matrices which give you A, right? So this is a special form. This captures normal and self-adjoint operators.

(Refer Slide Time: 21:41)



And I put a table here to tell you how you can pick e_i, and how you can pick λ_i to get different types of matrices, okay? So if you want to think of normal operators in complex vector spaces, you can allow e_i to be complex and you can pick λ_i complex. If you are thinking of self-adjoint operators in complex vector spaces, e_i's can be complex but these λ_i's have to be real, okay? You can't take λ_i being complex. λ_i have to be real, that makes it self-adjoint. Now if you're thinking of real operator, real self-adjoint operator, symmetric operators in real vector spaces, the e_i's are real, λ_i's are real, okay? So this lets you go one way or the other. Both ways this is all that you have to worry about when you think of normal and self-adjoint operators in complex and real vector spaces. This sort of conveys everything together, okay?

So what are the advantages of a form like this? First thing is: you can do powers of A, okay? So when you do powers of A, you can see, I mean the, any diagonalization gives you that, but this diagonalization is particularly easy to write down, you know? A^k is $(\lambda_1)^k$, everything else remains the same, just $\lambda_1^k, \ldots, \lambda_n^k$, okay? So usually what people do is: you would order these lambdas by magnitude, you know, biggest one comes first and smaller one comes next, etc. So these guys are bigger and all these guys are smaller. Remember, e_i is orthonormal, right? So I have normalized it, so keep that in mind. So as k becomes larger and larger, you see that if this λ_1 is the unique maximum, let's say, then A^k will tend to just a rank 1 matrix with $\lambda_1^k e_1 \overline{e_1^T}$ being the rank 1 matrix, okay? So it's also common to make a rank r approximation of A. Remember, if you order it in magnitude like this, and you want a rank r approximation, a low rank approximation of A for whatever reason, you can simply stop with the first r terms, isn't it? So these kind of ideas have powerful applications in engineering and all that, okay? Quite often you'll have, you'll be dealing with a very, very large matrix. And for various reasons, I mean, you don't care so much about being precise about every entry in the matrix, you want a sense of what it is and for that a low rank approximation might be very, very useful for you. So something like this can be done and that's very useful. You can go back and see in the example what happens when you make a low rank approximation for a symmetric matrix of this form, okay? So that sort of summarizes the spectral theorem.

So this sort of concludes the lectures in the week. And hopefully you get a good summary of what normal and self-adjoint operators are, and what the spectral theorems tell you in terms of characterizing them. Going forward in the next week, we will start looking at, you know, positive operators and isometries. So those are again two different types of operators which sort of complement these normal and self-adjoint operators and give you a good handle on how to, you know, work with more, larger class of operators in real and complex vector spaces. Thank you very much.