PROFESSOR: Going to get started right away. I want to make a comment about energy time uncertainty relations. And we talked last time about the fact that the energy time uncertainty relation really tells you something about how fast a state can change.

So an interesting way to try to evaluate that is to consider a system that has to state that time equals 0 and compute the overlap with a state at time equals t. Now this overlap is a very schematic thing. It's a bracket. It's a number. And you know what that means. You can keep in mind what that is. It's an integral over space possibly of psi star at t equals 0 x psi t x. It's a complete integral, its inner product.

So you may want to understand this, because, in a sense, this is telling you how quickly a state can change. At time equals 0, this overlap is just 1. A little later, this overlap is going to change and perhaps after some time the overlap is going to be 0. And we're going to say that we actually have changed a lot.

So this number is very interesting to compute. And in fact, we might as well square it, because it's a complex number. So to understand better what it is we'll square it. And we'll have to evaluate this. Now how could you evaluate this? Well, we'll assume that this system that governs this time evolution has a time independent Hamiltonian. Once this evolution is done by a time independent Hamiltonian, you can wonder what it is.

Now it's quite interesting, and you will have to discuss that in the homework because it will actually help you prove that version of the time energy uncertainty relationship that says that the quickest time state can turn orthogonal to itself is bounded by some amount. Cannot do it infinitely fast. So you want to know how fast this can change.

Now it's very surprising what it depends on, this thing. Because suppose you had an energy eigenstate, suppose psi at time equals 0 is an energy eigenstate. What would happen later on? Well, you know that energy eigenstates evolve with a phase, an exponential of e to the minus iht over h bar. So actually if you had an energy eigenstate, this thing would remain equal to 1 for all times.

So if this is going to be non-zero, it's because it's going to have-- you have to have a state is not an energy eigenstate. That you're going to have an uncertainty in the energy and energy uncertainty.

So the curious thing is that you can evaluate this, and expand it as a power in t, and go, say, to quadratic ordering in t evaluating what this is. And this only depends on the uncertainty of h, and t, and things like that. So only the uncertainty of h matters at this moment.

So this would be quite interested, I think, for you to figure out and to explore in detail. That kind of analysis has been the center of attention recently, having to do with quantum computation. Because in a sense, in a quantum computer, you want to change states quickly and do operations. So how quickly you can change a state is crucial.

So in fact, the people that proved these inequalities that you're going to find say that this eventually will limit the speed of a quantum computer, and more slow that computers become twice as fast every year or so, and double the speed. So this limit apparently, it's claimed, will allow 80 duplications of the speed until you hit this limit in a quantum computer due to quantum mechanics. So you will be investigating this in next week's homework.

Today we want to begin quickly with an application of the uncertainty principal to find exact bounds of energies for quantum systems, for ground states of quantum systems. So this will be a very precise application of the uncertainty principle. Then we'll turn to a completion of these things that we've been talking about having to do with linear algebra. We'll get to the main theorems of the subject. In a sense, the most important theorems of the subject. These are the spectral theorem that tells

you about what operators can be diagonalized. And then a theorem that leads to the concept of a complete set of commuting observables. So really pretty key mathematical ideas.

And the way we're good to do it, I think you will see, that we have gained a lot by learning the linear algebra concepts in a slightly abstract way. I do remember doing this proof that we're going today in previous years, and it was always considered the most complicated lecture of the course. Just taking the [? indices ?] went crazy, and lots of formulas, and the notation was funny. And now we will do the proof, and we'll write a few little things. And we'll just try to imagine what's going on. And will be, I think, easier. I hope you will agree.

So let's begin with an example of a use of the uncertainty principle. So example. So this will be maybe for 20 minutes. Consider this Hamiltonian for a one dimensional particle that, in fact, you've considered before, alpha x to the fourth, for which you did some approximations.

You know that the expectation value of the energy in the ground state. You've done it numerically. You've done it variationally. And variationally you knew that the energy at every stage was smaller at a given bound. The uncertainty principle is not going to give us an upper bound. It's going to give us a lower bound. So it's a really nice thing, because between the variational principal and the uncertainty principle, we can narrow the energy of this ground state to a window.

In one of the problems that you're doing for this week-- and I'm sorry only today you really have all the tools after you hear this discussion-- you do the same for the harmonic oscillator. You do a variational estimate for the ground state energy. You do the uncertainty principle bound for ground state energy. And you will see these two bounds meet. And therefore after you've done this to bounds, you found the ground state energy of the harmonic oscillator, so it's kind of a neat thing.

So we want to estimate the ground state energy. So we first write some words. We just say H in the ground state will be given by the expectation value of p squared in the ground state plus alpha times the expectation value of x to the fourth in the

ground state. Haven't made much progress, but you have, because you're starting to talk about the right variables.

Now this thing to that you have to know is that you have a potential that is like this, sort of a little flatter than x squared potential. And what can we say about the expectation value of the momentum on the ground state and the expectation value of x in the ground state? Well, the expectation value of x should be no big problem. This is a symmetric potential, therefore wave functions in a one dimensional quantum mechanics problems are either symmetric or anti symmetric. It could not be anti symmetric because it's a ground state and kind of have a 0. So it's asymmetric. Has no nodes. So there's the wave function of the ground state, the symmetric, and the expectation value of x in the ground state is 0.

Similarly, the expectation value of the momentum in the ground state, what is it? Is? 0 too. And you can imagine just computing it. It would be the integral of psi. Psi is going to be a real d dx h bar over i psi. This is a total derivative. If it's a bound state, it's 0 at the ends. This is 0.

So actually, we have a little advantage here. We have some control over what p squared is, because the uncertainty in p in the ground state-- well, the uncertainty in p squared is the expectation value of p squared minus the expectation value of p squared. So in the ground state, this is 0. So delta p squared in the ground state is just p squared on the ground state. Similarly, because the expectation value of x is equal to 0, delta x squared in the ground state is equal to expectation value of x squared in the ground state.

So actually, this expectation of p squared is delta p. And we want to use the uncertainty principle, so that's progress. We've related something we want to estimate to an uncertainty. Small complication is that we have an expectation value of x to the fourth. Now we learned-- maybe I can continue here. We learned that the expectations value for an operator squared is bigger than or equal to the expectation value of the operator squared. So the expectation value of x to the fourth is definitely bigger than the expectation value of x squared squared.

And this is true on any state. This was derived when we did uncertainty. We proved that the uncertainty squared is positive, because the norm of a vector, and that gave you this thing. So here you think of the operator as x squared. So the operator squared is x to the fourth. And here's the operator expectation value squared.

So this is true for the ground state. It's also here true for any state, so is the ground state. And this x squared now is delta x. So this is delta x on the ground state to the fourth.

So look what we have. We have that the expectation value of H on the ground state is strictly equal to delta p on the ground state squared over 2m plus alpha. And we cannot do a Priorean equality here, so we have this. This is so far an equality. But because of this, this thing is bigger than that. Well, alpha is supposed to be positive. So this is bigger than delta p ground state squared over 2m plus alpha delta x on the ground state to the fourth.

OK, so far so good. We have a strict thing, this. And the order of the inequality is already showing up. We're going to get, if anything, a lower bound. You're going to be bigger than or equal to something.

So what is next? Next is the uncertainty principle. We know that delta p delta x is greater than or equal to h bar over 2 in any state. So the delta p ground state and delta x on the ground state still should be equal to that. Therefore delta p ground state is bigger than or equal than h over 2 delta x in the ground state like that.

So this inequality still the right direction. So we can replace this by something that is bigger than this quantity day without disturbing the logic. So we have H ground state now is greater than or equal to replace the delta p by this thing here, h squared over 8, because this is squared and there's another 2m delta x ground state squared plus alpha delta x ground state to the fourth. And that's it. We've obtained this inequality.

So here you say, well, this is good but how can I use it? I don't know what delta x is in the ground state, so what have I gained?

Well, let me do a way of thinking about this that can help you. Plot the right hand side as a function of delta x on the ground. So you don't know how much it is, delta x on the ground state, so just plot it. So if you plot this function, there will be a divergence as this delta x goes to 0, then it will be a minimum. It will be a positive minimum, because this is all positive. And then it will go up again.

So the right hand side as a function of delta x is this. So here it comes. You see I don't know what delta x is. Suppose delta x happens to be this. Well, then I know that the ground state energy is bigger than that value. But maybe that's not delta x. Delta x may be is this on the ground state. And then if it's that, well, the ground state energy is bigger than this value over here.

Well, since I just don't know what it is, the worst situation is if delta x is here, and therefore definitely H must be bigger than the lowest value that this can take. So the claim is that H of gs, therefore is greater than or equal than the minimum of this function h squared over 8m, and I'll just write here delta x squared plus alpha delta x to the fourth over delta x.

The minimum of this function over that space is the bound. So I just have to do a calculus problem here. This is the minimum. I should take the derivative with respect to delta x. Find delta x and substitute. Of course, I'm not going to do that here, but I'll tell you a formula that will do that for you. A over x squared plus Bx to the fourth is minimized for x squared is equal to 1 over 2-- it's pretty awful numbers. 2 to the 1/3 A over B to the 1/3. And its value at that point is 2 to the 1/3 times 3/2 times A to the 2/3 times B to the 1/3. A little bit of arithmetic.

So for this function, it turns out that A is whatever coefficient is here. B is whatever coefficient is there, so this is supposed to be the answer. And you get H on the ground state is greater than or equal to 2 to the 1/3 3/8 h squared square root of alpha over m to the 2/3, which is about 0.4724 times h squared square root of alpha over m to the 2/3. And that's our bound.

How good or how bad is the bound? It's OK. It's not fabulous. The real answer is done numerically is 0.668. I think I remember variational principal gave you

something like 0.68 or 0.69. And this one says it's bigger than 0.47. It gives you something.

So the important thing is that it's completely rigorous. Many times people use the uncertainty principle to estimate ground state energies. Those estimates are very hand wavy. You might as well just do dimensional analysis. You don't gain anything. You don't know the factors.

But this is completely rigorous. I never made an approximation or anything here. Every step was logical. Every inequality was exact. And therefore, this is a solid result. This is definitely true. It doesn't tell you an estimate of the answer. If you dimensional analysis, you say the answer is this times 1, and that's as good as you can do with dimensional analysis. It's not that bad. The answer turns out to be 0.7. But the uncertainty principle really, if you're careful, sometimes, not for every problem, you can do a rigorous thing and find the rigorous answer.

OK, so are there any questions? Your problem in the homework will be to do this for the harmonic oscillator and find the two bounds. Yes?

AUDIENCE: How does the answer change if we don't look at the ground state?

PROFESSOR: How do they what?

PROFESSOR: How does the answer change if we look at a state different from the ground state?

PROFESSOR: Different from the ground state? So the question was how would this change if I would try to do something different from the ground state. I think for any state, you would still be able to say that the expectation value of the momentum is 0. Now the expectation value of x still would be 0. So you can go through some steps here. The problem here being that I don't have a way to treat any other state differently. So I would go ahead, and I would have said for any stationary state, or for any energy eigenstate, all of what I said is true. So I don't get a new one.

These things people actually keep working and writing papers on this stuff. People sometimes find bounds that are a little original. Yes?

7

AUDIENCE: How do you know the momentum expectation is 0 again?

PROFESSOR: The momentum expectation for a bound state goes like this. So you want to figure out what is psi p psi. And you do the following. That's integral. Now psi in these problems can be chosen to be real, so I won't bother. It's psi of x h bar over i d dx of psi. So this is equal to h bar over 2i the integral dx of d dx of psi squared.

So at this moment, you say well that's just h bar over 2i, the value of psi squared at infinity and at minus infinity. And since it's a bound state, it's 0 here, 0 there, and it's equal to 0. A state that would have expectation value of momentum you would expect it to be moving. So this state is there is static. It's stationary It doesn't have expectation value of momentum. Yes?

AUDIENCE: Is the reason that you can't get a better estimate for the things that are on the ground state, because if you consider the harmonic oscillator, the uncertainty delta x [? delta p ?] from the ground state [INAUDIBLE] you go up to higher states.

PROFESSOR: Right, I think that's another way.

AUDIENCE: [INAUDIBLE] higher state using the absolute.

PROFESSOR: Yeah. That's a [INAUDIBLE]. So the ground state of the harmonic oscillator saturates the uncertainty principal and the others don't. So this argument, I think, is just good for ground state energies. One more question.

AUDIENCE: It appears that this method really works. Doesn't particularly work well if we have a potential that has an odd power, because we can't use [? the packet ?], like x [INAUDIBLE] expectation value x to the fourth is something, some power, expectation.

PROFESSOR: Right, if it's an odd power, the method doesn't work well. But actually for an odd power, the physics doesn't work well either, because the system doesn't have ground states. And so let's say that if you had x to the fourth plus some x cubed, the physics could still make sense. But then it's not clear I can do the same. Actually

you can do the same for x to the eighth and any sort of powers of this type. But I don't think it works for x to the sixth. You can try a few things.

OK, so we leave the uncertainty principle and begin to understand more formerly the operators for which there's no uncertainty and you can simultaneously diagonalize them. So we're going to find operators like A and B, that they commute. And then sometimes you can simultaneously diagonalize them. Yes? You have a question.

AUDIENCE: So part of [INAUDIBLE] we use here is [INAUDIBLE], right?

PROFESSOR: Right.

AUDIENCE: If we get an asset-- is there any way that we can better our [INAUDIBLE] principle based on the wave function with non saturated? Can we get an upper bound for just [INAUDIBLE] principle with an h bar over it?

PROFESSOR: Can I get an upper bound? I'm not sure I understand your question.

AUDIENCE: [INAUDIBLE] the fact that the [INAUDIBLE] principle will not be saturated. Can you put the bound for just taking [INAUDIBLE]?

PROFESSOR: Yeah, certainly. You might have some systems in which you know that this uncertainty might be bigger than the one warranted by the uncertainty principles. And you use that information. But on general grounds, it's hard to know that a system might come very close to satisfy the uncertainty principle in its ground state. We don't know. There are systems that come very close in the ground state to satisfy this and some that are far. If they are far, you must have some reason to understand that to use it. So I don't know.

So let me turn now to this issue of operators and diagonalization of operators. Now you might be irritated a little even by the title. Diagonalization of operation. You'll be talking about diagonalization of matrices. Well, there's a way to state what we mean by diagonalizing an operator in such a way that we can talk later about the matrix. So what is the point here? You have an operator, and it's presumably an important

operator in your theory. You want to understand this operator better.

So you really are faced with a dilemma. How do I get some insight into this operator? Perhaps the simplest thing you could do is to say, OK let me choose some ideal basis of the vector space, such as that operator is as simple as possible in that basis. So that's the origin of this thing. Find a basis in the state space, so the operator looks as simple as possible.

So you say that you can diagonalize an operator if you can find the basis such that the operator has just diagonal entries. So let me just write it like this. So if you can find a basis in V where the matrix representing the operator is diagonal, the operator is said to be diagonalizable.

So to be diagonalizable is just a statement that there is some basis where you look at the matrix representation operator, and you find that it takes form as a diagonal. So let's try to understand this conceptually and see what actually it's telling us. It tells us actually a lot.

Suppose t is diagonal in some basis u1 up to un. So what does it mean for it to be diagonal? Well, you may remember all these definitions we had about matrix action. If T acting on a ui is supposed to be Tki uk in some basis sum over k. You act on ui, and you get a lot of u's. And these are the matrix elements of the operator.

Now the fact that it's diagonalizable means that in some basis, the u basis, this is diagonal. So ki in this sum only happens to work out when k is equal to i. And that's one number and you get back to the vector ui. So if it's diagonal in this basis, you have the T on u1 is lambda a number times u1 T on u2 is lambda 2 in u2. And Tun equal lambda n un.

So what you learn is that this basis vector-- so you learn something that maybe you thought it's tautological. It's not tautological. You learn that if you have a set of basis vectors in which the operator is diagonal, these basis vectors are eigenvectors of the operator. And then you learn something that is quite important, that an operator is diagonalizable if, and only if, it has a set of eigenvectors that span the space.

So the statement is very important. An operator T is that diagonalizable if it has a set of eigenvectors that span the space. Span V. If and only if. If this double f. If and only if. So here it's diagonalizable, and we have a basis, and it has a set of these are eigenvectors. So diagonalizable realizable really means that it has a set of eigenvectors that span the space.

On the other hand, if you have the set of eigenvectors that span the space, you have a set of u's that satisfy this, and then you read that, oh yeah, this matrix is diagonal, so it's diagonalizable. So a simple statement, but an important one, because there are examples of matrices that immediately you know you're never going to succeed to diagonalize.

So here is one matrix, 0 0 1 0. This matrix has eigenvalues, so you do the characteristic equation lambda squared equals 0. So the only eigenvalue is lambda equals 0. And let's see how many eigenvectors you would have for lambda equals 0.

Well, you would have if this is T, T on some vector a b must be equal to 0. So this is 0 1 0 0 on a b, which is b and 0, must be zero. So b is equal to 0. So the only eigenvector here-- I'll just write it here and then move to the other side. The only eigenvector for lambda equals 0, the only eigenvector is with b equals 0. So it's 1 0. One eigenvector only. No more eigenvectors. By the theorem, or by this claim, you know it's a two dimensional vector space you just can't diagonalize this matrix. It's impossible. Can't be done.

OK, a couple more things that I wish to say about this process of diagonalization. Well, the statement that an operator is diagonal is a statement about the existence of some basis. Now you can try to figure out what that basis is, so typically what is the problem that you face? Typically you have a vector spaces V. Sorry?

AUDIENCE: I have a question.

PROFESSOR: Yes? If you had an infinite dimensional space and you had an operator whose eigenvectors do not span the space, can it still have eigenvectors, or does it not

have any then?

PROFESSOR: No. You said it has some eigenvectors, but they don't span the space. So it does have some eigenvectors.

AUDIENCE: So my question is was what I just said a logical contradiction in an infinite dimensional space?

PROFESSOR: To have just some eigenvectors? I think--

PROFESSOR: I'm looking more specifically at a dagger for instance.

PROFESSOR: Yes.

AUDIENCE: In the harmonic oscillator, you I think mentioned at some point that it does not have--

PROFESSOR: So the fact that you can' diagonalize this thing already implies that it's even worse in higher dimensions. So some operator may be pretty nice, and you might still be able to diagonalize it, so you're going to lack eigenvectors in general. You're going to lack lots of them. And there are going to be blocks of Jordan. Blocks are called things that are above the diagonal, things that you can't do much about.

Let me then think concretely now that you have a vector space, and you've chosen some basis v1 vn. And then you look at this operator T, and of course, you chose an arbitrary basis. There's no reason why its matrix representation would be diagonal. So T on the basis v-- Tij. Sometimes to be very explicit we write Tij like that-- is not diagonal.

Now if it's not diagonal, the question is whether you can find a basis where it is diagonal. And then you try, of course, changing basis. And you change basis-- you've discussed that in the homework-- with a linear operator. So you use a linear operator to produce another basis, an invertible in your operator.

So that you get these vectors uk being equal to some operator A times vk. So this is going to be the u1's up to un's are going to be another basis. The n vector here is

the operator acting with the n vector on this thing. And then you prove, in the homework, a relationship between these matrix elements of T in the new basis, in the u basis. And the matrix elements of T in the v basis.

You have a relationship like this, or you have more explicitly T_{ij} in the basis u is equal to $A^{-1}_{ik} T_{kp}$ of v A_{pj}. So this is what happens. This is the new operator in this basis. And typically what you're trying to do is find this matrix A that makes this thing into a diagonal matrix. Because we say in the u basis the operator is diagonal.

I want to emphasize that there's a couple of ways in which you can think of diagonalization. Sort of a passive and an active way. You can imagine the operator, and you say look, this operator I just need to find some basis in which it is diagonal. So I'm looking for a basis. The other way of thinking of this operator is to think that A minus 1 TA is another operator, and it's diagonal in original basis.

So it might have seem funny to you, but let's stop again and say this again. You have an operator, and the question of diagonalization is whether there is some basis in which it looks diagonal, its matrix is diagonal. But the equivalent question is whether there is an operator A such that this is diagonal in the original basis.

To make sure that you see that, consider the following. So this is diagonal in the original basis. So in order to see that, think of $T u_i$ is equal to $\lambda_i u_i$. We know that the u's are supposed to be this basis of eigenvectors where the matrix is diagonal, so here you got it. Here the i not summed. It's pretty important.

There's a problem with this eigenvalue notation. I don't know how to do it better. If you have several eigenvalues, you want to write this, but you don't want this to think that you're acting on u1 and you get lambda 1 u1. Not the sum right here.

OK, but they u_i is equal to A on v_i. So therefore this is $\lambda_i A$ on v_i. And then you'll act with A minus 1. Act with A minus 1 from the left with the operator. So you get $A^{-1} T A v_i$ is equal to $\lambda_i v_i$.

So what do you see? You see an operator that is actually diagonal in the v basis. So

this operator is diagonal in the original basis. That's another way of thinking of the process of diagonalization.

There's one last remark, which is that the columns of A are the eigenvectors, in fact. Columns of A are the eigenvectors. Well, how do you see that? It's really very simple. You can convince yourself in many ways, but the uk are the eigenvectors. But what are uk's? I have it somewhere. There it is. A on vk. And A on vk is this matrix representation is sum over i Aik vi.

So now if this is the original basis, the vi's are your original basis, then you have the following, that the vi's can be thought as the basis vectors and represented by columns with a 1 in the ith entry. So this equation is saying nothing more, or nothing less than uk, in terms of matrices or columns, is equal to A1k v1 plus Ank vn, which is just A1k A2k Ank. Because vi is the ith basis vector. So 1 0's only in the ith position.

So these are the eigenvectors. And they're thought as linear combinations of the vi's. The vi's are the original basis vectors. So the eigenvectors are these numbers.

OK, we've talked about diagonlization, but then there's a term that is more crucial for our operators that we're interested in. We're talking about Hermitian operators. So the term that is going to be important for us is unitarily diagonalizable.

What is a unitarily diagonalizable operator? Two ways again of thinking about this. And perhaps the first way is the best. And I will say it. A matrix is set to be unitarily diagonalizable if you have an orthonormal basis of eigenvectors. Remember diagonalizable meant a basis of eigenvectors. That's all it means. Unitarily diagonalizable means orthonormal basis of eigenvectors.

So T has an orthonormal basis of eigenvectors. Now that's a very clear statement. And it's a fabulous thing if you can achieve, because you basically have broken down the space into basis spaces, each one of them with a simple thing before your operators. And they're orthonormal, so it's the simplest possible calculational tool. So it's ideal if you can have this.

Now the way we think of this is that you start with-- concretely, you start with a T of some basis v that is an orthonormal basis. Start with an orthonormal basis, and then pass to another orthonormal basis u. So you're going to pass to another orthonormal basis u with some operator. But what you have learned is that if you want to pass from v orthonormal to another basis u, a vector that is also orthonormal, the way to go from one to the other is through a unitary operator. Only unitary operators pass you from orthonormal to orthonormal basis.

Therefore really, when you start with your matrix in an orthonormal basis that is not diagonal, the only thing you can hope is that T of u will be equal to sum u dagger, or u minus 1, T of v u. Remember, for a unitary operator, where u is unitary, the inverse is the dagger. So you're doing a unitary transformation, and you find the matrix that is presumably then diagonal.

So basically, unitarily diagonalizable is the statement that if you start with the operator in an arbitrary orthonormal basis, then there's some unitary operator that takes you to the privilege basis in which your operator is diagonal, is still orthonormal. But maybe in a more simple way, unitarily diagonalizable is just a statement that you can find an orthonormal basis of eigenvectors.

Now the main theorem of this subject, perhaps one of the most important theorems of linear algebra, is the characterization of which operators have such a wonderful representation. What is the most general operator T that will have an orthonormal basis of eigenvectors?

Now we probably have heard that Hermitian operators do the job. Hermitian operators have that. But that's not the most general ones. And given that you want the complete result, let's give you the complete result.

The operators that have this wonderful properties are called normal operators, and they satisfy the following property. M is normal if M dagger, the adjoint of it, commutes with M. So Hermitian operators are normal, because M dagger is equal to M, and they commute. Anti Hermitian mission operators are also normal, because anti Hermitian means that dagger is equal to minus M, and it still

commutes with M.

Unitary operators have U dagger U equal to U U dagger equals to 1. So U and U dagger actually commute as well. So Hermitian, anti Hermitian, unitary, they're all normal operators.

What do we know about normal operators? There's one important result about normal operators, a lemma. If M is normal and W is an eigenvector, such that MW is equal to lambda W. Now normal operators need not have real eigenvalues, because they include unitary operators. So here I should write Hermitian, anti Hermitian, and unitary are normal.

So here is what a normal operator is doing. You have a normal operator. It has an eigenvector with some eigenvalue. Lambda is a complex number in principle. Then the following result is true, then M dagger omega is also an eigenvector of M dagger. And it has eigenvalue lambda star.

Now this is not all that easy to show. It's a few lines, and it's done in the notes. I ask you to see. It's actually very elegant. What is the usual strategy to prove things like that Is to say oh, I want to show this is equal to that, so I want to show that this binds that is 0. So I have a vector that is zero what is the easiest way to show that it's 0? If I can show it's norm is 0.

So that's a typical strategy that you use to prove equalities. You say, oh, it's a vector that must be 0 as my equality. Let's see if it's 0. Let's find its norm, and you get it. So that's a result.

So with this stated, we finally have the main result that we wanted to get to. And I will be very sketchy on this. The notes are complete, but I will be sketchy here. It's called the spectral theorem. Let M be an operator in a complex vector space. The vector space has an orthonormal basis of eigenvectors of M if and only if M is normal. So the normal operators are it. You want to have a complete set of orthonormal eigenvectors. Well, this will only happen if your operator is normal, end of story.

Now there's two things about this theorem is to show that if it's diagonalizable, it is normal, and the other thing is to show, that if it's normal, it can be diagonalized. Of course, you can imagine the second one is harder than the first. Let me do the first one for a few minutes. And then say a couple of words about the second. And you may discuss this in recitation. It's a little mathematical, but it's all within the kind of things that we do. And really it's fairly physical in a sense. We're accustomed to do such kinds of arguments.

So suppose it's unitarily diagonalizable, which means that M-- so if you have U dagger, MU is equal to a diagonal matrix, DM. I'm talking now matrices. So these are all matrices, a diagonal matrix. There's no basis to the notion of a diagonal operator, because if you have a diagonal operator, it may not look diagonal in another basis. Only the identity operator is diagonal in all basis, but not the typical diagonal operator.

So unitarily diagonalizable, as we said, you make it-- it's gone somewhere. Here. You act with an inverse matrices, and you get the diagonal matrix. So from this, you find that M is equal to U DM U dagger by acting with U on the left and with U dagger from the right, you solve for M, and it's this.

And then M dagger is the dagger of these things. So it's U to DM dagger U dagger. The U's sort of remain the same way, but the diagonal matrix is not necessarily real, so you must put the dagger in there.

And now M dagger M. To check that the matrix is normal that commutator should be 0. So M dagger M. You do this times that. You get U DM dagger. U dagger U. That's one. DM U dagger. And M M dagger you multiply the other direction you get U DM DM dagger U dagger. So the commutator of M dagger M is equal to U DM dagger DM minus DM Dm dagger U dagger.

But any two diagonal matrices commute. They may not be that simple. Diagonal matrices are not the identity matrices, but for sure they commute. You multiply elements along with diagonal so this is 0. So certainly any unitarily diagonalizable

matrix is normal.

Now the other part of the proof, which I'm not going to speak about, it's actually quite simple. And it's based on the fact that any matrix in a complex vector space has at least one eigenvalue. So what you do is you pick out that eigenvalue and it's eigenvector, and change the basis to use that eigenvector instead of your other vectors. And then you look at the matrix. And after you use that eigenvector, the matrix has a lot of 0's here and a lot of 0's here. And then the matrix has been reduced in dimension mansion, and then you go step by step.

So basically, it's the fact that any operator has at least one eigenvalue and at least one eigenvector. It allows you to go down. And normality is analogous to Hermiticity in some sense. And the statement that you have an eigenvector generally tells you that this thing is full of 0's, but then you don't know that there are 0's here. And either normality or Hermiticity shows that there are 0's here, and then you can proceed at lower dimensions. So you should look at the proof because it will make clear to you that you understand what's going on.

OK but let's take it for granted now you have these operators and can be diagonalized. Then we have the next thing, which is simultaneous diagonalization. What is simultaneous diagonalization? It's an awfully important thing.

So we will now focus on simultaneous diagonalization of Hermitian operators. So simultaneous diagonalization of Hermitian ops. Now as we will emphasize towards the end, this is perhaps one of the most important ideas in quantum mechanics. It's this stuff that allows you to label and understand your state system.

Basically you need to diagonalize more than one operator most of the time. You can say OK, you found the energy eigenstates. You're done. But if you find your energy eigenstates and you think you're done, maybe you are if you have all these energy eigenstates tabulated. But if you have a degeneracy, you have a lot of states that have the same energy.

And what's different about them? They're certainly different because you've got

several states, but what's different about them? You may not know, unless you figure out that they have different physical properties. If they're different, something must be different about them. So you need more than one operator, and your facing the problem of simultaneously diagonalizing things, because states cannot be characterized just by one property, one observable. Would be simple if you could, but life is not that simple.

So you need more than one observable, and then you ask when can they be simultaneously diagonalizable. Well, the statement is clear. If you have two operators, S and T that belong to the linear operators in a vector space, they can be simultaneously diagonalized if there is a basis for which every basis vector is eigenstate of this and an eigenstate of that. Common set of eigenstates.

So they can be simultaneously diagonalized. Diagonalizable is that there is a basis where this basis is comprised of the eigenvectors of the operator. So this time you require more, that that basis be at the same time a basis set of eigenvectors of this and a set of eigenvectors of the second one.

So a necessary condition for simultaneous diagonalization is that they commute. Why is that? The fact that two operators commute or they don't commute is an issue that is basis independent. If they don't commute, the order gives something different, and that you can see in every basis.

So if they don't commute and they're simultaneously diagonalizable, there would be a basis in which both are diagonal and they still wouldn't commute. But you know that diagonal matrices commute. So if two operators don't commute, they must not commute in any base, therefore there can't be a basis in which both are at the same time diagonal.

So you need, for simultaneous diagonalizable, you need that S and P commute. Now that may not be enough, because not all operators can be diagonalized. So the fact that they commute is necessary, but not everything can be diagonalizable. Well, you've learned that every normal operator, every Hermitian operator is diagonalizable.

And then you got now a claim of something that could possibly be true. Is the fact that whenever you have two Hermitian operators, each one can be diagonalized by themselves. And they commute. There is a simultaneous set of eigenvectors of 1 that are eigenvectors of the first and eigenvectors of the second.

So the statement is that if S and T are commuting Hermitian operators, they can be simultaneously diagonalized. So this theorem would be quite easy to show if there would be no degeneracies, and that's what we're going to do first. But then we'll consider the case of degeneracies.

So I'm going to consider the following possibilities. Perhaps neither one has a degenerate spectrum. What does it mean a degenerate spectrum? Same eigenvalue repeated many times. But that is a wishful thinking situation. So either both are non degenerate, either one is non degenerate and the other is degenerate, or both are degenerate. And that causes a very interesting complication.

So let's say there's going to be two cases. It will suffice. In fact, it seems that there are three, but two is enough to consider. There is no degeneracy in T. So suppose one operator has no degeneracy, and let's call it T. So that's one possibility. And then S may be degenerate, or it may not be degenerate. And the second possibility is that both S and T are degenerate.

So I'll take care of case one first. And then we'll discuss case two, and that will complete our discussion. So suppose there's no the degeneracy in the spectrum of T. So case one.

So what does that mean? It means that T is non degenerate. There's a basis U1 can be diagonalized to UM, orthonormal by the spectral theorem. And there's eigenvectors T U-- these are eigenvectors. Lambda I Ui. And lambda I is different to lambda j for i different from j.

So all the eigenvalues, again, it's not summed here. All the eigenvalues are different. So what do we have? Well, each of those eigenvectors, each of the Ui's

that are eigenvectors, generate invariant subspaces. There are T invariant subspaces.

So each one, each vector U1 you can imagine multiplying by all possible numbers, positive and negative. And that's an invariant one dimensional subspace, because if you act with T, it's a T invariant space, you get the number times a vector there.

So the question that you must ask now is you want to know if these are simultaneous eigenvectors. So you want to figure out what about S. How does S work with this thing? So you can act with S from the left. So you get STUi is equal to lambda i SUi.

So here each Ui generates an invariant subspace Ui T invariant. But S and T commute, so you have T SUi is equal to lambda i SUi. And look at that equation again. This says that this vector belongs to the invariant subspace Ui, because it satisfies exactly the property that T acting on it is equal to lambda i Ui. And it couldn't belong to any other of the subspaces, because all the eigenvalues are different.

So spaces that are in Ui are the spaces-- vectors that are in Ui are precisely all those vectors that are left invariant by the action of T. They're scaled only. So this vector is also in Ui. If this vector is in Ui, SUi must be some number Wi times Ui. And therefore you've shown that Ui is also an eigenvector of S, possibly with a different eigenvalue of course. Because the only thing that you know is that SUi is in this space. You don't know how big it is.

So then you've shown that, indeed, these Ui's that were eigenstates of T are also eigenstates of S. And therefore that's the statement of simultaneously diagonalizable. They have the common set of eigenvectors. So that's this part. And it's relatively straight forward.

Now we have to do case two. Case two is the interesting one. This time you're going to have degeneracy. We have to have a notation that is good for degeneracy. So if S is degeneracies, has degeneracies, what happens with this operator? It will have--

remember, a degenerate operator has eigenstates that form higher than one dimensional spaces.

If you have different eigenvalues, each one generates a one dimensional operator invariant subspace. But if you have degeneracies, there are operators-- there are spaces of higher dimensions that are left invariant.

So for example, let Uk denote the S invariant subspace of some dimension Dk, which is greater or equal than 1. I will go here first. We're going to define Uk to be the set of all vectors so that SU is equal to lambda k U. And this will have dimension of Uk is going to be Dk.

So look what's happening. Basically the fact is that for some eigenvalues, say the kth eigenvalue, you just get several eigenvectors. So if you get several eigenvectors not just scaled off each other, these eigenvectors correspond to that eigenvalue span of space. It's a degenerate subspace.

So you must imagine that as having a subspace of some dimensionality with some basis vectors that span this thing. And they all have the same eigenvector. Now you should really have visualized this in a simple way. You have this subspace like a cone or something like that, in which every vector is an eigenvector. So every vector, when it's acted by S is just scaled up. And all of them are scaled by the same amount. That is what this statement says.

And corresponding to this thing, you have a basis of vectors, of these eigenvectors, and we'll call them Uk1, the first one, the second, up to U Dk1, because it's a subspace that we say it has the dimensionality Dk. So look at this thing. Somebody tells you there's an operator. It has degenerate spectrum. You should start imagining all kind of invariant subspaces of some dimensionality. If it has degeneracy, it's a degeneracy each time the Dk is greater than 1, because if it's one dimensional, it's just one basis vector one eigenvector, end of the story.

Now this thing, by the spectral theorem, this is an orthonormal basis. There's no problem, when you have a degenerate subspace, to find an orthonormal basis. The

theorem guarantees it, so these are all orthonormal.

So at the end of the day, you have a decomposition of the vector space, V as U1 plus U2 plus maybe up to UM. And all of these vector spaces like U's here, they may have some with just no degeneracy, and some with degeneracy 2, degeneracy 3, degeneracy 4. I don't know how much degeneracy, but they might have different degeneracy.

Now what do we say next? Well, the fact that S is a Hermitian operator says it just can be diagonalized, and we're can find all these spaces, and the basis for the whole thing. So the basis would look U1 of the first base up to U d1 of the first base. These are the basis vectors of the first plus the basis vectors of the second. All the basis vectors U1 up to Udm of the mth space.

All this is the list. This is the basis of V. So I've listed the basis of V, which a basis for U1, all these vectors. U2, all of this. So you see, we're not calculating anything. We're just trying to understand the picture.

And why is this operator, S, diagonal in this basis? It's clear. Because every vector here, every vector is an eigenvector of S. So when you act with S on any vector, you get that vector times a number. But that vector is orthogonal to all the rest. So when you have some U and S and another U, this gives you a vector proportional to U. And this is another vector. The matrix element is 0, because they're all orthogonal.

So it should be obvious why this list produces something that is completely orthogonal-- a diagonal matrix. So S, in this basis, looks like the diagonal matrix in which you have lambda 1 d1 times up to lambda m dm times. Now I'll have to go until 2:00 to get the punchline. I apologize, but we can't stop right now. We're almost there, believe it or not.

Two more things. This basis is good, but actually another basis would also be good. I'll write this other basis would be a V1 acting on the U1 up to V1 acting on that U1 up to a Vm acting on this U1 up to Vm acting on that U1. This is m. m. This is dm. And here it's not U1. It's Ud1.

You see, in the first collection of vectors, I act with an operator V1 up to here with an operator Vm. All of them with Vk being a unitary operator in Uk. In every subspace, there are unitary operators. So you can have these bases and act with a unitary operator of the space U1 here. A unitary operator with a space U2 here. A unitary operator of the space Un here.

Hope you're following. And what happens if this operator is unitary, this is still an orthonormal basis in U1. These are still orthonormal basis in Um. And therefore this is an orthonormal basis for the whole thing, because anyway those different spaces are orthogonal to each other. It's an orthogonal decomposition. Everything is orthogonal to everything. So this basis would be equally good to represent the operator. Yes?

AUDIENCE: [INAUDIBLE] arbitrary unitary operators?

PROFESSOR: Arbitrary unitary operators at this moment. Arbitrary. So here comes the catch as to the main property that now you want to establish is that the spaces Uk are also T invariant. You see, the spaces Uk were defined to be S invariant subspaces. And now the main important thing is that they are also T invariant because they commute with that.

So let's see why that is the case. Suppose U belongs to Uk. And then let's look at the vector-- examine the vector Tu. What happens to Tu? Well, you want to act on S on Tu to understand it. But S and T commute, so this is T SU. But since U belongs to Uk, that's the space with eigenvalue lambda k. So this is lambda k times u, so you have Tu here. So Tu acted with S gives you lambda k Tu. So Tu is in the invariant subspace Uk.

What's happening here is now something very straightforward. You try to imagine how does the matrix T look in the basis that we have here. Here is this basis. how does this matrix T look? Well, this matrix keeps the invariant subspaces. So you have to think of it blocked diagonally. If it acts on it-- here are the first vectors that you're considering, the U1. Well if you act on it with T of the U1 subspace, you stay in the U1 subspace. So you don't get anything else. So you must have 0's all over

here. And you can have a matrix here.

And if you act on the second Uk U2, you get a vector in U2, so it's orthogonal to all the other vectors. So you get a matrix here. And you get a matrix here. So actually you get a blocked diagonal matrix in which the blocks correspond to the degeneracy. So if there's a degeneracy d1 here, it's a d1 times d1. And d2 times d2.

So actually you haven't simultaneously diagonalized them. That's the problem of degeneracy. You haven't, but you now have the tools, because this operator is Hermitian, therefore it's Hermitian here, and here, and here, and here. So you can diagonalize here.

But what do you need for diagonalizing here? You need a unitary matrix. Call it V1. For here you need another unitary matrix. Call it V2. Vn. And then this matrix becomes diagonal. But then what about the old matrix? Well, we just explained here that if you change the basis by unitary matrices, you don't change the first matrix. So actually you succeeded. You now can diagonalize this without destroying your earlier result. And you managed to diagonalize the whole thing. So this is for two operators in the notes. You'll see why it simply extends for three, four, and five, or arbitrary number of operators. See you next time.