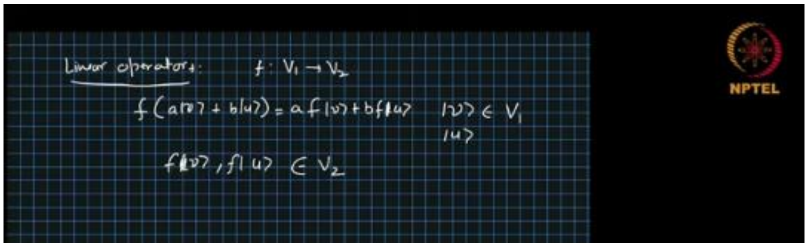
FOUNDATIONS OF QUANTUM THEORY: NON-RELATIVISTIC APPROACH

Dr. Sandeep K. Goyal
Department of Physical Sciences
IISER Mohali
Week-02
Lecture-04
Linear Operators - Part 01

Linear operators are mappings. Let us say we have a map f, which takes the vectors of a vector space V_1 to another vector space V_2. Now this map f or this function f which maps the vector of V_1 to V_2, this will be called a linear operator or it can be represented by a linear operator if it satisfies the following condition, $f(a|v\rangle+b|u\rangle)$, where $|v\rangle$ and $|u\rangle$ they both belong to the vector space V_1 and a and b are the scalars on which the vector space V_1 is defined, then $f(a|v\rangle+b|u\rangle)$ should be equal to $af|v\rangle+bf|u\rangle$. And $f|v\rangle$ and $f|u\rangle$, they belong to V_2. This is what the transformation is.

The f maps the vector of V_1 to V_2. But if you have a combination of vectors in V_1 on which f is acting, then it should satisfy in this fashion. Then we can call it a linear operator. Why we are interested in linear operator? Because linear operators are very important tool for quantum mechanics, which we will learn in this entire course.

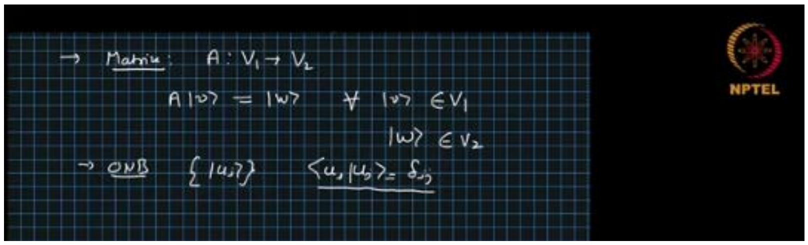
(Refer slide time: 1:38)



So, one interesting fact about linear operator is any linear operator can be represented in a matrix form. So, if we are given a linear operator A, which is again acting on vector space V_1 and results in vector space V_2, then we can write the matrix form of A if we know the action of A on the vectors of V_1 and yielding V_2. So for that, it will be like A acting on vector $|v\rangle$ gives us $|w\rangle$, let us say, where $|v\rangle$ is from V_1 and $|w\rangle$ is from V_2. So this is given to us. Now this is for every vector $|v\rangle$ given in V_1. We know the action of A, the operator A on $|v\rangle$ and we know what $|w\rangle$ we get.

Now let us say that we choose an orthonormal basis. Let me remind you what was orthonormal basis. It's the set of vectors $|u_i\rangle$ such that the inner product of $|u_i\rangle$ with $|u_j\rangle$ gives us δ_{ij}. So let us compress this statement. If i is equal to j, then this is the norm of the vector $|u\rangle$, and then it should be equal to 1. If i is not equal to j, then they should be orthogonal, so the result should be 0.

(Refer slide time: 3:19)

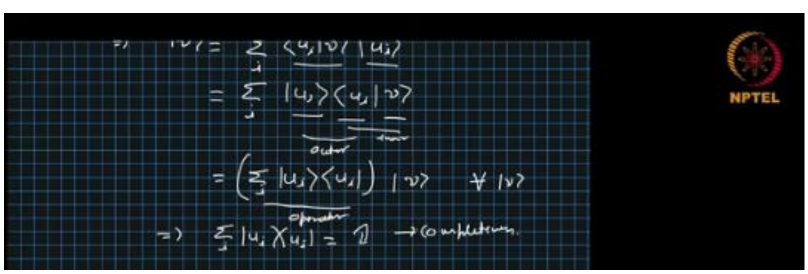


And this is the maximum set, maximum possible set in a given vector space. And we are taking ONB in V_1 vector space. Now, if we apply A, if we know the action of A on all the vectors of V_1, then we know the action of A on $|u_i\rangle$ also, and let us say we get $|w_j\rangle$ output.

So, if you remember what are the properties of orthonormal bases, they are complete and they are, so it means we can write any vector as a linear superposition of u_i's. In other words, any vector $|v\rangle$ can be written as $\sum_i \alpha_i |u_i\rangle$, for all $|v\rangle$. And we, of course, find different coefficients α_i's.

α_i's are given as $\langle u_i|v\rangle$. So, this implies that we can write $|v\rangle$ as $\sum_i \langle u_i|v\rangle |u_i\rangle$. We can reshuffle since this is a scalar and this is a vector. They commute in a manner of speaking. So $\sum_i |u_i\rangle\langle u_i|v\rangle$. Now we have product of three vectors.

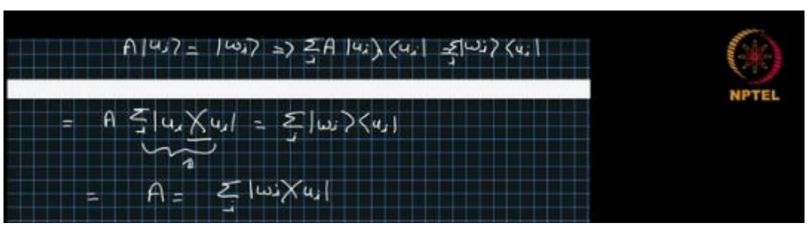
(Refer slide time: 4:50)



Okay in some or the other way this is the scalar product this is the operator product it's outer product it's called and this is called the inner product. We can redistribute it, $\sum_i(|u_i\rangle\langle u_i|)|v\rangle$, since the summation is over i, we can take all this and this is $|v\rangle$, now you see this is an operator and we have decomposed, we can, we have rewritten this equation as an operator acting on vector $|v\rangle$, the original vector $|v\rangle$, so this is true for all the vectors $|v\rangle$, so this implies that $\sum_i|u_i\rangle\langle u_i|$ must be equal to identity, okay, this is called completeness. So, if this is the case, then you can write every vector $|v\rangle$, as a sum of the orthonormal basis $|u_i\rangle$, then this outer product of $|u_i\rangle$, and sum over i should yield an identity operator. Identity operator is something which acts on a vector, gives you the same vector. Now, why we have gone so far into this mathematics ?

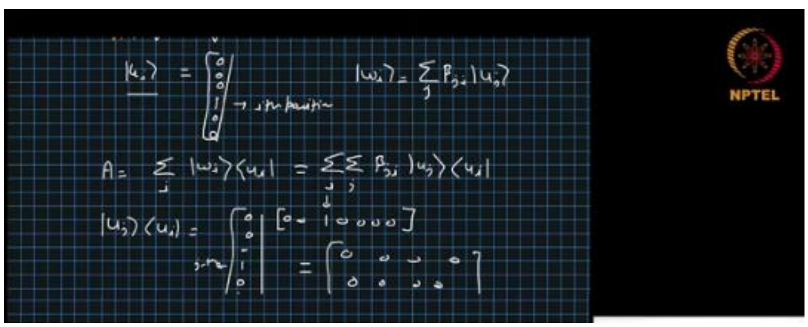
Because our original task was A acting on $|u_i\rangle$ is giving us $|w_i\rangle$. It means if we take the product of, if we multiply the whole equation from the right side by this quantity, which is the $\langle u_i|$, so that we can, $\sum_i A|u_i\rangle\langle u_i| = \sum_i|w_i\rangle\langle u_i|$. This is what we get. Now we can take sum over i, and we can write it as A acting on $\sum_i|u_i\rangle\langle u_i|$ equals $\sum_i|w_i\rangle\langle u_i|$, this I'm just writing short form as this I'm writing as this big cross don't get confused here, but now we use the completeness relation that $|u_i\rangle$ was orthonormal basis so this must be identity. So A times identity is A and we are getting $\sum_i|w_i\rangle\langle u_i|$.

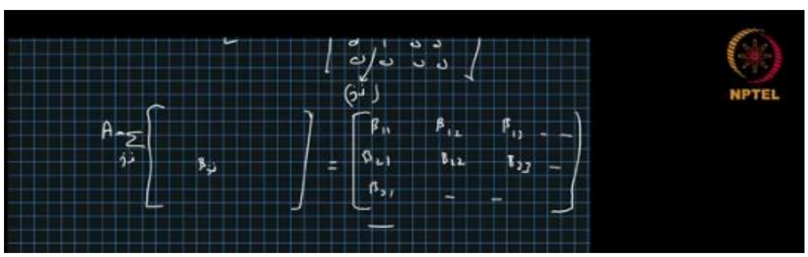
(Refer slide time: 7:35)



So, this is one expression for the operator A which is acting on the vectors of V_1 and yielding the vectors of V_2. Now you can just simply see A acting on $|v\rangle$ will be $\sum_i|w_i\rangle\langle u_i|v\rangle$ and this is a scalar, let us call it α_i, it becomes $\sum_i \alpha_i|w_i\rangle$. So, this is all good but where are the matrices here, okay, so to find the matrices what we do is let us take a simple case let us say A is acting on $|v\rangle$ and giving the vectors of V. It means V_1 and V_2 both are the same vector space. It will work for anything but let us take the simpler case. Then let us say the ONB, orthonormal basis $|u_i\rangle$ is written in the following way. The i-th vector is a column matrix, one dimensional column matrix. All of the elements are zero except one on the ith position, okay, so this is our $|u\rangle$ okay, we can have, we can choose more complicated $|u\rangle$ also but if simple works why not work with a simple one then and so since V is going to V then our $|w_i\rangle$ can also be written as, let us say, $\sum_j \beta_{ji}|u_j\rangle$. And every vector can be decomposed as a linear superposition of the orthonormal basis $|u_i\rangle$. So, every $|w_i\rangle$ also can be written as a superposition of u i's or u j's. So, now since A is $\sum_i|w_i\rangle\langle u_i|$, this can be written as $\sum_i\sum_j|u_j\rangle\langle u_i|$. Now, what is $|u_j\rangle\langle u_i|$? It is vector which is (0, 0, 0, 1, 0, 0..) somewhere. This is the jth position, and it's the Hermitian conjugate here that is (0, 0, 0, 1, 0, 0..), this is ith position. So, this will give us a matrix which is given matrix and this one is at ji-th position. So, this $|u_j\rangle\langle u_i|$ is a matrix with adding just one 1 and all other zeros and that one is at the ji location. So, if we put all everything together, then we get A to be β_{ji} and everything else is zero and sum over j and i location wise. So, this is a strange kind of sum. It's not a normal sum because it's operators also are changing location. So, it will become β_{11}, β_{12}, β_{13} and so on. β_{21}, β_{31}, β_{22} . β_{23} and so on.

(Refer slide time: 8:37)





So, this is the matrix representation we were aspiring for. So now let us recapitulate what we did. We are given a linear operator A, which acts on a vector space V_1 and results in V_2. Now we want to find a matrix representation of A. To find the matrix representation, first we realize that if we apply the matrix A or operator A on an orthonormal basis $|u_i\rangle$, we get a corresponding vector $|w_i\rangle$. And by using the completeness relation of the orthonormal basis, we can find that A is indeed $\sum_i|w_i\rangle\langle u_i|$.

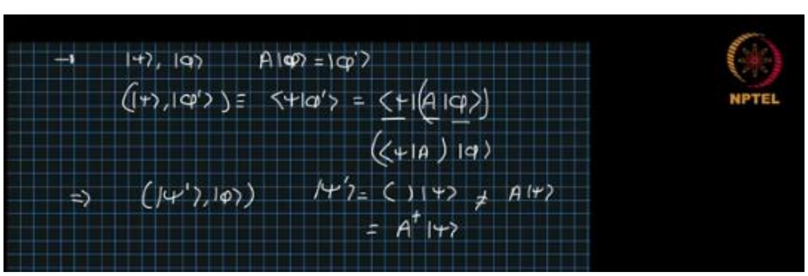
So, this is still abstract because we do not have any representation of the vectors $|u_i\rangle$ or $|w_i\rangle$. So, to get a proper matrix representation, first we represent the orthonormal basis $|u_i\rangle$ as a column vector. So, bra becomes the row vector. We decompose $|w_i\rangle$ in similar orthonormal basis. For simplicity, we assume V_1 equals V_2.

But if it is not equal, then $|w_i\rangle$ can be decomposed in some other basis of the vector space V_2. So, once we decompose $|w_i\rangle$ in the same orthonormal basis $|u_j\rangle$ with coefficient β_{ji}, then we have, we substitute this expression and this representation matrix representation of the vectors and we get a corresponding representation for the operator. So, this confirms our claim that every linear operator can be represented as a matrix. So, this type will be very useful. And since we are working, this course is on foundation of quantum mechanics. We will be dealing with only matrices most of the time and very few times we will be using abstract operators. So, some interesting things about operators, linear operators.

If you have two operators A and B such that, we say A equals B, what is the definition of a equality so when we say A is equal to B, then it means A acting on $|v\rangle$ should be equal to B acting on the same vector $|v\rangle$ for all the vectors $|v\rangle$ in the vector space V, okay. So, this is the only definition to say A = B that it should be the action of A and B on the same vector should be same for all the possible vectors. So, this is what we call the equality of two operators. If you have a vector $|\psi\rangle$ and another vector $|\phi\rangle$, then outer product $|\psi\rangle\langle\phi|$, which I write $|\psi\rangle\langle\phi|$, this is also a linear operator.

Now, if we have a vector $|\psi\rangle$ and a $|\phi\rangle$, and we apply a linear operator A on $|\phi\rangle$, which will give us $|\phi'\rangle$, then the inner product of $|\psi\rangle$ and $|\phi'\rangle$, $(|\psi\rangle,|\phi'\rangle)$ is $\langle\psi|\phi'\rangle$, which is $\langle\psi|A|\phi\rangle$. And this is basically a product of three matrices, one, two and three. And it doesn't matter in what order we multiply as long as we keep this order. So, whether we multiply it this way or we multiply it this way, it does not matter. So, we can also write it as $(|\psi'\rangle,|\phi\rangle)$, where $|\psi'\rangle$ is something acting on $|\psi\rangle$. This is not equal to $A|\psi\rangle$ but we write this something as, A this symbol which we call dagger or Hermitian conjugate or adjoint $|\psi\rangle$, so this $A^\dagger$ is called the adjoint operator.

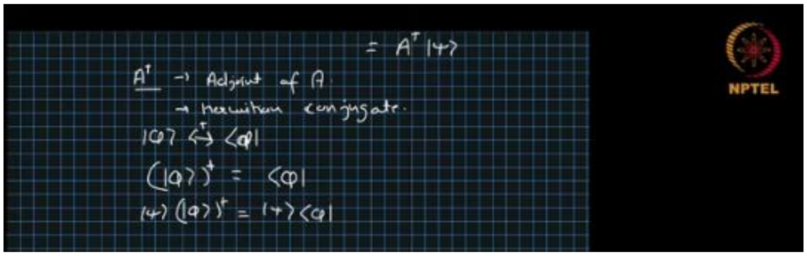
(Refer slide time: 14:31)



In quantum mechanics, we are more familiar with Hermitian conjugate. So, $A^\dagger$ is Hermitian conjugate of A. Since I am a physicist and we are doing the physics course, I can take the liberty of saying that $|\phi\rangle$ and $\langle\phi|$ are also Hermitian conjugate of each other. So, I can write $(|\phi\rangle)^\dagger$ equals $|\phi\rangle$. Operationally, this is correct, but probably some mathematicians will take it as an offense. But for us, this will be the same thing.

So, when we say an outer product, it is actually $|\psi\rangle$ and $(|\phi\rangle)^\dagger$ outer product, which is a product which is $|\psi\rangle\langle\phi|$.

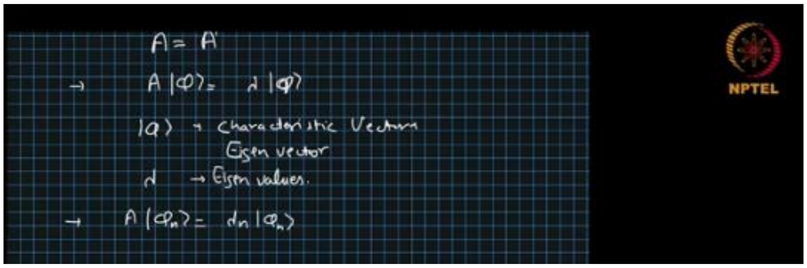
(Refer slide time: 14:31)



Another thing to understand about linear operators is the trace. How do we define a trace of an operator? We generally know that if we are given a matrix, we just take the sum of the diagonal elements, the sum of the diagonal elements of operator A, matrix A is the trace. But more generally, trace of A is defined as $\sum_i \langle u_i|A|u_i\rangle$,where $\{|u_i\rangle\}$ is the set of orthonormal basis, complete and orthonormal basis on a vector space.

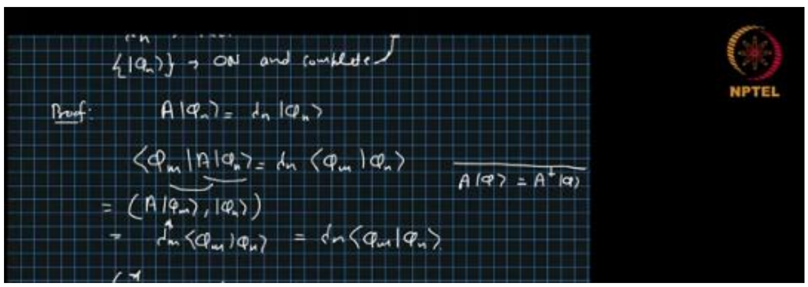
This is how we will define the trace of a matrix. Now there are few very interesting types of linear operators which we will be using throughout this course and that is important class of operators. So, in this class we have first is the Hermitian operators. They have another name which is called self adjoint. So, as it is evident from the name self-adjoint, it means that given operator A, if it is equal to its adjoint, then it's self-adjoint and it's Hermitian operator. So, what is so special about this?

(Refer slide time: 19:20)



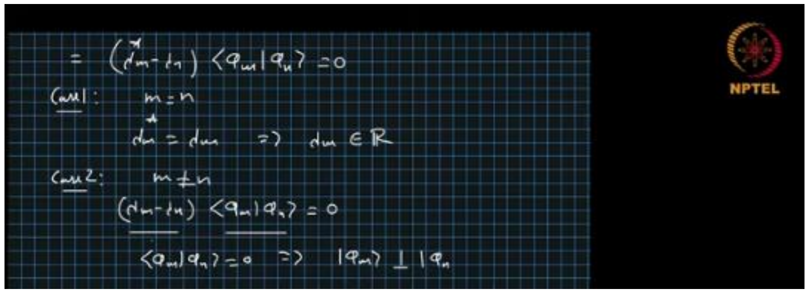
Before going to that, let us discuss. So, if you have operator A acting on a vector $|\phi\rangle$, and you get the same vector $|\phi\rangle$ with a scalar multiple, λ, then these vectors are called characteristic vector. We are we are more familiar with their other name which is the eigenvector, and the scalar λ is called eigenvalue. If you have $A|\phi_n\rangle$ equals $\lambda_n|\phi_n\rangle$, then the set, there are many eigenvectors and eigenvalues, then for a Hermitian operator, λ_n eigenvalues are always real not just that, $\{|\phi_n\rangle\}$, the set of all the eigenvectors of A, they form an orthonormal and complete basis. So let me say it again, for a Hermitian operator, for a self-adjoint operator A, the eigenvalues are always real and eigenvectors form a complete and orthonormal basis. Now proof of that is fairly straightforward. Let us say we have A acting on $|\phi_n\rangle$ gives us $\lambda_n|\phi_n\rangle$.

(Refer slide time: 20:58)



Let us take the inner product of this with some $|\phi_\mathrm{m}\rangle$, it gives $\lambda_n\langle\phi_\mathrm{m}|\phi_\mathrm{n}\rangle$. So, like I said, we can take the product like this or we can take the product like this. So, if we write this as inner product of $\mathrm{A}^\dagger$ acting on $|\phi_\mathrm{m}\rangle$, and $\mathrm{A}^\dagger$ is equal to A. So, $(\mathrm{A}|\phi_\mathrm{m}\rangle, |\phi_\mathrm{n}\rangle)$, then the right-hand side is lambda m phi m phi n and right-hand side is $\lambda *_\mathrm{m} \langle\phi_\mathrm{m}|\phi_\mathrm{n}\rangle$ or we can say $(\lambda *_\mathrm{m} - \lambda_\mathrm{n})\langle\phi_\mathrm{m}|\phi_\mathrm{n}\rangle = 0$, okay. So, now let us say case one when m=n okay. So, when m=n then $\lambda *_\mathrm{m} = \lambda_\mathrm{m}$, this implies that all the λ_m's are real. Okay, so it means all the eigenvalues are real, the first thing is proven. The case two, when m is not equal to n, so we can rewrite this equation as $(\lambda *_\mathrm{m} - \lambda_\mathrm{n})$ since all of them are real we don't need to keep the complex conjugate, $\langle\phi_\mathrm{m}|\phi_\mathrm{n}\rangle$ equal to zero so this is a number scalar. This is a scalar and the product of the two scalar is zero if and only if at least one of them is zero. Since m is not equal to n and let us assume that this is a non-degenerate case, it means all the eigenvalues are different then this quantity is not zero so it must be that $\langle\phi_\mathrm{m}|\phi_\mathrm{n}\rangle$ should be zero which implies that $|\phi_\mathrm{m}\rangle$ is orthogonal to $|\phi_\mathrm{n}\rangle$ whenever m is not equal to m and this proves our whole, our whole claim that the eigenvalues are always real for a Hermitian operator and eigenvectors are always complex, orthonormal and complete.

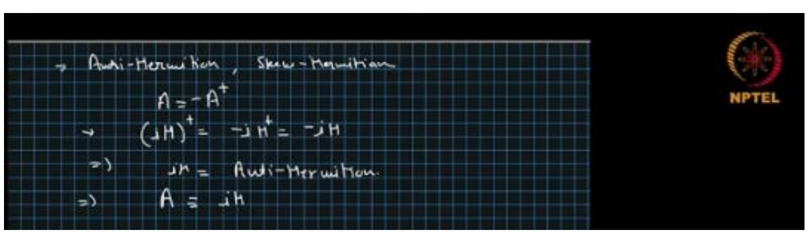
(Refer slide time: 22:01)



Just see that where have we used the property that it has to be a Hermitian operator. The property we have used that it has to be a Hermitian operator is here that A acting on $|\phi\rangle$ and $A^\dagger$ acting on $|\phi\rangle$ is same. It means $A^\dagger$=A, that is the definition of equality, $|\phi_n\rangle$ let us say. So, this is where we have used and we could arrive at this equation of course, we are not assuming the from the beginning that the eigenvalues are real so we have taken complex conjugate and this ultimately, we proved the whole claim. Another class of interesting operators is the anti-Hermitian, or skew-symmetric, skew Hermitian operator. So as the name suggests A, if it is an anti-Hermitian operator, then it should be equal to minus of $A^\dagger$. So, it's not equal to equal to its adjoint but it's negative of its adjoint.

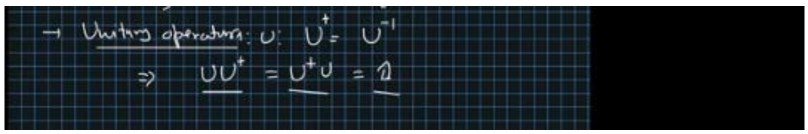
Now, very easily we can see that if we have a Hermitian operator H and we multiply it with the i, which is the square root of -1, the dagger of it is always -i$H^\dagger$ dagger, which is -iH. So, this implies that iH is always anti-Hermitian. We can claim here that all the anti-Hermitian operators can be written as i times Hermitian operators, and this is what we will be using throughout this coursework. Next class of operators, so if A is iH, then we can use all the properties of the Hermitian operators to redefine the properties of the anti-Hermitian. For example, if the eigenvalues of H are λ_i's, which are real, then eigenvalues of A will be i times λ_n. So, it is purely imaginary. So, in that way, we can define many other properties.

(Refer slide time: 24:40)



The eigenvectors will remain the same. So, eigenvectors over anti-Hermitian operator are always orthonormal and complete. And whatever property we have for Hermitian operator, we can see how it translates to anti-Hermitian operator. Next class of operators which are of interest to us are the unitary operators. Operator U, which is unitary, is defined as U is an operator which is, the adjoint of U is actually the U inverse.

(Refer slide time: 26:53)



The operator U is, unitary operator U is defined in such a way that the adjoint of the unitary, the operator is its inverse, or in other way, which is more familiar way of writing, $UU^{\dagger}$ equals $U^{\dagger}U$ equals identity. Let me restate it that $UU^{\dagger}$ equals $U^{\dagger}U$ equals identity. If only one of the conditions are satisfied, not both, then it is not a unitary. It need not be a unitary operator. For a unitary operator, both the condition must be satisfied that $U^{\dagger}U$ and $UU^{\dagger}$ dagger should be identity.