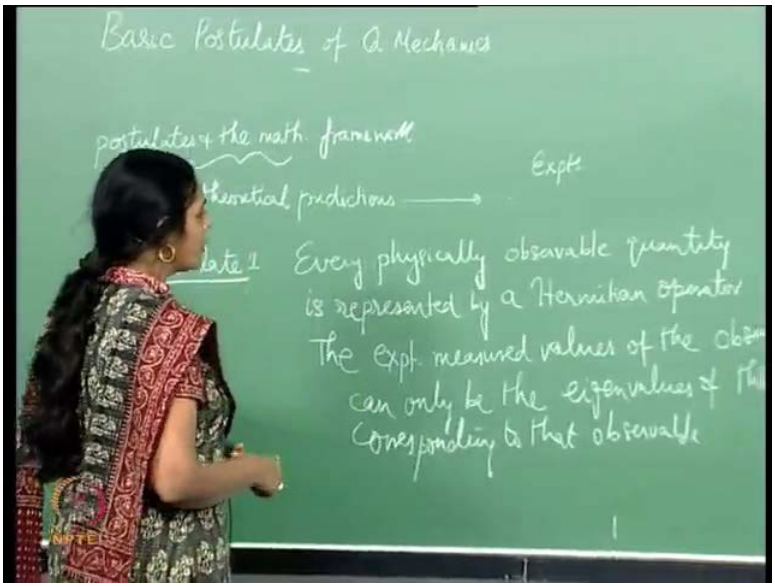
Quantum Mechanics- I
Prof. Dr. S. Lakshmi Bala
Department of Physics
Indian Institute of Technology, Madras

Lecture - 4
Postulates of Quantum Mechanics – I

In today's lecture I will essentially be talking about the basic postulates of quantum mechanics.

(Refer Slide Time: 00:19)

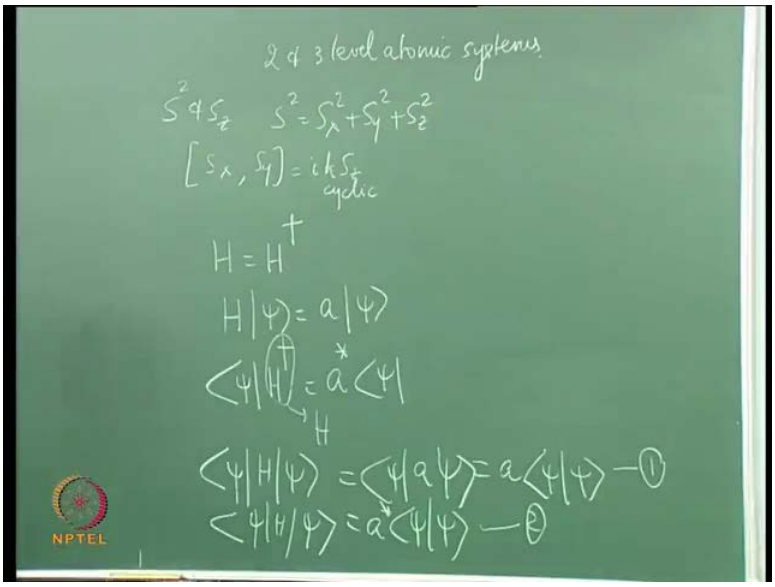


Postulates are a set of hypothesis. And then there is a theory, which is essentially these postulates together with the mathematical framework, which makes predictions, which are supposed to be experimentally verifiable. The theory is not meant to explain the postulates, but given the postulates and the mathematical framework you make theoretical predictions based on these postulates, and compare with experiments, the consequences of these predictions. Now, if the experimental results agree with the theoretical predictions, in a sense it is an indirect test of the postulates. So, quantum mechanics too works on a set of basic postulates.

Several of these postulates have been already told to you by way of examples. The two dimensional linear vector space illustrates these postulates quite well as I will explain to you right away. So, the first postulate it says that every physically observable quantity is represented by a Hermitian operator. If you wish to give matrix representations to

operators this would be a Hermitian matrix. The experimentally measured values of the observable can only be the Eigen values of the operator corresponding to that observable.

(Refer Slide Time: 04:07)



So, for instance in the familiar example of the two level atom the Hermitian operators that we thought up were S squared, and S z. So, too with the three level atom. These are Hermitian operators and we had matrix representation for these operators. In the case of the two level atom there were two by two Hermitian matrices or in the case of the three level atom there were three by three Hermitian matrices, so spin and the third component of spin if you thought about it in terms of the spin doublet which could be the electron or the proton or the neutron. In the case of the two and three level atomic systems it would be just S square and S z which satisfied the algebra S square is equal to S x square plus S y square plus S z square and S z satisfied along with S x and the S y.

The Lie algebra, the experimentally measured values of these observables can only be the Eigen values of the operators corresponding to the observables. And Hermitian operators are selected because Hermitian matrices have real Eigen values, and all measurable quantities will have to be real quantities. So, right away it is good to see why Hermitian matrices have real Eigen values. A Hermitian matrix has this property that H is equal to H dagger, dagger would just mean interchange rows and columns and take the complex conjugate of every element in the matrix.

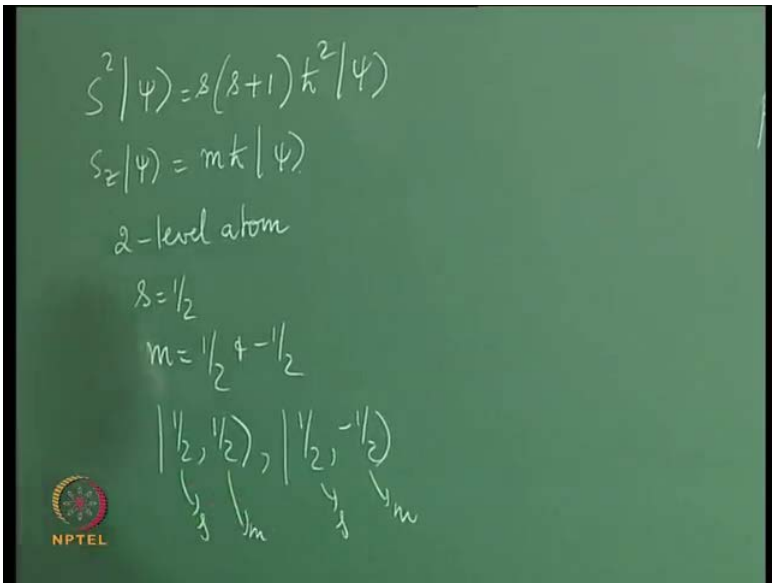
I give you an Eigen value equation, H psi is equal to a psi where a is the Eigen value and psi is the corresponding Eigen vector. Clearly in the Dirac notation, if I take the dagger the ket would become the bra and the number would simply be replaced by its complex conjugate, but H dagger is equal to H. So now, I could well find psi H psi h psi is a state because H is an operator that acted on psi to produce a new state H psi, and we are trying to find the inner product of psi with H psi. This quantity clearly from this equation is psi a psi, but a is a number and can be pulled out.

So, you have the first equation psi H psi is a, inner product of psi with itself, that is the first equation. Well, you could have done that with this. You could have started with psi H and you could have had a psi on this side that would have given me an a star psi psi and this is my second equation. But since both quantities are the same a must be equal to a star. In other words, the Eigen value of the Hermitian matrix is real and that is true for all Eigen values.

So, the set of Eigen values of Hermitian matrices are real and therefore, the experimentally measured values of the observable can only be the Eigen values of the Hermitian operator corresponding to that observable. We also assume that the Hermitian operator is a bounded operator. A concept of a bounded operator really needs explanation only when we deal with infinite dimensional spaces. And since, up to now I have only spoken about finite dimensional vector spaces. All finite dimensional vector spaces have this property that the operators are anyway bounded operators.

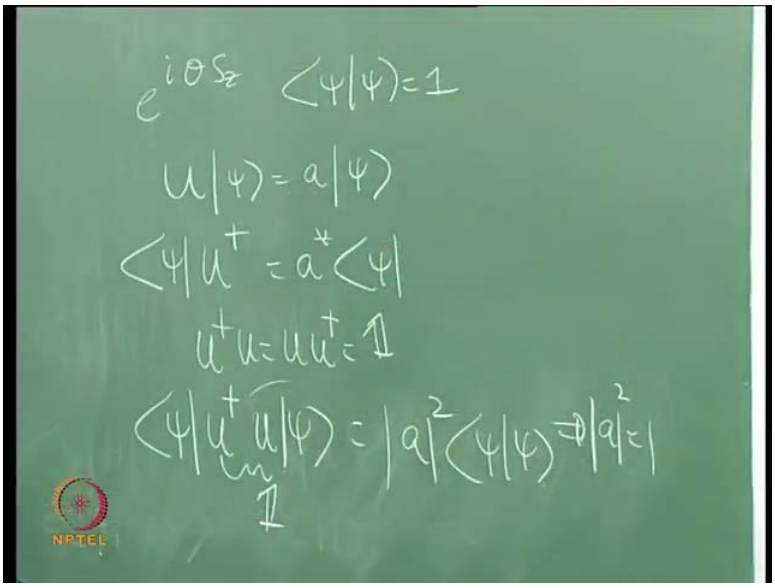
The concept of boundedness of an operator is intimately linked with the concept of continuity and both of them are best explained in the framework of infinite dimensional linear vector spaces, which I will do in a subsequent lecture. So, for the moment these examples the two and the three level atomic systems really have only bounded operators as relevant operators, having said that in general, there is no need to imagine that all the Eigen values of the Hermitian matrix should be a continuous set. In general there could be discrete Eigen values and that is precisely why the observable could be quantized with discrete values, in contrast to a classical system where the measurement outcome could take a continuous set of values. Here, for instance in the familiar example of the two and three level atom equivalently the spin doublet.

(Refer Slide Time: 10:03)



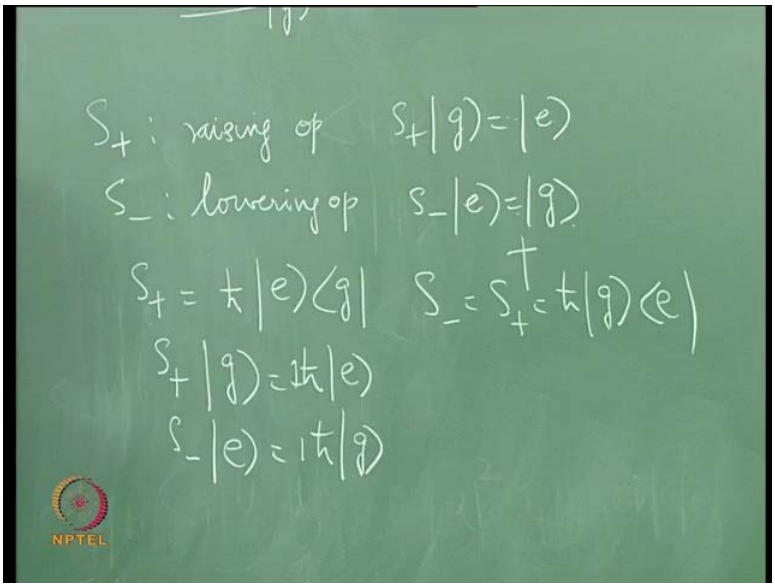
This was certainly true that S square acting on psi was S into s plus 1 h cross square psi, and S z on psi was m h cross psi. In the case of the two level atom s was half and m could take values minus half or plus half, and therefore, I defined for you in a previous lecture two basis states labelled by the s and m values and the states were half half and half minus half. It was on the basis of this, you will recall that in passing I mentioned Fermions and Bosons. Fermions have half integer spin that means s can take values half, three half and so on and Bosons have integer spin, which means s can take values 0 1 2 3 and so on. While Hermitian operators are certainly the only operators, that can be used to represent physically measurable objects. It is also true that other operators have their own importance in quantum mechanics.

(Refer Slide Time: 11:54)



For instance, we have already seen the operators S plus and S minus. We have seen unitary operators an example would be e to the i theta S z, which I spoke about in the last lecture. This is unitary because that is Hermitian and this is exponential i times a Hermitian operator. Talking of Eigen values it is good to digress at this point and see in general what kind of Eigen values a unitary operator can have. So, if u is a unitary operator this is an example: and if this is the Eigen value equation. Correspondingly I can take the dagger and get this equation. A unitary operator satisfies this and therefore, I can find the inner product of this bra with this ket. This is a ket and that is the bra vector corresponding to that ket vector. But this object is simply mod a square psi psi because u dagger u is identity, and if initially psi has been normalised to unity it is clear that mod a square is equal to 1.

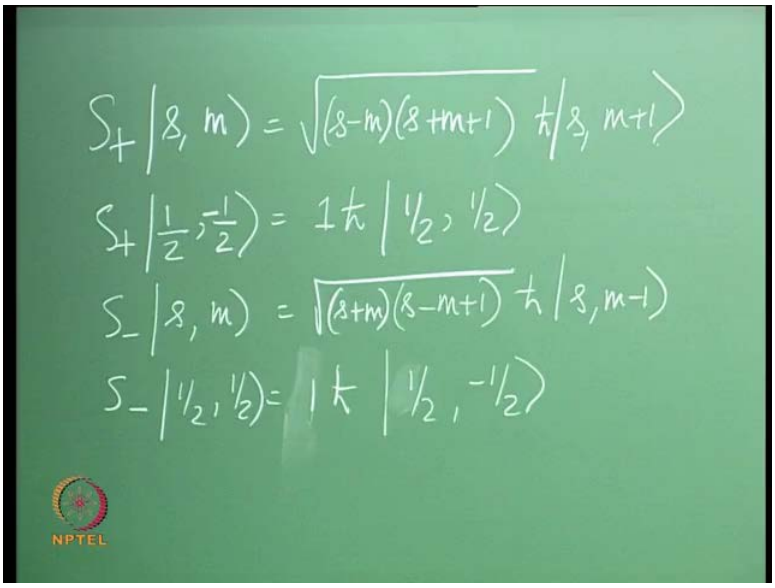
(Refer Slide Time: 13:55)



So the Eigen values of a unitary operator are complex numbers, whose modulus is unity. We also have non Hermitian operators in quantum mechanics which do important jobs. As for instance, S plus and S minus they were the raising and lowering operators. S minus was a lowering operator in the context of the atomic system, because S plus on the ground state of the atom took it to the excited state and S minus on the excited state of the atom gave us the ground state.

We can easily find out the physical significance of the coefficients here. If you go back to the two level atom problem you will see that S plus was defined as h cross e g. Let us recall that these were the two levels and S plus acting on g gave me 1 h cross e and S minus acting on e gave me 1 h cross g, S minus being the dagger of S plus.

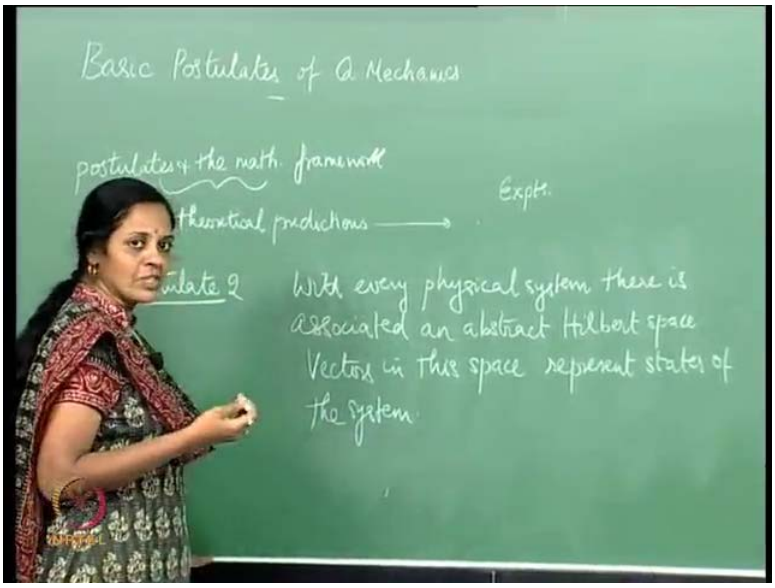
(Refer Slide Time: 15:24)



Now in general, because S plus is the raising operator it acts on the state given by labels s comma m takes it to the state s comma m plus 1 with the coefficient, which is s minus m times s plus m plus 1 h cross. Now, this is a general relation and we will derive this in greater detail later. And therefore, S plus in the case of the spin doublet or the two level atom would act on the state half minus half and give us 1 h cross as the coefficient, and take it to the state half half.

Now similarly, S minus acts on a state given by the labels s comma m, takes it to the state s m minus 1 with the coefficient, which is root of s plus m times s minus m plus 1. And therefore, in our example S minus acting on the state half half will take it to this coefficient which really turns out to be 1 h cross times s comma m minus 1, which is half minus half. This is true even for the three level atom and this is a general expression which can be derived from the angular momentum algebra, so much for the first postulate. The second postulate is the following.

(Refer Slide Time: 17:28)

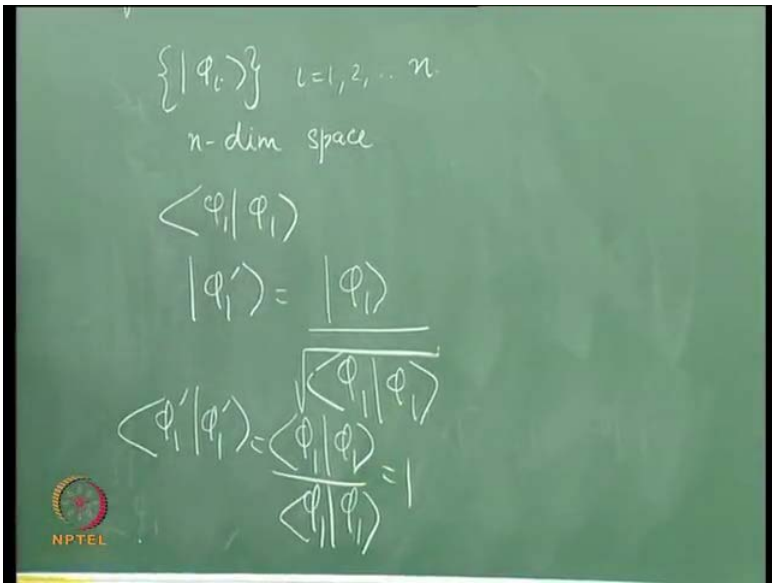


So there is postulate 2. And this postulate says with every physical system there is associated an abstract Hilbert space. Vectors in this space represent states of the system, while I have not introduced the concept of a Hilbert space. A Hilbert space is a linear vector space on which an inner product has been defined and which also has a concept of completeness. Again, completeness is best described and understood in the context of infinite dimensional spaces becomes non trivial there.

But as far as the finite dimensional linear vector spaces are concerned, the completeness relation has already been spelt out by me, and completeness is a concept that needs to be understood mainly in the case of infinite dimensional spaces, which I will do subsequently. The spin system that you have seen is certainly the concept of completeness is pretty much there already, so much for the Hilbert space. In general any state of the system would be a vector in this space and a general vector can be written in terms of the basis vectors as superpositions of the basis vectors.

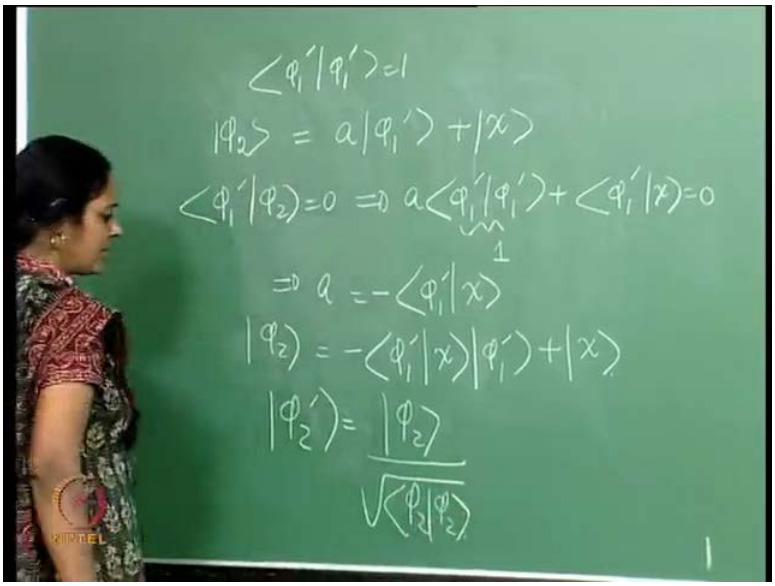
So, already there is the concept of basis vectors, I have this space it is spanned by the basis vectors. So, if it is an n dimensional space there are n basis vectors, by definition these are linearly independent vectors and every vector in the space can be expanded as the superposition of these basis vectors. Now, it is very convenient to choose an orthonormal basis, where the basis vectors are mutually orthogonal and normalised to unity.

(Refer Slide Time: 20:39)



The orthonormalization as I indicated in my last lecture is done through a procedure called the Gram-Schmidt orthonormalization procedure. And before we proceed it is good to understand this procedure. So, suppose I have a basis set let me call it phi i, i is equal to 1 to n because it is an n dimensional space. Let me start with the first of these basis states I can always normalise the state to 1 so, you consider phi 1 phi one. Suppose it is not equal to 1 I can always define the ket phi 1 prime, which is phi 1 by phi 1 phi 1 square root, because then phi 1 prime phi 1 prime is phi 1 phi 1 by phi 1 phi 1, which is equal to one. So, I have normalised one of the basis vectors to unity.

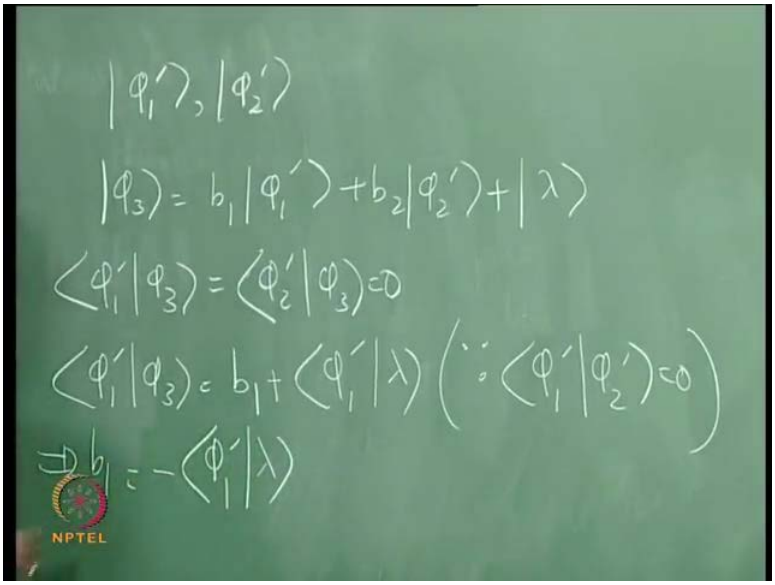
(Refer Slide Time: 22:35)



So now let me look at this second vector in this set. The second vector is phi 2 and I am given that phi 1 prime phi 1 prime is equal to 1. I need to construct from phi 2 a vector which is orthogonal to phi 1. In general, I expand phi 2 as a phi 1 prime plus some other ket chi. I require that phi 1 prime is orthogonal to phi 2. So, clearly from the first term I get this and from the second term I get this.

So, this should be equal to 0. But, this is 1 because I have already normalised phi 1 prime and therefore, a is equal to the inner product phi 1 prime chi. In other words, phi 2 should be selected to be minus phi 1 prime chi this inner product which is in general a complex number phi 1 prime plus chi. Now we need to normalise phi 2 to unity, and for that as before i will divide phi 2 given in this fashion by the square root of the inner product of phi 2 with itself. And therefore, phi 2 prime phi 2 prime inner product is 1.

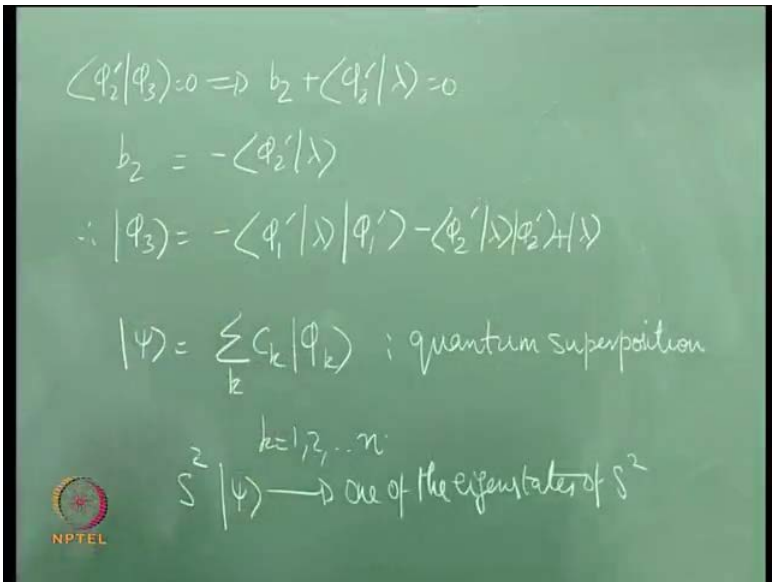
(Refer Slide Time: 24:48)



So, I have normalised the new vector which is phi 2 prime. I therefore, have two vectors in the space which are normalised to 1 and which are orthogonal to each other, namely phi 1 prime and phi 2 prime. I have to repeat this procedure with the 3rd vector phi 3, which I expand as b 1 phi 1 prime plus b 2 phi 2 prime plus some vector lambda. My requirements are the following; that phi 1 prime phi 3 is equal to 0 and phi 2 prime phi 3 is also equal to 0. It is evident, that if I first work with phi 1 prime phi 3 that just gives me b 1 plus phi 1 prime lambda, because I have already shown that this inner product is

0, which tells me that b 1 should be equal to the inner product phi 1 prime lambda with a negative sign.

(Refer Slide Time: 26:06)



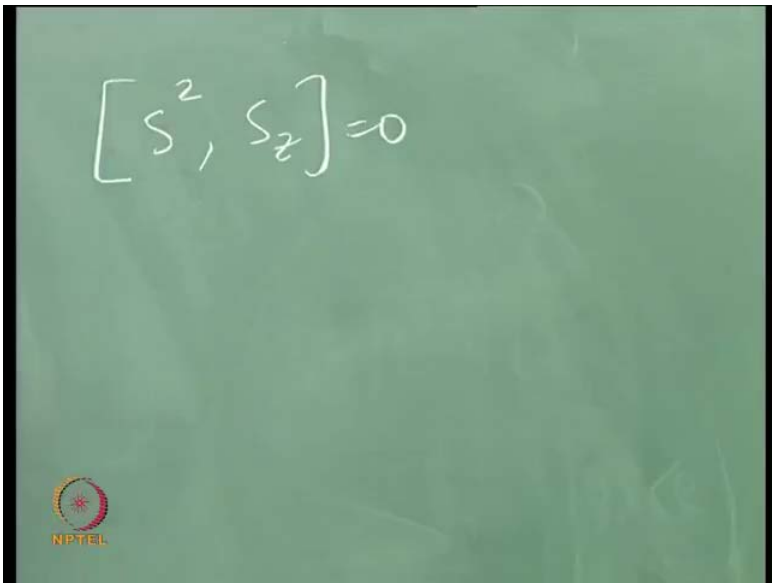
Similarly, the fact that phi 2 prime phi 3 is equal to 0 implies from here, that b 2 plus phi 2 prime lambda is equal to 0 or b 2 is equal to minus phi 2 prime lambda. Again I have used the fact that the inner product of phi 1 prime with phi 2 prime is equal to 0. Therefore, I expand phi 3 as minus phi 1 prime lambda. Phi 1 prime this is a ket and this is a number minus phi 2 prime lambda phi 2 prime plus lambda. In this manner I can proceed and get an orthonormal basis of n basis vectors. By definition they are linearly independent and now they are also mutually orthogonal, and each vector is normalised to 1, so much for the Gram-Schmidt procedure.

The second postulate clearly said that every vector in the Hilbert space represents a state of the system. So, a general vector in the Hilbert space which I will call psi can be expanded as a superposition of the phi's or the phi primes now in this context, because these are the orthonormal basis I can without loss of generality remove the prime, and say that I have a set of n mutually orthogonal vectors phi k, which are the basis set and any vector in that space can be expanded in terms of this basis set. This is the expansion postulate; this is quantum superposition because I can superpose basis states to produce all vectors in that state.

Now the question is the following: I have operators given by my first postulate which act on states what kind of basis states can I select? I have already demonstrated in the context of the two and three level atoms that you could choose different basis states and these are unitarily related to each other. They are related by a unitary transformation i have also demonstrated that the Lie algebra is preserved under such a unitary transformation. Suppose I were making a measurement of some physical observable. Let us take the familiar example of S square and if it acts on any state in that space, quantum mechanics tells us that the state will collapse to one of the basis states or Eigen states of S square in this case with the corresponding Eigen value. Because the measurement outcome is simply going to be one of the Eigen values of S square.

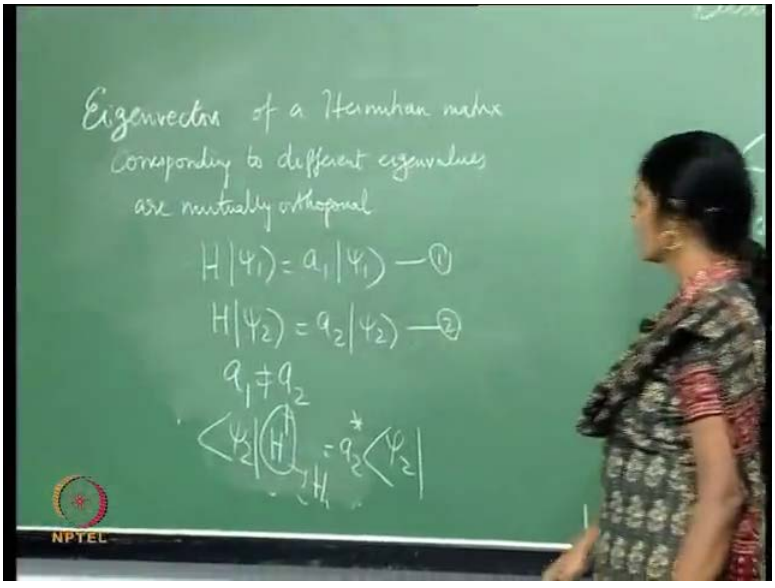
And therefore, the basis state in this context would be simply the Eigen states of the observable S square and the measurement of an arbitrary state will lead post measurement to one of the Eigen states of S square. With the measurement outcome which is the Eigen value, I could well make a measurement of S z as well. Now if I did that, again the system will collapse to one of the Eigen states of S z with the corresponding Eigen value. But the physical state of the system is the same. And therefore, the Eigen state that I have finally, a post measurement of S square and S z must be a common Eigen state of S square and S z. In other words, what happens to the system after measurement is this the system collapses to a state which is a common Eigen state of the various observables that are measured simultaneously giving corresponding Eigen values as the measurement output.

(Refer Slide Time: 31:33)



In this context, one needs to understand that, if two operators commute with each other, you can find a complete set of common Eigen states. First of all one needs to understand what one means by a complete set of Eigen vectors of each operator. As I said a bounded Hermitian operator has a complete set of Eigen vectors.

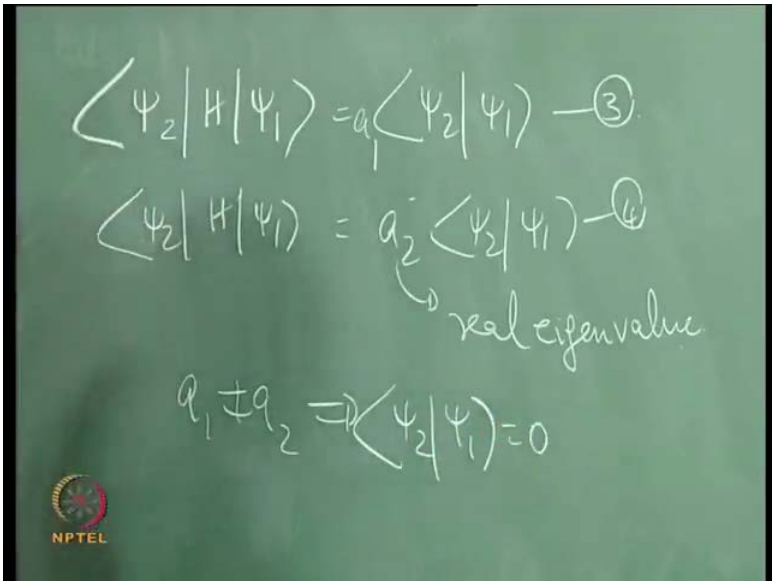
(Refer Slide Time: 32:22)



Now, to digress a little bit if you take a Hermitian operator we can show that Eigen vectors of a Hermitian operator or a Hermitian matrix of a Hermitian matrix, corresponding to distinctly different Eigen values are mutually orthogonal. So, the

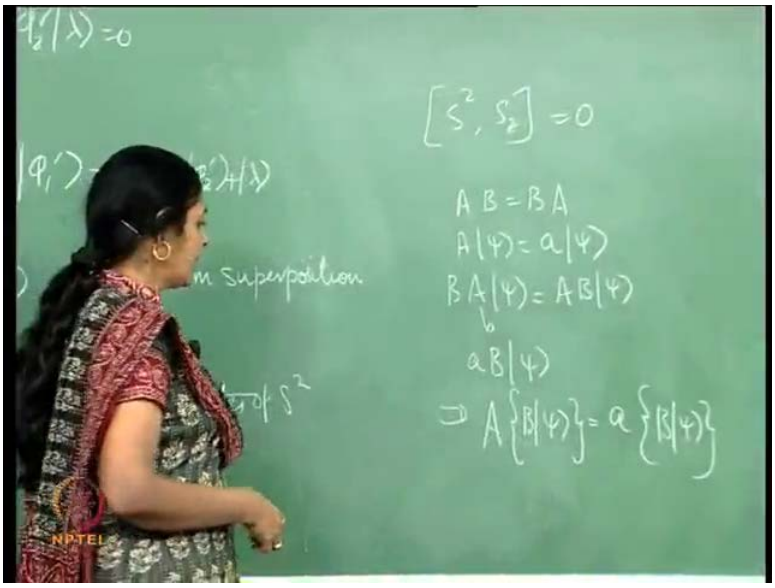
bounded Hermitian operator has a complete set of Eigen vectors and if these Eigen vectors correspond to distinctly different Eigen values, they are mutually orthogonal. This can be seen very simply because if I have a Hermitian matrix and suppose this is the first Eigen value equation where a 1 is the Eigen value and psi 1 is the Eigen vector. And I also have another Eigen vector satisfying this Eigen value equation and the Eigen values are different from each other.

(Refer Slide Time: 33:52)



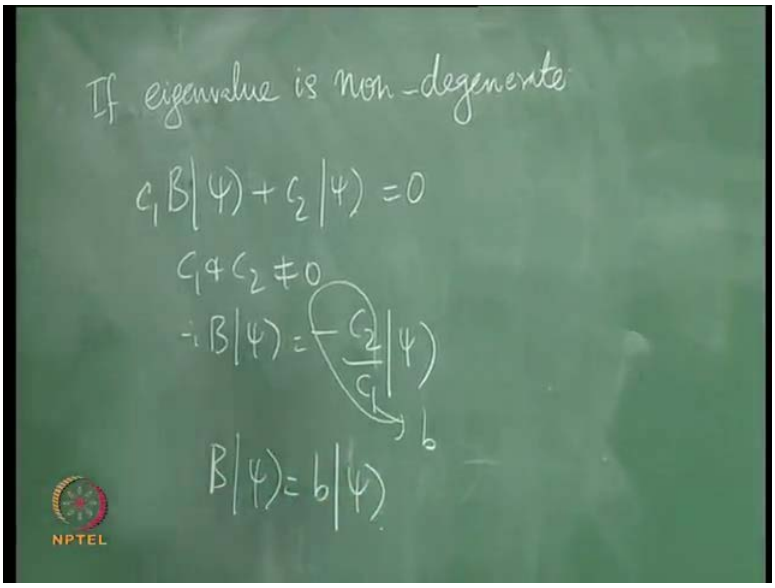
Then clearly I can do the following thing. I can find out psi 2 h psi 1. Now it is clear, that if this equation holds h dagger is the same as h is a 2 star psi 2. So, if i find out psi 2 h psi 1 that will be psi 2, but from my first equation h psi 1 is a 1 psi 1 and a 1 is a number. But, I could have used the second equation and that just gives me a 2 star. I know already I have proved that Hermitian matrices have real Eigen values. So i can drop this star and a 1 is not equal to a 2; implies if this has to be true psi 2 is orthogonal to psi 1. (Refer Slide Time: 32:22) So that is how you prove that Eigen vector is of a Hermitian matrix corresponding to different Eigen values are mutually orthogonal.

(Refer Slide Time: 35:35)



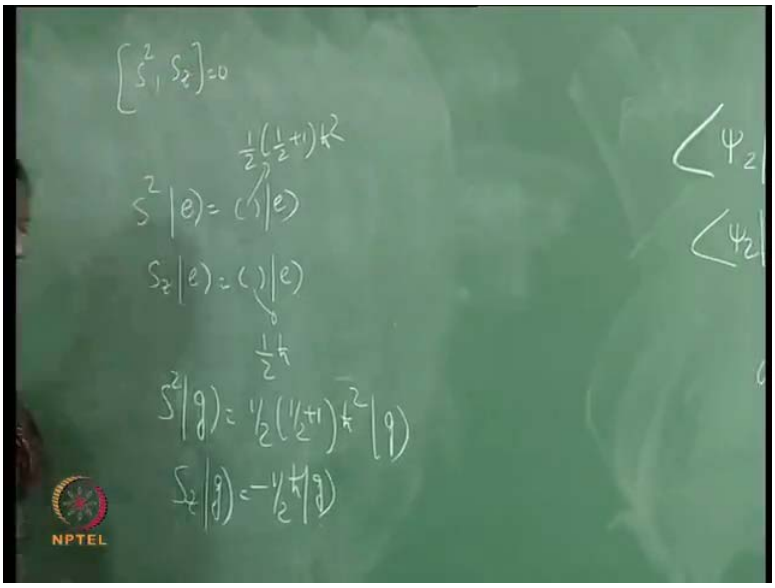
So, now the next part is to show that if I have two such Hermitian operators. Let us say S square and S z and they commute with each other. Then I should be able to find a complete set of common Eigen vectors. This is an example: In general if A B is equal to B A that is A and B commute with each other, and if psi is an Eigen vector of A with Eigen value a, B A psi is equal to A B psi. But I know that A psi pulls out an Eigen value a so this object is simply a b psi. And therefore, the state B psi is an Eigen state of A the operator A with Eigen value a. Now two cases arise, A could be non degenerate. In other words there is not more than one Eigen vector corresponding to the Eigen value A, if at all there is one more that is linearly dependent on the other Eigen vector. So, i consider that case that is an easy case to consider.

(Refer Slide Time: 37:13)



So, if the Eigen value is non degenerate it is clear that the new vector B psi is linearly dependent on the vector psi. So a constant C 1 B psi plus C 2 psi is equal to 0, if there is linear dependence C 1 and C 2 not equal to 0 therefore, B psi can be written as minus C 2 by C 1 psi. And his is just a number, which I will call b. And therefore, psi is also an Eigen state ket psi is an Eigen state of the operator B with Eigen value b. And therefore, I have found a common Eigen state ket psi for the two operators A and B which commute with each other. Now, in the even that there is a degeneracy you have to work a little bit more. Consider linear combinations of the degenerate Eigen vectors and show once more that there is a complete set of common Eigen vectors for the two operators that commute.

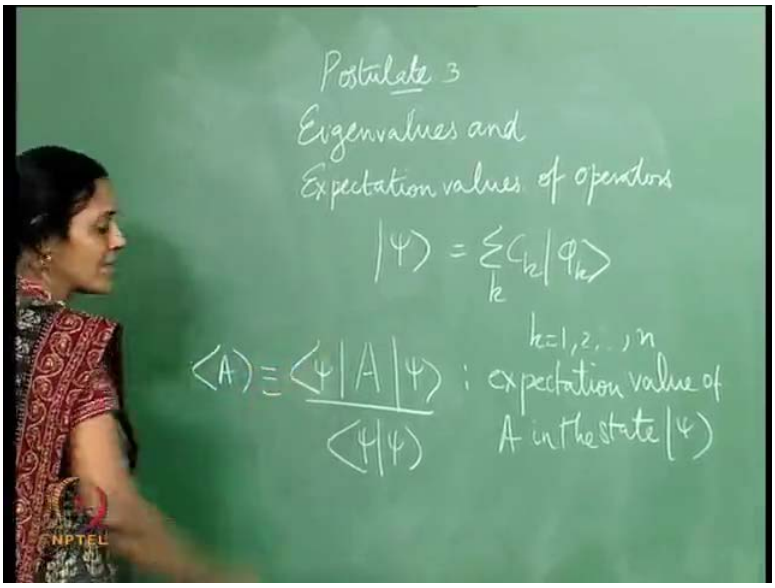
(Refer Slide Time: 38:41)



Taking the example that we are familiar with, S square commutes with S z and we know that they have common Eigen vectors S square e and S z e is an Eigen value equation. And this is simply going to be half into half plus 1 h cross square and this is just going to be half h cross. Similarly, S square on the ground state gives me half into half plus 1 h cross square g and S z g is minus half h cross g. S square and S z commute with each other. So this state collapses to a common Eigen state of these observables. The state itself is a very interesting concept in quantum physics. Because, in classical physics if you look at phase space you simply need to know the value of the generalised coordinate I call it x, and the generalised momentum p corresponding to x.

At any instant of time to completely know the state of the system, then there is the equation of motion. The equation of motion will tell you how exactly you could find the values of x and p at a later time. This state is completely determined. But here the state of the system is described by this ket in an abstract Hilbert space. In order to find out what is this state of the system, we need to first of all know what are the objects that we are measuring. In this context, we need to know that we are looking out for Eigen values of S square and S z. So, the dynamical variable becomes important and then one talks about what is the state of the system post measurement.

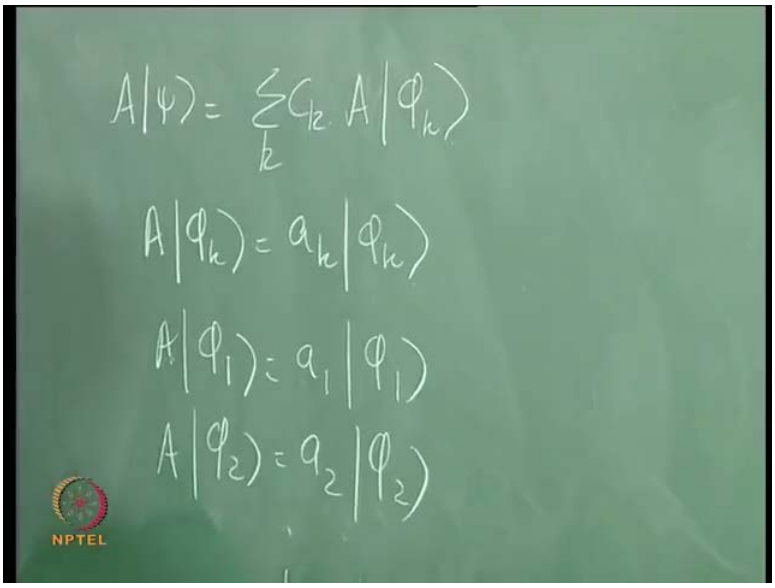
(Refer Slide Time: 40:52)



The 3rd postulate is a very important and interesting postulate about measurement itself, and this has to do with Eigen values and expectation values of operators. So, the postulate basically tells us this. So, this is postulate 3. It is clear what the Eigen values are we have already spoken about the Eigen values being the measurement outcomes. But since every state in the linear vector space or the Hilbert space is a possible state of the physical system, I could in general have a state psi which is a superposition of the basis states in an n dimensional linear vector space, and it need not be an Eigen state of the operator a.

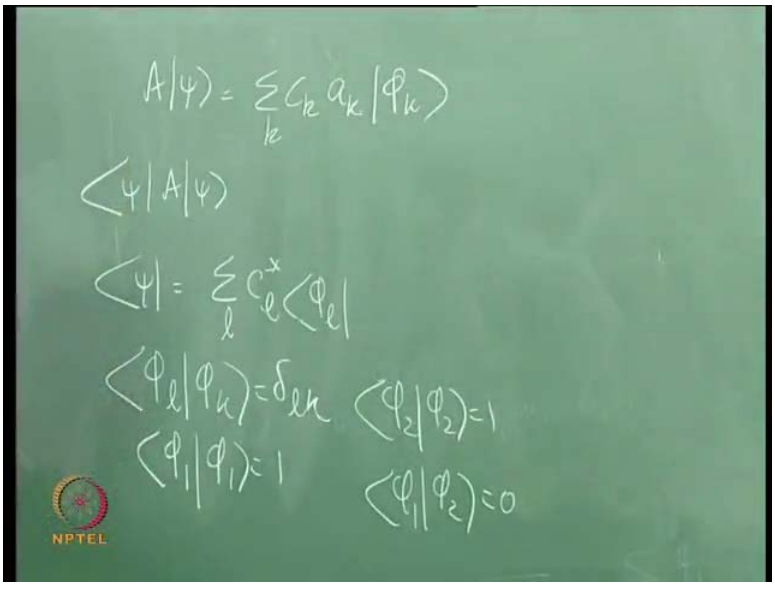
This is the physical observable I am interested in measuring and this is the operator corresponding to that physical observable. In that case one does not have an Eigen value equation, but one talks of the average or expectation value of A in the state psi. This should be suitably normalised because psi itself may not be normalised to unity and this has a short hand notation, A is identical to this. So, wherever the state is not an Eigen state of A, it is a different matter that it will collapse through the Eigen state after measurement. The value of the physical observable that I intend to measure is given by this expectation value or the average value and that is a symbol which denotes it. The denominator has been put in because the state need not in general be normalised to one, if it is normalised to one this of course, becomes one and the numerator will ((Refer Time: 43:10))

(Refer Slide Time: 43:18)



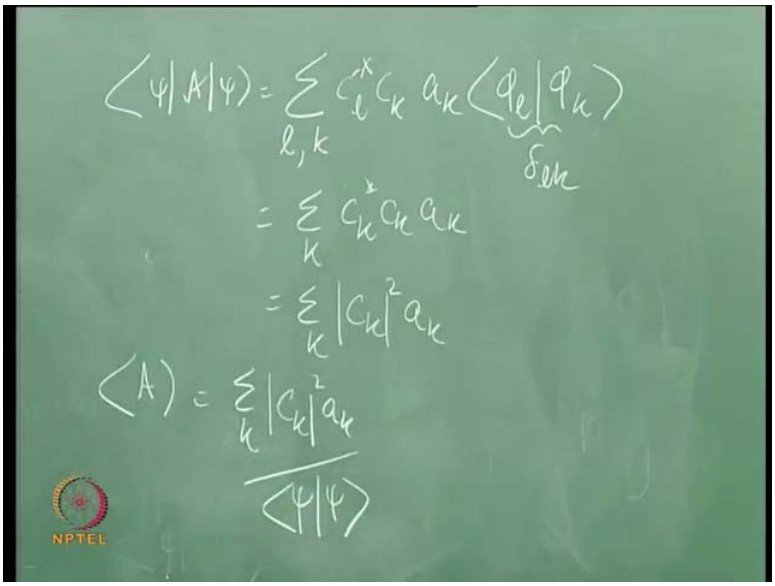
So, let us look at this expectation value in the context of the general state psi. So I have A psi is summation over k, C k A acting on phi k where phi k are the basis states corresponding to A. It should be read off like this. There are a set of states phi 1 phi 2 up to phi n. So, I have A phi 1 is a 1 phi 1 A phi 2 is equal to a 2 phi 2 and so on all the way to k is equal to 1 2 3 4 to n. So, in general if i take any one of these the Eigen value is this and since this is merely a number, and A acts linearly it just acts on every term in this expansion and therefore, I have this. But this object is clearly given in the following manner.

(Refer Slide Time: 44:25)



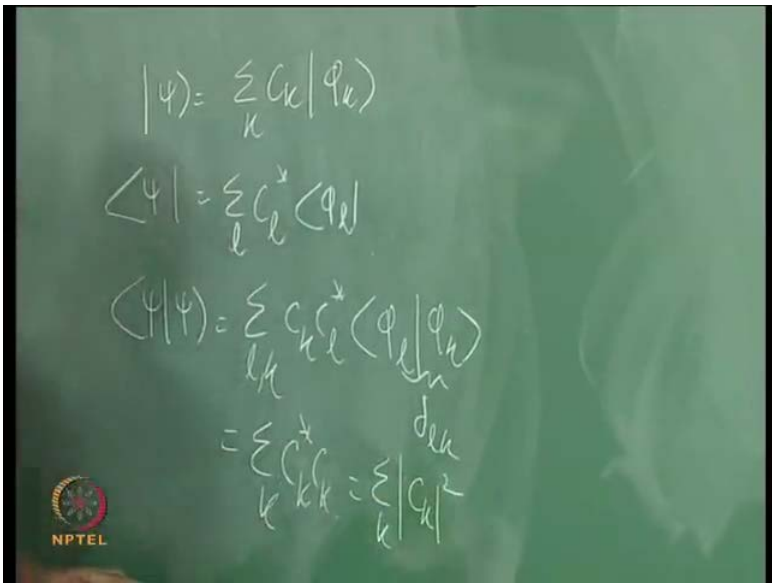
A on ket psi is summation over k C k and from these Eigen value equations it is clear that this is what I have. Now you see I need to find psi A psi that is the numerator. This object is simply equal to summation over l C l star phi l. The ket has become a bra vector these are numbers these are coefficients, so when I take the bra each coefficient becomes is replaced by its complex conjugate, and since I do not want to confuse indices I have used l here instead of k. And therefore, I sum over l here and I sum over k there. I need to use the fact that this is in an orthonormal basis. I can write the orthonormality condition as phi l phi k inner product is delta l k where delta l k is a Kronecker delta it means that when l is equal to k is equal to 1 say the answer is 1. Similarly, when l is equal to 2 the answer is one and so on, but if l and k are different the answer is 0. So this is what I have and these are numbers so I can well write it in the following manner.

(Refer Slide Time: 46:12)



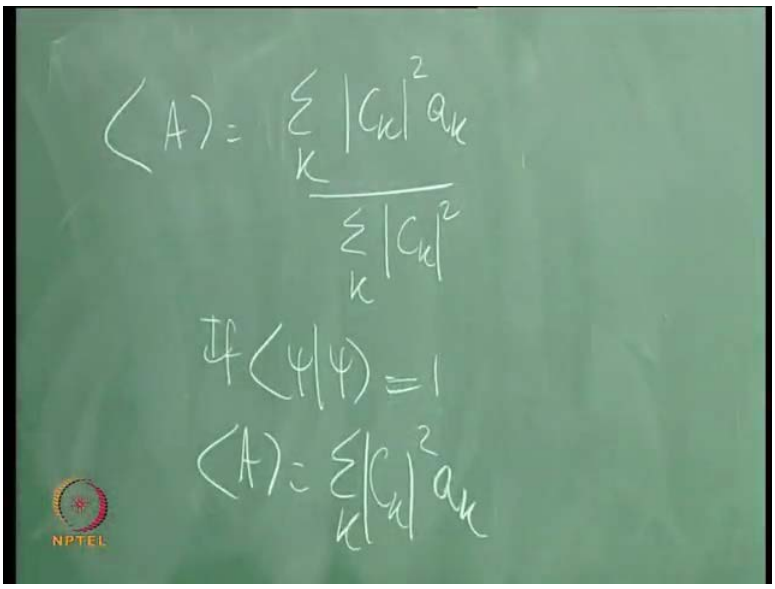
I have psi A psi is summation over l and k that is a double summation c l star C k a k. The inner product phi l with phi k and that was a delta l k, which means you can get rid of one of the summations and I have C k star C k a k. I can well write this as modulus of C k square a k therefore, the expectation value of A is summation over k mod C k square a k by this.

(Refer Slide Time: 47:15)



This object is easy to determine in the event that psi is not normalised to unity. And since psi is expandable in terms of the basis vectors in that fashion, I can find out the inner product psi psi. Use the fact that there is a delta l k out here, remove one of the summations is a number, and therefore I just have summation over k modulus of C k square.

(Refer Slide Time: 48:08)



And therefore, the expectation value of a, it is a average value of A as a measurement outcome is simply given by summation over k modulus of C k square a k by summation

over k modulus of C k square. Now, the event that psi is normalised to one it is clear that if this is 1 modulus C k square summation over k is 1. And then, it is clear that expectation value of a simply has the numerator because the denominator is one. This has to be properly interpreted.

It means the following. Suppose, I conduct a number of trials to experimentally determine A and I find the average value of these trials the measurement outcome would be one of the Eigen values of a, which is a k could be a 1, could be a 2, could be a 3 so one of the Eigen values with probability modulus of C k square. And therefore, the Eigen value a 1 will occur with probability in fact with probability modulus of C 1 square by summation over k modulus of C k square, but I am assuming that psi is normalised to 1 and then it makes it simpler to explain because a 1 will occur with probability mod c 1 square, a 2 with probability mod c 2 square and so on.

So, the measurement outcome really occurs with a certain probability. Each of these outcomes is possible with a certain probability and therefore, this is a weighted average that we have here. This is the sum and substance of the 3rd postulate which tells us about expectation values as oppose to Eigen values and measurement outcomes. I will continue to describe this and go on to the other postulates in my next lecture.