FOUNDATIONS OF QUANTUM THEORY: NON-RELATIVISTIC APPROACH

Dr. Sandeep K. Goyal
Department of Physical Sciences
IISER Mohali
Week-01
Lecture-02
Axiomatic Approach to Quantum Mechanics - Part 02

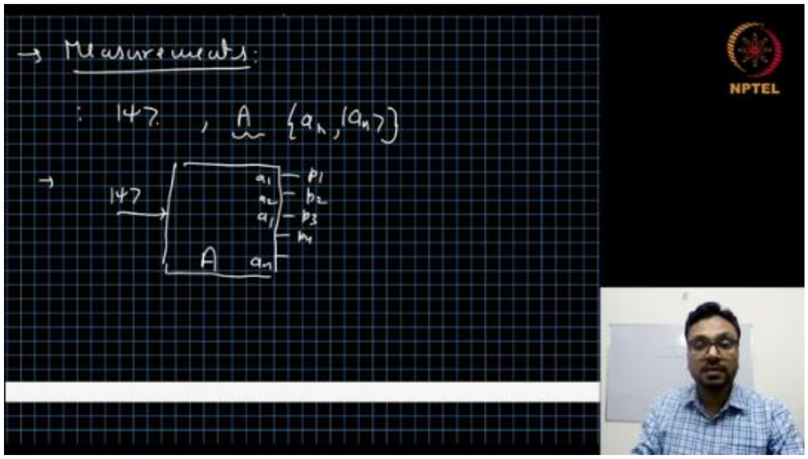
Our next and the last postulate is about measurement. So far, we have seen that states are the vectors in the Hilbert space, the observables are the Hermitian operators, the dynamics are given by Schrodinger equation, which is kind of related to the wave equation. So far, whatever we have studied, whatever the property we have seen or we will see, for example, the superposition of states, the entanglement which comes with the superposition of states and some such properties, they can be seen in the classical optics. Classical waves also have superposition properties.

If we have more than one degree of freedom for a wave, we can probably see the entanglement properties also. What makes quantum mechanics really special is the measurement postulate. So, measurement postulate is the most striking thing which gives us all the advantages and all the mysteries in the quantum mechanics. We see the quantum key distribution gives us like really absolute security in the communication because of the measurement postulate. We have quantum advantage because of entanglement, but measurement also plays a very significant role here.

Let us say we have a quantum system in the state $|\psi\rangle$ and we are performing the measurement of an observable A given by the eigenvalue a_n and eigenvectors, $|a_n\rangle$. So, we have a setup, measurement setup. Measurement setups do not look like this how I am drawing, but this is just a conceptual representation. In this setup, the state $|\psi\rangle$ goes in, it means the quantum particle in the state $|\psi\rangle$ goes in. So, we may or may not know what is the state that is irrelevant, but the experimental setup is designed to perform the measurement of the observable A. Then the outcomes will be a_1, a_2, a_3 and so on all the eigenvalues, whatever observable we have in quantum mechanics when we perform measurement then eigenvalues will be what we will observe in the lab, and they will occur with some probability, p1, p2, p3, p4.

In the sense that if we send one quantum particle, it can click on any of these outcomes. It can fall on any of these outcomes. And if we repeat this whole thing N number of times where N is sufficiently large, then it will click in the first one with N_1 number of times it will click in the first one, N_2 number of times it will click in the second one and so on. And the probability p1 will be close to N_1 over N, N_2 over N. So, let me repeat.

(Refer slide time: 3:42)

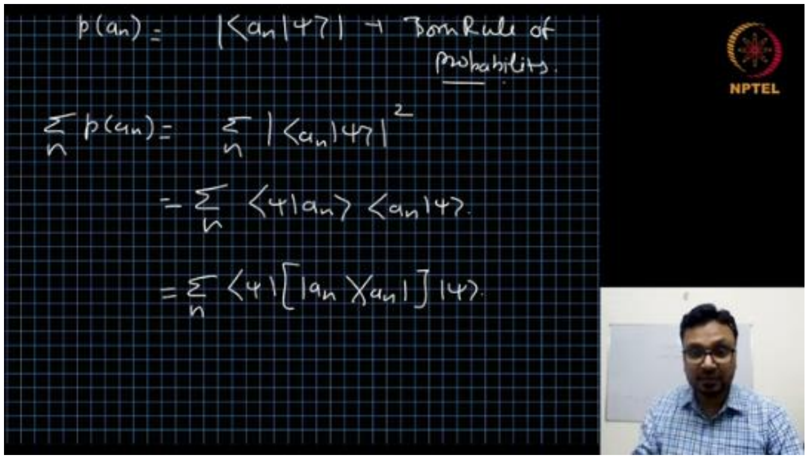


We have an experimental setup which is designed to perform measurement on the observable A. Then we send the quantum systems identically prepared in the state $|\psi\rangle$, N number of quantum systems. And we send them one by one through this setup and they will give us clicks over these outcomes. And the probability of these clicks is given by the N_1 over N. So, the particle reached the outcome a_1, N_1 number of times. The probability p_1 is N_1 over N. But what is, how is this probability p related to $|\psi\rangle$ and A? So, the probability of outcome a_n is given by the measurement of the eigenstate of the observable A as $|\langle a_n|\psi\rangle|^2$.

So, the probability of the outcome, $p(a_n)$ is given by $|\langle a_n|\psi\rangle|^2$. This is called the Born rule of probability. This is the first axiom in this is the first sub axiom of the measurement axiom measurement postulate that the probability in an experiment, the probability of the outcome a_n is given by the corresponding eigenvectors and $|\psi\rangle$ in a product mod square, now we can see that $\sum_n p(a_n)$ is $\sum_n |\langle a_n|\psi\rangle|^2$. We can write it as $\sum_n \langle\psi|a_n\rangle\langle a_n|\psi\rangle$. And this is $\sum_n \langle\psi|[|a_n\rangle\langle a_n|]|\psi\rangle$.

And since $|a_n\rangle$ are the normalized eigenvectors of the Hermitian operator A, they add up to identity. So, if we take the summation inside, we get $\langle\psi|\{\sum_n |a_n\rangle\langle a_n|\}|\psi\rangle$ and this is 1. So, if we add all the probabilities for all the outcomes, we get one. This is, and this is what any probability distribution should satisfy. Next is the expectation value, the average value of the observable A, $\langle A \rangle$. And average value of an observable in classical statistical mechanics can be written as $\sum_n a_n\, p(a_n)$.

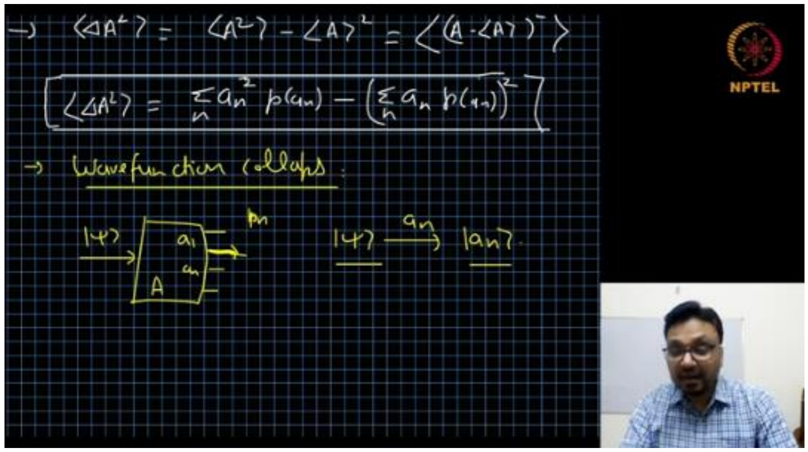
(Refer slide time: 5:22)



So, this is, in the statistical mechanics, this is the probability of an observable where a_n is the value of the observable and p_n is the probability of that outcome. Now, if we use the Born probability rule for $p(a_n)$, then we get $\sum_n a_n |\langle \psi | a_n \rangle|^2$, $|\langle \psi | a_n \rangle|^2$ and $|\langle a_n | \psi \rangle|^2$ are same things. And this we can write as $\sum_n a_n \langle \psi | a_n \rangle \langle a_n | \psi \rangle$. We can take summation and a_n inside the, we can rewrite it like $\langle \psi | [\sum_n a_n |a_n\rangle \langle a_n|] | \psi \rangle$, and $\sum_n a_n |a_n\rangle \langle a_n|$ is the spectral decomposition of the observable A and we can replace it with A. So, we get $\langle \psi | A | \psi \rangle$ as the expectation value. So, what we have done is we have taken the classical definition of the average of a measurable quantity and that is the value of the quantity times the probability of that outcome.

And for the probability, we have substituted the Born rule of probability for quantum systems and we got an expression for the average or expectation value of the observable in terms of the state of the system $|\psi\rangle$ and the operator A. So, this is the expression for the expectation value and with this derivation you can see that the definition of this expectation value is inspired from the classical statistical average of any quantity. Of course, we can generalize this definition to find the expectation value of A^2. So, to calculate this thing first we see that A^2 is nothing but $\sum_n a_n^2 |a_n\rangle\langle a_n|$. This can be seen by taking the square and realizing that $|a_n\rangle$ and $|a'_n\rangle$ are orthogonal vectors, so the average, the expectation of A^2 square will be $\langle A^2 \rangle = \langle \psi | A^2 | \psi \rangle$, and that will be $\sum_n a_n^2\, p(a_n)$, and from here we can calculate something called the variance of A, $\langle \Delta A^2 \rangle$ that will be $\langle A^2 \rangle - \langle A \rangle^2$ or that is $\langle (A - \langle A \rangle)^2 \rangle$, and that will be $\sum_n a_n^2\, p(a_n) - (\sum_n a_n p(a_n))^2$, so we'll leave it at this thing and we will use this expression for the variance.

(Refer slide time: 10:33)



The second part of the measurement postulate is the collapse, wave function collapse So we have the experimental setup, A. We have state |ψ⟩ going in, quantum systems in the state |ψ⟩ identically prepared, one by one we are sending them in the setup and we have clicks in certain outcome, a_1, a_n, with probability p_n given by the Born rule. What happens to the quantum system after the measurement? So, the force postulate, the measurement postulate or the collapse postulate of the measurement postulate states that the state of the system after the measurement collapses to the eigenvector of the observable corresponding to the outcome. That is, if |ψ⟩ goes in and we get a_n outcome, then the state of the system after the measurement will be $|a_n\rangle$.

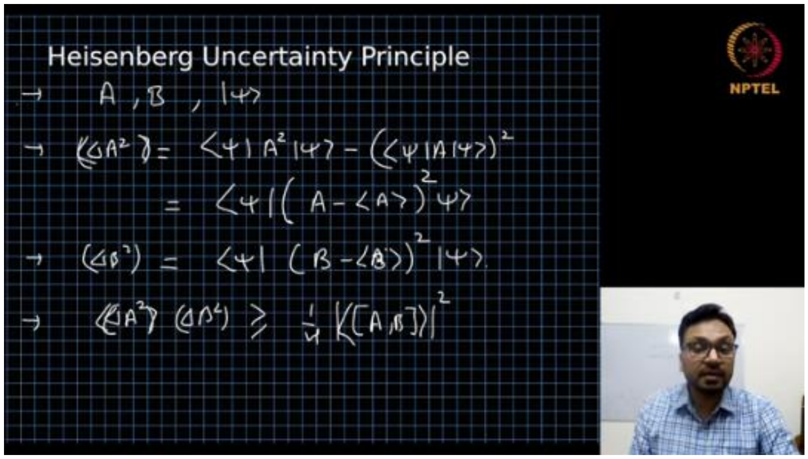
So, in this way, the process of performing measurement on the quantum system destroys or erases the information of the original state and replace it with the new state that is the eigen state of the observable. So, in that way once we get a click in certain outcome, once we get an outcome the in the quantum system has lost all the information about its original state so this state cannot be used to further perform measurement on it in order to get more information out of it because there is no information left in it other than that it's in the $|a_n\rangle$ eigen state of the observable A. In that way, when we want to perform measurement on a quantum system, we consider a large ensemble of quantum systems, each of them or every one of them is prepared in the same state |ψ⟩ and we send those states |ψ⟩ one by one through this experimental setup and we collect the statistics of the outcomes. And from those statistics, we can calculate the expectation value of the observable and that can be used for whatever processing we want to do. But we cannot use the same, we cannot infer any information from a single quantum system because once we send single quantum system in the experimental setup it collapses to one of the eigen states of the observable and it has lost all the information about it and we cannot use it for extracting any more information.

So, in that way, we cannot get any information out of a single quantum system because it will just give us one click. It will not give us any probability. It will not give us any information about the state $|\psi\rangle$. Hence, it is not possible to do any meaningful thing with a single quantum system. So, this was the detailed explanation or discussion about the four postulates of quantum mechanics. In this course, we will be coming back to these postulates again and again and we will be going in more and more details and find good applications or interesting applications of these postulates.

Our next topic is Heisenberg Uncertainty Principle. And this shows a striking difference between classical and quantum world. So, in classical picture, whenever we have, we perform a measurement on a system A and we get certain uncertainty, certain randomness in that system. This randomness does not affect the randomness of any other system under the situation, but that is not true in quantum mechanics. The meaning the full meaning of the statement will be clear soon but we'll start here with the following let us say we have two observables A and B, they are Hermitian observables, they are observable so they must be Hermitian so A has its own spectrum a_n and eigenvectors $|a_n\rangle$, B has its own spectrum and eigenvectors so the uncertainty in A that is $\langle \Delta A^2 \rangle$.

So, we have a quantum system with these two observables, we have quantum system with these two observables and the state of system is $|\psi\rangle$. So, the uncertainty in A is $\langle \Delta A^2 \rangle$ that can be written as $\langle \psi|A^2|\psi\rangle - (\langle \psi|A|\psi\rangle)^2$, or it can be written as $\langle \psi|(A - \langle A\rangle)^2|\psi\rangle$ that is the expression for the uncertainty in A, similarly we have uncertainty in B, $\langle \Delta B^2 \rangle$ and that can also be written as $\langle \psi|(B - \langle B\rangle)^2|\psi\rangle$. Now, the Heisenberg uncertainty principle states that $\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle$ will always be greater than or equal to 1 over 4 times $|\langle [A, B]\rangle|^2$, so what what does it say is if we have two observables A and B which do not commute, if we have such observable which do not commute then the uncertainty in A times uncertainty in B is lower bounded by the 1 by 4 times the uncertainty in the commutator of A,B. So, this is a very strong statement and it comes only because of the quantum mechanics because in classical world we do not have observables which do not commute.

(Refer slide time : 15:20)

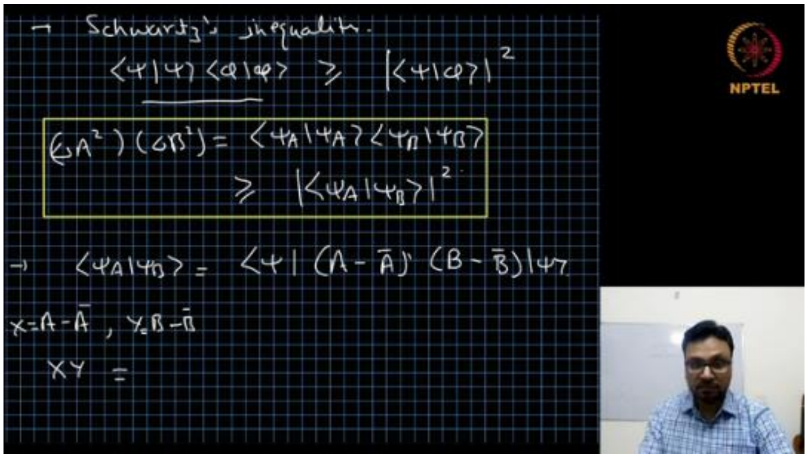


Every observable commute in classical world. In that way, the commutation is always zero. So, there is no lower bound on the uncertainties of the two observables. Just to give you an example, we have $\hat{\boldsymbol{x}}$, the position operator and the $\hat{\boldsymbol{p}}$, the momentum operator. From our quantum mechanics, we know the commutator of $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{p}}$, $[\widehat{\boldsymbol{x}}, \hat{\boldsymbol{p}}]$ is i$\hbar$. So, then our $\langle \Delta x^2 \rangle \, \langle \Delta p^2 \rangle \geq \hbar^2/4$. This is what it means.

So, it means in a quantum system, the uncertainty in $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{p}}$ cannot be product of those two, cannot be less than $\hbar^2/4$. Or, in other way, if we perform measurement in x, then we get very, very precise value without any uncertainty. Uncertainty tending to 0, that would mean that there will be infinite uncertainty in the momentum. So, the proof of this theorem goes as follows. Let us go back to $\langle \Delta A^2 \rangle$, that is $\langle \psi | (A - \langle A \rangle)^2 | \psi \rangle$.

From here, let us define $|\psi_A\rangle$, some vector which is not normalized as $(A - \langle A \rangle)|\psi\rangle$. We know it is not normalized but it is some vector and we will keep it like this. From here we can write $\langle \Delta A^2 \rangle$ as $\langle \psi_A | \psi_A \rangle$. Similarly, we can write $\langle \Delta B^2 \rangle$ as $\langle \psi_B | \psi_B \rangle$. It means that $\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle = \langle \psi_A | \psi_A \rangle \langle \psi_B | \psi_B \rangle$. Now, I would like to recall the recall an inequality and that is called Schwarz inequality, which says that if we have vectors $|\psi\rangle$ and $|\phi\rangle$, product of their norm will always be greater than or equal to $|\langle \psi | \phi \rangle|^2$. Let me repeat, if you have two vectors $|\psi\rangle$ and $|\phi\rangle$, then the product of their norm will always be greater than or equal to the mod square of their inner product.

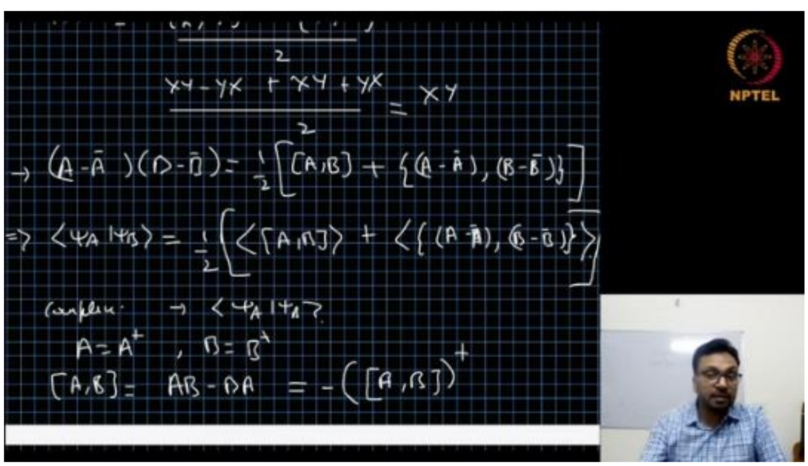
(Refer slide time: 18:20)



That we have proven in our mathematical supplementary lectures. So, we use this Schwarz inequality in the previous equation and we get $\langle \Delta A^2\rangle\langle \Delta B^2\rangle$, which is $\langle\psi_A|\psi_A\rangle\langle\psi_B|\psi_B\rangle$. This will be greater than or equal to $|\langle\psi_A|\psi_B\rangle|^2$. So, this is one statement. Now, consider $\langle\psi_A|\psi_B\rangle$, that will be $\langle\psi|(A-\bar{A})(B-\bar{B})|\psi\rangle$. Let me put $\bar{A}$ for the expectation value and $(B-\bar{B})$ so that it's easier to write.

Now we have two operators $(A-\bar{A})$ and $(B-\bar{B})$. So let us call them X and Y. So we are given X Y product, which we can write as ([X,Y]+{X,Y})/2, that we can check [X,Y] is XY-YX, plus {X,Y} is XY+YX. It means $(A-\bar{A})(B-\bar{B})$ can be written as half times commutator of $(A-\bar{A})$ with $(B-\bar{B})$. $\bar{A}$ and $\bar{B}$ are numbers, so we are left with only A and B plus anti-commutator of $(A-\bar{A})$ and $(B-\bar{B})$.

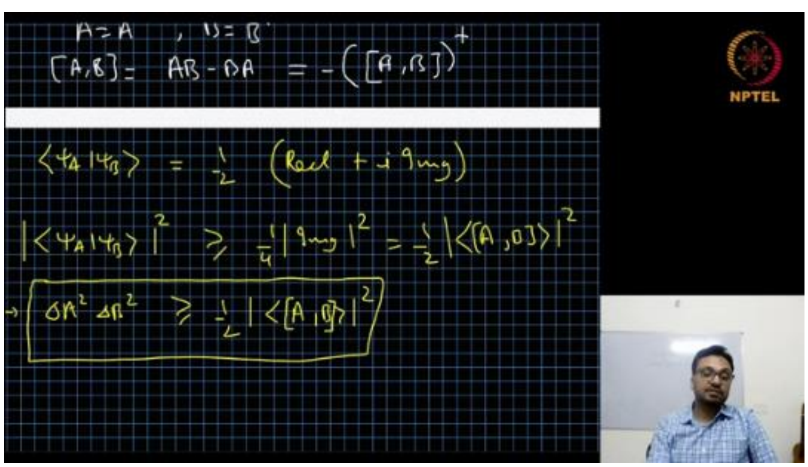
(Refer slide time: 22:14)



This implies that our $\langle\psi_A|\psi_B\rangle$ will be half expectation value of commutator of A, B plus expectation value of anti-commutator of A, B. Now, ψ_A and ψ_B is a complex number. A is Hermitian, so is B. It means commutator of A, B, which is AB-BA is $-([A,B])^{\dagger}$, you can check that the commutator if we take the dagger of the commutator, commutator is the operator but AB-BA is some matrix if A and B are some matrices and if A and B are both Hermitian, then the commutator is anti-Hermitian. Anti-Hermitian is nothing but i times Hermitian, but the important thing is the expectation value of the anti-commutator is always imaginary and the expectation value of A commutator anti commutator is always real.

So, we can write $\langle\psi_A|\psi_B\rangle$ as half times some real part plus i times imaginary part. So we can say $|\langle\psi_A|\psi_B\rangle|^2$ will always be greater than or equal to the imaginary part mod square, that is, ¼ $|\langle[A,B]\rangle|^2$and from our earlier equation we know that the product of the uncertainty $\langle\Delta A^2\rangle$ $\langle\Delta B^2\rangle$ is greater than $|\langle\psi_A|\psi_B\rangle|^2$which is greater than 1/2 $|\langle[A,B]\rangle|^2$ and this is the proof of the Heisenberg uncertainty principle which shows that if you have two observables which do not commute then the product of their uncertainty is lower bounded by their commutator, the expectation of their commutator. This is a technical statement. What it means is if you cannot measure two non-commuting observables simultaneously with perfection. So, that is what it says that if you have two non-commuting observables, you cannot measure them simultaneously with arbitrary accuracy.

(Refer slide time : 24:13)



One interesting thing which I like to point out here, in deriving the whole Heisenberg uncertainty principle, we have never used the collapse postulate of the quantum mechanics. We have only used the expectation value, we have used the commutation relation. The expectation values can be derived without invoking the collapse postulate, we just need the Born rule of probability.