Review of the Formalism of Quantum Mechanics

The "postulates of quantum mechanics" are often stated in textbooks. There are two main properties of "physics" upon which these postulates are based: 1)the **probability of interactions (measurements)** and 2) the **interference properties of the probability amplitudes**. This review is meant to help students understand the "physics" underlying the formalism of quantum mechanics.

Two classic experiments will guide us in understanding the required mathematical formalism of QM: Stern-Gerlach experiment and the double slit interference experiment. In the Stern-Gerlach experiment the outcomes are discrete, and one has continuous outcomes in the double slit experiment. We start with the Stern-Gerlach experiment.

Stern-Gerlach Experiment (Discrete Outcomes)

The experiment involves sending a neutral atom that has a magnetic moment into a non-uniform magnetic field. The simplest case is when the atom has a "spin" of 1/2. Let the axis of the Stern-Gerlach (SG) apparatus be in the direction of the gradient of the magnetic field. Then the force on the atom will be along the axis.

Experiment 1:
Send unpolarized atoms into an SG apparatus that has it's axis in the vertical direction (i.e. up). The result is that they are deflected up a certain amount or down the *same* amount. If they are allowed to impinge on a screen, one does not see a continuous smearing of the atoms, but rather **two distinct spots. The outcomes are discrete**. The discreteness of the outcomes was not expected, and is a remarkable result. If the atoms are truly unpolarized, then one can not determine what they will do. Half the time an atom will go up, and half the time the atom will go down. There is a true probabilty of 1/2 for each outcome. You might expect this, since the initial state was random. The randomness in this experiment is not astonishing, the discreteness is.

Experiment 2:
We now take the atoms that went up and send them though an identical SG apparatus. The result is that these atoms go up every time. Likewise, if we take the atoms that went down and send them though an identical SG apparatus, they will go down every time. We can label the atoms that will go up every time as being in the "up state". Likewise, we can label the atoms that will go down every time as being in the

"down state". Remember that "up" and "down" is in the direction of the axis of the SG setup.

Experiment 3:
We now take the atoms that went up through a vertical SG apparatus, and send them through an SG apparatus that has its axis tilted by an angle θ with the vertical in a plane perpendicular to the direction of the atoms. After passing through the tilted SG apparatus, the atoms will be deflected in a direction along the tilted axis. Whether they get deflected "up" or "down" the axis is random, with a probability equal to $cos^2(\theta/2)$ to be deflected "up". The probability to be deflected "down" the axis is $sin^2(\theta/2)$.

This is an interesting result. The atoms are initially not unpolarized, **they are in a definite state**. There is no randomness in the initial state, before they enter the tilted SG apparatus. However, the outcome is random. This is remarkable! The probability of outcomes from a definite state is demonstrated in every quantum experiment from decays to scattering experiments.

We note that if the atoms that went "up" through the tilted SG apparatus are sent though an identical tilted SG apparatus, they will always be deflected "up". The same will be true for the atoms deflected "down" through the tilted SG apparatus. Here "up" and "down" refer to along the tilted axis.

Double Slit Experiment (Continuous Outcomes)

In these experiments electrons (or photons) of a **definite energy** impinge on a screen with either one or two slits. After passing through the slit(s) the electrons (or photons) land on a wall.

Experiment 4:
Electrons impinge on a screen with one slit. Suppose the slit opening is in the "x" direction, and the electrons are traveling in the "y" direction. If the slit is small enough, then the electrons make a smeared out pattern when they hit the wall. The pattern is the same as a single slit diffraction pattern that waves would make. If the slit is made smaller, then the pattern is more spread out. The electrons arrive one at a time, but one cannot predict where a single electron will land. The outcome is probabilistic, with the probability proportional to the intensity of a single slit diffrac-

tion pattern.

Once an electron has arrived at the wall, then it's momentum in the x-direction, p_x, is determined. p_x is measured since the time to hit the wall is known (from p_y) as well as the distance in the x-direction from the slit. The further away the electron has traveled from the center of the slit center, the larger p_x is. Thus, the probability that an electron has momentum p_x is proportional to the intensity of a single slit diffraction pattern. The outcomes of this experiment are continuous, and the probability is a different mathematical quantity than in the case of discrete outcomes for which the probability is unitless. For continuous outcomes, the probability will be a **probability density**. The units of probability density is 1/(units of the continuous outcome).

We note that as the slit opening Δx gets smaller, then the probability distribution for p_x becomes more spread out. x and p_x are called congugate quantities. Every continuous outcome has a corresponding congugate quantity.

Experiment 5
In this experiment electrons impinge on a screen with two very thin slits that are closely spaced. If one slit is covered, the electrons arrive at the wall with a spread out probability distribution as in the last experiment. If the other slit is covered, electrons arrive at the wall with a spread out probability distribution that is shifted slightly from the previous one with the first slit covered.

Now when both slits are open, the electrons hit the wall one at a time randomly as before. However, the probability distribution is **NOT** equal to the normalized sum of the probability distributions of each slit separately! An interference effect occurs. The probability distribution is one of bands of high and low probability, similar to a two slit interference pattern for waves.

How can we understand the experimental fact that the electrons arrive at the wall one at a time, in a random manner with a probability distribution that exhibits "interference" properties? Adding the probability distributions from each slit does not agree with experiment. The interference pattern on the wall can be understood from the superposition of probability **amplitudes** as follows. The probability amplitude for the electron to land at a certain point on the screen is the sum of the probability amplitude for the electron to pass through the right slit plus the probability amplitude for the electron to pass through the left slit. The probability itself is the

absolute square of the total probability amplitude. For the double slip experiment the probability ampliude for passing though each slit is proportional to $e^{ipd/\hbar}$, where p is the momentum of the electron in the "x" direction and d is the total distance traveled from source to the wall. The total probability amplitude is:

$$Probamp \propto e^{ipd_1/\hbar} + e^{ipd_2/\hbar} \tag{1}$$

The probability (or probability density) is just the absolute square of the probability amplitude or:

$$\begin{aligned}
Prob &\propto |e^{ipd_1/\hbar} + e^{ipd_2/\hbar}|^2 && (2)\\
&\propto (e^{-ipd_1/\hbar} + e^{-ipd_2/\hbar})(e^{ipd_1/\hbar} + e^{ipd_2/\hbar}) && (3)\\
&\propto 2 + 2\,cos(\frac{p(d_2 - d_1)}{\hbar}) && (4)\\
&\propto cos^2(\frac{p\Delta d}{\hbar}) && (5)\\
Prob &\propto cos^2(\frac{\pi\Delta d}{\lambda}) && (6)
\end{aligned}$$

where $\lambda = \hbar/p_y$ is the DeBroglie wavelength of the electron, and $\Delta d = d_2 - d_1$ is the difference in the distance the particle takes going through the right versus the left slit. This formula agrees with the double slit interference experiment, as discussed in your sophmore level modern physics class. There is no chance for the particle to land at points on the screen where the path difference Δd is half a "wavelength", and maximum when Δd is an integral multiple of a "wavelength". Calling $\hbar/p_y$ a "wavelength" is somewhat misleading, since nothing is waving. However, we will use this term and remember that there is no waving going on. The distance $\hbar/p_y$ can be thought of as an interference-length.

Stern-Gerlach Revisited

Will one observe "interference" effects with the Stern-Gerlach apparatus? Yes. Consider the following experiment.

Experiment 6

Suppose we set up three successive Stern-Gerlach apparatus. The first and third (last) one have their axis pointing vertical, and the second (middle) apparatus with it's axis tilted 90° with respect to the other two. Now, send un-polarized atoms into the three Stern-Gerlach apparatus, which are in series. After passing through the first SG apparatus, stop all atoms that are forced downward. That is, let only "up" atoms continue through the other two SG setups.

After the second SG apparatus, half of the electrons will be forced "right" and half will go left. Stop those going "left". Then, after going though the final SG setup, half of those left will go "up" and half will go "down". So, of the "up" atoms entering the last two setups, 1/4 will go "up" and 1/4 will go "down". We lost half of the atoms since we blocked atoms going "left" in the middle SG setup.

Now, repeat the experiment, but this time block those going "right" in the middle SG setup. That is, let only those going left pass through onto the final SG setup. As before, after emerging from the last SG setup, 1/4 will be force "up" and 1/4 will go down. So, of the "up" atoms entering the last two setups, 1/4 will go "up" and 1/4 will go "down". We lost half of the atoms since we blocked atoms going "right" in the middle SG setup.

Finally, run the experiment but do not block any of the atoms in the middle setup. That is, let all the original "up" atoms go through both of the last two SG setups. What will happen? How many atoms will go "up" and/or down after passing though the last two SG setups? If one simply adds the probabilities of going right plus left through the middle SG setup, one would guess that half go "up" and half go "down": $1/4 + 1/4 = 1/2$. However, this is not what happens in the experiment. They ALL go "up" and none go "down". One has to add the probability amplitudes, and the amplitudes cancel for the atoms to go down. Note: a word of caution. The SG setups must be close enough so that if an atom goes up at the end, one cannot distinguish if it went "right" or "left" through the middle SG setup. That is, the atom must be able to arrive at the same spot by either going "right" or "left" through the middle SG setup. If this is the case, then the amplitudes will add.

Mathematical Formalism

What type of mathematical structure can we use to explain these and other experiments of modern physics. We need to include probability and "interference" in the formalism. The field of abstract linear algebra can be used to embody both of

these properties, as we try to demonstrate.

Throughout the discussion let A and B be two quantities that can be measured (observables). For simplicity, we start with measurements that yield values that are discrete, quantized. We will generalize later to measurements that yield continuous values. Label the possible values of observable A as α_i and, and of observable B as β_i. The integer i can take on a finite number or an infinite number of values.

Probability of Measurements

Our experiments indicate that one can not determine the exact outcome of a measurement before the measurement is carried out. The only information about the outcome of a measurement are the probabilities for the various outcomes. If a measurement of A is guaranteed to yield the result α_i, then we will call this state a "proper state" (eigenstate) of the observable A and label it as $|\alpha_i>$. For example in the SG setup, if the atom is in the "up" state, it will always be forced up after passing though a SG setup which has it's axis vertical. However, a system is not always in a proper state of a measurement that is guaranteed to give a particular result. For example, if a SG apparatus has its axis tilted at an angle θ from the vertical, and an atom passes through in the "vertical-up" state, then the atom will sometimes be forced at an angle θ and sometimes in the opposite direction. In general, a system will be in a state that has a certain "probability distribution" to be in the various α_i states. One can represent the state $|\Psi>$ of a system as a linear combination of the $|\alpha_i>$ states:

$$|\Psi>= a_1|\alpha_1> +a_2|\alpha_2> +... = \sum a_i|\alpha_i> \tag{7}$$

What is the meaning of the a_i'? One might be lead to believe that the a_i represent the probability that a measurement of A yields a result α_i. However, experiments also indicate that the probabilities are subject to interference effects. Thus, in order to have the probabilities include interference effects we are lead to interpret the a_i as a probability amplitude: The probability for obtaining a value of α_i if the observable A is measured is $|a_i|^2$.

Similar reasoning and notation can be used with any other observable. For example, for the observable B, we have

$$|\Psi>= b_1|\beta_1> +b_2|\beta_2> +... = \sum b_i|\beta_i> \tag{8}$$

where $|\Psi>$ represents the state of the system, and $|b_i|^2$ is the probability that a measurement of B will yield the value β_i. The above expansion is a result of the probabilistic nature of the outcome of a measurement.

For example, A could be a SG setup that has its axis aligned vertically and B could be a SG setup that has its axis aligned at an angle θ with the vertical. The state $|\alpha_+>$ could be the state that goes "up" through A and $|\alpha_->$ the state that goes "down" through A. The state $|\beta_+>$ could be the state that is forced in the θ direction after going through B and $|\beta_->$ the state that goes "away" from the θ direction.

Often in quantum mechanical experiments one measures the probability to obtain a value of β_i if B is measured when the system is in the state $|\alpha_j>$. We will label this probability as $|<\beta_i|\alpha_j>|^2$. We will discuss why this method of labeling is valuable in a moment. The quantity $<\beta_i|\alpha_j>$ is a probability amplitude, whose square is a probability. In terms of experiments, this would mean that the observable A is measured with a result $|\alpha_j>$. *Then* observable B is immediately measured. The probability of obtaining β_i when B is measured is $|<\beta_i|\alpha_j>|^2$.

Linear Vector Space of the States of a System

We would like to express the ideas above mathematically. Abstract linear algebra has the elements that we need. If observable A is measured, the outcome must be one of the α_i. Thus, in general one can express the state of the system in terms of the states $|\alpha_i>$. A linear combination of the proper states $|\alpha_i>$ is consistent with this property. If $|\Psi>$ is any general state of the system, then we expect that $|\Psi>$ can be expressed as

$$|\Psi>=\sum a_i|\alpha_i> \tag{9}$$

where the sum goes over all possible outcomes of A. The a_i are related to the probability that the measurement A will yield α_i. Since any linear combination of the $|\alpha_i>$ is a possible state of the system, **the set of states is a vector space**. The field of the vector space is the type of quantity that the a_i are. For the state space of quantum systems, the a_i must be complex (at least). Since any linear combination of states is also a state of the system, the set of all states of a system form a vector space over the field complex numbers.

For a set of objects to be an inner-product vector space, a scalar product must be defined. We write the scalar product between the state $|\Psi_1>$ and $|\Psi_2>$ as

$< \Psi_2|\Psi_1 >$. The scalar product must be an element of the field (i.e. a complex number). What is the physical meaning of the scalar product in the state space? To be consistent with the probability of outcomes, we want $< \alpha_i|\alpha_j >$ to be zero if $i \neq j$ and equal to one if $i = j$. Using this property, we see from the above equation that $a_i =< \alpha_i|\Psi >$, so $| < \alpha_i|\Psi > |^2$ can be interpretated as the probability to obtain α_i if the system is in the state $|\Psi >$ and A is measured. Thus, the important interpretation of the physics of the inner product is:

The inner product $< \Psi_2|\Psi_1 >$ is the probability amplitude that the system is in the state $|\Psi_2 >$ if it is in the state $|\Psi_1 >$.

This statement is one of the key ideas that connects the physics to our abstract linear algebra ansatz. Admitedly, the the statement is rather strange. It is like saying that the inner product of a cat and a dog is the probability amplitude that you have a cat if you have a dog. However, experiments show that if an atom is in the vertical "up" state, there is a certain probability amplitude that the atom is in a "right" state, etc.

We have discussed the main basic ideas of quantum pure states and measurements. We now expand on some of the properties of the vector state space.

Basis sets of the Vector Space

A basis for a vector space is a set of elements such that any state can be expressed as a linear combination of the basis elements. Since the vector space elements are states of a physical system, it is natural to use as a basis the states of an observable corresponding to a measurement outcome. For example, one could use the states of the observable A, $|\alpha_i >$, as a basis. This means that every state of the system can be expressed as combinations of states for all possible values of measurement A:

$$|\Psi >= \sum a_i|\alpha_i > \tag{10}$$

where the sum goes over all possible outcomes of A. This is a complete basis when it comes to measurements of A, since very possible outcome is included. However, this basis might be incomplete when it comes to all possible measurements and outcomes. Our physical interpretation of the scalar product allows a nice interpretation of the a_i's. The scalar product $< \alpha_i|\alpha_j >$ is the probability amplitude that a measurement

of A will result in a value α_j when the system is in the state $|\alpha_i>$. The probability must be zero unless $i = j$. If $i = j$, the probability is 1. Thus, we have:

$$<\alpha_i|\alpha_j>=\delta_{ij} \tag{11}$$

In the terms of linear algebra, the basis $|\alpha_i>$ is an orthonormal basis. If we take the scalar product of the previous equation with $|\alpha_j>$, one obtains

$$<\alpha_j|\Psi>=\sum a_i<\alpha_j|\alpha_i>=a_j \tag{12}$$

Thus, a_j is interpreted as the probability amplitude that the state $|\Psi>$ will be found in the state $\alpha_j>$, and $|a_j|^2$ is the probability that a measurement will yield the value α_j.

One could use the proper states of any observable as a basis. For example, one could also use the proper states of the observable B: $|\beta_i>$ which correspond to values of β_i if B is measured.

$$|\Psi>=\sum b_i|\beta_i> \tag{13}$$

For these basis states the following is also true: $<\beta_j|\beta_i>=\delta_{ij}$, $b_j=<\beta_j|\Psi>$, with $|b_j|^2$ being the probability that a measurement B on the system yields β_j.

If one knows all the probability amplitudes $<\beta_i|\alpha_j>$ for all i and j, then all measurement probabilities of the two observables A and B can be determined.

The "energy" basis

Very important states for non-relativistic applications are the states of definate values of the system's energy. We label these states as $|E_i>$, corresponding to an energy of the system of E_i. For bound systems, the energy is discrete (quantized). In practice, if the energy gets large enough the particles become unbound and there is a continuous range of possible energies. In order to use only sums, lets just consider systems with only discrete values of allowed energy (i.e. the harmonic oscillator or infinite square well). We will generalize to continuous spectra later. "Proper states" of definate energy are important, because their time evolution is simple. If the system is in the state $|E_k>$ at time $t = t_0$, then for times $t > t_0$ the state of the system is:

$$|\Psi>_{time\ t>t_0}=e^{-i\frac{E_k(t-t_0)}{\hbar}}|E_k> \tag{14}$$

If the scalar products $< E_i|\alpha_j >$ are known for all i and j, then the time evolution of the measurement probabilities for observable A can be determined. This can be seen as follows. Suppose at time $t = 0$ the system is in the state $|\Psi >_0$ with probability amplitudes a_i for the observable A.

$$|\Psi >_0 = \sum_i a_i |\alpha_i > \tag{15}$$

The states $|\alpha_i >$ can be expressed in terms of the E_j as $|\alpha_i >= \sum_j < E_j|\alpha_i > |E_j >$. Substuting into the equation above one obtains:

$$|\Psi >_0 = \sum_{ij} a_i < E_j|\alpha_i > |E_j > \tag{16}$$

This is just a long sum with the only "proper energy" states of the system on the right side of the equation. The time evolution of the states of definate energy are simple: multiply by the factor $e^{-iE_j t/\hbar}$. Thus, the state of the system for times greater than $t = 0$ is

$$|\Psi >_{t>0} = \sum_{ij} a_i e^{-i\frac{E_j t}{\hbar}} < E_j|\alpha_i > |E_j > \tag{17}$$

The probability amplitude for obtaining the value α_k if A is measured is just $< \alpha_k|\Psi >_{t>0}$ or $< \beta_k|\Psi >_{t<0}$ for β_k if B is measured. The same mathematics can be used for any other observable (e.g. B). The scalar products between all the "proper states" of the observables (e.g. $< \alpha_i|E_j >$, $< \beta_i|E_j >$, and $< \alpha_i|\beta_j >$ contains all the "physics" of the system. In words, this means that **the probability amplitudes between all outcomes of all possible observables contains all the "physics"**.

Since the scalar product and the "proper states" are so important we should review a few of their properties.

Completeness

If every state can be expessed as a linear combination of the basis states, the basis is said to be complete. That is, if every state $|\Psi >$ can be expressed as

$$|\Psi >= \sum_{all\ i} a_i |\alpha_i > \tag{18}$$

for the basis $|\alpha_i >$, then the basis $|\alpha_i >$ is a complete basis and "spans" the state space. If $< \alpha_i|\alpha_j >= \delta ij$, then the basis is ortho-normal.

If the basis is ortho-normal, the a_j can be determined by taking the inner product with $< \alpha_j|$ on both sides of the above equation.

$$\begin{aligned} < \alpha_j|\Psi > &= \sum_{all\ i} a_i < \alpha_j|\alpha_i > \\ &= \sum_{all\ i} a_i \delta_{ij} \\ < \alpha_j|\Psi > &= a_j \\ a_i &= < \alpha_i|\Psi > \end{aligned}$$

replacing j with i.

If we substitute the expression into the equation above we have

$$\begin{aligned} |\Psi > &= \sum_{all\ i} a_i|\alpha_i > \\ |\Psi > &= \sum_{all\ i} |\alpha_i > a_i \\ |\Psi > &= \sum_{all\ i} |\alpha_i >< \alpha_i|\Psi > \\ |\Psi > &= (\sum_{all\ i} |\alpha_i >< \alpha_i|)\Psi > \\ 1 &= \sum_{all\ i} |\alpha_i >< \alpha_i| \end{aligned}$$

This unusual way to expess the number one is called the completeness relation, since it embodies the completeness property of an orthonormal basis.

The completeness relation helps one understand how interference effects enter in the "vector space formalism". For example, suppose the state space is three dimensional, and the system starts out in the state $|\alpha_1 >$. Suppose that measurement B is done, then measurement A.

The final state is the found by summing the amplitudes of the possible paths, then squaring the sum:

$$|\Psi_{final} > = \sum_{i=1}^{3} < \beta_i|\alpha_1 > |\beta_i >$$

$$
\begin{aligned}
&= (\sum_{i=1}^{3} |\beta_i><\beta_i|)|\alpha_1> \\
|\Psi_{final}> &= |\alpha_1>
\end{aligned}
$$

the final state is the same as the initial state, because the sum was over all the possible states of the measurement B. Physically, this means that the measurement B did not affect the state. There was an equal chance to go through any of the states of B. In our SG example, this is what occured when we did not observe which way (right or left) the atom went.

Now, suppose that we block the passage of state $|\beta_1>$, and do not observe which state $|\beta_2>$ or $|\beta_3>$ is the intermediate state. Then, the final state will be

$$
\begin{aligned}
|\Psi_{final}> &= \sum_{i=2}^{3} <\beta_i|\alpha_1>|\beta_i> \\
&= <\beta_2|\alpha_1>|\beta_2> + <\beta_3|\alpha_1>|\beta_3>
\end{aligned}
$$

Now, the probability amplitude for the final state to be $|\alpha_2>$ (say) is

$$
<\alpha_2|\Psi_{final}> = <\beta_2|\alpha_1><\alpha_2|\beta_2> + <\beta_3|\alpha_1><\alpha_2|\beta_3> \tag{19}
$$

which, in general, is the sum of two complex numbers. The complex numbers can have cancellations, and hence interference properties. The probability is the absolute square of this sum.

Unitary Transformations

Since the scalar product corresponds to a probability amplitude, its magnitude must be less than or equal to one: $|<\alpha_i|\beta_j>|^2 \leq 1$. Also, between any scalar product one can insert a sum over a complete basis: $<\Psi_1|\Psi_2> = \sum_i <\Psi_1|\alpha_i><\alpha_i|\Psi_2>$. If $|\Psi_1>$ and $|\Psi_2>$ are proper states of the B measurement, we have:

$$
<\beta_i|\beta_j> = \sum_k <\beta_i|\alpha_k><\alpha_k|\beta_j> \tag{20}
$$

Since $|\beta_i>$ and $|\beta_j>$ are orthonormal, this gives:

$$
\sum_k <\beta_i|\alpha_k><\alpha_k|\beta_j> = \delta_{ij} \tag{21}
$$

Thus, the matrix $< \alpha_i|\beta_j >$ is a unitary matrix. This requirement puts constraints on the transition probabilities and guarantees the probabilities add up to one. The unitary matrix $< \alpha_i|\beta_j >$ is the transformation matrix from the B basis to the A basis. Unitary matrices can be associated with symmetries of the system, and play an important role in fundamental physics.

Operators

It is not always useful to express the "physics" in terms of the various probability amplitudes $< \beta_j|\alpha_i >$. Operators can be defined on a linear vector space, and are important quantities. An operator "operating" on a vector produces another vector in the space: $|\Psi_2 >= \hat{O}|\Psi_1 >$ where $\hat{O}$ represents an operator, $|\Psi_1 >$ is an element in the vector space, and so is $|\Psi_2 >$. Some operators will not have any physical significance, however there are many that will. In particular, it is possible to associate an operator with each physical observable. Such an operator can be defined in the following way:

Consider the observable A. The operator associated with A in the $|\alpha_i >$ basis is

$$A_{ij} \equiv \alpha_i \delta_{ij} \tag{22}$$

To find the representation of the operator A in the basis of the proper states of B, one only needs to transform via the unitary transformation $< \beta_i|\alpha_j >$. In the $|\beta_i >$ basis we have:

$$A_{ij} = \sum_k \alpha_k < \beta_i|\alpha_k >< \alpha_k|\beta_j > \tag{23}$$

since $\delta_{ij} =< \beta_i|\beta_j >= \sum_k < \beta_i|\alpha_k >< \alpha_k|\beta_j >$ from completeness. If we want to write down the operator A in a basis-independent way, we put a carrot over the symbol: $\hat{A}$.

A linear operator on a vector space is basis independent. Once one picks a basis, say $|\gamma_i >$, then the matrix representation of the operator **in terms of the basis** $|\gamma_i >$ is $A_{ij} =< \gamma_i|(\hat{A}|\gamma_j >)$. If the operator is associated with an observable as defined above, then the eigenvalues of the operator are the possible outcomes of the measurement and the "proper states" are the eigenstates corresponding to the eigenvalue. Often the "physics" of the system is best expressed in terms of the properties

of the operators corresponding to the systems observables.

In general terms, from the definition of the operator associated with an observable A, we have $\hat{A}|\alpha_i >= \alpha_i|\alpha_i >$ and $< \alpha_i|\hat{A}|\alpha_j >= \alpha_i\delta_{ij}$.

Postulates of Quantum Mechanics

Most textbooks (ours included) on quantum mechanics for undergraduate seniors introduce the mathematical formalism of quantum mechanics with a chapter on "The Postulates of Quantum Mechanics". The texts then proceed to apply these postulates to quantum experiments to justify their validity. The postulates, which we have tried to motivate with our discussion, are usually stated as follows:

1. The states of a system form a complete inner product vector space, which is a Hilbert Space.

2. Observables are represented by hermitian operators on the Hilbert Space.

3. If a quantity is measured (or observed), the only possible outcomes are the eigenvalues of the hermitian operator corresponding to the observable.

4. The inner product of two (normalized) states $|\psi_1 >$ and $|\psi_2 >$, $< \psi_2|\psi_1 >$, is the probability amplitude that the system is in the state $|\psi_2 >$ if it is in the state $|\psi_1 >$.

That any part of the universe would behave according to the above postulates seems to defy "common sense". However, our sense of what the laws of nature should be are formed from our limited physical senses, which are blind to the atomic and subatomic world. As far as we know, the above postulates are the common sense of modern physics. To understand the universe at the atomic and smaller scale, quantum mechanical logic is required.

Continuous Spectra

We have limited the discussion to discrete spectra. The generalization to continuous spectra can be done using the Dirac Delta function. As an example, we will consider the position observable. The measurement values are continuous, which we label as x. The eigenstate that has the value of x when position is measured will be labeled as $|x >$, and the position operator as $\hat{x}$. That is $\hat{x}|x >= x|x >$. The orthonormal property of the eigenstates is

$$< x|x' >= \delta(x - x') \tag{24}$$

The "completeness relation" $\sum_{all\, i} |\alpha_i >< \alpha| = 1$ becomes:

$$\int_{-\infty}^{+\infty} |x >< x|\, dx = 1 \tag{25}$$

We note that the scalar product $< x|\Psi >$ is the position probability amplitude *density* that the particle will be found at the position x. This is the function we have been labeling as $\Psi(x)$. So, $\Psi(x) \equiv < x|\Psi >$.

Another important observable that has a continuous spectra is momentum. The operator is labeled as $\hat{p}$, with eigenstates $|p >$ for eigenvalues p: $\hat{p}|p >= p|p >$. The orthonormality property is $< p|p' >= \delta(p-p')$. The momentum probability amplitude density $\varphi(p)$ is $< p|\Psi >$. The completeness relationship for the momentum eigenstates is similar to that of the position eigenstates:

$$\int_{-\infty}^{+\infty} |p >< p|\, dp = 1 \tag{26}$$

Some important scalar products

Since position and momentum are often measured, the scalar product between the state of a system $|\psi >$ and an eigenstate of definate position x, $|x >$, and definate momentum p, $|p >$ is an important quantity. The quantity $< x|\psi >$ is a function which represents the probability amplitude to "find" the particle at the position x if the position is measured. Since x is a continuous variable, $< x|\psi >$ is a probability amplitude density. It is often given the symbol $\psi(x)$:

$$\psi(x) \equiv < x|\psi > \tag{27}$$

and is the function that is the solution to the Schroedinger equation in coordinate space.

The quantity $< p|\psi >$ is a function of p which represents the probability amplitude density to obtain a value of p if the momentum of a particle is measured when in the state $|\psi >$. It is often given the symbol $\phi(p)$:

$$\phi(p) \equiv < p|\psi > \tag{28}$$

Some important unitary transformations and scalar products

A very important unitary transformation (inner product) is between the position and momentum eigenstates: $< x|p >$. Interference experiments indicate that $< x|p > \propto e^{ipx/\hbar}$. Using Dirac normalization, we have:

$$< x|p >= \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} \tag{29}$$

From this inner product, one can transform from the position probability amplitude density function, $< x|\Psi >$, to the momentum probability amplitude density function, $< p|\Psi >$:

$$< p|\Psi >= \int_{-\infty}^{+\infty} < p|x >< x|\Psi > dx \tag{30}$$

giving

$$\varphi(p) = \int_{-\infty}^{+\infty} \Psi(x) \frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}} dx \tag{31}$$

Using the inner product $< x|p >$, one can also determine the form of the momentum operator in coordinate space as follows:

$$\begin{aligned} < x|\hat{p}|\psi > &= \int < x|p >< p|\hat{p}|\psi > dp & (32) \\ &= \int \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} p < p|\psi > dp & (33) \\ &= -i\hbar \frac{d}{dx} \int \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} < p|\psi > dp & (34) \end{aligned}$$

which reduces to

$$< x|\hat{p}|\psi >= -i\hbar \frac{d}{dx} \int < x|p >< p|\psi > dp \tag{35}$$

or

$$< x|\hat{p}|\psi >= -i\hbar \frac{d}{dx} < x|\psi > \tag{36}$$

For the position operator in coordinate space we have:

$$< x|\hat{x}|\psi >= x < x|\psi > \tag{37}$$

From the above equations, one can derive the commutator: $[\hat{x}, \hat{p}]$. First $< x|\hat{x}\hat{p}|\psi >$:

$$\begin{aligned} < x|\hat{x}\hat{p}|\psi > &= x < x|\hat{p}|\psi > & (38) \\ &= x(-i\hbar\frac{d}{dx}) < x|\psi > & (39) \end{aligned}$$

Next $< x|\hat{p}\hat{x}|\psi >$:

$$\begin{aligned} < x|\hat{p}\hat{x}|\psi > &= \int < x|p >< p|\hat{p}\hat{x}|\psi > dp & (40) \\ &= \int < x|p > p < p|\hat{x}|\psi > dp & (41) \\ &= \int \frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} p < p|\hat{x}|\psi > dp & (42) \\ &= -i\hbar\frac{d}{dx}\int < x|p >< p|\hat{x}|\psi > dp & (43) \\ &= -i\hbar\frac{d}{dx}(< x|\hat{x}|\psi >) & (44) \end{aligned}$$

Operating with $\hat{x}$ on the left gives:

$$< x|\hat{p}\hat{x}|\psi >= -i\hbar\frac{d}{dx}(x < x|\psi >) \quad (45)$$

Using the product rule for the derivative gives:

$$< x|\hat{p}\hat{x}|\psi >= -i\hbar < x|\psi > +x(-i\hbar\frac{d}{dx}) < x|\psi > \quad (46)$$

Subtracting $\hat{p}\hat{x}$ from $\hat{x}\hat{p}$ gives:

$$< x|\hat{x}\hat{p} - \hat{p}\hat{x}|\psi >= i\hbar < x|\psi > \quad (47)$$

Since this expression is true for all states $|\psi >$,

$$\hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \quad (48)$$

This important operator equation derives from $< x|p >= \frac{e^{ipx/\hbar}}{2\pi\hbar}$.

Double Slit Experiment Revisited

In our previous discussion of the double slit experiment, the pattern on the wall was understood as interference between the amplitude for the electron to pass through one slit plus the amplitude to pass through the other. How can the "interference" pattern on the wall be understood with the linear algebra formalism?

As follows: *The slits on the screen correspond to the "x" probability distribution of the electron just after it passes through the screen. The pattern on the wall is proportional to the x-component of the electron's momentum after it passes through the screen.*

That is, the state of the electron in the position basis, $< x|\Psi >= \psi(x)$, has equal amplitudes at the location of the slits, say $x = 0$ and $x = d$. Or, $\psi(x) \propto (\delta(x) + \delta(x - d))$. After the electron passes through the slits, the more momentum it has in the x-direction then the further away it lands on the wall. So the probability for the electron to land at a particular distance from the center is proportional to the momentum distribution of the electron, $| < p_x|\Psi > |^2 = |\phi(p_x)|^2$.

As discussed, the inner product $< x|p >$ contains all the information for transforming from the position basis to the momentum basis.

$$\begin{aligned} < p_x|\Psi > &= \int_{-\infty}^{+\infty} < p|x >< x|\Psi > dx \\ \phi(p_x) &\propto \int_{-\infty}^{+\infty} e^{ip_x x/\hbar}\psi(x)dx \\ &\propto \int_{-\infty}^{+\infty} e^{ip_x x/\hbar}(\delta(x) + \delta(x - d))dx \\ &\propto 1 + e^{ip_x d/\hbar} \\ &\propto e^{ip_x d/(2\hbar)}(e^{-ip_x d/(2\hbar)} + e^{ip_x d/(2\hbar)}) \\ &\propto e^{ip_x d/(2\hbar)} cos(p_x d/(2\hbar)) \end{aligned}$$

Taking the absolute square, one obtains the probability distribution for the electrons momentum in the x-direction, the direction in which the slits are separated.

$$|\phi(p_x)|^2 \propto cos^2(p_x d/(2\hbar)) \tag{49}$$

Thus, the probability distribution on the wall after the electron has passed through two very narrow slits will be a cosine squared function as observed experimentally. This equation is the same as we derived earlier using the interference of amplitudes

from each slit, as follows. The sin of the scattering angle θ equals p_x/p_y for small angles. Here, p_y is the electrons momentum in a direction towards the slits and the wall. So the difference in distance Δd as defined before equals $\Delta d \approx d\ sin(\theta) \approx dp_x/p_y$. Thus, $dp_x \approx \Delta dp_y$. With this substitution we have

$$\begin{aligned} |\psi(p_x)|^2 &\propto cos^2(\Delta dp_y/(2\hbar)) \\ &\propto cos^2(\frac{\pi \Delta dp_y}{h}) \\ &\propto cos^2(\frac{\pi \Delta d}{\lambda}) \end{aligned}$$

since the DeBroglie wavelength λ equals h/p. This is the same result that we obtained before using "interference".