

Quantum Mechanics Chapter 1.5: An illustration using measurements of particle spin.

I Introduction.

Quantum mechanics is a theory of physics that has been very successful in explaining and predicting many physical phenomenon, particularly those that occur on a molecular, atomic, and subatomic scale. Its philosophical interpretation has been the subject of debate.

In order to understand the theory of quantum mechanics we will explore its application to the physical property of spin. Spin is an observable quantity of a particle (such as an electron or atom). Although it is not strictly correct, one can think of spin in terms of the particle spinning around its axis like the earth rotates (spins) around its axis. Thus measuring spin would be equivalent to measuring how fast, and about what axis (vertical, horizontal or other) the particle is spinning.

We use the observable Spin because the mathematics involved is relatively simple. Nevertheless, the theory of quantum mechanics applies to Spin the same way as it does to any other observable such as the position or velocity of the particle.

II. Mathematical Background.

A. A 2x2 matrix.

A matrix is an ordered array of numbers. Here we need to introduce the concept of a 2x2 matrix. Here is an example:

$$\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}. \tag{1}$$

Notice the following:

- 1) The matrix consists of 4 numbers.
- 2) The numbers are arranged in a specific manner.

We say that the numbers 3 and 7 are in column 1. Likewise, 2 and 5 are in column 2. The numbers 3 and 2 are in row 1 and 7 and 5 are in row 2. Thus the number 3 is in row 1 column 1 and 5 is in row 2 column 2.

B. A 2 x 1 vector. A 2 x1 vector is like a matrix with two rows and 1 column. Here is an example:

$$\begin{bmatrix} 9 \\ 4 \end{bmatrix}. \tag{2}$$

Here we have two numbers that are arranged in a single column.

C. Matrix Operations.

1. Multiplication by a scalar.

A scalar is just a number. When multiplying a matrix (or a vector) by a scalar, the scalar just multiplies each number in the matrix.

For example if we multiply the matrix in Equation 1 above by the scalar (number) 4 we have

$$4 \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} (4)(3) & (4)(2) \\ (4)(7) & (4)(5) \end{bmatrix} = \begin{bmatrix} 12 & 8 \\ 28 & 20 \end{bmatrix}. \quad (3)$$

Notice that multiplying a matrix by a scalar yields another matrix. Here's what happens when we multiply the vector in Equation 2 by the scalar 4:

$$4 \begin{bmatrix} 9 \\ 4 \end{bmatrix} = \begin{bmatrix} (4)(9) \\ (4)(4) \end{bmatrix} = \begin{bmatrix} 36 \\ 16 \end{bmatrix}. \quad (4)$$

Thus if we multiply a vector by a scalar we get another vector.

Please complete the following exercises

$$\begin{aligned} 7 \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix} &= \\ 9 \begin{bmatrix} 0 & 1 \\ 2 & 7 \end{bmatrix} &= \\ 6 \begin{bmatrix} 1 \\ 3 \end{bmatrix} &= \\ 7 \begin{bmatrix} -2 \\ 3 \end{bmatrix} &= \end{aligned} \quad (5)$$

2. Operation of a 2x2 matrix on a vector.

A 2x2 matrix can operate on a (2x1) vector to produce another vector. We say that the 2x2 matrix is an operator. The operation is illustrated below.

$$\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} 9 \\ 4 \end{bmatrix} = \begin{bmatrix} (3)(9) + (2)(4) \\ (7)(9) + (5)(4) \end{bmatrix} = \begin{bmatrix} 27 + 8 \\ 63 + 20 \end{bmatrix} = \begin{bmatrix} 35 \\ 83 \end{bmatrix}. \quad (6)$$

The element in the first row and first column of the resultant vector (63) comes from multiplying the first row of the matrix with the first (and only) column of the vector. The element in the second row and first column of the resultant vector (83) comes from multiplying the second row of the matrix with the first (and only) column of the vector. Here is another example:

$$\begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} (2)(-1) + (3)(2) \\ (3)(-1) + (0)(2) \end{bmatrix} = \begin{bmatrix} -2 + 6 \\ -3 + 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}. \quad (7)$$

Please complete the following exercises:

$$\begin{aligned} \begin{bmatrix} -2 & 1 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} &= \\ \begin{bmatrix} 0 & 1 \\ 2 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} &= \\ \begin{bmatrix} 5 & 3 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} &= \end{aligned} \quad .(8)$$

3. Multiplication of two vectors.

Two vectors can be multiplied using the dot or scalar product to produce a scalar (number). For example,

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \end{bmatrix} = (1)(4) + (2)(3) = 4 + 6 = 10. \quad (9)$$

Here we have changed the column vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ to row vector $\begin{bmatrix} 1 & 2 \end{bmatrix}$. Whenever we multiply two vectors we always change the first one to a row vector. Then the multiplication is the same as for matrices. Let

$$A = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, B = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, C = \begin{bmatrix} 0 \\ 2 \end{bmatrix}. \quad (10)$$

Find the scalar product of AB, BA, AC, BC .

The norm of a vector is its dot product with itself. If a vector is complex, One takes the complex conjugate of the first, transposed vector to calculate the scalar product, so if

$$D = \begin{bmatrix} i \\ 2 \end{bmatrix}, \quad (11)$$

Find AD, DA , so that $AD = (DA)^*$

D. Probability.

When you toss a quarter there is a 50% chance that it will land on "heads" and a 50 % chance it will land on "tails". We can say that the probability that it will land on heads is 0.5 and the probability that it will land on tails is also 0.5. This means that if we toss a lot of quarters that half of them will land on heads and the other half will land on tails. Thus we can predict with some certainty that if we make 1000 tosses 500 will land on heads and 500 will land on tails. However, we can not predict the outcome of a single toss.

The probability that the coin will land on either heads or tails during one toss is the sum of these two probabilities $(0.5 + 0.5) = 1$. A probability of 1 means that it is certain to occur. The probability that we will get two heads in a row is $(0.5)(0.5) = 0.25$. That is there is a 25% chance of this occurring.

The probability that a die that you throw will land with 6 is 1/6. If you made 1000 tosses 166 to 167 of them would most likely land on 6. The higher the number of tosses you make the more predictable will be the outcome. That is, if you make only one toss you know that there is 1/6 chance of getting 6 but can only make a weak prediction. If you make 100 tosses it is likely that you will get 16 to 17 that land on 6 but you might get only 14 or 15.

E. The eigenvalues and eigenvectors of a matrix.

We can refer to a 2x2 matrix as an operator in the sense that it can operate on a vector and give a new vector.

$$\hat{O}V_1 = V_2. \tag{12}$$

An operator is said to be Hermitian iff

$$\hat{O}^\dagger = \hat{O}, \tag{13}$$

where the dagger refers to the adjoint which is the transpose conjugate.

For all the matrices we will use in examining Spin, there are associated two eigenvectors and two eigenvalues. An eigenvector is a vector (like in Equation 2) and an eigenvalue is a number or scalar. The eigenvectors and eigenvalues obey the following relation:

$$(\text{Matrix})(\text{eigenvector}) = (\text{eigenvalue})(\text{eigenvector})$$

Let's define three matrices,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{14}$$

The eigenvectors associated with σ_x are $\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$ with the eigenvalue 1, and $\begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$

with the eigenvalue -1.

The eigenvectors associated with σ_z are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ with the eigenvalue 1, and $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ with the

eigenvalue -1.

Let's verify that $\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$ is an eigenvector of σ_x with the eigenvalue 1:

We need to see if

$$\begin{aligned}
& \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \stackrel{?}{=} 1 \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \\
& \rightarrow \begin{bmatrix} (0)(\sqrt{2}/2) + (1)(\sqrt{2}/2) \\ (1)(\sqrt{2}/2) + (0)(\sqrt{2}/2) \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} (1)(\sqrt{2}/2) \\ (1)(\sqrt{2}/2) \end{bmatrix} \\
& \rightarrow \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}
\end{aligned} \tag{15}$$

Let's also verify that $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$ is an eigenvector of σ_z with an eigenvalue of -1:

$$\begin{aligned}
& \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} \stackrel{?}{=} -1 \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\
& \rightarrow \begin{bmatrix} (1)(0) + (0)(-1) \\ (0)(0) + (-1)(-1) \end{bmatrix} \stackrel{?}{=} \begin{bmatrix} (-1)(0) \\ (-1)(-1) \end{bmatrix} \\
& \rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\end{aligned} \tag{16}$$

As an exercise, show that $\begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$ is an eigenvector of σ_x with the eigenvalue -1, and that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of σ_z with an eigenvalue of 1:

One can use methods of linear algebra to calculate the eigenvectors and eigen values of a matrix. One usually chooses the normalized eigenvectors. Also, we require that the eigenvectors be orthogonal*. For hermitian operators, eigenvalues are obtained from the equation:

$$\text{Det}(\hat{O} - \omega \hat{I}) = 0 \tag{17}$$

Once these have been found, the orthonormal eigenvectors can be determined from the eigenvalue equation.

Exercise: Show that the above eigenvectors and eigenvalues are correct by using equation 17.

* All eigenvectors are only uniquely determined to within a factor of $e^{i\phi}$, that is to within a phase factor.

III. The postulates of Quantum Mechanics.

- (1) The state of a system (that is what we know about it) is completely described by its state vector, which is generally represented by ψ . When we are dealing only with spin, the state vector will be a 2 by 1 vector such as those above (Equation 2).[□]
- (2) Every observable (a measurable quantity) is represented by an operator. For spin, the operators are 2x2 matrices like in Equation 1, We will generally represent an operator by \hat{S} .
- (3) The only possible outcomes of a measurement of an observable are the eigenvalues of the operator associated with the observable. [We will describe what an eigenvalue is below along with a description of eigenvectors]. We will represent the eigenvalues as ω .
- (4) The probability that a measurement will result in a value of ω is $|\Psi\Omega|^2$, where Ω is the eigenvector associated with ω .
- (5) After a measurement resulting in ω the state vector becomes (collapses) into the eigenvector Ω .
- (6) The state vector evolves in time in a specific way. [For our purposes in this chapter, we can ignore this postulate because the state vectors describing spin will be constant in time. In general, the state vector obeys an equation called the time-dependent Schrödinger equation.]

III. Spin.

A. The operators.

The two observables we will consider are the spin of a certain type of particle measured in the x-direction and the spin measured in the z-direction. You can think of these observables as an extent to which a particle is spinning around an axis oriented up and down (z) and oriented in a horizontal direction (x). The operators associated with these observables are

$$\hat{S}_x = \frac{\hbar}{2}\sigma_x = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \hat{S}_z = \frac{\hbar}{2}\sigma_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \hat{S}_y = \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \quad (18)$$

Here \hbar is planck's constant, $h = 6.63 \times 10^{-27}$ erg-sec divided by 2π .

B. The eigenvectors and eigenvalues.

[□] A state vector, as well as an eigenvector, is usually written so that its magnitude is one. That is the first term squared plus the second squared is one. This is necessary so that the total probability of getting some value from a measurement be one.

The eigenvectors for \hat{S}_x are the same as those for σ_x but now the eigenvalues are $+\hbar/2$ for $\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$ and $-\hbar/2$ for $\begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$.

The eigenvectors for \hat{S}_z are the same as those for σ_z but now the eigenvalues are $+\hbar/2$ for $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and $-\hbar/2$ for $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$.

C. Direct interpretation of the postulates for the observable spin.

(1) The first postulate says that everything we know about a system will be given by its state vector. In this case, for spin, that is a vector such as:

$$\Psi = \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}. \tag{19}$$

(2) The two operators are already given above in equation 18. The eigenvectors and eigenvalues of these operators give us the power to make predictions about the outcome of a measurement.

(3) The only possible values of a spin measurement are the eigenvalues $+\hbar/2$ and $-\hbar/2$. That means that if we were to measure the spin of a particle along a given axis, we could only get either one of these values. This is like saying that if we were to measure how high we can throw a ball, we would only get certain values like 10 feet or one hundred feet, but not any value in between. This is what is meant by quantization. The possible values of spin are not any value between $+\hbar/2$ and $-\hbar/2$ but only these exact values. This is completely different than anything classical mechanics predicts.

(4) Let's take as our state vector that given in equation 19. As stated in postulate (1) all our knowledge of the system is contained in the state vector. Let's calculate the probability that if we make a measurement of the spin of the particle along the x-axis that we will get the value of $+\hbar/2$. This probability is given by the product of the state vector and the \hat{S}_x eigenvector associated with the eigenvalue $+\hbar/2$. Thus we have that the probability of obtaining a value of $+\hbar/2$ for the spin along the x-axis is

$$\begin{aligned} & \left([0.6 \quad 0.8] \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \right)^2 \\ &= ((0.6)(\sqrt{2}/2) + (0.8)(\sqrt{2}/2))^2 \\ &= ((0.6)(0.707) + (0.8)(0.707))^2, \\ &= (0.424 + 0.566)^2 \\ &= (0.990)^2 \\ &= 0.980 \end{aligned} \tag{20}$$

where I have rounded off at the third decimal point.

Likewise the probability of obtaining a value of $-\hbar/2$ for the spin along the x-axis is

$$\begin{aligned}
& ([0.6 \quad 0.8] \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix})^2 \\
& = ((0.6)(\sqrt{2}/2) + (0.8)(-\sqrt{2}/2))^2 \\
& = ((0.6)(0.707) + (0.8)(-0.707))^2 \quad (21) \\
& = (0.424 - 0.566)^2 \\
& = (-0.142)^2 \\
& = 0.020
\end{aligned}$$

Taken together, this means that if a particle begins in a state given by equation 16, we will have a 2% chance of measuring $-\hbar/2$ for its spin in the x-direction and a 98% chance of measuring $+\hbar/2$. If we had many particles, all prepared in the same state given by equation 16, then, upon measuring the spin along x, 98% of them would have $+\hbar/2$ and 2% would have $-\hbar/2$.

These results are totally different than what one gets from classical mechanics. Notice that we can only predict the probabilities of the outcome of a given measurement. In classical mechanics, if we know enough about the system, we can predict the exact outcome of every measurement. Here, in Quantum Mechanics, there is an inherent indeterminism in measurement.

As an exercise, calculate the probabilities that we will measure $+\hbar/2$ and $-\hbar/2$ for the spin along the z-axis, if a particle starts in a state described by the state vector of equation 19.

(5) Let's say that we make a measurement of the spin along the x-axis as we did above and get the result $+\hbar/2$. According to postulate 5, our state vector now becomes the

eigenvector associated with the operator \hat{S}_x and the eigenvalue $+\hbar/2$: $\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$. This

result is independent of the state of the system before the measurement. That is the system could have been in the state described in equation 19, $\psi = \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix}$, or some

other state but after measuring $+\hbar/2$ for the spin along the x-axis the state vector collapses into $\begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$.

This result is totally different than anything encountered in Classical Mechanics. Classically, the act of measuring something, does not alter the thing you are measuring. Here, in Quantum Mechanics, the state of the system is drastically and fundamentally altered.

IV. Quantum Weirdness.

A. Experimental measurements of Spin: The Stern-Gerlach device

The device used to measure Spin in the real world is called a Stern-Gerlach device. When a particle is injected into the device moving horizontally, say in the y-direction, and the device is oriented in the z-direction (vertically) then the particle will be deflected vertically as illustrated in Figure 1 below. A stream of particles is shown incident on the device.

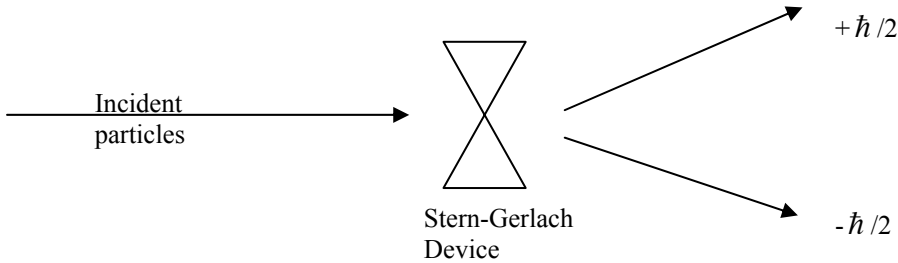


Figure 1

The particles are deflected either up or down a certain amount, but not in between. This is because there are only two values of the Spin, $+\hbar/2$ and $-\hbar/2$. If a screen were detected behind the device the particles would show up at only two positions. Classical mechanics predicted that they would show up as all possible positions.

B. A series of measurements

Let us now see what would happen if we make a series of three measurements using Stern-Gerlach devices. Let us first measure the Spin along the x-direction, then (selecting those with Spin, $x = -\hbar/2$) make a measurement along the z-direction, and finally (selecting the particles with Spin $z = +\hbar/2$), make a measurement another measurement in the x-direction. This series of measurements is illustrated in the Figure 2.

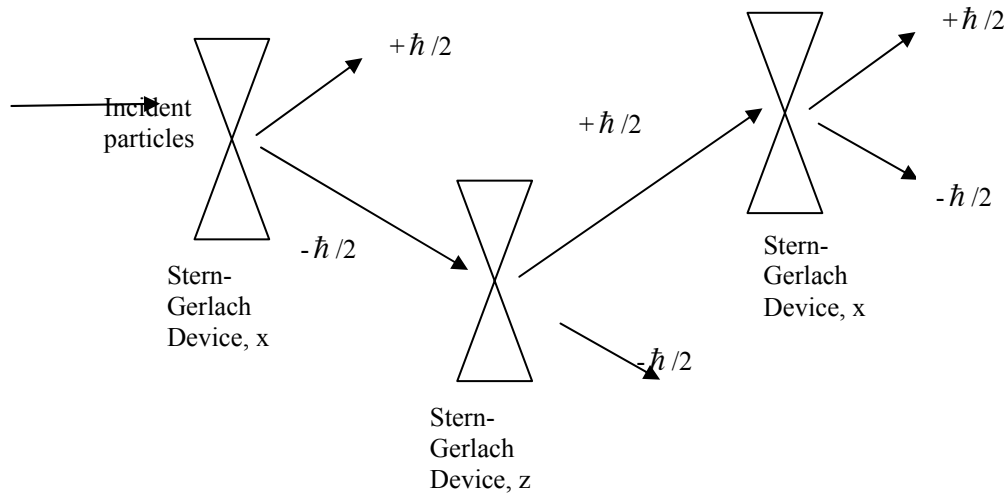


Figure 2

One might predict (in accordance with a classical view) that since we have already selected particles that have Spin in the x-direction of $-\hbar/2$ after the first device, that all of the particles coming out of the third device will also have $-\hbar/2$ Spin in the x-direction. Classically, if we measure a property of something, say its color or its Spin along x, then it possesses this property. Furthermore, classical mechanics holds that the act of measuring it or another property, such as Spin along z, does not change its properties. As shown here, Quantum Mechanics predicts that only half of the particles coming out of the third Stern-Gerlach device have Spin $-\hbar/2$ along the x-direction.

After the first Stern Gerlach device, the state vector of the particles with Spin $= -\hbar/2$ along x is the eigenvector associated with this eigenvalue : $\begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$. Now let's calculate the percentage of these particles that will have a Spin along z $= +\hbar/2$. That fraction is:

$$\begin{aligned} & \left([1 \ 0] \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} \right)^2 \\ &= ((1)(\sqrt{2}/2) + (0)(-\sqrt{2}/2))^2 \\ &= ((1)(0.707) + (0)(-0.707))^2 \\ &= (0.707)^2 \\ &= 0.5 \end{aligned} \tag{22}$$

So half the particles that enter the second Stern-Gerlach device also go into the third one. Now let's calculate the quantum mechanical prediction of what fraction entering the third device will have Spin $= -\hbar/2$. Remember classically that fraction should be 1 (or 100%). According to Quantum mechanics the state vector of the particles entering the third Stern-Gerlach device is the eigenvector associated with a value of $+\hbar/2$ along z: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

According to Quantum Mechanics the fraction of particles leaving the third device with Spin $-\hbar/2$ along x is

$$\begin{aligned} & \left([1 \ 0] \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix} \right)^2 \\ &= ((1)(\sqrt{2}/2) + (0)(-\sqrt{2}/2))^2 \\ &= ((1)(0.707) + (0)(-0.707))^2 \\ &= (0.707)^2 \\ &= 0.5 \end{aligned} \tag{23}$$

Thus the Quantum Mechanical prediction is different than the classical one. This is largely because the measurement of Spin by the second device alters the system we are measuring.

Exercise: Repeat this exercise with a sequence of Stern-Gerlach devices as follows: z, x, z, finding the particles that have first up then down then up spin.

IV.5 Heisenberg Uncertainty Principle (HUP)

The HUP says that there is a minimum ontological indeterminacy in conjugate attributes. For example $\Delta X \Delta P \geq \hbar/2$, where X and P are the position and momentum. The delta refers to the indeterminacy. This relation says, for example, that if we have no uncertainty in the position ($\Delta X = 0$) then the indeterminacy of the momentum (and hence velocity) is infinite. Thus, if we know exactly where something is at some moment, we will have no idea where it will be afterwards.

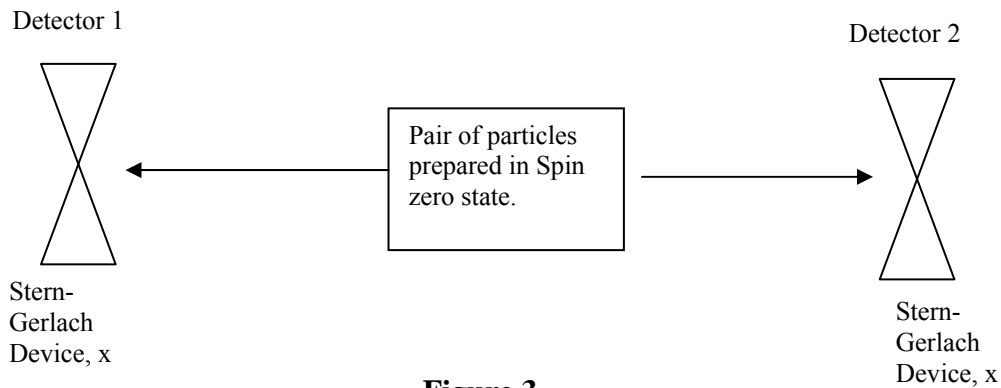
Similarly, we cannot know both the spin in the x-direction and in the y-direction. This is illustrated in our example of Figure 2. After the first detector the spin in the x-direction, for the particles that are deflected downwards is, $-\hbar/2$. At this point we can say that $\Delta \hat{S}_x$ is zero, so we *know* what the spin in the x-direction is. Now if we were to measure \hat{S}_z , quantum mechanics tells us that the best we can do is predict the probability of getting up or down and that this is the same (0.5/0.5). This is what we saw above. So if we *know* the spin in the x-direction, we cannot know its value in the z-direction.

V. The EPR Paradox

A. Conservation of angular momentum and Spin.

In order to describe the important argument made by Einstein, Podolsky and Rosen (Phys. Rev. **47**, 777, 1935) we need to expand our description of Spin to include that of a system of two particles. Classically, and quantum mechanically, the total angular momentum or Spin of a system, along any direction, is conserved. The angular momentum is a measure of how much matter is spinning around a certain axis at a certain speed. This is also, how we might think of Spin.

If a system starts out with a total Spin of zero, then the total spin of the system will be zero for all time afterwards. Thus if we prepare two of our particles (that can have Spin = $+\hbar/2$ or Spin = $-\hbar/2$) in a state that starts with total Spin of 0, then when one has Spin $+\hbar/2$ in a certain direction, the other will have Spin $-\hbar/2$ along that direction (so the total Spin is $(+\hbar/2) + (-\hbar/2) = 0$). The situation of interest is shown below in Figure 3.



Spin is conserved in each direction. Whenever a pair of particles are produced and one is incident on Detector 1 (oriented along x) and the other is incident at Detector 2 (also oriented along x) then the two detectors must measure opposite spins. In order to simplify our discussion, from now on when a particle has a Spin of $+\hbar/2$ we will simply say its Spin is $+1$ or that it is up. When a particle has a Spin of $-\hbar/2$ we will say that it has a Spin of -1 or that it is down.

If the two detectors are not oriented along the same direction then the sum of the detected Spins is not necessarily equal to zero. That is, if Detector 1 is oriented along x and Detector 2 along z then the sum of the detected Spins may be other than zero. In fact, in this case, the probability that the two spins will be the same is $1/2$ and the probability that they are different is also $1/2$. So, for example, if Detector 1 measures Spin $+1$ (up) along x then there is a 50/50 chance that Detector 2 will measure either $+1$ or -1 if it is oriented along z . Thus if we measure Spin up at Detector 1 along x , Quantum Mechanics cannot predict with certainty the outcome of a measurement along z . Quantum Mechanics can only predict the probability that it will be up or down (in this case the probability for each is $1/2$).

B. EPR's argument

Einstein, Podolsky and Rosen devised a thought experiment to show that Quantum Mechanics is an incomplete description of reality. The argument presented here is similar to that made by EPR but is based on the situation depicted in Figure 3.

EPR defined a theory to be complete if and only if it includes all real quantities. A quantity was defined to be real if it could be predicted with certainty, without disturbing the system.

If Detector 1 is oriented along x and we measure the Spin to be up then we can predict (with certainty) that the Spin along x at Detector 2 is down. Thus, the Spin at Detector 2 along x can be said to be real quantity. EPR went on to argue that we *could have* made a measurement along z in Detector 1 along the z -direction, if we liked. If we had done so, we could have predicted (with certainty and without disturbing the system) the outcome

of the measurement of Spin along the z-direction for the particle at Detector 2. Thus, EPR argued, we can predict with certainty the value of the Spin of the particle incident upon Detector 2 along both x and z directions simultaneously. However, according to Quantum Mechanics, if we know the spin along x then we can only predict a probability of the outcome of a measurement of the Spin of the particle along z. So Quantum Mechanics, according to the Copenhagen Interpretation, says that we cannot simultaneously know (with certainty) both the Spin in the x-direction and the Spin in the z-direction. So, according to EPR, Quantum Mechanics does not include the two real quantities Spin along x and z simultaneously, so quantum Mechanics is incomplete.

C. The Copenhagen interpretation

The realistic view of EPR was contrary to the widely held Copenhagen interpretation of Quantum Mechanics, promoted largely by Niels Bohr. Bohr argued that although the statement "I could have measured ..." makes no sense. For him properties of systems are only defined with respect to their measurement devices.

VI Bell's Inequality

A. Generalization of Spin measurements on a pair of particles.

The choice of axes is arbitrary. We can define x and z however we like, as long as they are perpendicular to each other. Furthermore, we need not orient our Stern-Gerlach device along either of these axes. Consider this generalization depicted in Figure 4.

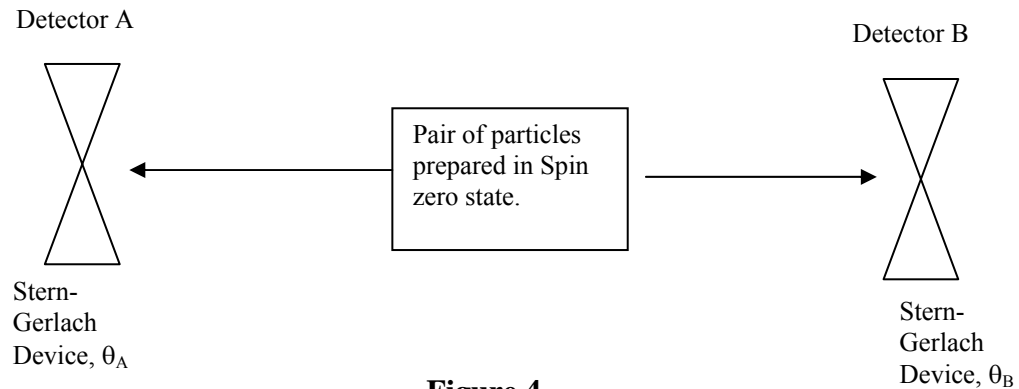


Figure 4

We now depict the two detectors by Detector A (oriented along θ_A) and Detector B (oriented along θ_B). Quantum Mechanics predicts that the probability that the Spin measured at A is equal that at B as

$$P(A=B) = \frac{1}{2}[1 - \cos(\theta_A - \theta_B)]. \quad (24)$$

Notice that if $\theta_A = \theta_B$ then $P(A=B) = (1 - \cos(0))/2 = 0$, which says that when both detectors are oriented along the same direction they never give the same result. When one is up the other cannot also be up. If θ_A is 0° and θ_B is 180° then $P(A=B) = (1 - \cos(-180))/2 = 1$. When we say that the detector is oriented at 180° we have flipped the detector so Spins that were measured up are now down and those that were down are now up. Thus our result for θ_A is 0° and θ_B is 180° giving $P(A=B) = 1$ means that when one is up the other is down. If θ_A is 90° and θ_B is 0° then $P(A=B) = (1 - \cos(90))/2 = 0.5$. This says that when the detectors are oriented perpendicular to each other there is a 50/50 chance that they will both have the same value.

The probability that the Spin measured at A is **not** equal that at B is just

$$\begin{aligned} P(A \neq B) &= 1 - P(A=B) \\ &= 1 - \frac{1}{2}[1 - \cos(\theta_A - \theta_B)] \\ &= \frac{1}{2}[1 + \cos(\theta_A - \theta_B)] \end{aligned} \quad (25)$$

To ease the notation in following discussions, let us define $AB = P(A \neq B)$.

B. Background

The completeness of quantum mechanics was debated a lot after EPR's paradox was proposed. Some physicists claimed that the indeterminism of quantum mechanics was an indeterminism in our knowledge of reality. If only we knew more about the system then we could predict everything with certainty. Some of these physicists developed other theories that made the same predictions as quantum mechanics but added a "hidden" variable that determined the outcome the measurement. These ad hoc hidden variable theories were consistent with a realistic metaphysics. They added nothing new in the way of prediction (since the added variable was hidden) but were attractive to many who wanted to preserve a realistic view.

In 1965 John S. Bell proved that any local, realistic theory is inconsistent with the predictions of quantum mechanics. Locality refers to the inability of any cause or information to be able to travel faster than the speed of light. Einstein's theory of special relativity shows that if a causal element can go faster than the speed of light, cause and effect as we know it is undermined. A realistic theory is one that assumes that quantities exist even when we are not observing them. Thus, in EPR's thought experiment, a realistic theory would hold that each particle has Spin both in the x and z direction. That quantum mechanics cannot predict both of these quantities shows that it is incomplete. If only we knew more about the system, the value of some hidden variable, then we would be able to predict the Spin in both directions. Bell showed that such a train of thought, coupled with locality, is inconsistent with quantum mechanics.

C. An example of Bell's Inequality

Consider the experiment depicted in Figure 4. In addition to making measurements at Detector A at angle θ_A , let us be able to make a measurement at Detector A at another angle $\theta_{A'}$. We will denote this measurement as A'. Likewise let us be able to make another measurement Detector B at a new angle $\theta_{B'}$ that we will denote as B'. Thus we can make two measurements at Detector A, denoted by A and A', that differ in how we have oriented our detector. We can also make two measurements at detector B, denoted by B and B', that differ in how we have oriented detector B. Note that we cannot make any two measurements at one detector at the same time, the detector can only have one orientation at any time. In a realistic model, however, we assume that a particle possesses Spin in both directions at once so we can talk about these properties belonging to the particle.

We have, according to quantum mechanics

$$\begin{aligned}
 AB &= P(A \neq B) = \frac{1}{2}[1 + \cos(\theta_A - \theta_B)] \\
 A'B &= P(A' \neq B) = \frac{1}{2}[1 + \cos(\theta_{A'} - \theta_B)] \\
 AB' &= P(A \neq B') = \frac{1}{2}[1 + \cos(\theta_A - \theta_{B'})] \\
 A'B' &= P(A' \neq B') = \frac{1}{2}[1 + \cos(\theta_{A'} - \theta_{B'})]
 \end{aligned} \tag{26}$$

Assuming the reality of the Spin of each particle along each direction we will now show that

$$AB' + A'B + A'B' \geq AB \tag{27}$$

To arrive at equation 27 we will treat each measurement pair as a circle whose area represents its probability. For example let us consider AB, which is the probability that the spin measured at Detector A at angle θ_A is not equal to that measured at Detector B oriented along θ_B . This is illustrated in Figure 5, below.

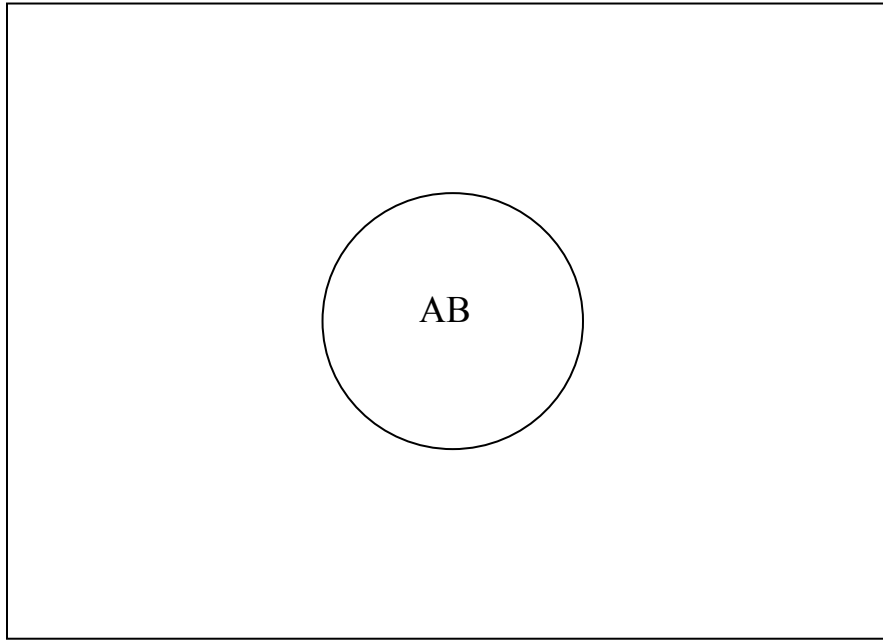


Figure 5.

The area of the square is taken to be 1. The area inside the circle represents $P(A \neq B)$. Thus the area outside of the circle but inside the square is $1 - P(A \neq B)$ which is $P(A = B)$, since either $A = B$ or $A \neq B$. The sum of these possibilities must be 1.

Let's illustrate this with an analogous situation. Let say that instead of measuring Spin at two positions (A and B) we have two people (at stations A and B) flipping pennies. The result AB corresponds to the situation when person A does not get the same result as person B (that is when A has heads, B has tails or when A has tails, B has heads). This analogy is not perfect in the sense that in quantum mechanics there are defined relations between the results at A and B that are not applicable to the flipping of pennies, but to illustrate the use of diagrams like Figure 5, the analogy is adequate. To figure out the probability of AB we can figure out the fraction of possibilities that correspond to this result. All together we have four possibilities: HH, HT, TH, TT, where the first value is that at A and the second one is that at B. Then $AB = 2/4 = 0.5$. This is illustrated in Figure 6.

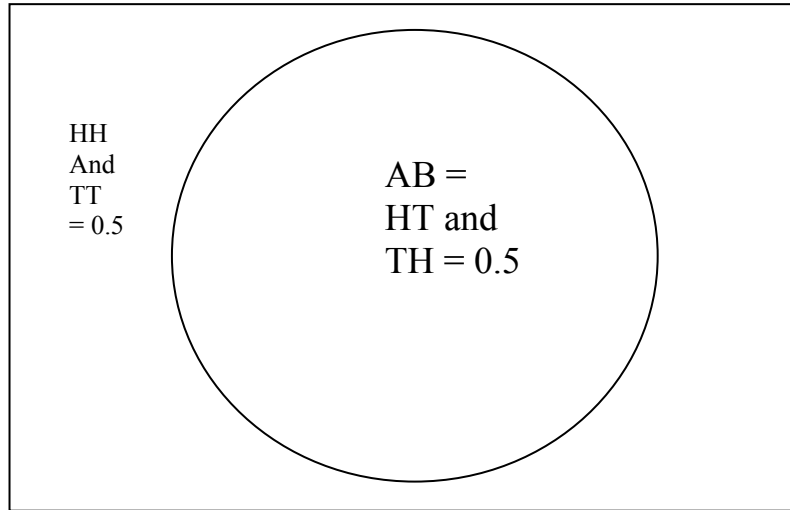


Figure 6.

Now let us consider the probabilities of the three measurements AB' , $A'B$ and $A'B'$, depicted in Figure 7.

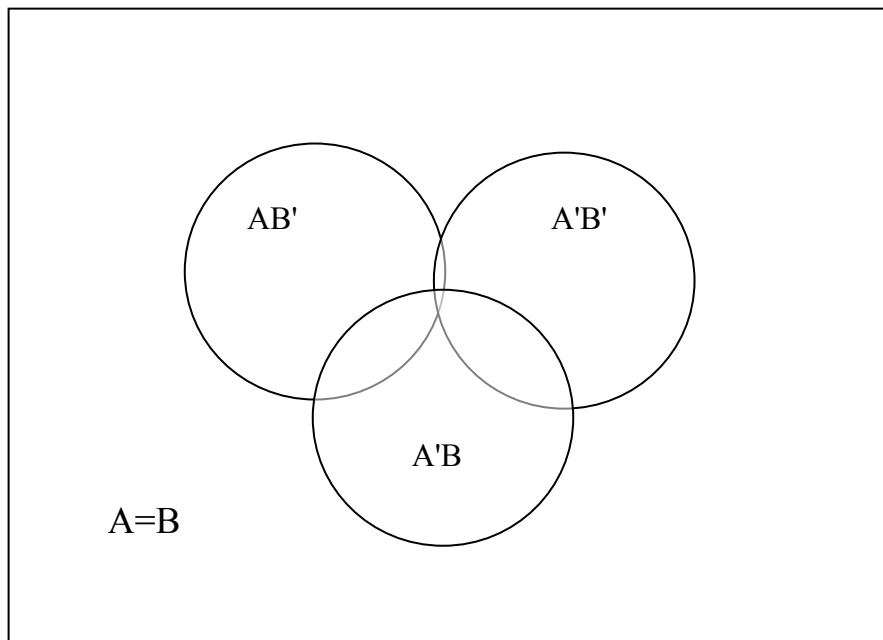


Figure 7

When $A' \neq B$ and $A \neq B'$ then the circles representing these overlap. We are considering a general case here, where the intersections of the outcomes these three sets of measurements may overlap or not.

For our penny analogy, let's introduce the prime measurements (A' and B') as when, instead of using pennies, we use nickels. Then the A' measurement is when a nickel is flipped at position A and a B' measurement is when a nickel is flipped at position B. AB' is when the person at A gets a different value for his or her penny than the person at B gets for his or her nickel (TH or HT). An example of overlap for the measurements $A'B$

and $A' B'$ would be when you have $H_n T_p$ and $H_n T_n$, where the subscripts denote whether a penny or nickel was flipped.

The area outside the circle for AB' is the probability that $A = B'$, that outside $A'B$ is $P(A'=B)$, and that outside $A'B'$ is $P(A'=B')$. Therefore, the area that is completely outside all 3 circles satisfies $A = B'$, $A' = B'$, and $A' = B$. Thus, in that area outside all three circles, $A=B$. Then the area inside the area covered by the three circles satisfies $1 - P(A=B) = P(A \neq B) = AB$. Thus we get Equation 27. If the three circles do not intersect at all then $AB' + A'B + A'B' = AB$. If they do intersect some, $AB' + A'B + A'B' > AB$. Now we show that Equation 27 can be inconsistent with Quantum Mechanics. Let $\theta_A = 90^\circ$, $\theta_B = 90^\circ$, $\theta_{B'} = 210^\circ$, $\theta_{A'} = 330^\circ$ (with all these angles being measured counterclockwise from the horizontal direction). Then Let $\theta_A - \theta_B = 0^\circ$, $\theta_{A'} - \theta_B = 240^\circ$, $\theta_A - \theta_{B'} = -120^\circ$, and $\theta_{A'} - \theta_{B'} = 120^\circ$. Then, from equation 26,

$$\begin{aligned}
 AB &= \frac{1}{2}[1 + \cos(0)] = 1 \\
 A'B &= \frac{1}{2}[1 + \cos(240)] = 0.25 \\
 AB' &= \frac{1}{2}[1 + \cos(-120)] = 0.25 \\
 A'B' &= \frac{1}{2}[1 + \cos(120)] = 0.25
 \end{aligned} \tag{28}$$

Now putting these into the relation obtained from our realistic model (Equation 24), we get

$$0.25 + 0.25 + 0.25 \geq 1 \tag{29}$$