Introduction to Quantum Mechanics

S9528

Grace H. Zhang • Summer 2015 • MIT HSSP

Last Revision: July 26, 2015

Table of Contents

Invitation

What's special about quantum mechanics? It describes the world that is too small for us to see. The particles of interest on this size scale include atoms on the large end and electrons, which we will be focusing on for most of this course, on the small end. We cannot perceive the motion of these particles with our eyes; we may be able to detect the presence of electrons experimentally using certain equipments, but we cannot observe them like we can do with macroscopic objects in our everyday lives. Therefore, we will use mathematical abstractions to understand them. As physicists, we are not new to mathematical abstractions. In classical mechanics (CM), we use point charges to study electrostatics, and center of mass to study motion, etc., but the mathematical tools we develop for quantum mechanics (QM) will turn out to be very different.

What are some properties a particle can have?

- location
- velocity
- charge
- mass
- ...

Let's look at location in space.

In CM, an object can be uniquely specified by its location in space at a certain time. Can an electron be uniquely specified by its location?

No. Electrons have an extra property that is completely quantum mechanical: spin. We can (somewhat naively) imagine spin as an arrow $\uparrow$ point in a certain direction. This classical picture will help us visualize spin and draw analogies to concepts that we are familiar with. This property must be specified in addition to location. To study the property of spin, we can forget about the electron that it is attached to. That is, we will be studying spin as its own system.

States v. Measurement

Before we start studying spin, we will need to distinguish between two concepts and their different meanings in CM and QM - states and measurements.

Classical Case: Our system consists of a coin that could either be heads up or tails up. We label the state of the coin (the quality describing which side is facing up) as σ. σ could be assigned two values: 1 if the coin is heads up, or 0 if the coin is tails up. We can "measure" σ by simply looking at the coin. If we see heads up, then we have measured σ to be 1. If we see tails up, then we have measured σ to be 0. In CM, the state and the measurement are the same thing, identically represented by σ. The relationship is so trivial that we don't think about it.

Quantum Case: Our system consists of a spin $\uparrow$. We label the state of spin as σ. To measure this quality however, we cannot simply look at it (too small!), we need an apparatus that interacts with the spin and record the value of σ for us. Let's call this apparatus A. The apparatus A has an orientation, as indicated by the arrow in the following figures. The coordinate space that our spin system and apparatus are situated in is indicated by the axes on the left of the figures. We will make a series of measurements to figure out the properties of spin. In the following experiments, your friend prepares a spin state $\uparrow$, and you will be measuring the spin with apparatus A.

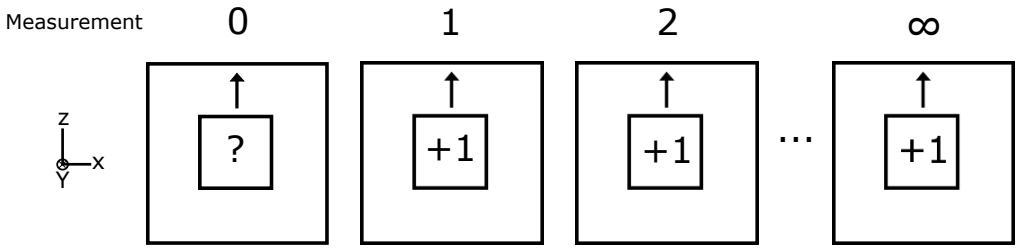

Figure 0.1: Experiment 1

Your apparatus is pointing in the same direction as the spin. You measure $\sigma = +1$. After the first measurement, you measure again and get $+1$ a second time. You keep measuring, without doing anything to change the system or the apparatus A in between the measurements except reset A for the next measurement. You get $+1$ every single time.

This situation is the same as that of a classical system, eg. if you see a coin facing up, and you do not do anything to the coin, you will see that it is still facing up every time you look at it after that.

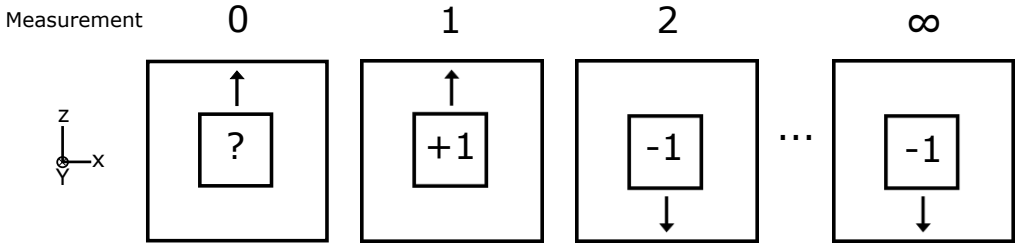

Figure 0.2: Experiment 2

You make the first measurement like you did in the first experiment, with A pointing up. Likewise, you measure $\sigma = +1$. However, after the first measurement, you flip A upside down and measure $\sigma = -1$. We start to wonder whether spin σ may be associated with a sense of direction in space. We think this because reversing A, thus reversing its direction relative to spin $\uparrow$ negates the value of our measurement of σ.

From this, we hypothesize that σ is a unit vector that gives the component of the spin along the direction A is orientated.

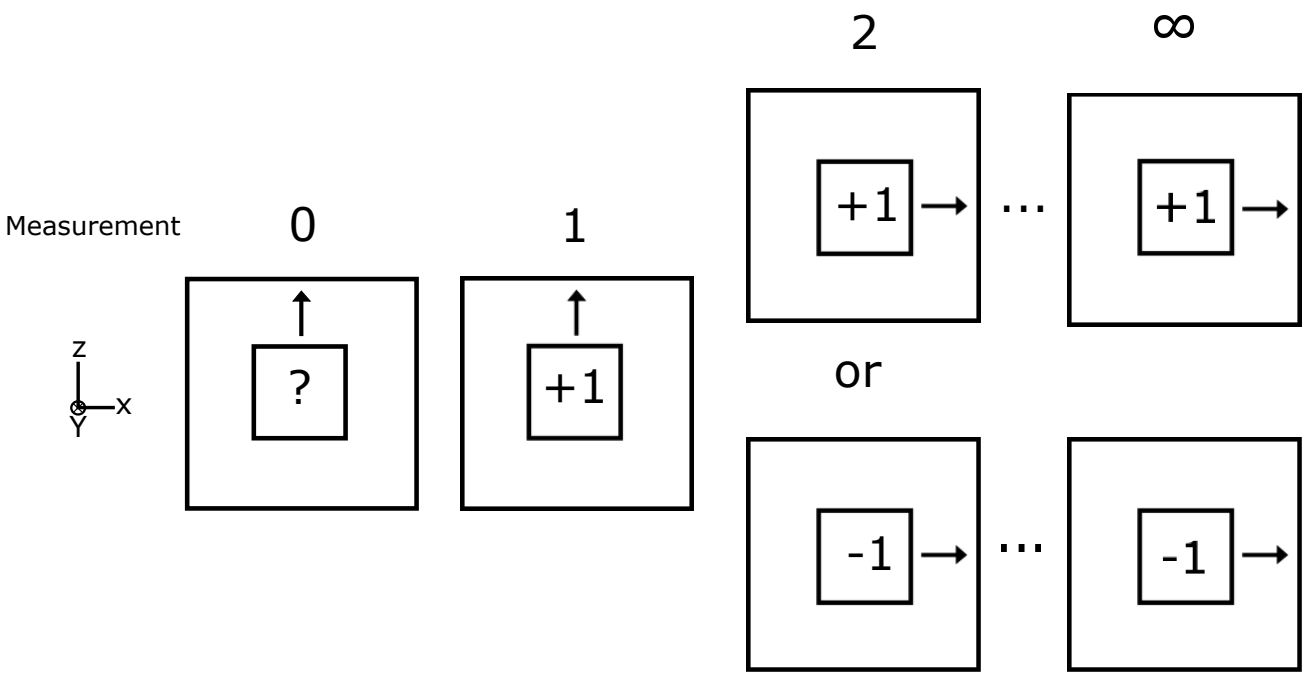

Figure 0.3: Experiment 3

You make the first measurement as in Experiment 1 and Experiment 2 and measure $\sigma = +1$ again.

After the first measurement, you rotate A by 90^o clockwise. According to your hypothesis from Experiment 2, you expect that σ should measure to be 0 because the spin $\uparrow$ is now perpendicular the the direction A is pointing $\rightarrow$. However, your second measurement gives $\sigma = +1$.

Unsatisfied and perplexed, you repeat the experiment from step 0 with the exact same procedure (rotating A after the first measurement by 90^o). You find that your second measurement (with A pointing right) gives either $+1$ **or** -1, but never 0. However, after repeating this experiment many times, you find the the number of times you measure $\sigma = +1$ is approximately equal to the number of times you measure $\sigma = -1$. In other words, the average value you measure for σ is 0, which coincides with your hypothesis based on a classical vector!

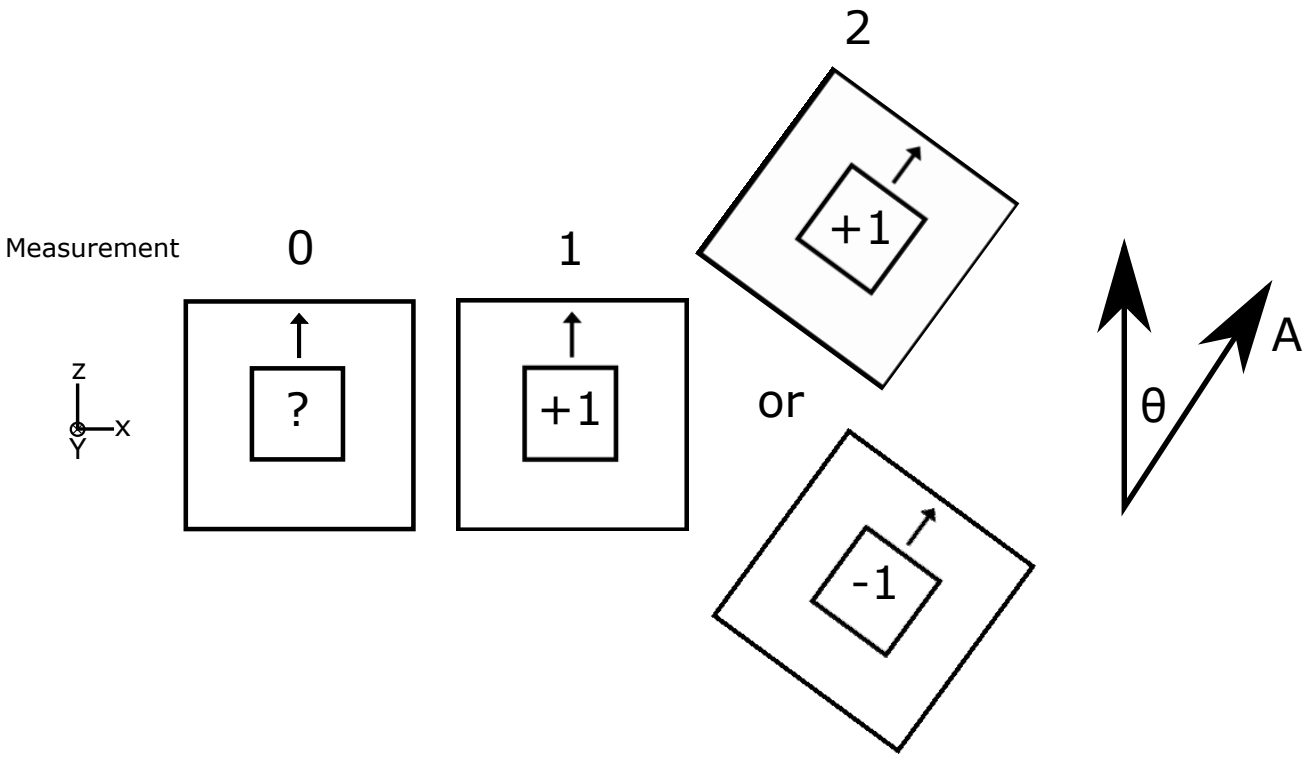

Figure 0.4: Experiment 4

You first measure with A pointed up as you have in the first step of all previous experiments.

Then, instead of rotating A by 90^o as you did in Experiment 3, you rotate A by an arbitrary angle θ. Now, the component of spin $\uparrow$ pointing in the direction of A should be $\cos\theta$. Like all previous experiments, however, A only outputs +1 or -1 and never any value in between. After repeating the experiment many times, always rotating A by the same angle θ after the first measurement, you get that the average value you measure for σ is equal to $\cos\theta$, which is the component of the up vector along the direction A is pointing in.

Properties realized about our Quantum Mechanical Spin system so far:

- System is not deterministic; We cannot predict what σ will measure to be with A pointing any arbitrary direction by measuring σ with A pointing in any other direction. In a classical system, the system is deterministic eg., for a truly classical unit vector, if we measure its component along z (up) to be 1, then we would deduce with complete certainty that its component along x (right) to be 0. This is obviously not true with our quantum mechanical spin.
- The average of our measurements of σ after many repetitions equal the classical result based on our classical vector hypothesis.

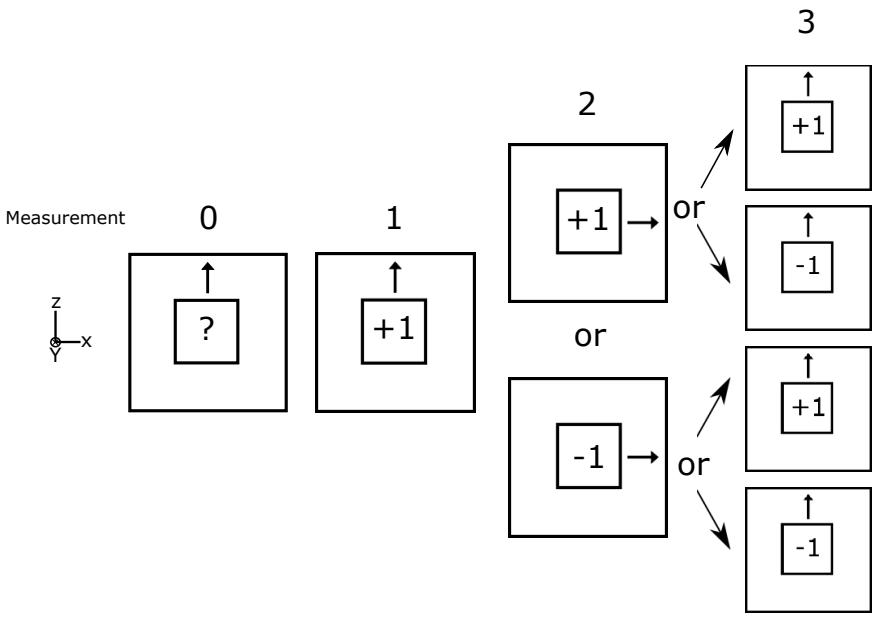

Figure 0.5: Experiment 5

After turning A by 90^o for the second measurement, you turn it back so that it points up just like before. You expect that you should measure $\sigma = +1$ again, because that's the value you got when you measured with A pointing up the first time. However, this is not the case. Instead, you get, in your third measurement, $\sigma = +1$ **or** -1.

What's more? You repeat these steps many times and find that regardless of whether you measure $\sigma = +1$ or $\sigma = -1$ in your second measurement (with A pointing right), in your third measurement (with A point back up), you measure $\sigma = +1$, 50% of the time and $\sigma = -1$, 50% of the time. The second measurement with A pointing right has left our state in a random configuration.

Let us label our measurements σ_z to be the quantum mechanical component of the spin along the z axis (up), measured with A pointing up, and σ_x to be the quantum mechanical component of the spin along the x axis (right), measured with A pointing right. We have shown that measuring σ_z randomizes the measurement we get for σ_x and measuring σ_x randomizes the measurement we get for σ_z. Therefore, the act of measuring σ_z must have disrupted the state of σ_x and measuring σ_x must have disrupted the state of σ_z.

From Experiment 5, we have learned two more things about our spin system:

- We cannot simultaneously know σ_z and σ_x, the component of the spin along the z axis and the x axis, at the same time.
- The above statement can be generalized to any quantum mechanical system in the following way: Any interaction strong enough to measure a property of a system is strong enough to disrupt another property of the system. You can't learn something without changing something else.

Math Interlude I

Complex Numbers

	Cartesian Form	Polar Form
z	$= x + iy$	$= re^{i\theta}$ $= r(\cos\theta + i\sin\theta)$
z^*	$= x - iy$	$= re^{-i\theta}$ $= r(\cos\theta - i\sin\theta)$
z^*z	$= x^2 + y^2$	$= r^2$
z_1z_2	$= (x_1 + iy_1)(x_2 + iy_2)$	$= r_1r_2e^{i(\theta_1+\theta_2)}$

Vector Spaces

CM states live in a set, where Boolean logic applies.
QM states live in a vector space, where a different set of logic applies.

Comparison between a vector we're used to in CM and a general vector

	Classical 3-vector $\vec{v}$	Quantum ket vector $\|v\rangle$
value of components	real numbers	complex numbers integers harmonic polynomials ...
number of components (dimensions)	3	any number from 1 to ∞

Axioms that define a complex vector space (a Hilbert Space) in QM

1. $|A\rangle + |B\rangle = |C\rangle$ — Sum of two kets is another ket
2. $|A\rangle + |B\rangle = |B\rangle + |A\rangle$ — Vector addition is commutative
3. $(|A\rangle + |B\rangle) + |C\rangle = |A\rangle + (|B\rangle + |C\rangle)$ — Vector addition is associative
4. $|A\rangle + |0\rangle = |A\rangle$ — There is a unique vector $|0\rangle$ such that no ket is changed by addition of $|0\rangle$
5. $|A\rangle + (-|A\rangle) = 0$ — There is a unique vector $-|A\rangle$ for every $|A\rangle$ such that their sum is $|0\rangle$
6. $z|A\rangle = |z\rangle A = |B\rangle$ — Any complex number z times a ket is another ket*
7. $z(|A\rangle + |B\rangle) = z|A\rangle + z|B\rangle$ — Distributive Property

*The fact that z can be any complex number is what makes a vector space a complex vector space. This principle does not apply to real 3D vectors, as a complex number times a real 3D vector will give back a complex vector whose components are not necessarily purely real. Therefore, the real 3D space is not a complex vector space (a Hilbert space).

Example of vector space: 2D column vectors, whose components are complex numbers: $\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$ → satisfies all 7 axioms.

Bras and Kets

For every ket (QM vector), there is a bra (QM dual vector). Given a ket, we should know how to find its corresponding bra. In the 2D column vector representation, if $|A\rangle = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$, then $\langle A|$ is the complex conjugate row vector: $\langle A| = \begin{pmatrix}\alpha_1^* & \alpha_2^*\end{pmatrix}$

Axioms for finding the corresponding bra of a ket

1. $|A\rangle + |B\rangle \quad \rightarrow \quad \langle A| + \langle B|$
2. $z\,|A\rangle \quad \rightarrow \quad \langle A|\,z^*$

Inner Product

The inner product can be thought of as a functional $f(A, B)$ that takes in a bra and a ket and produces a complex number, eg. $\langle A|B\rangle = z$, that indicates the overlap between $|A\rangle$ and $|B\rangle$. The functional $f(A, B)$ that we use is dependent on our vector space.

For 3-vectors in the 3D coordinate space, the inner product of $\vec{A}$ and $\vec{B}$ is simply the dot product of vectors $\vec{A}$ and $\vec{B}$. That is, $f(A, B) = \vec{A}\vec{B}$.

For ket vectors in a complex vector space, we are free to choose what $f(A, B)$ is based on the specific complex vector space we're using, and denote the inner product of $\langle A|$ and $|B\rangle$ in bra-ket notation as $\langle A|B\rangle$. In the case of column vectors, $f(A, B)$ is simply the matrix multiplication of $\langle A|$ and $|B\rangle$. For $|A\rangle = \begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix}$ and $|B\rangle = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$, the inner product of $\langle A|$ and $|B\rangle$ is,

$$\langle A|B\rangle = \begin{pmatrix}\alpha_1^* & \alpha_2^*\end{pmatrix}\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$$

Axioms for Inner Products

1. $\langle C|\,(|A\rangle + |B\rangle) = \langle C|A\rangle + \langle C|B\rangle$
2. $\langle B|A\rangle = \langle A|B\rangle^*$ Interchanging bras and kets = complex conjugation

We also use inner products to define properties between two vectors. The two properties we are most concerned with in QM are:

Normalized vector:
$\langle A|A\rangle = 1$
A ket is normalized if its inner product with its own corresponding bra is 1.

Orthogonal vectors:
$\langle B|A\rangle = \langle A|B\rangle = 0$
Two kets are orthogonal if their inner products are equal to 0.

Orthonormal Bases

Basis vectors of a vector space are a set of vectors that can be used to construct any vector in that vector space. A set of basis vectors are orthonormal if for any two distinct basis vectors $|i\rangle$ and $|j\rangle$:

- $\langle i|j\rangle = 0$ They are orthogonal to each other.
- $\langle i|i\rangle = \langle j|j\rangle = 1$ Each basis vector is normalized.

Let's take a look at some examples.

3-vectors
Dimensions: 3
Basis vectors: unit vectors $\hat{x}, \hat{y}, \hat{z}$,
We can form linear combinations $\vec{r} = a\hat{x} + b\hat{y} + c\hat{z}$ to construct any vector $\vec{r}$ in 3D coordinate space. a, b, and c must be real numbers, because these 3-vectors in live a real vector space.

Kets represented by a column vector
Dimensions: N, where N is the # of entries in the column vector
Basis vectors: N column vectors labeled $|1\rangle, |2\rangle, ... |N\rangle$, where $|i\rangle$ for any number i from 1 to N indicates a column vector with 1 in the ith entry and 0 in all the other entires: $|1\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $|2\rangle = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, $|N\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$.
We can form linear combinations of $|i\rangle$ to construct any state $|A\rangle$: $|A\rangle = \sum_{i=1}^{N} \alpha_i |i\rangle$, where α_i is the component of $|A\rangle$ along $|i\rangle$: $\langle j|A\rangle = \sum_{i=1}^{N} \alpha_i \langle j|i\rangle = \alpha_j$ since $\langle j|i\rangle = 0$ unless $i = j$.