Lecture 1 (Sep. 6, 2017)

1.1 A "Weird" Example in Quantum Mechanics

Suppose we have an atom with spin-$\frac{1}{2}$, and we measure the z-component of its spin, S^z. The first claim of quantum mechanics that differs from classical mechanics is that we will always find one of the two values, $\pm\frac{\hbar}{2}$, where $\hbar = \frac{h}{2\pi}$.

The first thing we should wonder is, "How can we reasonably make this measurement?" There is a standard setup for such a measurement known as the Stern–Gerloch filter. We shoot a beam of these atoms through a region of spatially-varying magnetic field $\boldsymbol{B} = B(z)\hat{\boldsymbol{z}}$, such as the region between a square north-pole magnet and a pointed south-pole magnet. When travelling through this region, spin-up and spin-down particles will deflect in opposite directions, allowing us to filter the particles by their spin.

The reason this works is that the energy of the magnetic dipole moment of the atom in the magnetic field is

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\mu_z B_z \,, \tag{1.1}$$

and the force on the atom in the z-direction is thus

$$F_z = -\frac{\partial E}{\partial z} = \mu_z \frac{\mathrm{d}B_z}{\mathrm{d}z} \,. \tag{1.2}$$

Thus, states with different values of μ_z will deflect in different directions along the z-axis.

Now, let's imagine we have such an apparatus that can filter out the components of spin. We could modify this apparatus by changing its axis of orientation, and in this way we could create filters for spin along each of the x-, y-, and z-axes. Imagine then a sequence of filtering protocols to which we can subject our beam of atoms.

First, suppose we use this device to select out the $S^z = +\frac{\hbar}{2}$ atoms, and then measure S^z once again. This can be done by feeding the spin-up output of one z-oriented Stern–Gerloch filter into the input of a second z-oriented Stern–Gerloch filter. With this setup, we will find that 100% of the output of the second Stern–Gerloch filter will be spin-up.

Now consider altering this setup so that the second filter is oriented along the x-axis, so we are sending the atoms through a z-axis filter and then an x-axis filter. The claim in quantum mechanics is that we will now find that $S^x = +\frac{\hbar}{2}$ 50% of the time, and $S^x = -\frac{\hbar}{2}$ 50% of the time.

Now let's complicate matters further by taking the $S^x = +\frac{\hbar}{2}$ output of the x-axis filter and feeding it into yet another z-axis filter. The claim is that we will find $S^z = +\frac{\hbar}{2}$ 50% of the time, and $S^z = -\frac{\hbar}{2}$ 50% of the time. We have filtered out atoms that are only spin-up along the z-axis, then from among these we have filtered out only those that are spin-up along the x-axis, and then measured the spin along the z-axis again, and found that the atoms are equally likely to be spin-up and spin-down along the z-axis.

The reason for this phenomenon is that in quantum mechanics, S^z and S^x are *incompatible observables*, and cannot be measured simultaneously with complete accuracy. As far as we know, this behavior is inconsistent with a classical description of the state of a spin-$\frac{1}{2}$ particle.

1.2 The Fundamental Postulates of Quantum Mechanics

We will take the following as postulates of a quantum mechanical system:

1. The state of a quantum mechanical system at a fixed time t is given by a mathematical object known as a *vector* (or more precisely, a *ray*) $|\psi\rangle$, which is an element of a *complex Hilbert space* $\mathcal{H}$.

2. Observers cannot measure the state directly, but rather can only measure *observables*, which are *Hermitian operators* that act in the Hilbert space $\mathcal{H}$, and whose eigenvectors form a complete set.

3. A measurement of an observable $\hat{A}$ returns one of its eigenvalues.

   a. If $\hat{A}$ is measured in a state $|\psi\rangle$, then the probability that $A = a$ is given by

   $$\langle\psi|M_a|\psi\rangle\,, \tag{1.3}$$

   where

   $$M_a = \sum_{j:a_j=a} |a_j\rangle\langle a_j| \tag{1.4}$$

   is the *measurement operator*. Here, the sum is over all states $|a_j\rangle$ with eigenvalue a, i.e., $\hat{A}|a_j\rangle = a|a_j\rangle$.

   b. After measurement, the system is in the state

   $$|\tilde{\psi}\rangle \propto M_a|\psi\rangle\,. \tag{1.5}$$

   This is an often-discussed feature of quantum mechanics, referred to as the "collapse" of the wavefunction.

4. Time evolution is carried out by a map

   $$|\psi(t)\rangle \mapsto |\psi\big(t'\big)\rangle = U(t',t)|\psi(t)\rangle\,, \tag{1.6}$$

   where $U(t',t)$ is the *time-evolution operator*, which is a *unitary operator*, $U^\dagger U = \mathbb{1}$. The infinitesimal form of the time-evolution operator, when t' and t are infinitesimally separated, takes the form

   $$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle\,, \tag{1.7}$$

   where $\hat{H}$ is a Hermitian operator called the *Hamiltonian*.

1.3 Mathematical Preliminaries

1.3.1 Hilbert Spaces

We will begin by recalling the definition of a vector space. A *vector space* V is a collection of objects $\{|\alpha\rangle\}$ with several properties:

1. There is a well-defined *addition* operator '+' defined on this collection, so that for any $|\alpha\rangle, |\beta\rangle \in V$, there exists a unique state

   $$|\alpha\rangle + |\beta\rangle := |\gamma\rangle \in V\,. \tag{1.8}$$

   a. We demand that the addition operator is *commutative*: for all $|\alpha\rangle, |\beta\rangle \in V$,

   $$|\alpha\rangle + |\beta\rangle = |\beta\rangle + |\alpha\rangle\,. \tag{1.9}$$

   b. We demand that the addition operator is *associative*: for all $|\alpha\rangle, |\beta\rangle, |\gamma\rangle \in V$,

   $$(|\alpha\rangle + |\beta\rangle) + |\gamma\rangle = |\alpha\rangle + (|\beta\rangle + |\gamma\rangle)\,. \tag{1.10}$$

2. We require the existence of a *null vector* $|0\rangle \in V$ with the property that

$$|0\rangle + |\alpha\rangle = |\alpha\rangle \tag{1.11}$$

for all $|\alpha\rangle \in V$.

3. For all $|\alpha\rangle \in V$, we require the existence of an *inverse* state $-|\alpha\rangle \in V$ with the property

$$|\alpha\rangle + (-|\alpha\rangle) = |0\rangle\,. \tag{1.12}$$

4. There is a well-defined *scalar multiplication* operation: for some field $\mathbb{F}$ (such as $\mathbb{R}$ or $\mathbb{C}$), we can multiply any state $|\alpha\rangle \in V$ by any *scalar* $c \in \mathbb{F}$, yielding a unique state $c|\alpha\rangle \in V$.

   a. For all $c, d \in \mathbb{F}$ and $|\alpha\rangle \in V$, scalar multiplication must satisfy

$$c(d|\alpha\rangle) = (cd)|\alpha\rangle\,, \tag{1.13}$$

   where cd indicates multiplication in the field $\mathbb{F}$.

   b. For all $|\alpha\rangle \in V$, the multiplicative identity element $1 \in \mathbb{F}$ must satisfy

$$1|\alpha\rangle = |\alpha\rangle\,. \tag{1.14}$$

   c. Scalar multiplication must distribute over addition in the vector space: for all $c \in \mathbb{F}$ and $|\alpha\rangle, |\beta\rangle \in V$, we must have

$$c(|\alpha\rangle + |\beta\rangle) = c|\alpha\rangle + c|\beta\rangle \tag{1.15}$$

   d. Scalar multiplication must distribute over addition in the field of scalars: for all $c, d \in F$ and $|\alpha\rangle \in V$, we must have

$$(c + d)|\alpha\rangle = c|\alpha\rangle + d|\alpha\rangle\,. \tag{1.16}$$

If the field of scalars is $\mathbb{F}$, then we refer to V as a "vector space over $\mathbb{F}$." A vector space V is called a *real vector space* if its field of scalars is the field of real numbers, $\mathbb{F} = \mathbb{R}$, and a *complex vector space* if its field of scalars is the field of complex numbers, $\mathbb{F} = \mathbb{C}$.

As a familiar example, we might consider the vectors in $\mathbb{R}^n$ of the form

$$\boldsymbol{v} = (v_1, \ldots, v_n)\,. \tag{1.17}$$

These form a real vector space under addition. We could also consider the space of states of a spin-$\frac{1}{2}$ particle,

$$c_+|+\rangle + c_-|-\rangle\,, \tag{1.18}$$

with $c_\pm \in \mathbb{C}$, which form a complex vector space. A more interesting example of a vector space is the space of all real functions $f(x)$ on the interval $[0, 1]$.

A subset $V \subset W$ of a vector space W is a *subspace* of W if it itself is a vector space over the field of scalars of W. A *ray* is a subspace of a vector space V of the form $\{c|\alpha\rangle \mid c \in \mathbb{F}\}$, for some $|\alpha\rangle \in V$.

A collection of vectors $|\alpha_1\rangle, \ldots, |\alpha_n\rangle \in V$ is *linearly independent* if and only if the statement

$$\sum_{j=1}^{n} c_j|\alpha_j\rangle = 0 \tag{1.19}$$

implies that $c_j = 0$ for all $j = 1, \ldots, n$. A collection of vectors $\{|\alpha_j\rangle\}$ is *maximally linearly independent* if it is linearly independent and there exists no state $|\beta\rangle \in V$ such that the collection $\{|\alpha_j\rangle, |\beta\rangle\}$ is linearly independent. In this case, $\{|\alpha_j\rangle\}$ is called a *basis* for V, and the number of elements n in the collection is called the *dimension* of V, denoted $\dim V$. The dimension of a vector space can be finite or infinite, and if infinite, can be either countable (also called denumerable) or uncountable.

Let's consider some examples. A finite-dimensional vector space with which we are already familiar is $\mathbb{R}^n$, which has $\dim \mathbb{R}^n = n$. An example of a infinite-dimensional vector space is the space of square-integrable functions on the interval $[0, 1]$ with $f(0) = f(1) = 0$. This is a countable vector space: we can expand such a function $f(x)$ as

$$f(x) = \sum_{k=1}^{\infty} c_k \sin(\pi k x)\,, \tag{1.20}$$

which provides a countable basis of functions for this space.

Similarly, we can consider the vector space of square integrable functions on $(-\infty, \infty)$. More surprisingly, this vector space also has countable dimension. A clean way to see this is to produce a countable basis for this space: one such basis is the set of eigenfunctions of the simple harmonic oscillator. These eigenfunctions are all square-integrable, and any other wavefunction can be expanded in this basis. The oscillator eigenfunctions are labelled by integers, so this shows that this space is countable. This may seem confusing, because another basis we might use is the position basis, where our basis elements are Dirac delta functions at each position; the problem with this approach is that the Dirac delta function is not itself square-integrable.

8.321 Quantum Theory I

Fall 2017