

Lecture 1. Probability in

The central spike occurs because the two paths form an isosceles triangle. Where does the maximum of first fringe occur? Call the fringe spacing d , the separation between the slits a , and distance from the slits to the screen l . For $l \gg a$, we can show with a little geometry that

$$\frac{a}{l} = \frac{\lambda}{d},$$

which is a well-known result for diffraction in optics. The corresponding law for combining probabilities for particles is

$$p_{tot} = |f_1 + f_2|^2,$$

where

$$p_1 = |f_1|^2,$$

$$p_2 = |f_2|^2$$

and f_1 and f_2 are the (generally complex) *probability amplitudes* for paths 1 and 2.

The conclusion that $p_{tot} = p_1 + p_2$ implies that the particle does not simply pass through either one hole or the other, because if it did one could simply tabulate the outcome for each hole and add the probabilities. In other words, the particle must have wave-like behavior that causes interference. But what is to stop us from simply observing which hole each electron passes through? In other words, suppose we “tag” the electrons and construct two distinct probability density distributions, $p_1(y)$ and $p_2(y)$, particle by particle. The trouble is that when we do this, we destroy the interference; i.e. when we tag the electrons we get

$$p_{exclusive}(y) = p_1(y) + p_2(y) = |f_1|^2 + |f_2|^2 = p_{non-osc},$$

whereas if we don't tag the particles, we get a different result,

$$p_{interfering}(y) = |f_1 + f_2|^2 = p_1(y) + p_2(y) + f_1^* f_2 + f_1 f_2^* = p_{non-osc}.$$

The reason this happens is that “looking” requires us to shine light on the particles, and the scattered light cause the particles to recoil and blur the fringes. Reducing the intensity simply reduces the number of photons, thereby dividing the particles into two classes, tagged and untagged, so that

$$P_{tot} = p_{tagged} + p_{untagged} = p_{non-osc} + p_{osc}.$$

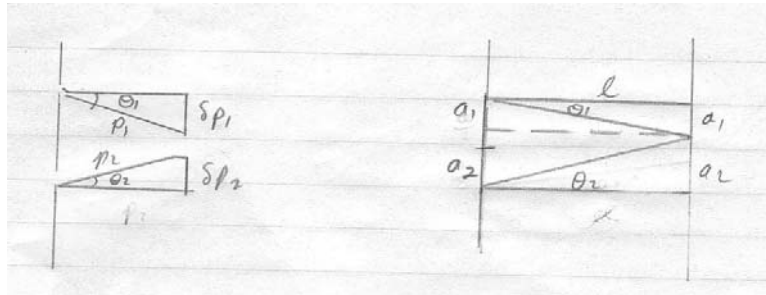
$$\lim_{l \rightarrow 0} p_{tot} = p_{osc}$$

$$\lim_{l \rightarrow \text{large}} p_{tot} = p_{non-osc}$$

Suppose we were to reduce the photon energy so as to reduce the amount of scattering. This means increasing $l = c/u = hc/E$, where we have used Planck's equation $E = h\nu$ for the photoelectric effect. But it turns out that in order to prevent blurring of the oscillations, we require $l > d/2$, which prevents us from knowing which hole the electrons passed through.

Suppose we try an entirely different strategy. Instead of tagging the particles, we measure the recoil of the plate that houses the slits. From the direction of the recoil we can deduce which slit the electron passed through without actually looking at the electron. Let the difference between momentum changes for slit 1 and slit 2 be δp_y . With a little geometry, you can show that

$$\frac{\delta p_y}{p_y} = \frac{a}{l}$$



But we know from the deBroglie experiment that the probability amplitude of a free particle oscillates very rapidly in space, with a wavelength that is given by $l p_y = h$. Invoking the diffraction equation,

$$\frac{a}{l} = \frac{\lambda}{d},$$

it follows that

$$\delta p_y = p_y \left(\frac{a}{l} \right) = p_y \left(\frac{\lambda}{d} \right) = h/d$$

But again it turns out that any attempt to reconstruct from the momentum distributions which hole the electrons pass through wipes out the interference pattern. That is, we need to know not only the momentum of the plate but also the position of the holes while the

plate is moving. What we find is that the uncertainty the latter is $\delta y > d/2$. It follows that the best we can do is

$$\delta p_y \delta y > h/2,$$

which is the well known uncertainty principle.

Conclusion: *When there exist two or more interfering alternatives, we must add the probability amplitudes for each path and then take the square of the sum to get the total probability.* Interfering alternatives are defined as mutually exclusive paths for which no experimental attempt has been made to distinguish between them.

Superposition states.

Consider a beam of light propagating in the z-direction and linearly polarized in the xy plane. The beam is incident on a polarizer cut so that it transmits 100% of the light that is polarized parallel to the crystal axis and 0% that is polarized perpendicular to the crystal axis. We find that (i) when the crystal axis is aligned at an angle a with respect to the y-axis, 100% of the light is transmitted. But (ii) when the crystal is aligned along the y-axis, a fraction $\cos^2 a$ is transmitted, and (iii) when the crystal is aligned along the x-axis, a fraction $\sin^2 a$ is transmitted. These observations are readily understood if the electric vector of the light has the amplitude

$$\mathbf{E}_0 = E_0 \cos a \mathbf{j} + E_0 \sin a \mathbf{i}.$$

How do we explain these observations for a beam of particles (i.e. photons)? From observation (i) we deduce that 100% of the particles are polarized at an angle a . In experiment (ii) only a fraction $\cos^2 a$ of the photons *reach* the detector, but 100% of the detected particles are polarized along the y-axis, and in experiment (iii) a fraction $\sin^2 a$ of the photons reach the detector, and all of those that are detected are polarized along the x-axis. Experiment (i) shows that the photons are in a pure polarization state with an angle a . But experiments (ii) and (iii) show that the photons are in a superposition of perpendicular and parallel polarization states with amplitudes proportional to $\cos a$ and $\sin a$.

Conclusion: *Whenever a system is in one state, it can always be considered to be partly in two or more states, and vice versa.*

Interpretation of the wave function: The square wave function of a system gives the probability of finding the system in a particular location in space. It is incorrect to think of it as giving the number density of particles in a small volume that somehow interfere with each other. The reason why the latter interpretation is false is that the interference pattern persists even if only one particle is measured at a time, with the density so low that two particles are never present in the apparatus simultaneously.

Lecture 2. Probability Distributions

Discrete variables

Consider a property (an “observable”) of the system that can have only discrete values, E_1, E_2, \dots, E_n .

Examples: No. of people in a family, number of vibrational quanta in an oscillator.

We make n measurements, and observe $E = E_1$ in n_1 of those measurements, $E = E_2$ in n_2 measurements, etc., such that

$$\sum_{i=1}^N n_i = n.$$

The probability of observing $E = E_i$ is given by

$$P_i = n_i/n, \quad 0 \leq P_i \leq 1$$

With normalization

$$\sum_{i=1}^N P_i = 1.$$

The set $\{P_i\}$ is called a probability distribution.

If for some s , $P_s = 1$, then all other $P_i \neq s = 0$, and the system is said to be in a definite or “pure” state s of the observable.

The probability of observing $E_1 \leq E \leq E_m$ is

$$P(l, m) = \sum_{i=l}^m P_i$$

The most probable value of E is E_{mp} , where $P_{mp} = \max\{P_i\}$.

The average value, mean value, or expectation value of E is

$$\langle E \rangle = \sum_{i=1}^N P_i E_i$$

which is also called the first moment of $\{P_i\}$. The second moment is given by

$$\langle E^2 \rangle = \sum_{i=1}^N P_i E_i^2.$$

The root mean square value of E is given by

$$E_{rms} = \sqrt{\langle E^2 \rangle}.$$

The variance of E is given by

$$Var(E) = \sigma^2 = \langle E^2 \rangle - \langle E \rangle^2.$$

$\sigma = 0$ for a pure state of E.

The average or expectation value for any function of E is given by

$$\langle f(E) \rangle = \sum_{i=1}^N P_i f(E_i).$$

Lecture 3. Probability Distributions, Part 2.

Continuous variables

Consider a continuous variable, x defined over some domain $a \leq x \leq b$.

Examples: position, velocity, lifetime

The probability of observing a value of the variable between x and x+dx is $p(x)dx$, where $p(x)$ is a probability density function with possible values

$$0 \leq p(x) \leq \infty$$

with dimensions x^{-1} .

Normalization:
$$\int_a^b p(x)dx = 1$$

Total probability:
$$P(x_1, x_2) = \int_{x_1}^{x_2} p(x)dx.$$

The domain of x depends on the nature of x. For example,

Speed:	$0 \leq v \leq \infty$
Velocity:	$-\infty \leq v_x \leq \infty$
Radial distance:	$0 \leq r \leq \infty$
Polar angle:	$0 \leq \theta \leq \pi$
Azimuthal angle:	$0 \leq \phi \leq 2\pi$

Properties of the distribution function:

x_{mp} satisfies $\frac{dp(x)}{dx} = 0$

$$\langle x \rangle = \int_a^b xp(x)dx$$

$$\langle f(x) \rangle = \int_a^b f(x)p(x)dx$$

Example discussed in class: the 1D Boltzmann distribution function.

$$p(v)dv = \frac{1}{\sqrt{\pi}\alpha} \exp(-v^2 / \alpha^2) dv$$

$$p(z)dz = \frac{1}{\sqrt{\pi}} e^{-z^2} dz$$

$$\begin{aligned}\vec{v} &= v_x \hat{i} + v_y \hat{j} \\ v &= \sqrt{v_x^2 + v_y^2} \\ \tan \phi &= v_y / v_x \\ dv_x dv_y &= v dv d\phi\end{aligned}$$

Define a reduced variable:

$$z = v/a$$

$$\begin{aligned}p(v_x, v_y) dv_x dv_y &= F(v, \phi) v dv d\phi \\ &= \frac{1}{\pi \alpha^2} \exp\left\{-\left(v_x^2 + v_y^2\right) / \alpha^2\right\} dv_x dv_y = \frac{2v}{\alpha^2} e^{-v^2 / \alpha^2} dv \frac{d\phi}{2\pi} \\ &= 2ze^{-z^2} dz \frac{d\phi}{2\pi} \\ &= F_z(z) dz \cdot F_\phi(\phi) d\phi\end{aligned}$$

where

$$F_z(z) = 2ze^{-z^2}$$

$$F_\phi(\phi) = \frac{1}{2\pi}$$

Normalization:

$$\begin{aligned}\int_0^\infty F_z(z) dz &= 1 \\ \int_0^{2\pi} F_\phi(\phi) d\phi &= 1\end{aligned}$$

The factor of z (or v) is a Jacobian that arises from the transformation. It may either be included in the distribution function, as it is here, or it may be inserted separately into the normalization integral, as in the case of the wave function.

Example of the hydrogen wave function.

Lecture 4. Hilbert Spaces

Let's start with the familiar example of Cartesian vectors:

$$\begin{aligned} \hat{i}, \hat{j}, \hat{k} \\ \hat{x}, \hat{y}, \hat{z} \\ |1\rangle, |2\rangle, |3\rangle \end{aligned}$$

These are called “basis vectors” or a “basis set.” An arbitrary vector may be written as a linear combination of the basis vectors.

$$\begin{aligned} \hat{v} &= a_1 \hat{i} + a_2 \hat{j} + a_3 \hat{k} \\ &= a_1 |1\rangle + a_2 |2\rangle + a_3 |3\rangle \\ &= \sum_{n=1}^3 a_n |n\rangle. \end{aligned}$$

The coefficients a_n are in general complex.

$$\begin{aligned} a &= \text{Re } a + i \text{Im } a = |a| e^{i\delta} \\ |a|^2 &= \text{mod}^2(a) = (\text{Re } a)^2 + (\text{Im } a)^2 \\ \delta &= \tan^{-1}(\text{Im } a / \text{Re } a) = \arg(a) \end{aligned}$$

Orthonormality of the basis vectors:

$$\begin{aligned} \hat{i} \cdot \hat{i} &= 1, \\ \hat{i} \cdot \hat{j} &= 0, \text{ etc} \end{aligned}$$

Dirac notation for a vector and its dual:

The dual basis set is written as $\langle 1|, \langle 2|, \langle 3|$

Orthonormality in Dirac notation:

$$\begin{aligned} \langle m|n\rangle &= 1, \text{ for } m=n \\ \langle m|n\rangle &= 0, \text{ for } m \neq n \\ \langle m|n\rangle &= \delta_{mn} \text{ in general.} \end{aligned}$$

An arbitrary column vector may be written as

$$|v\rangle = \sum_{n=1}^3 a_n |n\rangle.$$

Its dual is a row vector that is written as

$$\langle v| = \sum_{n=1}^3 a_n^* \langle n|.$$

Similarly, another row vector may be written as

$$\langle u| = \sum_{n=1}^3 b_n^* \langle n|.$$

The “dot” or “inner” has the scalar value

$$\begin{aligned} \langle u|v\rangle &= \vec{u} \cdot \vec{v} \\ &= \sum_m \sum_n \langle m|b_m^* a_n|n\rangle \\ &= \sum_m \sum_n b_m^* a_n \langle m|n\rangle \\ &= \sum_m \sum_n b_m^* a_n \delta_{mn} \\ &= \sum_n b_n^* a_n = b_1^* a_1 + b_2^* a_2 + b_3^* a_3 \end{aligned}$$

Dot product of an arbitrary column vector with a basis row vector:

$$\langle n|v\rangle = \langle n|\sum_i a_i|i\rangle = a_n$$

It follows that

$$|v\rangle = \sum_i a_i|i\rangle = \sum_i \langle i|v\rangle|i\rangle = \sum_i |i\rangle\langle i|v\rangle$$

$\langle i|v\rangle$ = projection of $|v\rangle$ onto $|i\rangle$

$|i\rangle\langle i|$ = projection operator.

Numerical example. $|v\rangle = 5|1\rangle + 3i|2\rangle - \sqrt{2}|3\rangle$. Calculate $|2\rangle\langle 2|v\rangle$

Dot product of a vector with itself:

$$\langle v|v\rangle = \sum_i \sum_j \langle j|a_j^* a_i|i\rangle = \sum_i a_i^* a_i = \sum_i |a_i|^2$$

Normalized vectors:

$$|\hat{v}\rangle = \frac{1}{\sqrt{\sum_i |a_i|^2}} |v\rangle$$
$$\langle \hat{v} | \hat{v} \rangle = 1$$

Two vectors are orthogonal if $\langle u | v \rangle = 0$.

Two vectors are orthonormal if

$$\langle u | v \rangle = 0,$$
$$\langle u | u \rangle = 1,$$
$$\langle v | v \rangle = 1.$$

Basis vectors are orthonormal. A complete set of basis vectors are said to “span” the vector space in which they are defined because there do not exist any vectors that are orthogonal to all of the basis vectors.

Matrix notation:

$$|v\rangle = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$
$$\langle u| = (b_1^* \quad b_2^* \quad b_3^*)$$
$$\langle u | v \rangle = (b_1^* \quad b_2^* \quad b_3^*) \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = b_1^* a_1 + b_2^* a_2 + b_3^* a_3$$

Lecture 5. Orthonormal Functions

Two key generalizations:

1. There may be N dimensions, and N may be infinite.
2. The basis vectors and their dot products may be functions of some variable, x , defined over an interval $[a,b]$.

$$|n\rangle = \phi_n(x)$$

$$\langle m| = \phi_m^*(x)$$

$$\langle m|n\rangle = \int_a^b \phi_m^*(x)\phi_n(x)dx$$

More generally, $\langle m|n\rangle = \int_a^b \phi_m^*(x)\phi_n(x)r(x)dx$, where $r(x)$ is a weighting function.

Example of a complete set of orthonormal functions:

$$|n\rangle = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

where $0 \leq x \leq L$, $n = 1, \dots, \infty$.

$$\langle n|n\rangle = \frac{2}{L} \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) dx = 1$$

$$\langle m|n\rangle = \frac{2}{L} \int_0^L \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx = \delta_{mn}$$

Consider some arbitrary, but well-behaved, function $f(x)$ defined over $[a,b]$.

Recall that $|v\rangle = \sum_n a_n |n\rangle = \sum_n |n\rangle \langle n|v\rangle$

Similarly,

$$f(x) = \sum_{n=1}^N a_n \phi_n$$

$$a_n = \int_a^b \phi_n^*(x) f(x) dx$$

In the example given here,

$$a_n = \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{n\pi x}{L}\right) f(x) dx$$

Example in class: $f(x) = x$.

A complete set of orthogonal functions may be found by solving a suitable differential equation, along with the appropriate boundary equations. The example of the sine functions may be derived by solving the following differential equation,

$$\frac{d^2 y}{dx^2} + \lambda y = 0, \quad 0 \leq x \leq L$$

subject to the boundary conditions,

$$\begin{aligned} y(0) &= 0 \\ y(L) &= 0 \end{aligned}$$

Solution of the differential equation:

$$y(x) = A \sin(ax + \phi).$$

The boundary conditions require that

$$a = \frac{n\pi}{L},$$

where n is an integer, and $\phi = 0$, so that

$$y_n(x) = A \sin \frac{n\pi x}{L}.$$

The boundary conditions also impose constraints on λ :

$$\frac{d^2 y}{dx^2} = -\left(\frac{n\pi}{L}\right)^2 y(x) \Rightarrow \lambda_n = \left(\frac{n\pi}{L}\right)^2.$$

Normalization:

$$\int_0^L y_n^2(x) dx = 1$$

$$A^2 \int_0^L \sin^2 \frac{n\pi x}{L} dx = A^2 \frac{L}{2} = 1$$

$$y_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$$

Orthogonality:

$$\frac{2}{L} \int_0^L \sin \frac{n\pi x}{L} \sin \frac{m\pi x}{L} dx = \delta_{mn}$$

Changing the boundary conditions for the same differential equations gives a different set of orthogonal functions:

$$y'(0) = y'(L) = 0 \Rightarrow y_n(x) = \sqrt{\frac{2}{L}} \cos \frac{n\pi x}{L}$$

Generalization: The Sturm-Liouville Equation

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + [q(x) + \lambda r(x)] y(x) = 0$$

$$a \leq x \leq b$$

p , q , and r are real functions. They could be constants.

λ is a constant called an eigenvalue.

$r(x)$ is a weighting function. It may be lumped into $y(x)$ as $\sqrt{r(x)}y(x)$.

Possible boundary conditions:

1. $y(a) = y(b) = 0$
2. $y^f(a) = y^f(b) = 0$
3. $y(a) + a y^f(a) = y(b) + y^f(b) = 0$
4. $p(a) = 0$ or $p(b) = 0$ may replace any of the above.

Properties of the solutions:

1. $f_n(x)$ is a real solution associated with a real, non-negative eigenvalue λ_n

2. Normalization: $\int_a^b [\phi_n(x)]^2 r(x) dx = 1$

3. Orthogonality: $\int_a^b \phi_m(x)\phi_n(x)r(x) dx = \delta_{mn}$

4. Completeness: $|f\rangle = \sum_i |i\rangle\langle i|f\rangle$

An even more general problem is the solution of the following differential equation,

$$a_0(x) y'' + a_1(x) y' + [a_2(x) + \lambda a_3(x)]y = 0$$

It may be converted into the standard Sturm-Liouville form by means of the following identifications:

$$p(x) = \exp\left\{\int \frac{a_1(x)}{a_0(x)} dx\right\}$$

$$q(x) = \frac{a_2(x)}{a_0(x)} p(x)$$

$$r(x) = \frac{a_3(x)}{a_0(x)} p(x)$$

You may verify this by direct substitution into the Sturm-Liouville equation. Note that

$$\frac{dp}{dx} = p(x) \frac{a_1(x)}{a_0(x)}$$

Lecture 6. Boundary Value Problems and Quantum Mechanics

Let's re-examine the Sturm-Liouville problem by collecting all the terms except for $\lambda y(x)$. We will label them as $-L y(x)$, where L is a differential operator. The equation then becomes

$$L y(x) = \lambda y(x).$$

This equation says that if $y(x)$ happens to be an eigenfunction of L , then operating on $y(x)$ does change it apart from a multiplicative constant. If $y(x)$ is not an eigenfunction, we can still operate on it with L , but we will end up with an entirely new function. To see this, let's expand $y(x)$ as a series of eigenfunctions:

$$y(x) = \sum_i a_i |i\rangle = \sum_i a_i \phi_i(x)$$

$$L(x) y(x) = L(x) \sum_i a_i \phi_i(x) = \sum_i a_i L(x) \phi_i(x) = \sum_i a_i \lambda_i \phi_i(x).$$

A key property of L that is essential for the last step to work is that it is a linear operator. That is,

$$L \{a f(x) + b g(x)\} = a L f(x) + b L g(x).$$

An example is $L = d/dx$. What are its eigenfunctions? Is it linear?

We can determine an individual coefficient a_s by taking a dot-product:

$$\int \phi_s^*(x) L(x) y(x) r(x) dx = \langle s | L | y \rangle = a_s$$

We could describe this process by saying that the operator $\langle i | L$ projects out component $|i\rangle$ from an arbitrary function of x .

To see the connection with quantum mechanics, recall the Stern-Gerlach experiment, or the photon polarization experiment. If the filter is aligned along the y -axis and the photon is y -polarized, then the filter transmits the photon without changing it. But for any other polarization, the filter will transmit only the y -component of the photon state. Similarly, if the filter is aligned along the x -axis, it will transmit only the x -component. The filter behaves like a projection operator.

This leads us to make some postulates:

For every observable L , there exists a linear operator $L(x)$ with eigenvalues λ_i . The x variable may be a spatial coordinate or a momentum. The only possible outcomes of a measurement are the eigenvalues of L . For example, if the system happens to be in a

“pure” state of L , described by the function $\phi_i(x)$, then the only possible outcome of the measurement is λ_i . The expectation value of L is then simply

$$\langle L \rangle = \langle i | L | i \rangle = \lambda_i$$

If system is in some other state, it can be written as a superposition of $|i\rangle$ with coefficients a_i ,

$$f(x) = \sum_i a_i \phi_i(x),$$

If the probability of obtaining outcome λ_i is P_i , then the expectation value of L is

$$\langle L \rangle = \sum_i P_i \lambda_i$$

We can evaluate P_i by operating on $f(x)$ with the projection operator for each eigenstate and then sum over all possible outcomes:

$$\langle L \rangle = \sum_i \sum_j \langle i | a_i^* L a_j | j \rangle = \sum_i |a_i|^2 \lambda_i$$

We see that $P_i = |a_i|^2$.

It remains to find the operator L , and to assign a meaning to the eigenfunctions, $f_i(x)$. We will postpone the former to a later lecture. We can gain insight into $f_i(x)$ by considering the special case that L is simply a multiplicative (rather than a differential) operator. An example is the coordinate operator which is just x itself. In this case,

$$\langle L \rangle = \int \phi_i^*(x) L \phi_i(x) dx = \int |\phi_i|^2 L dx.$$

$$\langle x \rangle = \int |\phi_i(x)|^2 x dx$$

In this case it is reasonable to identify the probability density $p(x)$ of finding the system located between x and $x+dx$ to be given by $|\phi_i|^2$. This is also true if the system is not in an eigenstate,

$$\langle x \rangle = \int f^*(x) x f(x) dx = \int |f|^2 x dx.$$

The function $f(x)$ is a probability amplitude, also known as a wave function.

We see that there are two types of probability distributions. The first is the probability that an object is in a particular eigenstate of L . That probability is the square of the

expansion coefficient, a_i . The second probability is the probability density that the object is located at position x , given by the square of the wave function. (There is also another wave function that gives the probability density that the object has a particular momentum. We will learn about this later.) $|f_i(x)|^2$ is a conditional probability in the sense that if the system is in state i , then $|f_i(x)|^2$ is the probability density for observing the system at coordinate x .

A key property of L is that it has the Sturm-Liouville form. This property not only assures that L is linear, but also that its eigenfunctions comprise a complete set of orthogonal functions.

Example: The spin angular momentum has projections

$$s_z = \pm \frac{1}{2} \hbar$$

$$S_z |1\rangle = \frac{1}{2} \hbar |1\rangle$$

$$S_z |2\rangle = -\frac{1}{2} \hbar |2\rangle$$

Consider the state $|s\rangle = \frac{1}{\sqrt{10}} |1\rangle + \frac{3}{\sqrt{10}} |2\rangle$

What are the probabilities of observing $s_z = \frac{1}{2} \hbar$ and $s_z = -\frac{1}{2} \hbar$? What is the expectation value of S_z ?

Summary of the above arguments:

1. If a system is in a “pure” state of a particular observable, then a measurement of that observable leaves the system in its original state. This is described mathematically by operating on the state:

$$\hat{L}|i\rangle = \lambda_i |i\rangle$$

2. We may determine the value of the observable by projecting the above result back onto the original state:

$$\langle i|\hat{L}|i\rangle = \langle i|\lambda_i|i\rangle = \lambda_i$$

3. The above results apply even if the system is not in a pure state of the observable.

We can calculate the expectation value of the observable by first expanding the state in the basis set of the observable, then operating on the expansion, and finally projecting the result back onto the original state:

$$\begin{aligned} |f\rangle &= \sum_i a_i |i\rangle \\ \hat{L}|f\rangle &= \sum_i a_i \lambda_i |i\rangle \\ \langle f|\hat{L}|f\rangle &= \sum_i |a_i|^2 \lambda_i = \sum_i P_i \lambda_i \end{aligned}$$

4. The above arguments apply also if the eigenvalue happens to be continuous. (This is not a Sturm-Liouville problem, but the principles are the same). In this case the sum over i becomes an integral over the eigenvalue, which we will denote by x :

$$\langle x \rangle = \int f^*(x) \hat{x} f(x) dx$$

5. If x happens to be a coordinate, then $\hat{x} = x$, and

$$\langle x \rangle = \int f^*(x) x f(x) dx = \int x |f(x)|^2 dx$$

We see in this case that $f(x)$ is the probability amplitude for finding the system between x and $x + dx$.

6. If the observable is a momentum, then the corresponding operator is

$$\hat{p} = -i\hbar \frac{d}{dx}.$$

We may continue to describe the system in the coordinate basis, in which case the expectation value of p is

$$\langle p \rangle = \int f^*(x) \hat{p} f(x) dx = -i\hbar \int f^*(x) \frac{df(x)}{dx} dx$$

7. The system cannot always be in a pure state of two observables simultaneously. If it cannot, and if it is in a pure state of one observable, then it is necessarily in a superposition of states of the second observable. This is why we spoke before about two types of probabilities. They are the probability of being in a particular pure state of the first observable and the distribution of states for the second observable. The “compatibility” of the two observables (that is, the possibility of being in a pure state of both at the same time), is explained by the Uncertainty Principle, which we will defer to a later discussion.

Lecture 7. Particle in a Box

Consider a particle bound in a one-dimensional well:

$$V=0 \text{ for } 0 \leq x \leq L, V = \infty \text{ elsewhere}$$

We would like to know the possible energy levels of the particle. To find these, we solve the eigenvalue equation,

$$H y = E y$$

where H is the energy operator and E is its eigenvalue. The recipe for constructing operators, to be rationalized later, is as follows:

$$\begin{array}{ll} \text{coordinates:} & x \leftrightarrow x \\ \text{momenta:} & p \leftrightarrow -i\hbar \frac{\partial}{\partial x} \end{array}$$

The energy operator is $H = \text{kinetic energy} + \text{potential energy} = \frac{p^2}{2m} + V(x)$

In the present case, $V(x) = 0$ in the regions where the particle can be found. The operator for the kinetic energy is

$$\frac{p^2}{2m} = \left(-i\hbar \frac{d}{dx} \right)^2 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2},$$

with boundary conditions $y(0) = y(L) = 0$.

The Sturm-Liouville equation that we must solve is

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} = E \psi(x)$$

We already showed that the eigenfunctions are

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

The eigenvalues are obtained by substituting y_n back into the differential equation:

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} = \frac{\hbar^2}{2m} \left(\frac{n\pi}{L}\right)^2 \psi(x)$$

It follows that

$$E_n = \frac{\hbar^2 n^2}{8mL^2} = E_1 n^2$$

E_1 is the zero point energy.

Example of an electron in a 1-D quantum dot 0.5 nm long.

Properties of the eigenstates:

n-1 nodes

$$\langle x \rangle = \langle n | x | n \rangle = \frac{2}{L} \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) x dx = \frac{L}{2}$$

$$\langle x^2 \rangle = \langle n | x^2 | n \rangle = \frac{2}{L} \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) x^2 dx = \frac{L^2}{3}$$

The variance of x is $\frac{L^2}{3} - \left(\frac{L}{2}\right)^2 = \frac{L^2}{12}$; $\sigma = \frac{L}{2\sqrt{3}}$

$$\langle p \rangle = \langle n | -i\hbar \frac{d}{dx} | n \rangle = 0$$

$$\langle p^2 \rangle = 2mE_n$$

Suppose the particle is not in a pure state. There are two possibilities, an incoherent mixture and a coherent superposition.

Example of the former is a Boltzmann mixture. The property that makes it an incoherent mixture is that the wave functions for each eigenstate have no definite phase relation with each other.

$$P_n = \frac{e^{-E_n/kT}}{\sum_s e^{-E_s/kT}}$$

A density function, rather than a wave function, is needed to describe it.

A coherent superposition is described by a single wave function:

$$\psi(x) = \sum_n a_n \psi_n(x)$$

For example,

$$\psi(x) = \sqrt{\frac{1}{3}} \psi_1(x) + \sqrt{\frac{2}{3}} \psi_3(x)$$

$$\langle E \rangle = \sum_m \sum_n a_m^* a_n \langle m | H | n \rangle = \sum_m \sum_n a_m^* a_n E_n \delta_{mn} = \sum_m |a_m|^2 E_m$$

In the example, $\langle E \rangle = \frac{1}{3} E_1 + \frac{2}{3} E_3$

Here is a more complicated example:

$$\psi(x) = cx \sin \frac{\pi x}{L}$$

First find the normalization constant:

$$c^2 \int_0^L x^2 \sin^2 \left(\frac{\pi x}{L} \right) dx = 1$$

$$c = \sqrt{\frac{6}{L^3}}$$

Here are two ways to calculate the expectation value of the energy.

Method 1:

$$\langle E \rangle = \int_0^L \psi^*(x) H \psi(x) dx$$

We'll work out the integral in detail.

Method 2. Expand $y(x)$ in a Fourier sine series.

$$\psi(x) = \sqrt{\frac{2}{L}} \sum_n a_n \sin \left(\frac{n\pi x}{L} \right)$$

$$\langle E \rangle = \sum_n a_n^2 E_n = E_1 \sum_n n^2 a_n^2$$

We will also work this out in detail in class.

Lecture 8. Particle in a 3D Box

Consider a box of dimensions L_x, L_y, L_z . A particle of mass m is placed in this box with boundary conditions such that the wave function vanishes on all the walls. The infinite potential outside the box guarantees that the wave function is zero there as well.

Schrodinger equation:

$$-\frac{\hbar^2}{2m} \left(\frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} \right) \psi(x, y, z) = E \psi(x, y, z)$$

Let us make the following assumption:

$$\psi(x, y, z) = \psi_x(x) \psi_y(y) \psi_z(z),$$

with boundary conditions

$$\begin{aligned} \psi_x(0) &= \psi_x(L_x) = 0, \\ \psi_y(0) &= \psi_y(L_y) = 0, \\ \psi_z(0) &= \psi_z(L_z) = 0. \end{aligned}$$

The Schrodinger eq. now becomes:

$$-\frac{\hbar^2}{2m} (\psi_x'' \psi_y \psi_z + \psi_x \psi_y'' \psi_z + \psi_x \psi_y \psi_z'') = E \psi_x \psi_y \psi_z.$$

Divide both sides by $\psi_x \psi_y \psi_z$:

$$-\frac{\hbar^2}{2m} \left(\frac{\psi_x''}{\psi_x} + \frac{\psi_y''}{\psi_y} + \frac{\psi_z''}{\psi_z} \right) = E.$$

Rearrange:

$$\frac{\hbar^2}{2m} \left(\frac{\psi_y''}{\psi_y} + \frac{\psi_z''}{\psi_z} \right) + E = -\frac{\hbar^2}{2m} \frac{\psi_x''}{\psi_x}$$

The LHS is independent of x . The only way that the RHS can also be independent of x is if it equals a constant. Call that constant E_x . It follows that the LHS satisfies a 1D Schrodinger equation:

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi_x}{dx^2} = E_x \psi_x$$

The total energy is given by

$$E = E_x + E_y + E_z = \frac{h^2}{8m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right)$$

For a cube of side L,

$$E = \frac{h^2}{8mL^2} (n_x^2 + n_y^2 + n_z^2)$$

Degeneracy of a Cube

n_x, n_y, n_z	n^2	g
1 1 1	3	1
(112), (121), (211)	6	3
(122), (212), (221)	9	3
(113), (131), (311)	11	3
2 2 2	12	1
(123), (132), (213), (231), (312), (321)	14	6
...		
(511), (151), (115), (333)	27	4

The last example is a case of accidental degeneracy. Can you find more examples?

Density of States

How many states with energy $\leq E$ are there in a cube?

Construct the volume of an octant with radius n in a space with axes n_x, n_y, n_z .

$$N(E) = \frac{1}{8} \cdot \frac{4}{3} \pi n^3 = \frac{\pi}{6} n^3,$$

where

$$E = \frac{h^2 n^2}{8mL^2} = E_1 n^2.$$

It follows that

$$n = \sqrt{E/E_1}$$

and

$$N(E) = \frac{\pi}{6} \left(\frac{E}{E_1} \right)^{3/2} \propto E^{3/2}.$$

The density of states is given by

$$\rho(E) = \frac{dN(E)}{dE} = \frac{\pi}{4E_0^{3/2}} \sqrt{E}$$

Example of a Bose-Einstein condensate of ^{87}Rb at 100 nK with $L = 50 \mu\text{m}$.