Quantum Mechanics

Muchang Bahng

Spring 2024

Contents

1   Systems

Imagine a particle of mass m constrained to move along the x-axis, subject to some applied force $F(x,t)$ (time-dependent). In classical mechanics, we want to determine the position of the particle at any given time t by finding the function $x(t)$. How do we do this? We simply apply *Newton's second law*.

$$\mathbf{F} = m\mathbf{a} \tag{1}$$

and solve the ordinary differential equation (with some initial conditions), either analytically or numerically. Once we have $x(t)$, we can find other metrics of interest, such as the velocity $v(t)$, kinetic energy $T = \frac{1}{2}mv^2$, or others. For conservative systems (the only kind we'll consider, and fortunately the only kind that exists at the microscopic level), the force can be expressed as the gradient of a potential energy function $U(\mathbf{x})$.

1.1   Schrödingers Equation

Quantum mechanics approaches the same problem differently. Rather than looking for position function $\mathbf{x}(t)$, we are looking for a **wave function** $\mathbf{\Psi}(\mathbf{x},t)$ of the particle, which we can get by solving the *Schrödinger equation*.

Definition 1.1 (Schrödinger Equation)

The **Schrödinger equation** is defined

$$i\hbar\frac{\partial\mathbf{\Psi}}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\mathbf{\Psi}}{\partial\mathbf{x}^2} + V\mathbf{\Psi} \tag{2}$$

where $\hbar$ is the **reduced Planck's constant**, defined

$$\hbar = \frac{h}{2\pi} = 1.054573 \times 10^{-34}\mathrm{J}s \tag{3}$$

Therefore, given suitable initial conditions $\mathbf{\Psi}(\mathbf{x},0)$, the Schrödinger equation determines $\mathbf{\Psi}(\mathbf{x},t)$ for all future time. Now let's talk about this wave function and its physical interpretation, starting with **Born's statistical interpretation**. This says that the wavefunction $\mathbf{\Psi}(\mathbf{x},t)$ determines the probability of finding th particle at point $\mathbf{x}$ at time t. That is, the probability density function of the particle's position at time t is given by

$$f_{X,t}(\mathbf{x},t) = |\mathbf{\Psi}(\mathbf{x},t)|^2 \tag{4}$$

Unfortunately, Schrödinger's equation is a linear system, so the set of solutions to this equation forms a vector space. So if $\mathbf{\Psi}$ is a solution, then $c\mathbf{\Psi}$ is also a solution for all $c \in \mathbb{C}$. This is a problem since it must be normalized. This is why we have an extra condition that

$$\int |\mathbf{\Psi}(\mathbf{x},t)|^2 \, d\mathbf{x} = 1 \tag{5}$$

which means that the probability of finding a particle somewhere in the space X at a certain point t must integrate to 1. This gives us a normalization condition, and any functions that have an integral of infinity is not within our search space. Therefore, we're really just trying to find a function in the L^2 space of integrable functions.

Definition 1.2 (Superposition)

Now the solutions of Schrodinger's equation, as a linear PDE, forms a vector space, and so given two solutions ψ_1, ψ_2, any linear combination of them (after normalization), known as a superposition, is also a solution.

There's two things that the reader may realize: First, it seems that multiple wavefunctions Ψ may induce the same probability measure. It turns out that within this set of normalized wavefunctions, there

is an equivalence class denoted by $\Psi \sim e^{i\theta}\Psi$ for all $\theta \in \mathbb{R}$, which means that $|e^{i\theta}\Psi|^2 = (e^{i\theta}\Psi)^*(e^{i\theta}\Psi) = \Psi^* e^{-i\theta} e^{i\theta}\Psi = |\Psi|^2$. This is a consequence of the fact that the probability of finding a particle at a certain point is not affected by the phase of the wavefunction.

> **Definition 1.3 (Global Phase Factor)**
>
> Two wavefunctions are said to differ by a **global phase factor** iff they differ by a normalized complex scalar $e^{i\theta}$.

For now, we can think of two wavefunctions that differ by a global phase factor as being the same. The important distinction is the relative phase factor, which is the difference between two wavefunctions at a certain point.

> **Definition 1.4 (Relative Phase)**
>
> Two wavefunctions are said to differ by a **relative phase** if they differ by a normalized complex scalar $e^{i\theta(x)}$ that is dependent on the position x.

Second, this normalization restriction may not be preserved in Shrodinger's equation. Fortunately for us, Schrödinger's equation keeps this normalization condition as time passes. Let's prove this.

> **Theorem 1.1 ()**
>
> Given a solution $\boldsymbol{\Psi}$ that has been normalized, the function will stay normalized as time passes.

> **Proof.**
>
> Let us take the time-derivative of the total probability and show that is is 0. We show that
>
> $$\frac{d}{dt}\int |\boldsymbol{\Psi}(\mathbf{x},t)|^2 \, dx = \int \frac{\partial}{\partial t}|\boldsymbol{\Psi}(\mathbf{x},t)|^2 \, dx \tag{6}$$
>
> and by the product rule we have
>
> $$\frac{\partial}{\partial t}|\boldsymbol{\Psi}|^2 = \frac{\partial}{\partial t}(\boldsymbol{\Psi}^*\boldsymbol{\Psi}) = \boldsymbol{\Psi}^*\frac{\partial \boldsymbol{\Psi}}{\partial t} + \frac{\partial \boldsymbol{\Psi}^*}{\partial t}\boldsymbol{\Psi} \tag{7}$$
>
> We can take the complex conjugate of the Schrödinger equation to get
>
> $$\frac{\partial \boldsymbol{\Psi}}{\partial t} = \frac{i\hbar}{2m}\frac{\partial^2 \boldsymbol{\Psi}}{\partial x^2} - \frac{i}{\hbar}V\boldsymbol{\Psi} \tag{8}$$
>
> $$\frac{\partial \boldsymbol{\Psi}^*}{\partial t} = -\frac{i\hbar}{2m}\frac{\partial^2 \boldsymbol{\Psi}^*}{\partial x^2} + \frac{i}{\hbar}V\boldsymbol{\Psi}^* \tag{9}$$
>
> and substituting both equations into the product rule gives
>
> $$\frac{\partial}{\partial t}|\boldsymbol{\Psi}|^2 = \frac{i\hbar}{2m}\left(\boldsymbol{\Psi}^*\frac{\partial^2 \boldsymbol{\Psi}}{\partial x^2} - \frac{\partial^2 \boldsymbol{\Psi}^*}{\partial x^2}\boldsymbol{\Psi}\right) = \frac{\partial}{\partial x}\left[\frac{i\hbar}{2m}\left(\boldsymbol{\Psi}^*\frac{\partial \boldsymbol{\Psi}}{\partial x} - \frac{\partial \boldsymbol{\Psi}^*}{\partial x}\boldsymbol{\Psi}\right)\right] \tag{10}$$
>
> and now we can evaluate the integral to be
>
> $$\frac{d}{dt}\int |\boldsymbol{\Psi}(x,t)|^2 \, dx = \frac{i\hbar}{2m}\left(\boldsymbol{\Psi}^*\frac{\partial \boldsymbol{\Psi}}{\partial x} - \frac{\partial \boldsymbol{\Psi}^*}{\partial x}\boldsymbol{\Psi}\right)\Big|_{-\infty}^{+\infty} \tag{11}$$
>
> which evaluates to 0 since $\boldsymbol{\Psi}$ must go to 0 as it is a probability density.

Now that we have settled on what a wavefunction is, we can now introduce the braket notation.

1.2   Ket Notation

We can talk about another paradigm of representing the states of a system: as a ket. We can equivalently say that the state of a system is represented by a normalized vector $|\psi\rangle$ in the Hilbert space L^2. Now it seems that both ψ and $|\psi\rangle$ both live in L^2, so what is the difference between them? This is something that is not constructed very rigorously in introductory quantum mechanics courses, so let's try to settle this by constructing $|\psi\rangle$ from ψ.

     Note that we have two spaces here: the function space L^2 and the state space X. In 1939, Dirac wanted to unify these 2 spaces in order to more easily work with them. He does this by first constructing some Hilbert space $\mathcal{H}$ and identifying two mappings $x \in X \mapsto |x\rangle \in \mathcal{H}$ and $\psi \in L^2 \mapsto |\psi\rangle \in \mathcal{H}$. The goal was to be able to evaluate $\psi(x)$ with an inner product $\langle x|\psi\rangle$ in $\mathcal{H}$. Let's talk about each construction separately:

1. To do this, note that ψ can be reduced to an uncountable sum of δ-functions.

$$\psi(x) = \int_{x \in X} \psi(x')\delta_{x'}(x)\, dx' \tag{12}$$

   This is analogous to an uncountable sum of basis vector functions $\delta_{x'}$ and their coefficients $\psi(x')$. Each $\delta_{x'}$ can be though of as an uncountably long vector

$$|\delta_{x'}\rangle = (\ldots, 0, \ldots, 0, 1, 0, \ldots, 0, \ldots) \tag{13}$$

   where the 1 is in the x' index, and each function is a vector

$$|\psi\rangle = (\ldots \psi(0) \ldots \psi(\pi) \ldots) = \int_{x \in X} \psi(x)\,|\delta_x\rangle\, dx \tag{14}$$

   Therefore, we have defined the mapping $\psi \mapsto |\psi\rangle$.

2. To define $x \mapsto |x\rangle$, we can also analogously say that each $x \in X$ can also be represented as

$$|x\rangle = (\ldots, 0, \ldots, 0, 1, 0, \ldots, 0, \ldots) = |\delta_x\rangle \tag{15}$$

   like an uncountable version of one-hot encoding (each real number gets mapped to its own dimension). So,

$$|\psi\rangle = \int_{x \in X} \psi(x)\,|x\rangle\, dx \tag{16}$$

Therefore, we can represent

$$\langle x|\psi\rangle = \langle \delta_x|\psi\rangle = \int_{x' \in X} \psi(x')\,\delta_{x'}(x)\, dx' = \psi(x) \tag{17}$$

We can verify that this is bilinear since

$$(\psi + \phi)(x) = \langle \delta_x|\psi + \phi\rangle = \langle \delta_x|\psi\rangle + \langle \delta_x|\phi\rangle = \psi(x) + \psi(x) \tag{18}$$

and

$$\langle \delta_x + \delta_y|\psi\rangle = \langle \delta_x|\psi\rangle + \langle \delta_y|\psi\rangle = \psi(x) + \psi(y) \tag{19}$$

Note that in here $\delta_x + \delta_y \neq \delta_{x+y}$ since we are not talking addition in the reals. Rather,

$$(\delta_x + \delta_y)(q) = \begin{cases} 1 & \text{if } x = q \text{ or } y = q \\ 0 & \text{if else} \end{cases} \tag{20}$$

1.3 Evolution

1.4 Composite Systems

2 Measurements

When we have a (classical or quantum mechanical) system, there is a configuration space M that describes the complete state of the system. In classical mechanics, this is usually the space $\mathcal{M} = \mathcal{X} \times \mathcal{P}$ where $\mathcal{X}$ is the position and $\mathcal{P}$ is the momentum. In quantum mechanics, $\mathcal{M} = L^2(\mathcal{X})$ is the space of wavefunctions defined over $\mathcal{X}$.

> **Definition 2.1 (Observable in Classical Mechanics)**
>
> In classical mechanics, an observable is a function $Q : \mathcal{M} \to \mathbb{R}$ of some dynamical quantity that you want to measure from the system.

> **Example 2.1 ()**
>
> Some examples are
>
> 1. The position of the vector is simply the x value of the tuple.
>
> 2. The kinetic energy is $K(x, p, t) = \frac{1}{2m} p^2$.
>
> 3. The potential energy is $U(x, p, t) = U(x)$, which is dependent only on x for conservative systems.
>
> 4. The Hamiltonian is defined as the total (kinetic plus potential) energy, defined $H(x, p, t) = \frac{1}{2m} p^2 + V(x)$.

In quantum mechanics, an observable is similar as in you want to define the specific type of quantity you want to measure, but now this is probabilistic: we now have a random variable Q. Broadly speaking, if we measure this observable, then it will cause Q to realize into some real number q.

> **Definition 2.2 (Observable in Quantum Mechanics)**
>
> However, there is a slight catch: $\mathcal{X}$ must be some. A **quantum mechanical observable** Q is a real valued random variable, and its **associated Hamiltonian** (often called the observable) is a Hermitian operator $\hat{Q} : L^2(\mathcal{X}) \to L^2(\mathcal{X})$.

Two things to note: First, the output of the operator doesn't necessarily have to be a wavefunction (doesn't need to be normalized, look at the position operator), but it should be normalizable.

Second, the Hermitian operator $\hat{H}$ is relevant since it turns out that its set of eigenvalues $\mathcal{E} = \{E\}$ of $\hat{Q}$ determines the support of Q, and the set of eigenvectors $\{\psi_E\}$ will determine the probability distribution of Q, which will be elaborated soon. Note that the Hermitian constraint happens because the outcomes must be real, since they are physical measurements that must be compared. Another technicality is that not all Hermitian operators are proper observables in the infinite dimensional case, but this is a technicality and can be ignored for now.

> **Example 2.2 ()**
>
> Some examples of quantum mechanical observables can be obtained by the canonical substitution
>
> $$p \to \frac{\hbar}{i} \frac{\partial}{\partial x} \tag{21}$$
>
> from the classical observables.

1. Position: $\hat{Q} : \psi \to x\psi$

2. Momentum: $\hat{Q} : \psi \to \frac{\partial}{\partial x}\psi$

3. Kinetic energy: $\hat{Q} : \psi \to \frac{1}{2m}\frac{\hbar^2}{-1}\frac{\partial^2}{\partial x^2}\psi$

4. Hamiltonian, or total energy, denoted with a special character $\hat{H} : \psi \mapsto \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V$.

Now an observable simply describes the type of quantity that you want to measure. A *measurement*, on the other hand, is an action on the system, as if we would measure with a ruler or an instrument, and therefore is different from an observable. Once you do a measurement, the random variable Q realizes onto a certain number $q \in \mathbb{R}$, and the state ψ of the system is affected by a collapse of the wavefunction ψ into the Delta function centered at q. Therefore, there are two evolutionary processes: the unitary evolution of ψ according to Shrodinger's equation and measurement operators.

Definition 2.3 (Measurement)

A **measurement** of a system in state ψ, with respect to observable Q, is defined by two things.

1. First, a measurement will cause a realization of Q into a certain real quantity $E \in \mathcal{E}$. The probability distribution of Q is determined by taking the state ψ, expanding it into a linear combination of the eigenvectors

$$\psi = \int_{E \in \mathcal{E}} \alpha_E \psi_E \, dE \tag{22}$$

and finally taking $|\alpha_E|^2$ to be the probability (PMF for countable $\mathcal{E}$ and PDF for uncountable).

2. While q is realized, the wavefunction ψ will collapse into δ_q.

Now let's put these things all together in one formal process.

1. You have a system in a state Ψ and want to measure it in some way. But measuring something does not make sense without knowing what quantity you want to measure, so you choose an observable $\hat{Q}$, a Hermitian operator that corresponds to the value that you want to measure.

2. As soon as you choose this operator, by the spectral theorem it has some set of eigenvalues $\mathcal{E} \subset \mathbb{R}$ and corresponding orthonormal eigenfunctions $\{\psi_E\}_{E \in \mathcal{E}}$, which may be finite, countable, or uncountable. It turns out that the measurement outcome *must* be one of these eigenvalues, and it is random, so you're essentially drawing from this set an E with some random variable Q. This makes $\mathcal{E}$ a probability space with some measure μ.

3. To determine the actual probabilities of landing on each eigenvalue, you must first take the state Ψ and expand it out into the eigenbasis.

$$\Psi = \sum_{E \in \mathcal{E}} \alpha_E \psi_E \text{ or } \Psi = \int_{E \in \mathcal{E}} \alpha_E \psi_E \, dE \tag{23}$$

and the amplitude squared of the coefficients gives you the respective probabilities.

4. Once you measure and get result E at time t, the wavefunction $\psi(t)$ collapses onto the delta at E (similar to a realization of a random variable) and immediately begins to "smear" out again. If you measure really soon after, you can still sample $\psi(t + \delta t)$ with approximately probability 1 at the point you just measured, since the state is so concentrated there and therefore has a very high coefficient corresponding to ψ_E when you expand it out.

Example 2.3 (Measurement of the Position Observable)

We have said that given a wavefunction $\Psi(t)$, the probability distribution of the particle's position is determined by the PDF at time t is $p_t(x) = |\Psi(x,t)|^2$. This is consistent with what we have described here if we talk about the position observable. Say that I have a wavefunction Ψ. I want to measure position with the observable $\hat{Q} : \psi \mapsto x\psi$. We first have to find the eigendecomposition of $\hat{Q}$. It turns out that the set of eigenfunctions is the set of all Dirac delta functions indexed by $\mathbb{R}$.

$$\delta_k(x) = \begin{cases} 1 & x = k \\ 0 & x \neq k \end{cases} \tag{24}$$

and we can see that every k is an eignvalue with eigenfunction δ_k since

$$(\hat{Q}\delta_k)(x) = (k\delta_k)(x) = \begin{cases} k & x = k \\ 0 & x \neq k \end{cases} = k\delta_k \tag{25}$$

So the eigendecomposition is

$$\hat{Q} = \sum_{x \in X} k\delta_k \otimes \delta_k^* \text{ or } \hat{Q} = \int_{k \in \mathbb{R}} k\delta_k \otimes \delta_k^* \, dx \tag{26}$$

and so ψ can be decomposed as the uncountable linear combination of the eigenfunctions

$$\psi = \sum_x \alpha_x \delta_x \implies \mathbb{P}(Q = x) = |\alpha_x|^2 = |\psi(x)|^2$$

$$\psi = \int_{x \in \mathbb{R}} \alpha_x \delta_x \, dx \implies p_Q(x) = |\alpha_x|^2 = |\psi(x)|^2$$

Example 2.4 (Measurement of the Momentum Observable)

Say that I have a wavefunction ψ. I want to measure momentum with the observable Q and corresponding operator $\hat{Q} : \psi \to \frac{\hbar}{i} \frac{\partial}{\partial x} \psi$. We can verify that the spectrum consists of the eigenvalue/eigenfunction pair

$$\left\{ p, \psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \right\}_{p \in \mathbb{R}} \tag{27}$$

and so

$$\hat{Q} = \int_{p \in \mathbb{R}} p\psi_p \otimes \psi_p^* \, dp \tag{28}$$

and so we can decompose

$$\psi = \int_{x \in \mathbb{R}} \alpha_x \psi_x \, dx \implies p_Q(x) = |\langle \psi_x | \psi \rangle|^2 \tag{29}$$

Some properties of the these measurements are as follows.

Exercise 2.1 ()

Compute the commutator of the position and momentum observables (we will need this for the uncertainty principle). What does the result mean?

2.1   Eigenvalues are just Symbols

Note that the probability measure of the measurement outcome is not dependent on the actual value of the eigenvalues themselves. We can think of them as some sort of symbol that we use to index the eigenfunctions. This is because the probability measure is determined by the coefficients of the expansion of the wavefunction into the eigenbasis, and not the eigenvalues themselves. Sometimes, there are natural quantities where it may be easy to assign the eigenvalues. The position operator in $\mathbb{R}^n$ gives a natural way since it has a natural ordering, but the following example does not.

> **Example 2.5 (Position Measurement of Qubit)**
>
> Given a qubit in state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$, let us measure it with two observables.
>
> 1. The observable
> $$\hat{Q} = 1 \cdot |0\rangle \langle 0| + 0 \cdot |1\rangle \langle 1| \qquad (30)$$
>   gives $\mathbb{P}(Q = 1) = |\alpha|^2$ and $\mathbb{P}(Q = 0) = |\beta|^2$.
>
> 2. The observable
> $$\hat{Q} = 1 \cdot |0\rangle \langle 0| + (-1) \cdot |1\rangle \langle 1| \qquad (31)$$
>   gives $\mathbb{P}(Q = 1) = |\alpha|^2$ and $\mathbb{P}(Q = -1) = |\beta|^2$.
>
> Either way, these eigenvalues are just symbols that we use to index the eigenfunctions, and it does not matter whether we use 0 or -1 to represent the collapse to the $|1\rangle$ state. The state of the system after measurement is also not affected, since it's only dependent on the eigenvector, not the value.

Another observable is the spin, which was first found by the Stern-Gerlach experiment.

2.2   Basis Orthogonality

Now let's talk a little about practical methodology. It turns out that in practice, preparing states of a quantum system is quite hard, and in other cases we may not even know what the state can be. This can be formulated in the language of density operators, but for now, let's talk about the problem of distinguishing quantum states. We start off with an exercise.

> **Exercise 2.2 (Neilson 2.58)**
>
> If the state vector $|\psi\rangle$ is an eigenvector of the observable operator $\hat{Q}$ with eigenvalue, say E, then what can we say about the measurement outcome? Furthermore, what is the average observed value of M, and the standard deviation?

> **Solution 2.1**
>
> The measurement outcome will always be E with probability 1.

Say that we are given some state that is one of $\{|\psi\rangle_i\}_{i=1}^n$ and we must construct some way to find which state it comes from. We would like to find some measurement operator $\hat{Q}$, measure the system, and deduce which state it came from based on what value Q realized on. Can we do this perfectly? It turns out that we can under certain conditions.

> **Theorem 2.1 (Distinguishing Quantum States)**
>
> Given a set of states $\{|\psi\rangle_i\}_{i=1}^n$, there exists a measurement operator that can distinguish them perfectly if and only if they are orthogonal. If they are not orthogonal, then there exists not measurement operator.

> **Proof.**
>
> Assume that the states are orthogonal. Then, we can construct the measurement operator
>
> $$\hat{Q} = \sum_{i=1}^{n} \lambda_i \ket{\psi_i} \bra{\psi_i} \tag{32}$$
>
> and so if Q realizes onto λ_i, then we know that the state was $\ket{\psi_i}$ with probability 1.

If they are not orthogonal, then we can't construct an orthonormal basis and cannot construct a measurement operator. Generally, if we have two non-orthogonal states $\ket{\psi}$ and $\ket{\phi}$, say that we want to construct $\hat{Q}$ with just two eigenvectors $\ket{u}, \ket{v}$ (the rest are irrelevant).

1. If none of its eigenvectors contain either $\ket{\psi}$ or $\ket{\phi}$. Then

$$\mathbb{P}(\ket{u} \text{ after } | \ket{\psi} \text{ before}) = |\braket{u|\psi}|^2 > 0 \tag{33}$$
$$\mathbb{P}(\ket{v} \text{ after } | \ket{\psi} \text{ before}) = |\braket{v|\psi}|^2 > 0 \tag{34}$$
$$\mathbb{P}(\ket{u} \text{ after } | \ket{\phi} \text{ before}) = |\braket{u|\phi}|^2 > 0 \tag{35}$$
$$\mathbb{P}(\ket{v} \text{ after } | \ket{\phi} \text{ before}) = |\braket{v|\phi}|^2 > 0 \tag{36}$$

   and so no matter whatever eigenvector that the system realizes onto, there is a nonzero probability of it coming from either. So from using Bayes rule, we can't deduce anything.

2. Therefore, we must choose the eigenvectors to be the states themselves. Since they are not orthogonal, we can do it for one state, and let the eigenvectors be $\ket{\psi}, \ket{\psi^\perp}$.

$$\mathbb{P}(\ket{\psi} \text{ after } | \ket{\psi} \text{ before}) = 1 \tag{37}$$
$$\mathbb{P}(\ket{\psi^\perp} \text{ after } | \ket{\psi} \text{ before}) = 0 \tag{38}$$
$$\mathbb{P}(\ket{\psi} \text{ after } | \ket{\phi} \text{ before}) = |\braket{\psi|\phi}|^2 \tag{39}$$
$$\mathbb{P}(\ket{\psi^\perp} \text{ after } | \ket{\phi} \text{ before}) = 1 - |\braket{\psi|\phi}|^2 \tag{40}$$

   This is slightly better than before, since now when the system realizes onto $\ket{\psi^\perp}$, we can actually deduce where it came from. Using Bayes rule,

$$\mathbb{P}(\ket{\phi} \text{ bef } | \ket{\psi^\perp} \text{ aft}) = \frac{\mathbb{P}(\ket{\psi^\perp} \text{ aft } | \ket{\phi} \text{ bef})\mathbb{P}(\ket{\phi} \text{ bef})}{\mathbb{P}(\ket{\psi^\perp} \text{ aft})} \tag{41}$$

$$= \frac{\mathbb{P}(\ket{\psi^\perp} \text{ aft } | \ket{\phi} \text{ bef})\mathbb{P}(\ket{\phi} \text{ bef})}{\mathbb{P}(\ket{\psi^\perp} \text{ aft } | \ket{\phi} \text{ bef})\mathbb{P}(\ket{\phi} \text{ bef}) + \mathbb{P}(\ket{\psi^\perp} \text{ aft } | \ket{\psi} \text{ bef})\mathbb{P}(\ket{\psi} \text{ bef})} \tag{42}$$

$$= \frac{(1 - |\braket{\psi|\phi}|^2)\mathbb{P}(\ket{\phi} \text{ bef})}{(1 - |\braket{\psi|\phi}|^2)\mathbb{P}(\ket{\phi} \text{ bef}) + 0 \cdot \mathbb{P}(\ket{\psi} \text{ bef})} = 1 \tag{43}$$

   but the problem comes from when the state is initialized at $\ket{\phi}$. Then, we have

$$\mathbb{P}(\ket{\psi} \text{ before } | \ket{\psi} \text{ after}) = \frac{\mathbb{P}(\ket{\psi} \text{ before}, \ket{\psi} \text{ after})}{\mathbb{P}(\ket{\psi} \text{ after})} \tag{44}$$

$$= \frac{\mathbb{P}(\ket{\psi} \text{ aft } | \ket{\psi} \text{ bef})\,\mathbb{P}(\ket{\psi} \text{ bef})}{\mathbb{P}(\ket{\psi} \text{ aft } | \ket{\psi} \text{ bef})\mathbb{P}(\ket{\psi} \text{ bef}) + \mathbb{P}(\ket{\psi} \text{ aft } | \ket{\phi} \text{ bef})\mathbb{P}(\ket{\phi} \text{ bef})} \tag{45}$$

$$= \frac{\mathbb{P}(\ket{\psi} \text{ bef})}{\mathbb{P}(\psi \text{ bef}) + |\braket{\psi|\phi}|^2\mathbb{P}(\phi \text{ bef})} \tag{46}$$

   which is neither 0 nor 1, given that our prior distributions are non-degenerate.

Therefore, there is no perfect way to distinguish non-orthogonal states.

2.3   Expectation and Variance of Observables

Sometimes, we are interested in specific statistical properties of these distributions of observables. In this case, we can define the expectation and variance of the operators.

> **Definition 2.4 (Expectation of Operator)**
>
> The expectation value of the operator (as a function of time) is
>
> $$\langle Q \rangle(t) := \int \Psi(x,t)^*(\hat{Q}\Psi)(x,t)\,dx = \langle \Psi | \hat{Q}\Psi \rangle(t) \tag{47}$$

Since physical observables must be real (to see why, look here), it must be the case that

$$\langle \Psi | \hat{Q}\Psi \rangle = \langle \Psi | \hat{Q}\Psi \rangle^* = \langle \hat{Q}\Psi | \Psi \rangle \tag{48}$$

for all vectors $|\Psi\rangle$, so it must follow that $\hat{Q}$ must be a Hermitian operator.

> **Example 2.6 (Expectatation of Position Observable)**
>
> The expected position of the particle can be written
>
> $$\langle \mathbf{x} \rangle := \int \Psi(x,t)^*(\hat{Q}\Psi)(x,t)\,dx = \int \Psi(x,t)^*(x\Psi)(x,t)\,dx = \int x\Psi(x,t)^*\Psi(x,t)\,dx = \int x|\mathbf{\Psi}(\mathbf{x},t)|^2\,d\mathbf{x} \tag{49}$$
>
> which is consistent with the original definition of expectation in probability.

> **Theorem 2.2 ()**
>
> The velocity of the expected value can be evaluated to:
>
> $$\frac{d}{dt}\langle \mathbf{x} \rangle = -\frac{i\hbar}{m}\int \mathbf{\Psi}^*\frac{\partial \mathbf{\Psi}}{\partial x}\,dx \tag{50}$$

> **Proof.**
>
> We can see that using integration by parts in the penultimate step,
>
> $$\frac{d\langle x \rangle}{dt} = \int x\frac{\partial}{\partial t}|\mathbf{\Psi}|^2\,dx \tag{51}$$
>
> $$= \frac{i\hbar}{2m}\int x\frac{\partial}{\partial x}\left(\mathbf{\Psi}^*\frac{\partial \mathbf{\Psi}}{\partial x} - \frac{\partial \mathbf{\Psi}^*}{\partial x}\mathbf{\Psi}\right) \tag{52}$$
>
> $$= -\frac{i\hbar}{2m}\int \left(\mathbf{\Psi}^*\frac{\partial \mathbf{\Psi}}{\partial x} - \frac{\partial \mathbf{\Psi}^*}{\partial x}\mathbf{\Psi}\right)dx \tag{53}$$
>
> $$= -\frac{i\hbar}{m}\int \mathbf{\Psi}^*\frac{\partial \mathbf{\Psi}}{\partial x}\,dx \tag{54}$$

Note that this does not mean that if we take a single particle and measure it multiple times, we will get the expected value, since it will first evolve and second we will get the exact same measurement. Rather, if we take an *ensemble* of particles all in the same state $\mathbf{\Psi}$ and measure them all at once, then the histogram of measurements can be used as an unbiased estimator of $\langle x \rangle$. Therefore, we can define momentum as the following.

Definition 2.5 (Momentum)

The momentum is defined as

$$\langle p \rangle := m \frac{d\langle x \rangle}{dt} = -i\hbar \int \left(\mathbf{\Psi}^* \frac{\partial \mathbf{\Psi}}{\partial x} \right) dx \tag{55}$$

Example 2.7 (Expectation of Total Energy)

The expectation value of the total energy is

$$\langle H \rangle = \int \psi^* \hat{H} \psi \, dx = E \int |\psi|^2 \, dx = E \tag{56}$$

Theorem 2.3 ()

Observable quantities, $Q(x, p, t)$, are represented by Hermitian operators $\hat{Q}(x, \frac{\hbar}{i} \frac{\partial}{\partial x}, t)$. The expectation value of Q in the state Ψ at time t is $\langle \Psi | \hat{Q} \Psi \rangle(t)$.

Definition 2.6 (Variance)

The variance of the observable Q is defined as

$$\sigma_Q^2 := \langle \Psi | \hat{Q}^2 \Psi \rangle - \langle \Psi | \hat{Q} \Psi \rangle^2 \tag{57}$$

and the standard deviation is

$$\sigma_Q = \Delta(Q) := \sqrt{\sigma_Q^2} \tag{58}$$

2.4   Heisenberg Uncertainty Principle

The Heisenberg uncertainty principle is often very misunderstood. It's blown up in the popular literature as some mysterious force of quantum mechanics, but under the things we have constructed, it is simply an inequality no more sophisiticated than the Cauchy-Schwarz inequality. It is also not the same thing as the observer effect.

Theorem 2.4 (Heisenberg Uncertainty Principle)

Suppose that $\hat{A}$ and $\hat{B}$ are Hermitian operators corresponding to observables A and B. Then, the following inequality holds.

$$\sigma_A \sigma_B \geq \frac{|\langle \Psi \mid [\hat{A}, \hat{B}] \mid \Psi \rangle|}{2} \tag{59}$$

Proof.

Given $\hat{A}, \hat{B}$, by linearity we can write

$$|\langle \Psi \mid [\hat{A}, \hat{B}] \mid \Psi \rangle|^2 + |\langle \Psi \mid \{\hat{A}, \hat{B}\} \mid \Psi \rangle|^2 = 4|\langle \Psi \mid \hat{A}\hat{B} \mid \Psi \rangle|^2 \tag{60}$$

By Cauchy-Schwarz,

$$|\langle \Psi \mid \{\hat{A}, \hat{B}\} \mid \Psi \rangle|^2 \leq \langle \Psi \mid \hat{A}^2 \mid \Psi \rangle \langle \Psi \mid \hat{B}^2 \mid \Psi \rangle \tag{61}$$

and substituting this to the first equation gives us

$$|\langle \Psi \mid [\hat{A}, \hat{B}] \mid \Psi \rangle|^2 \leq 4\langle \Psi \mid \hat{A}^2 \mid \Psi \rangle \langle \Psi \mid \hat{B}^2 \mid \Psi \rangle \tag{62}$$

> and substituting $A = A - \langle A \rangle$ and $B = B - \langle B \rangle$ gives us the desired result.

3 Time Independent Schrödinger Equation

Let's revisit Shrodinger's equation again.

$$i\hbar \frac{\partial \boldsymbol{\Psi}}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \boldsymbol{\Psi}}{\partial \mathbf{x}^2} + V\boldsymbol{\Psi} \tag{63}$$

If we pay attention to the right hand side, the operator acting on $\boldsymbol{\Psi}$ is really the Hamiltonian, and the left hand side is also another linear operator, say A. Then, the equation reduces to a linear equation

$$A\psi = \hat{H}\psi \tag{64}$$

which we solve for ψ.

Let's talk about solving the Schrödinger function itself by looking at the simple case when the Schrödinger equation is time independent and we can thus solve it by the method of separation of variables. That is, we look for solutions of the form

$$\boldsymbol{\Psi}(x, t) = \psi(x)f(t) \tag{65}$$

We have

$$\frac{\partial \Psi}{\partial t} = \psi\frac{df}{\partial t}, \quad \frac{\partial^2}{\partial x^2} = \frac{d^2\psi}{dx^2} = \frac{d^2\psi}{dx^2}f \tag{66}$$

and the Schrödinger equation becomes

$$i\hbar\psi\frac{df}{dt} = -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}f + V\psi f \tag{67}$$

and dividing by f on both sides gives us

$$i\hbar\frac{1}{f}\frac{df}{dt} = -\frac{\hbar^2}{2m}\frac{1}{\psi}\frac{d^2\psi}{dx^2} + V \tag{68}$$

The LHS as a function of t alone and RHS as a function of x alone. This is only possible if both sides are constant (since we can simply change t or x to get different values). Now setting the LHS as the constant E, we can rewrite the partial differential equation as a system of two ODEs.

$$\frac{df}{dt} = -\frac{iE}{\hbar}f \tag{69}$$

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi = E\psi \tag{70}$$

The first equation is easy to solve since we can just integrate, which gives $f(t) = e^{-iEt/\hbar}$. The second equation is simply the Hamiltonian observable $\hat{H}$ acting on ψ, and so it reduces to an eigenvalue equation.

$$\hat{H}\psi = E\psi \tag{71}$$

Therefore, once we solve by getting the eigenfunction ψ_E corresponding to this eigenvalue E, the final solution of the time-independent Shrodinger equation is of form

$$\boldsymbol{\Psi}(x, t) = \psi(x)e^{-iEt/\hbar} \tag{72}$$

Definition 3.1 (Stationary State)

A wavefunction $\boldsymbol{\Psi}(x,t)$ is a **stationary state** if its corresponding probability density does not depend on t. That is,
$$|\boldsymbol{\Psi}(x,t)|^2 = |\boldsymbol{\Psi}(x,t')|^2 \text{ for all } t,t' \in \mathbb{R} \tag{73}$$
This means that the expected position is always constant since the probability density function never changes with time.

It turns out that solutions to separable equations all stationary states since we have
$$|\boldsymbol{\Psi}(x,t)|^2 = \boldsymbol{\Psi}^* \boldsymbol{\Psi} = \psi^* e^{+iEt/\hbar} \psi e^{-iEt/\hbar} = |\psi(x)|^2 \tag{74}$$

4 Other

Definition 4.1 (Pauli Matrices)

The **Pauli matrices** are defined
$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Lemma 4.1 (Decomposition of Pauli Matrices)

The Pauli matrices are self-adjoint (Hermitian) and unitary.

1. They have the outer product decomposition of:
$$X = |1\rangle \langle 0| + |0\rangle \langle 1|, Y = i|1\rangle \langle 0| - i|0\rangle \langle 1|, Z = |0\rangle \langle 0| - |1\rangle \langle 1|$$

2. They have an eigendecomposition of:
$$X = |+\rangle \langle +| - |-\rangle \langle -|, Y = i|+\rangle \langle +| - i|-\rangle \langle -|, Z = |0\rangle \langle 0| - |1\rangle \langle 1| \tag{75}$$

Theorem 4.1 ()

These three matrices are very important because it turns out that
$$\frac{1}{2}iX, \frac{1}{2}iY, \frac{1}{2}iZ$$
forms the basis for the Lie algebra $\mathfrak{u}(2)$, which exponentiates to the unitary group U(2). Therefore, by exponentiating each Pauli matrix, we have

$$e^{-i\beta X/2} = \cos\frac{\beta}{2}I - i\sin\frac{\beta}{2}X = \begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix},$$
$$e^{-i\gamma Y/2} = \cos\frac{\gamma}{2}I - i\sin\frac{\gamma}{2}Y = \begin{pmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{pmatrix},$$
$$e^{-i\delta Z/2} = \cos\frac{\delta}{2}I - i\sin\frac{\delta}{2}Z = \begin{pmatrix} e^{-i\delta/2} & 0 \\ 0 & e^{i\delta/2} \end{pmatrix}.$$

and so every rotation matrix $U \in U(2)$, which represents single qubit operations, can be decomposed as the following products for real values of β, γ, δ:

$$U = \begin{pmatrix} \cos\frac{\beta}{2} & -i\sin\frac{\beta}{2} \\ -i\sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix} \begin{pmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{pmatrix} \begin{pmatrix} e^{-i\delta/2} & 0 \\ 0 & e^{i\delta/2} \end{pmatrix}.$$

Theorem 4.2 (Commutation and Anti-Commutation)

The Pauli matrices satisfy the following properties.

1. Commutation properties:

$$[X, Y] = 2iZ, \tag{76}$$
$$[Y, Z] = 2iX, \tag{77}$$
$$[Z, X] = 2iY. \tag{78}$$

2. Anticommutation properties:

$$\{X, Y\} = 0, \tag{79}$$
$$\{Y, Z\} = 0, \tag{80}$$
$$\{Z, X\} = 0. \tag{81}$$