Quantum Mechanics

Lecture Notes

Sergio Charles

Contents

Chapter 1

Introduction

"God does not play dice"

Albert Einstein, 1926

In 1926, Albert Einstein wrote a letter to Max Born in which be stated "the theory produces a good deal but hardly brings us closer to secret of the Old One. I am at all events convinced He does not play dice." A dissent to Born's probabilistic approach to quantum mechanics, Einstein thought such an incongruence with classical theory could not be possible. In these notes, I have followed the Modern Physics Quantum Mechanics Lectures of Leonard Susskind from 2008 at Stanford [1]. The fundamentally probabilistic and discrete nature of reality is perplexing. Herein, I hope to un-confuse myself by systematically presenting quantum mechanics with mathematical rigor.

1.1 The Double Slit Experiment

Suppose God randomly hits the moon. The slight deviations cause slight changes in energy, which eventually builds up into statistical randomness in energy; however, this would violate the conservation of energy in classical physics. In quantum mechanics, such randomness still conserves energy.

Consider a single slit and a laser that shoots out a photon. If we use a wall to detect where the photon lands, we'd expect to see a normal distribution, as illustrated in Figure 1.1.

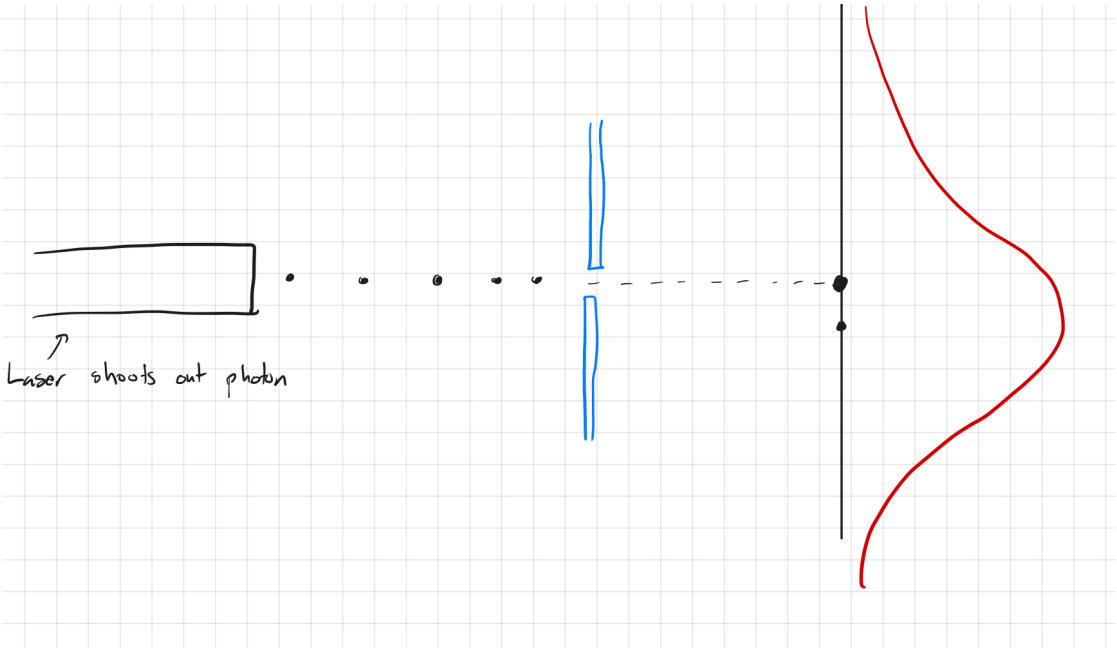


Figure 1.1: Single slit experiment.

Now, suppose we add a second slit, as shown in Figure 1.2 such that photons come through one at a time sparsely. Namely, what happens to the next photon doesn't dependent on the current photon in a way that is statistically independent. If we open both holes, in classical physics, the probability distribution would be the sum of the two distributions seen when only one slit is open. That is, it would simply be the sum of the two Gaussian distributions. However, in a real experiment, we see an interference pattern. There will be places, as shown in the diagram, where no photons get to despite there being such a long period of time between the firing of photons. That is, we observe a strange phenomenon, whereby *photons are aware of other photons despite being fired independently.*

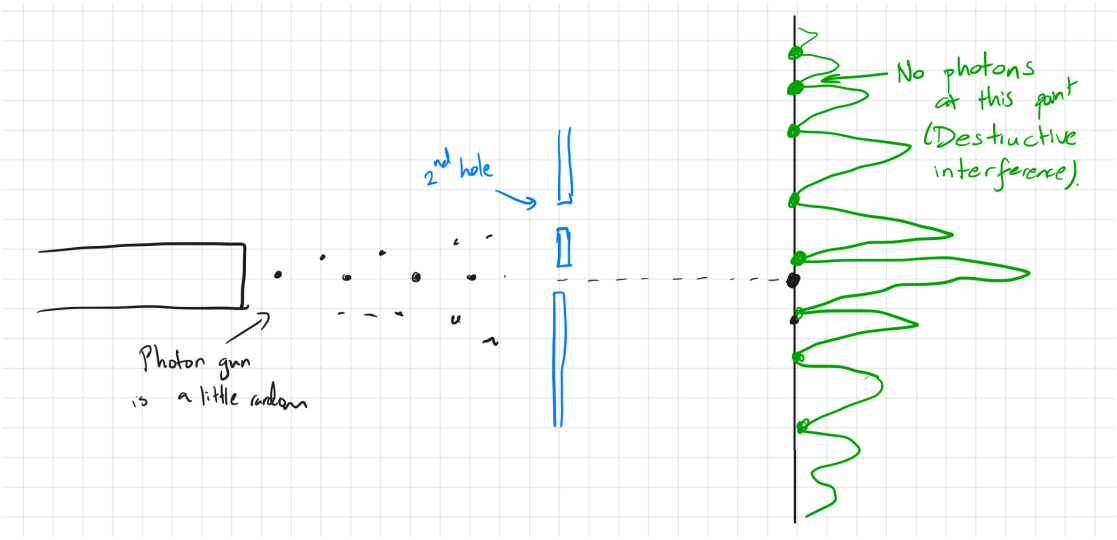
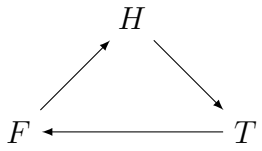
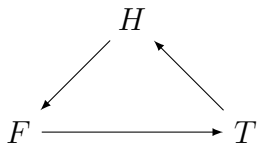


Figure 1.2: Double slit experiment.

Next, consider a very simple classical systems where the laws of physics are specified by the rule $H \to T \to F \to H$:



The reverse law of physics is $F \to T \to H \to F$:



Definition 1 ***Determinism*** *means if we run the original law of physics for many steps, i.e. cycles of the triangle, stop it, and let it evolve with the reversed law backwards for the same time, it will come back to where it started. In particular, no information is lost, as it is deterministic. However, if there was any randomness injected in the transition, like in a Markov Chain on a finite state space, it would destroy the conservation of information.*

Consider the one slit experiment again. Suppose we run it forward in time with the laws of physics and then run it backward by reversing time. Note, we assume we do not have a screen on which to observe or detect the arrival of photons. In classical mechanics, we will observe it running backward along the same trajectory because observing it shouldn't change the state. However, in quantum mechanics, if you do anything to detect the photon and then run the law backward, the probabilistic nature is compounded and the test of reversibility will fail.

Remark 1 *Note, the interference pattern seen in Figure 1.2 only holds if nothing records where the particle goes through. If you record the way the particle went through, then the interference pattern is destroyed and the probabilities simply add like in classical physics–this is known as* ***wave function collapse****.*

1.2 The Uncertainty Principle

There is a fundamental obstruction to observing both position and momentum simultaneously. From the classical theory of light, we know the energy of a photon is:

$$E = hf = \hbar\omega \tag{1.1}$$

where h is Planck's constant, f is the frequency of light (in units of cycles/sec or Hertz), $\omega = 2\pi f$ is the angular frequency and $\hbar = \frac{h}{2\pi}$. Furthermore, Einstein tells us beams of light have not just energy, but also momentum such that:

$$E = cp \tag{1.2}$$

where c is the speed of light and $p = mc$ is momentum. Note, in the non-relativistic Newtonian setting:

$$\begin{aligned} E &= \frac{p^2}{2m} \\ &= \frac{1}{2}p\frac{p}{m} \\ &= \frac{1}{2}pv. \end{aligned} \tag{1.3}$$

Re-writing $E = pc$, we find:

$$p = \frac{E}{c} = \frac{hf}{c}. \tag{1.4}$$

By **wave-particle duality**, we can view the photon as a wave moving with the speed of light. If it has wavelength λ, the wave moves λ meters in one cycle. A period of the cycle is $1/f$, i.e. the time it takes to move distance λ. Thus it takes $t = 1/f$ time to go λ distance. Since it's traveling at the speed of light,

$$c = \frac{\lambda}{t} = \lambda f. \tag{1.5}$$

Thus $f = c/\lambda$ so the momentum can be written as:

$$p = \frac{hf}{c} = \frac{h}{c}\frac{c}{\lambda} = \frac{h}{\lambda}, \tag{1.6}$$

which is known as **de Broglie's equation**. It states that momentum and wavelength are inverse to each other.

Suppose we want to take a photograph of the electron and wait it to be non-fuzzy on a scale of Δx. To form such an image, you have to use wavelengths that are shorter than Δx. That is we require: $\lambda < \Delta x$. Recall de Broglie's equation: $p = h/\lambda$. Thus, if we want Δx to be small, λ needs to be smaller, which implies p must be larger.

So if we try to measure an electron, we bombard it with a high momentum photon, which will knock it off into a random direction. Thus, immediately after we try to measure it's position, the electron is kicked hard and the momentum becomes very uncertain. *Measuring position necessarily imparts a random momentum kick to the particle.* That is, there is no such thing as gentle determination in quantum mechanics.

Remark 2 *We have used the fact that light comes in discrete, indivisible photons (quanta). There is a minimal amount of momenta that is compatible with that wavelength, namely one photon.*

In classical physics, energy does not come in discrete multiples of some basic unit. In quantum mechanics, the energy of a light wave comes in discrete packets.

Next, suppose we try to measure a particle's velocity. We use a long wavelength so as not displace the electron, since as we previously determined, if λ is too small, it will kick the electron with large momentum. Hence, we observe $x \pm \lambda$ where $\pm\lambda$

is the uncertainty in position. This will change the velocity by a small amount and, then, at a later time, we discover the particle at position:

$$x \pm \lambda + vt \tag{1.7}$$

where t is the time between measurements. The distance the particle moves is:

$$vt \pm \lambda \sim d \tag{1.8}$$

so $v \pm \lambda/t \sim d/t$, where λ/t is the disorder in measurement of velocity. If we wait long enough, this sloppiness can be made small.

To be more precise, suppose we measure the position of a particle with some poor accuracy, i.e. within a fuzzy region of radius Δ as shown in Figure 1.3. Then suppose we measure the position again after time t. The distance of separation is $L \pm \Delta$. It is moving at a velocity of:

$$\frac{L \pm \Delta}{t} = \frac{L}{t} \pm \frac{\Delta}{t} \tag{1.9}$$

One may think by making t large, you can get an accurate measurement of velocity and, hence, momentum. However, to measure the first location within Δ accuracy, we will hit the particle with a photon of wavelength $\lambda < \Delta$ and momentum $p = h/\lambda$, which kicks the velocity by $\hbar/\Delta$. Thus, we may think we have a true measurement of velocity, but it is not.

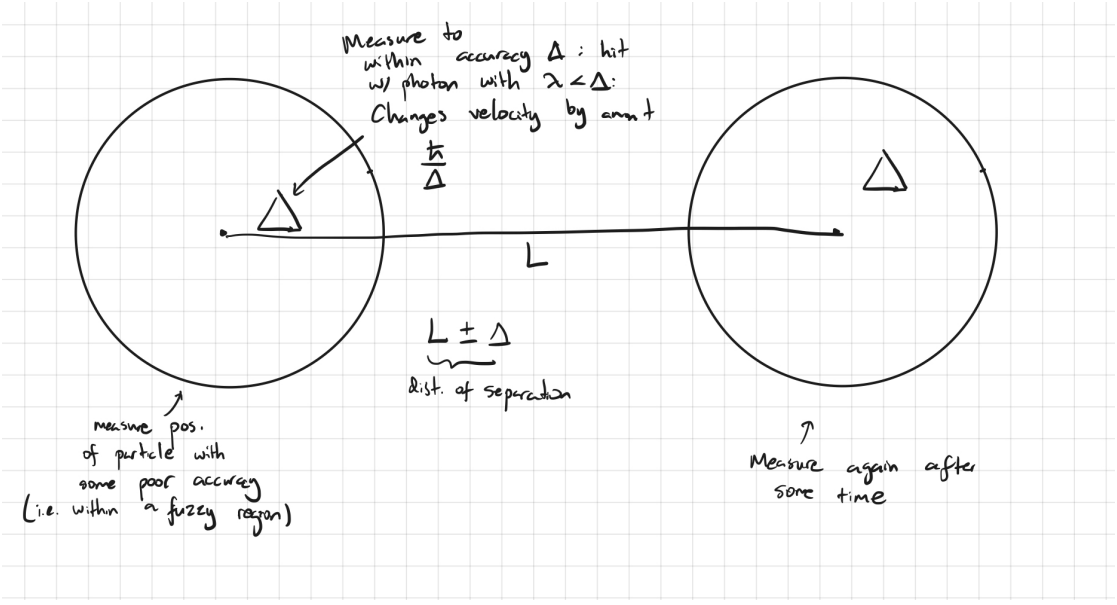


Figure 1.3: Measure particle within fuzzy region Δ, then measured the location again after time t.

1.3 State Space

In classical physics, the phase space is the set of points of the system, e.g. $\{H, T\}$ representing Heads and Tails in a coin toss. The transition between states in the

system describe the dynamics. As an example, momentum and position (p, x) can be thought of as points in a phase space of a set of points. The takeaway is that states in classical physics are points in a set. However, *in quantum mechanics, states are not sets, but vectors in a vector space.*

Definition 2 *A **vector space** V (elements of V are **vectors**) over a field $\mathbb{F}$ (elements of $\mathbb{F}$ are called scalars) is equipped with vector addition and scalar multiplication and is closed under these operations:*

- $|a\rangle + |b\rangle = |c\rangle \in V$ *for all* $|a\rangle, |b\rangle \in V$.
- $\alpha |a\rangle = |b\rangle \in V$ *for all* $\alpha \in \mathbb{F}$ *and* $|a\rangle \in V$.

Together, these two properties imply closure under addition and scalar multiplication, i.e. linear combinations. If $|a\rangle, |b\rangle \in V$ then $\alpha |a\rangle + \beta |b\rangle \in V$ for all $\alpha, \beta \in \mathbb{F}$. In quantum mechanics, the state space is a vector space over $\mathbb{F} = \mathbb{C}$.

Consider the collection of complex-valued functions:

$$\mathcal{F} = \{\psi : \mathcal{X} \to \mathbb{C} : \psi(x) = \mathrm{Re}(\psi)(x) + \mathrm{Im}(\psi)(x)\}. \tag{1.10}$$

We can see that this forms a vector space over over $\mathbb{C}$:

- $\alpha\psi(x) \in \mathcal{F}$ for all $\alpha \in \mathbb{C}, \psi \in \mathcal{F}$.
- $\psi(x) + \phi(x) \in \mathcal{F}$ for all $\psi, \phi \in \mathcal{F}$.

Consider another example:

$$\mathcal{W} = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} : a_1, a_2, a_3, a_4 \in \mathbb{C} \right\} \tag{1.11}$$

We can see that $\mathcal{W}$ is a vector space of dimension 4 over $\mathbb{C}$:

- Addition: $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \\ a_4 + b_4 \end{bmatrix}$ for all $a_i, b_i \in \mathbb{C}$ $i = 1, \ldots, 4$.
- Scalar multiplication:
$$\alpha \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \\ \alpha a_4 \end{bmatrix}$$
for all $\alpha \in \mathbb{C}$.

1.4 Bra and Ket Vectors

First, we recall a complex number $z = x + iy$ has complex conjugate $z^* = x - iy$. One can think of the ket vector $|z\rangle$ as a representation of z and the bra vector $\langle z|$ as a representation of z^*. That is, the bra vector is a complex conjugate of the ket vector.

To be more precise, the ket $|v\rangle$ is a vector in a complex vector space V and $\langle f|$ is a linear functional $f : V \to \mathbb{C}$ in the dual space V^*. One can think of interchanging bras and kets as a complex conjugation:

$$\langle B|A\rangle = \langle A|B\rangle^* \tag{1.12}$$

In quantum mechanics, the state is an element of a complex **Hilbert space**, i.e. the infinite-dimensional vector space of square-integrable wave functions:

$$\mathcal{H} = \left\{ |\psi\rangle : \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty \right\}. \tag{1.13}$$

Definition 3 *The minimum number of basis vectors needed to span the vector space is its* ***dimension****.*

As an example, one can represent any vector on a line as $r\,|v\rangle$ where $|v\rangle$ is its basis vector. As another example,

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \tag{1.14}$$

is a basis for a 4-dimensional vector space. As another less intuitive example, the Hilbert space $\mathcal{H} = \left\{ |\psi\rangle : \int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty \right\}$ is ∞-dimensional.

Remark 3 *We can think of a dual space V^* of a complex vector space V as the 'complex conjugated' version of V. Similar to how we can map $z \to z^*$, we can map bras to keys under isomorphism of a vector space and its dual. As a mnemonic: if a vector space is a space of complex-valued functions, the dual vector space is made out of the complex conjugates of the functions. In particular, there exist a bijective mapping: $\langle A| \leftrightarrow |A\rangle$.*

Remark 4 *Suppose we have a finite-dimensional ket vector:*

$$|A\rangle = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \tag{1.15}$$

then the bra vector can be identified with it's complex conjugated transpose:

$$\langle A| = \begin{bmatrix} a_1^* & \cdots & a_n^* \end{bmatrix}. \tag{1.16}$$

Similarly, since the complex conjugate of a product is the product of complex conjugates, there exists another mapping between vector space and dual space representatives under scalar multiplication and addition:

- $\alpha\,|A\rangle \in V \leftrightarrow \langle A|\,\alpha^* \in V^*$
- $|A\rangle + |B\rangle \in V \leftrightarrow \langle A| + \langle B| \in V^*$

1.5 Inner Products

The inner product of a ket vector $|B\rangle$ with bra vector $\langle A|$ is:

$$\langle A|B\rangle \tag{1.17}$$

such that:

- $\langle A|\beta B\rangle = \beta\langle A|B\rangle$
- $\langle A|\left[|B\rangle + |C\rangle\right] = \langle A|B\rangle + \langle A|C\rangle$.
- $\langle A|B\rangle = \langle B|A\rangle^*$ (thought of as $(z_1 z_2)^* = z_2^* z_1^*$).

Notice, the inner product is a linear operator on B. Furthermore, these properties imply:

$$\langle A|A\rangle = \langle A|A\rangle^* \implies \langle A|A\rangle \text{ is real.} \tag{1.18}$$

We also axiomatically enforce that $\langle A|A\rangle > 0$ (analogous to $z^* z > 0$).

Example 1.5.1 *We define an inner product on the vector space of complex-valued functions as:*

$$\langle \phi|\psi\rangle = \int dx \phi^*(x)\psi(x). \tag{1.19}$$

Example 1.5.2 *The k-dimensional space of column vectors with complex entries $V \subseteq \mathbb{C}^k$ has inner product:*

$$\langle b|a\rangle = \sum_{i=1}^{k} b_i^* a_i. \tag{1.20}$$

1.5.1 Bases

We define a D-dimensional orthonormal basis as follows:

- *Basis: $|b\rangle_i$ $i = 1, \ldots, D$ basis vectors.*
- *Normality: $\langle b_i|b_i\rangle = 1$ unit vectors.*
- *Orthogonality: $\langle b_i|b_j\rangle = 0$ for all $j \neq i$.*

Example 1.5.3 *We can form a basis of $\mathbb{R}^4$ under the identifications:*

$$|b_1\rangle \implies \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, |b_2\rangle \implies \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, |b_3\rangle \implies \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, |b_4\rangle \implies \begin{bmatrix}0\\0\\0\\1\end{bmatrix}. \tag{1.21}$$

Using Kronecker delta notation, we write the orthonormality property as:

$$\langle b_i | b_j \rangle = \delta_{ij}. \tag{1.22}$$

Furthermore, we can write $|v\rangle$ *in a basis as:*

$$|v\rangle = \sum_i v_i \, |b\rangle_i \, . \tag{1.23}$$

Notice, we can extract the coefficient v_j *by taking the inner product with the corresponding basis vector* b_j*:*

$$\begin{aligned} \langle b_j | v \rangle &= \sum_i v_i \langle b_j | b_i \rangle \\ &= v_j \langle b_j | b_j \rangle \\ &= v_j. \end{aligned} \tag{1.24}$$

Therefore the ***basis vector expansion*** *is:*

$$|v\rangle = \sum_i |b_i\rangle \, v_i = \sum_i |b_i\rangle \langle b_i | v \rangle \, . \tag{1.25}$$

Chapter 2

Observables

2.1 Postulates

A coin is a two-state system described by Heads (H) and Tails (T), forming a set called the phase space. Now, suppose each state has an associated pointer as shown in Figure 2.1. Consider a **superposition of states**:

$$\alpha \left|H\right\rangle + \beta\beta T. \tag{2.1}$$

While this doesn't have an interpretation in classical mechanics, we can entertain such a state in quantum mechanics. That is, states of a system are vectors and, what's more, we can linearly combine them.

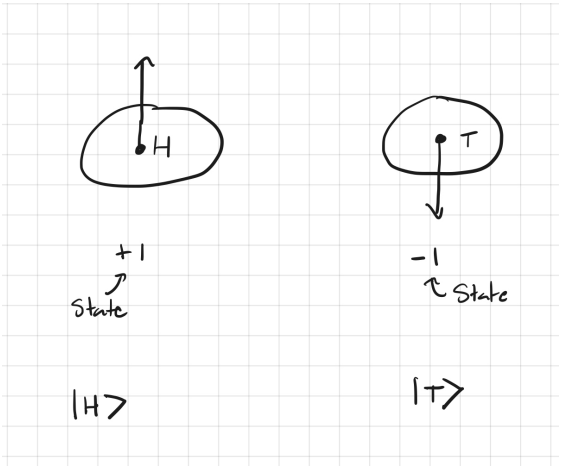


Figure 2.1: Two-state coin system.

2.1.1 Postulate: Orthogonality

The first postulate of quantum mechanics is as follows: for states of a system that are easily distinguished by a simple experiment (e.g. look to see if the state is heads or tails in this case), the vectors that go along with those two configurations are **orthogonal**.

Example 2.1.1 *For instance, in the two-state $|H\rangle, |T\rangle$ system, we have that:*

$$\langle H|T\rangle = 0. \tag{2.2}$$

We also choose:

$$\langle H|H\rangle = \langle T|T\rangle = 1. \tag{2.3}$$

Thus, $|H\rangle, |T\rangle$ form a basis.

Example 2.1.2 *Consider a die tossed on a table, as shown in Figure 2.2. Here, $|1\rangle, \ldots, |6\rangle$ are basis vectors. They are mutually orthogonal because one can tell the difference between them in a single experiment. As before, we can consider a superposition of states:*

$$\alpha_1 |1\rangle + \cdots + \alpha_6 |6\rangle, \tag{2.4}$$

which has no classical analog.

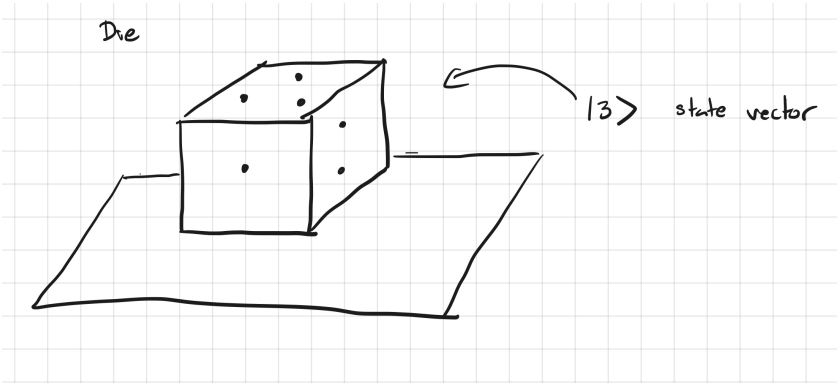


Figure 2.2: Six state die.

2.1.2 Postulate: Unit Length

Consider a confused coin of a two-state system:

$$|CC\rangle = \alpha_H |H\rangle + \alpha_T |T\rangle. \tag{2.5}$$

The probability of heads (resp. tails) is:

$$\begin{aligned} p_H &= \alpha_H^* \alpha_H \\ p_T &= \alpha_T^* \alpha_T \end{aligned} \tag{2.6}$$

such that

$$p_H + p_T = 1. \tag{2.7}$$

This implies that:

$$\alpha_H^* \alpha_H + \alpha_T \alpha_T^* = 1 = \langle CC|CC\rangle. \tag{2.8}$$

We call α_H and α_T **probability amplitudes**. An important property we notice is $p_H = \alpha_H^* \alpha_H$ must be positive and real, which aligns with our discussion of inner products and probability.

Example 2.1.3 *For a confused die:*

$$|CD\rangle = \alpha_1 |1\rangle + \dots \alpha_6 |6\rangle, \tag{2.9}$$

we require that:

$$\alpha_1^* \alpha_1 + \cdots + \alpha_6^* \alpha_6 = 1 = \langle CD|CD\rangle. \tag{2.10}$$

Another postulate of quantum mechanics is that all linear superpositions should themselves be of unit length, which means probabilities sum to 1. That is, for a finite-dimensional state space, suppose $|v\rangle = \sum_i v_i |b_i\rangle$. Then $\langle v|v\rangle = 1$.

2.2 Linear Operators

A linear operator $\hat{L}$ acts on a ket to give another ket:

$$\hat{L} |A\rangle = |B\rangle. \tag{2.11}$$

It is linear in scalar multiplication and vector addition:

- $\hat{L}\alpha |A\rangle = \alpha \hat{L} |A\rangle$
- $\hat{L}[\langle A| + \langle B|] = \hat{L}\langle A| + \hat{L}\langle B|$.

Examples of linear operators include ration by an angle, multiplying vectors by a number, and reflection. In quantum mechanics, linear operators correspond to quantities you can measure, i.e. *observables*, e.g. the "headness" or "tailness" of a coin.

2.3 Observables

Measurable quantities (**observables**) are represented by Hermitian linear operators. One can act on kets with linear operators to get a ket:

$$\hat{K} |A\rangle = |C\rangle. \tag{2.12}$$

Consider taking the inner product with $\langle B|$:

$$\langle B| [\hat{K} |A\rangle] = \langle B|\hat{K}|A\rangle = \hat{K}_{AB}. \tag{2.13}$$

This can be thought of as $\hat{K}$ first acting on $|A\rangle$ to give $\hat{K} |A\rangle$, followed by an inner product with $\langle B|$. This is a scalar, such that $\hat{K}_{AB}$ can be interpreted as the matrix element of $\hat{K}$ between vector B and vector A.

Consider a D-dimensional basis of vectors $\{|n\rangle\}$. Then we can enumerate he matrix for all $|m\rangle, |n\rangle$ basis vectors drawn from the same basis:

$$\begin{gathered} \langle m|\hat{K}|n\rangle = K_{mn} \\ \begin{bmatrix} K_{11} & \cdots & K_{1D} \\ \vdots & \ddots & \vdots \\ K_{D1} & \cdots & K_{DD} \end{bmatrix}. \end{gathered} \tag{2.14}$$

Let's expand $|A\rangle$ in a basis:

$$|A\rangle = \sum_n A_n |n\rangle , \tag{2.15}$$

where we sum over basis vectors and $\langle m|A\rangle = A_m$. As such, giving the coefficients A_m is a representation of the abstract vector $|A\rangle$, i.e. there exists an isomorphism between the two. Then:

$$|A\rangle = \sum_n |n\rangle\langle n|A\rangle . \tag{2.16}$$

(The implicit mnemonic is that $|n\rangle\langle n|$ does nothing, i.e. it's an identity.) To find the components of $\hat{K}|A\rangle$, we compute the n-th component as:

$$\begin{aligned} \langle n|\hat{K}|A\rangle &= \sum_m \langle n|\hat{K}|m\rangle\langle m|A\rangle \\ &= \sum_m \hat{K}_{mn} A_m \end{aligned} \tag{2.17}$$

where we use the identity $|A\rangle = \sum_m |m\rangle\langle m|A\rangle$ in the first equality. In the second equality, we can think of matrix $\hat{K}$ acting on a column vector A.

2.3.1 Composition of Linear Operators

The composition of linear operators is also linear:

$$\hat{K}\hat{L}|A\rangle = \hat{K}[\hat{L}|A\rangle] \tag{2.18}$$

where we interpret the chain as $\hat{L}$ first acting on A to product $\hat{L}|A\rangle$, which is then acted on by $\hat{K}$ to return $\hat{K}\hat{L}|A\rangle$. To find the matrix elements of $\hat{K}\hat{L}$, we first apply the identity trick:

$$\hat{L}|m\rangle = \sum_r |r\rangle\langle r|\hat{L}|m\rangle . \tag{2.19}$$

It follows that:

$$\langle n|\hat{K}\hat{L}|m\rangle = \sum_r \langle n|\hat{K}|r\rangle\langle r|\hat{L}|m\rangle , \tag{2.20}$$

where $\hat{K}_{nr} = \langle n|\hat{K}|r\rangle$ and $\hat{L}_{rm} = \langle r|\hat{L}|m\rangle$, such that:

$$(\hat{K}\hat{L})_{nm} = \sum_r \hat{K}_{nr}\hat{L}_{rm} \tag{2.21}$$

is the nm-th matrix element of $\hat{K}\hat{L}$.

Chapter 3

Postulates of Quantum Mechanics

3.1 Hermitian Operators

A **Hermitian operator** $\hat{H}$ satisfies:

(A)

$$\langle B|\hat{H}|A\rangle = \langle A|\hat{H}|B\rangle^* \tag{3.1}$$

(B)

$$\langle A|\hat{H}|A\rangle = \langle A|\hat{H}|A\rangle^* \implies \langle A|\hat{H}|A\rangle \text{ is real.} \tag{3.2}$$

Notably, property (A) follows (B); thus, we may use (B) as the true definition of Hermitian. Property (A) is equivalent to $H_{AB} = H^*_{BA}$ in components. One can also think of the complex conjugate in (A) as $\langle A|\hat{H}|B\rangle^* = \langle B|^* \hat{H}^* |A\rangle^* = \langle B|\hat{H}^*|A\rangle$. Therefore, if $\hat{H}$ is Hermitian, we have that:

$$H_{mn} = H^*_{nm}. \tag{3.3}$$

3.1.1 Eigenvalues and Eigenvectors

Eigenvectors $|\lambda\rangle$ of an operator $\hat{H}$ are vectors whose direction does not change when one applies $\hat{H}$. More precisely,

$$\hat{H}\,|\lambda\rangle = \lambda\,|\lambda\rangle \tag{3.4}$$

where $|\lambda\rangle$ is an eigenvector of $\hat{H}$ with eigenvalue λ. Note, if we multiply an eigenvector by $\alpha \in \mathbb{C}$, it still has the sample eigenvalue since:

$$\hat{H}(\alpha\,|\lambda\rangle) = \lambda(\alpha\,|\lambda\rangle), \tag{3.5}$$

which means we can always multiply eigenvectors to be on unit length.

We present three key theorems related to eigenvectors and eigenvalues of Hermitian operators.

Theorem 1 *All eigenvalues of a Hermitian operator $\hat{H}$ are real.*

Theorem 2 *The eigenvectors of $\hat{H}$ are orthogonal for $\lambda_1 \neq \lambda_2$.*

Theorem 3 *There exists D mutually orthogonal eigenvectors.*

Since they are mutually orthogonal, *the eigenvectors form a **basis** for the complex vector space.* Furthermore, *all of their eigenvalues are **real**.*

Remark 5 *Symmetric operators are the real-analog of Hermitian operators:*

$$\hat{H}_{AB} = \hat{H}^*_{BA} = H_{BA}. \tag{3.6}$$

Remark 6 *We can think of Hermitian operators $\hat{H}$ as matrices and ket vectors $|A\rangle$ as column vectors. Thus, $\hat{H}|A\rangle$ can be simply thought of as matrix-vector multiplication.*

3.2 The Postulates of Quantum Mechanics

(1) **Postulate 1** *States correspond to collections of ket vectors:*

$$States \implies \{|A\rangle\} \tag{3.7}$$

E.g. For a coin $|H\rangle, |T\rangle$ are two mutually orthogonal states for the two-dimensional system.

(2) **Postulate 2** *Observables (measured in experiments) correspond to collections of Hermitian operators:*

$$Observables \implies \{\hat{H}\} \tag{3.8}$$

Examples of observables include: momentum, position, the electric field.

(3) **Postulate 3** *The values of observables are the **eigenvalues** of $\hat{H}$. In an experiment, eigenvalues are what we get when we measure an observable $\hat{H}$.*

Remark 7 *When we measure an observable, e.g. angular momentum, we get a real number so it's important that Hermitian operators have real eigenvalues.*

(4) **Postulate 4** *The state for which observable $\hat{H}$ is **definite** (certain)) are the **eigenvectors** of $\hat{H}$. E.g. suppose we create an electron in an eigenstate of an observable such as position. If we measure position, the measurement will always yield and eigenvalue of the position operator.*

Remark 8 *The mnemonic is: we observe the observable $\hat{H}$ and the result we get is one of the eigenvalues of the observable.*

Remark 9 *Suppose we put an electron in an arbitrary state, i.e. spin along a particular axis. If we then measure the spin along some other axis, with a statistical probability, we will get different answers. However, if we if we measure spin along the axis we prepared it in, we will get a definite answer (i.e. no statistical uncertainty). Namely, if the system has been prepared in an eigenstate of a particular variable and we measure that variable, we always get an eigenvalue.*

(5) **Postulate 5** *Take an arbitrary state of the system $|A\rangle$. If we measure the $\hat{H}$-ness of the electron, then in a basis of eigenvectors $|\lambda\rangle_1, \ldots, |\lambda\rangle_n$ of $\hat{H}$, the possible answers we get are $\lambda_1, \ldots, \lambda_n$. The probability of measuring λ_n if the system has been prepared in state $|A\rangle$ is:*

$$\mathbb{P}(\lambda_n) = |\langle\lambda_n|A\rangle|^2 = \langle A|\lambda_n\rangle\langle\lambda_n|A\rangle \tag{3.9}$$

where $\langle\lambda_n|A\rangle$ is usually a complex number and measures the component of $|A\rangle$ along direction $\langle\lambda_n|$. Notably, $\langle A|\lambda_n\rangle\langle\lambda_n|A\rangle$ is a product of complex conjugates and, hence, real.

Example 3.2.1 (Particle on a Line, Position Observable) *Consider a particle on a line. It is described by its* ***wavefunction*** *$\psi(x)$, corresponding to a vector in a vector space of complex functions (Hilbert space). The inner product on the Hilbert space is:*

$$\langle\phi(x)|\psi(x)\rangle = \int dx \phi^*(x)\psi(x) \tag{3.10}$$

Consider the position observable:

$$\hat{X}\,|\psi\rangle \implies x\psi(x) \tag{3.11}$$

where " $\implies$ " means "corresponds to". This is equivalent to:

$$\hat{X}\,|\psi(X)\rangle = |X\psi(X)\rangle\,. \tag{3.12}$$

We confirm the position operator is Hermitian by checking whether $\langle A|\hat{H}|A\rangle$ is real, which is a necessary and sufficient condition to be Hermitian.

$$\begin{aligned}\langle\psi(X)|\hat{X}|\Psi(X)\rangle &= \langle\psi(X)|\left[\hat{X}\,|\psi(X)\rangle\right] \\ &= \int \psi^*(x)x\psi(x)dx \\ &= \int \psi^*(x)\psi(x)xdx,\end{aligned} \tag{3.13}$$

where we always assume convergence of $\int\psi^*(x)\psi(x)dx$. Clearly, this is real, since $\psi^*(x)\psi(x) \in \mathbb{R}$. Thus, $\hat{X}$ is Hermitian.

The eigenvectors are solutions to:

$$\hat{X}\,|\psi(X)\rangle = \lambda\,|\psi(X)\rangle \tag{3.14}$$

which is equivalent to:

$$\begin{aligned} x\psi(x) &= \lambda\psi(x) \\ (x-\lambda)\psi(x) &= 0. \end{aligned} \tag{3.15}$$

Thus, $\psi(x) = 0$ everywhere but $x = \lambda$. In particular, $\psi(x) = \delta(x-\lambda)$ where δ is the Dirac-delta function. The ***Dirac-delta function*** *is defined as non-zero over a tiny interval ϵ, as seen in Figure 3.1. That is, $\delta(x-\lambda)$ is an eigenvector of $\hat{X}$ with eigenvalue λ.*

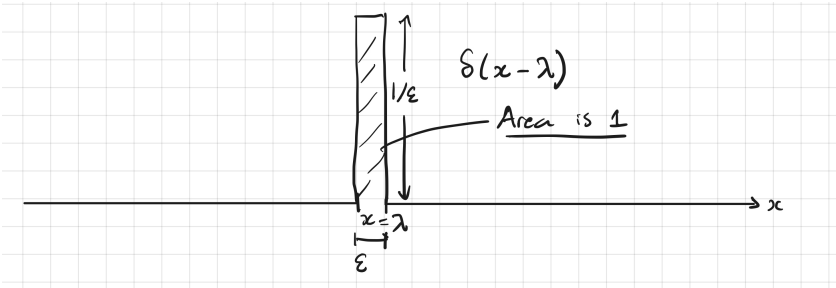


Figure 3.1: Dirac-delta function $\delta(x-\lambda)$.

As shown in Figure 3.2, this is an orthogonal family of functions. Let $\phi(x) = \delta(x-\lambda')$, $\psi(x) = \delta(x-\lambda)$, then:

$$\int_{-\infty}^{\infty} \phi^*(x)\psi(x)dx = 0. \tag{3.16}$$

Notice that we can erect and eigenvector at any λ on the real axis, which means the spectrum of eigenvalues is the entire real line $\mathbb{R}$.

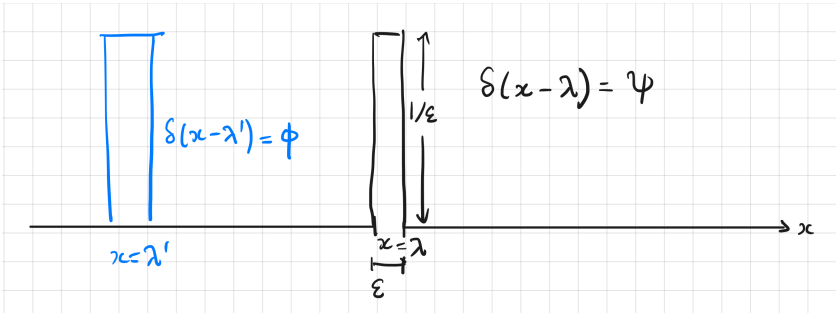


Figure 3.2: Two orthogonal delta functions $\delta(x-\lambda)$ and $\delta(x-\lambda')$.

Suppose $|\lambda\rangle = \delta(x-\lambda)$ is an eigenvector. We find:

$$\begin{aligned} \langle\lambda|\psi\rangle &= \int \delta(x-\lambda)\psi(x)dx \\ &= \psi(\lambda)\int_{[\lambda-\varepsilon,\lambda+\varepsilon]} \delta(x-\lambda)dx \\ &= \psi(\lambda) \end{aligned} \tag{3.17}$$

as $\int_{[\lambda-\varepsilon,\lambda+\varepsilon]} \delta(x-\lambda)dx = 1$*, by definition. It follows that* $\langle\lambda|\psi\rangle = \psi(\lambda)$ *or*

$$\langle x|\psi\rangle = \psi(x). \tag{3.18}$$

That is, the state localized at position x *corresponds to the wave function localized at* x *where* $\psi(x)$ *is the component of* $|\psi\rangle$ *along basis vector* $|x\rangle$*. The probability of detecting the particle at position* x *is:*

$$\mathbb{P}(x) = |\langle x|\psi\rangle|^2 = \psi^*(x)\psi(x). \tag{3.19}$$

Example 3.2.2 (Momentum Observable) *Consider the momentum observable* $\frac{\partial}{\partial x}\psi(x)$*, a linear operator acting on the space of functions* $\psi(x)$*. It turns out this operator isn't Hermitian–it's anti-Hermitian. Recall the definition of a Hermitian operator.*

Definition 4 *An operator* H *is Hermitian if:*

$$\langle A|H|B\rangle = \langle B|H|A\rangle^* . \tag{3.20}$$

We can define an anti-Hermitian operator as follows:

Definition 5 *An operator* $\mathcal{H}$ *is* ***anti-Hermitian*** *if*

$$\langle A|\mathcal{H}|B\rangle = -\langle B|\mathcal{H}|A\rangle^* . \tag{3.21}$$

Remark 10 *We note that if* $\mathcal{H}$ *is anti-Hermitian, then* $-i\mathcal{H}$ *is Hermitian.*

Following this remark, we claim that

$$\hat{K}\,|\psi(x)\rangle = -i\frac{\partial}{\partial x}\psi(x) \tag{3.22}$$

is a linear Hermitian operator. It suffices to show

$$\begin{aligned}\langle\psi|\hat{K}|\psi\rangle &= \int \psi^* \left(-i\frac{\partial}{\partial x}\psi(x)\right) dx \\ &= \int \psi^* \left(-i\frac{\partial\psi}{\partial x}\right) dx\end{aligned} \tag{3.23}$$

is real, per property (B) of Section 3.1. We will show it is real by show it is it's own complex conjugate. Using integration by parts $\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$ *and ignoring endpoints* $\int f(x)g'(x)dx = -\int g(x)f'(x)dx$*, we get that:*

$$\int \psi^* \left(-i\frac{\partial\psi}{\partial x}\right) dx = i\int \frac{\partial\psi^*}{\partial x}\psi(x)dx. \tag{3.24}$$

Taking the complex conjugate of the right hand side (RHS):

$$-i\int \psi^*(x)\frac{\partial\psi}{\partial x}dx, \tag{3.25}$$

which is just the LHS=RHS. This implies that the LHS is real. Thus,* $-i\frac{\partial}{\partial x}$ *is Hermitian, i.e.* $\hat{K} = -i\frac{\partial}{\partial x}$ *is a Hermitian operator.*

Consider the eigenvector equation:

$$-i\frac{\partial\psi(x)}{\partial x} = k\psi(x) \tag{3.26}$$

with eigenvalue k*. The complex exponential eigenfunction* $\psi(x) \propto e^{ikx} = \cos(kx) + i\sin(kx)$ *is a solution to this equation. Evidently:*

$$\frac{\partial\psi}{\partial x} = ike^{ikx} = ik\psi(x) \tag{3.27}$$

so $-i\frac{\partial\psi}{\partial x} = k\psi(x)$*. That is,* $\psi(x) = e^{ikx}$ *is an eigenvector of* $\hat{K} = i\frac{\partial}{\partial x}$ *with eigenvalue* k*.*

Remark 11 *The probability of finding such a particle in position* x *is:*

$$|\psi(x)|^2 = |e^{ikx}|^2 = 1. \tag{3.28}$$

In other words, there is a uniform probability distribution of finding the particle anywhere in space, i.e. there is no information about where the particle is.

We can reconcile this wavefunction solution with classical wave theory. Consider the real part of the wave $\cos(kx)$*. For a wave with wavelength* L*,* $kL = 2\pi$ *as* $\cos(2\pi) = 0$*. That is* $L = \frac{2\pi}{k}$ *relates the wavelength* L *to eigenvalue* k*. From de Broglie, we know a shorter wavelength* L *means greater momentum. Under this interpretation, if we have a wave of particles corresponding to waves of a given wavelength, these particles have momentum:*

$$\begin{aligned} p &= \frac{h}{L} \\ &= \frac{h}{\frac{2\pi}{k}} \\ &= \frac{h}{2\pi}k \\ &= \hbar k \end{aligned} \tag{3.29}$$

Therefore, the eigenvalue k *is the momentum of a particle in units of* $\hbar$*. Hence, we can write the eigenvector-eigenvalue equation* $-i\frac{\partial\psi(x)}{\partial x} = k\psi(x)$ *as:*

$$-i\hbar\frac{\partial\psi(x)}{\partial x} = p\psi(x). \tag{3.30}$$

Remark 12 *For* $\hat{X}$*, the eigenfunction is* $\delta(x-\lambda)$*–a peaked wavefunction. That is, we know it's position with certainty. On the other hand, for the momentum operator* $-i\frac{\partial}{\partial x}$*, the wave function is spread out all over space. This tradeoff is the source of the Heisenberg uncertainty principle.*

Remark 13 *The wave function* $\psi(x) = e^{ikx}$ *can be interpreted as follows. If we fix* k*, then* $\psi(x)$ *winds around the unit circle in* $\mathbb{C}$*. For large* k *(equivalently small* L*), we wind quickly as we move along. Likewise, for small* k*, we wind slowly.*

Chapter 4

Quantization & Heisenberg Uncertainty

4.1 From Discrete to Continuous: Inner Products

Let N be a discrete random variable with probability mass function $p(n)$ such that $\sum_n p(n) = 1$. Recall that basis vectors form an orthonormal basis if:

$$\langle m|n\rangle = \delta_{mn}. \tag{4.1}$$

This is really a function of a difference between variables $n - m$:

$$\begin{cases} \delta(n-m) = 0, \text{ if } n - m \neq 0 \\ \delta(n-m) = 1, \text{ if } n - m = 0 \end{cases}. \tag{4.2}$$

For a continuous random variable X, we have a probability density function, such that $\int dx p(x) = 1$. The probability of finding a particle between x and $x + \Delta x$ is:

$$\int_x^{x+\Delta x} p(x)dx = \mathbb{P}(x, x + \Delta x). \tag{4.3}$$

Recall, the Dirac-delta function is the continuous version of the Kronecker-delta function. We define a generalized function $\delta(x - y) = 0$ if $x \neq y$ and $\int \delta(x - y)dx = 1$. That is, the delta-function is an operation that we apply to functions with a *convolution* as follows:

$$\begin{aligned} \int \delta(x-y)F(x)dx &= \int \delta(x-y)F(y)dx \\ &= F(y)\int \delta(x-y)dx \\ &= F(y) \end{aligned} \tag{4.4}$$

where in the first equality we used the fact that in a small ϵ neighborhood of y, $F(x) \approx F(y)$ is constant. Hence, rule defining the Dirac-delta function is:

$$\int \delta(x-y)F(x)dx = F(y). \tag{4.5}$$

We said that $|y\rangle$, $|x\rangle$ are orthogonal $\langle y, x\rangle = 0$ for $|y\rangle \neq |x\rangle$ if two configurations are measurably different. The continuous analog of $\langle m|n\rangle = \delta_{mn}$ is:

$$\langle y|x\rangle = \delta(x - y). \tag{4.6}$$

Note, $\langle x|x\rangle = \delta(0)$ isn't defined since it's infinitely-high; however we don't end up using this quantity.

4.2 Rules of Probability in Quantum Mechanics

For discrete spaces, the probability of detecting n for an observable labeled by n is:

$$p(n) = |\langle n|\psi\rangle|^2 = \langle\psi|n\rangle\langle n|\psi\rangle. \tag{4.7}$$

In continuous space, the probability of finding particle at y given it's in a state $|\psi\rangle$ is:

$$p(y) = |\langle y|\psi\rangle|^2 = \langle\psi|y\rangle\langle y|\psi\rangle \tag{4.8}$$

Consider the vector space of complex functions $\psi(x)$ of x. The inner product for the vector space of complex functions is:

$$\langle\phi|\psi\rangle = \int dx\phi^*(x)\psi(x). \tag{4.9}$$

The wave function associated with a particle known to be (localized) at y is $\phi(x) = \delta(x - y)$. Hence,

$$\langle y|\psi\rangle = \int dx\psi(x)\delta(x - y) = \psi(y) \tag{4.10}$$

which is the amplitude of discovering the particle at point y. The probability density is simply $p(y) = \psi^*(y)\psi(y)$.

4.2.1 Limiting Construction

To go from the discrete to continuous space, we consider a one-dimensional dense set of discrete points with separation a between all points. Then $\langle m|n\rangle = \delta_{mn}$ in the discrete setting. We can identify x with n as follows:

$$|x\rangle = \frac{1}{\sqrt{a}}|n\rangle \tag{4.11}$$

such that

$$\langle x|y\rangle = \frac{\delta(n - m)}{a} = \delta(x - y), \tag{4.12}$$

i.e. $\frac{\delta(n-m)}{a}$ is a delta function with height $1/a$. When $a \to 0$, this approaches infinite height, i.e. the delta function.

4.3 Particle on a Circle

Consider a particle on a circle of radius $2\pi r$. We can cut the circle at the bottom and lay the circle into a line of length $2\pi r$. The circular nature of the line restricts wave functions $\psi(x)$ to be periodic over the diameter of the circle:

$$\psi(x) = \psi(x + 2\pi r). \tag{4.13}$$

This special class of functions is still a vector space as:

- $x\psi(x) = c\psi(x + 2\pi r)$
- $\phi(x) + \psi(x) = \phi(x + 2\pi r) + \psi(x + 2\pi r)$.

The position operator $\hat{X}\,|\psi\rangle \implies x\psi(x)$ and the momentum operator $P = -i\hbar\frac{\partial}{\partial x}$ act on kets $|\psi\rangle \iff \psi(x)$. By $|\psi\rangle \iff \psi(x)$ we mean there exists a 1-to-1 correspondence between the ket vector and wave function representation. The eigenvectors $\psi_p(x) = e^{i\frac{p}{\hbar}x}$ satisfy $-i\hbar\frac{\partial}{\partial x}\psi(x) = p\psi(x)$ with eigenvalue p. We normalize this solution such that $\int_{\text{Circle}} \psi_p^*(x)\psi_p(x)dx = 1$. Since $\psi_p^*(x)\psi_p(x) = |e^{i\frac{p}{\hbar}x}| = 1$, we find:

$$1 = \int_{\text{Circle}} \psi_p^*(x)\psi_p(x)dx = \int_0^{2\pi r} 1 = 2\pi r. \tag{4.14}$$

Hence, we normalize ψ:

$$\psi_p(x) = \frac{e^{i\frac{p}{\hbar}x}}{\sqrt{2\pi r}}. \tag{4.15}$$

We note that $\psi_p^*\psi_q = 0$ because eigenvectors of Hermitian operators with different eigenvalues are always orthogonal.

Furthermore, we must also guarantee that our solution is periodic:

$$\begin{aligned} e^{i\frac{p}{\hbar}(x+2\pi r)} &= e^{i\frac{p}{\hbar}x} \\ e^{i\frac{p}{\hbar}2\pi r} &= 1, \end{aligned} \tag{4.16}$$

which is a restriction on the possible values of eigenvalues p so we are in the appropriate space of functions. Since $e^{2\pi ni} = 1$ for all $n \in \mathbb{Z}$, we have that $\frac{p}{\hbar}2\pi r = 2\pi n$ or

$$\boxed{p = n\frac{\hbar}{r},} \tag{4.17}$$

i.e. the allowable values of momentum are integer multiples of $\frac{\hbar}{r}$. Thus, although position is a continuous variable on the circle, momentum is discrete, i.e. *quantized*. Note, as $r \to \infty$, the spacing $\frac{\hbar}{r} \to 0$, so possible momenta become continuous. Similarly, since angular momentum of a particle on a circle is $L = pr$, the *spectrum of angular momentum* is:

$$\boxed{L = r\frac{n\hbar}{r} = n\hbar,\ n \in \mathbb{Z},} \tag{4.18}$$

i.e. angular momentum is also quantized.

As $r \to \infty$, the discrete spectrum of momentum (eigenvalues) becomes continuous and the circle approximates the infinite line. Viewing momentum as a continuous variable,

$$\langle q|p\rangle = \delta(p-q). \tag{4.19}$$

The takeaway is *the quantization of momentum and angular momentum is due to the periodicity of the wave function.*

Remark 14 *Recall eigenstates of position look like $\delta(x-\lambda)$. In contrast, the eigenstates of momentum take the form $e^{i\frac{p}{\hbar}x} = \cos(\frac{p}{\hbar}x) + i\sin(\frac{p}{\hbar}x)$, where each function is uniformly spread over the circle with $|e^{i\frac{p}{\hbar}x}|^2 = 1$. As $r \to \infty$, the wave becomes infinitely spread over the line. As a consequence, no wave function is simultaneously an eigenvector of position and momentum, unlike in classical mechanics.*

4.4 Compatibility of Observables

Definition 6 *The operators $\hat{A}$, $\hat{B}$ are said to be **simultaneously specifiable** if they share eigenvectors. If there is a complete basis of vectors $\{|n\rangle\}$ simultaneously eigenvectors of $\hat{A}$ and $\hat{B}$, then both $\hat{A}$ and $\hat{B}$ can be specified and definite, and are said to be **compatible**.*

Consider operators $\hat{A}$, $\hat{B}$ with a basis of eigenvectors $\{|n\rangle\}$:

$$\begin{aligned} \hat{A}\,|n\rangle &= \alpha_n\,|n\rangle \\ \hat{B}\,|n\rangle &= \beta_n\,|n\rangle \end{aligned} \tag{4.20}$$

Then

$$\begin{aligned} \hat{A}\hat{B}\,|n\rangle &= \hat{A}[\hat{B}\,|n\rangle] \\ &= \hat{A}[\beta_n\,|n\rangle] \\ &= \beta_n A\,|n\rangle \\ &= \beta_n\alpha_n\,|n\rangle\,. \end{aligned} \tag{4.21}$$

The order we multiply $\hat{A}$ and $\hat{B}$ does not matter if $\{|n\rangle\}$ are simultaneously eigenvectors of $\hat{A}, \hat{B}$. Thus, the necessary and sufficient conditions that there exists a basis of vectors that are simultaneous eigenvectors of $\hat{A}$ and $\hat{B}$ is:

$$\hat{A}\hat{B} = \hat{B}\hat{A}. \tag{4.22}$$

We say that $\hat{A}$ and $\hat{B}$ **commute** if:

$$[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A} = 0, \tag{4.23}$$

where $[\hat{A}, \hat{B}]$ is called the **commutator**.

4.4.1 Heisenberg Uncertainty: Compatibility of Position and Momentum

We can see whether position and momentum are compatible by checking their commutator:

$$\hat{X}\hat{P} - \hat{P}\hat{X} = X\left(-i\hbar\frac{\partial}{\partial X}\right) - \left(-i\hbar\frac{\partial}{\partial X}\right)X, \tag{4.24}$$

which is an operator acting on functions:

$$[\hat{X}\hat{P} - \hat{P}\hat{X}]\psi(X) = \left[X\left(-i\hbar\frac{\partial}{\partial X}\right) - \left(-i\hbar\frac{\partial}{\partial X}\right)X\right]\psi(X). \tag{4.25}$$

Recall the order of operators acting on a ket vector: $\hat{A}\hat{B}\,|\psi\rangle = \hat{A}[\hat{B}\,|\psi\rangle]$. Applying the order of operations to the above equation:

$$\left[X\left(-i\hbar\frac{\partial}{\partial X}\right) - \left(-i\hbar\frac{\partial}{\partial X}\right)X\right]\psi(X) = -i\hbar X\frac{\partial\psi}{\partial X} + i\hbar\frac{\partial}{\partial X}(X\psi(X)) \tag{4.26}$$

where

$$\frac{\partial}{\partial X}(X\psi(X)) = \psi(X) + X\frac{\partial\psi}{\partial X}. \tag{4.27}$$

Substituting, we find that

$$\begin{aligned} -i\hbar X\frac{\partial\psi}{\partial X} + i\hbar\frac{\partial}{\partial X}(X\psi(X)) &= -i\hbar X\frac{\partial\psi}{\partial X} + i\hbar\psi(X) + i\hbar X\frac{\partial\psi}{\partial X} \\ &= i\hbar\psi(X). \end{aligned} \tag{4.28}$$

Therefore,

$$(\hat{X}\hat{P} - \hat{P}\hat{X})\psi(X) = i\hbar\psi(X) \tag{4.29}$$

so the commutator is simply:

$$[\hat{X}, \hat{P}] = i\hbar. \tag{4.30}$$

Therefore, $\hat{X}$ and $\hat{P}$ are operators that do not commute and have no common eigenvectors! The non-commutativity of $\hat{X}$ and $\hat{P}$ is known more widely as the **Heisenberg uncertainty principle**.

4.4.2 Double Slit Experiment Revisited

Consider a beam of particles described by a wave function with momentum $p/\hbar$: $\psi(x) = e^{i\frac{p}{\hbar}x}$ passing through one slit. Over a small interval, on the screen, we can approximate the wave by $|\psi_1\rangle = e^{i\frac{p}{\hbar}y}$ with a probability density $p(y) = |e^{i\frac{p}{\hbar}y}|^2 = 1$. That is, the probability is uniform and the intensity is just a blob for one hole. Now consider the double slit experiment, as shown in Figure **??** We add states if the particles go through the second pinhole. The second wave has slightly different variation over the same interval, i.e. a slightly different momentum. Hence, the wave function is:

$$\psi(x) = \psi_1(x) + \psi_2(x) = e^{i\frac{p}{\hbar}y} + e^{i\frac{q}{\hbar}y}. \tag{4.31}$$

Then probability of observing $|\psi\rangle$ in position y is:

$$\begin{aligned}\psi(y)\psi^*(y) &= (e^{i\frac{p}{\hbar}y} + e^{i\frac{q}{\hbar}y})(e^{-i\frac{p}{\hbar}y} + e^{-i\frac{q}{\hbar}y}) \\ &= 1 + 1 + e^{i\frac{(p-q)}{\hbar}y} + e^{i\frac{(q-p)}{\hbar}y} \\ &= 2 + 2\cos\left(\frac{p-q}{\hbar}y\right) \\ &= 2\left(1 + \cos\left(\frac{p-q}{\hbar}y\right)\right)\end{aligned} \tag{4.32}$$

which oscillates. This probability is always positive and when $\frac{p-q}{\hbar}y = 2\pi$, the probability is 0. Hence, if we open only one hole, the probability distribution of observing a particle in a position is uniform. If we open both holes, we find a destructive interference pattern, i.e. there are places that the particles simply cannot go.

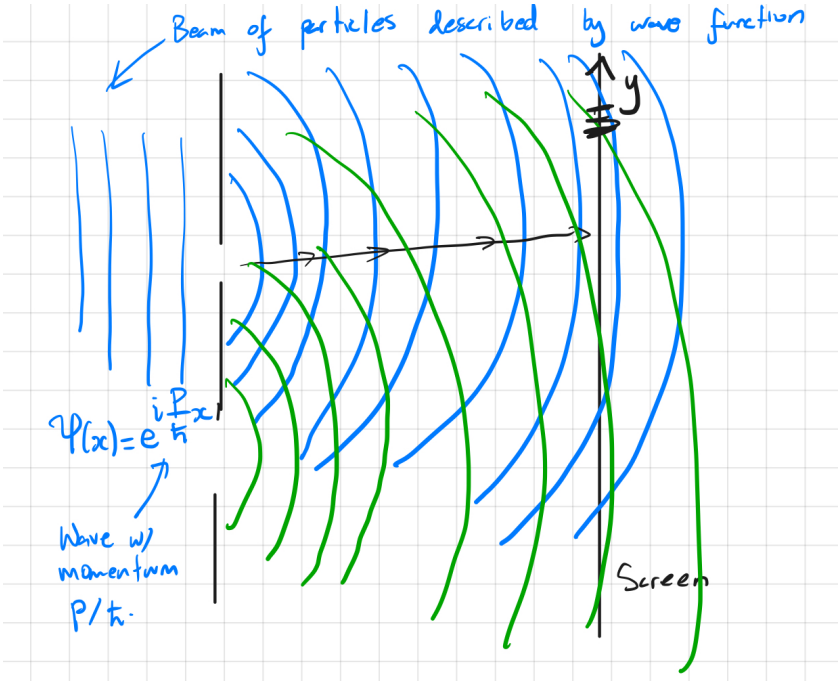


Figure 4.1: Double slit.

Note, there will be shorter wavelengths the further you move away from the vertical location of the slits. If one hole is closed and the other is open, then the further you depart from the vertical position of that hole, the shown the wave lengths, the greater the vertical momentum, and the faster the oscillations. The vertical frequency due to the bottom hole is slightly different than the vertical frequency due to the top hole because of their difference in height, hence why $p \neq q$. Notice that if $p = q$, the probability becomes uniform again. To formalize this, we can re-write the wave function in terms of a radial distance r and horizontal distance between the wall the particles emerge from and the screen L, as shown in

Figure 4.2. In this setting, $r = \sqrt{L^2 + y^2}$. Therefore,

$$\frac{e^{i\frac{p}{\hbar}r}}{r} = \frac{e^{i\frac{p}{\hbar}\sqrt{L^2+y^2}}}{\sqrt{L^2+y^2}} \tag{4.33}$$

such that oscillations get faster as y increases. This is the idea of beats in the probability distribution where there is a shifted wave function and the sum of them have an oscillation probability that individually neither of them have, as they individually have uniform probabilities. Since $\psi\psi^* = 2\left(1 + \cos(\frac{p-q}{\hbar}y)\right)$ if the two holes are closer together, $p \approx q$, it will have shorter wavelength.

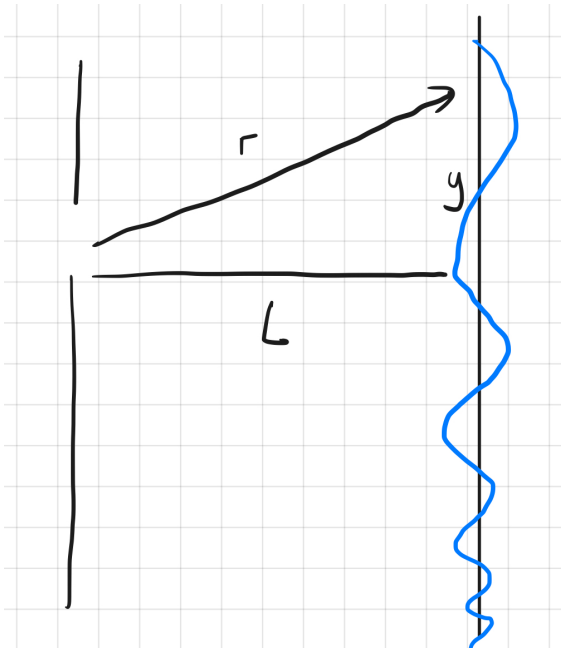


Figure 4.2: Reformulation in terms of radial distance r and vertical separation L.

In the one-hole setting, the wave function generally has the form

$$e^{i\frac{p}{\hbar}y}\rho(y) \tag{4.34}$$

where $e^{i\frac{p}{\hbar}y}$ are fast oscillations, as $\frac{p}{\hbar}$ is usually large, and $\rho(y)$ is a smooth real function. Then it has probability:

$$|e^{i\frac{p}{\hbar}y}\rho(y)|^2 = \rho(y)^2, \tag{4.35}$$

such that there are no oscillations. When we open a second hole, it will have wave function:

$$e^{i\frac{q}{\hbar}y}\rho'(y) \tag{4.36}$$

so their superposition:

$$e^{i\frac{p}{\hbar}y}\rho(y) + e^{i\frac{q}{\hbar}y}\rho'(y) \tag{4.37}$$

gives oscillations or *beats*.

Chapter 5

Periodic Wavefunctions & Polarizers

5.1 Particle on a Circle

Consider a particle on a circle of diameter $2\pi r$, such that the wave function satisfies:

$$\psi(x) = \psi(x + L). \tag{5.1}$$

Assume the notation $e^{i\frac{p}{\hbar}x} = e^{ikx}$. The normalized periodic wave function is:

$$\psi(x) = \frac{e^{ikx}}{\sqrt{2\pi r}} \tag{5.2}$$

such that

$$\int \psi^*(x)\psi(x)dx = \int_0^{2\pi r} \left|\frac{e^{ikx}}{\sqrt{2\pi r}}\right|^2 dx = \frac{1}{2\pi r}\int_0^{2\pi r} dx = 1. \tag{5.3}$$

The periodicity $\psi(x) = \psi(x + 2\pi r)$ of the wave function implies a constraint on the eigenvalue of momentum k:

$$\begin{aligned} \frac{e^{ikx}}{\sqrt{2\pi r}} &= \frac{e^{ikx}}{\sqrt{2\pi r}}e^{i2\pi rk} \\ 1 &= e^{2\pi rki} \\ e^{2\pi ni} &= e^{2\pi rki} \\ k &= \frac{n}{r} \end{aligned} \tag{5.4}$$

for $n \in \mathbb{Z}$. That is, momentum is quantized. Notice, the spectrum of possible values of k are spaced by $1/r$. Hence, when $r \to \infty$, $1/r \to 0$, so k is no longer quantized. That is, momentum is continuous in the limit.

Since wavefunctions are eigenvectors of a Hermitian operator, for different values of k, they must be orthogonal, by the postulates of quantum mechanics. More

precisely,

$$\begin{aligned}\frac{1}{2\pi r}\int_0^{2\pi r} e^{ikx}e^{-ik'x}dx &= \frac{1}{2\pi r}\int_0^{2\pi r} e^{i(k-k')x}dx \\ &= \begin{cases}0, & \text{if } k \neq k' \\ 1, & \text{if } k \neq k'\end{cases} \\ &= \delta_{nn'}\end{aligned} \tag{5.5}$$

where n and n' are integers associated with k and k' respectively, such that $k = n/r$, $k' = n'/r$. These wave functions form an orthonormal basis for the vector space.

5.1.1 The Identity Operator

The outer product of two vectors $\langle A|B\rangle$ is a complex number. The *outer product* $|B\rangle\langle A|$ is a linear operator, represented as a matrix. In particular, it maps ket vectors to ket vectors:

$$[|B\rangle\langle A|]\,|C\rangle = |B\rangle\langle A|C\rangle\,. \tag{5.6}$$

It is a linear operator or a *dyad*.

Observe, we can expand a ket vector in a basis $\{|n\rangle\}$ as:

$$|v\rangle = \sum_n v_n\,|n\rangle \tag{5.7}$$

where $v_n = \langle n|v\rangle$ are the expansion coefficients. This can be re-written as:

$$\begin{aligned}|v\rangle &= \sum_n |n\rangle\langle n|v\rangle \\ &= \sum_n [|n\rangle\langle n|]\,|v\rangle \\ &= \left(\sum_n |n\rangle\langle n|\right)|v\rangle\,.\end{aligned} \tag{5.8}$$

Hence, we define the **identity (unit) operator** as:

$$I = \sum_n |n\rangle\langle n| \tag{5.9}$$

such that

$$I\,|v\rangle = \sum_n |v\rangle\langle n|v\rangle = |v\rangle\,. \tag{5.10}$$

If we consider the particle on a circle from the previous section, as the circle becomes large, the momenta form a continuum and we can define its identity operator as:

$$I = \int dk\,|k\rangle\langle k| \tag{5.11}$$

and likewise for position, which is known to be continuous:

$$I = \int dx\,|x\rangle\langle x|\,. \tag{5.12}$$

5.1.2 Probability of Momenta

Let $|\psi\rangle$ be the state of a particle on a line with $|x\rangle$ the eigenfunction of the position operator $\hat{X}$ with eigenvalue x. Then, $\langle x|\psi\rangle = \psi(x)$ or, equivalently, $\langle x|k\rangle = \frac{e^{ikx}}{\sqrt{2\pi}}$. Note, since we are dealing with the continuum limit, r is omitted. We square the probability amplitudes

$$\tilde{\psi}(k) = \langle k|\psi\rangle \tag{5.13}$$

to get the probability of different momenta k. Since $I = \int dx\, |x\rangle\langle x|$, we can write this as:

$$\begin{aligned}\tilde{\psi}(k) &= \langle k|\psi\rangle \\ &= \int \langle k|x\rangle\langle x|\psi\rangle\, dx \\ &= \int \left(\frac{e^{ikx}}{\sqrt{2\pi}}\right)^* \psi(x) dx \\ &= \frac{1}{\sqrt{2\pi}} \int e^{-ikx}\psi(x) dx,\end{aligned} \tag{5.14}$$

which is the *Fourier Transform* of $\psi(x)$. That is,

$$\begin{aligned}\tilde{\psi}(k) &= \mathcal{F}[\psi(x)] \\ \mathbb{P}(k) &= |\tilde{\psi}(k)|^2\end{aligned} \tag{5.15}$$

where $\mathcal{F}[\psi(x)] := \frac{1}{\sqrt{2\pi}} \int e^{-ikx}\psi(x)dx$ is the Fourier transform, $\tilde{\psi}(k)$ is a momentum space wave function and $\psi(x)$ is a position space wave function. Similarly, we can go from position to momentum space by using the identity operator $I = \int dk\, |k\rangle\langle k|$:

$$\begin{aligned}\psi(x) &= \langle x|\psi\rangle \\ &= \int dk\, \langle x|k\rangle\langle k|\psi\rangle \\ &= \int \frac{e^{ikx}}{\sqrt{2\pi\}}\tilde{\psi}(k) dk.\end{aligned} \tag{5.16}$$

We can write this reciprocal relationship of momentum and position as:

$$\begin{aligned}\tilde{\psi}(k) &= \int \frac{e^{-ikx}}{\sqrt{2\pi}}\psi(x) dx \\ \psi(x) &= \int \frac{e^{ikx}}{\sqrt{2\pi}}\tilde{\psi}(k) dk.\end{aligned} \tag{5.17}$$

That is, there is an isomorphism between the $\psi(x)$ and $\tilde{\psi}(k)$ representations for state $|\psi\rangle$. Recall, the position and momentum operator for $\psi(x)$ is defined as:

$$\begin{aligned}\hat{X}\psi(x) &= x\psi(x) \\ \hat{K}\psi(x) &= -i\frac{\partial}{\partial x}\psi(x)\end{aligned}. \tag{5.18}$$

There is an equivalent formalism for the momentum space wave function:

$$\begin{aligned} \hat{K}\tilde{\psi}(k) &= k\tilde{\psi}(k) \\ \hat{X}\tilde{\psi}(k) &= i\frac{\partial\tilde{\psi}}{\partial k}. \end{aligned} \tag{5.19}$$

A *wave packet* has a wave function $\psi(x)$ that is concentrated in a region with a smooth shape, as shown in Figure 5.1. By the symmetry arguments in the previous section, $\tilde{\psi}(k)$ also has a nice concentrated wave function.

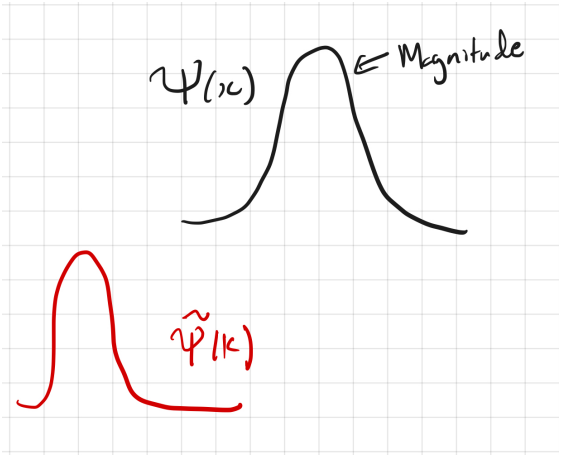


Figure 5.1: Wave packet.

On the other hand, wave functions that describe particles with fairly precise momentum and position take the form:

$$e^{ikx} f(x) \tag{5.20}$$

where the complex exponential oscillates and f is an envelope function, concentrated in x, as shown in Figure 5.2. Such a description is as certain as possible in both position and momentum, saturating the Heisenberg uncertainty principle. In fact, the center of these wave packets will move like a particle in accordance with the Schrodinger wave equation.

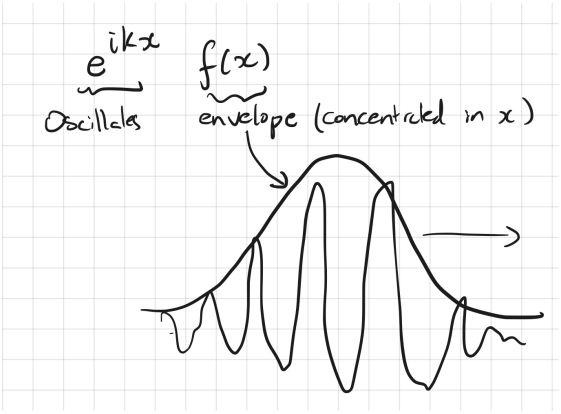


Figure 5.2: Wave function for particle with fairly precise momentum and position.

5.2 Polarizers

Light is made of photons. In the wave theory of light, there is an idea of *polarization*, which is the direction of the electric field $\vec{E}$. Namely, we say light is polarized along the axis of the direction of $\vec{E}$. In Figure 5.3, the electric field oscillates in the vertical plane and, hence, is vertically polarized.

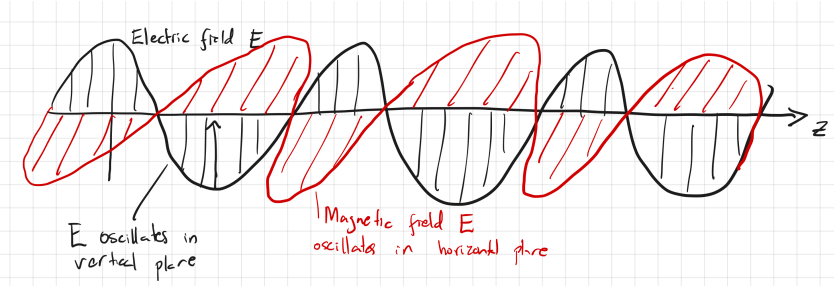


Figure 5.3: Electromagnetic wave.

More generally, the direction of the polarization vector is perpendicular to the direction of motion of a photon, as seen in Figure 5.4.

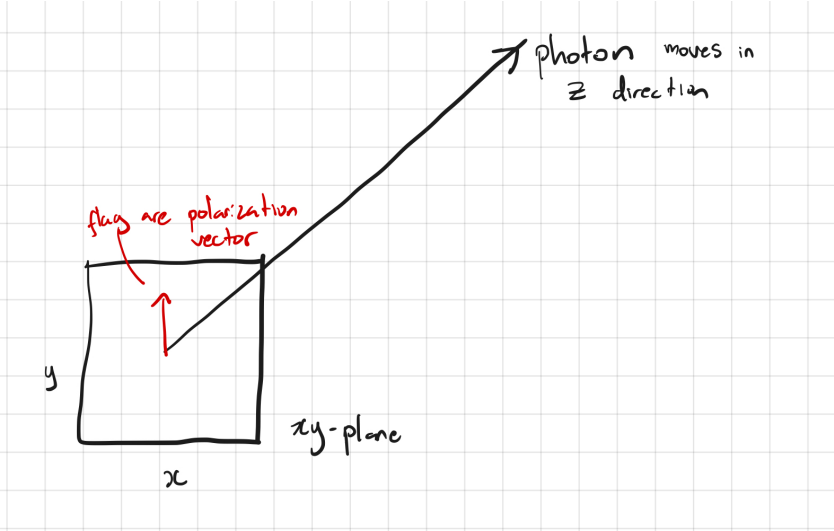


Figure 5.4: Polarization vector perpendicular to direction of motion of photon.

For instance, if one sent a radio wave in the direction perpendicular to a set of parallel wires, the wave would come out polarized along the axis perpendicular to the wires. As in Figure 5.5, current can only oscillate in the vertical direction, which implies the grid is effectively a mirror in the vertical direction, i.e. it reflects waves whose $\vec{E}$ field is vertical. When an electromagnetic wave falls on a mirror, current flows due to the electromagnetic field in any firection. The flow produces it's own reflected wave. In this example, the wave with vertically polarized $\vec{E}$ field will start current in the wires and get reflected since the $\vec{E}$ field cannot drive current in the horizontal direction. On the other hand, if the $\vec{E}$ field is horizontal,

it doesn't start current and passes through. Hence, the field goes through, i.e. the $\vec{E}$ field component that goes through is perpendicular to the direction of the grid.

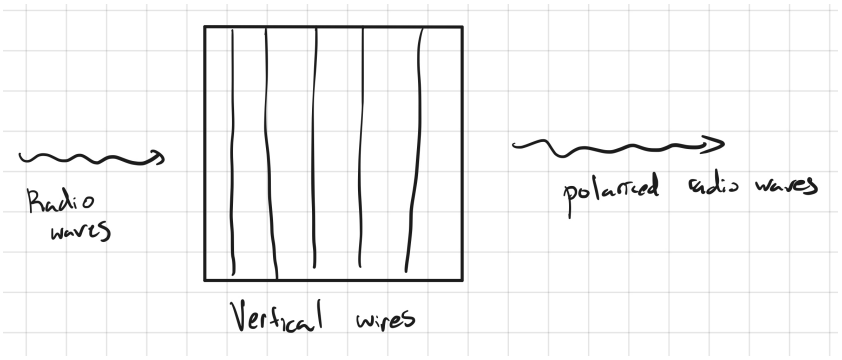


Figure 5.5: Vertical wire polarizer.

The *axis of a polarizer* is the direction of transmission. As an example, we can use a horizontal polarizer to prepare a photon into a state of horizontal polarization $|x\rangle$, as per Figure 5.6.

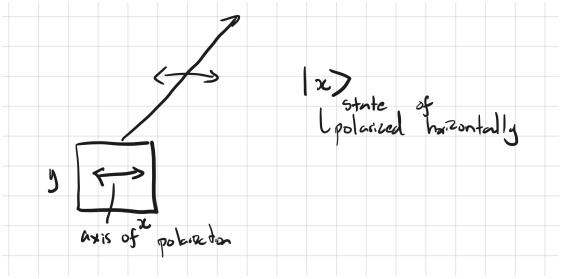


Figure 5.6: Horizontal state of polarization.

Polarizers can prepare a photon in a direction of polarization. We can also use them to *detect polarization*. Suppose we use a horizontal polarizer to prepare a horizontally polarized photon. Then we send it through another polarizer. If it is a horizontal polarizer, it will come out horizontally polarized with probability 1. However, if it a vertical polarizer, it will get blocked, as shown in Figure 5.7. Hence, we can record whether or not the photon passes as a form of detection. No matter how the photon is prepared (i.e. the direction of polarization of the light wave), it will either be transmitted or reflected–a quantity we can observe.

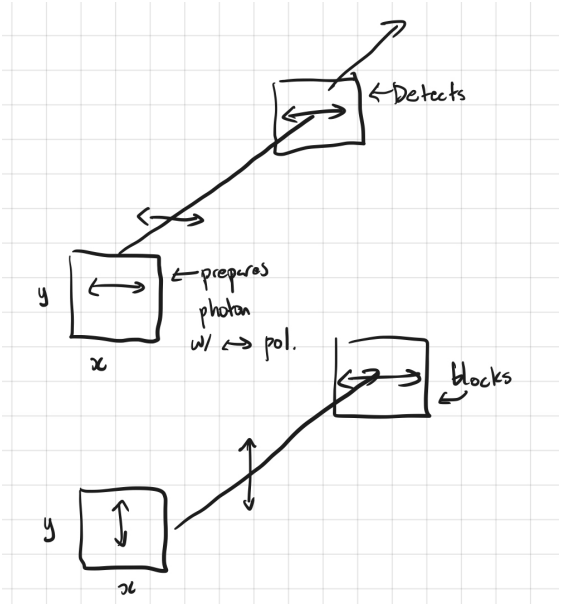


Figure 5.7: Detection with a polarizer.

If we polarize the photon at a slightly different angle, as shown in Figure 5.8, then when we detect horizontal polarization in the second polarizer, the photon may or may not go through with a certain probability. However, there are only two possible states–either it goes through or it doesn't! Indeed, we can rotate the polarization angle of the preparation polarizer continuously, *yet there will only be two states when we go to detect*; this is the strangeness of quantum mechanics.

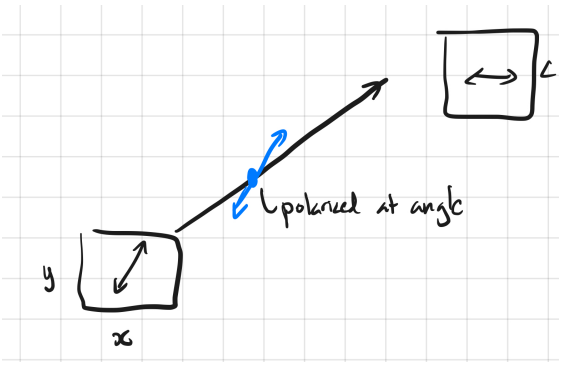


Figure 5.8: Two-state system.

In quantum mechanics, there are two orthogonal states of polarization along the x-axis and y-axis:

$$\begin{aligned} |\leftrightarrow\rangle = |x\rangle &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ |\updownarrow\rangle = |y\rangle &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{aligned} \tag{5.21}$$

A *single experiment* can tell you which of these two states the system is in, which means $|x\rangle$, $|y\rangle$ must be orthogonal by the postulates. That is, $\langle x|y\rangle = 0$ and we

normalize the basis vectors $\langle x|x\rangle = \langle y|y\rangle = 1$, such that $|x\rangle, |y\rangle$ are an orthonormal basis for a two-dimensional quantum mechanical system, i.e. a **qubit**.

5.2.1 Polarizer Observables

The polarization operator $\hat{\mathbb{P}}_{\oplus}$ for observable polarization in the xy-plane is defined such that we observe $+1$ eigenvalue for vertical polarization and -1 eigenvalue for horizontal polarization. That is, the eigenvector-eigenvalue definition is:

$$\begin{aligned} \hat{\mathbb{P}}_{\oplus}|x\rangle &= +|x\rangle \\ \hat{\mathbb{P}}_{\oplus}|y\rangle &= -|y\rangle \end{aligned}. \tag{5.22}$$

The matrix representation is:

$$\hat{\mathbb{P}}_{\oplus} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5.23}$$

We can check that it, indeed, satisfies the eigenvector-eigenvalue formulation:

$$\hat{\mathbb{P}}_{\oplus}|x\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |x\rangle \tag{5.24}$$

and

$$\hat{\mathbb{P}}_{\oplus}|x\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} = -|y\rangle. \tag{5.25}$$

Recall from the postulates, eigenvectors are the states of an observable with definite values, which are the eigenvalues. In this setting, $\hat{\mathbb{P}}_{\oplus}$ is an observable and $|x\rangle$, $|y\rangle$ are the eigenvector states of definite polarization.

5.2.2 Polarizers Along 45°.

Suppose we orient a polarizer along the 45° axis. Again, we can use it to prepare the photon in a 45°-axis polarized state. Then, if we use a second 45°-axis polarizer, the photon will pass through with probability one. Alternatively, if we use a -45° polarizer, the photon will be reflected or absorbed, i.e. it does not pass through. If we instead chose a second polarizer along the x-axis, which is halfway between 45° and -45°, then it is either reflected or transmitted with probability $1/2$.

More precisely, when a photon passes through a 45° polarizer, it comes out in a state $|x\rangle$ or $|y\rangle$ with equal probability. Hence the state of the photon is polarized half way in between and can be written as a superposition:

$$|\nearrow\rangle = \frac{|x\rangle}{\sqrt{2}} + \frac{|y\rangle}{\sqrt{2}} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}. \tag{5.26}$$

Note, if we use a -45° polarizer, then the state vector describing the polarization along the -45° axis is orthogonal to the state vector along the 45° axis:

$$|\nwarrow\rangle = \frac{|x\rangle}{\sqrt{2}} - \frac{|y\rangle}{\sqrt{2}} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \tag{5.27}$$

such that $\langle \nearrow | \nwarrow \rangle = 0$.

Suppose we prepare a photon in state $|\nearrow\rangle$ with a 45° polarizer. If we sent it through a second vertical y-axis polarizer (see Figure 5.9), then the probability it will go through is:

$$|\langle y|\nearrow\rangle|^2 = \left|\left\langle y \left| \left(\frac{|x\rangle}{\sqrt{2}} + \frac{|y\rangle}{\sqrt{2}}\right)\right\rangle\right|^2 = \left|\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}\right|^2 = \frac{1}{2}. \tag{5.28}$$

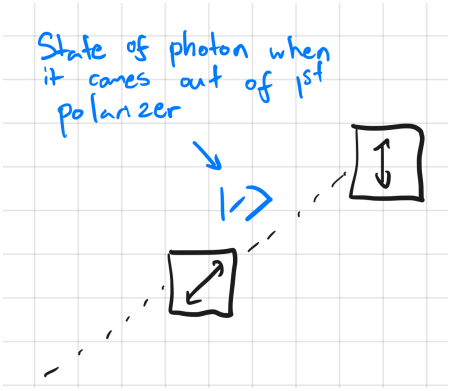


Figure 5.9: Thought experiment: 45° preparation polarizer, followed by vertical detection polarizer.

Similarly, the probability it's polarized along the x-axis after we prepare it in state $|\nearrow\rangle$ is $||x|\nearrow\rangle|^2 = 1/2$. By the same logic, if photons

$$|\langle y|\nwarrow\rangle|^2 = \left|\left\langle y \left| \left(\frac{|x\rangle}{\sqrt{2}} - \frac{|y\rangle}{\sqrt{2}}\right)\right\rangle\right|^2 = \frac{1}{2}. \tag{5.29}$$

We can now define a new observable for $\otimes$ polarization along the 45° and $-45°$ directions. As before, we will observe $+1$ and -1 for 45°-axis and $-45°$-axis polarization, respectively. Namely, the eigenvector-eigenvalue equation is:

$$\begin{aligned}\hat{\mathbb{P}}_\otimes |\nearrow\rangle &= +|\nearrow\rangle \\ \hat{\mathbb{P}}_\otimes |\nwarrow\rangle &= -|\nwarrow\rangle\end{aligned}. \tag{5.30}$$

The operator $\hat{\mathbb{P}}_\otimes$ has matrix representation and satisfies the eigenvector-eigenvalue formulation as follows:

$$\begin{aligned}\hat{\mathbb{P}}_\otimes &= \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \\ \hat{\mathbb{P}}_\otimes |\nearrow\rangle &= \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix} = \begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix} = |\nearrow\rangle \\ \hat{\mathbb{P}}_\otimes |\nwarrow\rangle &= \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix} = -\begin{bmatrix}1/\sqrt{2}\\ -1/\sqrt{2}\end{bmatrix} = -|\nwarrow\rangle.\end{aligned} \tag{5.31}$$

Remark 15 *There is no vector that is simultaneously an eigenvector of $\hat{\mathbb{P}}_{\oplus}$ and $\hat{\mathbb{P}}_{\otimes}$, as these two polarizations are incompatible. Recall, an observable is a quantity you can measure with an apparatus in a single experiment on a system. Therefore, there is no experiment that can tell one both $\oplus$-axis and $\otimes$-axis polarization simultaneously.*

Remark 16 *If $|\psi\rangle$ the state of a system with $|x\rangle$ the eigenvector of the observable, then the probability of measuring that eigenvalue is*

$$p(x) = \langle\psi|x\rangle\langle x|\psi\rangle = \psi^*(x)\psi(x). \tag{5.32}$$

In the case of polarization, $\langle\psi|x\rangle\langle x|\psi\rangle$ is the probability to find x-polarization, whereas $\langle\psi|y\rangle\langle y|\psi\rangle$ is the probability to find y-polarization. It turns out one can write the $\oplus$-axis polarization operator as a difference of dyads:

$$\hat{\mathbb{P}}_{\oplus} = |x\rangle\langle x| - |y\rangle\langle y|\,. \tag{5.33}$$

5.3 Polarizers at Arbitrary Angles

Let $|\theta\rangle$ denote the state of a photon once passed through a θ polarizer. In the $|x\rangle$, $|y\rangle$ basis, we can write this state as a linear combination:

$$|\theta\rangle = \cos\theta\,|x\rangle + \sin\theta\,|y\rangle = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}. \tag{5.34}$$

The orthogonal state for the photon polarized in $\theta + \frac{\pi}{2}$ is $|\theta + \frac{\pi}{2}\rangle$, which can be similarly expressed in the basis as:

$$\begin{aligned}|\theta + \frac{\pi}{2}\rangle &= \cos\left(\theta + \frac{\pi}{2}\right)|x\rangle + \sin\left(\theta - \frac{\pi}{2}\right)|y\rangle\\ &= -\sin\theta\,|x\rangle + \cos\theta\,|y\rangle\\ &= \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}.\end{aligned} \tag{5.35}$$

As defined, $|\theta\rangle$ and $|\theta + \frac{\pi}{2}\rangle$ are, indeed, orthogonal states:

$$\langle\theta|\theta + \frac{\pi}{2}\rangle = \begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}\begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix} = 0 \tag{5.36}$$

and they are normalized to unity $\langle\theta|\theta\rangle = \langle\theta + \frac{\pi}{2}|\theta + \frac{\pi}{2}\rangle = 1$. Hence, $|\theta\rangle$, $|\theta + \frac{\pi}{2}\rangle$ is another basis for the vector space of states.

Suppose we use a θ-axis polarizer to put a photon into state $|\theta\rangle$, as per Figure 5.10. The probability it goes through a second horizontal detector polarizer is:

$$|\langle x|\theta\rangle|^2 = \left|\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}\right|^2 = \cos^2\theta. \tag{5.37}$$

Similarly, the probability it passes through a y-polarizer is:

$$|\langle y|\theta\rangle|^2 = \left|\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}\right|^2 = \sin^2\theta. \tag{5.38}$$

Hence, this obeys the law of total probability: $|\langle x|\theta\rangle|^2+|\langle y|\theta\rangle|^2 = \cos^2\theta+\sin^2\theta = 1$.

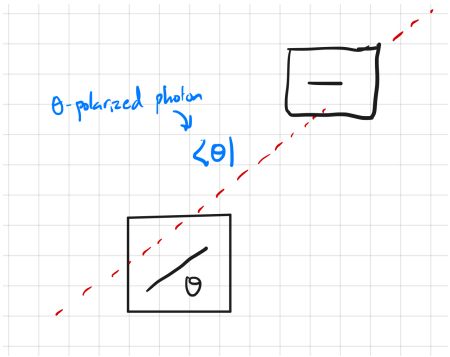


Figure 5.10: θ Polarization.

5.3.1 Arbitrary Angle Detection Polarizer

Suppose we polarize a photon with a preparation θ-axis polarizer to state $|\theta\rangle$. If we use a β-axis polarizer as a detector, as shown in Figure 5.11, we would like to compute the probability it will pass through the polarizer.

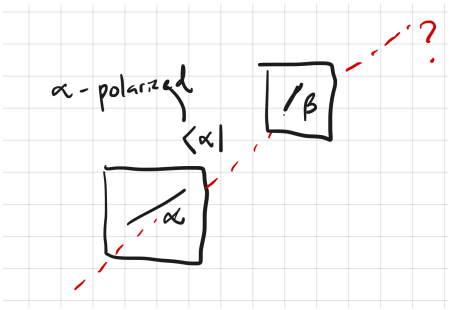


Figure 5.11: Setup: θ preparation polarization, followed by β detector polarizer.

The α-polarized photon has state:

$$|\alpha\rangle = \begin{bmatrix}\cos\alpha\\ \sin\alpha\end{bmatrix}. \tag{5.39}$$

Likewise, a β-polarized state has ket and bra vectors:

$$|\beta\rangle = \begin{bmatrix}\cos\beta\\ \sin\beta\end{bmatrix}, \langle\beta| = \begin{bmatrix}\cos\beta & \sin\beta\end{bmatrix}. \tag{5.40}$$

Hence, to test for β-polarization, we compute the amplitude:

$$\begin{aligned}\langle\beta|\alpha\rangle &= \cos\alpha\cos\beta + \sin\alpha\sin\beta\\ &= \cos(\alpha-\beta)\end{aligned} \tag{5.41}$$

and the probability of observing β-polarization for a photon in state α:

$$\mathbb{P}(\beta) = |\langle\beta|\alpha\rangle|^2 = \cos^2(\alpha-\beta). \tag{5.42}$$

Notice, if $\beta = \alpha$, the probability is 1. As β is rotated away from α, the probability gets smaller, and when $\beta = \alpha + \frac{\pi}{2}$, the probability is 0.

Consider the following thought experiment, displayed in Figure 5.12. We insert a horizontal polarizer between a θ and $\theta + \pi/2$ polarizer. If there was no horizontal polarizer in the middle, we would expect the photon to not pass through the $\theta + \pi/2$ polarizer because the first polarizer puts it into state θ. Once prepared into state $|\theta\rangle$ by the first θ-polarizer, the probability the photon goes through the horizontal polarizer is $p = \cos^2\theta$, as previously derived. The probability it goes through the third polarizer is $p = \cos^2(\alpha - \beta) = \cos^2(0 - (\theta + \pi/2)) = \cos^2(\theta + \pi/2) = \sin^2\theta$. Hence, by inserting a horizontal polarizer, the probability it goes through becomes non-zero!

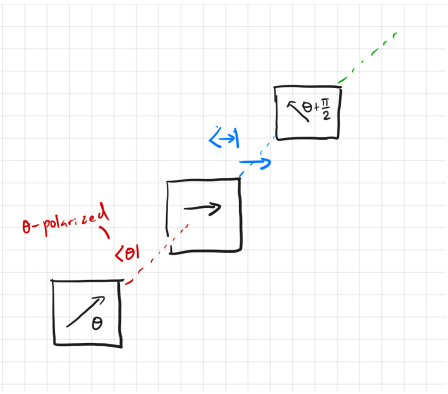


Figure 5.12: Three polarizer setup.

5.3.2 Observable for polarizer in θ-direction

For polarization in the θ-direction, the observable $\hat{\mathbb{P}}_\theta$ should satisfy:

$$\begin{aligned} \hat{\mathbb{P}}_\theta\,|\theta\rangle &= +\,|\theta\rangle \\ \hat{\mathbb{P}}_\theta\,|\theta + \frac{\pi}{2}\rangle &= -\,|\theta + \frac{\pi}{2}\rangle \end{aligned} \tag{5.43}$$

with eigenvectors $|\theta\rangle$, $|\theta + \frac{\pi}{2}\rangle$ and respective eigenvalues $+1$, -1. The matrix representation of the operator is:

$$\mathbb{P}_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}. \tag{5.44}$$

We can check that it satisfies the first eigenvector-eigenvalue equation:

$$\begin{aligned}
\mathbb{P}_\theta |\theta\rangle &= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta \\ 2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta(\cos^2\theta + \sin^2\theta) \\ 2\sin\theta\cos^2\theta + \sin^2\sin\theta - \cos^2\theta\sin\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\
&= |\theta\rangle .
\end{aligned} \tag{5.45}$$

Similarly, the observable for the orthogonal $\theta + \frac{\pi}{2}$ direction is:

$$\mathbb{P}_{\theta+\frac{\pi}{2}} = \begin{bmatrix} -\cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}. \tag{5.46}$$

5.3.3 Circular Polarization

Left-circular polarization can be represented with a state vector with complex numbers:

$$\begin{aligned}
|\circlearrowleft\rangle &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{\sqrt{2}} |x\rangle + \frac{i}{\sqrt{2}} |y\rangle \\
|\circlearrowright\rangle &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \frac{1}{\sqrt{2}} |x\rangle - \frac{i}{\sqrt{2}} |y\rangle
\end{aligned} \tag{5.47}$$

with inner product:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i^* \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = 0. \tag{5.48}$$

Consider a vertically polarized light wave with $\vec{E}$ field:

$$\begin{aligned}
E_y &= \cos(z - ct) \\
E_x &= \sin(z - ct)
\end{aligned} \tag{5.49}$$

which moves down the z-axis with the speed of light. If $z = 0$, then:

$$\begin{aligned}
E_y &= \cos(ct) \\
E_x &= -\sin(ct)
\end{aligned} \tag{5.50}$$

so the $\vec{E}$ field rotates around in a circle. This is a circularly polarized light wave, whereby each photon of the light wave is circularly polarized. Suppose we send a circularly polarized photon through a θ polarizer. The two relevant states are represented as:

$$|\circlearrowleft\rangle = \begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix}, |\theta\rangle = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \tag{5.51}$$

with amplitude given by:

$$\langle\theta|\circlearrowleft\rangle = \begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}\begin{bmatrix}1/\sqrt{2}\\ i/\sqrt{2}\end{bmatrix} = \frac{\cos\theta}{\sqrt{2}} + i\frac{\sin\theta}{\sqrt{2}}. \tag{5.52}$$

Hence, the probability of the circularly polarized photon being observed in eigenstate $|\theta\rangle$ is:

$$\begin{aligned}|\langle\theta|\circlearrowleft\rangle|^2 &= \frac{1}{2}(\cos\theta + i\sin\theta)\frac{1}{\sqrt{2}}(\cos\theta + i\sin\theta)^*\\ &= \frac{1}{2}(\cos\theta + i\sin\theta)(\cos\theta - i\sin\theta)\\ &= \frac{1}{2}(\cos^2\theta + \sin^2\theta)\\ &= \frac{1}{2}.\end{aligned} \tag{5.53}$$

This means, for a circularly polarized photon, there is a 1/2 probability it goes through a linear polarizer.

Remark 17 *The polarizer observables do not commute for different directions or types of polarization. For instance, $\hat{\mathbb{P}}_\otimes$ and $\hat{\mathbb{P}}_\oplus$ have different eigenvectors. Hence, one cannot simultaneously measure, say, horizontal and 45° polarization. On the other hand, $\hat{\mathbb{P}}_\theta$ and $\hat{\mathbb{P}}_{\theta+\frac{\pi}{2}}$ do commute, so these polarizations can be simultaneously measured, as expected.*

Chapter 6

Expectation & Conservation of Information

Let $\hat{K}$ be a observable Hermitian operator, $\{\lambda_n\}$ the collection of eigenvalues of $\hat{K}$, and $\{|n\rangle\}$ the collection of eigenvectors of $\hat{K}$. The eigenvector-eigenvalue equation satisfied by this operator is:

$$\hat{K}\,|n\rangle = \lambda_n\,|n\rangle\,. \tag{6.1}$$

We will now investigate $\langle\psi|\hat{K}|\psi\rangle$ for $|\psi\rangle$ a state of the system. Recall $\sum_n |n\rangle\langle n| = I$. Then we can write this expression as:

$$\begin{aligned}\langle\psi|\hat{K}|\psi\rangle &= \sum_n \langle\psi|\hat{K}|n\rangle\langle n|\psi\rangle \\ &= \sum_n \langle\psi|n\rangle\langle n|\psi\rangle\,\lambda_n \\ &= \sum_n |\,\langle\psi|n\rangle\,|^2\lambda_n \\ &= \sum_n \mathbb{P}_n\lambda_n,\end{aligned} \tag{6.2}$$

where in the second equality we used the fact that $\hat{K}\,|n\rangle = \lambda_n\,|n\rangle$. Notice that $\mathbb{P}_n = |\,\langle\psi|n\rangle\,|^2$ is the probability that if we were to measure $\hat{K}$, we would get the n-th eigenvalue λ_n. In particular, this is the average value of λ:

$$\bar{\lambda} = \sum_n \mathbb{P}_n\lambda_n. \tag{6.3}$$

More precisely, the **expected value (average)** of $\hat{K}$ observable is:

$$\langle\hat{K}\rangle = \sum_n \mathbb{P}_n\lambda_n = \langle\psi|\hat{K}|\psi\rangle \tag{6.4}$$

where $|\psi\rangle$ is the vector state we have prepared.

As an example, consider the xy-plane polarization observable:

$$\hat{\mathbb{P}}_{\oplus} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \tag{6.5}$$

and prepared a photon in a θ-polarized state $|\theta\rangle$. The average value of the observable is:

$$\begin{aligned}\langle\theta|\mathbb{P}_{\oplus}|\theta\rangle &= \begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}\\ &= \begin{bmatrix}\cos\theta & \sin\theta\end{bmatrix}\begin{bmatrix}\cos\theta\\ -\sin\theta\end{bmatrix}\\ &= \cos^2\theta - \sin^2\theta\\ &= \cos 2\theta.\end{aligned} \tag{6.6}$$

6.1 Expectation of Observables & Phase Shift

In classical mechanics, a state is a point among a set of points, known as the phase space. The observables are functions that label states. When the state space becomes continuous, for instance, it is labeled by momenta and positions. Hence, observable functions include position, momenta, functions of either, or functions of both.

In quantum mechanics, the state of the system is described by the ket vector $|\psi\rangle$. The observable $\hat{K}$ is a Hermitian operator, whereby we compute probabilities of $|\psi\rangle$ assuming distinct eigenvalues when measured. Due to the probabilistic nature of observation, we can calculate the average value of the observable $\langle\psi|\hat{K}|\psi\rangle$.

Suppose we have two states $|A\rangle$, $|B\rangle$ that are eigenvectors of two different observables:

$$\begin{aligned}\hat{K}\,|A\rangle &= \alpha\,|A\rangle\\ \hat{L}\,|B\rangle &= \beta\,|B\rangle\,.\end{aligned} \tag{6.7}$$

Suppose we prepare a system with a definite value of the observable $\hat{K}$, namely α. The probability that we get β when we measure $\hat{K}$ is:

$$|\langle A|B\rangle|^2 = \langle B|A\rangle\langle A|B\rangle = P_{AB} \tag{6.8}$$

where we normalize the vectors $\langle A|A\rangle = \langle B|B\rangle = 1$.

Recall from elementary complex analysis, we can represent a complex number in polar coordinate form as a complex exponential:

$$z = x + iy = r\cos\theta + r\sin\theta = re^{i\theta}, \tag{6.9}$$

where $r = |z|$ is the magnitude of z and θ is its **phase**. Then multiplying two complex numbers $z_1 = r_1e^{i\theta_1}$, $z_2 = r_2e^{i\theta_2}$ adds phases:

$$z_1z_2 = r_1r_2e^{i(\theta_1+\theta_2)}. \tag{6.10}$$

Likewise, taking the complex conjugate simply negates the phase $z^* = (re^{i\theta})^*re^{-i\theta}$.

As a result of this, in quantum mechanics, $|A\rangle$ and $e^{i\theta}\,|A\rangle$ are different vectors, but have *identical* physical properties. Consider the average value $\langle\psi|\hat{K}|\psi\rangle$ of an observable $\hat{K}$. Suppose we multiply $|\psi\rangle$ by a phase $e^{i\theta}$:

$$\begin{aligned}\langle e^{-i\theta}\psi|\hat{K}|e^{i\theta}\psi\rangle &= \langle\psi|\hat{K}|\psi\rangle\,e^{i\theta}e^{-i\theta}\\ &= \langle\psi|\hat{K}|\psi\rangle\,.\end{aligned} \tag{6.11}$$

That is, $\langle \hat{K} \rangle$ is **invariant to phase shift**. It turns out probabilities $P_{AB} = \langle B|A\rangle\langle A|B\rangle$ are also unchanged. By complex conjugation, if we multiply a ket $|B\rangle$ to get $e^{i\theta}|B\rangle$, a bra $\langle B|$ becomes $e^{-i\theta}\langle B|$. Therefore,

$$\begin{aligned} P_{AB} &= \langle B|A\rangle\langle A|B\rangle \\ &= \langle e^{-i\theta}B|A\rangle\langle A|e^{i\theta}B\rangle \\ &= \langle B|A\rangle\langle A|B\rangle \\ &= P_{AB}, \end{aligned} \tag{6.12}$$

so probabilities are also unchanged. Hence, we conclude **there is no physical significance of the phase of a wave function**.

6.1.1 Polarization Example

The basis vectors for xy-polarization are:

$$|x\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |y\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{6.13}$$

Per the conclusion in the previous section, multiplying the state by a phase shift does not change the underlying physical significance, which we denote by $\dot{=}$:

$$|x\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \dot{=} \begin{bmatrix} e^{i\theta} \\ 0 \end{bmatrix}, |y\rangle = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \dot{=} \begin{bmatrix} 0 \\ -e^{i\theta} \end{bmatrix}. \tag{6.14}$$

6.1.2 Parameter Constraints for Polarized States

Consider the complex 2-vector

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{6.15}$$

with 2 complex parameters α, β or, equivalently 4 real parameters. We impose the constraint that the probabilities sum to 1:

$$\alpha^*\alpha + \beta^*\beta = 1, \tag{6.16}$$

which means we have $4 - 1 = 3$ parameters left over. We use the additional fact that multiplying by phase does not change the content of the state. Hence, we can always multiply by a phase to make the upper entry α be real:

$$\begin{bmatrix} a \\ \beta \end{bmatrix} \tag{6.17}$$

where we require:

$$\begin{aligned} a^2 + \beta^*\beta &= 1 \\ \beta^*\beta &= 1 - a^2 \\ \beta &= \sqrt{1-a^2}e^{i\phi}. \end{aligned} \tag{6.18}$$

It follows that the most general form for the polarization states of a photon is a two-parameter family:

$$\begin{bmatrix} a \\ \sqrt{1-a^2}e^{i\phi} \end{bmatrix}. \tag{6.19}$$

Consider a polarized electromagnetic wave with $\vec{E}$ field:

$$\begin{aligned} E_x &= \cos(x-ct) \\ E_y &= \sin(x-ct). \end{aligned} \tag{6.20}$$

When $x = 0$, we get $E_x = \cos(ct)$, $E_y = -\sin(ct)$, so the electric field moves around a circle. The more general form described by Equation 6.19 described an electric field that moves around an ellipse. In fact, we can create elliptical polarization along any axis, where the two parameters are (1) the angle of the major axis, measured from the horizontal θ and (2) the eccentricity of the ellipse (aspect ratio of the minor to major diameters). When eccentricity approaches zero, the ellipse becomes infinitely thin and collapses to a plane, i.e. the $\vec{E}$ field becomes plane-polarized. When eccentricity approaches 1, the field becomes circularly polarized so θ doesn't matter. Hence, it's not surprising a, ϕ specify the state, as there is a one-to-one correspondence between $(\theta, \text{eccentricity})$ and (a, ϕ) in this example.

Consider an elliptically polarized photon in state:

$$|\psi\rangle = \begin{bmatrix} a \\ \sqrt{1-a^2}e^{i\phi} \end{bmatrix}. \tag{6.21}$$

Let $|\theta\rangle$ be the state of θ plane polarization:

$$|\theta\rangle = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{6.22}$$

The probability amplitude the elliptically polarized photon is measured along the θ plane polarization axis is:

$$\langle\theta|\psi\rangle = a\cos\theta + \sqrt{1-a^2}e^{-i\phi}\sin\theta, \tag{6.23}$$

so the probability is simply $|\langle\theta|\psi\rangle|^2$. If we fix ϕ, we can find the major axis by computing:

$$\theta^* = \arg\max_\theta \langle\theta|\psi\rangle\langle\psi|\theta\rangle. \tag{6.24}$$

6.2 Evolution in Quantum Mechanics

In classical mechanics, evolution is simply a permutation of the states. The conservation property is if two states are different and we evolve time, they remain different. In quantum mechanics, for two different states $|A\rangle$ and $|B\rangle$, the notion of similarity and difference is encoded in the inner product $\langle B|A\rangle$. If two states are the same, up to phase, we write $|A\rangle \triangleq |B\rangle$. On the other hand, if they are different, they must be orthogonal with inner product $\langle B|A\rangle = 0$. That is, if $|A\rangle$

and $|B\rangle$ are similar, $\langle B|A\rangle$ is close to 1. Similarly, if $|B|A\rangle = 1$ then $|A\rangle = |B\rangle$. Hence, the inner product is the basic relationship between states.

Suppose the system evolves in time:

$$\begin{aligned} |A\rangle &\overset{\text{evolve time}}{\longrightarrow} |A'\rangle \\ |B\rangle &\overset{\text{evolve time}}{\longrightarrow} |B'\rangle \end{aligned} \tag{6.25}$$

then we insist that relationships between states remain fixed. That is, the **conservation of information** (relationships between states) is:

$$\langle B'|A'\rangle = \langle B|A\rangle\,. \tag{6.26}$$

The mnemonic is that the *angles between vectors in the space of states stays fixed in time, even if the vectors themselves vary in time.*

6.2.1 Hermitian Conjugates and Operators

Let $\hat{L}$ be a linear operator acting on kets as $\hat{L}\,|A\rangle = |C\rangle$. Using the $\sum_n |n\rangle\langle n| = I$ identity, we can determine its action on bras:

$$\sum_n \langle B|\hat{L}|n\rangle\langle n| = \langle B|\,\hat{L}. \tag{6.27}$$

We can view this as

$$\sum_n \langle B|\hat{L}|n\rangle\langle n| = \sum_n \langle B|[\hat{L}|n\rangle]\,\langle n|\,, \tag{6.28}$$

whereby $\hat{L}\,|n\rangle$ produces a ket which gives a complex number when we take its inner product with $\langle B|$, namely $\langle B|[\hat{L}|n\rangle]$. These are the coefficients in the bra basis $\{\langle n|\}$. Recall, we can extract the BA element of $\hat{L}$ as follows:

$$\langle B|\hat{L}|A\rangle = \langle B|[\hat{L}|A\rangle], \tag{6.29}$$

which we can now compute with the bra basis expansion.

While we have established the notion of a linear operator acting from the left on kets $\hat{L}\,|A\rangle = |C\rangle$, there is an analogous mapping for bra vectors, where a linear operator $\hat{L}^\dagger$ acts from the left:

$$\langle A|\,\hat{L}^\dagger = \langle C|\,. \tag{6.30}$$

Taking its inner product of $\hat{L}\,|A\rangle = |C\rangle$ with $\langle B|$ and $\langle A|\,\hat{L}^\dagger = \langle C|$ with $|B\rangle$ yields:

$$\begin{aligned} \langle B|\hat{L}|A\rangle &= \langle B|C\rangle \\ \langle A|\hat{L}^\dagger|B\rangle &= \langle C|B\rangle\,, \end{aligned} \tag{6.31}$$

which are complex conjugates of each other. Hence, we find the defining identity for the **Hermitian conjugate** $\hat{L}^\dagger$:

$$\langle B|\hat{L}|A\rangle = \langle A|\hat{L}^\dagger|B\rangle^* . \tag{6.32}$$

The matrix representation specifies the full operator, whose elements can be written as:

$$\langle n|\hat{L}^\dagger|m\rangle = L^\dagger_{nm} = \langle m|L|n\rangle^* = L^*_{mn}. \tag{6.33}$$

It follows that the matrix elements of the Hermitian conjugate $\hat{L}^\dagger$ can be expressed as:

$$L^\dagger_{mn} = L^*_{mn}. \tag{6.34}$$

As an example, if we have the following linear operator matrix representation:

$$\hat{L} = \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{bmatrix}, \tag{6.35}$$

then the Hermitian conjugate transposes and applies entry-wise complex conjugation:

$$\hat{L}^\dagger = \begin{bmatrix} L^*_{11} & L_{21} & L^*_{31} \\ L^*_{12} & L^*_{22} & L^*_{32} \\ L^*_{13} & L^*_{23} & L^*_{33} \end{bmatrix}. \tag{6.36}$$

Remark 18 *The* ***Hermitian operator*** $\hat{L}$ *is a special linear operator that satisfies:*

$$\hat{L} = \hat{L}^\dagger = (\hat{L}^\top)^*. \tag{6.37}$$

This means $L_{ii} = L^*_{ii}$ *so its diagonal elements* L_{ii} *must be real. As before, its off-diagonal elements are reflected about the diagonal and complex conjugated such that:*

$$\hat{L} = \begin{bmatrix} L_{11} & L^*_{21} & L^*_{31} \\ L^*_{12} & L_{22} & L^*_{32} \\ L^*_{13} & L_{23} & L^*_{33} \end{bmatrix}, \tag{6.38}$$

for a 3×3 *example.*

Example 6.2.1 *The following operators are Hermitian and satisfy* $\mathbb{P} = \mathbb{P}^\dagger$*:*

$$\begin{aligned} \mathbb{P}_\oplus &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ \mathbb{P}_\otimes &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \mathbb{P}_\circlearrowleft &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \end{aligned} \tag{6.39}$$

Remark 19 *If the expectation* $\langle\psi|\hat{L}|\psi\rangle$ *of* $\hat{L}$ *is real for all* ψ*, it follows that* $\hat{L}$ *is Hermitian.*

Remark 20 *The eigenvectors of a Hermitian operator span the vector space (state space), i.e.* ***every Hermitian operator has a complete orthonormal collection of eigenvectors with real eigenvalues.***

6.2.2 Unitary Operators

Let $|A\rangle$, $|B\rangle$ be states in a vector space. Their similarity is measured by the inner product $\langle A|B\rangle$. Suppose we apply a linear operator U that changes, i.e. evolves, the states:

$$\begin{aligned} U|B\rangle &= |B'\rangle \\ U|A\rangle &= |A'\rangle \end{aligned} \tag{6.40}$$

As discussed in the previous section, the same operation has an image among bra vectors, namely the Hermitian conjugate:

$$\langle A| U^\dagger = \langle A'| . \tag{6.41}$$

We require that such an evolution under U does not disturb the relationship between states by enforcing the following condition:

$$\begin{aligned} \langle A|B\rangle &= \langle A'|B'\rangle \\ &= \langle A|U^\dagger U|B\rangle \end{aligned} \tag{6.42}$$

for *any* pair $|A\rangle$, $|B\rangle$. Therefore,

$$U^\dagger U = I,\ U^\dagger = U^{-1}. \tag{6.43}$$

Definition 7 *A **unitary operator** U satisfies $U^\top U = I$. That is, unitary operators preserve the inner product between states and, hence, their relationships.*

Consequently, as seen in the introduction of Section 6.1, *unitary operators preserve expectations and probabilities.*

Chapter 7

Evolution & The Schrödinger Equation

In classical mechanics, each state is a point in a phase space. The time evolution of the states must be made by discretizing time because there is no continuous way to permute states from one configuration to another. The **Hamiltonian** $H(x,p)$ is the energy of system as a function of position and momenta. For instance, one could define the Hamiltonian as the sum of kinetic and potential energy:

$$H(x,p) = \frac{p^2}{2m} + U(x). \tag{7.1}$$

The law of evolution is dictated by the set of flow lines in phase space, as seen in Figure 7.1. In particular, **Hamilton's equations** prescribe the flow:

$$\begin{aligned} \frac{\partial H}{\partial p} &= \frac{p}{m} = v = \dot{x} \\ \frac{\partial H}{\partial x} &= \frac{\partial U}{\partial x} = -F(x) = -\dot{p} \end{aligned} \tag{7.2}$$

where momentum is $p = mv$ and its derivative is force $\dot{p} = ma = F$. Therefore, the Hamilton equations are:

$$\begin{aligned} \frac{\partial H}{\partial p} &= \dot{x} \\ \frac{\partial H}{\partial x} &= -\dot{p} \end{aligned} \quad . \tag{7.3}$$

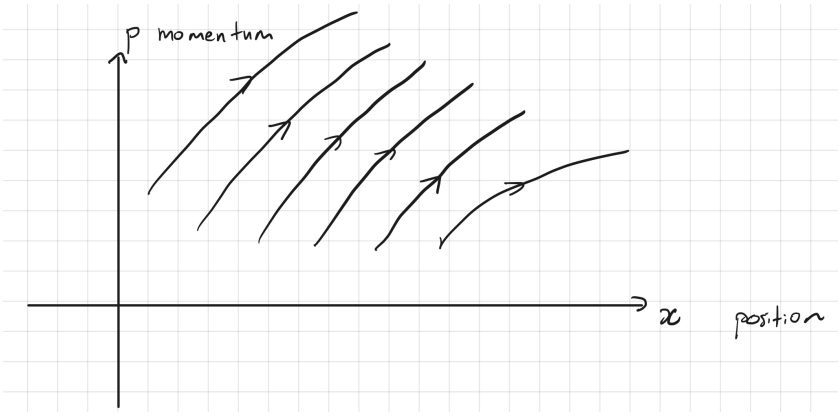


Figure 7.1: Flow lines in phase space described by Hamiltonian.

Observables are arbitrary functions of position and momentum $F(p, x)$. We will analyze its evolution by computing the first time derivative and substituting Hamilton's equations:

$$\begin{aligned}\frac{d}{dt}F(p,x) &= \frac{\partial F}{\partial p}\dot{p} + \frac{\partial F}{\partial x}\dot{x} \\ &= \frac{\partial F}{\partial p}\left(-\frac{\partial H}{\partial x}\right) + \frac{\partial F}{\partial x}\frac{\partial H}{\partial p} \\ &= -\frac{\partial F}{\partial p}\frac{\partial H}{\partial x} + \frac{\partial F}{\partial x}\frac{\partial H}{\partial p} \\ &:= \{F, H\}\end{aligned} \tag{7.4}$$

where $\{F, H\} = -\frac{\partial F}{\partial p}\frac{\partial H}{\partial x} + \frac{\partial F}{\partial x}\frac{\partial H}{\partial p}$ is the **Poisson bracket**. Therefore, the Hamiltonian determines a flow and if we have an observable quantity F, we can determine how it changes with time along the flow by Poisson brackets:

$$\frac{d}{dt}F(p,x) = F, H. \tag{7.5}$$

The **Hamiltonian is a generator of time evolution**.

In quantum mechanics, the basic rule of time evolution is the relationships between states does not change with time. Suppose states $|A\rangle$, $|B\rangle$ evolves over a period of time T by way of a unitary operator U:

$$\begin{aligned}|A\rangle &\xrightarrow{T} U(T)\,|A\rangle \\ |B\rangle &\xrightarrow{T} U(T)\,|B\rangle\end{aligned} \tag{7.6}$$

where $U^\dagger U = I$. For bra vectors, we effectively complex conjugate the equations, replacing U with its Hermitian conjugate $U^\dagger$:

$$\begin{aligned}\langle A| &\xrightarrow{T} \langle A|\,U(T)^\dagger \\ \langle B| &\xrightarrow{T} \langle B|\,U(T)^\dagger.\end{aligned} \tag{7.7}$$

We would like to analyze how the inner product $\langle B|A\rangle$ evolves. We note that $|\langle B|A\rangle|^2$ is the probability of B occurring if we are in state A–similar to a transition probability in Markov decision processes. The inner product evolves as:

$$\langle B|A\rangle \xrightarrow{T} \langle B|U^\dagger U|A\rangle = \langle B|A\rangle, \tag{7.8}$$

i.e. it is preserved. Hence, states that are orthogonal stay orthogonal as we evolve the system.

7.0.1 The Schrödinger Equation Derivation

Let $|\psi(t)\rangle$ be the state vector at time t. Suppose we evolve by T seconds with the unitary operator $U(T)$:

$$|\psi(t)\rangle \longrightarrow |\psi(t+T)\rangle = U(T)\,|\psi(t)\rangle\,. \tag{7.9}$$

If $T = 0$, $|\psi(t)\rangle = U(0)\psi(t)$ so $U(0) = I$. Suppose we divide the evolution into smaller time intervals. The corresponding **evolution operator** is:

$$U(\epsilon) = I - \frac{i\epsilon}{\hbar}H \tag{7.10}$$

where $-\frac{i\epsilon}{\hbar}H$ is a $O(\epsilon)$ correction to $U(0) = I$. The Hermitian conjugate is thus:

$$U^\dagger(\epsilon) = I + \frac{i\epsilon}{\hbar}H^\dagger. \tag{7.11}$$

As before, we require that relationships are invariant under time evolution:

$$\begin{aligned} U^\dagger(\epsilon)U(\epsilon) &= I \\ \left(I + \frac{i\epsilon}{\hbar}H^\dagger\right)\left(I - \frac{i\epsilon}{\hbar}H\right) &= I. \end{aligned} \tag{7.12}$$

We expand this further, ignoring $O(\epsilon^2)$ terms since we only care for first order $O(\epsilon)$ terms:

$$\begin{aligned} I + \frac{i\epsilon}{\hbar}(H^\dagger - H) &= I \\ \frac{i\epsilon}{\hbar}(H^\dagger - H) &= 0 \\ H^\dagger &= H. \end{aligned} \tag{7.13}$$

Hence, **for U to be unitary, H must be Hermitian**. Since H is Hermitian, it must be an observable. That is, H has an orthonormal basis of eigenstates, all of its eigenvalues are real, and it can be measured in a lab. In fact, H is a Hamiltonian and its eigenvalues are called **energy**.

Evolving the system ϵ seconds into the future:

$$\begin{aligned} |\psi(t+\epsilon)\rangle &= U(\epsilon)\,|\psi(t)\rangle \\ &= \left(I - \frac{i\epsilon}{\hbar}H\right)|\psi(t)\rangle\,. \end{aligned} \tag{7.14}$$

Taking the difference between states at different times we find:

$$\begin{aligned} |\psi(t+\epsilon)\rangle - |\psi(t)\rangle &= -\frac{i\epsilon H}{\hbar}|\psi(t)\rangle \\ \frac{|\psi(t+\epsilon)\rangle - |\psi(t)\rangle}{\epsilon} &= -\frac{iH}{\hbar}|\psi(t)\rangle \\ \frac{\partial|\psi\rangle}{\partial t} &= -\frac{iH}{\hbar}|\psi\rangle . \end{aligned} \tag{7.15}$$

Hence, we know how $|\psi\rangle$ changes assuming we know the Hamiltonian, given by:

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle , \tag{7.16}$$

which is known as **the Schrödinger equation**.

7.1 Special Material Example

Suppose we have created a material whereby a photon of some wavelength has energy E_1 or E_2 *depending on which polarization it has*. This has a translation into a statement about the Hamiltonian:

$$\begin{aligned} H|x\rangle &= E_1|x\rangle \\ H|y\rangle &= E_2|y\rangle . \end{aligned} \tag{7.17}$$

That is, the eigenvalues of H are E_1, E_2 depending on the corresponding polarization states $|x\rangle$, $|y\rangle$. Start with an initial state of photon polarization at time t:

$$|\psi(t)\rangle = \begin{bmatrix}\alpha(t)\\ \beta(t)\end{bmatrix} = \alpha(t)|x\rangle + \beta(t)|y\rangle \tag{7.18}$$

where $\alpha(t)$ and $\beta(t)$ are the amplitudes the photon has x and y polarization, respectively. Similarly, $\alpha^*(t)\alpha(t)$ and $\beta^*(t)\beta(t)$ are the probabilities the photon has x and y polarization, respectively. The Schrödinger equation is:

$$\begin{aligned} i\hbar\frac{\partial|\psi\rangle}{\partial t} &= H|\psi\rangle \\ i\hbar\begin{bmatrix}\dot\alpha(t)\\ \dot\beta(t)\end{bmatrix} &= H\begin{bmatrix}\alpha(t)\\ \beta(t)\end{bmatrix} \end{aligned} \tag{7.19}$$

where the Hamiltonian is:

$$H = \begin{bmatrix}E_1 & 0\\ 0 & E_2\end{bmatrix} \tag{7.20}$$

wit E_1, E_2 eigenvalues since $H|x\rangle = E_1|x\rangle$ and $H|y\rangle = E_2|y\rangle$. Thus, the Schrödinger equation becomes:

$$i\hbar\begin{bmatrix}\dot\alpha(t)\\ \dot\beta(t)\end{bmatrix} = \begin{bmatrix}E_1 & 0\\ 0 & E_2\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \begin{bmatrix}E_1\alpha\\ E_2\beta\end{bmatrix} . \tag{7.21}$$

Hence, we have a 2×2 system of ordinary differential equations:

$$\begin{aligned} \frac{d\alpha}{dt} &= -\frac{iE_1}{\hbar}\alpha(t) \\ \frac{d\beta}{dt} &= -\frac{iE_2}{\hbar}\beta(t). \end{aligned} \tag{7.22}$$

Let $\alpha(0) = \alpha$, $\beta(0) = \beta$, then we find $\alpha(t) = \alpha e^{-\frac{iE_1}{\hbar}t}$, $\beta(t) = \beta e^{-\frac{iE_2}{\hbar}t}$ are solutions. Hence, the state of the photon is:

$$|\psi(t)\rangle = \begin{bmatrix} \alpha e^{-\frac{iE_1}{\hbar}t} \\ \beta e^{-\frac{iE_2}{\hbar}t} \end{bmatrix}. \tag{7.23}$$

The probability the photon is polarized along the x-axis is:

$$\begin{aligned} p_x &= \langle x|\psi(t)\rangle\langle\psi(t)|x\rangle \\ &= \left(\alpha e^{-\frac{iE_1}{\hbar}t}\right)^* \left(\alpha e^{-\frac{iE_1}{\hbar}t}\right) \\ &= \alpha^*\alpha \end{aligned} \tag{7.24}$$

and, likewise, the probability it is polarized along the y-axis is $p_y = \beta^*\beta$. That is, *the probabilities haven't changed even after evolving the state in time.*

Suppose we run the photon through a $\otimes$ polarizer instead with state:

$$|\nearrow\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}. \tag{7.25}$$

The amplitude the photon will go through a 45°-axis polarizer is:

$$\begin{aligned} \langle\nearrow|\psi\rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \alpha e^{-\frac{iE_1}{\hbar}t} \\ \beta e^{-\frac{iE_2}{\hbar}t} \end{bmatrix} \\ &= \frac{\alpha}{\sqrt{2}} e^{-\frac{iE_1}{\hbar}t} + \frac{\beta}{\sqrt{2}} e^{-\frac{iE_2}{\hbar}t} \end{aligned}. \tag{7.26}$$

The probability it goes through the 45° polarizer is $|\langle\nearrow|\psi\rangle|^2$.

Suppose we have two 45° polarizers with the strange material completely filling the space between the polarizers. Ordinarily, without the special material, one would expect the photon to go right through with probability one. Suppose $\alpha(0) = \alpha = 1/\sqrt{2}$, $\beta(0) = \beta = 1/\sqrt{2}$, i.e.

$$|\psi(0)\rangle = \begin{bmatrix} \alpha(0) \\ \beta(0) \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \tag{7.27}$$

is the state right after coming out of the first 45° polarizer. Then the photon propagates for time t in the special material with different energies depending on polarization:

$$|\psi(t)\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} e^{-\frac{iE_1}{\hbar}t} \\ \frac{1}{\sqrt{2}} e^{-\frac{iE_2}{\hbar}t} \end{bmatrix}. \tag{7.28}$$

This corresponds to the experiment shown in Figure 7.2.

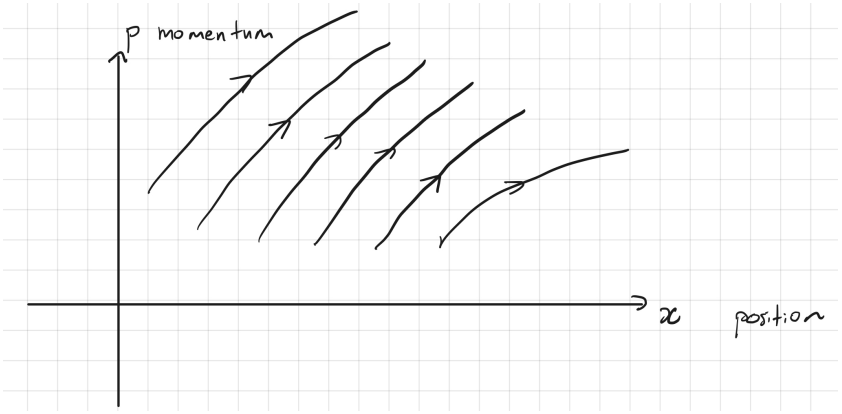


Figure 7.2: Special material permeating space between two 45° polarizers.

At the second 45° polarizer, we would like to know the probability the photon passes through. That is, the amplitude it is polarized exactly like the first 45° polarizer and passes through the second polarizer is:

$$\begin{aligned}\langle \nearrow | \psi(t) \rangle &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} e^{-\frac{iE_1}{\hbar}t} \\ \frac{1}{\sqrt{2}} e^{-\frac{iE_2}{\hbar}t} \end{bmatrix}. \\ &= \frac{1}{2} e^{-\frac{iE_1}{\hbar}t} + \frac{1}{2} e^{-\frac{iE_2}{\hbar}t}\end{aligned} \tag{7.29}$$

The probability it passes through the second polarizer is:

$$\begin{aligned} p &= |\langle \nearrow | \psi(t) \rangle|^2 \\ &= \langle \psi(t) | \nearrow \rangle \langle \nearrow | \psi(t) \rangle \\ &= \frac{1}{2} \left(e^{-\frac{iE_1}{\hbar}t} + e^{-\frac{iE}{\hbar}t} \right)^* \frac{1}{2} \left(e^{-\frac{iE_1}{\hbar}t} + e^{-\frac{iE}{\hbar}t} \right), \\ &= \frac{1}{4} \left(e^{-\frac{iE_1}{\hbar}t} + e^{-\frac{iE}{\hbar}t} \right) \left(e^{\frac{iE_1}{\hbar}t} + e^{\frac{iE}{\hbar}t} \right) \\ &= \frac{1}{4} \left(2 + e^{\frac{i(E_2 - E_1)}{\hbar}t} + e^{-\frac{i(E_2 - E_1)}{\hbar}t} \right) \end{aligned} \tag{7.30}$$

where $E_2 - E_1$ is the energy difference of the two polarization states. That is the probability the photon passes through the second polarizer only depends on energy differences, which is a deep idea in physics. Let $\Delta E = E_2 - E_1$, then the probability is:

$$\begin{aligned} p &= \frac{1}{4} \left(2 + e^{\frac{i\Delta E}{\hbar}t} + e^{-\frac{i\Delta E}{\hbar}t} \right) \\ &= \frac{1}{2} + \frac{1}{2} \frac{\cos \Delta E}{\hbar} t \end{aligned} . \tag{7.31}$$

At $t = 0$, the probability to get through the second polarizer is $p = 1$ because the system has not been given time to evolve so it is still 45°-polarized. Hence, the

probability a $-45°$-polarized photon passes is zero. At a time t such that $\frac{\Delta E}{\hbar}t = \pi$, a $-45°$-polarized photon passes with probability 1, whereas a $45°$-polarized photon does not since $p = 0$. Hence, in some sense, there is a continuous rotation of the angle of polarization.

Remark 21 *Note, if $\Delta E = 0$, then $p = 1$. Hence, if the two energies are the same, the system will behave as if no material was there at all. If one adds a constant number to the energies E_1, E_2 which satisfy the eigenvector-eigenvalue equations $H|x\rangle = E_1|x\rangle$, $H|y\rangle = E_2|y\rangle$, there is no effect on what is measured. That is, only energy differences are important!*

7.2 Evolution of the Average Observable

Suppose we have an observable $\hat{K}$ operator. The average value in state ψ is $\langle K\rangle = \langle\psi|\hat{K}|\psi\rangle$. Since $\hat{K}$ is fixed, the only component of this that changes with time is the state $|\psi\rangle$. We apply the product rule to compute its time derivative:

$$\frac{d}{dt}\langle\psi|\hat{K}|\psi\rangle = \langle\psi|\hat{K}|\dot{\psi}\rangle + \langle\dot{\psi}|\hat{K}|\psi\rangle\,. \tag{7.32}$$

Notice, in the product rule, we decomposed the average as first the linear operator acting on a ket followed by an inner product with a bra vector: $\langle\psi|\hat{K}|\psi\rangle = \langle\psi|[\hat{K}|\psi\rangle]$. The Schrödinger equation

$$|\dot{\psi}\rangle = -\frac{iH}{\hbar}|\psi\rangle \tag{7.33}$$

has the corresponding formulation for a bra vector:

$$\begin{aligned}\langle\dot{\psi}| &= \left(-\frac{iH}{\hbar}|\psi\rangle\right)^{\dagger}\\ &= \frac{i}{\hbar}\langle\psi|\,H^{\dagger}\\ &= \frac{i}{\hbar}\langle\psi|\,H.\end{aligned} \tag{7.34}$$

Returning to the derivative calculation, we find:

$$\begin{aligned}\frac{d}{dt}\langle\psi|\hat{K}|\psi\rangle &= \langle\psi|\hat{K}|\dot{\psi}\rangle + \langle\dot{\psi}|\hat{K}|\psi\rangle\\ &= \left\langle\psi\middle|\,\hat{K}\left(\frac{-i}{\hbar}H\right)\middle|\,\psi\right\rangle + \left\langle\frac{i}{\hbar}\psi\middle|\,H\hat{K}\,\middle|\,\psi\right\rangle\\ &= -\frac{i}{\hbar}\langle\psi|\hat{K}H|\psi\rangle + \frac{i}{\hbar}\langle\psi|H\hat{K}|\psi\rangle\\ &= -\frac{i}{\hbar}\langle\psi|\hat{K}H - H\hat{K}|\psi\rangle\\ &= -\frac{i}{\hbar}\langle\psi|[\hat{K},H]|\psi\rangle\\ &= -\frac{i}{\hbar}\langle[\hat{K},H]\rangle\,,\end{aligned} \tag{7.35}$$

where $[K, H] = \hat{K}H - H\hat{K}$.

Theorem 4 *The evolution of the average value of an observable is given by the Hamiltonian H:*

$$\frac{d}{dt}\langle\psi|\hat{K}|\psi\rangle = -\frac{i}{\hbar}\langle\psi|[\hat{K}, H]|\psi\rangle \tag{7.36}$$

or, equivalently:

$$\frac{d}{dt}\langle\hat{K}\rangle = -\frac{i}{\hbar}\langle[\hat{K}, H]\rangle. \tag{7.37}$$

If the observable commutes with the Hamiltonian, then $\frac{d}{dt}\langle\hat{K}\rangle = 0$, which is a *weak conservation of energy law.* That is, the expectation value of the quantity $\hat{K}$ does not change if $\hat{K}$ commutes with the Hamiltonian.

Remark 22 *In the special case where $\hat{K} = H$, $\frac{d}{dt}\langle H\rangle = 0$, i.e.* ***the Hamiltonian is a conserved quantity.***

Remark 23 *Our evolution law is analogous to classical mechanics, where one defines a Poisson bracket:*

$$\{A, B\} = \frac{\partial A}{\partial x}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial x}. \tag{7.38}$$

It follows that $\{H, H\} = 0$ for the Hamiltonian H, so energy is conserved. Likewise, if $\{A, H\} = 0$ for a quantity A, then A is conserved.

7.3 Particle on a Line

Consider a particle on a line described by its wave function $\psi(x, t)$, which encodes information about the probability of being in position x at time t. We would like to study the evolution $\frac{\partial\psi(x,t)}{\partial t}$. The wave function $\psi(x, t)$ can be thought of as a representation of the state vector $|\psi(x, t)\rangle$ in Hilbert space. Recall, the Schrödinger equation is:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = H\psi(x, t). \tag{7.39}$$

Classically, we know a particle on a line has Hamiltonian (energy:

$$E = H = \frac{p^2}{2m}. \tag{7.40}$$

In quantum mechanics, the momentum becomes the operator $\hat{P} = -i\hbar\frac{\partial}{\partial x}$. Therefore,

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = H\psi(x, t) = \frac{\hat{P}^2}{2m}\psi(x, t) \tag{7.41}$$

where $\hat{P}^2\psi = -i\hbar\frac{\partial}{\partial x}\left(-i\hbar\frac{\partial\psi}{\partial x}\right) = -\hbar^2\frac{\partial^2\psi}{\partial x^2}$. It follows that

$$\frac{\partial\psi}{\partial t} = i\frac{\hbar}{2m}\frac{\partial^2\psi}{\partial x^2}, \tag{7.42}$$

which is the Schrödinger wave equation for a free particle on a line with no forces on it.

Remark 24 *The momentum operator $\hat{P}$ is Hermitian, so $\hat{P}^2$ is Hermitian and, thus, so is the Hamiltonian H.*

The eigenvectors of the momentum operator are wave functions with definite momentum $e^{i\frac{p}{\hbar}x}$. Suppose we makke an ansatz for the time-dependent solution of the Schrödinger wave equation for a free particle on a line:

$$\psi(x,t) = f(t)e^{i\frac{p}{\hbar}x}. \tag{7.43}$$

This is still an eigenvector of momentum since $\hat{P}$ only differentiates with respect to position. Computing the first time derivative:

$$\dot{\psi} = \dot{e}^{i\frac{p}{\hbar}x}, \tag{7.44}$$

we plug the ansatz into the Schrödinger equation:

$$\begin{aligned} i\frac{\hbar}{2m}\frac{\partial^2\psi}{\partial x^2} &= i\frac{\hbar}{2m}f(t)\left(\frac{ip}{\hbar}\right)^2 e^{i\frac{p}{\hbar}x} \\ \dot{f}e^{i\frac{p}{\hbar}x} &= i\frac{\hbar}{2m}f(t)\left(\frac{ip}{\hbar}\right)^2 e^{i\frac{p}{\hbar}x} \\ \dot{f} &= -\frac{i}{2m\hbar}p^2 f(t) \\ f(t) &= e^{-\frac{i}{\hbar}\frac{p^2}{2m}t}. \end{aligned} \tag{7.45}$$

Therefore, the time-dependent solution of the Schrödinger wave equation corresponding to a momentum eigenstate is:

$$\psi(x,t) = e^{i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t}. \tag{7.46}$$

Chapter 8

General Solutions to The Schrödinger Equation

The general Schrödinger equation is:

$$i\hbar\frac{\partial\left|\psi\right\rangle}{\partial t}=\hat{H}\left|\psi\right\rangle. \tag{8.1}$$

In classical mechanics, a particle moving along a line has Hamiltonian $H=\frac{p^2}{2m}$. In quantum mechanics, we write the Hamiltonian as:

$$\hat{H}=\frac{\hat{P}^2}{2m} \tag{8.2}$$

where the momentum operator is $\hat{P}=-\hbar\frac{\partial}{\partial x}$. As in the previous chapter, we identify the wave function representation $\psi(x,t)$ with the state $|\psi\rangle$. Hence, the Schrödinger wave equation becomes:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=\frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}\right)\left(-i\hbar\frac{\partial}{\partial x}\right)\psi(x,t)=-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}. \tag{8.3}$$

We make an ansatz that $\psi(x,t)$ is separable in space and time dependence:

$$\psi(x,t)=e^{i\frac{p}{\hbar}x}e^{i\omega t} \tag{8.4}$$

where, recall, the eigenvectors of $\hat{P}$ have the form $|P\rangle=e^{i\frac{p}{\hbar}x}$ in wave function form. We compute the time derivative of the ansatz:

$$\frac{\partial\psi}{\partial t}=i\omega\psi(x,t). \tag{8.5}$$

Likewise, we compute the second order space derivative:

$$\begin{aligned}\frac{\partial\psi}{\partial x}&=i\frac{p}{\hbar}\psi(x,t)\\ \frac{\partial^2\psi}{\partial x^2}&=\left(i\frac{p}{\hbar}\right)^2\psi(x,t).\end{aligned} \tag{8.6}$$

Hence, the Schrödinger wave equation simplifies to:

$$\begin{aligned} i\hbar\frac{\partial\psi}{\partial t} &= -\omega\hbar\psi(x,t) \\ &= -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} \\ &= -\frac{\hbar^2}{2m}\left(i\frac{p}{\hbar}\right)^2\psi \\ &= \frac{p^2}{2m}\psi(x,t). \end{aligned} \tag{8.7}$$

Solving for ω, we find:

$$\omega = -\frac{p^2}{2m\hbar} \tag{8.8}$$

such that

$$\psi(x,t) = e^{i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t}. \tag{8.9}$$

Note, this is the wave function for a positive momentum. The negative p solution is given by:

$$\psi(x,t) = e^{-i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t}. \tag{8.10}$$

This will have a different direction for momenta, but the same energy $\frac{p^2}{2m}$ as the positive p solution. Hence, we can combine solutions in a superposition:

$$\psi(x,t) = \alpha e^{i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t} + \beta e^{-i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t}. \tag{8.11}$$

The probability the particle is moving to the left (negative p) and right (positive p) is $\alpha^*\alpha$ and $\beta^*\beta$, respectively. The probability it has energy $\frac{p^2}{2m}$, however, is one.

The general solution of Schrödinger's equation is the integral of these solutions over all possible momenta p:

$$\psi(x,t) = \int_{-\infty}^{\infty}\frac{dp}{\sqrt{2\pi}}\left[e^{i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t}\tilde{\psi}(p)\right], \tag{8.12}$$

i.e. a quantum superposition of states with different momenta and, consequently, energies, where $\tilde{\psi}$ is the Fourier transform. The probability density that a particle has momentum q is:

$$p(q) = \tilde{\psi}^*(q)\tilde{\psi}(q). \tag{8.13}$$

We are also interested in the probability the energy is E. Recall $E = \frac{p^2}{2m}$ so $p^2 = 2mE$. Hence, there are two possible momenta corresponding to E:

$$p_\pm = \pm\sqrt{2mE}, \tag{8.14}$$

so, by the law of total probability we add the respective probabilities corresponding to p_+ and p_-:

$$p(E) = \tilde{\psi}^*(p_+)\tilde{\psi}(p_+) + \tilde{\psi}^*(p_-)\tilde{\psi}(p_-). \tag{8.15}$$

Remark 25 *When $t = 0$:*

$$\int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}} e^{ipx} \tilde{\psi}(p) = \psi(x, 0) := \psi(x) \tag{8.16}$$

, which is the definition of the Fourier transform of $\psi(x)$. Indeed, we can reconstruct $\tilde{\psi}(p)$ from $\psi(x, 0) = \psi(x)$ as:

$$\tilde{\psi}(p) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ipx} \psi(x) \tag{8.17}$$

which we can substitute into the general solution:

$$\psi(x, t) = \int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}} \left[e^{i\frac{p}{\hbar}x - i\frac{p^2}{2m\hbar}t} \tilde{\psi}(p) \right]. \tag{8.18}$$

In this setting,

$$\tilde{\psi}(p, t) = e^{-i\frac{p^2}{2m\hbar}t} \tilde{\psi}(p) \tag{8.19}$$

is a Fourier transform that evolves with time.

8.1 Time Evolution of Expectation

As per Theorem 4, for a fixed observable $\hat{K}$, the time evolution of the expectation value $\langle \psi | \hat{K} | \psi \rangle$ is given by the Hamiltonian. In particular, the Schrödinger equation gives:

$$|\dot{\psi}\rangle = -\frac{i}{\hbar} H |\psi\rangle, \quad \langle \dot{\psi}| = \frac{i}{\hbar} \langle \psi | H \tag{8.20}$$

which we use when computing the derivative of the expected value of $\hat{K}$ in state $|\psi\rangle$:

$$\begin{aligned} \frac{d}{dt} \langle \psi | \hat{K} | \psi \rangle &= \langle \dot{\psi} | \hat{K} | \psi \rangle + \langle \psi | \hat{K} | \dot{\psi} \rangle \\ &= \frac{i}{\hbar} \langle \psi | H \hat{K} | \psi \rangle - \frac{i}{\hbar} \langle \psi | \hat{K} H | \psi \rangle \\ &= \frac{i}{\hbar} \langle \psi | H \hat{K} - \hat{K} H | \psi \rangle \\ &= \frac{i}{\hbar} \langle [H, \hat{K}] \rangle. \end{aligned} \tag{8.21}$$

That is,

$$\frac{d}{dt} \langle \hat{K} \rangle = \frac{i}{\hbar} \langle [H, \hat{K}] \rangle. \tag{8.22}$$

Remark 26 *In classical mechanics, there is an analogous equation with Poisson brackets. Consider the observable function $K(x, p)$. Then,*

$$\frac{d}{dt} K(x, p) = \frac{\partial K}{\partial x} \dot{x} + \frac{\partial K}{\partial p} \dot{p}. \tag{8.23}$$

We substitute Hamilton's equations

$$\dot{x} = \frac{\partial H}{\partial p}, \ \dot{p} = -\frac{\partial H}{\partial x} \tag{8.24}$$

to find:

$$\frac{d}{dt}K(x,p) = \frac{\partial K}{\partial x}\frac{\partial H}{\partial p} - \frac{\partial K}{\partial p}\frac{\partial H}{\partial x} = \{K, H\}. \tag{8.25}$$

Notice, we do not take expectations because classical mechanics is deterministic, unlike quantum mechanics.

8.2 Poisson Brackets and Commutators

In classical mechanics, we have the following Poisson bracket identity:

$$\{F(x), p\} = \frac{\partial F}{\partial x}\frac{\partial p}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial p}{\partial x} = \frac{\partial F}{\partial x}. \tag{8.26}$$

To make a connection with commutators in quantum mechanics, consider the commutator of operator $F(x)$ (producing a ket when applied to a ket) with P:

$$\left[F(x), -i\hbar\frac{\partial}{\partial x}\right] = F(x)\left(-i\hbar\frac{\partial}{\partial x}\right) - \left(i\hbar\frac{\partial}{\partial x}\right)F(x). \tag{8.27}$$

We apply this operator to a state vector $|\psi(x)\rangle$ using its wave representation:

$$\begin{aligned}\left[F(x), -i\hbar\frac{\partial}{\partial x}\right]\psi(x) &= \left[F(x)\left(-i\hbar\frac{\partial}{\partial x}\right) - \left(i\hbar\frac{\partial}{\partial x}\right)F(x)\right]\psi(x) \\ &= -i\hbar F(x)\frac{\partial\psi}{\partial x} + i\hbar\left[F(x)\frac{\partial\psi}{\partial x} + \frac{\partial F}{\partial x}\psi(x)\right] \\ &= i\hbar\frac{\partial F}{\partial x}\psi(x),\end{aligned} \tag{8.28}$$

where we invoked the product rule in line two. It follows that

$$[F(x), P] = i\hbar\frac{\partial F}{\partial x}. \tag{8.29}$$

When $F(x) = x$, we recover the Heisenberg uncertainty principle:

$$[X, P] = i\hbar\frac{\partial x}{\partial x} = i\hbar. \tag{8.30}$$

Through similar derivation, one finds:

$$[X, F(p)] = i\hbar\frac{\partial F}{\partial p}. \tag{8.31}$$

Again, it recovers the Heinsenberg uncertainty principle when we set $F(p) = p$:

$$[X, P] = i\hbar\frac{\partial p}{\partial p} = i\hbar. \tag{8.32}$$

Remark 27 *In classical mechanics,* $\{F(x), p\} = \frac{\partial F}{\partial x}$. *Setting* $F(x) = x$, *we find* $\{X, P\} = 1$.

Remark 28 *Poisson brackets and commutators seem to be the same, except for a* $i\hbar$ *factor:*

$$i\hbar\{X, P\} = [X, P]. \tag{8.33}$$

8.3 Properties of Poisson Brackets and Commutators

Poisson brackets are antisymmetric:

$$\{A, B\} = -\{B, A\}. \tag{8.34}$$

Similarly, commutators are antisymmetric:

$$[A, B] = AB - BA = -(BA - AB) = -[B, A]. \tag{8.35}$$

Consider the following Poisson bracket:

$$\{AB, C\} = \frac{\partial AB}{\partial x}\frac{\partial C}{\partial p} - \frac{\partial AB}{\partial p}\frac{\partial C}{\partial x} \tag{8.36}$$

where

$$\begin{aligned} \frac{\partial AB}{\partial x} &= A\partial_x B + B\partial_x A \\ \frac{\partial AB}{\partial p} &= A\partial_p B + B\partial_p A. \end{aligned} \tag{8.37}$$

Thus, we simplify the bracket as:

$$\begin{aligned} \{AB, C\} &= (A\partial_x B + B\partial_x A)\partial_p C - (A\partial_p B + B\partial_p A)\partial_x C \\ &= A(\partial_x B\partial_p C - \partial_p B\partial_x C) + (\partial_x A\partial_p C - \partial_p A\partial_x C)B \\ &= A\{B, C\} + \{A, C\}B \end{aligned} \tag{8.38}$$

so

$$\{AB, C\} = A\{B, C\} + \{A, C\}B. \tag{8.39}$$

We check that the same property holds for commutators:

$$\begin{aligned} [AB, C] &\stackrel{?}{=} A[B, C] + [A, C]B \\ ABC - CAB &\stackrel{?}{=} A(BC - CB) + (AC - CA)B \\ &= ABC - ACB + ACB - CAB \\ &= ABC - CAB. \end{aligned} \tag{8.40}$$

Hence, we have confirmed commutators also obey the identity:

$$[AB, C] = A[B, C] + [A, C]B. \tag{8.41}$$

That is, commutators are the natural analog of Poisson brackets in quantum mechanics.

8.4 Schrödinger Equation for Particle in Potential

The Schrödinger wave equation for a particle is:

$$i\frac{\partial\psi(x,t)}{\partial t} = H\psi(x,t). \tag{8.42}$$

If the particle is subject to an external potential, then the Hamiltonian is

$$H = \frac{p^2}{2m} + U(x) \tag{8.43}$$

so the Schrödinger wave equation becomes:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + U(x)\psi(x). \tag{8.44}$$

We want to study how the center of this wave function moves. Indeed, it turns out that if the potential is smooth–that is, it varies smoothly on the scale of oscillations of the wave function, then the center moves according to classical phyiscs. To be more precise, we characterize the time-evolution of the average position observable:

$$\frac{d}{dt}\langle X\rangle = \frac{i}{\hbar}\langle[H,X]\rangle = \frac{i}{\hbar}\left\langle\left[\frac{P^2}{2m} + U(X), X\right]\right\rangle. \tag{8.45}$$

Since $[U(X), X] = 0$ (to see an example of this, take $U(X) = X^2$), we find that:

$$\frac{d}{dt}\langle X\rangle = \frac{i}{\hbar}\left\langle\left[\frac{P^2}{2m}, X\right]\right\rangle = \frac{i}{2m\hbar}\langle[P^2, X]\rangle. \tag{8.46}$$

We apply the commutator identity $[AB, C] = A[B, C] - [A, C]B$ and and Heisenberg uncertainty principle $[P, X] = -i\hbar$ to show:

$$[P^2, X] = P[P, X] + [P, X]P = -2i\hbar P. \tag{8.47}$$

Consequently,

$$\begin{aligned}\frac{d}{dt}\langle X\rangle &= \frac{i}{2m\hbar}\langle -2i\hbar P\rangle \\ &= \frac{\langle P\rangle}{m},\end{aligned} \tag{8.48}$$

or, equivalently:

$$\frac{d}{dt}\langle X\rangle = \left\langle\frac{P}{m}\right\rangle, \tag{8.49}$$

which is the average velocity.

Remark 29 *This is not surprising since $p/m = v$ classically.*

We are also interested in the time evolution of the average momentum observable $\frac{d}{dt}\langle P\rangle$. We use the fact that $[P^2, P] = 0$ to simplify it as follows:

$$\begin{aligned}\frac{d}{dt}\langle P\rangle &= \frac{i}{\hbar}\langle [H, P]\rangle \\ &= \frac{i}{\hbar}\left\langle \left[\frac{P^2}{2m} + U(X), P\right]\right\rangle \\ &= \frac{i}{\hbar}\langle [U(X), P]\rangle .\end{aligned} \tag{8.50}$$

Next, invoking the identity $[F(X), P] = i\hbar\frac{\partial F}{\partial x}$, we find:

$$[U(X), P] = i\hbar\frac{\partial U}{\partial x}. \tag{8.51}$$

Substituting this quantity into Equation 8.50, the derivative of average momentum is:

$$\frac{d}{dt}\langle P\rangle = \frac{i}{\hbar}\left\langle i\hbar\frac{\partial U}{\partial X}\right\rangle = -\left\langle \frac{\partial U}{\partial X}\right\rangle . \tag{8.52}$$

Remark 30 *This is the analog of Newton's $F = ma$ since $F = -\frac{\partial U}{\partial x}$ and if $p = mv$ then $\frac{d}{dt}p = ma = F$ classically. In quantum mechanics, one can only obtain these rules in expectation, as the laws are fundamentally probabilistic.*

Chapter 9

The Quantum Harmonic Oscillator

9.1 The Classical Harmonic Oscillator

Consider a spring harmonic oscillator, subject to gravitation, in an equilibrium state as shown in Figure 9.1. If you pull and release the spring, it will exhibit simple harmonic oscillation.

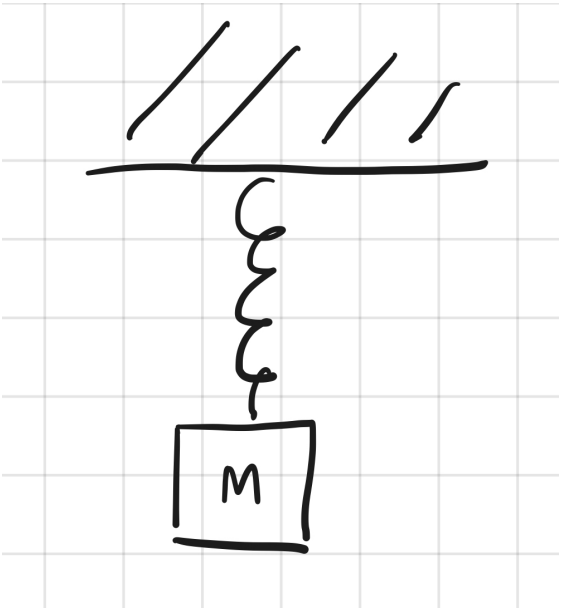


Figure 9.1: Simple harmonic oscillator.

Let x denote the displacement from its equilibrium position. The kinetic energy is

$$\frac{1}{2}m\dot{x}^2. \tag{9.1}$$

Let $y = \sqrt{m}x$ be a transformation of the coordinate system such that:

$$\dot{y}^2 = m\dot{x}^2. \tag{9.2}$$

Then the kinetic energy can be simply written as $\frac{1}{2}\dot{x}^2$ in the transformed coordinates. Let $k = \omega^2$ be the spring constant. In this system, potential energy is

proportional to displacement squared. Hence, the Lagrangian, which is the difference of kinetic and potential energy, can be written as:

$$\mathcal{L} = \frac{1}{2}\dot{x}^2 - \frac{1}{2}kx^2 = \frac{1}{2}\dot{x}^2 - \frac{\omega}{2}x^2. \tag{9.3}$$

The canonical momentum conjugate to x is:

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = \dot{x}. \tag{9.4}$$

The Euler-Lagrange equations of motion are

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0. \tag{9.5}$$

Plugging in the canonical momentum conjugate to x:

$$\begin{aligned} \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}} &= \ddot{x} \\ &= \frac{\partial \mathcal{L}}{\partial x} \\ &= -\omega^2 x. \end{aligned} \tag{9.6}$$

or, succinctly:

$$\ddot{x} = -\omega^2 x \tag{9.7}$$

This is precisely *Hooke's law*, which states that $F = -kx$ for spring constant $k = \omega^2$. That is, F is a restoring force that pulls upwards if you displace the spring downwards. The classical solution to this ordinary differential equation is:

$$x = a\cos(\omega t) + b\sin(\omega t). \tag{9.8}$$

9.2 Quantum Harmonic Oscillator

Let $\psi(x)$ be the wave function representation for a particle moving on a line. Then $\psi^*(x)\psi(x) = p(x)$ is the probability density for finding the particle at position x. As discussed in the previous section, the canonical momentum conjugate to x for the harmonic oscillator is

$$\hat{P} = \frac{\partial \mathcal{L}}{\partial \dot{X}} = \dot{X}. \tag{9.9}$$

Recall, the Lagrangian and Hamiltonian can be expressed in terms of kinetic and potential energy as:

$$\begin{aligned} \mathcal{L} &= \text{KE} - \text{PE} \\ H &= \text{KE} + \text{PE}. \end{aligned} \tag{9.10}$$

Hence, the sum of kinetic energy $\frac{1}{2}\dot{x}^2$ and potential energy $\frac{1}{2}\omega^2 x^2$ yields the Hamiltonian operator:

$$\begin{aligned} H &= \hat{P}\dot{\hat{X}} - \mathcal{L} \\ &= \frac{1}{2}\dot{\hat{X}}^2 + \frac{1}{2}\omega^2\dot{\hat{X}}^2 \\ &= \frac{1}{2}\hat{P}^2 + \frac{1}{2}\omega^2\hat{X}^2. \end{aligned} \tag{9.11}$$

We can treat $\hat{P}$ and $\hat{X}$ as observables acting on the state vector $\psi(x)$. That is, we identify the actions of $\hat{X}$, $\hat{P}$ and H as follows:

$$\begin{aligned} \hat{X}\,|\psi(x)\rangle &\overset{\text{corresponds to}}{\longrightarrow} x\psi(x) \\ \hat{P}\,|\psi(x)\rangle &\overset{\text{corresponds to}}{\longrightarrow} -i\frac{\partial}{\partial x}\psi(x) \\ H\,|\psi\rangle &\overset{\text{corresponds to}}{\longrightarrow} \frac{1}{2}\left(-i\frac{\partial}{\partial x}\right)\left(-i\frac{\partial\psi}{\partial x}\right)+\frac{1}{2}\omega^2x^2\psi(x) \end{aligned} \tag{9.12}$$

where $\frac{1}{2}\left(-i\frac{\partial}{\partial x}\right)\left(-i\frac{\partial\psi}{\partial x}\right)+\frac{1}{2}\omega^2x^2\psi(x) = -\frac{1}{2}\frac{\partial^2\psi}{\partial x}+\frac{1}{2}\omega^2x^2\psi(x)$. That is, we have the correspondence:

$$H\,|\psi\rangle \longrightarrow -\frac{1}{2}\frac{\partial^2\psi}{\partial x^2}+\frac{1}{2}\omega^2x^2\psi(x) \tag{9.13}$$

Hence, the Schrödinger equation for a quantum harmonic oscillator is:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi = -\frac{1}{2}\frac{\partial^2\psi}{\partial x^2}+\frac{1}{2}\omega^2x^2\psi(x). \tag{9.14}$$

9.2.1 Calculating Energy Eigenvector and Eigenvalues

Consider the eigenvectors $|\psi_E\rangle$ and eigenvalues E of the Hamiltonian, satisfying:

$$H\,|\psi_E\rangle = E\,|\psi_E\rangle\,. \tag{9.15}$$

Substituting this into the Schrödinger equation, we find:

$$-\frac{1}{2}\frac{\partial^2\psi_E(x)}{\partial x^2}+\frac{1}{2}\omega^2x^2\psi_E(x) = E\psi_E(x). \tag{9.16}$$

This is a second-order differential equation so it will have two independent solutions for $\psi_E(x)$. In this case, each solution is an exponential $e^{\xi x^2}$, $e^{-\xi x^2}$.

Remark 31 *The solution $\psi_E(x) = e^{\xi x^2}$ will exponentially increase as $x \to \infty$, which implies the probability density exponentially increases. Hence, the relative probability of being anywhere versus near ∞ is 0. Since the Hamiltonian (energy) contains X^2, if the probability density is concentrated overwhelmingly far away, it means the energy is enormous. Hence, the probability of finding the particle anywhere other than ∞ is zero. This solution is* ***degenerate****.*

Indeed, we must enforce the true space of states belongs to the state space of **square integrable functions**. That is, we require:

$$\int \psi^*(x)\psi(x)dx = 1 \tag{9.17}$$

to preclude functions that exponentially blow up far away.

Notice, the non-commutativity of X and P due to the Heisenberg uncertainty principle implies

$$H = \frac{1}{2}P^2+\frac{1}{2}\omega^2X^2 \tag{9.18}$$

is positive. Classically, to minimize energy, we set $P = 0$ and $X = 0$. However, this is not possible in quantum mechanics because one has to find a compromise between P and X. Thus, due to the uncertainty principle, **the lowest ground state energy of a system will not be zero**.

9.2.2 Ground State Solution

Consider the following ansatz for an oscillator concentrated near the origin (setting $\hbar$=1):

$$\psi(x) = e^{-\frac{\omega x^2}{2}} \tag{9.19}$$

as a solution to

$$-\frac{1}{2}\frac{\partial^2 \psi_E(x)}{\partial x^2} + \frac{1}{2}\omega^2 x^2 \psi_E(x) = E\psi_E(x). \tag{9.20}$$

We compute the first-order space derivative:

$$\frac{\partial \psi}{\partial x} = (-\omega x)\psi(x) \tag{9.21}$$

and the second-order space derivative:

$$\begin{aligned} \frac{\partial^2 \psi}{\partial x^2} &= -\omega \left[\psi(x) + x\frac{\partial \psi}{\partial x} \right] \\ &= -\omega\psi(x) - \omega x(-\omega x\psi(x)) \\ &= -\omega\psi(x) + \omega^2 x^2 \psi(x). \end{aligned} \tag{9.22}$$

Substituting these derivatives into Equation 9.20, we find:

$$\begin{aligned} \frac{\omega}{2}\psi(x) - \frac{\omega^2}{2}x^2\psi(x) + \frac{1}{2}\omega^2 x^2 \psi(x) &= E\psi_E(x) \\ \frac{\omega}{2} &= E. \end{aligned} \tag{9.23}$$

Hence, the eigenvector is $\psi(x) = e^{-\frac{\omega}{2}x^2}$ with eigenvalue $E = \omega\hbar/2$ (when we re-introduce $\hbar$). Since $\hbar$ is very small, the energy ground state is very small. The conventional notation to denote a ground state solution is:

$$\psi_0(x) = e^{-\frac{\omega}{2}x^2}. \tag{9.24}$$

9.2.3 Other Energy Levels

We are interested in the solutions for other energy levels. Recall the quantum mechanical (and classical) Hamiltonian is:

$$H = \frac{1}{2}(P^2 + \omega^2 X^2). \tag{9.25}$$

We could try to use the property that $a^2 + b^2 = (a + ib)(a - ib)$ to re-write the Hamiltonian. However, it turns out that this does not work due to the non-commutativity of P and X. That is, consider the following:

$$\begin{aligned}\frac{1}{2}(P + i\omega X)(P - i\omega X) &= \frac{1}{2}(P^2 + \omega^2 X^2 + i\omega XP - i\omega PX) \\ &= \frac{1}{2}(P^2 + \omega^2 X^2 + i\omega(XP - PX)) \\ &= \frac{1}{2}(P^2 + \omega X^2 - \omega) \\ &= \frac{1}{2}(P^2 + \omega^2 X^2) - \frac{\omega}{2}\end{aligned} \tag{9.26}$$

where we used the fact that $[X, P] = i\hbar = i$ in the $\hbar = 1$ convention. Therefore, we add a correction to recover the Hamiltonian:

$$\begin{aligned}H &= \frac{1}{2}(P^2 + \omega^2 X^2) \\ &= \frac{1}{2}(P + i\omega X)(P - i\omega X) + \frac{\omega}{2}\end{aligned} \tag{9.27}$$

where we effectively add $\omega/2$ to every eigenvalue of H, corresponding to the zero point (ground state) energy of the system.

Recall the ground state is $\psi_0(x) = e^{-\frac{\omega}{2}x^2}$ and the momentum operator is $P = -i\frac{\partial}{\partial x}$. hence, we apply $H = \frac{1}{2}(P + i\omega X)(P - i\omega X) + \frac{\omega}{2}$ to ψ_0 by first computing:

$$\begin{aligned}(P - i\omega X)\psi_0(x) &= \left(-i\frac{\partial}{\partial x} - i\omega X\right) e^{-\frac{\omega}{2}x^2} \\ &= -i(-\omega x - \omega x)e^{-\frac{\omega}{2}x^2} \\ &= 0.\end{aligned} \tag{9.28}$$

It follows that

$$\frac{1}{2}(P + i\omega X)(P - i\omega X)\psi_0 = 0 \tag{9.29}$$

so that

$$H\psi_0 = \frac{\omega}{2}\psi_0, \tag{9.30}$$

as desired.

9.3 Creation and Annihilation Operators

The Hamiltonian for the quantum harmonic oscillator is

$$H = \frac{1}{2}(P + i\omega X)(P - i\omega X) + \frac{\omega}{2}. \tag{9.31}$$

Definition 8 *The **creation operator** is $b^+ = P + i\omega X$ and the **annihilation operator** is $b^- = P - i\omega X$. They are Hermitian conjugates of each other; however, they are not observables because they are not individually Hermitian.*

We can compute the commutator of the creation and annihilation operators:

$$\begin{aligned}[b^+, b^-] &= [P + i\omega X, P - i\omega X] \\ &= [P, P] + [i\omega X, P] + [P, -i\omega X] + [i\omega X, -i\omega X] \\ &= [P, P] + i\omega[X, P] - i\omega[P, X] + \omega^2[X, X] \\ &= -2\omega\end{aligned} \tag{9.32}$$

where we used the fact that $[P, P] = [X, X] = 0$, $[X, P] = i$, and $[P, X] = -i$. Consequently,

$$[b^+, b^-] = -2\omega. \tag{9.33}$$

By convention, we define

$$a^+ := \frac{b^+}{\sqrt{2\omega}}, \; a^- := \frac{b^-}{\sqrt{2\omega}} \tag{9.34}$$

such that $[a^+, a^-] = -1$. We can now write the Hamiltonian in terms of the canonical creation and annihilation operators:

$$\begin{aligned}H &= \frac{1}{2}(P + i\omega X)(P - i\omega X) + \frac{\omega}{2} \\ &= \frac{1}{2}b^+b^- + \frac{\omega}{2} \\ &= \frac{1}{2}(2\omega)a^+a^- + \frac{\omega}{2} \\ &= \omega a^+a^- + \frac{\omega}{2}.\end{aligned} \tag{9.35}$$

The Hamiltonian in terms of the a^+ and a^- is:

$$H = \omega\left(a^+a^- + \frac{1}{2}\right). \tag{9.36}$$

In essence, we have reduced the process of finding energy levels to finding eigenvalues of a^+a^- given that $[a^+, a^-] = -1$.

Let $|0\rangle$ be the eigenvector of the ground state. Recall that $(P - i\omega X)\psi_0(x) = b^-\psi_0(x) = 0$, so we know $a^- |0\rangle = 0$. Now consider the creation operator applied to the ground state $a^+ |0\rangle$. We would like to know whether this is an eigenvector of the energy and, if so, its corresponding eigenvalue λ. That is, we are interested in the eigenvector-eigenvalue equation:

$$\begin{aligned}(a^+a^-)a^+ |0\rangle &\stackrel{?}{=} \lambda a^+ |0\rangle \\ a^+a^-a^+ |0\rangle &\stackrel{?}{=} \lambda a^+ |0\rangle\end{aligned} \tag{9.37}$$

We can re-write the commutator identity

$$\begin{aligned}[a^+, a^-] = a^+a^- - a^-a^+ &= -1 \\ a^+a^- + 1 &= a^-a^+.\end{aligned} \tag{9.38}$$

Substituting this into the left-hand side of 9.37, we get

$$a^+(a^+a^- + 1)\,|0\rangle\,. \tag{9.39}$$

Since $a^-\,|0\rangle = 0$, the left-hand side simplifies to

$$a^+\,|0\rangle\,. \tag{9.40}$$

Equating the left-hand side and right-hand side of Equation 9.37, we find:

$$a^+\,|0\rangle = \lambda a^+\,|0\rangle\,. \tag{9.41}$$

Therefore, $a^+\,|0\rangle$ is an eigenvector of a^+a^- with eigenvalue 1:

$$(a^+a^-)a^+\,|0\rangle = 1 \cdot a^+\,|0\rangle\,. \tag{9.42}$$

It follows that $a^+\,|0\rangle$ is an eigenvector of the Hamiltonian

$$H = \omega(a^+a^- + \frac{1}{2}) \tag{9.43}$$

with eigenvalue $\omega + \frac{\omega}{2} = \frac{3}{2}\omega$. Thus, **we added one unit of energy to the ground state by acting with the operator** a^+.

Let's continue to add energy with the creation operator. Suppose

$$a^+a^-\,|n\rangle = n\,|n\rangle\,. \tag{9.44}$$

We have already found eigenvectors for $n = 0$ and $n = 1$. Let

$$|n+1\rangle := a^+\,|n\rangle \tag{9.45}$$

then

$$a^+a^-\,|n+1\rangle = a^+a^-a^+\,|n\rangle\,. \tag{9.46}$$

Substituting the identity $a^-a^+ = a^+a^- + 1$,

$$\begin{aligned} a^+a^-\,|n+1\rangle &= a^+(a^+a^- + 1)\,|n\rangle \\ &= a^+a^+a^-\,|n\rangle + a^+\,|n\rangle \\ &= a^+(a^+a^-\,|n\rangle) + a^+\,|n\rangle \\ &= a^+(|n+1\rangle) + |n+1\rangle \\ &= (n+1)\,|n+1\rangle\,. \end{aligned} \tag{9.47}$$

Theorem 5 *Let $|n\rangle$ be an eigenvector of a^+a^- with eigenvalue n such that:*

$$a^+a^-\,|n\rangle = n\,|n\rangle\,, \tag{9.48}$$

then $a^+\,|n\rangle$ is also an eigenvector of a^+a^- with eigenvalue $n+1$:

$$a^+a^-(a^+\,|n\rangle) = (n+1)a^+\,|n\rangle \tag{9.49}$$

or, in the notation $|n+1\rangle := a^+\,|n\rangle$, we find:

$$a^+a^-\,|n+1\rangle = (n+1)\,|n+1\rangle\,. \tag{9.50}$$

The takeaway is that a^+ acts on any eigenvector of a^+a^-, say $|n\rangle$, to raise the eigenvalue by 1. Hence, there is a spectrum of eigenvalues of energy, namely:

$$\frac{\omega}{2}, \frac{3\omega}{2}, \frac{5\omega}{2}, \ldots, \frac{\omega(n+1/2)}{2}, \ldots. \tag{9.51}$$

The creation operator a^+ increases the energy in integer multiples of ω each time it acts such that:

$$\omega\left(n+\frac{1}{2}\right)-\frac{\omega}{2}=\omega n \tag{9.52}$$

where $\omega/2$ is the ground state. Hence, we have the following table of eigenvectors and eigenvalues:

Eigenvectors	**Eigenvalues**
$\|0\rangle$	$\omega\left(\frac{1}{2}\right)$
$a^+\|0\rangle$	$\omega\left(\frac{3}{2}\right)$
$(a^+)^2\|0\rangle$	$\omega\left(\frac{5}{2}\right)$
$\vdots$	$\vdots$
$(a^+)^n\|0\rangle$	$\omega\left(\frac{n+1}{2}\right)$

Table 9.1: Eigenvectors and eigenvalues of the system

This explains why *energy of oscillations is quantized*, in units of $\omega\hbar$ (re-introducing $\hbar$).

9.3.1 Shape of Wavefunctions for Harmonic Oscillator

Consider the ground state function for the quantum harmonic oscillator:

$$\psi_0(x)=e^{-\frac{\omega}{2}x^2}, \tag{9.53}$$

which is a symmetric Gaussian-like wave function. Consider the shape of the wave function at the next energy level. Apply the creation operator $b^+ = P + i\omega X = -i\frac{\partial}{\partial x} + i\omega x$:

$$\begin{aligned} b^+\psi_0 &= i\left(-\frac{\partial}{\partial x}+\omega x\right)e^{-\frac{\omega}{2}x^2} \\ &= i(\omega x+\omega x)e^{-\frac{\omega}{2}x^2} \\ &= 2i\omega x e^{-\frac{\omega}{2}x^2}. \end{aligned} \tag{9.54}$$

This is an anti-symmetric wave function, as shown in Figure 9.2. It has one zero or **node**, such that the probability the particle appears at that point is nil.

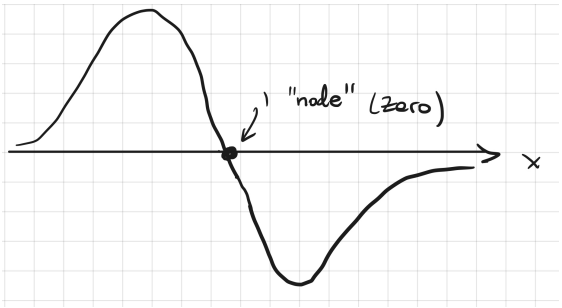


Figure 9.2: Antisymmetric wave function for first non-zero point energy.

If we apply the creation operator again, the wave function will be symmetric and have two nodes, as shown in Figure 9.3:

$$\begin{aligned} b^+\psi_1 &\propto \left(-i\frac{\partial}{\partial x} + \omega x\right) xe^{-\frac{\omega}{2}x^2} \\ &\propto (\text{quadratic polynomial})e^{-\frac{\omega}{2}x^2}. \end{aligned} \tag{9.55}$$

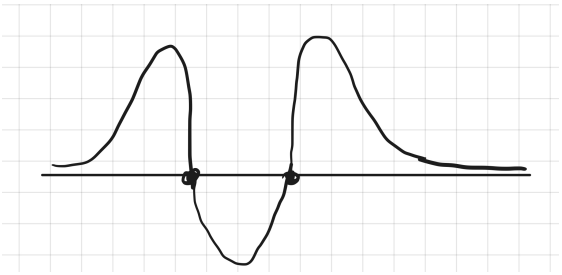


Figure 9.3: Wave function with two nodes, corresponding to the second energy level.

By induction, as we apply the creation operator b^+ to the wave function, the wave will oscillate more rapidly. That is, it will have high momentum and the wave function will spread out over greater distances, swinging back and forth. Likewise, the annihilation operator b_- lowers the energy states.

Bibliography

[1] L. Susskind, "Modern physics: Quantum mechanics." https://www.youtube.com/watch?v=JzhlfbWBuQ8, Jan. 2008. [Video]. YouTube.