Quantum Mechanics Formalism

Yu Lu

September 7, 2022

1 States and Observables

1.1 States and wavefunctions

The most general thing describing a quantum system is its **state** $|S(t)\rangle$, which is often represented in the position space by the **wavefunction** $\Psi(x,t)$ (the exact process for decomposition will be discussed later). The wavefunction $\Psi(x,t)$ is a complex-valued function living in the **Hilbert space** so that $\int_{-\infty}^{\infty} |\Psi|^2 \, \mathrm{d}x$ is finite. The **observables** (such as position and momentum) are operators that can be represented by (sometimes infinite dimensional) matrices. To make a measurement, we need to act the operator on the wavefunction (though the exact process is more subtle). In the finite-dimensional case, this is simply matrix multiplication, though it is more complicated in the infinite-dimensional case.

The wavefunctions are like vectors. For example, we define the **inner product** of two functions $f(x)$ and $g(x)$ as

$$\langle f|g \rangle \equiv \int_{-\infty}^{\infty} f(x)^* g(x) \, \mathrm{d}x,$$

which satisfies

$$\langle f|g \rangle = \langle g|f \rangle^*.$$

A function $f(x)$ can be decomposed into a set of basis functions $\{f_n(x)\}$ as

$$f(x) = \sum_{i=1}^{N} c_n f_n(x).$$

The basis is **complete** if any function in the Hilbert space could be represented as above and is **orthonormal** if any two member satisfy

$$\langle f_m|f_n \rangle = \delta_{mn}.$$

Sometimes the basis $f_p(x)$ are continuous and satisfy Dirac orthonormality

$$f(x) = \int_{-\infty}^{\infty} c(p) f_p(x) \, \mathrm{d}p, \quad \langle f_p|f_p' \rangle = \delta(p - p'),$$

where $\delta(x)$ is the Dirac delta function satisfying

$$\int_{-\infty}^{\infty} e^{ipx} \, \mathrm{d}x = \delta(p), \quad \int_{-\infty}^{\infty} \delta(x' - x) f(x') \, \mathrm{d}x = f(x), \quad \delta(ax) = \frac{\delta(x)}{|a|}.$$

Assuming orthonormality, the coefficients c_n could be given by Fourier's trick

$$c_n = \langle f_n|f \rangle.$$

1.2 Operators

Operators $\hat{Q}$ act on wavefunction $\Psi(x,t)$ and represent observables only if they are **Hermitian**, satisfying

$$\left\langle f \middle| \hat{Q}f \right\rangle = \left\langle \hat{Q}f \middle| f \right\rangle \quad \forall f(x).$$

This, in turn, defines the Hermitian conjugate $\hat{Q}^\dagger$ of $\hat{Q}$ such that $\left\langle f \middle| \hat{Q}g \right\rangle = \left\langle \hat{Q}^\dagger f \middle| g \right\rangle$.

The expectation value of $\hat{Q}$ could be computed as

$$\boxed{\left\langle \hat{Q} \right\rangle = \langle \Psi | \hat{Q} | \Psi \rangle .}$$

We would need the Copenhagen interpretation to make sense of this equation. Also, the standard deviation of $\hat{Q}$ is defined as

$$\boxed{\sigma^2 = \left\langle (\hat{Q} - \langle Q \rangle)^2 \right\rangle = \langle Q^2 \rangle - \langle Q \rangle^2 .}$$

We are mostly concerned with the eigenfunction expansion of the operator $\hat{Q}$, for reasons that will become clear in the Copenhagen interpretation. Luckily, such expansion is always possible for Hermitian operators, since they have only real eigenvalues and orthonormal eigenfunctions in the case of non-degenerate discrete spectrums. Even if the spectrum is degenerate, it is always possible to choose a set of orthonormal eigenfunctions. Moreover, the eigenfunction spans the entire Hilbert space. In the case of a continuous spectrum, it is possible to choose a set of complete eigenfunctions with Dirac orthonormality and real eigenvalues.

Furthermore, the eigenfunctions Ψ are ***deterministic states*** of the operator $\hat{Q}$, in the sense that repeated measurements yield the same result, with $\sigma = 0$ and $\left\langle \hat{Q} \right\rangle = q$, which is the eigenvalue.

Like matrices, operators do not generally commute. This gives rise to the uncertainty principles. However it could be demonstrated that two operators can be expanded in the same set of eigenfunctions when they do commute.

1.3 General statistical interpretation

By Hamiltonian mechanics, a system can be fully specified by its position and momentum. Therefore, any observable Q can be expressed as $Q(\hat{x}, \hat{p})$. Upon observation, the wavefunction $\Psi(x, t)$ collapses into one of its eigenfunction $f_n(x)$ with probability

$$|c_n|^2 , \quad \text{where } c_n = \langle f_n | \Psi \rangle ,$$

yielding the associated eigenvalue q_n as a result. The extension to a continuous spectrum is trivial. The probability is properly normalised in a orthonormal basis, since $\sum_{n=1}^{N} |c_n|^2 = \langle \Psi | \Psi \rangle = 1$.

1.4 Schrödinger's equation

Between measurements, the wavefunction $\Psi(x, t)$ evolves smoothly in time according to the Schrödinger's equation

$$\boxed{i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi,} \tag{1}$$

where $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$ is the Hamiltonian. By the separation of variables $\Psi(x, t) = \psi(x)\phi(t)$, the equation could be simplified into the time-independent Schrödinger's equation and the time evolution

$$\hat{H}\Psi = E\psi \quad \phi(t) = e^{-iEt/\hbar},$$

where $\Psi(x)$ is the eigenfunction of the Hamiltonian and E is the associated energy.

1.5 Uncertainty principles

1.5.1 General uncertainty principle

Consider two observables A and B that do not commute such that $[\hat{A}, \hat{B}] = AB - BA \neq 0$. Their variances $\sigma^2 = \left\langle (\hat{Q} - \langle Q \rangle)\Psi \big| (\hat{Q} - \langle Q \rangle)\Psi \right\rangle$ satisfy the Cauchy Schwarz inequality

$$\sigma_A^2 \sigma_B^2 = \left\langle (\hat{A} - \langle A \rangle)\Psi \big| (\hat{A} - \langle A \rangle)\Psi \right\rangle \left\langle (\hat{B} - \langle B \rangle)\Psi \big| (\hat{B} - \langle B \rangle)\Psi \right\rangle$$

$$\geq \left| \left\langle (\hat{A} - \langle A \rangle)\Psi \big| (\hat{B} - \langle B \rangle)\Psi \right\rangle \right|^2 (\hat{A}^\dagger = \hat{A})$$

$$= \left| \left\langle \Psi \big| (\hat{A} - \langle A \rangle)(\hat{B} - \langle B \rangle)\Psi \right\rangle \right|^2 \equiv |z|^2 .$$

Simplifying z gives

$$z = \left\langle \hat{A}\hat{B} \right\rangle - \left\langle \hat{B} \right\rangle \left\langle \hat{A} \right\rangle - \left\langle \hat{A} \right\rangle \left\langle \hat{B} \right\rangle + \langle A \rangle \langle B \rangle \langle \Psi | \Psi \rangle = \left\langle \hat{A}\hat{B} \right\rangle - \langle A \rangle \langle B \rangle.$$

Noting that $|z|^2 \geq \text{Im}(z)^2 = \left[\frac{1}{2i}(z - z^*) \right]^2$, we obtain the generalised uncertainty principle.

$$\sigma_A^2 \sigma_B^2 \geq \left(\frac{1}{2i} \left\langle [\hat{A}, \hat{B}] \right\rangle \right)^2 \quad \Leftrightarrow \quad \sigma_A \sigma_B \geq \frac{1}{2} \left| \left\langle \hat{A}, \hat{B} \right\rangle \right|. \tag{2}$$

1.5.2 Heisenburg principle

As will be shown later, the commutation relation between position x and momentum p is

$$[\hat{x}, \hat{p}] = i\hbar.$$

Applying this directly to Eq.(2) gives the famous Heisenberg uncertainty principle

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}. \tag{3}$$

By forcing equality in the deriving process (see the previous section), we could show that the equality is achieved for a Gaussian wave packet.

1.5.3 Prelude to the Heisenburg picture

In classical mechanics, we are concerned with the time derivative of observables. In quantum mechanics, we could relate them to commutators. Taking the total derivative on an operator $Q(x, p, t)$ that depends explicitly on time gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle Q \rangle = \frac{\mathrm{d}}{\mathrm{d}t} \left\langle \hat{Q} \right\rangle = \left\langle \frac{\partial \Psi}{\partial t} \middle| \hat{Q} \middle| \Psi \right\rangle + \langle \Psi | \frac{\partial \hat{Q}}{\partial t} | \Psi \rangle + \left\langle \Psi \middle| \hat{Q} \middle| \frac{\partial \Psi}{\partial t} \right\rangle$$
$$= \frac{1}{i\hbar} (\left\langle \hat{Q}\hat{H} - \hat{H}\hat{Q} \right\rangle) + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle,$$

where the second line uses the time-dependent Schrödinger's equation Eq.(1). This gives us the generalised Ehrenfest theorem, which is the prelude to Heisenberg picture (differs by an expected value)

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \hat{Q} \right\rangle = \frac{i}{\hbar} \left\langle [\hat{H}, \hat{Q}] \right\rangle + \left\langle \frac{\partial \hat{Q}}{\partial t} \right\rangle. \tag{4}$$

Usually, the observable does not depend explicitly on time. Thus, observables that commutes with $\hat{H}$ have a constant expected value.

1.5.4 Other uncertainty principles

Similarly, it could be shown that

$$\sigma_x \sigma_H \geq \frac{\hbar}{2m} |\langle p \rangle|,$$

where $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$ is the Hamiltonian.

We could also consider the uncertainty principle from a different perspective. Applying the general uncertainty principle to the Hamiltonian H and a general operator Q gives

$$\sigma_H \sigma_Q \geq \frac{1}{2} \left| [\hat{H}, \hat{Q}] \right| = \frac{\hbar}{2} \left| \frac{\mathrm{d} \langle Q \rangle}{\mathrm{d}t} \right|.$$

Taking

$$\Delta t \equiv \frac{\sigma_Q}{|\mathrm{d} \langle Q \rangle / \mathrm{d}t|}$$

gives the energy-time uncertainty principle

$$\sigma_H \Delta t \geq \frac{\hbar}{2}, \tag{5}$$

where Δt could be interpreted as the time it takes for *any* observable $\hat{Q}$ to change substantially. In other words, the fluctuation in energy σ_H places an upper bound on the rate of change of all observables. It should be warned that Δt should not be interpreted as the standard deviation in time, which is not an observable in non-relativistic quantum mechanics, nor should it be interpreted as "the duration for which an energy σ_H could be borrowed from the system," since this is simply a violation of the energy conservation.

2 Dirac Notation

2.1 The notation

Dirac broke the two components in a inner product $\langle\alpha|\beta\rangle$ into a **bra** $\langle\alpha|$ and a **ket** $|\beta\rangle$. While the ket is simply a vector in the Hilbert space, a bra is in fact a function defined by the inner product $(\langle f| = \int f * [\ldots]\mathrm{d}x)$. Just like the ket lives in a vector space, the bra lives in the dual vector space.

With this new notation, we could define operators in terms of outer products $|\alpha\rangle\langle\beta|$. For example, the **projection operator** onto $|\alpha\rangle$ is

$$\hat{P} \equiv |\alpha\rangle\langle\alpha|.$$

In this way, the completeness of orthonormal basis could be compactly expressed as

$$\boxed{\sum_{n=1} |e_n\rangle\langle e_n| = 1 \quad \text{or} \quad \int |e_z\rangle\langle e_z|\, \mathrm{d}z = 1.} \tag{6}$$

Insertion of this identity are sometimes the first line of attack in proving quantum mechanical relationships. The following discussion of change of basis in Dirac notation shows the real power of Dirac notation—it serves as a quick visual aid so we could play with vectors and operators like playing Lego.

2.2 Basis

We haven't explicitly pointed out the distinction between the wavefunction $\Psi(x,t)$ and the state $|S(t)\rangle$. The wavefunction $\Psi(x,t)$ is the representation of $|S(t)\rangle$ in the position space. In each eigenbasis $|x\rangle$ of the position operator $\hat{x}$, the component of $S(t)$ can be evaluated using the inner product $\Psi(x,t) = \langle x|S(t)\rangle$. In analogy to a vector in $\mathbb{R}^3$, $|S(t)\rangle$ is like the physical vector itself, $\Psi(x,t)$ at each x is like individual components of the vector (essentially a number) while the collective of $\Psi(x,t)$ at all possible x forms an infinite-dimensional coordinate representing the vector in the position space, like $(1,2,3,\ldots)$. It should be noted that the both $|S(t)\rangle$ and $\Psi(x,t)$ reside in a vector space, it is technically only $|S(t)\rangle$ that is in the Hilbert space (basis independent) while $\Psi(x,t)$ lives in the infinite-dimensional vector space spanned by $|x\rangle$ (basis dependent).

Similarly, we could also form the momentum wavefunction as $\Phi(p,t) = \langle p|S(t)\rangle$ or decompose it into a discrete basis $c_n(t) = \langle n|S(t)\rangle$. The representation of $S(t)$ in these basis could be achieved by inserting the identity as

$$|S(t)\rangle = \left(\int \mathrm{d}x\, |x\rangle\langle x|\right)|S(t)\rangle = \int |x\rangle\, \langle x|S(t)\rangle\, \mathrm{d}x \equiv \int \Psi(x,t)\, |x\rangle\, \mathrm{d}x,$$

$$|S(t)\rangle = \left(\sum_{n=1} |n\rangle\langle n|\right)|S(t)\rangle = \sum_{n=1} |n\rangle\, \langle n|S(t)\rangle \equiv \sum_{n=1} c_n(t)\, |n\rangle,$$

where neighbouring bras and kets can be naturally assembled into an inner product. It can be seen that bras and kets cannot be freely moved around, since that might turn an inner product into an outer product or commute two random operators. However, when a bra-ket (inner product) sticks together, they as a whole can move freely.

Operators are abstract constructs which act on a ket to yield another ket. In calculation, we normally express operators in a basis as well. Like states, operators are represented in different ways in different basis. For example, since the action of a position operator on a state can be expressed in the p basis as $\langle p|\hat{x}|S(t)\rangle$, the operator can be expressed as $\langle p|\,\hat{x}$ in the p basis. Detailed examples will be given in later sections.

2.3 Exemplification: position and momentum

2.3.1 Eigenfunction

The position and momentum operator could be *represented* in the position space as $\hat{x} \to x$ and $\hat{p} \to -i\hbar\frac{\partial}{\partial x}$. Their eigenfunction $|x\rangle$ and $|p\rangle$ could be found from the definition $Q\Psi = \lambda\Psi$.

For the position operator, the eigenfunction equation is

$$\hat{x}\,|S(t)\rangle = y\,|S(t)\rangle \implies \langle x|\hat{x}|S(t)\rangle = y\,\langle x|S(t)\rangle \implies x\Psi = y\Psi,$$

where we have used the definition $\Psi = \langle x|S(t)\rangle$ and the facts that $x = \langle x|\,\hat{x}$. Holding y constant, the position x can still take any value. Therefore, the eigenfunction $\Psi_y(x,t)$ can only be

$$\Psi_y(x,t) = \delta(x-y),$$

where y can take any real number to satisfy the normalisation condition. Any wavefunction could also be represented in this eigenfunction basis as

$$f(x) = \int_{-\infty}^{\infty} c(y)\delta(x-y)\,\mathrm{d}y, \quad c(y) = f(y).$$

Similarly, we can formulate the eigenfunction equation for the momentum operator as (p is the eigenvalue)

$$-i\hbar\frac{\partial\Psi}{\partial x} = p\Psi,$$

which is a first order differential equation with normalised solution

$$\Psi_p(x,t) = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}, \quad p \in \mathbb{R}.$$

This gives rise to the Fourier relation between x and p, since the change of basis is now

$$\boxed{\begin{aligned}
\Psi(x,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\Phi(p,t)e^{ipx/\hbar}\,\mathrm{d}p, \\
\Phi(p,t) &= \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\Psi(x,t)e^{-ipx/\hbar}\,\mathrm{d}x,
\end{aligned}}$$

where $\Phi(p,t)$ is the **momentum wavefunction** which is like $\Psi(x,t)$ in the momentum space.

2.3.2 Commutator

Apparently, $\hat{x}$ and $\hat{p}$ does not commute. Since the commutator should not depend on the basis, we calculate it as

$$[\hat{x},\hat{p}]\Psi = -i\hbar\left(x\frac{\partial\Psi}{\partial x} - \frac{\partial(x\Psi)}{\partial x}\right) = -i\hbar\left(x\frac{\partial\Psi}{\partial x} - x\frac{\partial\Psi}{\partial x} - \Psi\right),$$

giving the operator form

$$\boxed{[\hat{x},\hat{p}] = i\hbar.} \tag{7}$$

Note that the commutators are in general still operators.

Using the above relationship, together with the properties of commutators

$$\boxed{[\hat{A},\hat{B}] = -[\hat{B},\hat{A}], \quad [\hat{A}+\hat{B},\hat{C}] = [\hat{A},\hat{C}]+[\hat{B},\hat{C}], \quad [\hat{A}\hat{B},\hat{C}] = \hat{A}[\hat{B},\hat{C}]+[\hat{A},\hat{C}]\hat{B},}$$

we could arrive at two more general conclusions

$$[f(\hat{x}),\hat{p}] = i\hbar\frac{\partial f}{\partial x}, \quad [x,f(p)] = i\hbar\frac{\partial f}{\partial p},$$

which actually extends to all variables that are Fourier transform of each other.

2.3.3 Position and momentum basis

Since x and p are Fourier pairs, they should be symmetric. In other words, we could use $\Phi(x,t)$ to perform calculation in the momentum space as well, leaving the question of how to represent $\hat{x}$ and $\hat{p}$ in the momentum space. In this change of basis, we would need the inner product

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \delta(x-x')e^{ipx'/\hbar}\,\mathrm{d}x' = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar},$$

which is independent of the basis. Therefore, the position operator expressed in the momentum space is

$$\begin{aligned}
\langle p|\hat{x}|S(t)\rangle &= \int_{-\infty}^{\infty} \langle p|x\rangle \, \langle x|\hat{x}|S(t)\rangle \,\mathrm{d}x \qquad \text{(inserting the identity)} \\
&= \int_{-\infty}^{\infty} \frac{e^{-ipx/h}}{\sqrt{2\pi\hbar}} x\Psi(x,t)\,\mathrm{d}x \\
&= i\hbar\frac{\partial}{\partial p} \int_{-\infty}^{\infty} \frac{e^{-ipx/h}}{\sqrt{2\pi\hbar}}\Psi(x,t)\,\mathrm{d}x \qquad \text{(using the definition of $\Phi(p,t)$)} \\
&= i\hbar\frac{\partial \Phi}{\partial p}.
\end{aligned}$$

Therefore, we conclude that $\hat{x} \to i\hbar\frac{\partial}{\partial p}$ in the momentum space. The transform from $\Psi(x,t)$ to $\Phi(p,t)$ could also be achieved in a similar fashion. The results we have obtained regarding $\hat{x}$ and $\hat{p}$ are summarised in Table 1.

Operator	Representation in $|x\rangle$	Representation in $|p\rangle$	Eigenfunction in x basis
$\hat{x}$	x	$i\hbar\frac{\partial}{\partial p}$	$|x\rangle = \delta(x-x')$
$\hat{p}$	$-i\hbar\frac{\partial}{\partial x}$	p	$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx'/\hbar}$

Table 1: The position and momentum operators in different basis.