

Week 1

This week we will cover chapter 1 of Sakurai specifically sections 1.2, 1.3, 1.4, 1.5, 1.6 1.7

Representing physics with math

The connection between math and physics has always been close and we are used to the idea of using math to describe physical reality. We use the natural numbers to describe the number of physical things, like the number of atoms. Fractions also make sense in certain situations, we can have $\frac{1}{2}$ an apple, but it is less clear what half an atom is! Mathematical concepts may have physical limits.

One the ways physics has developed has been to try to find the most appropriate way to represent physical quantities mathematically. A good example is velocity. In Newtonian mechanics we can represent our velocity as a 3 component vector and if we want to know our velocity relative to other objects which may be moving, we simply add or subtract these velocity vectors to get a new velocity vector.

However, if we move to relativistic mechanics, we know that the 3-vector representation of velocity is no longer valid. Instead, we must move to 4-vectors and although we can add two 4-vectors together, math tells us that the sum of two 4-vectors is not a 4-vector. Thus, by choosing the right mathematical representation we can gain a better understanding on our physical theories work.

Our aim is to find out the correct mathematical description of quantum states.

Building a mathematical representation of quantum mechanics – Quantum States

The strangeness of quantum mechanics in part comes about because the wavefunction, which describes the system cannot be directly measured and this is because when we do make a measurement we change the wavefunction.

Dirac used the symbol

$$|\psi\rangle$$

To represent a general state of the system. This state, to be defined, must encode all possible measurement results.

When we make a measurement we then go from our initial state to a new state with some probability $|c_i|^2$

$$|\psi\rangle \rightarrow |\psi_i\rangle$$

The measurement results are then said to form a **basis**. If we have a complete set of probability amplitudes, i.e. we can specify c_i for each possible measurement outcome, we have a **complete set of probability amplitudes**. Then we can write

$$|\psi\rangle = \sum_i c_i |\psi_i\rangle$$

The probability of finding the system in state $|\psi_i\rangle$ is given by $|c_i|^2$.

One of the postulates of quantum mechanics is that c_i is complex. This is another unusual aspect of quantum mechanics. While we are used to using imaginary numbers in physics, it is normally a trick to save us time. For example, it is often easier to solve equations in electromagnetism by saying the electric field is the real part of

$$E(t) = A e^{i\omega t + \phi}$$

Than trying to solve equations with $E(t) = A \cos \omega t + B \sin \omega t$

However, as we do not directly measure c_i , it can be imaginary¹.

The choice of notation by Dirac was deliberate to point out the connection with vectors

$$\mathbf{p} = a \hat{\mathbf{x}} + b \hat{\mathbf{y}} + c \hat{\mathbf{z}}$$

Where the point p lies in a vector space defined by the vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$

We can now say that the quantum mechanical **state vector** “lives” is a generalized vector space known as a **complex Hilbert space**.

We are often very interested in the coefficients c_i which determine the probability of measurements, so we need some mathematical method to **map** our state vector onto the set of complex numbers, i.e. we need some set of functionals

$$f_i(|\psi\rangle) \rightarrow c_i \forall i$$

We can achieve this mapping if

$$f_i(|\psi_j\rangle) \rightarrow \delta_{ij}$$

The complete set of functionals f_i also forms a basis its own vector space known as the **Dual Space**.

In Dirac notation we denote a state in the dual space as a bra

$$\langle \psi_i |$$

And an arbitrary state in the dual space can be written in as

$$\langle \rho | = \sum_i b_i^* \langle \psi_i |$$

For every ket $|\psi_i\rangle$ there is a corresponding bra $\langle \psi_i |$. This is known as **Dual correspondence**

The dual vector for an arbitrary state

$$|\psi\rangle = \sum_i c_i |\psi_i\rangle \rightarrow \langle \psi | = \sum_i c_i^* \langle \psi_i |$$

Finally, we define the **inner product** between two states as

¹ Note that the use of complex numbers in quantum mechanics is still debated and it is possible to make a quantum mechanical theory based on real numbers if the state vector becomes a matrix. See [Quantum theory based on real numbers can be experimentally falsified | Nature](#)

$$\langle \rho | \psi \rangle = \sum_i \sum_j b_i^* c_j \langle i | j \rangle = \sum_j b_j^* c_j$$

And so

$$\langle \rho | \psi \rangle = \langle \psi | \rho \rangle^*$$

The inner product is a generalization of the dot product between two vectors.

The “length” of a quantum state

Just like a vector, we can determine the “length” of a quantum state by performing the inner product with itself

$$l^2 = \langle \psi | \psi \rangle = \sum_i |c_i|^2$$

In a generalized vector space, $l = \sqrt{\langle \psi | \psi \rangle} \geq 0$ is the norm. This is a useful property as we want to work with normalized states so that the probability of finding the system in $|\psi\rangle$ in state $|\psi\rangle$ should be 1. Therefore we can always normalize our states.

Summary:

Quantum states are complex vectors in a Hilbert space. They are made up of basis vectors and are characterized by a complete set of probability amplitudes, which quantify the probability of finding the state with a specific value. These are the ket states. For any quantum state we can define a dual state known as a bra and the product of a bra and a ket state returns the probability amplitude of finding the ket state in the state of the bra.

Operators

In quantum mechanics we know that when we take a measurement, we disturb the system, so we need a formalism to understand this process.

Operators are a way of mapping one quantum state onto another. i.e. the operator X takes the state $|\psi\rangle$ to the state $|\psi_x\rangle$ ²

$$|\psi_x\rangle = \hat{X} |\psi\rangle$$

As we are working in a vector space state kets can be considered vectors, therefore operators are like matrices. This means that operators can be added together multiplied etc. but like matrices the order is important.

In general operators are linear so

$$\hat{X} c |\psi\rangle = c \hat{X} |\psi\rangle = c |\psi_x\rangle$$

² We will use capital letters and/or \hat{x} to represent operators and lower case letters to represent complex scalars

With the exception of one operator, the time-reversal operator, which will discuss later.

An operator can either act to a ket to the right or a bra to the left

$$\langle \psi | \hat{X} | \psi \rangle = \langle \psi | (\hat{X} | \psi \rangle) = (\langle \psi | \hat{X}) | \psi \rangle$$

But note if $|\psi_x\rangle = \hat{X} |\psi\rangle$, $\langle \psi | X \neq \langle \psi_x |$ in general.

This can be seen from

$$\langle \psi | (\hat{X} | \psi \rangle) = (\langle \psi | \hat{X}) | \psi \rangle$$

$$\langle \psi | \phi' \rangle = \langle \phi | \psi \rangle$$

$$\langle \phi' | \psi \rangle^* = \langle \phi | \psi \rangle$$

So $\langle \phi' | \neq \langle \phi |$

Instead, due to dual correspondence

$$\langle \psi_x | = \langle \psi | X^\dagger$$

Where X^\dagger is the Hermitian conjugate.

Hermitian operators

Operators that satisfy $X^\dagger = X$ are said to be Hermitian and they represent an important class of operators.

Let the ket $|a\rangle$ and $|a'\rangle$ be eigenkets of the Hermitian operator A

$$A|a\rangle = a|a\rangle$$

$$A|a'\rangle = a'|a\rangle$$

We can create the corresponding eigenbra equation:

$$\langle a' | A = a'^* \langle a |$$

$$\langle a' | A | a \rangle = a'^* \langle a' | a \rangle$$

$$a \langle a' | a \rangle = a'^* \langle a' | a \rangle$$

$$(a'^* - a) \langle a' | a \rangle = 0$$

If $|a'\rangle = |a\rangle$

Then $\langle a | a \rangle = 1$ and $(a^* - a) = 0$

So the eigenvalues must be real – This is good for physical quantities because we can only measure real numbers.

If $|a'\rangle \neq |a\rangle$, assuming no degeneracy, so $(a^* - a) \neq 0$ this means that $\langle a' | a \rangle = 0$ and the eigensates are orthogonal. Even if there is degeneracy, it can be shown that you can always construct an orthogonal basis.

Finally, the basis is complete – This means that any point in my Hilbert space can be represented by a superposition of my basis states.

Outer products

The inner product between two states is a scalar so we can write

$$(\langle \alpha | \gamma \rangle) |\beta \rangle = |\beta \rangle \langle \alpha | \gamma \rangle$$

We can then identify the quantity $|\beta\rangle\langle\alpha|$ as an operator acting on the state ket $|\gamma\rangle$. The operator in this case takes the ket $|\gamma\rangle$ and gives back the ket $|\beta\rangle$ with some probability. When we construct an operator from a ket and a bra, we say that we are taking the outer product.

The Identity operator

A general state may not be an eigen state of a Hermitian operator, but we can always write it in terms of the eigen states

$$|\alpha\rangle = \sum_n c_n |a_n\rangle$$

We can find the coefficient c_m by multiplying by the bra $\langle a_m |$

$$\langle a_m | \alpha \rangle = \sum_n c_n \langle a_m | a_n \rangle = c_m$$

So we can write

$$|\alpha\rangle = \sum_n \langle a_n | \alpha \rangle |a_n\rangle = \sum_n |a_n\rangle \langle a_n | \alpha \rangle$$

So we can define the identity operator as

$$1 = \sum_n |a_n\rangle \langle a_n|$$

We will see that this operator is extremely useful.

For example, we can show (assuming $|\alpha\rangle$ is normalized)

$$\langle \alpha | \alpha \rangle = \langle \alpha | \left(\sum_n |a_n\rangle \langle a_n| \right) | \alpha \rangle = \sum_n |\langle a_n | \alpha \rangle|^2 = \sum_n |c_n|^2 = 1$$

i.e there is 100% certainty we will measure one of the eigen values of the operator.

Projection Operator

We can also take one element of the Identity operator

$$\hat{\Pi}_n = |a_n\rangle \langle a_n|$$

Applying this operator to a state will give me back a specific eigen state multiplied by the probability. This is important operator for understanding measurements.

Operators as matrix

Using the identity operator we can write a general operator as

$$X = \sum_m \sum_n |x_m\rangle \langle x_m| X |x_n\rangle \langle x_n|$$

Here we are expressing the operator X in the $|x_n\rangle$ basis

We can recognize $\langle x_m| X |x_n\rangle$ as elements of a matrix X_{mn}

$$X \doteq \begin{pmatrix} \langle x_1| X |x_1\rangle & \langle x_1| X |x_2\rangle & \dots \\ \langle x_2| X |x_1\rangle & \langle x_2| X |x_2\rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

The size of this matrix will be NxN, where N is the number of states needed to describe the Hilbert space, which could be infinite.

In this representation we can identify state kets as column vectors and state bras as row vectors

This is just one possible matrix representation, but we can use a different basis state and the matrix elements will change. If we work in the eigen bases of the operator, the matrix will be **diagonal**.

Example: Spin ½ system

Spin half systems are strange, they can either be aligned or anti-aligned along any given direction, so there are two possible eigen values $\pm \frac{\hbar}{2}$, so the dimension of our Hilbert space is 2.

Let \hat{S}_z be the operator that probes the spin orientation which has eigen states

$$|S_z, +\rangle = |+\rangle, |S_z, -\rangle = |-\rangle$$

Sometimes we label the kets with both the eigenvalue ($\pm \frac{\hbar}{2}$) and the operator it is the eigen value of (S_z). As S_z is the standard choice, we often drop the operator.

we can represent our kets as 2 component column vectors

$$|+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note they are orthogonal and normalized.

We can then write operator \hat{S}_z in the bases of S_z as

$$\hat{S}_z = |+\rangle \langle +| S_z |+\rangle \langle +| + |+\rangle \langle +| S_z |-\rangle \langle -| + |-\rangle \langle -| S_z |+\rangle \langle +| + |-\rangle \langle -| S_z |-\rangle \langle -|$$

Or

$$\hat{S}_z = \begin{pmatrix} \langle +| S_z |+ \rangle & \langle +| S_z |-\rangle \\ \langle -| S_z |+ \rangle & \langle -| S_z |-\rangle \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Now, we could get the exact same expression for \widehat{S}_x in the $|S_x, \pm\rangle$ basis. However, we cannot use them together. We need to pick a basis and stick with it. This means that we need to calculate

$$\widehat{S}_x = \begin{pmatrix} \langle +|S_x|+ \rangle & \langle +|S_x|- \rangle \\ \langle -|S_x|+ \rangle & \langle -|S_x|- \rangle \end{pmatrix}$$

We do not know (yet) what \widehat{S}_x does on $|+\rangle$, only $|\widehat{S}_x, +\rangle$.

We note that the $|\mathcal{S}_x; \pm\rangle$ state can be expressed as

$$|\mathcal{S}_x; \pm\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)$$

We will discuss this more later in the course, a justification is given by Equation 1.110 page 25 Sakurai.

Note this means that $|\mathcal{S}_x; \pm\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$, i.e. it is a rotation.

From this we can now write

$$\begin{aligned} |\pm\rangle &= \frac{1}{\sqrt{2}}(|\mathcal{S}_x; +\rangle \pm |\mathcal{S}_x; -\rangle) \\ S_x|+\rangle &= S_x \frac{1}{\sqrt{2}}(|\mathcal{S}_x; +\rangle + |\mathcal{S}_x; -\rangle) = \frac{1}{\sqrt{2}}\left(\frac{\hbar}{2}|\mathcal{S}_x; +\rangle - \frac{\hbar}{2}|\mathcal{S}_x; -\rangle\right) \\ \langle +|S_x|+\rangle &= \frac{\hbar}{4}(\langle S_x, +| + \langle S_x, -|)(|\mathcal{S}_x; +\rangle - |\mathcal{S}_x; -\rangle) \\ &= \frac{\hbar}{4}(1 - 0 + 0 - 1) = 0 \end{aligned}$$

Repeating these step gives

$$\widehat{S}_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

For completeness

$$\widehat{S}_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Unitary matrices

We have already discussed one important class of matrices, Hermitian Matrices.

Another important class are unitary matrices. A unitary matrix is one that satisfies $U^\dagger = U^{-1}$

Let's generate a new ket $|\alpha'\rangle$ from the ket $|\alpha\rangle$ via a unitary operation U

$$|\alpha'\rangle = U|\alpha\rangle$$

Then the corresponding bra $\langle \alpha'| = \langle \alpha|U^\dagger$

Then the norm of this state is $\langle \alpha' | \alpha' \rangle = \langle \alpha | U^\dagger U | \alpha \rangle = \langle \alpha | \alpha \rangle$

So a unitary operation preserves the norm of a state vector. This is a useful property because it means we preserve probability, we didn't "lose" any of the state. If it was 100% in state $|\alpha\rangle$ it is 100% in state $|\alpha'\rangle$.

Unitary matrices can be used to transform between eigenstates of two observables, (or Hermitian operators).

To prove this we assume that we have two sets of eigenkets $|a_n\rangle, |b_n\rangle$ belonging to two different operators.

We can then write an operator that transforms the eigenstates of $|a_n\rangle$ onto $|b_n\rangle$ as

$$U = \sum_n |b_n\rangle\langle a_n|$$

So

$$U|a_m\rangle = |b_m\rangle$$

We check that it is unitary

$$\begin{aligned} U^\dagger &= \sum_n |a_n\rangle\langle b_n| \\ U^\dagger U &= \sum_n \sum_m |a_m\rangle\langle b_m|b_n\rangle\langle a_n| = \sum_n |a_n\rangle\langle a_n| = 1 \end{aligned}$$

Sometimes we might have an operator expressed in one basis set and we need to express it in another (like S_x in S_z basis)

The elements of X in the $|b_n\rangle$ basis are

$$X'_{mn} = \langle b_m | X | b_n \rangle$$

We can insert two identity operators

$$X'_{mn} = \sum_k \sum_l \langle b_m | a_k \rangle \langle a_k | X | a_l \rangle \langle a_l | b_n \rangle$$

Then removing the $|b_n\rangle$ with a unitary operation

$$X'_{mn} = \sum_k \sum_l \langle a_m | U^\dagger | a_k \rangle \langle a_k | X | a_l \rangle \langle a_l | U | a_n \rangle$$

Which can be written more compactly as

$$\hat{X}' = \hat{U}^\dagger \hat{X} \hat{U}$$

In quantum mechanics, we have a choice, we can either “rotate” the states to the operator, or rotate the operator to the states. We will see that this is the key difference between Schrödinger and Heisenberg interpretations of quantum mechanics.

Trace of an operator

The trace of an operator is an important way to characterize them it is defined by

$$tr(\hat{X}) = \sum_n \langle a_n | \hat{X} | a_n \rangle$$

This is useful because it doesn’t matter what basis I use to calculate the trace, to see this we use the identity operator to change the basis.

$$\begin{aligned} tr(\hat{X}) &= \sum_l \sum_m \sum_n \langle a_n | b_l \rangle \langle b_l | \hat{X} | b_m \rangle \langle b_m | a_n \rangle \\ tr(\hat{X}) &= \sum_l \sum_m \sum_n \langle b_m | a_n \rangle \langle a_n | b_l \rangle \langle b_l | \hat{X} | b_m \rangle \\ tr(\hat{X}) &= \sum_l \sum_m \langle b_m | b_l \rangle \langle b_l | \hat{X} | b_m \rangle \\ tr(\hat{X}) &= \sum_l \langle b_l | \hat{X} | b_l \rangle \end{aligned}$$

Measurements

Measurements are major events in quantum mechanics, they fundamentally change the wavefunction. The assumption of quantum mechanics is that, when you measure a quantum state the system collapses onto one of the eigen states.

If we start with a state $|\psi\rangle = \sum_n c_n |a_n\rangle$, and make a measurement corresponding to the observable operator \hat{A}

$$|\psi\rangle \rightarrow |a_n\rangle$$

This is another postulate of quantum mechanics

We can use the projection operator to find the probability of finding the system in a specific eigen state as

$$P_n = \langle \psi | \hat{\Pi}_n | \psi \rangle = \langle \psi | a_n \rangle \langle a_n | \psi \rangle = |\langle a_n | \psi \rangle|^2$$

And this represents the 3rd postulate.

The eigen values of observable are the possible measurement outcomes, but often we measure an ensemble of quantum systems, thus we measure an average. The average value measured, or expected value is

$$\langle A \rangle \equiv \langle \psi | A | \psi \rangle = \sum_n \sum_m \langle \psi | a_m \rangle \langle a_m | X | a_n \rangle \langle a_n | \psi \rangle$$

$$\begin{aligned}
&= \sum_n \sum_m a_n \langle \psi | a_m \rangle \langle a_m | a_n \rangle \langle a_n | \psi \rangle \\
&= \sum_n \sum_m a_n \langle \psi | a_m \rangle \delta_{mn} \langle a_n | \psi \rangle \\
&= \sum_n a_n |\langle a_n | \psi \rangle|^2
\end{aligned}$$

i.e. The expectation value is the weighted sum of the eigenvalues, weighted by the probability of being in that state.

Compatibility

As we are using matrices to represent observables, it is no surprise that in general the operation $AB \neq BA$.

Two **observables** are said to be compatible if the **commutator**

$$[A, B] = AB - BA = 0$$

Operators are compatible if they are diagonal in the same basis.

Proof:

Lets take the eignestates of an **Hermitian** operator \hat{A} with non degenerate eigenvalues

$$\langle a' | [\hat{A}, \hat{B}] | a \rangle = a' \langle a' | \hat{B} | a \rangle - a \langle a' | \hat{B} | a \rangle = (a' - a) \langle a' | \hat{B} | a \rangle = 0$$

The left hand side is non zero for all values except $a' = a$ so the only allowed non zero values of \hat{B} are $\langle a | \hat{B} | a \rangle$ i.e. B is diagonal. Note that \hat{B} can have degeneracies and this proof still works.

In this case it makes sense to label states with all commuting variables.

$$|a, b, \dots \rangle$$

Note, that if we do have degeneracy the following happens in a measurement. Assume that a state with quantum number a is degenerate with multiple values of b

We first measure A, and get an, but we are in a superposition. If we now measure B we will get a specific b state with probability $|c_i|^2$. However, remeasuring A again will still give us an with 100% certainty. In addition, the value of B will not change.

$$\hat{A}, \hat{B}, \hat{A}$$

$$|\psi\rangle \rightarrow \sum c_i |a_n, b_i\rangle \rightarrow |a_n, b_m\rangle \rightarrow |a_n, b_m\rangle$$

Incompatibility

If two observables are incompatible they have to satisfy an uncertainty principle. For this we define the operator,

$$\Delta A = A - \langle A \rangle$$

which is just the operator minus its expectation value (which is a scalar).

The expectation value of this operator squared is

$$\langle(\Delta A)^2\rangle = \langle A^2 - 2A\Delta A + \langle A \rangle^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2$$

which is the variance of the distribution.

To obtain the inequality we note the Schwarz inequality

$$\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq |\langle \alpha | \beta \rangle|^2$$

We then say $|\alpha\rangle = |\Delta A\rangle$, $|\beta\rangle = |\Delta B\rangle$, where $|\dots\rangle$ means any state, as ΔA is Hermitian $\langle \alpha | = \langle | A$

Which gives

$$\langle(\Delta A)^2\rangle \langle(\Delta B)^2\rangle \geq |\langle \Delta A \Delta B \rangle|^2$$

$\Delta A \Delta B$ can be written as $\frac{1}{2}[\Delta A, \Delta B] + \frac{1}{2}\{\Delta A, \Delta B\}$

Where

$$\{\Delta A, \Delta B\} = \Delta A \Delta B + \Delta B \Delta A$$

is the anticommutator,

It can be trivially shown that $[\Delta A, \Delta B] = [A, B]$

The commutator of two Hermitian operators is **anti Hermitian**, which means that $O^\dagger = -O$

$$[A, B] = AB - BA$$

$$[A, B]^\dagger = (AB)^\dagger - (BA)^\dagger = BA - AB = -[A, B]$$

These operators have imaginary expectation values

$$\langle a | O | a \rangle^\dagger = -\langle a | O | a \rangle$$

This can only be true if the expectation value is *imaginary*

$\{\Delta A, \Delta B\}$ is Hermitian and thus has real expectation values.

Thus $\langle \Delta A \Delta B \rangle = \frac{1}{2}\langle [A, B] \rangle + \frac{1}{2}\{\Delta A, \Delta B\}$ is a real number + imaginary number

So $|\langle \Delta A \Delta B \rangle|^2 = \frac{1}{4}|\langle [A, B] \rangle|^2 + \frac{1}{4}|\langle \{\Delta A, \Delta B\} \rangle|^2$

The minimum value this function can take is $\frac{1}{4}|[A, B]|^2$ because it is the difference of two operators

So $\langle(\Delta A)^2\rangle \langle(\Delta B)^2\rangle \geq \frac{1}{4}|\langle [A, B] \rangle|^2$

Commutator rules

Commutators are very important in QM and there are a few rules that are worth remembering

$$[A, A] = 0$$

$$\begin{aligned}
[A, cB] &= c[A, B] \\
[A + B, C] &= [A, C] + [B, C] \\
[A, BC] &= [A, B]C + B[A, C] \\
[A, [B, C]] + [B, [C, A]] + [C, [A, B]] &= 0
\end{aligned}$$

Postulates of quantum mechanics

To summarize, we have built the mathematical foundation of quantum mechanics and we have claimed that:

1. Any quantum state ϕ can be described by a point in a complex Hilbert space, \mathcal{H}_s , with a norm $\langle \phi | \phi \rangle = 1$
2. A measurement collapses the state onto one of the eigenstates of the operator (and that there is a projection operator for all possible states in the system)
3. The probability of the system being in a state after a measurement is given by $P_n = \langle \psi | \hat{\Pi}_n | \psi \rangle$

There is also a 4th which is needed and we will come to when we discuss angular momentum which is

4. The Hilbert space of a composite system is $\mathcal{H}_s \otimes \mathcal{H}_t$

Note, these are very different postulates than those of general or special relativity, which are based on physical statements. Here the postulates are entirely mathematical and in part explain why it is hard to completely comprehend quantum systems. However, it is not entirely maths driven, these statements are of course coming from our ability to construct a theory that is consistent with what we observe in experiments.

Continuous states

So far we have been discussing discrete Hibert spaces. That is one in which the eigen values are discrete or numerable (but can still be infinite). However, we can also have continuous eigenvalues. In this case the maths needs a slight modification.

We define the position operator \hat{x} which when acting on a state of a particle returns the position of the particle with the eigen states defined as

$$\hat{x}|x\rangle = x|x\rangle$$

The position can be continuous, rather than a discrete set. This means that when we have orthogonal states the Kronecker delta is replace by a Dirac delta function:

$$\langle x' | x \rangle = \delta(x' - x)$$

This means that the definition of the identity operator becomes:

$$\hat{I} = \int_{-\infty}^{\infty} |x\rangle \langle x|,$$

Where $|x\rangle$ can be any orthonormal basis ket.

Definition of the wavefunction

Let's take a general state and expand it in terms of the eigenfunctions of the position operator.

$$|\psi\rangle = \int_{-\infty}^{\infty} dx |x\rangle \langle x|\psi\rangle$$

$\langle x|\psi\rangle$ previously represented a discrete value, but now it represents a function. In fact an important function called the wavefunction

$$\psi(x) = \langle x|\psi\rangle$$

As the eigenvalues of x are continuous, $\psi(x)$ is continuous and presents the probability amplitude of finding a particle at position x . i.e. **it is a complete set of continuous eigenvalues** in the position basis.

We can see that the overlap of two states is

$$\langle \phi|\psi\rangle = \int dx \langle \phi|x\rangle \langle x|\psi\rangle = \int dx \phi^*(x)\psi(x)$$

is the same as the overlap of their wavefunction. This also shows that the probability of the two states being overlapped is independent of how I represent the state because I could equally write the overlap in the momentum representation.

$$\langle \phi|\psi\rangle = \int dp \langle \phi|p\rangle \langle p|\psi\rangle = \int dx \phi'^*(p)\psi'(p)$$

Where $\psi'(p)$ is the momentum wavefunction of the particle. This stresses the point of Dirac notation and kets, they describe a state without us having to specify the basis in advance.

We can also work out expectation values of operators in this way

$$\langle \psi|\hat{A}|\psi\rangle = \int_{\infty}^{\infty} \int_{\infty}^{\infty} dx' dx \langle \psi|x'\rangle \langle x'|\hat{A}|x\rangle \langle x|\psi\rangle = \int_{\infty}^{\infty} \int_{\infty}^{\infty} dx' dx \psi^*(x') \langle x'|\hat{A}|x\rangle \langle x'|\psi\rangle$$

From here, it is easy to show that the expectation of the position operator is

$$\begin{aligned} \langle \psi|\hat{x}|\psi\rangle &= \int_{\infty}^{\infty} \int_{\infty}^{\infty} dx' dx \langle \psi|x'\rangle \langle x'|\hat{x}|x\rangle \langle x|\psi\rangle = \int_{\infty}^{\infty} \int_{\infty}^{\infty} dx' dx \psi^*(x') x \langle x'|\psi\rangle \\ &= \int_{\infty}^{\infty} dx \psi^*(x) x \psi(x) \end{aligned}$$

However, in many cases we have operators that commute with the position operator. For example we often have a potential operator which is a function of position, $\hat{V}(\hat{x}) \propto \hat{x}^2$. The most general operator can be written as

$$\hat{A}(\hat{x}) = \sum_n b_n \hat{x}^n$$

Now

$$[\hat{A}(\hat{x}), \hat{x}] = 0$$

So the eigenvalues of the A operator are eigenvalues of the position operator, which allows us to write

$$\langle \psi | \hat{A}(\hat{x}) | \psi \rangle = \int_{-\infty}^{\infty} dx \psi^*(x) a(x) \psi(x)$$

Where $a(x)$ are the eigenvalues for a value of x.

3D representation

To extend to three dimensions we can define a position operator $\hat{r} = \hat{x} + \hat{y} + \hat{z} \equiv \sum_i x_i$

We then make the assumption that

$$[\hat{x}_i, \hat{x}_j] = 0$$

i.e the different position operators commute. The assumptions comes from our understanding of classical mechanics, that if I first translate along x, then along y I am in the same state as if I first move along y and then along x. This does not apply to rotations as we see later.

As the observables are compatible, we can see that the general position eigenstate is and eigenstate of all operators

$$\hat{x}|r\rangle = x|r\rangle, \hat{y}|r\rangle = y|r\rangle, \hat{z}|r\rangle = z|r\rangle$$

Translations

We can define an infinitesimal translation operator as one which takes the eigen state of the position operator $|x\rangle$ and moves it to $|x + \delta x\rangle$

write this as $\hat{\zeta}(\Delta x)|x\rangle = |x + \delta x\rangle$

$|x\rangle$ is clearly not an eigenstate of the translation operator.

The translation operator on a general state $|\alpha\rangle$ gives

$$|\alpha'\rangle = \hat{\zeta}(\delta x)|\alpha\rangle = \hat{\zeta}(\delta x) \int_{-\infty}^{\infty} dx' |x'\rangle \langle x'|\alpha\rangle = \int_{-\infty}^{\infty} dx' |x' + \delta x\rangle \langle x'|\alpha\rangle = \int_{-\infty}^{\infty} dx' |x' + \delta x\rangle \psi_{\alpha}(x')$$

We can also define a new variable $x' = x - \delta x$ so

$$\hat{\zeta}(\delta x)|\alpha\rangle = \int_{-\infty}^{\infty} dx |x\rangle \psi_{\alpha}(x - \delta x)$$

$$\psi_{\alpha'}(x) = \langle x | \hat{\zeta}(\delta x) | \alpha \rangle = \psi_{\alpha}(x - \delta x)$$

This shows I have two options when defining my translations. I can either change my basis and keep the wavefunction fixed $|x\rangle \rightarrow |x + \delta x\rangle$ or I can keep my basis fixed and shift the wavefunction in the opposite direction $\psi_{\alpha}(x) \rightarrow \psi_{\alpha}(x - \delta x)$. These are equivalent operations.

What is $\hat{\zeta}(\delta x)$?

1. We need it to conserve probability i.e. $\langle \alpha' | \alpha' \rangle = \langle \alpha | \alpha \rangle$. This means that $\hat{\zeta}(\Delta x)$ should be unitary.

2. We should be able to do two translations in a single translation i.e.

$$\zeta(\delta x_2)\zeta(\delta x_1) = \zeta(\delta x_1 + \delta x_2)$$

3. We should be able to invert the translation by going backwards i.e.

$$\zeta^{-1}(\delta x) = \zeta(\delta x)$$

4. The translation operator should be the indemnity in the limit $\delta x \rightarrow 0$

$$\lim_{\Delta x \rightarrow 0} \hat{\zeta}(\delta x) = \hat{I}$$

We can then guess the form of the translation operator as

$$\hat{\zeta}(\delta x) = \hat{I} - i\hat{\mathbf{K}} \cdot \delta x$$

To satisfy 1. We need:

$$\zeta(\delta x)\zeta^\dagger(\delta x) = \hat{I}$$

$$\zeta(\delta x)\zeta^\dagger(\delta x) = (\hat{I} - i\hat{\mathbf{K}} \cdot \delta x)(\hat{I} + i\hat{\mathbf{K}}^\dagger \cdot \delta x)$$

As δx is infinitesimal, we only keep the linear terms so

$$\zeta(\delta x)\zeta^\dagger(\delta x) = (\hat{I} + i(\hat{\mathbf{K}} - \hat{\mathbf{K}}^\dagger) \cdot \delta x) = 1$$

Therefore $\hat{\mathbf{K}}$ should be a Hermitian operator. It should also have units of inverse length.

It is straightforward to show that conditions 2-4 are also fulfilled by this operator.

Commutation relations

Lets looking at how $\hat{\zeta}$ commutes with \hat{x}

$$[\hat{x}, \hat{\zeta}(\delta x)]|x\rangle = (\hat{x}\hat{\zeta}(\delta x) - \hat{\zeta}(\delta x)\hat{x})|x\rangle = (x + \delta x)|x + \delta x\rangle - x|x + \delta x\rangle = \delta x|x + \delta x\rangle$$

$$[\hat{x}, \hat{\zeta}(\delta x)]|x\rangle = \delta x|x + \delta x\rangle$$

$$[\hat{x}, \hat{\zeta}(\delta x)]|x\rangle \approx \delta x|x\rangle$$

$$[\hat{x}, \hat{\zeta}(\delta x)] = \delta x$$

We substitute in for $\hat{\zeta}$

$$[\hat{x}, \hat{I} - i\hat{\mathbf{K}} \cdot \delta x] = \delta x$$

$$[\hat{x}, -i\hat{\mathbf{K}} \cdot \delta x] = \delta x$$

$$[\hat{x}, \hat{\mathbf{K}}] = i$$

We can now redefine $\hat{\mathbf{K}} = \frac{\hat{\mathbf{p}}}{\hbar}$

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

$\hat{\mathbf{p}}$ really is the momentum operator. This is known as the **canonical commutator relationship**.

Dirac noted that quantum commutators can be obtained from classical Poisson brackets found in Hamiltonian mechanism by added an $i\hbar$! $[,]_c \rightarrow \frac{[,]_q}{i\hbar}$

If we put this into our uncertainty relation earlier we get the Famous Heisenberg relation

$$\langle (\Delta x)^2 \rangle \langle (\Delta p)^2 \rangle \geq \frac{\hbar^2}{4}$$

Finally, if we write out the translation operator,

$$\hat{\zeta}(\delta x) = \left(\hat{I} - \frac{i\mathbf{P}}{\hbar} \cdot \delta x \right)$$

We see that momentum is the generator of translations.

Finite Translations

So far we have only discussed infinitesimal translations. To do a finite translation we need to do a series of infinitesimal translations. To move a finite distance Δx , we split the journey up into N steps of size $\Delta x/N$ and let $N \rightarrow \infty$

$$\hat{\zeta}(\Delta x) = \lim_{N \rightarrow \infty} \left(\hat{I} - \frac{i\mathbf{P}}{\hbar} \cdot \frac{\Delta x}{N} \right)^N = \exp \left(- \frac{i\mathbf{P}}{\hbar} \cdot \Delta x \right)$$

Momentum Basis

Earlier we had that the translated wave function can be written as

$$\psi_{\alpha'}(x) = \psi_{\alpha}(x - \delta x)$$

If we Taylor expand the old wavefunction we get

$$\psi_{\alpha'}(x) = \psi_{\alpha}(x - \delta x) \approx \psi_{\alpha}(x) - \delta x \frac{\partial \psi_{\alpha}}{\partial x} = \left(1 - \delta x \frac{\partial}{\partial x} \right) \psi_{\alpha}$$

We can re-write this in Dirac notation as

$$\begin{aligned} \langle x | \hat{I} - \frac{i\hat{p}}{\hbar} \cdot \delta x | \alpha \rangle &= \left(1 - \delta x \frac{\partial}{\partial x} \right) \langle x | \alpha \rangle \\ \langle x | \hat{p} | \alpha \rangle &= -i\hbar \frac{\partial}{\partial x} \langle x | \alpha \rangle \end{aligned}$$

And

$$\langle \beta | \hat{p} | \alpha \rangle = \int \langle \beta | x \rangle \langle x | \hat{p} | \alpha \rangle = -i\hbar \int \psi_{\beta}^{*} \frac{\partial}{\partial x} \psi_{\alpha}$$

using

$$\langle \beta | = \int \langle \beta | x \rangle \langle x |$$

Which are two important relations.

Momentum eigenstates

So far we have been discussing base gets in terms of the eigen kets of the position operator, but we can also use eigen kets of the momentum operator

$$\hat{p}|p\rangle = p|p\rangle$$

We can also define a momentum wavefunction as

$$\Psi_\alpha(p) = \langle p|\alpha\rangle$$

We can also convert from position and momentum with the following

$$\psi_\alpha(x) = \langle x|\alpha\rangle = \int \langle x|p\rangle\langle p|\alpha\rangle dp = \int \langle x|p\rangle\Psi(p)dp$$

$\langle x|p\rangle$ is a particularly useful overlap. To determine it we make use of

$$\langle x|\hat{p}|\alpha\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|\alpha\rangle$$

And take the state $|\alpha\rangle$ to be a momentum eigenstate

$$\langle x|\hat{p}|p\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|p\rangle$$

$$p \langle x|p\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|p\rangle$$

$$\langle x|p\rangle = N \exp\left(\frac{ipx}{\hbar}\right)$$

To find N we note that

$$\langle x'|x\rangle = \delta(x' - x) = \int dp \langle x'|p\rangle\langle p|x\rangle = N^2 \int dp \exp\left(\frac{ip(x' - x)}{\hbar}\right) = 2\pi\hbar N^2 \delta(x' - x)$$

So

$$N = \frac{1}{\sqrt{2\pi\hbar}}$$

Giving

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \exp\left(\frac{ipx}{\hbar}\right)$$

Putting this back in to the expression for $\psi(x)$ we get

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int dp \psi(p) \exp\left(\frac{ipx}{\hbar}\right)$$

Similarly

$$\psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dp \psi(x) \exp\left(-\frac{ipx}{\hbar}\right)$$

i.e. These functions are just the Fourier transform of each other!

Minimum uncertainty state

Now lets assume that our wavefunction is a gaussian distribution, centered at the origin.

$$\psi(x) = \langle x | \psi \rangle = \frac{1}{\pi^{\frac{1}{4}} \sqrt{d}} \exp\left(-\frac{x^2}{2d^2}\right) \exp(ikx)$$

Using

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

$$\langle \hat{x} \rangle = \langle \psi | \hat{x} | \psi \rangle = \int dx' x' |\psi(x')|^2$$

$|\psi(x')|^2$ is an even function, so $\langle \hat{x} \rangle = 0$

Then

$$\begin{aligned} \langle x^2 \rangle &= \frac{1}{\sqrt{\pi} d} \int dx x^2 \exp\left(-\frac{x^2}{d^2}\right) = \frac{1}{\sqrt{\pi} d} \int dx x x \exp\left(-\frac{x^2}{d^2}\right) \\ &= \frac{1}{\sqrt{\pi} d} \left\{ \left[\frac{d^2}{2} x \exp\left(-\frac{x^2}{d^2}\right) \right]_{-\infty}^{\infty} + \frac{d^2}{2} \int dx \exp\left(-\frac{x^2}{d^2}\right) \right\} \\ &= \frac{1}{\sqrt{\pi} d} \left\{ 0 + \frac{d^2}{2} d \sqrt{\pi} \right\} \end{aligned}$$

Alternatively,

$$\int dx' |\psi(x')|^2 = 1$$

$$\text{Let } a = \frac{1}{d}$$

$$\frac{\partial}{\partial a} \int dx' |\psi(x')|^2 = \frac{1}{a} \int dx' |\psi(x')|^2 - 2a \int dx' x^2 |\psi(x')|^2 = 0$$

$$\int dx' x^2 |\psi(x')|^2 = \frac{1}{2a^2}$$

$$\langle x^2 \rangle = \frac{d^2}{2}$$

So

$$\langle (\Delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle = \frac{d^2}{2}$$

For the momentum we have

$$\langle p \rangle = -i\hbar \int dx \psi^*(x) \frac{\partial}{\partial x} \psi(x)$$

It is easy to see that

$$\frac{\partial}{\partial x} \psi(x) = \left(ik - \frac{x}{d^2} \right) \psi(x)$$

so

$$\langle p \rangle = -i\hbar \int dx \psi^*(x) \left[ik - \frac{x}{d^2} \right] \psi(x)$$

The last term is zero because it is odd leaving

$$\begin{aligned} \langle p \rangle &= \hbar k \int |\psi(x)|^2 = \hbar k \\ \langle p^2 \rangle &= -\hbar^2 \int dx \psi^*(x) \frac{\partial^2}{\partial x^2} \psi(x) \\ \frac{\partial^2}{\partial x^2} \psi(x) &= \left\{ \left(ik - \frac{x}{d^2} \right)^2 - \frac{1}{d^2} \right\} \psi(x) \end{aligned}$$

We only need to keep the even terms

$$\int \left(-k^2 - \frac{1}{d^2} + \frac{x^2}{d^4} \right) |\psi(x)|^2 dx = -k^2 - \frac{1}{d^2} + \frac{1}{2d^2}$$

So

$$\langle p^2 \rangle = \hbar^2 k^2 + \frac{\hbar^2}{2d^2}$$

$$\text{And } \langle (\Delta p)^2 \rangle = \frac{\hbar^2}{2d^2}$$

$$\text{And so, finally we are at the uncertainty relationship } \langle (\Delta x)^2 \rangle \langle (\Delta p)^2 \rangle = \frac{\hbar^2}{4}$$

Or

$$\langle \Delta x \rangle \langle \Delta p \rangle = \frac{\hbar}{2}$$

So the lowerbound of the uncertainty principle is an equality for a Gaussian wavepacket. Therefore, Gaussian wavepackets are often called the minimum uncertainty wavepacket.

We can the more localized we make a state in x, the broader it becomes in p, and p tends to infinity as the wavepacket is localized to a delta function.

In non-relativistic QM, particles can be completely localized. However in relativistic QM this cannot be the case. The energy of the particle will increase the more localized it becomes, at some point it will start to create pairs of particles, at which point non-relativistic QM breaks down... More on that later.