SCPY 323: Quantum Mechanics II **2nd Semester 2022**

Lecture 1: Concepts in Quantum Mechanics

Kit Matan

Mahidol University

Reading assignment: *Handout from "Feynman lectures on physics"*

1.1 Young's Double Slit Experiment for Electrons

We probably all heard of Young's famous double-slit experiment, which illustrates the interference of light when going through double slits. Here we will imagine a similar experiment but performed on electrons. By now, we know that when a light goes through double slits, we will see the interference pattern on a screen behind the slits. You can imagine that light that goes through the first slit interferes with light that goes through the second slit, giving rise to the interference. However, if you use bullets instead of light, you also know that we will not see the interference pattern on the screen. Instead, you will see a lot of bullet holes right behind each slit and nowhere else. And you know right away which slit each bullet went through. Therefore, we can conclude that the difference between the bullets and light is that for the bullets we know which slit they went through but for light we have no way to distinguish light that went through one slit from the other. So, in this case, we can use this fact (be able to tell or not be able to tell) to distinguish between particles and waves.

Now let us shift our focus to electrons, suppose that we have an electron gun, which can be a tungsten that is heated up so that some free electrons are emitted and we can apply voltage difference to force those electrons to move toward the slits. Behind the double slits on the screen, suppose that we have a detector that moves parallel to the separation of the slits to count the number of electrons that land on that spot.

After performing the experiment by letting a lot of electrons go through the double slits and counting the number of electrons at each point along the line, we will observe an interference pattern similar to the one we obtain from the same experiment but using light. However, we believe that electrons are particles; hence they should behave like bullets, that is, we should be able to tell which slit each electron went through. How can we tell which slit each electron went through? One way is to use a "flashlight" (of course not a regular flashlight but I think you get the analogy here). That is, we will shine the light just behind each slit and when an electron goes through one of the slits we will see light reflected off that electron so you know which slit that electron went through. So, now you know for all electrons that land on the screen which slit they went through. However, to your surprise, you will no longer find the interference pattern on the screen. What you see is that those electrons that went through the first (top) slit will land on the screen right behind that slit; the same thing happens for the second (bottom) slit.

But you are clever, and you know that by shining light on the electrons to see which slit they went through, you probably have perturbed the trajectory of the electrons. Therefore, you decide to do two things to fix this problem:

1. You decrease the intensity of your flashlight so that fewer photons would hit electrons and you hope that probably the trajectory of the electrons does not change.

2. You decrease the energy of each photon by using long-wave-length light so that less energy is transferred from photons to electrons.

The First Case

If you decrease the intensity of light, which is equivalent to decreasing the number of photons. Therefore, there will be some electrons that landed on the screen but you don't know which slit it went through since no photon hits that electron when it went through the slit. Therefore, in this case, you will have two kinds of electrons. The first kind is those that reflect off the light and you know which slit they went through while the other is those that did not reflect light and you do not know which slit they went through. What you will find experimentally is that if you only plot the positions on the screen where those electrons in the first group landed, you find the interference pattern while if you only plot the position of those electrons in the second group, you will find no interference pattern and those electrons will simply land right behind each slit. Therefore, you cannot tell which slit the electrons went through and at the same time get the interference pattern.

The Second Case

What about if you decrease the energy of photons so that the electrons do not get perturbed that much? You can decrease the energy of photons by decreasing their frequencies, that is, choosing long-wavelength light. And you hope that if the energy of photons is small energy then the trajectory of the electrons will not get perturbed that much and you would be able to tell which slit they went through and at the same time get the interference pattern.

However, what you would find experimentally is that if you use long-wavelength light, then at some point you will not be able to distinguish the two slits, that is, you will not be able to tell exactly where those two slits are. Therefore, you will not be able to tell for sure which slit the electrons went through. So, you could be able to get the interference pattern on the screen if you use the flashlight with a long-enough wavelength but then you will not be able to tell which slit the electrons went through. On the other hand, you will not get the interference pattern if the wavelength of the light you use is too short since it will perturb the trajectory of electrons just enough to eradicate the interference pattern. Therefore, again, you cannot tell which slit the electrons went through and at the same time get the interference pattern.

The conclusion of this experiment is the fundamental concept of quantum mechanics, that is, the uncertainty principle. This version of the uncertainty principle states that *we cannot design a double-slit experiment that allows us to know exactly which slit the electrons go through and at the same time observe the interference.*

In classical mechanics, the world is deterministic, that is, we know exactly what the system will be in the future providing that we know its initial conditions and the force acting on the system. This is not true for the world of quantum mechanics.

1.1.1 Interpretations of Quantum Mechanics

When we do the measurements on the quantum mechanical system, we cannot know exactly what the outcome will be. The best we can do is to give the probability of each possible outcome of the measurements. The system is described by a wavefunction, which is used to determine the probability of each outcome. Although the time evolution of this wavefunction is deterministic, we have no way again to tell what the future outcome of the measurement will be. And once the measurement is done, the system will stay in the particular system; this phenomenon is called the collapse of the wave function. This probabilistic interpretation of quantum mechanics is called the Copenhagen interpretation. The essence of this interpretation was proposed

by Niels Bohr and Werner Heisenberg.

Another interpretation of quantum mechanics is "many world interpretation", where people believe that each time we perform a measurement our world split into many, one for each possible outcome of the measurement. This interpretation may sound absurd to you but physicists have no way to disprove these interpretations so far.

The last attempt to understand the probabilistic aspect of quantum mechanics was supported by Einstein. Einstein believes that there are "hidden variables" embedded in the theory of quantum mechanics. These hidden variables prevent the theory to become deterministic. And the probabilistic aspect of quantum mechanics arises from these hidden variables, which give the appearance of randomness to the outcome. Once these hidden variables are known then the quantum mechanics will be the deterministic theory. However, this argument is at odd with all experimental results so far.

1.2 Classical vs. Quantum Framework

When physicists first studied the physics of atoms, they extended on physics theories they already knew at the time; those includes Newtonian mechanics and electromagnetism. Therefore, in order to get some intuition about quantum mechanics, we will make a comparison between the classical physics of Newton and quantum physics, which the 20^{th}-century physicists use to explain the physics of atoms.

Classical Physics Framework

- A state at a fixed time t can be defined by a point with a coordinate $(\vec{x}, \vec{p})$ in a phase space.
- An observable is a function of points on the phase space, *i.e.* $\Theta(\vec{x}, \vec{p})$. Some examples of the observable are kinetic or potential energy, angular momentum, total energy, and etc.
- A Hamiltonian $\mathcal{H}$ defines all dynamics of the system through Hamilton's equations of motion, which are:

$$\begin{aligned} \dot{x}_i &= \frac{\partial \mathcal{H}}{\partial p_i} \\ -\dot{p}_i &= \frac{\partial \mathcal{H}}{\partial x_i} \end{aligned}$$

- This framework described all classical physics, such as classical mechanics, electromagnetism, fluids dynamics, wave mechanics, etc.
- It is a simple and intuitive framework.
- The formulation is deterministic. It is completely determined by the Hamiltonian and initial conditions.
- Time dynamics is reversible.

Now we will turn our attention to the quantum physics framework.

Quantum Physics Framework

- A state in quantum mechanics is represented by a vector in a Hilbert vector space.
- An observable is a Hermitian operator, where $\Theta^\dagger = \Theta$ (self-adjoint).

- The dynamics or time evolution is governed by the Hamiltonian (Hermitian operator) $\mathcal{H}$. A state at time t can be described by
$$|v(t)\rangle = e^{-i\mathcal{H}t/\hbar}\,|v(0)\rangle\,,$$
where $|v(0)\rangle$ is an initial state at $t = 0$.

- **Collapse Postulate**: This phenomenon has no classical counterpart. It can be simply stated that a state after the measurements will be one of the eigenstates of the operator (the observable) that has been measured. This state can be different from the initial state. In other words, the state collapses into one of the eigenstates. In a mathematical form, this can be expressed as follows. Let $|\psi\rangle$ be an initial state. We can write $|\psi\rangle$ as a superposition of a complete set of the eigenstates:
$$|\psi\rangle = \sum C_i\,|a_i\rangle,$$
where C_i is complex and $|a_i\rangle$ is a eigenstate of an operator (observable) A, which we want to measure; $A\,|a_i\rangle = a_i\,|a_i\rangle$, where a_i is an eigenvalue. We will assume here that $a_i \neq a_j$ if $i \neq j$. Now if A is measured, then
  1. The probability of obtaining the eiganvalue a_i is equal to $|C_i|^2$. Note that the measurement can only yield one of the eigenvalues of A.
  2. After the measurement, $|\psi\rangle \to |a_i\rangle$, the collapse of the state.

- This framework describes all physical systems except those that involve the gravitational force.

- It is rather difficult and not intuitive.

- It is non-deterministic, in the sense that one cannot predict for a certain outcome of the measurement. We can only calculate the probability of each outcome.

- Time dynamics is not reversible due to the collapse of the state.

The nature of quantum mechanics, which we will explore further in a future lecture, can be well illustrated by Young's double-slit experiment of electrons. From the discussion of the experiment, we discovered two fundamental aspects of quantum mechanics; those are

1. **Uncertainty principle:** we learned that no matter how much we try, we cannot design an experiment so that we can tell which slit an electron goes through and at the same time observe interference.
2. **Probability amplitudes (wave functions)**: we learned that in order to describe the results of the experiment, we have to create a quantity called probability amplitude ϕ, which is a complex number. To compare with the experiment, we have to calculate the probability $P = \phi^*\phi = |\phi|^2$, where ϕ^* is the complex conjugate of ϕ.

First, I will show how we can explain the double slit experiment of electrons using the probability amplitude.

Rough Mathematical description of the double experiment of electrons

We found that in the double-slit experiment, the system can have two states for each electron that goes through the slits. The first state is where the electron goes through the first slit (top one or S1) and the second state is where the electron goes through the second slit (bottom one or S2). We can represent these two states by two probability amplitudes; ϕ_1 for the former and ϕ_2 for the latter. Therefore, when the electrons arrive at the two slits, they can be in the state that is a superposition of these two states, that is, they can go through S1 or S2. To calculate the probability of an event when we do the measurement, we

have to take the amplitude squared of the probability amplitude. The main point is that you only have to calculate the probability only when you do the measurement. After doing the measurement, the system will stay in that state; this is called the "collapse of the wavefunction" or in this case the collapse of the state.

In terms of the probability amplitude, we can explain the double-slit experiment as follows:

1. As we mentioned before, for the probability amplitude ϕ, we can calculate the probability from

$$P = \phi^* \phi = |\phi|^2,$$

where ϕ is a complex number.

2. When you do the measurements to find out whether the electron goes through S1 or S2, the probability amplitude collapses at the point of measurement, that is, right behind each slit. Therefore, the probability of finding the electron at S1 and S2, which we will call P_1 and P_2 respectively, is equal to

$$\begin{aligned} P_1 &= |\phi_1|^2 \\ P_2 &= |\phi_2|^2 \end{aligned}$$

Since ϕ_1 and ϕ_2 are complex numbers, we can represent them as

$$\begin{aligned} \phi_1 &= a_1 + ib_1 \\ \phi_2 &= a_2 + ib_2 \end{aligned}$$

where $a_{1,2}$ and $b_{1,2}$ are real numbers. Therefore, we find that

$$\begin{aligned} P_1 &= |\phi_1|^2 = a_1^2 + b_1^2 \\ P_2 &= |\phi_2|^2 = a_2^2 + b_2^2 \end{aligned}$$

We assume that P_1 and P_2 are "normalized" such that $P_1 + P_2 = 1$, that is, the electron has to either go through S1 or S2. After the measurement, the probability amplitude collapse to S1 or S2 state, and what we observe is that the electron landed on the screen right behind S1 or S2 and no interference.

3. However, if we decide to do the measurement on the screen, the electron can reach the screen from two paths; one path through S1 and the other path through S2. Therefore, the probability amplitude at the point on the screen is the superposition of the probability amplitude ϕ_1 and ϕ_2, that is,

$$\phi_{\rm sc} = \phi_1 + \phi_2,$$

where $\phi_{\rm sc}$ is the probability amplitude at some particular point on the screen. Now, if you decide to do the measurement at that point on the screen, we find that the probability of finding the electron there is equal to

$$\begin{aligned} P_{12} &= |\phi_{\rm sc}|^2 = |\phi_1 + \phi_2| = |(a_1 + a_2) + i(b_1 + b_2)|^2 \\ &= (a_1 + a_2)^2 + (b_1 + b_2)^2 \\ &= a_1^2 + b_1^2 + a_2^2 + b_2^2 + 2(a_1 a_2 + b_1 b_2) \\ &= P_1 + P_2 + 2(a_1 a_2 + b_1 b_2), \end{aligned}$$

where we can see that P_{12} is not equal to $P_1 + P_2$ and the extra term in the expression above gives rise to the interference. Therefore, what we observe in this case is that there is an interference pattern on the screen.

As you can see, the superposition of the two probability amplitude can create interference, and the description of the quantum state using the probability amplitude appears to work very well and gives a prediction that is consistent with the experimental results.

Now we know that we cannot tell exactly what slit the electrons go through without doing the measurement and in the process collapsing the system's probability amplitude. In other worlds, we are not able to tell the position of the electron exactly without disturbing it. However, what we can predict now is the probability P of finding the electron at that position given that we know the probability amplitude ϕ of the electron at that point, that is, $P = |\phi|^2$.

But ϕ is just a probability amplitude of just a single point. What if you want to find the probability of finding the electron at any point in space? To be able to do that, we need to calculate the probability amplitude at any point in space. Not only that, but since the probability amplitude can vary with time, we have to know the probability amplitude as a function of time as well. Therefore, ϕ is now a function of both x and t, that is, $\phi(x,t)$. We call $\phi(x,t)$ the *probability wave function* or *wave function* of a particle (in this case an electron). And knowing $\phi(x,t)$, we can calculate the *probability density*, that is, $p(x,t) = |\phi(x,t)|^2$. We note here that $p(x,t)$ is NOT the probability since you cannot really find a particle at one particular point; a point is not well defined in quantum mechanics since we cannot locate a particle to a single point. Therefore, what we can calculate is the probability of finding a particle in some given range, say, between x_1 and x_2 at time t, which we will call $P_{[x_1,x_2]}$. The expression for $P_{[x_1,x_2]}$ is given by

$$P_{[x_1,x_2]} = \int_{x_1}^{x_2} |\phi(x,t)|^2 \, dx.$$

Note that since the probability of finding a particle at all space must be equal to 1, to be physically meaningful, our wave function has to be normalizable, that is,

$$P_{[-\infty,\infty]} = \int_{-\infty}^{\infty} |\phi(x,t)|^2 \, dx = 1.$$

We assume here and from now on that our wave function is defined from $-\infty$ to ∞. The range of integration can change to $[a,b]$ if the wave function is only defined in that range, for example; note that if that is the case, then x_1 and x_2 must be in this range $[a,b]$ as well. This is an important point. You have to make sure that your wave function has to normalizable and is normalized before you calculate the probability.

1.3 Mathematics of Vector Space

From the first postulate of quantum mechanics, we learn that:

> A state of a quantum mechanical system at time t is given by a vector $|v(t)\rangle$ in a Hilbert space H.

Therefore, in order to understanding the physical meaning of quantum mechanics, we need to first understand the underlying mathematical vehicle that we will use to describe it. There are two key-words that we need to study here; the first one is *vector* and the other is *Hilbert space*.

Vector Space

Normally in physics, if you hear the word vector, you would visualize an arrow, which indicates both direction and length. However, here we will talk about a vector in a more general, abstract, and mathematical sense but it is useful to have a picture of an arrow in mind in some cases.

Definition 1: A linear vector space $\mathbb{V}$ is a collection of objects called vectors, which will be represented by $|v_1\rangle$, $|v_2\rangle$, ..., (the symbol $|\ \rangle$ is called a *ket*) and for which there exists

1. the vector sum denoted by $|v_1\rangle + |v_2\rangle$
2. multiplication by scalars a, b, ..., denoted by $a\,|v\rangle$,

with the following properties:

1. *Closure*: $|v_1\rangle + |v_2\rangle \in \mathbb{V}$.
2. *Distributive in the vectors*: $a\,(|v_1\rangle + |v_2\rangle) = a\,|v_1\rangle + a\,|v_2\rangle$.
3. *Distribution in the scalars*: $(a+b)\,|v\rangle = a\,|v\rangle + b\,|v\rangle$.
4. *Associative of scalar multiplication*: $a(b\,|v\rangle) = ab\,|v\rangle$.
5. *Commutative of vector addition*: $|v_1\rangle + |v_2\rangle = |v_2\rangle + |v_1\rangle$
6. *Associative of vector addition*: $|v_1\rangle + (|v_2\rangle + |v_3\rangle) = (|v_1\rangle + |v_2\rangle) + |v_3\rangle$.
7. There exist a *null vector* $|0\rangle$ such that $|v\rangle + |0\rangle = |v\rangle$.
8. For $|v\rangle$ there exist an inverse under addition $|-v\rangle$ such that $|v\rangle + |-v\rangle = 0$.

Definition 2: The scalar numbers a, b, ..., are called the field over which $\mathbb{V}$ is defined.

If these numbers are real, we have a real vector space normally denoted by $\mathbb{V}(\mathbb{R})$. On the other hand, if these numbers are complex, we have a complex vector space denoted by $\mathbb{V}(\mathbb{C})$. In our case, we only deal with the complex vector space.

Definition 3: The set of vector is said to be *linearly independent* if

$$\sum_{i=1}^{n} a_i\,|i\rangle = |0\rangle\,,$$

then all $a_i = 0$.

Definition 4: A vector space has *dimension* n if it has a maximum of n linearly independent vector, which will be denoted by $\mathbb{V}^n(\mathbb{R})$ if the field is real and $\mathbb{V}^n(\mathbb{C})$ is the field is complex.

Theorem 1: A vector $|v\rangle$ in an n-dimensional vector space can be written as

$$|v\rangle = \sum_{i=1}^{n} a_i\,|i\rangle\,,$$

where $|i\rangle$ are n linearly independent vectors. We call a set of these n linearly independent vectors *a basis* and say that this basis *spans* the vector space.

So now I think you know what the vector and vector space is. We will next define the Hilbert space. Hilbert space is a vector space that has the following properties:

1. It is complete, which means all Cauchy sequence converges to an element in the vector space.
2. It contains an inner product.

We will not concern with completeness but will only focus on the concept of the inner product in Hilbert space.

Inner Product

When talking of the inner product, I would like you to think of the dot product in a usual three-dimensional space. We will denote $\langle v_1|v_2\rangle$ for the inner product. The inner product gives a scalar number and if the field is real, this scalar number is real; on the other hand, if the field is complex, then this scalar number is complex. Again, in our case, the inner product gives a complex number. Furthermore, the inner product has to have the following properties

1. *Skew-symmetry*: $\langle v_1|v_2\rangle = \langle v_2|v_1\rangle^*$, where * denotes complex conjugate.
2. Positive semi-definiteness: $\langle v|v\rangle \geqslant 0$ and is equal to 0 if and only if $|v\rangle = |0\rangle$.
3. *Linearity in ket*: $\langle v_1|\left(a\,|v_2\rangle + b\,|v_3\rangle\right) = a\,\langle v_1|v_2\rangle + b\,\langle v_1|v_3\rangle$.

We call the symbol $\langle\ |$ a *bra*. A vector space with the inner product is called an *inner product space*. Therefore, the Hilbert space is the complete inner product space with *a norm* defined by the inner product, that is, the norm $|v| = \sqrt{\langle v|v\rangle}$. For any inner product that has these properties, it must also satisfy *Schwarz and Triangle Inequalities*, which state that

- *Schwarz Inequality*: $|\,\langle v|w\rangle\,| \leqslant |v||w|$.
- *Triangle Inequality*: $|v + w| \leqslant |v| + |w|$.

Two vectors $|v\rangle$ and $|w\rangle$ are *orthogonal* if their inner product vanishes, that is, $\langle v|w\rangle = 0$, and a set of the basis of all unit norm that is pairwise orthogonal is called an *orthogonal basis*. We can create the orthogonal basis in a n-dimensional vector space from the n linearly independent vectors by using the *Gram-Schmidt Procedure*, which consists of the following:

- Rescale the first vector to create a unit vector, which will be our first basis vector.
- Subtract the component of the second vector along the first basis vector from the second vector, leaving only the perpendicular component. Then, normalize the length of this vector to one.
- Create the next basis vector by subtracting from it the components along all previous basis vectors and normalizing its length to one.
- Continue this process until you have n orthogonal unit vectors.

The inner product of the orthogonal basis is very simple. Let $|i\rangle$ and $|j\rangle$ be orthogonal basis vectors, then

$$\langle i|j\rangle \equiv \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

where δ_{ij} is called the *Kronecker delta*. Since

Now we will only focus on a complex vector space. Let $|v\rangle$ and $|w\rangle$ be vectors in the complex vector space with n- dimensions and let $|i\rangle$ be a set of the orthogonal basis. We know that $|i\rangle$ spans the vector space; therefore we can write

$$|v\rangle = \sum_{i=1}^{n} v_i\,|i\rangle \quad \text{and} \quad |w\rangle = \sum_{i=1}^{n} w_i\,|i\rangle\,.$$

To calculate the inner product of these two vectors, we introduce a dual vector space whose element is denoted by a *bra* $\langle|$;

$$\langle v| = \sum_{i=1}^{n} v_i^* \langle i| ,$$

where v_i^* is a complex conjugate of v_i. And the inner product can be readily calculated as

$$\langle v|w\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i^* w_j \langle i|j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i^* w_j \delta_{ij} = \sum_{i=1}^{n} v_i^* w_i$$

Alternatively, we can represent a ket $|w\rangle$ by a matrix such that

$$|w\rangle = \sum_{i=1}^{n} w_i |i\rangle \leftrightarrow \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \begin{matrix} |1\rangle \\ |2\rangle \\ \vdots \\ |n\rangle \end{matrix}$$

and a bra $\langle v|$ by

$$\langle v| = \sum_{i=1}^{n} v_i^* \langle i| \quad \leftrightarrow \quad \begin{bmatrix} v_1^*, & v_2^*, & \cdots & v_n^* \end{bmatrix} \\ \langle 1|, \quad \langle 2|, \quad \cdots \quad \langle n|$$

which is just a complex transpose (or adjoint) of $|v\rangle$. Hence, the inner product is just the simple multiplication of these two matrices, that is,

$$\langle v, w\rangle \leftrightarrow \begin{bmatrix} v_1^*, & v_2^*, & \cdots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = v_1^* w_1 + v_2^* w_2 + \cdots v_n^* w_n = \sum_{i=1}^{n} v_i^* w_i .$$

1.3.1 Linear Operators

A linear operator basically takes one vector and transforms it to another vector, which can be the same as or different from the input vector. The action of a operator Ω on a vector $|V\rangle$ is represented by

$$\Omega |V\rangle = |V'\rangle .$$

We will only consider a case where both $|V\rangle$ and $|V'\rangle$ are in the same vector space $\mathbb{V}$. Similarly, the operator can act on *a bra*, that is,

$$\langle W| \Omega = \langle W'| .$$

Again, both $\langle W|$ and $\langle W'|$ must be in the same *dual* space.

For the linear operator, which we will only concern with here, there exist the following properties for a ket state:

$$\begin{aligned} \Omega\alpha |V\rangle &= \alpha\Omega |V\rangle \\ \Omega\{\alpha |V\rangle + \beta |W\rangle\} &= \alpha\Omega |V\rangle + \beta\Omega |W\rangle , \end{aligned}$$

and similarly for a bra state, that is,

$$\begin{aligned} \Omega\alpha \langle V| &= \alpha\Omega \langle V| \\ \{\langle V| \alpha + \langle W| \beta\}\Omega &= \langle V| \alpha\Omega + \langle W| \beta\Omega, \end{aligned}$$

Some examples of linear operators are the projection operator that we talked about in the previous class and the rotation operator of a unit vector in three-dimensional space.

Given an operator Ω, if its action on the basis vectors is known then its action on any vectors is also known, that is, if

$$\Omega |i\rangle = |i'\rangle ,$$

for the basis $|1\rangle$, $|2\rangle$, $|3\rangle$, $\ldots$, $|n\rangle$ in $\mathbb{V}^n$, then for any $|V\rangle \in \mathbb{V}^n$, $|V\rangle = \sum v_i |i\rangle$

$$\Omega |V\rangle = \Omega \sum v_i |i\rangle = \sum v_i \Omega |i\rangle = \sum v_i |i'\rangle .$$

This is also true if you interchange $|\ \rangle$ for $\langle\ |$.

A product of two operators corresponds to the action of each operator on a vector carried out in sequence starting from the operator closest to the vector first, that is, in the case of a ket from the right and in the case of a bra from the left.

$$\Lambda\Omega |V\rangle = \Lambda(\Omega |V\rangle)$$

$$\langle V| \Lambda\Omega = (\langle V| \Lambda)\Omega$$

As you might already suspect, the product of two operators is not commutative, that is, $\Omega\Lambda$, in general, is not necessarily equal to $\Lambda\Omega$. (You can think of rotation in three dimensions about two orthogonal axes.) We can define the *commutator* of Ω and Λ as

$$[\Omega, \Lambda] \equiv \Omega\Lambda - \Lambda\Omega.$$

Some useful identities and properties of commutators are listed without proven below

- $[\Omega, \Lambda\Theta] = \Lambda[\Omega, \Theta] + [\Omega, \Lambda]\Theta$
- $[\Lambda\Omega, \Theta] = \Lambda[\Omega, \Theta] + [\Lambda, \Theta]\Omega$
- *Inverse of* Ω: $\Omega\Omega^{-1} = \Omega^{-1}\Omega = I$
- *Inverse of product* : $(\Omega\Lambda)^{-1} = \Lambda^{-1}\Omega^{-1}$

1.3.2 Matrix Representation of Operators

For an operator acting on a finite-dimensional Hilbert space $\mathbb{V}^n$, we can represent that operator in a form of a matrix. The elements of this matrix can be calculated if we know how the operator acts on the basis, that is, if

$$\Omega |i\rangle = |i'\rangle$$

then

$$\langle j|i'\rangle = \langle j|\Omega|i\rangle \equiv \Omega_{ji}.$$

The action of Ω on a state $|V\rangle$

$$\Omega |V\rangle = |V'\rangle$$

can be represented by

$$\begin{bmatrix} v_1' \\ v_2' \\ \vdots \\ v_n' \end{bmatrix} = \begin{bmatrix} \langle 1|\Omega|1\rangle & \langle 1|\Omega|2\rangle & \cdots & \langle 1|\Omega|n\rangle \\ \langle 2|\Omega|1\rangle & & & \\ \vdots & & \ddots & \vdots \\ \langle n|\Omega|1\rangle & \cdots & & \langle n|\Omega|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

where we can write v_i' as

$$v_i' = \langle i|V'\rangle = \langle i|\Omega|V\rangle = \langle i|\,\Omega\left(\sum_j v_j\,|j\rangle\right) = \sum_j v_j\,\langle i|\Omega|j\rangle = \sum_j \Omega_{ij}v_j$$

As you can see an element of the vector $|V'\rangle$ is obtained by simple matrix multiplication. Therefore, it is quite useful to represent the operator in a matrix form and to have it act as a vector represented also by a matrix.

Similarly, we can represent the product of two operators in terms of matrix multiplication, that is, we can write an element of $\Omega\Lambda$ as

$$(\Omega\Lambda)_{ij} = \langle i|\Omega\Lambda|j\rangle = \sum_k \langle i|\Omega|k\rangle\,\langle k|\Lambda|j\rangle = \sum_k \Omega_{ik}\Lambda_{kj},$$

which corresponds to how we normally multiply two matrices.

Adjoint of Operator

For an operator Ω acting on a vector $|V\rangle$,

$$\Omega\,|V\rangle = |\Omega V\rangle$$

we can define the corresponding bra of this vector $|\Omega V\rangle$ as

$$\langle \Omega V| = \langle V|\,\Omega^\dagger,$$

where $\Omega^\dagger$ called *adjoint operator* is the transpose conjugate of Ω, the same way as $\langle \alpha V| = \langle V|\,\alpha^\star$, where $\alpha^\star$ is a complex conjugate of α. Some important properties of the adjoint are listed below:

1. $\Omega_{ij}^\dagger = \Omega_{ji}^*$ (This is just a restatement of the definition of adjoint)
2. $(\Omega\Lambda)^\dagger = \Lambda^\dagger\Omega^\dagger$

1.3.3 Hermitian and Unitary Operators

Now, we will discuss what kind of an operator is physically relevant or what we call *an observable*. Remember here that the measurable values must always be real. For that purpose, we will consider only special classes of operators, which will concern us in quantum mechanics.

Definitions:

- An operator Ω is *Hermitian* if $\Omega^\dagger = \Omega$.
- An operator Ω is *anti-Hermitian* if $\Omega^\dagger = -\Omega$.

- An operator U is *unitary* if $UU^\dagger = U^\dagger U = I$, that is, $U^\dagger = U^{-1}$.

As you can see that there is an obvious analogy between the adjoint of the operator and the complex conjugate of a complex number, that is, the Hermitian operator is analogous to a purely real number while the anti-Hermitian is analogous to a purely imaginary number. We will see this analogy play out when we consider the eigenvalues of these operators. The Hermitian will play the role of the Hamiltonian in our quantum mechanics formulation since the measurable quantities, which correspond to the eigenvalues of the Hermitian operator, are real.

We will state some properties of the unitary matrix again without proving them (if you are interested in their proof, we can read Chapter 1 of Shankar).

- A unitary operator preserves the inner product between the vector they act on.
- The columns (or rows) of an $n \times n$ unitary matrix from orthogonal vectors.

Basically, we can think of the unitary operator as a "rotational operator" since it transforms a vector without changing its "length". The unitary operator will play an important role when we study time evolution of the wave function. In that case, we can think of the time evolution as a "rotation" of the eigenstates.

We can think of the unitary operator acting on the vector as mentioned above, which we can represent as

$$|V\rangle \to U\,|V\rangle \text{ and } \langle W| \to \langle W|\,U^\dagger. \tag{1.1}$$

Under this transformation of the vector, we will find that the operator Ω will change accordingly as

$$\langle W|\Omega|V\rangle \to \langle W|U^\dagger \Omega U|V\rangle\,,$$

that is, the transform of the operator Ω is expressed by

$$\Omega \to U^\dagger \Omega U, \tag{1.2}$$

which is the transformation of Ω under the unitary operator U. Therefore, we can equivalently think of the unitary operator transforming the operator Ω instead of the vector $|V\rangle$. We call the transformation of the vector $|V\rangle$ (Eq. 1.1) *an active transformation* and the transformation of the operator Ω (Eq. 1.2) *a passive transformation*. And these two points of view are equivalent.

1.3.4 Eigenvalue Problem

Now we come to the most important aspect of operators for doing calculations in quantum mechanics. To summarize in one sentence, a problem in quantum mechanics is basically to solve for eigenvalues and eigenvectors of a given Hermitian operator. Once both eigenvalues and eigenvectors of the system are known, we basically know everything about that system, that is, first we know what the possible outcomes of our measurements (Postulate III) are and secondly we know how the system evolves as a function of time (Postulate IV). These are possible due to two important properties of the eigenvalues and eigenvector of the Hermitian operator, those are,

1. The eigenvalues of a Hermitian operator are real.
2. For every Hermitian operator Ω, there exists a basis consisting of its orthogonal eigenvectors, where Ω is diagonal in this basis and its diagonal elements are its eigenvalues.

Again, we are not going to prove these statements. If you are interested, you can read about the proofs in Shankar. We will instead discuss their consequences on quantum mechanics.

The first point affirms that the measurable quantities, which are the eigenvalues of the operator Ω, have to be real. The second point indicates that a vector in this Hilbert space can be written as an expansion of the eigenvectors, that is, the basis formed by the eigenvectors spans the vector space. Later, we will learn that we can obtain the time evolution of the eigenvectors from the Schrödinger equation (Postulate IV). Hence, we can get the time evolution of any vectors in this vector space.

In an eigenvalue problem, let Ω be a Hermitian operator, the equation that we have to solve is:

$$\Omega \left|\omega_i\right\rangle = \omega_i \left|\omega_i\right\rangle ,$$

where we have to solve for both eigenvectors $|\omega_i\rangle$ and eigenvalues ω_i. In the case of a finite Hilbert space, this problem is equivalent to solving for eigenvalues and eigenvectors of a matrix Ω_{ij}, which I am certain that you are familiar with. In this case, a vector $|\psi\rangle$ in the Hilbert space can be written in terms of $|\omega_i\rangle$ as

$$\left|\psi\right\rangle = \sum_i \left\langle\omega_i|\psi\right\rangle \left|\omega_i\right\rangle .$$

Note that we use the eigenvalues to label the corresponding eigenstates.

Degenerate States

In solving the eigenvalue problem, you will sometimes encounter the case where the two distinct states have the same eigenvalue, for example

$$\Omega \left|\omega_i, \lambda\right\rangle = \omega_i \left|\omega_i, \lambda\right\rangle ,$$

where $i = 1, 2, 3, \ldots, n$. We call $|\omega_i, \lambda\rangle$ the degenerate states. We will use two indices to label these states. We choose ω_i so that $\langle\omega_i, \lambda|\omega_i, \lambda'\rangle = \delta_{\lambda\lambda'}$, that is, they are orthogonal. Note that there can be an infinite number of degenerate states. And these degenerate states or vectors form a *subspace*. These degenerate states are the result of the symmetry of the Hamiltonian, and if this symmetry is broken, then the degeneracy will also be lifted. Furthermore, we can find an operator Λ where $[\Lambda, \Omega] = 0$. which gives

$$\Omega\Lambda \left|\omega_i\right\rangle = \Lambda\Omega \left|\omega_i\right\rangle = \omega_i \Lambda \left|\omega_i\right\rangle ,$$

which means $\Lambda |\omega_i\rangle$ is the eigenstate of Ω with the eigenvalue ω_i. But then we can solve the eigenvalue problem for Λ:

$$\Lambda \left|\omega_i\right\rangle = \lambda \left|\omega_i\right\rangle ,$$

i.e. we can index the state $|\omega_i\rangle$ using λ, that is, $|\omega_i, \lambda\rangle$ and with respect to Λ this state becomes non-degenerate since it now has different eigenvalue. In general, if states are degenerate with respect to Ω, we can always find a set of Λ_i, which commutes with Ω and with one another, such that those states are non-degenerate and we can label those non-degenerate as $|\omega_i, \lambda_1, \lambda_2, \ldots, \lambda_n\rangle$, where λ_i denotes the eigenvalue of the operator Λ_i.