Lecture 2

Eslam Khalaf

September 8, 2023

1 Quantum mechanics in finite-dimensional Hilbert spaces

Last lecture, we discussed the Stern-Gerlach experiment which demonstrated the need to use a complex linear vector space to describe quantum state. This means that the main formalism needed to describe quantum mechanics is linear algebra. In this lecture, we will discuss the correspondence between the main concepts we encounter in linear algebra and their physical interpretation in the quantum theory. We will use the bra-ket notation developed by Dirac which is particularly suited for calculations in the quantum theory.

In the SG experiment, we needed two basis vectors to described the outcome of any given experiment which means that the vector space was two-dimensional. In the following, we will focus on the case where the vector space is finite-dimensional which means that we need a finite number of basis vectors to describe any given state. The more subtle case of infinite dimensional space will be discussed next week.

1.1 Bras and Kets

The basic objects in a linear vector space are vectors. In the Dirac notation, vectors are called kets and denoted by $|u\rangle$. Quantum states correspond to *rays* in the vector space which means that the two kets $|u\rangle$ and $\alpha|u\rangle$, with α being a non-zero complex number $\alpha! = 0$, describe the same physical state. This means that the Hilbert space description of quantum mechanics has a redundancy i.e. a given ket defines a unique state but a given state can be described by several kets. Such situations are quite common in physics with the most prominent example being electromagnetism where the gauge potential (consisting of the scalar and vector potentials) uniquely determines the physical electric field and the magnetic flux through any region, but a physical electric and magnetic field configuration corresponds to many gauge potentials. This is called a gauge ambiguity and it leads to some subtle effects that we will discuss later. For $\alpha = 0$, $\alpha|v\rangle$ for any v is called the null-vector. We can form linear superposition of states as $\alpha|u\rangle + \beta|v\rangle$ where α and β are complex scalars. An example from the SG setup was how we wrote the beam polarized in the $+x$ direction as a superposition of the $+z$ and $-z$ beams: $|+x\rangle = \frac{1}{\sqrt{2}}(|+z\rangle + |-z\rangle)$.

Vectors in a Hilbert space are equipped with an inner product. The inner product is the analog of the scalar product of two vectors which gives the components of a vector u in the direction of another vector v scaled by the length of v. The inner product satisfies the following properties

1. Positive definiteness: $\langle u|u\rangle > 0$ unless $|u\rangle$ is the null ket.
2. Linearity in the second argument: $\langle u|\alpha v_1 + \beta v_2\rangle = \alpha\langle u|v_1\rangle + \beta\langle u|v_2\rangle$
3. Conjugate symmetry $\langle u|v\rangle = \langle v|u\rangle^*$.

The first property is the obvious generalization of the fact that the length of a non-zero real vector is positive. The positive number $\|u\| = \sqrt{\langle u|u\rangle}$ is known as the norm of $|u\rangle$. For any ket $|u\rangle$, we can define a normalized ket $|\tilde{u}\rangle$:

$$|\tilde{u}\rangle = \frac{|u\rangle}{\|u\|} \tag{1}$$

Since the physical state does not change under multiplication by a scalar, we can always define our states to be normalized (notice that the phase ambiguity still remains). The second property has also an obvious analog for real vectors where the projection of a sum of vectors is the sum of projections and the projection of a scaled vector is the scaled projection. The last property is the analog of the fact that scalar product is symmetric. For complex inner products, we need conjugate symmetry instead to guarantees that the norm

$$\|u + v\|^2 = \langle u + v|u + v\rangle = \|u\|^2 + \|v\|^2 + \langle u|v\rangle + \langle v|u\rangle \tag{2}$$

is real. [1]. The conjugate symmetry also implies $\langle \alpha u_1 + \beta u_2|v\rangle = \alpha^*\langle u_1|v\rangle + \beta^*\langle u_2|v\rangle$.

A vector in an N-dimensional vector space can always be expanded in a set of basis which we denote by $|a\rangle$, $a = 1, \ldots, N$. For example, in the SG device, we can describe any beam as a linear superposition of $|+z\rangle$ and $|-z\rangle$. For a general ket $|u\rangle$, this means we can write

$$|u\rangle = \sum_a u_a|a\rangle \tag{3}$$

We can always choose these basis vectors to be orthonormal which means that

$$\langle a|a'\rangle = \begin{cases} 1 & a = a' \\ 0 & a \neq a' \end{cases} \tag{4}$$

If we now take the inner product of the vector $|u\rangle$ with the basis vectors $|a\rangle$, we find that the components of $|u\rangle$ in the $|a\rangle$ basis are $u_a = \langle a|u\rangle$.

It is very useful to think about the left part of the inner product $\langle v|u\rangle$ as an independent object called a bra $\langle v|$ that acts on the ket vector $|u\rangle$ to give us the inner product. In other words, the bra (dual vector) defines a linear map from the kets (vectors) to the complex numbers $f_v(|u\rangle) = \langle v|u\rangle$. To satisfy the property $\langle \alpha v|u\rangle = \alpha^*\langle v|u\rangle$, the bras satisfy $\langle \alpha v| = \alpha^*\langle v|$. If we use the basis expansion (3), we see that $f_v(|u\rangle) = \sum_a u_a f_v(|a\rangle) = \sum_a u_a\langle v|a\rangle = \sum_a u_a v_a^*$. This gives us the analog for the scalar product formula in terms of components for complex vectors and implies that the components of a bra in a given basis are given by the complex conjugate of the components of the corresponding ket.

1.2 Operators

An operator in a vector space maps vectors to vectors. In the bra-ket notation, operators act on a ket from the left $A|u\rangle$ and on a bra from the right $\langle u|A$ (always next to the vertical line). We will sometimes use the notation $|Au\rangle := A|u\rangle$. Under the duality map, the ket $|Au\rangle$ gets mapped to the bra $\langle Au|$. We now write $\langle Au|$ as the action of some operator acting on $\langle u|$ which we denote by $A^\dagger$, i.e. $\langle Au| = \langle u|A^\dagger$ where $A^\dagger$ is called the adjoint of A. From the definition of a dual vector $f_{|Av\rangle}(|u\rangle) = \langle Av|u\rangle = \langle v|A^\dagger|u\rangle = \langle v|A^\dagger u\rangle$.

To preserve the linear structure of the vector space, we consider linear operators satisfying

$$A(\alpha|u\rangle + \beta|v\rangle) = \alpha A|u\rangle + \beta A|v\rangle \tag{5}$$

However, it turns out that in the quantum theory, since we do not care about the overall phase of the wave-function, we can also have *anti-linear* operators satisfying

$$A(\alpha|u\rangle + \beta|v\rangle) = \alpha^* A|u\rangle + \beta^* A|v\rangle \tag{6}$$

We will discuss such operators later in the course. For the time-being, we are going to restrict ourselves to linear operators.

You can think of the example of an SG device when thinking about quantum operators. For example, SG$z+$ can be represented by an operator that selects the $|+z\rangle$ component of an incoming beam. Using

[1] It suffices to assume $\mathrm{Re}\langle u|v\rangle = \mathrm{Re}\langle v|u\rangle$ which together with positive definiteness implies $\langle u|v\rangle = \langle v|u\rangle^*$ due to Eq. 2

this example, we can deduce the properties of operators in quantum mechanics. For example, two SG devices are considered identical, if they give exactly the same output given the same input for any input. The corresponding statement is that two operators A and B are equal if $A|u\rangle = B|u\rangle$ for any $|u\rangle$. Next, consider what happens when we cascade SG devices. Let us take an input beam described by a ket $|u\rangle$ and have it go through a cascade of two SG devices described by operators B and A. The output of B is $B|u\rangle$ is the input of A. Thus, the final output is $AB|u\rangle$. This means that we can identify the operator AB as the operator representation of the cascaded device. Let us try to understand its properties. First, consider a cascade of three SG devices C, B, and A. We can first think of the output of the C device, being fed into the cascade of B and A, which is described by the operator AB, or we can first combine C and B, described by BC, then consider this as an input to A. Since these are two descriptions for the same physical setup, the operator multiplication has to be associative

$$(AB)C = A(BC) \tag{7}$$

Finally, let us consider the setup from last lecture where we had SG$z+$ (let's call it C), followed by SG$x+$ (let's call it B), followed by SG$z-$ (let's call it A). We found that this configuration generally has a non-zero output since the $+x$ beams coming from the middle SG has forgotted it was originally in the $|+z\rangle$ state. On the other hand, if we do the cascade SG$z+$, SG$z-$, SG$x+$ (described by the operator BAC), we find that the output is zero. From this we conclude that $ABC \neq BAC$ which implies that, in general, for any two operators

$$AB \neq BA \tag{8}$$

This property is responsible for many of the counter-intuitive aspects of quantum mechanics.

The power of the bra-ket notation comes from the fact that we can always place a bra to the left of a ket to obtain a meaningful object. For example, we said that we can understand the bra $\langle u|$ as a linear map that acts on a ket $|v\rangle$ to give the complex number $\langle u|v\rangle$. Using the same reasoning, we can define the object $|u\rangle\langle v|$ which acts on a ket $|w\rangle$ to give us $|u\rangle\langle v|w\rangle$. Since $\langle v|w\rangle$ is a complex number, the whole object $|u\rangle\langle v|w\rangle$ is a ket. Thus $|u\rangle\langle v|$ acts as an operator. It is easy to see from the linearity of the inner product that this will be a linear operator.

Let us now define the identity operator $\mathbb{1}$ as the operator that leaves every vector invariant: $\mathbb{1}|u\rangle = |u\rangle$ for any $|u\rangle$. Given an arbitrary ket $|u\rangle$, we can write the basis expansion (3) with $u_a = \langle a|u\rangle$. This means that

$$|u\rangle = \sum_a |a\rangle\langle a|u\rangle = \left(\sum_a |a\rangle\langle a|\right)|u\rangle \tag{9}$$

Since this identity holds for every $|u\rangle$, it implies

$$\sum_a |a\rangle\langle a| = \mathbb{1} \tag{10}$$

This relation is called resolution of unity and it is one of the most important relations in quantum mechanics. It tells us that if we take the projection on a vector relative to each basis vector, multiply it by that basis vector, then add all these components, we can reconstruct the full vector.

To demonstrate the usefulness of this identity, let us consider an operator A. Since the identity operator does nothing, we can insert it anywhere in an expression

$$A = \mathbb{1}A\mathbb{1} = \left(\sum_a |a\rangle\langle a|\right) A \left(\sum_b |b\rangle\langle b|\right) = \sum_{a,b} |a\rangle\langle a|A|b\rangle\langle b| = \sum_{a,b} A_{ab}|a\rangle\langle b| \tag{11}$$

where $A_{ab} = \langle a|A|b\rangle$ are the matrix elements of the operator A which specify the decomposition of $A|a\rangle$ in the $|a\rangle$ basis

$$A|a\rangle = \sum_b A_{ba}|b\rangle \tag{12}$$

For an N-dimensional space, A_{ab} describes an $N \times N$ matrix. The operator product AB can be written as

$$AB = \sum_{a,b,c} |a\rangle\langle a|A|b\rangle\langle b|B|c\rangle\langle c| = \sum_{a,c} |a\rangle\langle c| \sum_{b} A_{ab}B_{bc} = \sum_{a,c} (AB)_{ac}|a\rangle\langle c| \tag{13}$$

We see that the matrix elements of the operator AB are given by the matrix product of the matrix elements of the operators A and B. This means that, for the case of finite-dimensional Hilbert spaces, we can always think of the operators as matrices. However, it is important to distinguish the abstract operator A which describes a physical device or observable from its matrix elements A_{ab} which require us to specify a basis and is thus basis-dependent.

The identification of operators in the quantum theory with matrices (for finite-dimensional Hilbert space) allows us to deduce a lot of their properties from what we know about matrices. For instance, we can define eigenkets of an operator as $A|u\rangle = \lambda|u\rangle$ where λ is a complex number called the eigenvalue. We also see that we can identify the matrix elements of the hermitian adjoint of an operator $A^\dagger$ with the hermitian matrix adjoint obtained by transposing and complex conjugating the matrix elements $(A^\dagger)_{ab} = A^*_{ba}$.

There are three classes of operators that play a special role in quantum mechanics

1. Hermitian operators ($A = A^\dagger$): these operators have the very important property that their eigenvalues are real. To see this consider an eigenket $|u\rangle$ with eigenvalue λ:

$$A|u\rangle = \lambda|u\rangle \tag{14}$$

   Taking the dual of both sides, we get

$$\langle u|A^\dagger = \lambda^*\langle u| \tag{15}$$

   If we now take the inner product of (14) with $\langle u|$ and of (15) with $|u\rangle$, we get

$$\lambda\langle u|u\rangle = \langle u|A|u\rangle = \langle u|A^\dagger|u\rangle = \lambda^*\langle u|u\rangle \tag{16}$$

   which implies $\lambda = \lambda^*$. Another important property of Hermitian operators is that their eigenkets provide a complete basis of the vector space that can be chosen to be orthonormal. This implies that an operator is diagonal when expressed in its own eigenbasis. If we denote the eigenkets and eigenvalues of an operator A by $|a\rangle$ and λ_a [2]

$$A = \sum_a \lambda_a |a\rangle\langle a| \tag{17}$$

   Notice the absence of off-diagonal terms $|a\rangle\langle b|$ with $a \neq b$.

2. Unitary operators ($UU^\dagger = U^\dagger U = \mathbb{1}$): Unitary operators correspond to norm-preserving maps since the norm of $|v\rangle = U|u\rangle$ is $\|v\| = \sqrt{\langle v|v\rangle} = \sqrt{\langle u|U^\dagger U|u\rangle} = \sqrt{\langle v|v\rangle} = \|u\|$. Thus, they map normalized kets to normalized kets. They usually describe basis transformations in the quantum theory.

3. Projection operators ($P^2 = P = P^\dagger$): these are Hermitian operators which square to themselves. For example, the SG device which filters one polariztion, e.g. SG$z+$, is described by a projector since passing its output by an identical device SG$z+$ does not change the result. Recall that cascading SG devices corresponds to operator multiplication, we see that i.e. $P^2 = P$.

[2] Here, we are assuming all eigenvalues are distinct

1.3 Observables

In the quantum theory, physical observables are represented by hermitian operators and the possible values of a given observable correspond to the eigenvalues of this operator. Since the eigenvalues of a Hermitian operator are real, this guarantees that the allowed values of a physical observable are real. Since the eigenkets of a hermitian operator provide a complete basis, we can expand any ket in terms of the eigenbasis of a given hermitian operator.

Let us take our SGz device as an example. This device measures the magnetic moment in the z direction. It has two possible outcomes $\mu_z = \pm\mu_0$. As an aside, we have so far been discussing the magnetic moment without specifying its value μ_0. To understand this value, we invoke a classical analogy where we consider a charged spinning classical particle. For such a particle, the magnetic moment is proportional to the spin angular momentum $\boldsymbol{S}$, $\boldsymbol{\mu} = \gamma \boldsymbol{S}$, where the proportionality constant is called gyromagnetic ratio. If we assign a value for the spin angular momentum corresponding to the measured magnetic moment in the SG device, we find it corresponds to $S_z = \pm\hbar/2$ where $\hbar$ is the Planck's constant. From now onwards, we will refer to the different results of the SG device in terms of the values of spin rather than magnetic moment. This means that this device should be described by a Hermitian operator, which we call S_z acting on a two dimensional Hilbert space whose eigenkets are $|\pm z\rangle$ with corresponding eigenvalues $\pm\hbar/2$. Using Eq. 17, we can write

$$S_z = \frac{\hbar}{2}(|+z\rangle\langle +z| - |-z\rangle\langle -z|) \tag{18}$$

If we use a vector notation where

$$|+z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{19}$$

we see that S_z is the diagonal matrix

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{20}$$

The operators describing S_x has the same form in the $|\pm x\rangle$ basis

$$S_x = \frac{\hbar}{2}(|+x\rangle\langle +x| - |-x\rangle\langle -x|) \tag{21}$$

To write the same operator in the $|\pm z\rangle$ basis, we use

$$|\pm x\rangle = \frac{1}{\sqrt{2}}(|+z\rangle \pm |-z\rangle) \tag{22}$$

Substituting in (21), we find

$$S_x = \frac{\hbar}{2}(|+z\rangle\langle -z| + |-z\rangle\langle +z|) = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{23}$$

Performing a similar calculation for S_y, we find

$$S_y = \frac{i\hbar}{2}(-|+z\rangle\langle -z| + |-z\rangle\langle +z|) = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{24}$$

Thus, the three components have the matrix representation $S_i = \frac{\hbar}{2}\sigma_i$ where σ_i are called the Pauli matrices and play a central role in the theory of spin.