Lecture 0.2: State Vectors in Quantum Mechanics

Boyi Zhang

January 13, 2020

You will most likely have been introduced to state vectors in the past, but we will go through a more thorough treatment of them here because without an understanding of the basic manipulations on state vectors here, later on concepts in this course will be more difficult to understand. The concept of state vectors and operators on state vectors are used in Hartree-Fock and the second-quantization formalism for correlated methods (the majority of our class!) Without a concrete understanding of state vectors, you may be able to do the math, but you might not understand what you are doing.

1 Space vectors in the Cartesian coordinate

1.1 Defining a space vector

Let's start by reviewing the concept of vectors. In a given coordinate representation, we can describe the position of a particle with a set of coordinates. For example, in the Cartesian coordinate representation, we have x, y and z. We can do mathematical manipulations explicitly with these coordinates, but it is much easier if we can represent the set of these coordinates together. It is with this motivation that the concept of vectors was invented. We can represent the position of a particle a by a vector, written in this way:

$$\mathbf{A} = (A_x, A_y, A_z)$$

(In matrix notation, we can write $\mathbf{A} = [A_x \ A_y \ A_z]$) The boldface denotes the vector, while the regular text denotes a number. Since this vector is representing a particle in the Cartesian coordinate space, we refer to it as a space vector. Along with the idea of a vector, we also learn the idea of unit vectors, or base vectors. In the 3D Cartesian coordinate space, there are 3 unit vectors, corresponding to the 3 directions:

$$\mathbf{e_x} = (1, 0, 0)$$
$$\mathbf{e_y} = (0, 1, 0)$$
$$\mathbf{e_z} = (0, 0, 1)$$

In a physics class, these unit vectors are referred to as $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.

1.2 The dot product

Now, let us consider the concept of a dot product, or an inner product. Qualitatively, a dot product measures how much two vectors lie in the same direction. Alternatively, you can think about it as the component of one vector which is in the direction of another vector. The dot product gives a scalar value. Between space vectors $\mathbf{A}$ and $\mathbf{B}$, where $\mathbf{A} = (A_x, A_y, A_z)$ and $\mathbf{B} = (B_x, B_y, B_z)$, we write the dot product as $\mathbf{A} \cdot \mathbf{B}$. Mathematically, we can define this dot product as:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = A_x B_x + A_y B_y + A_z B_z \tag{1}$$

or written as a summation:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \sum_{all\ i} A_i B_i \tag{2}$$

Note the commutative property of the dot product. The dot product of base vectors have the relation $\mathbf{e_i} \cdot \mathbf{e_j} = \delta_{ij}$. That is to say, they are orthogonal, and no part of one unit vector lies in the same direction as another unit vector.

There is another way we can represent the dot product using our base vectors:

$$\mathbf{B} \cdot \mathbf{A} = \sum_{all\ i} (\mathbf{B} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{A}) \tag{3}$$

It's a little strange, but we can see it is equal to Equations 1 and 2.

Now, let us remove $\mathbf{B}$ from both sides of Equation 3. Why can we do this? Consider Newton's equation of motion, $\mathbf{F} = m\mathbf{a}$. We work with this equation in physics, but the equation does not give a number. If we wanted numbers, say, the x component of force F, we obtain it by applying the base vector $\mathbf{e}_x$ to both sides of the equation:

$$\mathbf{e}_x \cdot \mathbf{F} = \mathbf{e}_x \cdot (m\mathbf{a})$$

to get $F_x = ma_x$, which is an equation that gives us numbers. Qualitatively, we are finding the component of the force that lies in the x direction. If we wanted to find the component of the force that lies in the y direction, we would simply apply the vector $\mathbf{e}_y$. Thus we see that by itself the equation $\mathbf{F} = m\mathbf{a}$ is only an abstraction. If we want numbers, we need to apply another vector $\mathbf{C}$:

$$\mathbf{C} \cdot \mathbf{F} = \mathbf{C} \cdot (m\mathbf{a})$$

We don't write in this way because since $\mathbf{F} = m\mathbf{a}$ is true for *any* vector $\mathbf{C}$, then it is redundant to write $\mathbf{C}$ explicitly. We can use the abstract vector equation $\mathbf{F} = m\mathbf{a}$, with the understanding that if we want numbers we take its dot product with another vector.

1.3 A vector is a linear combination of base vectors

Since Equation 3 is true *for all vectors* $\mathbf{B}$, we can leave it out. Our remaining equation becomes:

$$\mathbf{A} = \sum_{all\ i} \mathbf{e}_i(\mathbf{e}_i \cdot \mathbf{A}) \tag{4}$$

Remember, Equation 4 is no longer giving a number; it gives a vector. It is an abstract equation. If we want numbers, we take its dot product with another vector. Another equivalent way to write it is:

$$\mathbf{A} = \sum_{all\ i} A_i\mathbf{e}_i = A_x\mathbf{e}_x + A_y\mathbf{e}_y + A_z\mathbf{e}_z$$

This alternative way of writing leads to an important conclusion. *Any* vector $\mathbf{A}$ in the Cartesian coordinate space can be represented as a linear combination of 3 "base vectors" $\mathbf{e_i}$. The 3 coefficient of the base vectors $(\mathbf{e}_i \cdot \mathbf{A}$ or $A_i)$ are the components of the vector $\mathbf{A}$ in the direction of each base vectors. If you have the 3 base vectors and the coefficients, you know everything about vector $\mathbf{A}$.

2 State vectors in quantum mechanics

Now that we have reviewed the concept of a space vector, I want to introduce a type of vector that we use in quantum mechanics, called a state vector. As you will see, the abstractions we use to define space vectors are not so different from the abstractions we use to define state vectors.

2.1 Defining a state vector

The notation we will use for defining a state vector ϕ is $|\phi\rangle$. Just as $\mathbf{A} = [A_x\ A_y\ A_z]$ is an abstract entity made up to describe the position of a particle in space, $|\phi\rangle$ is an abstract entity made up to represent a quantum mechanical state, hence why it is called a state vector. Like the space vector, it follows its own set of rules. Just like how space vectors have a set of base vectors $(\mathbf{e}_x, \mathbf{e}_y,$ and $\mathbf{e}_z)$ the state vectors have base states. We represent these base states as $|i\rangle$.

2.2 The inner product

Like the space vectors, we can define an inner product of two states which gives a scalar value. For states χ and ϕ, the inner product is written as $\langle \chi | \phi \rangle$. What qualitative understanding can we give to the inner product of two quantum states? Remember that in the framework of quantum mechanics, any event is described by an amplitude, whose square of the absolute value gives the probability of the event at a particular time and space (See Lecture 0.1 notes). We can thinking about the inner product $\langle \chi | \phi \rangle$ as the amplitude to start in state ϕ and end up in state χ. Note that we are reading from right to left. Mathematically, the amplitude is just a complex number.

As with the unit vectors $\mathbf{e}_i$, the base states $|i\rangle$ are orthogonal to each other, and have the relationship:

$$\langle i | j \rangle = \delta_{ij}$$

Just as we wrote $\mathbf{B} \cdot \mathbf{A}$ as $\sum (\mathbf{B} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{A})$, we can represent the inner product of state vectors as:

$$\langle \chi | \phi \rangle = \sum_{all\ i} \langle \chi | i \rangle \langle i | \phi \rangle \tag{5}$$

How do we interpret Equation 5? Let's use our qualitative definition of the inner product of state vectors. The amplitude to start in state ϕ and end up in state χ can be written as the sum over a complete set of base states of the amplitude to go from ϕ into one of the base states, multiplied by the amplitude to go from that base state into state χ. Equation 5 returns a complex number.

The notation of state vectors is also known as Dirac bra–ket notation. $|\phi\rangle$ is referred to as the "ket", while $\langle \chi |$ is referred to as the "bra".

2.3 A state vector is a linear combination of base states

Just as we could stop writing $\mathbf{B}$ from Equation 3 because it is true for all vectors $\mathbf{B}$, Equation 5 is true for all states χ, and so we can leave it out. This gives us an equation for the state vector itself:

$$|\phi\rangle = \sum_{all\ i} |i\rangle \langle i | \phi \rangle \tag{6}$$

We emphasize that this is a abstract equation, and does not give a number. If we want a number, we need to finish it out, or project on the left by another state. $\langle i | \phi \rangle$ is the amplitude to go from state ϕ into one of the base states i. It is just a complex number, which we can call C_i. $|\phi\rangle$ now becomes:

$$|\phi\rangle = \sum_{all\ i} C_i |i\rangle$$

Written this way, it is clear to see that any state vector can be represented as a linear combination of base vectors i, with the coefficients being the the amplitude to go from state ϕ into the base states. We can also think about it as representing ϕ as a superposition of base states. [1] This is a very similar way to thinking about space vectors.

Going back to Equation 6, one can ask, if it is true for all ϕ, why can we not just leave it out as well? Indeed, we can do that, and are left with:

$$| = \sum_{all\ i} |i\rangle \langle i|$$

This is known as the resolution of the identity. This does not seem very useful, but it helps us understand the notation a little better. The right side of the equation can be inserted in place of a bar at will, since it is the equivalent of a factor of 1. In fact, this trick is used very often in the mathematical manipulations of quantum mechanics. Additionally, it reminds us that we can put any 2 states on the left and right, and get back Equation 5.

---

[1] We see that for the state vectors to be represented this way, the base states have to form a complete set with the property that any state vector can be represented by a linear combination of base vectors.

3   What base states?

We have established that any quantum state can be represented as a superposition of base states, or to put it another way, a linear combination of base vectors with suitable coefficients. The remaining question is, what do we mean by a base state? Stated simply, base states are used to describe the "space" of what we want to represent. If we are representing a 3D space, we can use different coordinate systems. The Cartesian coordinate space is one. Spherical coordinates are another example. In each coordinate representation, there is a corresponding set of base vectors to describe that representation. For fundamental particles, like the electron, we want to work in a space that describes the nature of that particle. A particular set of base states is one representation of that space. In order to find a set of base states, we need to understand the "space" we are working with, and what the base state means physically.

For an electron, for example, we say that in the nonrelativistic view, a complete description of the electron requires a description of 1. spin and 2. momentum. Thus, the set of base states we choose must describe both the spin and momentum, where one particular base state corresponds to a particular spin and momentum. The combination of the complete set of base states is what describes an electron. The coefficients dictate how much of each base state goes into the final description of the actual electron. When we start talking about multiple electrons, each with a set of base states, it is not clear how we correctly put them together. The main problem in the study of fundamental particles is to discover the correct representation for the description of nature. The main problem in quantum chemistry is to find the correct representation for multi-electron systems.

3.1   Differences between space and state vectors

There are 2 differences between space and state vectors that I want to point out at this time:

- $\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{B}$, but $\langle \chi | \phi \rangle = \langle \phi | \chi \rangle^*$. Because the state vectors involve complex numbers, the order of operations matter.

- Space vectors have 3 orthogonal base vectors. The base states i must range over the complete set applicable to a particular problem. This means that there could be 1, 2, or an infinite number of base states.

4   Operators

4.1   What is an operator?

You are probably familiar with the concept of an operator, written this way: $|\psi\rangle = A |\phi\rangle$. Let's talk about the concept of operators from a physical motivation. In the state vector formalism, we represent the amplitude of a particle in state ϕ, going through an apparatus A, and arriving in state χ as:

$$\langle \chi | A | \phi \rangle$$

It is not clear to us at this point exactly how to interpret the quantity $A |\phi\rangle$. Can we represent $A |\phi\rangle$ as another state $|\psi\rangle$? Such a state ψ would require that the amplitude to get from ψ to χ ($\langle \chi | \psi \rangle$) is identical and the same everywhere as the amplitude $\langle \chi | A | \phi \rangle$. Mathematically,

$$\langle \chi | \psi \rangle = \langle \chi | A | \phi \rangle$$

The first step we take is to use the resolution of the identity:

$$\langle \chi | \psi \rangle = \sum_i \langle \chi | i \rangle \langle i | \psi \rangle$$

4

We see that if $\langle i|\psi\rangle = \langle i|A|\phi\rangle$, then

$$\langle\chi|\psi\rangle = \sum_i \langle\chi|i\rangle\langle i|\psi\rangle$$
$$= \sum_i \langle\chi|i\rangle\langle i|A|\phi\rangle$$
$$= \langle\chi|A|\phi\rangle$$

It may seem like we are going in circles, but through these manipulations we have defined a key term, $\langle i|\psi\rangle$. Remember that $\langle i|\psi\rangle$ is the coefficient to relate ψ to base state i. If you have a set of base states i and you know the coefficients to relate ψ to i, then that's all you need to uniquely define ψ.

Since $\langle i|\psi\rangle = \langle i|A|\phi\rangle$ is true for all states i, we can just write:

$$|\psi\rangle = A\,|\phi\rangle$$

We see that the entity A is not an amplitude, nor a vector. It is something that we call an operator. It operates on a quantum mechanical state, to produce a new state. From now on, we will denote operators with a hat: $\hat{A}$.

Let's use the resolution of the identity to further elucidate some things about $\hat{A}$. Remember, if you don't know what to make of a state vector equation, you can always use the resolution of the identity to put the equation in terms of amplitudes to give some qualitative meaning.

$$\sum_i \langle i|\psi\rangle = \sum_i \langle i|\hat{A}|\phi\rangle$$
$$= \sum_i \sum_j \langle i|\hat{A}|j\rangle\langle j|\phi\rangle$$

Now we can see that each amplitude from ψ to base state i is just a linear superposition of the amplitudes from state ϕ to base states j. $\langle i|\hat{A}|j\rangle$ acts as a coefficient to tell you how much of $\langle j|\phi\rangle$ go into each sum. The set of $\langle i|\hat{A}|j\rangle$ for all i and j, which we can also call A_{ij} completely (numerically) describes the operator $\hat{A}$ in terms of any set of base vectors. This set of A_{ij} can be collected into a matrix form.

The value of this exercise helps us see how, if we had absolutely no knowledge about operators, we go from experimental measurables to mathematical expressions. Specifically, it tells us if there is something we can measure in a lab through an apparatus, we can express the values we can obtain from that apparatus as a matrix of numbers. This is hugely meaningful.

One last note about the operator $\hat{A}$. It can be shown that

$$\langle\chi|\hat{A}|\phi\rangle* = \langle\phi|\hat{A}^\dagger|\chi\rangle$$

where the adjoint of the matrix A ($A_{ij}^\dagger$) is its conjugate transpose $[(A_{ji})^*]$. For operators we encounter in quantum chemistry, it turns out that they are self-adjoint ($\hat{A}^\dagger = \hat{A}$). We call self-adjoint operators Hermitian operators.

5 Arriving at the wavefunction formulation

The operator way of writing (e.g., $|\psi\rangle = A\,|\phi\rangle$) avoids making any particular choice; there is no reference to a set of base states. It is completely general. When we want definite numbers, we then make a choice of base states, and form the matrix A_{ij}. (It turns out if we know the matrix for one set of base states, we can transform the matrix into another representation of base states).

5.1 The relationship between $|\psi\rangle$ and $\psi(x)$

Let's see how the ideas we have discussed translates when we choose a particular base state. For simplicity, let's use the 1D x-representation. Remember that a state of a particle at a particular space and time can be completely described by a complex number ϕ which is called the probability amplitude. The square of the absolute value of ϕ gives the probability of finding that particle at that space and time. ϕ varies with space and time, and the function that gives ϕ over a certain space/time is what we call the wavefunction. For a 1D x-representation, our wavefunction can be written as $\psi(x)$. Qualitatively, it gives the amplitude of the particle at every x.

How does this relate to our notion of the state vector? The particle that we are talking about can be described by the state vector $|\psi\rangle$. $|\psi\rangle$ is completely general, and by itself is an abstract concept. Let's consider the term $\langle\psi|\psi\rangle$. This term is general, and tells us nothing about a specific representation. We can apply one however. In this case, we apply the 1D x-representation. Our base states $|i\rangle$ can be represented by $|x\rangle$. We can use the resolution of the identity:

$$\langle\psi|\psi\rangle = \sum_{all\ x} \langle\psi|x\rangle\langle x|\psi\rangle$$

Space is continuous, and since there are infinite values of x, the sum over infinite values of x can be rewritten as an integral:

$$\langle\psi|\psi\rangle = \int_x \langle\psi|x\rangle\langle x|\psi\rangle\,dx$$

By our previous definition of inner products, $\langle x|\psi\rangle$ means the amplitude to go from state ψ to all the base state x. To put it another way, it is giving the amplitude of the particle at every x. We see that this is now equivalent to our wavefunction to $\psi(x)$. And since $\langle\psi|x\rangle = \langle x|\psi\rangle^*$, then $\langle\psi|x\rangle = \psi^*(x)$. We can now write:

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

The expression above is probably how you were introduced to the state vector notation. However, it is crucial that you understand that the intermediate step. $\langle\psi|\psi\rangle$ by itself says nothing about a specific representation. It is only when we apply the resolution of the identity in a specific basis that we can make the connection to $\psi(x)$. (Note, $\langle\psi|\psi\rangle$ can be interpreted as a probability, since it is the square of the absolute value of the probability amplitude $\langle x|\psi\rangle$. We can set this probability to be one by normalizing the wavefunction).

5.2 Average energies

It can be shown (Appendix II) that the expectation value of a operator A corresponding to a physical observable can be expressed as:

$$\langle A\rangle_{av} = \langle\psi|\hat{A}|\psi\rangle \tag{7}$$

Remember, Equation 7 is completely general. There is no reference to a base representation. Let's use this equation along with the ideas of section 5.1 to derive an algebraic expression for the average energy in the 1D x-representation. To get the average expectation value for the energy, we use the Hamiltonian operator:

$$\langle E\rangle_{av} = \langle\psi|\hat{H}|\psi\rangle \tag{8}$$

Our base states $|i\rangle$ can be represented by $|x\rangle$. We can use the resolution of the identity applied twice:

$$\langle E\rangle_{av} = \int_x \langle\psi|x\rangle\langle x|\hat{H}|\psi\rangle\,dx \tag{9}$$

$$= \int_x\int_{x'} \langle\psi|x\rangle\langle x|\hat{H}|x'\rangle\langle x'|\psi\rangle\,dxdx' \tag{10}$$

$$= \int_x\int_{x'} \psi^*(x)\hat{\mathscr{H}}(x,x')\psi(x')dxdx' \tag{11}$$

To make the jump to the final algebraic expression, we recall that in nature, the amplitude for the electron to go from x to x' is zero except for points x' very close to x. Therefore, the righthand sight of the Equation 11 can be expressed completely assuming evaluation at positions of x:

$$\langle E \rangle_{av} = \int_x \psi^*(x) \hat{\mathscr{H}}(x) \psi(x) dx \tag{12}$$

Note the difference between $\hat{H}$ and $\hat{\mathscr{H}}(x, x')$. The quantum-mechanical $\hat{H}$ is a operator that operates on a state vector to give a new state vector. The algebraic operator $\hat{\mathscr{H}}(x)$ operates on a function of position $\psi(x) = \langle x | psi \rangle$ to give a new function of position. When you work with the x-representation, the difference between the algebraic terms and state operator notation appears similar. Indeed, most introductory quantum chemistry classes introduces Equations 8 and 12 as equivalent. However, this is not generally true for all operators, and you cannot always make such a direct correlation. Equation 7 is always true, and if you understand the notation of state vectors, you can always work from there.

6   Hilbert Space

One last note, in quantum mechanics literature you might come across the term "Hilbert space". Do not be intimidated. This is simply an abstract "space" for vectors for which an inner product (denoted $\langle x, y \rangle$ for now) can be defined and with the following properties:

- A real or complex inner product which is conjugate symmetric: $\langle x, y \rangle = \langle y, x \rangle^*$

- The inner product is linear in its first argument: $\langle ax_1 + bx_2, y \rangle = a \langle x_1, y \rangle + b \langle x_2, y \rangle$ for all complex numbers a and b

- The inner product of an element with itself is positive definite: $\langle x, x \rangle > 0 \, \text{if} \, x \neq 0$

The state vectors that we talk about follow these properties, and "live" in the Hilbert space. The space vectors also follow these properties, and "live" in a subset of the Hilbert space, referred to as the Euclidean space.

Appendix I. Recommended reading

This discussion on state operators was based mostly on *The Feynman Lectures on Physics*, Volume III, chapters 3, 8, and 20. This is freely available online and I recommend you read those chapters, as well as the ones in between. They are extremely readable and instructive.

Appendix II. Arriving at the definition of average energies

If a system is in a stationary state η, the measured energy of that state would be consistent every time you take a measurement. These are states that have a definite energy:

$$\hat{H} | \eta_i \rangle = E_i | \eta_i \rangle \tag{13}$$

What if we have a state ψ which is not a stationary state? ψ can be thought of as a superposition of base states η, each with their own energy. Every time you take a measurement, ψ collapses to one of these base states. If you take a collection of energies, you would get a range of energies corresponding to the energies of the base states. You can define an average energy over all the number of measurements you take.

To find out how to obtain this average energy, we first project ψ onto the set of states η which have a definite energy:

$$|\psi \rangle = \sum_i |\eta_i \rangle \langle \eta_i | \psi \rangle$$

$$= \sum_i C_i |\eta_i \rangle$$

We then observe that the probability one observes energy E_i is just the absolute value of the amplitude $\langle \eta_i | \psi \rangle = C_i$ squared:

$$P_i = |C_i|^2$$

We additionally note that the average energy is the sum over all i of the total number of times E_i is observed (N_i) multiplied by E_i over the total number of measurements taken:

$$E_{av} = \frac{\sum_i N_i E_i}{N}$$

Using a statistical relation, we can approximate $\frac{N_i}{N}$ as just P_i. Using this approximation and the previous equations, we then get an expression for the expectation value of the average energy $\langle E_{av} \rangle$:

$$\langle E_{av} \rangle = \sum_i P_i E_i$$

$$= \sum_i |C_i|^2 E_i$$

$$= \sum_i C_i^* C_i E_i$$

$$= \sum_i \langle \psi | \eta_i \rangle E_i \langle \eta_i | \psi \rangle$$

We can then take ψ out of the summation and using equation 13:

$$\langle E_{av} \rangle = \langle \psi | \left\{ \sum_i | \eta_i \rangle E_i \langle \eta_i | \psi \rangle \right\}$$

$$= \langle \psi | \left\{ \sum_i E_i | \eta_i \rangle \langle \eta_i | \psi \rangle \right\}$$

$$= \langle \psi | \left\{ \sum_i \hat{H} | \eta_i \rangle \langle \eta_i | \psi \rangle \right\}$$

$$= \langle \psi | \hat{H} | \psi \rangle$$

This gives our final equation for the expectation value for the energy:

$$\langle E_{av} \rangle = \langle \psi | \hat{H} | \psi \rangle$$

We can extend this equation to other physical measurements that you can express as an operator:

$$\langle A_{av} \rangle = \langle \psi | \hat{A} | \psi \rangle$$