1
Vectors and Functions

There's a great deal of interesting physics in the Schrödinger equation and its solutions, and the mathematical underpinnings of that equation can be expressed in several ways. It's been my experience that students find it helpful to see a combination of Erwin Schrödinger's wave mechanics approach and the matrix mechanics approach of Werner Heisenberg, as well as Paul Dirac's bra and ket notation. So these first two chapters provide the mathematical foundations that will help you understand these different perspectives and "languages" of quantum mechanics, beginning with the basics of vectors in Section 1.1. With that basis in place, you can move on to Dirac notation in Section 1.2 and abstract vectors and functions in Section 1.3. The rules pertaining to complex numbers, vectors, and functions are reviewed in Section 1.4, followed by an explanation of orthogonal functions in Section 1.5, and using the inner product to find components in Section 1.6. The final section of this chapter (as in all later chapters) is a set of problems that will allow you to exercise your understanding of the concepts and mathematical techniques presented in this chapter. Remember that you can find full, interactive solutions to every problem on the book's website.

And since it's easy to lose sight of the architectural plan of an elaborate structure when you're laying the foundation, as mentioned in the Preface you'll find in each section a plain-language statement of the main ideas of that section as well as a short paragraph explaining the relevance of that development to the Schrödinger equation and quantum mechanics.

As you look through this chapter, don't forget that this book is modular, so if you have a good understanding of the included topics and their relevance to quantum mechanics, you should feel free to skip over this chapter and jump into the discussions of operators and eigenfunctions in Chapter 2. And if you're

already up to speed on those topics, the Schrödinger equation and quantum wavefunctions await your attention in later chapters.

1.1 Vector Basics

If you pick up any book about quantum mechanics, you're sure to find lots of discussion about wavefunctions and the solutions to the Schrödinger equation. But the language used to describe those functions, and the mathematical techniques used to analyze them, are rooted in the world of **vectors**. I've noticed that students who have a thorough understanding of basis vectors, inner products, and vector components are far more likely to succeed when they encounter the more advanced aspects of quantum mechanics, so this section is all about vectors.

When you first learned about vectors, you probably thought of a vector as an entity that has both magnitude (length) and direction (angles from some set of axes). You may also have learned to write a vector as a letter with a little arrow over its head (such as $\vec{A}$), and to "expand" a vector like this:

$$\vec{A} = A_x \hat{\imath} + A_y \hat{\jmath} + A_z \hat{k}. \tag{1.1}$$

In this expansion, A_x, A_y, and A_z are the components of vector $\vec{A}$, and $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are directional indicators called "basis vectors" of the coordinate system you're using to expand vector $\vec{A}$. In this case, that's the Cartesian (x, y, z) coordinate system shown in Fig. 1.1. It's important to understand that vector $\vec{A}$ exists independently of any particular basis system; the same vector may be expanded in many different basis systems.

The basis vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ are also called "unit vectors" because they each have length of one unit. And what unit is that? Whatever unit you're using to express the length of vector $\vec{A}$. It may help you to think of a unit vector as defining one "step" along a coordinate axis, so an expression such as

$$\vec{A} = 5\hat{\imath} - 2\hat{\jmath} + 3\hat{k}, \tag{1.2}$$

tells you to take five steps in the (positive) x-direction, two steps in the (negative) y-direction, and three steps in the (positive) z-direction to get from the start to the end of the vector $\vec{A}$.

You may also recall that the magnitude (that is, the length or "norm") of a vector, usually written as $|\vec{A}|$ or $\|\vec{A}\|$, can be found from its Cartesian components using the equation
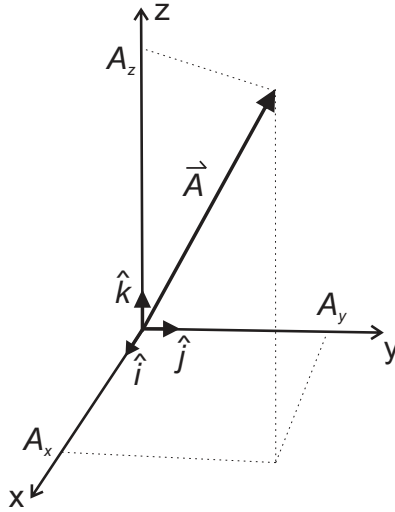
Figure 1.1 Vector $\vec{A}$ with its Cartesian components A_x, A_y, and A_z and the Cartesian unit vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$.

$$|\vec{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2},\qquad(1.3)$$

and that the negative of a vector (such as $-\vec{A}$) is a vector of the same length as $\vec{A}$ but pointed in the opposite direction.

Adding two vectors together can be done graphically, as shown in Fig. 1.2, by sliding one vector (without changing its direction or length) so that its tail is at the head of the other vector; the sum is a new vector drawn from the tail of the undisplaced vector to the head of the displaced vector. Alternatively, vectors may be added analytically by adding the components in each direction:

$$\vec{A} = A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}$$
$$+\ \underline{\vec{B} = B_x\hat{\imath} + B_y\hat{\jmath} + B_z\hat{k}}$$
$$\vec{C} = \vec{A}+\vec{B} = (A_x + B_x)\hat{\imath} + (A_y + B_y)\hat{\jmath} + (A_z + B_z)\hat{k}.\qquad(1.4)$$

Another important operation is multiplying a vector by a scalar (that is, a number with no directional indicator), which changes the length but not the direction of the vector. So if α is a scalar, then

$$\vec{D} = \alpha\vec{A} = \alpha(A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k})$$
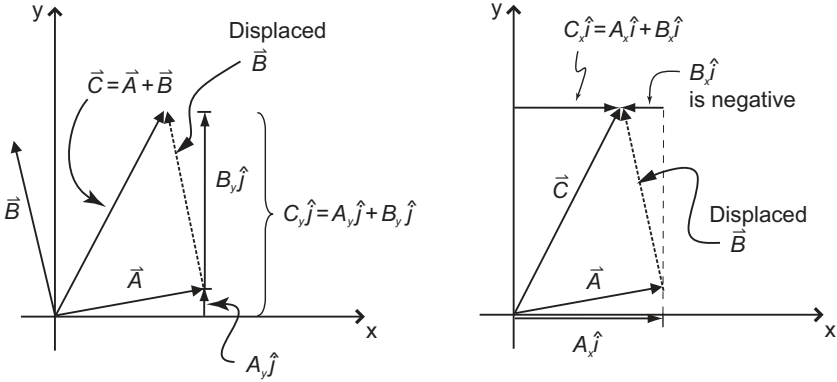$$= \alpha A_x\hat{\imath} + \alpha A_y\hat{\jmath} + \alpha A_z\hat{k}.$$

Figure 1.2 Adding vectors $\vec{A}$ and $\vec{B}$ graphically by sliding the tail of vector $\vec{B}$ to the head of vector $\vec{A}$ without changing its length or direction.

Scaling each component equally (by factor α) means that vector $\vec{D}$ points in the same direction as $\vec{A}$, but the length of $\vec{D}$ is

$$|\vec{D}| = \sqrt{D_x^2 + D_y^2 + D_z^2}$$
$$= \sqrt{(\alpha A_x)^2 + (\alpha A_y)^2 + (\alpha A_z)^2}$$
$$= \sqrt{\alpha^2 (A_x^2 + A_y^2 + A_z^2)} = \alpha |\vec{A}|.$$

So the vector's length is scaled by the factor α, but its direction remains the same (unless α is negative, in which case the direction reverses, but the vector still lies along the same line).

___

Relevance to Quantum Mechanics

___

As you'll see in later chapters, the solutions to the Schrödinger equation are quantum wavefunctions that behave like generalized higher-dimensional vectors. That means they can be added together to form a new wavefunction and they can be multiplied by scalars without changing their "direction." How functions can have "length" and "direction" is explained in Chapter 2.

___

In addition to summing vectors, multiplying vectors by scalars, and finding the length of vectors, another important operation is the scalar[1] product

___

[1] Note that this is called the scalar product because the result is a scalar, not because a scalar is involved in the multiplication.

(also called the "dot product") of two vectors, usually written as $(\vec{A}, \vec{B})$ or $\vec{A} \circ \vec{B}$. The scalar product is given by

$$(\vec{A}, \vec{B}) = \vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos\theta, \qquad (1.5)$$

in which θ is the angle between $\vec{A}$ and $\vec{B}$. In Cartesian coordinates, the dot product may be found by multiplying corresponding components and summing the results:

$$(\vec{A}, \vec{B}) = \vec{A} \circ \vec{B} = A_x B_x + A_y B_y + A_z B_z. \qquad (1.6)$$

Notice that if vectors $\vec{A}$ and $\vec{B}$ are parallel, then the dot product is

$$\vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos(0°) = |\vec{A}||\vec{B}|, \qquad (1.7)$$

since $\cos(0°) = 1$. Alternatively, if $\vec{A}$ and $\vec{B}$ are perpendicular, then the value of the dot product is zero:

$$\vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos(90°) = 0, \qquad (1.8)$$

since $\cos(90°) = 0$.

The dot product of a vector with itself gives the square of the magnitude of the vector:

$$\vec{A} \circ \vec{A} = |\vec{A}||\vec{A}| \cos(0°) = |\vec{A}|^2. \qquad (1.9)$$

A generalized version of the scalar product called the "inner product" is extremely useful in quantum mechanics, so it's worth a bit of your time to think about what happens when you perform an operation such as $\vec{A} \circ \vec{B}$. As you can see in Fig. 1.3a, the term $|\vec{B}| \cos\theta$ is the projection of vector $\vec{B}$ onto the direction of vector $\vec{A}$, so the dot product gives an indication of "how much" of $\vec{B}$ lies along the direction of $\vec{A}$.[2] Alternatively, you can isolate the $|\vec{A}| \cos\theta$ portion of the dot product $\vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos\theta$, which is the projection of $\vec{A}$ onto the direction of $\vec{B}$, as shown in Fig. 1.3b. From this perspective, the dot product indicates "how much" of vector $\vec{A}$ lies along the direction of $\vec{B}$. Either way, the dot product provides a measure of how much one vector "contributes" to the direction of another.

To make this concept more specific, consider what you get by dividing the dot product by the magnitude of $\vec{A}$ times the magnitude of $\vec{B}$:

---

[2] If you find the phrase "lies along" troubling (since vector $\vec{A}$ and vector $\vec{B}$ lie in different directions), perhaps it will help to imagine a tiny traveler walking from the start to the end of vector $\vec{B}$, and asking "In walking along vector $\vec{B}$, how much does a traveler advance in the direction of vector $\vec{A}$?"
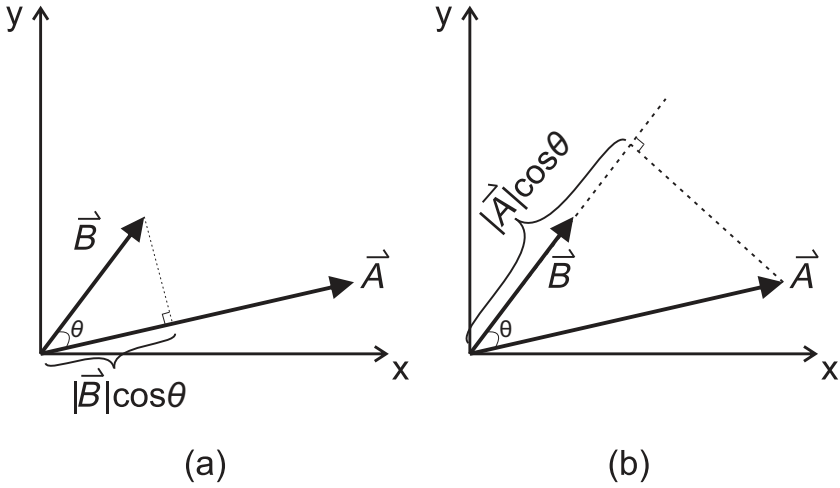
Figure 1.3 (a) The projection of vector $\vec{B}$ onto the direction of vector $\vec{A}$ and (b) the projection of vector $\vec{A}$ onto the direction of vector $\vec{B}$.

$$\frac{\vec{A} \circ \vec{B}}{|\vec{A}||\vec{B}|} = \frac{|\vec{A}||\vec{B}|\cos\theta}{|\vec{A}||\vec{B}|} = \cos\theta, \tag{1.10}$$

which ranges from one to zero as the angle between the vectors increases from $0°$ to $90°$. So if two vectors are parallel, each contributes its entire length to the direction of the other, but if they're perpendicular, neither makes any contribution to the direction of the other.

This understanding of the dot product makes it easy to comprehend the results of taking the dot product between pairs of the Cartesian unit vectors:

Each of these unit vectors lies entirely along itself
$$\begin{cases} \hat{\imath} \circ \hat{\imath} = |\hat{\imath}||\hat{\imath}|\cos 0° = (1)(1)(1) = 1 \\ \hat{\jmath} \circ \hat{\jmath} = |\hat{\jmath}||\hat{\jmath}|\cos 0° = (1)(1)(1) = 1 \\ \hat{k} \circ \hat{k} = |\hat{k}||\hat{k}|\cos 0° = (1)(1)(1) = 1 \end{cases}$$

No part of these unit vectors lies along any other
$$\begin{cases} \hat{\imath} \circ \hat{\jmath} = |\hat{\imath}||\hat{\jmath}|\cos 90° = (1)(1)(0) = 0 \\ \hat{\imath} \circ \hat{k} = |\hat{\imath}||\hat{k}|\cos 90° = (1)(1)(0) = 0 \\ \hat{\jmath} \circ \hat{k} = |\hat{\jmath}||\hat{k}|\cos 90° = (1)(1)(0) = 0 \end{cases}$$

The Cartesian unit vectors are called "orthonormal" because they're orthogonal (each is perpendicular to the others) as well as normalized (each has

magnitude of one). They're also called a "**complete set**" because any vector in three-dimensional Cartesian space can be made up of a weighted combination of these three basis vectors.

Here's a very useful trick: orthonormal basis vectors make it easy to use the dot product to determine the components of a vector. For a vector $\vec{A}$, the components A_x, A_y, and A_z can be found by dotting the basis vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ into $\vec{A}$:

$$
\begin{aligned}
A_x = \hat{\imath} \circ \vec{A} &= \hat{\imath} \circ (A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}) \\
&= A_x(\hat{\imath} \circ \hat{\imath}) + A_y(\hat{\imath} \circ \hat{\jmath}) + A_z(\hat{\imath} \circ \hat{k}) \\
&= A_x(1) + A_y(0) + A_z(0) = A_x.
\end{aligned}
$$

Likewise for A_y

$$
\begin{aligned}
A_y = \hat{\jmath} \circ \vec{A} &= \hat{\jmath} \circ (A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}) \\
&= A_x(\hat{\jmath} \circ \hat{\imath}) + A_y(\hat{\jmath} \circ \hat{\jmath}) + A_z(\hat{\jmath} \circ \hat{k}) \\
&= A_x(0) + A_y(1) + A_z(0) = A_y.
\end{aligned}
$$

And for A_z

$$
\begin{aligned}
A_z = \hat{k} \circ \vec{A} &= \hat{k} \circ (A_x\hat{\imath} + A_y\hat{\jmath} + A_z\hat{k}) \\
&= A_x(\hat{k} \circ \hat{\imath}) + A_y(\hat{k} \circ \hat{\jmath}) + A_z(\hat{k} \circ \hat{k}) \\
&= A_x(0) + A_y(0) + A_z(1) = A_z.
\end{aligned}
$$

This technique of digging out the components of a vector using the dot product and basis vectors is extremely valuable in quantum mechanics.

Main Ideas of This Section

Vectors are mathematical representations of quantities that may be expanded as a series of components, each of which pertains to a directional indicator called a basis vector. A vector may be added to another vector to produce a new vector, and a vector may be multiplied by a scalar or by another vector. The dot or scalar product between two vectors produces a scalar result proportional to the projection of one of the vectors along the direction of the other. The components of a vector in an orthonormal basis system may be found by dotting each basis vector into the vector.

Relevance to Quantum Mechanics

Just as a vector can be expressed as a weighted combination of basis vectors, a quantum wavefunction can be expressed as a weighted combination of basis wavefunctions. A generalized version of the dot product called the inner product can be used to calculate how much each component wavefunction contributes to the sum, and this determines the probability of various measurement outcomes.

1.2 Dirac Notation

Before making the connection between vectors and quantum wavefunctions, it's important for you to realize that vector components such as A_x, A_y, and A_z have meaning only when tied to a set of basis vectors (A_x to $\hat{\imath}$, A_y to $\hat{\jmath}$, and A_z to $\hat{k}$). If you had chosen to represent vector $\vec{A}$ using a different set of basis vectors (for example, by rotating the x-, y-, and z-axes and using basis vectors aligned with the rotated axes), you could have written the same vector $\vec{A}$ as

$$\vec{A} = A_{x'}\hat{\imath}' + A_{y'}\hat{\jmath}' + A_{z'}\hat{k}',$$

in which the rotated axes are designated x$'$, y$'$, and z$'$, and the basis vectors pointing along those axes are $\hat{\imath}'$, $\hat{\jmath}'$, and $\hat{k}'$.

When you expand a vector such as $\vec{A}$ in terms of different basis vectors, the vector components of the vector may change, but the new components and the new basis vectors add up to give the same vector $\vec{A}$. You may even choose to use a non-Cartesian set of basis vectors such as the spherical basis vectors $\hat{r}$, $\hat{\theta}$, and $\hat{\phi}$; expanding vector $\vec{A}$ in this basis looks like this:

$$\vec{A} = A_r\hat{r} + A_\theta\hat{\theta} + A_\phi\hat{\phi}.$$

Once again, different components, different basis vectors, but the *combination* of components and basis vectors gives the same vector $\vec{A}$.

What's the advantage of using one set of basis vectors or another? Depending on the geometry of the situation, it may be simpler to represent or manipulate vectors in a particular basis. But once you've specified a basis, a vector may be represented simply by writing its components in that basis as an ordered set of numbers.

For example, you could choose to represent a three-dimensional vector by writing its components into a single-column matrix

$$\vec{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix},$$

as long as you remember that vectors may be represented in this way only when the **basis system** has been specified.

Since they're vectors, the Cartesian basis vectors ($\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$) themselves can be written as column vectors. To do so, it's necessary to ask "In what basis?" Students sometimes find this a strange question, since we're talking about representing a basis vector, so isn't the basis obvious?

The answer is that it's perfectly possible to expand any vector, including a basis vector, using whichever basis system you choose. But some choices will lead to simpler representation than others, as you can see by representing $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ using their own Cartesian basis system:

$$\hat{\imath} = 1\hat{\imath} + 0\hat{\jmath} + 0\hat{k} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \hat{\jmath} = 0\hat{\imath} + 1\hat{\jmath} + 0\hat{k} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

and

$$\hat{k} = 0\hat{\imath} + 0\hat{\jmath} + 1\hat{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Such a basis system, in which each basis vector has only one nonzero component, and the value of that component is +1, is called the "standard" or "natural" basis.

Here's what it looks like if you express the Cartesian basis vectors ($\hat{\imath}, \hat{\jmath}, \hat{k}$) using the basis vectors ($\hat{r}, \hat{\theta}, \hat{\phi}$) of the spherical coordinate system

$$\hat{\imath} = \sin\theta\cos\phi\,\hat{r} + \cos\theta\cos\phi\,\hat{\theta} - \sin\phi\,\hat{\phi}$$
$$\hat{\jmath} = \sin\theta\sin\phi\,\hat{r} + \cos\theta\sin\phi\,\hat{\theta} + \cos\phi\,\hat{\phi}$$
$$\hat{k} = \cos\theta\,\hat{r} - \sin\theta\,\hat{\theta}.$$

So the column-vector representation of $\hat{\imath}, \hat{\jmath}, \hat{k}$ in the spherical basis system is

$$\hat{\imath} = \begin{pmatrix} \sin\theta\cos\phi \\ \cos\theta\cos\phi \\ -\sin\phi \end{pmatrix} \qquad \hat{\jmath} = \begin{pmatrix} \sin\theta\sin\phi \\ \cos\theta\sin\phi \\ \cos\phi \end{pmatrix} \qquad \hat{k} = \begin{pmatrix} \cos\theta \\ -\sin\theta \\ 0 \end{pmatrix}.$$

The bottom line is this: whenever you see a vector represented as a column of components, it's essential that you understand the basis system to which those components pertain.

Relevance to Quantum Mechanics

Like vectors, quantum wavefunctions can be expressed as a series of components, but those components have meaning only when you've defined the basis functions to which they pertain.

In quantum mechanics, you're likely to encounter entities called "ket vectors" or simply "kets," written with a vertical bar on the left and angled bracket on the right, such as $|A\rangle$. The ket $|A\rangle$ can be expanded in the same way as vector $\vec{A}$:

$$|A\rangle = A_x |i\rangle + A_y |j\rangle + A_z |k\rangle = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = A_x \hat{i} + A_y \hat{j} + A_z \hat{k} = \vec{A}. \quad (1.11)$$

So if kets are just a different way of representing vectors, why call them "kets" and write them as column vectors? This notation was developed by the British physicist Paul Dirac in 1939, while he was working with a generalized version of the dot product called the **inner product**, written as $\langle A|B\rangle$. In this context, "generalized" means "not restricted to real vectors in three-dimensional physical space," so the inner product can be used with higher-dimensional abstract vectors with complex components, as you'll see in Sections 1.3 and 1.4. Dirac realized that the inner product bracket $\langle A|B\rangle$ could be conceptually divided into two pieces, a left half (which he called a "**bra**") and a right half (which he called a "**ket**"). In conventional notation, an inner product between vectors $\vec{A}$ and $\vec{B}$ might be written as $\vec{A} \circ \vec{B}$ or $(\vec{A}, \vec{B})$, but in Dirac notation the inner product is written as

$$\text{Inner product of } |A\rangle \text{ and } |B\rangle = \langle A| \text{ times } |B\rangle = \langle A|B\rangle. \quad (1.12)$$

Notice that in forming the bracket $\langle A|B\rangle$ as the multiplication of bra $\langle A|$ by ket $|B\rangle$, the right vertical bar of $\langle A|$ and the left vertical bar of $|B\rangle$ are combined into a single vertical bar.

To calculate the inner product $\langle A|B\rangle$, begin by representing vector $\vec{A}$ as a ket:

$$|A\rangle = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \quad (1.13)$$