

Supplemental Material for *A Student's  
Guide to Laplace Transforms*

Vectors and Functions

Daniel A. Fleisch



# 1

## Vectors and Functions

This supplemental document provides an overview of vectors, abstract vectors, and functions that goes beyond the discussion in *A Student's Guide to Laplace Transforms*. The reason for providing this information is that you may find it helpful in understanding why and how integral transforms such as the Fourier transform and the Laplace transform and the discrete-time Z-transform work. This document is a revised and condensed version of the first chapter of *A Student's Guide to the Schrödinger Equation*, in which the relationship between quantum wavefunctions and abstract vectors is explained.

This document contains five sections, beginning with the basics of vectors in Section 1.1. Dirac notation is then introduced in Section 1.2, and abstract vectors and functions are discussed 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.

### 1.1 Vector Basics

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:

$$\vec{A} = A_x \hat{i} + A_y \hat{j} + A_z \hat{k}. \quad (1.1)$$

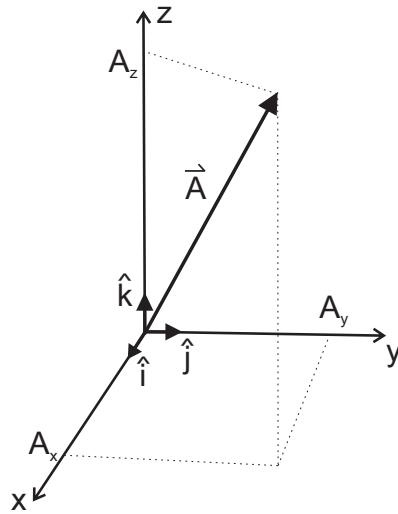


Figure 1.1 Vector  $\vec{A}$  with its Cartesian components  $A_x$ ,  $A_y$ , and  $A_z$  and the Cartesian unit vectors  $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$ .

In this expansion,  $A_x$ ,  $A_y$ , and  $A_z$  are the components of vector  $\vec{A}$ , and  $\hat{i}$ ,  $\hat{j}$ , 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  $\hat{i}$ ,  $\hat{j}$  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

$$\vec{A} = 5\hat{i} - 2\hat{j} + 3\hat{k}, \quad (1.2)$$

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 Carte-

sian components using the equation

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

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.

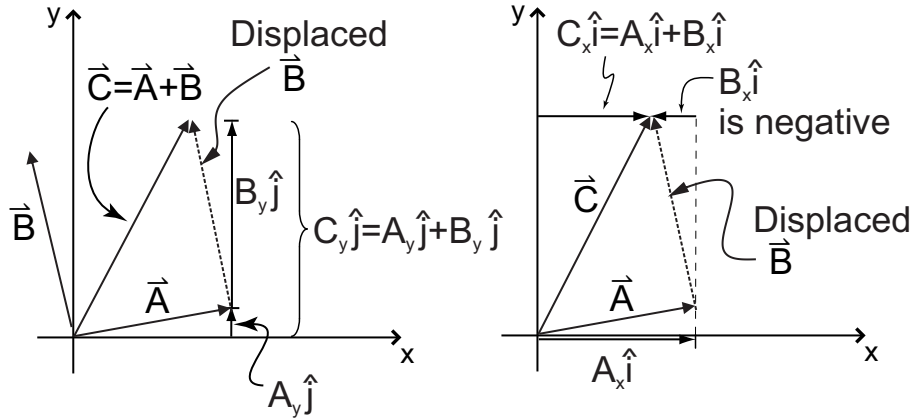


Figure 1.2 Adding vectors graphically.

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:

$$\begin{aligned} \vec{A} &= A_x \hat{i} + A_y \hat{j} + A_z \hat{k} \\ + \vec{B} &= B_x \hat{i} + B_y \hat{j} + B_z \hat{k} \\ \vec{C} = \vec{A} + \vec{B} &= (A_x + B_x) \hat{i} + (A_y + B_y) \hat{j} + (A_z + B_z) \hat{k}. \end{aligned} \quad (1.4)$$

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  $\alpha$  is a scalar, then

$$\begin{aligned} \vec{D} &= \alpha \vec{A} = \alpha(A_x \hat{i} + A_y \hat{j} + A_z \hat{k}) \\ &= \alpha A_x \hat{i} + \alpha A_y \hat{j} + \alpha A_z \hat{k}. \end{aligned}$$

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

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

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

In addition to summing vectors, multiplying vectors by scalars, and finding the length of vectors, another important operation is the scalar<sup>1</sup> product (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

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

in which  $\theta$  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:

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

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

$$\vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos(0^\circ) = |\vec{A}||\vec{B}| \quad (1.7)$$

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

$$\vec{A} \circ \vec{B} = |\vec{A}||\vec{B}| \cos(90^\circ) = 0 \quad (1.8)$$

since  $\cos(90^\circ) = 0$ .

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

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

A generalized version of the scalar product called the “inner product” is extremely useful in several areas of physics and engineering, so it's worth a bit of your time to think about what happens when you perform

<sup>1</sup> Note that this is called the scalar product because the result is a scalar, not because a scalar is involved in the multiplication.

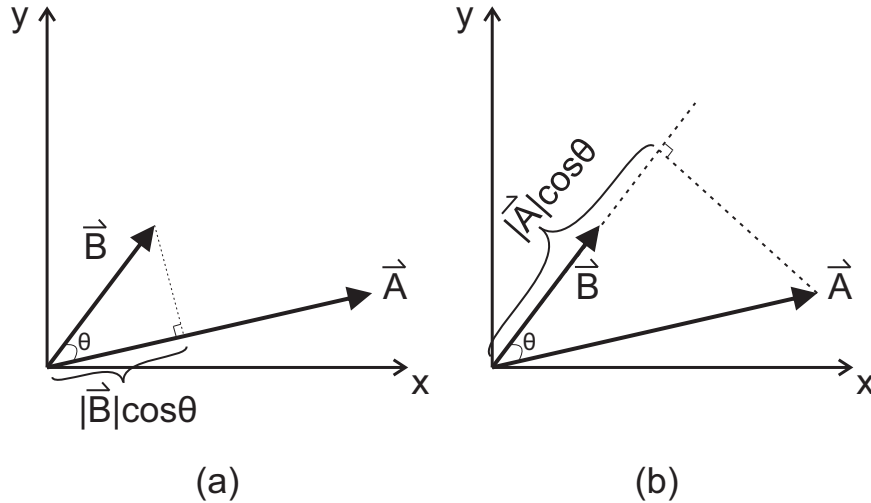


Figure 1.3 (a) The projection of vector  $\vec{B}$  onto the direction of vector  $\vec{A}$  and (b) the projection of vector  $\vec{A}$  onto the direction of vector  $\vec{B}$ .

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  $\vec{A}$ .<sup>2</sup> 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  $\vec{B}$ :

$$\frac{\vec{A} \circ \vec{B}}{|\vec{A}||\vec{B}|} = \frac{|\vec{A}||\vec{B}| \cos \theta}{|\vec{A}||\vec{B}|} = \cos \theta \quad (1.10)$$

which ranges from one to zero as the angle between the vectors increases

<sup>2</sup> 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  $\vec{A}$ ?”

from  $0^\circ$  to  $90^\circ$ . 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:

$$\begin{array}{l} \text{Each of these unit vectors} \\ \text{lies entirely along itself} \end{array} \left\{ \begin{array}{l} \hat{i} \circ \hat{i} = |\hat{i}||\hat{i}| \cos 0^\circ = (1)(1)(1) = 1 \\ \hat{j} \circ \hat{j} = |\hat{j}||\hat{j}| \cos 0^\circ = (1)(1)(1) = 1 \\ \hat{k} \circ \hat{k} = |\hat{k}||\hat{k}| \cos 0^\circ = (1)(1)(1) = 1 \end{array} \right.$$

$$\begin{array}{l} \text{No part of these unit vectors} \\ \text{lies along any other} \end{array} \left\{ \begin{array}{l} \hat{i} \circ \hat{j} = |\hat{i}||\hat{j}| \cos 90^\circ = (1)(1)(0) = 0 \\ \hat{i} \circ \hat{k} = |\hat{i}||\hat{k}| \cos 90^\circ = (1)(1)(0) = 0 \\ \hat{j} \circ \hat{k} = |\hat{j}||\hat{k}| \cos 90^\circ = (1)(1)(0) = 0 \end{array} \right.$$

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  $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$  into  $\vec{A}$ :

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

Likewise for  $A_y$

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

And for  $A_z$

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



This technique of digging out the components of a vectors using the dot product and basis vectors is extremely valuable in several areas of physics, including quantum mechanics.

## 1.2 Dirac Notation

When making the connection between vectors and functions, 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  $\hat{i}$ ,  $A_y$  to  $\hat{j}$ , 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

$$\vec{A} = A'_x \hat{i}' + A'_y \hat{j}' + A'_z \hat{k}'$$

in which the rotated axes are designated  $x'$ ,  $y'$ , and  $z'$ , and the basis vectors pointing along those axes are  $\hat{i}'$ ,  $\hat{j}'$ , 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 ( $\hat{i}$ ,  $\hat{j}$ , 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  $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$  using their own Cartesian basis system:

$$\hat{i} = 1\hat{i} + 0\hat{j} + 0\hat{k} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \hat{j} = 0\hat{i} + 1\hat{j} + 0\hat{k} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

and

$$\hat{k} = 0\hat{i} + 0\hat{j} + 1\hat{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Such a basis system, in which each basis vector has only one non-zero 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 ( $\hat{i}$ ,  $\hat{j}$ ,  $\hat{k}$ ) using the basis vectors ( $\hat{r}$ ,  $\hat{\theta}$ ,  $\hat{\phi}$ ) of the spherical coordinate system

$$\begin{aligned} \hat{i} &= \sin \theta \cos \phi \hat{r} + \cos \theta \cos \phi \hat{\theta} - \sin \phi \hat{\phi} \\ \hat{j} &= \sin \theta \sin \phi \hat{r} + \cos \theta \sin \phi \hat{\theta} + \cos \phi \hat{\phi} \\ \hat{k} &= \cos \theta \hat{r} - \sin \theta \hat{\theta}. \end{aligned}$$

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

$$\hat{i} = \begin{pmatrix} \sin \theta \cos \phi \\ \cos \theta \cos \phi \\ -\sin \phi \end{pmatrix} \quad \hat{j} = \begin{pmatrix} \sin \theta \sin \phi \\ \cos \theta \sin \phi \\ \cos \phi \end{pmatrix} \quad \hat{k} = \begin{pmatrix} \cos \theta \\ -\sin \theta \\ 0 \end{pmatrix}.$$

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.

When you're dealing with functions as abstract vectors, 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:

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

in which the subscripts indicate that these components pertain to the Cartesian basis system. Now form the bra  $\langle A|$  by taking the complex conjugate<sup>3</sup> of each component and writing them as a row vector:

$$\langle A| = (A_x^* \ A_y^* \ A_z^*). \quad (1.14)$$

<sup>3</sup> The reason for taking the complex conjugate is explained in Section 1.4, where you’ll also find a refresher on complex quantities.

The inner product  $\langle A|B\rangle$  is thus

$$\langle A| \text{ times } |B\rangle = \langle A|B\rangle = (A_x^* \ A_y^* \ A_z^*) \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}. \quad (1.15)$$

By the rules of matrix multiplication this gives

$$\langle A|B\rangle = (A_x^* \ A_y^* \ A_z^*) \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = A_x^* B_x + A_y^* B_y + A_z^* B_z \quad (1.16)$$

as you'd expect for a generalized version of the dot product.

So kets can be represented by column vectors, and bras can be represented by row vectors, but a common question among students new to Dirac notation is “What exactly *are* kets, and what are bras?” The answer to the first question is that kets are mathematical objects that are members of a “vector space” (also called a “linear space”). If you’ve studied any linear algebra, you’ve already encountered the concept of a vector space, and you may remember that a vector space is just a collection of vectors that behave according to certain rules. Those rules include the addition of vectors to produce new vectors (which live in the same space), and multiplying a vector by a scalar, producing a scaled version of the vector (which also lives in that space).

Since we’ll be dealing with generalized vectors rather than vectors in three-dimensional physical space, instead of labeling the components  $x, y, z$ , we’ll number them. And instead of using the Cartesian unit vectors  $\hat{i}, \hat{j}, \hat{k}$ , we’ll use the basis vectors  $\vec{\epsilon}_1, \vec{\epsilon}_2 \dots \vec{\epsilon}_N$ . So the equation

$$|A\rangle = A_x |i\rangle + A_y |j\rangle + A_z |k\rangle \quad (1.17)$$

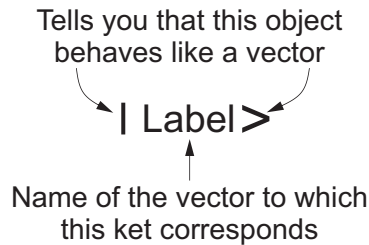
becomes

$$|A\rangle = A_1 |\epsilon_1\rangle + A_2 |\epsilon_2\rangle + \dots + A_N |\epsilon_N\rangle = \sum_{i=1}^N A_i |\epsilon_i\rangle \quad (1.18)$$

in which  $A_i$  represents the ket component for the basis ket  $|\epsilon_i\rangle$ .

But just as the vector  $\vec{A}$  is the same vector no matter which coordinate system you use to express its components, the ket  $|A\rangle$  exists independently of any particular set of basis kets (kets are said to be “basis independent”.) So ket  $|A\rangle$  behaves just like vector  $\vec{A}$ .

It may help you to think of a ket like this:



Once you’ve picked a basis system, why write the components of a ket as a column vector? One good reason is that it allows the rules of matrix multiplication to be applied to form scalar products, as in Eq. 1.16.

The other members of those scalar products are bras, and the definition of a bra is somewhat different from that of a ket. That’s because a bra is a “linear functional” (also called a “covector” or a “one-form”) that combines with a ket to produce a scalar; mathematicians say bras “map vectors to the field of scalars”.

So what’s a linear functional? It’s essentially a mathematical device (some authors refer to it as an instruction) that operates on another object. Hence a bra operates on a ket, and the result of that operation is a scalar. How does this operation map to a scalar? By following the rules of the scalar product, which you’ve already seen for the dot product between two real vectors. In Section 1.4 you’ll learn the rules for taking the inner product between two complex abstract vectors.

Bras don’t inhabit the same vector space as kets — they live in their own vector space that’s called the “dual space” to the space of kets. Within that space, bras can be added together and multiplied by scalars to produce new bras, just as kets can in their space.

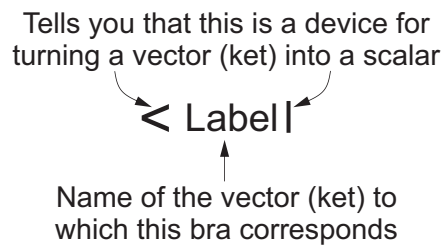
One reason that the space of bras is called “dual” to the space of kets is that for every ket there exists a corresponding bra, and when a bra operates on its corresponding (dual) ket, the scalar result is the square of the norm of the ket:

$$\langle A|A \rangle = (A_1^* \quad A_2^* \quad \dots \quad A_N^*) \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix} = |\vec{A}|^2,$$

just as the dot product of a (real) vector with itself gives the square of the vector's length (Eq. 1.9).

Note that the bra that is the dual of ket  $|A\rangle$  is written as  $\langle A|$ , not  $\langle A^*|$ . That's because the symbol inside the brackets of a ket or a bra is simply a name. For a ket, that name is the name of the vector that the ket represents. But for a bra, the name inside the brackets is the name of the ket to which the bra corresponds. So the bra  $\langle A|$  corresponds to the ket  $|A\rangle$ , but the components of bra  $\langle A|$  are the complex conjugates of the components of  $|A\rangle$ .

You may want to think of a bra like this:



### 1.3 Abstract Vectors and Functions

To understand the use of bras and kets, it's necessary to generalize the concepts of vector components and basis vectors to functions. I think the best way to do that is to change the way you graph vectors. Instead of attempting to replicate three-dimensional physical space as in Fig. 1.4a, simply line up the vector components along the horizontal axis of a two-dimensional graph, with the vertical axis representing the amplitude of the components, as in Fig. 1.4b.

At first glance, a two-dimensional graph of vector components may seem less useful than a three-dimensional graph, but its value becomes clear when you consider spaces with more than three dimensions.

And why would you want to do that? Because higher-dimensional abstract spaces turn out to be very useful tools for solving problems in several areas of physics, including classical and quantum mechanics. These spaces are called “abstract” because they're non-physical — that is, their dimensions don't represent the physical dimensions of the universe we inhabit. For example, an abstract space might consist of all

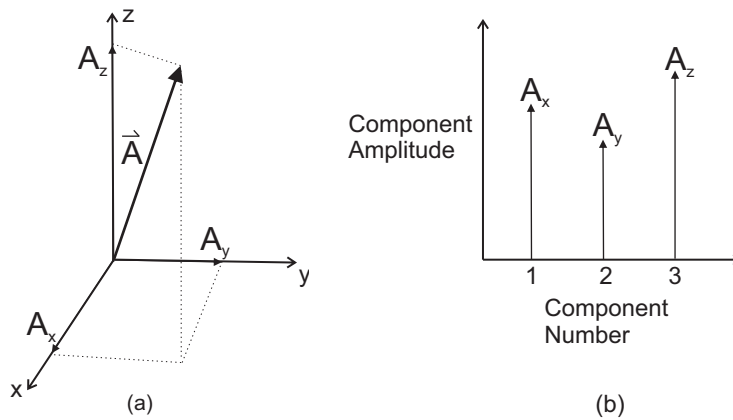


Figure 1.4 Vector components graphed in 3-D and 2-D.

of the values of the parameters of a mathematical model, or all of the possible configurations of a system. So the axes could represent speed, momentum, acceleration, energy, or any other parameter of interest.

Now imagine drawing a set of axes in an abstract space and marking each axis with the values of a parameter. That makes each parameter a “generalized coordinate”; “generalized” because these are not spatial coordinates (such as  $x$ ,  $y$ , and  $z$ ), but a “coordinate” nonetheless because each location on the axis represents a position in the abstract space. So if speed is used as a generalized coordinate, an axis might represent the range of speeds from zero to 20 meters per second, and the “distance” between two points on that axis is simply the difference between the speeds at those two points.

Physicists sometimes refer to “length” and “direction” in an abstract space, but you should remember that in such cases “length” is not a physical distance, but rather the difference in coordinate values at two locations. And “direction” is not a spatial direction, it’s an angle relative to an axis along which a parameter changes.

The multi-dimensional space most useful in several areas of physics is an abstract vector space called “Hilbert space”, after the German mathematician David Hilbert.

To understand the characteristics of Hilbert space, recall that vector spaces are collections of vectors that behave according to certain rules, such as vector addition and scalar multiplication. In addition to those rules, an “inner product space” also includes rules for multiplying two

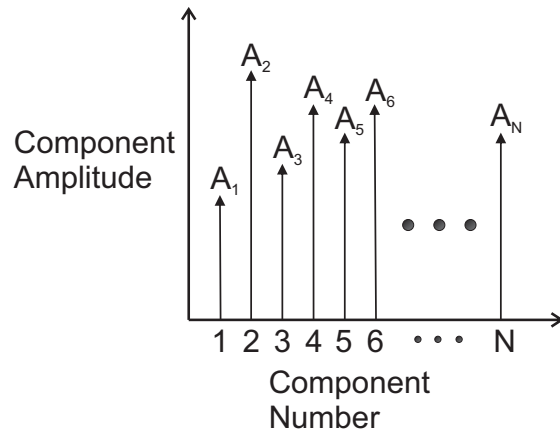


Figure 1.5 Vector components of an N-dimensional vector.

vectors together (the generalized scalar product). But an issue arises when forming the inner product between two higher-dimensional vectors, and to understand that issue, consider the graph of the components of an N-dimensional vector shown in Fig. 1.5.

Just as each of the three components ( $A_x$ ,  $A_y$ , and  $A_z$ ) pertains to a basis vector ( $\hat{i}$ ,  $\hat{j}$ , and  $\hat{k}$ ), each of the N components in Fig. 1.5 pertains to a basis vector in the N-dimensional abstract vector space inhabited by the vector.

Now imagine how such a graph would appear for a vector with an even larger number of components. The more components that you display on your graph for a given range, the closer together those components will appear along the horizontal axis, as shown in Fig. 1.6. If you're dealing with a vector with an extremely large number of components, the components may be treated as a continuous function rather than a set of discrete values. That function (call it "f") is depicted as the curvy line connecting the tips of the vector components in Fig. 1.6. As you can see, the horizontal axis is labeled with a continuous variable (call it "x"), which means that the amplitudes of the components are represented by the continuous function  $f(x)$ <sup>4</sup>.

So the continuous function  $f(x)$  is composed of a series of amplitudes, with each amplitude pertaining to a different value of the continuous

<sup>4</sup> We're dealing with functions of a single variable called  $x$ , but the same concepts apply to functions of multiple variables.



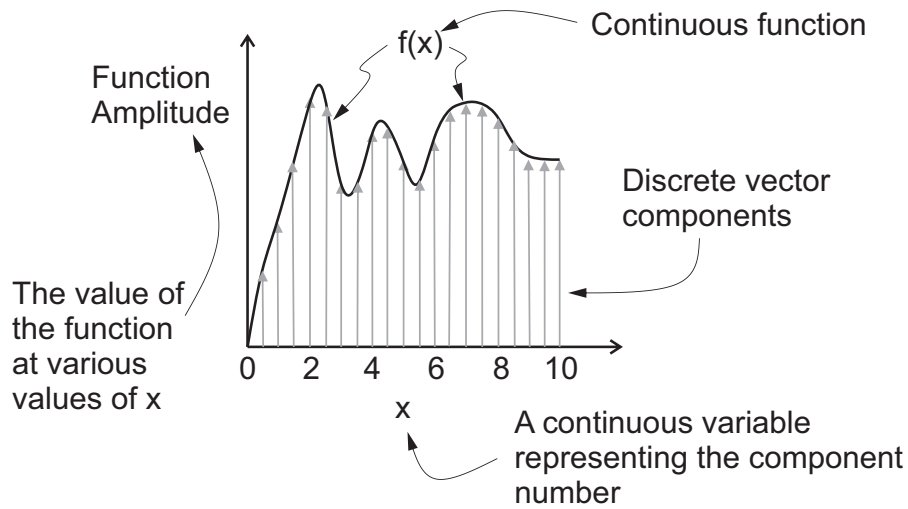


Figure 1.6 Relationship between vector components and continuous function.

variable  $x$ . And a vector is composed of a series of component amplitudes, with each component pertaining to a different basis vector.

In light of this parallel between a continuous function such as  $f(x)$  and the components of a vector such as  $\vec{A}$ , it's probably not surprising that the rules for addition and scalar multiplication apply to functions as well as vectors. So two functions  $f(x)$  and  $g(x)$  add to produce a new function, and that addition is done by adding the value of  $f(x)$  to the value of  $g(x)$  at every  $x$  (just as the addition of two vectors is done by adding corresponding components for each basis vector). Likewise, multiplying a function by a scalar results in a new function which has a value at every  $x$  of the original function  $f(x)$  times the scalar multiplier (just as multiplying a vector by a scalar produces a new vector with each component amplitude multiplied by the scalar).

But what about the inner product? Is there an equivalent process for continuous functions? Yes, there is. Since you know that for vectors the dot product in an orthonormal system can be found by summing the products of corresponding components in a given basis (such as  $A_x B_x + A_y B_y + A_z B_z$ ), a reasonable guess is that the equivalent operation for continuous functions such as  $f(x)$  and  $g(x)$  involves multiplication of the functions followed by integration rather than discrete summation. That works — the inner product between two functions  $f(x)$  and  $g(x)$  (which,

like vectors, may be represented by kets) is found by integrating their product over  $x$ :

$$(f(x), g(x)) = \langle f(x)|g(x)\rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx. \quad (1.19)$$

in which the asterisk after the function  $f(x)$  in the integral represents the complex conjugate, as in Eq. 1.16. The reason for taking the complex conjugate is explained in the next section.

And what's the significance of the inner product between two functions? Recall that the dot product between two vectors uses the projection of one vector onto the direction of the other to tell you how much one vector "lies along" the direction of the other. Similarly, the inner product between two functions uses the "projection" of one function onto the other to tell you how much of one function "lies along" the other (or, if you prefer, how much one function gets you in the "direction" of the other function)<sup>5</sup>.

Obedying the rules for addition, scalar multiplication, and the inner product means that functions like  $f(x)$  can behave like vectors — they are not members of the vector space of three-dimensional physical vectors, but they are members of their own abstract vector space.

There is, however, one more condition that must be satisfied before we can call that vector space a Hilbert space. That condition is that the functions must have a finite norm:

$$|f(x)|^2 = \langle f(x)|f(x)\rangle = \int_{-\infty}^{\infty} f^*(x)f(x)dx < \infty. \quad (1.20)$$

In other words, the integral of the square of every function in this space must converge to a finite value. Such functions are said to be "square summable" or "square integrable".

## 1.4 Complex Numbers, Vectors, and Functions

The motivation for the sequence of Figures 1.4, 1.5, and 1.6 is to help you understand the relationship between vectors and functions, and that understanding is helpful in several areas of physics and engineering. In some applications, the relevant functions and equations are complex —

<sup>5</sup> The concept of the "direction" of a function may make more sense after you've read about orthogonal functions in Section 1.5.

that is, they include the imaginary unit “ $i$ ” (the square root of minus one).<sup>6</sup> So this section contains a short review of complex numbers and their use in the context of vector components and Dirac notation.

As mentioned in the previous section, the process of taking an inner product between vectors or functions is slightly different for complex quantities. How can a vector be complex? By having complex components. To see why that has an effect on the inner product, consider the length of a vector with complex components. Remember, complex quantities can be purely real, purely imaginary, or a mixture of real and imaginary parts. So the most general way of representing a complex quantity  $z$  is

$$z = x + iy \quad (1.21)$$

in which  $x$  is the real part of  $z$  and  $y$  is the imaginary part of  $z$  (be sure not to confuse the imaginary unit  $i = \sqrt{-1}$  in this equation with the  $\hat{i}$  unit vector — you can always tell the difference by noting the caret hat on the unit vector  $\hat{i}$ ).

One important difference between real and imaginary numbers is that they lie along different number lines. The number line for imaginary numbers is perpendicular to the real number line, and a two-dimensional plot of both number lines represents the “complex plane” shown in Fig. 1.7.

As you can see from this figure, knowing the real and imaginary parts of a complex number allows you to find the magnitude or norm of that number. The magnitude of a complex number is the distance between the point representing the complex number and the origin in the complex plane, and you can find that distance using the Pythagorean theorem

$$|z|^2 = x^2 + y^2. \quad (1.22)$$

But if you try to square the complex number  $z$  by multiplying by itself, you find

$$z^2 = z \times z = (x + iy) \times (x + iy) = x^2 + 2ixy - y^2 \quad (1.23)$$

which is a complex number, and which may be negative. But a distance should be a real and positive number, so this is clearly not the way to find the distance of  $z$  from the origin.

<sup>6</sup> Mathematicians say that complex functions are members of an abstract linear vector space “over the field of complex numbers”. That means that the components may be complex, and that the rules for scaling a function by multiplying by a scalar apply not only to real scalars, but complex numbers as well.

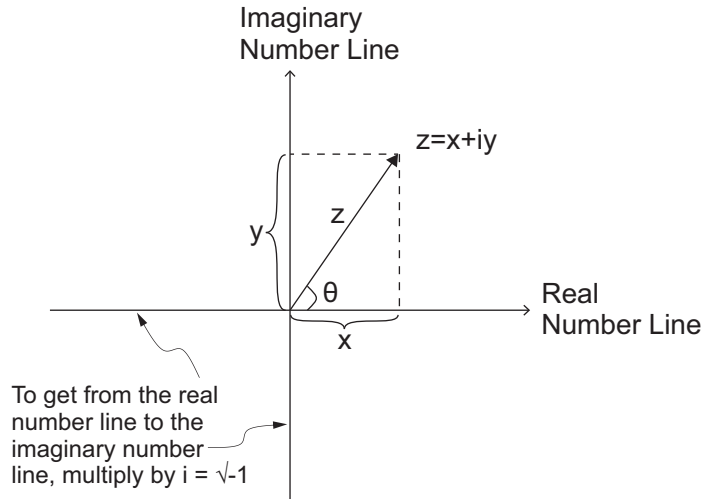


Figure 1.7 Complex number  $z=x+iy$  in the complex plane.

To correctly find the magnitude of a complex quantity, it's necessary to multiply the quantity not by itself, but by its complex conjugate. To take the complex conjugate of a complex number, just change the sign of the imaginary part of the number. The complex conjugate is usually indicated by an asterisk, so for the complex quantity  $z = x + iy$ , the complex conjugate is

$$z^* = x - iy. \quad (1.24)$$

Multiplying by the complex conjugate ensures that the magnitude of a complex number will be real and positive (as long as the real and the imaginary parts are not both zero). You can see that by writing out the terms of the multiplication:

$$|z|^2 = z \times z^* = (x + iy) \times (x - iy) = x^2 - xiy + iyx + y^2 = x^2 + y^2 \quad (1.25)$$

as expected. And since the magnitude (or norm) of a vector  $\vec{A}$  can be found by taking the square root of the inner product of the vector with itself, the complex conjugate is built into the process of taking the inner product between complex quantities:

$$|A| = \sqrt{\vec{A} \circ \vec{A}} = \sqrt{A_x^* A_x + A_y^* A_y + A_z^* A_z} = \sqrt{\sum_{i=1}^N A_i^* A_i}. \quad (1.26)$$

This also applies to complex functions:

$$|f(x)| = \sqrt{\langle f(x)|f(x) \rangle} = \sqrt{\int_{-\infty}^{\infty} f^*(x)f(x)dx}. \quad (1.27)$$

So it's necessary to use the complex conjugate to find the norm of a complex vector or function. If the inner product involves two different vectors or functions, by convention the complex conjugate is taken of the *first* member of the pair:

$$\begin{aligned} \vec{A} \circ \vec{B} &= \sum_{i=1}^N A_i^* B_i \\ \langle f(x)|g(x) \rangle &= \int_{-\infty}^{\infty} f^*(x)g(x)dx. \end{aligned} \quad (1.28)$$

This is the reason for the complex conjugation in the earlier discussion of the inner product using bras and kets (Eqs. 1.16 and 1.19).

The requirement to take the complex conjugate of one member of the inner product for complex vectors and functions means that the order matters, so  $\vec{A} \circ \vec{B}$  is not the same as  $\vec{B} \circ \vec{A}$ . That's because

$$\begin{aligned} \vec{A} \circ \vec{B} &= \sum_{i=1}^N A_i^* B_i = \sum_{i=1}^N (A_i B_i^*)^* = \sum_{i=1}^N (B_i^* A_i)^* = (\vec{B} \circ \vec{A})^* \\ \langle f(x)|g(x) \rangle &= \int_{-\infty}^{\infty} f^*(x)g(x)dx = \int_{-\infty}^{\infty} [g^*(x)f(x)]^* dx = (\langle g(x)|f(x) \rangle)^*. \end{aligned} \quad (1.29)$$

So reversing the order of the complex vectors or functions in an inner product produces a result that is the complex conjugate of the inner product without switching.

The convention of applying the complex conjugate to the first member of the inner product is common but not universal in physics texts, so you should be aware that you may find some texts and on-line resources that apply the complex conjugate to the second member.

## 1.5 Orthogonal Functions

For vectors, the concept of orthogonality is straightforward: two vectors are orthogonal if their scalar product is zero, which means that the projection of one of the vectors onto the direction of the other has zero

length. Simply put, orthogonal vectors lie along perpendicular lines, as shown in Fig. 1.8a for the two-dimensional vectors  $\vec{A}$  and  $\vec{B}$  (which we'll take as real for simplicity).

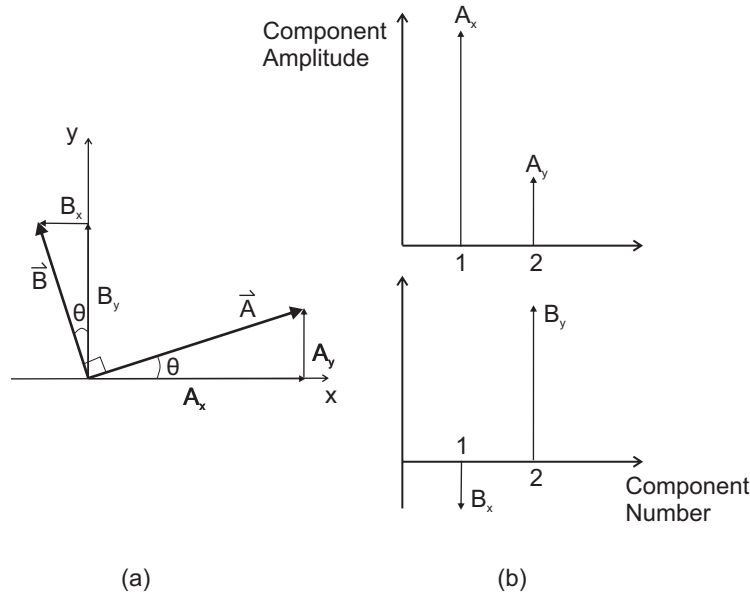


Figure 1.8 Components of perpendicular vectors.

Now consider the plots of the Cartesian components of vectors  $\vec{A}$  and  $\vec{B}$  in Fig. 1.8b. You can learn something about the relationship between these components by writing out the scalar product of  $\vec{A}$  and  $\vec{B}$ :

$$\begin{aligned}\vec{A} \circ \vec{B} &= A_x B_x + A_y B_y = 0 \\ A_x B_x &= -A_y B_y \\ \frac{A_x}{A_y} &= -\frac{B_y}{B_x}.\end{aligned}$$

This can only be true if one (and only one) of the components of  $\vec{A}$  has the opposite sign of the corresponding component of  $\vec{B}$ . In this case, since  $\vec{A}$  points up and to the right (that is,  $A_x$  and  $A_y$  are both positive), to be perpendicular,  $\vec{B}$  must point either up and to the left (with  $B_x$  negative and  $B_y$  positive, as shown in Fig. 1.8a), or down and to the right (with  $B_x$  positive and  $B_y$  negative). Additionally, since the angle

between the x- and y-axes is  $90^\circ$ , if  $\vec{A}$  and  $\vec{B}$  are perpendicular, the angle between  $\vec{A}$  and the positive x-axis (shown as  $\theta$  in Fig. 1.8a) must be the same as the angle between  $\vec{B}$  and the positive y-axis (or negative y-axis had we taken the “down and to the right” option for  $\vec{B}$ ). For those angles to be the same, the ratio of  $\vec{A}$ 's components ( $A_x/A_y$ ) must have the same magnitude as the inverse ratio of  $\vec{B}$ 's components ( $B_y/B_x$ ). You can get an idea of this inverse ratio in Fig. 1.8b.

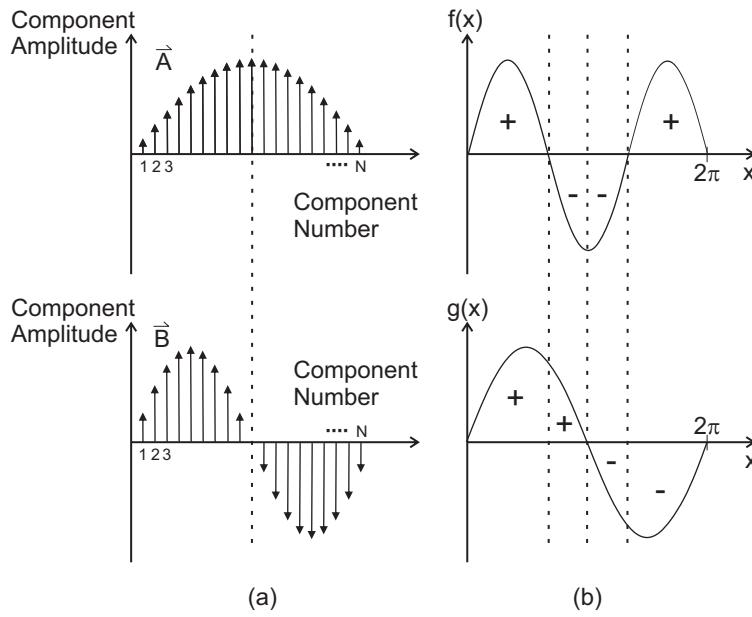


Figure 1.9 Orthogonal N-dimensional vectors (a) and functions (b).

Similar considerations apply to N-dimensional abstract vectors as well as continuous functions, as shown in Fig. 1.9a and b. If the N-dimensional abstract vectors  $\vec{A}$  and  $\vec{B}$  in Fig. 1.9a (again taken as real) are orthogonal, then their inner product  $(\vec{A}, \vec{B})$  must equal zero:

$$(\vec{A}, \vec{B}) = \sum_{i=1}^N A_i^* B_i = A_1 B_1 + A_2 B_2 + \dots + A_N B_N = 0.$$

For this sum to be zero, it must be true that some of the component products have opposite signs of others, and the total of all the negative products must equal the total of all the positive products. In the case of

the two  $N$ -dimensional vectors shown in Fig. 1.9a, the components in the left half of  $\vec{B}$  have the same sign as the corresponding components of  $\vec{A}$ , so the products of those left-half components ( $A_i B_i$ ) are all positive. But the components in the right half of  $\vec{B}$  have the opposite sign of the corresponding components in  $\vec{A}$ , so those products are all negative.

Since the magnitudes of these two vectors are symmetric about their midpoints, the magnitude of the sum of the left-half products equals the magnitude of the sum of the right-half products. With equal magnitudes and opposite signs, the sum of the products of the components from the left half and the right half is zero.

So although  $\vec{A}$  and  $\vec{B}$  are abstract vectors with “directions” only with respect to generalized rather than spatial coordinates, these two  $N$ -dimensional vectors satisfy the requirements of orthogonality, just as the two spatial vectors did in Fig. 1.8. Stated another way, even though we have no way of drawing the  $N$  dimensions of these vectors in different physical directions in our three-dimensional space, the zero inner product of  $\vec{A}$  and  $\vec{B}$  means that the projection of vector  $\vec{A}$  onto vector  $\vec{B}$  (and of  $\vec{B}$  onto  $\vec{A}$ ) has zero “length” in their  $N$ -dimensional vector space.

By this point, you’ve probably realized how orthogonality applies to functions such as  $f(x)$  and  $g(x)$  shown in Fig. 1.9b. Since these functions (also taken as real for simplicity) are continuous, the inner-product sum becomes an integral, as described in the previous section of this chapter. For these functions, the statement of orthogonality is

$$(f(x), g(x)) = \langle f(x) | g(x) \rangle = \int_{-\infty}^{\infty} f^*(x)g(x)dx = \int_{-\infty}^{\infty} f(x)g(x)dx = 0.$$

Just as in the case of discrete vectors  $\vec{A}$  and  $\vec{B}$ , the product  $f(x)g(x)$  can be thought of as multiplying the value of the function  $f(x)$  by the value of the function  $g(x)$  at each value of  $x$ . Integrating this product over  $x$  is the equivalent of finding the area under the curve formed by the product  $f(x)g(x)$ .

In the case of the functions  $f(x)$  and  $g(x)$  in Fig. 1.9b, you can estimate the result of multiplying the two functions and integrating (continuously summing) the result. To do that, notice that for the first one-third of the range of  $x$  shown on the graph (left of the first dashed vertical line),  $f(x)$  and  $g(x)$  have the same sign (both positive). For the next one-sixth of the graph (between the first and second dashed vertical lines), the two functions have opposite signs ( $f(x)$  negative and  $g(x)$  positive). For the next one-sixth of the graph (between the second and third dashed



vertical lines), the signs of  $f(x)$  and  $g(x)$  are again the same (both negative), and for the final one-third of the graph (right of the third dashed vertical line), the signs are opposite. Due to the symmetry of the regions in which the product  $f(x)g(x)$  is positive and negative, the total sum is zero, and these two functions qualify as orthogonal over this range of  $x$ . So these two functions are orthogonal in this region in exactly the same way as vectors  $\vec{A}$  and  $\vec{B}$  are orthogonal.

If you prefer a more mathematical approach to determining the orthogonality of these functions, notice that  $g(x)$  is the  $\sin x$  function over the range of  $x = 0$  to  $2\pi$  and that  $f(x)$  is the  $\sin \frac{3}{2}x$  function over the same range. The inner product of these two functions is

$$\begin{aligned}\langle f(x)|g(x)\rangle &= \int_{-\infty}^{\infty} f^*(x)g(x)dx = \int_0^{2\pi} \sin\left(\frac{3}{2}x\right)\sin(x)dx \\ &= \left[\sin\frac{x}{2} - \frac{1}{5}\sin\frac{5x}{2}\right]_0^{2\pi} = 0,\end{aligned}$$

which is consistent with the result obtained by estimating the area under the curve of the product  $f(x)g(x)$ .

Orthogonal functions play an important role in Fourier and Laplace transforms, as described in Chapter 1 of *A Student's Guide to Laplace Transforms*.

