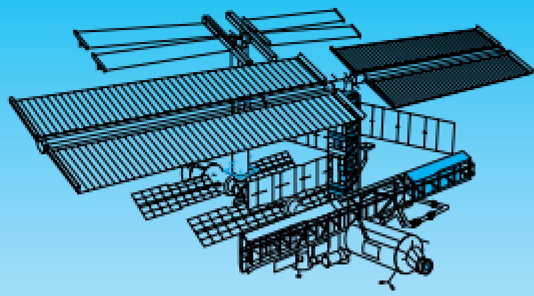
CHAPTER 2



Vectors and Vector Spaces

Engineers, scientists, and physicists need to work with systems involving physical quantities that, unlike the density of a solid, cannot be characterized by a single number. This chapter is about the algebra of important and useful quantities called vectors that arise naturally when studying physical systems, and are defined by an ordered group of three numbers (a, b, c). Vectors are of fundamental importance and they play an essential role when the laws governing engineering and physics are expressed in mathematical terms.

A scalar quantity is one that is completely described when its magnitude is known, such as pressure, temperature, and area. A vector is a quantity that is completely specified when both its magnitude and direction are given, such as force, velocity, and momentum. A vector can be described geometrically as a directed straight line segment, with its length proportional to the magnitude of the vector, the line representing the vector parallel to the line of action of the vector, and an arrow on the line showing the direction along the line, or the sense, in which the vector acts.

This geometrical interpretation of a vector is valuable in many ways, as it can be used to add and subtract vectors and to multiply them by a scalar, since this merely involves changing their magnitude and sense, while leaving the line to which they are parallel unchanged. However, to perform more general algebraic operations on vectors some other form of representation is required. The one that is used most frequently involves describing a vector in terms of what are called its components along a set of three mutually orthogonal axes, which are usually taken to be the axes $O\{x, y, z\}$ in the cartesian coordinate system. Here, by the component of a vector along a given line l, we mean the length of the perpendicular projection of the vector onto the line l.

We will see later that this cartesian representation of a vector identifies it completely in terms of three components and enables algebraic operations to be performed on it. In particular, it allows the introduction of the scalar product, or dot product, of two vectors that results in a scalar, and a vector product, or cross product, of two vectors that leads to a vector.

Finally, vectors and their algebra will be generalized to n space dimensions, leading to the concept of a vector space and to some related ideas.

2.1 Vectors, Geometry, and Algebra

scalar

Many quantities are completely described once their magnitude is known. A typical example of a physical quantity of this type is provided by the temperature at a given point in a room that is determined by the number specifying its value measured on a temperature scale, such as degrees F or degrees C. A quantity such as this is called a **scalar** quantity, and different examples of mathematical and physical scalar quantities are real numbers, length, area, volume, mass, speed, pressure, chemical concentration, electrical resistance, electric potential, and energy.

vector

Other physical quantities are only fully specified when both their magnitude and direction are given. Quantities like this are called **vector** quantities, and a typical example of a vector quantity arises when specifying the instantaneous motion of a fluid particle in a river. In this case both the particle speed and its direction must be given if the description of its motion is to be complete. Speed in a given direction is called **velocity**, and velocity is a vector quantity. Some other examples of vector quantities are force, acceleration, momentum, the heat flow vector at a point in a block of metal, the earth's magnetic field at a given location, and a mathematical quantity called the gradient of a scalar function of position that will be defined later. By definition, the magnitude of a vector quantity is a nonnegative number (a scalar) that measures its size without regard to its direction, so, for example, the magnitude of a velocity is a speed.

directed straight line segment

translation

A convenient geometrical representation of a vector is provided by a straight line segment drawn in space parallel to the required direction, with an arrowhead indicating the **sense** in which the vector acts along the line segment, and the length of the line segment proportional to the magnitude of the vector. This is called a **directed straight line segment**, and by definition all directed straight line segments that are parallel to one another and have the same sense and length are regarded as equal. Expressed differently, moving a directed straight line segment parallel to itself so that its length remains the same and its arrow still points in the same direction leaves the vector it represents unchanged. A shift of a directed straight line segment of this type is called a **translation** of the vector it represents. For this reason the terms *directed straight line segment* and *vector* can be used interchangeably. Some examples of vectors that are equal through translation are shown in Fig. 2.1.

It must be emphasized that geometrical representations of vectors as directed straight line segments in space are defined without reference to a specific coordinate system. This purely geometrical interpretation of vectors finds many applications, though a different form of representation is necessary if an effective vector algebra is to be developed for use with the calculus. An analytical representation of vectors that allows a vector algebra to be constructed with this purpose in mind can be based on a general coordinate system. However, throughout this chapter only rectangular cartesian coordinates will be used because they provide a simple and natural way of representing vectors.

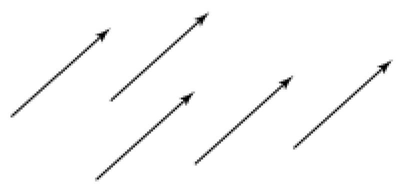


FIGURE 2.1 Equal geometrical vectors.

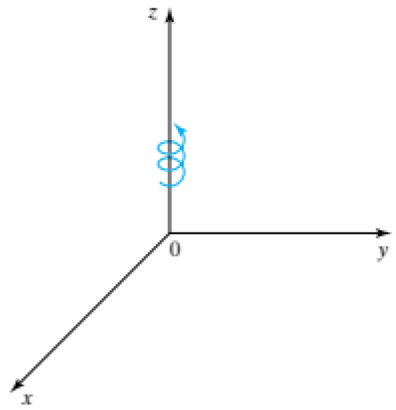


FIGURE 2.2 A right-handed rectangular cartesian coordinate system.

In rectangular cartesian coordinates the x-, y-, and z-axes are all mutually orthogonal (perpendicular), and the positive sense along the axes is taken to be in the direction of increasing x, y, and z. The orientation of the axes will always be such that the positive direction along the z-axis is the one in which a right-handed screw (such as a corkscrew) aligned with the z-axis will advance when rotated from the positive x-axis to the positive y-axis, as shown in Fig. 2.2. A system of axes with this property is called a **right-handed system**.

right-handed system

The end of a vector toward which the arrow points will be called the **tip** of the vector, and the other end its **base**. Because a vector is invariant under a translation, there is no loss of generality in taking its base to be located at the origin O of the coordinate system, and its tip at a point P with the coordinates (a_1, a_2, a_3), say, as shown in Fig. 2.3. An application of the Pythagoras theorem to the triangle OPP'

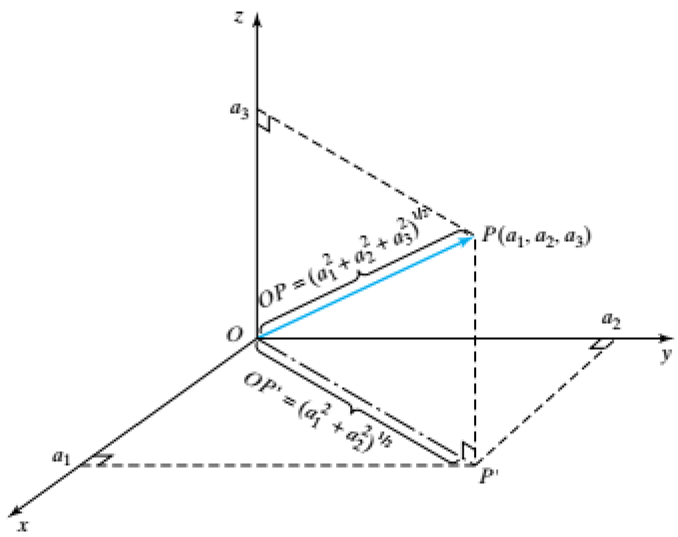


FIGURE 2.3 The vector from O to P and its components a_1, a_2, and a_3 in the x-, y-, z-coordinate system.

magnitude, unit vector, and components

shows the length of the line from O to P to be $(a_1^2 + a_2^2 + a_3^2)^{1/2}$. This length is *proportional* to the **magnitude** of the vector it represents, and as the base of the vector is at O, the sense of the vector is from O to P. For convenience, the constant of proportionality will be taken to be 1, so a directed straight line segment of unit length will represent a vector of magnitude 1 and so will be called a **unit vector**. Using this convention, the vector represented by the line from O to P in Fig. 2.3 has magnitude $(a_1^2 + a_2^2 + a_3^2)^{1/2}$. The three numbers a_1, a_2, and a_3, in this order, that define the vector from O to P are called its **components** in the x, y, and z directions, respectively.

ordered number triple

A set of three numbers a_1, a_2, and a_3 in a given order, written (a_1, a_2, a_3), is called an **ordered number triple**. As the coordinates (a_1, a_2, a_3) of point P in Fig. 2.3 completely define the vector from O to P, this ordered number triple may be taken as the definition of the vector itself. In general, changing the order of the numbers in an ordered number triple changes the vector it defines.

Sometimes it is necessary to consider a vector whose base does not coincide with the origin. Suppose that when this occurs the base C is at the point (c_1, c_2, c_3) and the tip D is at the point (d_1, d_2, d_3). Then Fig. 2.4 shows the components of this vector in the x, y, and z directions to be $d_1 - c_1$, $d_2 - c_2$, and $d_3 - c_3$. These components determine both the magnitude and direction of the vector. The vector is described by the ordered number triple $(d_1 - c_1, d_2 - c_2, d_3 - c_3)$, and the length of CD that is equal to the magnitude of the vector is $[(d_1 - c_1)^2 + (d_2 - c_2)^2 + (d_3 - c_3)^2]^{1/2}$.

norm and modulus

For convenience, it is usual to represent a vector by a single boldface character such as **a**, and its **magnitude** (length) by $\|\mathbf{a}\|$, called the **norm** of **a**. It is necessary to say here that in applications of vectors to mechanics, and in some purely geometrical applications of vectors, the norm of vector **r** is often called its **modulus** and written $|\mathbf{r}|$. When this convention is used, because $|\mathbf{r}|$ is a scalar it is usual to denote it by the corresponding ordinary italic letter r, so that $r = |\mathbf{r}|$.

If the base and tip of a vector need to be identified by letters, a vector such as the one from C to D in Fig. 2.4 is written $\underline{CD}$, with underlining used to indicate that a vector is involved, and the ordering of the letters is such that the first shows the

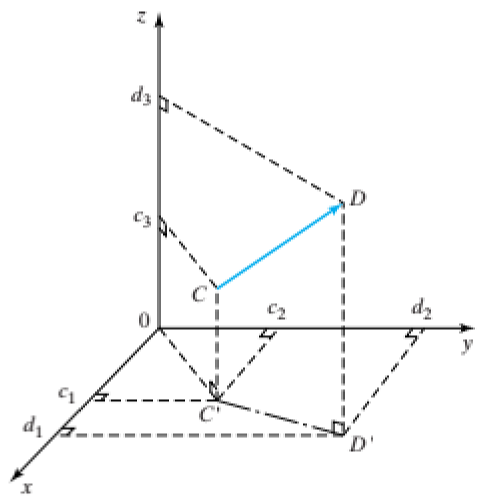


FIGURE 2.4 Vector directed from point C at (c_1, c_2, c_3) to point D at (d_1, d_2, d_3).

base and the second the tip of the vector. Thus, $\underline{CD}$ and $\underline{DC}$ are vectors of equal magnitude but opposite sense, and when these vectors are represented by arrows, the arrows are parallel and of equal length, but point in opposite directions.

EXAMPLE 2.1

If, in Fig. 2.4, C is the point $(-3, 4, 9)$ and D the point $(2, 5, 7)$, the vector $\underline{CD}$ has components $2-(-3)=5$, $5-4=1$, and $7-9=-2$, and so is represented by the ordered number triple $(5, 1, -2)$, whereas vector $\underline{DC}$ has components -5, -1, and 2 and is represented by the ordered number triple $(-5, -1, 2)$.

Having illustrated the concepts of scalars and vectors using some familiar examples, we now develop the algebra of vectors in rather more general terms.

Vectors

A **vector** quantity $\mathbf{a}$ is an ordered number triple (a_1, a_2, a_3) in which a_1, a_2, and a_3 are real numbers, and we shall write $\mathbf{a}=(a_1, a_2, a_3)$. The numbers a_1, a_2, and a_3, in this order, are called the first, second, and third **components** of vector $\mathbf{a}$ or, equivalently, its x-, y-, and z-components.

Null vector

The **null (zero)** vector, written $\mathbf{0}$, has neither magnitude nor direction and is the ordered number triple $\mathbf{0}=(0, 0, 0)$.

Equality of vectors

Two vectors $\mathbf{a}=(a_1, a_2, a_3)$ and $\mathbf{b}=(b_1, b_2, b_3)$ are **equal**, written $\mathbf{a}=\mathbf{b}$, if, and only if, $a_1=b_1$, $a_2=b_2$, and $a_3=b_3$.

EXAMPLE 2.2

If $\mathbf{a}=(a_1, -5, 6)$, $\mathbf{b}=(3, b_2, b_3)$ and $\mathbf{c}=(3, -5, 1)$, then $\mathbf{a}=\mathbf{b}$ if $a_1=3$, $b_2=-5$ and $b_3=6$, and $\mathbf{b}=\mathbf{c}$ if $b_2=-5$ and $b_3=1$, but $\mathbf{a}\neq\mathbf{c}$ for any choice of a_1 because $6\neq 1$.

Norm of a vector

The **norm** of vector $\mathbf{a}=(a_1, a_2, a_3)$, denoted by $\|\mathbf{a}\|$, is the non-negative real number

$$\|\mathbf{a}\| = \left(a_1^2+a_2^2+a_3^2\right)^{1/2},$$

and in geometrical terms $\|\mathbf{a}\|$ is the *length* of vector $\mathbf{a}$. The norm of the null vector $\mathbf{0}$ is $\|\mathbf{0}\|=0$. For example, if $\mathbf{a}$ is in m/sec, "length" of $\mathbf{a}$ is in m/sec.

EXAMPLE 2.3

If $\mathbf{a}=(1, -3, 2)$, then $\|\mathbf{a}\| = [1^2+(-3)^2+2^2]^{1/2}=\sqrt{14}$, as illustrated in Fig. 2.5.

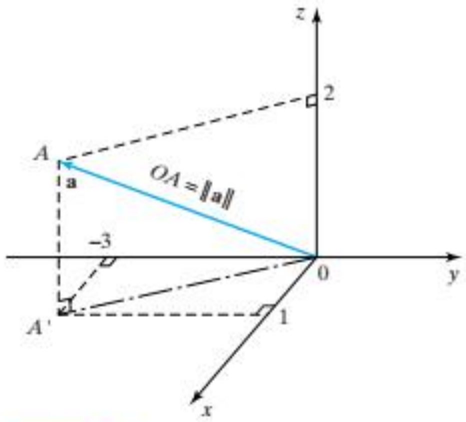


FIGURE 2.5 Vector **a** and its norm ‖**a**‖.

The sum of two vectors

If $\mathbf{a} = (a_1, a_2, a_3)$ and $\mathbf{b} = (b_1, b_2, b_3)$ have the same dimensions, say, both are m/sec, their **sum**, written $\mathbf{a} + \mathbf{b}$, is defined as the ordered number triple (vector) obtained by adding corresponding components of **a** and **b** to give

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

EXAMPLE 2.4

If $\mathbf{a} = (1, 2, -5)$ and $\mathbf{b} = (-2, 2, 4)$, then

$$\mathbf{a} + \mathbf{b} = (1 + (-2), 2 + 2, -5 + 4) = (-1, 4, -1).$$

Multiplying a vector by a scalar

Let $\mathbf{a} = (a_1, a_2, a_3)$ and λ be an arbitrary real number. Then the product $\lambda\mathbf{a}$ is defined as the vector

$$\lambda\mathbf{a} = (\lambda a_1, \lambda a_2, \lambda a_3).$$

EXAMPLE 2.5

Let $\mathbf{a} = (2, -3, 5)$, $\mathbf{b} = (-1, 2, 4)$. Then $2\mathbf{a} = (4, -6, 10)$, $4\mathbf{b} = (-4, 8, 16)$, and $2\mathbf{a} + 4\mathbf{b} = (4 + (-4), -6 + 8, 10 + 16) = (0, 2, 26)$.

This definition of the product of a vector and a scalar, called **scaling** a vector, shows that when vector **a** is multiplied by a scalar λ, the norm of **a** is multiplied by $|\lambda|$, because

$$\|\lambda\mathbf{a}\| = \left(\lambda^2 a_1^2 + \lambda^2 a_2^2 + \lambda^2 a_3^2\right)^{1/2} = |\lambda| \cdot \|\mathbf{a}\|.$$

It also follows from the definition that the sense of vector **a** is reversed when it is multiplied by -1, though its norm is left unaltered. The definition of the **difference**

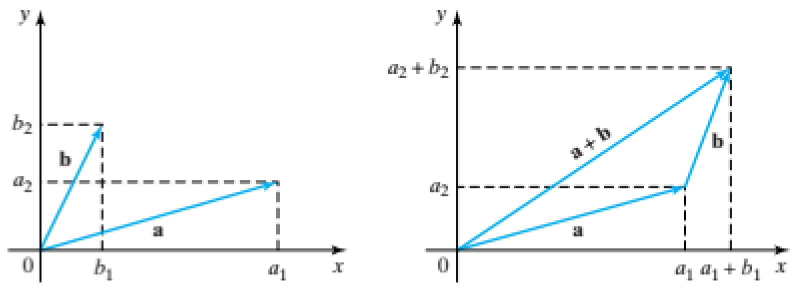


FIGURE 2.6 The vector sum $\mathbf{a} + \mathbf{b}$.

of two vectors is seen to be contained in the definition of their sum, because $\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$. In particular, when $\mathbf{a} = \mathbf{b}$, we find that that $\mathbf{a} - \mathbf{a} = \mathbf{0}$, showing that $-\mathbf{a}$ is the *additive inverse* of $\mathbf{a}$.

The geometrical interpretations of the sum $\mathbf{a} + \mathbf{b}$, the difference $\mathbf{a} - \mathbf{b}$, and the **scaled** vector $\lambda\mathbf{a}$ in terms of their components are shown in Figs. 2.6 to 2.8, though to simplify the diagrams only the two-dimensional cases are illustrated. This involves no loss of generality, because it is always possible to choose the (x, y)-plane to coincide with the plane containing the vectors $\mathbf{a}$ and $\mathbf{b}$.

Vector Addition by the Triangle Rule

Consideration of Fig. 2.6 shows that the addition of vector $\mathbf{b}$ to vector $\mathbf{a}$ is obtained geometrically by translating vector $\mathbf{b}$ until its base is located at the tip of vector $\mathbf{a}$, and then the vector representing the sum $\mathbf{a} + \mathbf{b}$ has its base at the base of vector $\mathbf{a}$ and its tip at the tip of the repositioned vector $\mathbf{b}$. Because of the triangle involving vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{a} + \mathbf{b}$, this geometrical interpretation of a vector sum is called the **triangle rule** for vector addition. The triangle rule also applies to the difference of two vectors, as may be seen by considering Fig. 2.7, because after obtaining $-\mathbf{b}$ from $\mathbf{b}$ by reversing its sense, the difference $\mathbf{a} - \mathbf{b}$ can be written as the vector sum $\mathbf{a} + (-\mathbf{b})$, where $-\mathbf{b}$ is added to vector $\mathbf{a}$ by means of the triangle rule.

triangle rule for addition

The algebraic results discussed so far concerning the addition and scaling of vectors, together with some of their consequences, are combined to form the following theorem.

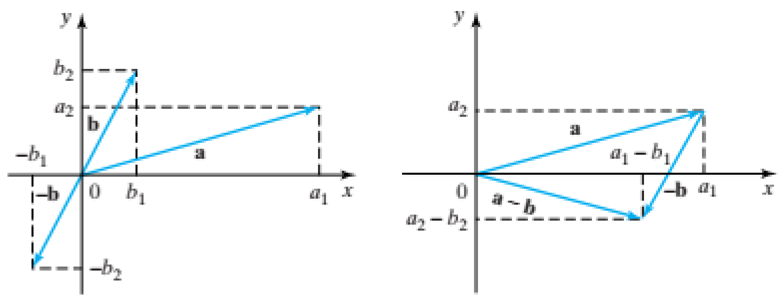


FIGURE 2.7 The vector difference $\mathbf{a} - \mathbf{b}$.

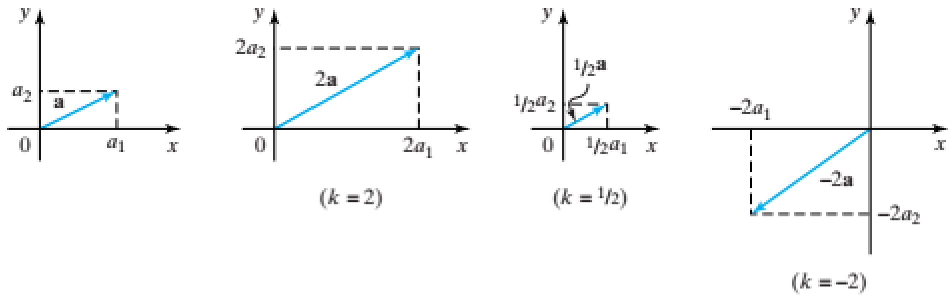


FIGURE 2.8 The vector $k\mathbf{a}$ for different values of k.

THEOREM 2.1

Addition and scaling of vectors Let $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ be arbitrary vectors and let α and β be arbitrary real numbers. Then:

1. $\mathbf{P} + \mathbf{Q} = \mathbf{Q} + \mathbf{P}$ (vector addition is **commutative**);
2. $\mathbf{P} + \mathbf{0} = \mathbf{0} + \mathbf{P} = \mathbf{P}$ (**0** is the **identity element** in vector addition);
3. $(\mathbf{P} + \mathbf{Q}) + \mathbf{R} = \mathbf{P} + (\mathbf{Q} + \mathbf{R})$ (vector addition is **associative**);
4. $\alpha(\mathbf{P} + \mathbf{Q}) = \alpha\mathbf{P} + \alpha\mathbf{Q}$ (multiplication by a scalar is **distributive** over **vector addition**);
5. $(\alpha\beta)\mathbf{P} = \alpha(\beta\mathbf{P}) = \beta(\alpha\mathbf{P})$ (multiplication of a vector by a product of scalars is **associative**);
6. $(\alpha + \beta)\mathbf{P} = \alpha\mathbf{P} + \beta\mathbf{P}$ (multiplication of a vector by a sum of scalars is **distributive**);
7. $\|\alpha\mathbf{P}\| = |\alpha| \cdot \|\mathbf{P}\|$ (scaling $\mathbf{P}$ by α scales the norm of $\mathbf{P}$ by $|\alpha|$).

Proof The results of this theorem are all immediate consequences of the above definitions so as the proofs of results 1 to 6 are all very similar, and result 7 has already been established, we only prove result 4.

$$\begin{aligned}
\text{Let } \mathbf{P} = (p_1, p_2, p_3) \text{ and } \mathbf{Q} &= (q_1, q_2, q_3);\ \text{then} \\
\alpha(\mathbf{P} + \mathbf{Q}) &= \alpha(p_1 + q_1, p_2 + q_2, p_3 + q_3) \\
&= \alpha[(p_1, p_2, p_3) + (q_1, q_2, q_3)] \\
&= \alpha(p_1, p_2, p_3) + \alpha(q_1, q_2, q_3) \\
&= \alpha\mathbf{P} + \alpha\mathbf{Q},
\end{aligned}$$

as was to be shown. ■

The Representation of Vectors in Terms of the Unit Vectors i, j, and k

The components of a vector, together with vector addition, can be used to describe vectors in a very convenient way. The idea is simple, and it involves using the standard convention that $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are vectors of unit length that point in the positive sense along the x-, y-, and z-axes, respectively. Vectors such as $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ that have a unit norm (length) are called **unit vectors**, so $\|\mathbf{i}\| = \|\mathbf{j}\| = \|\mathbf{k}\| = 1$.

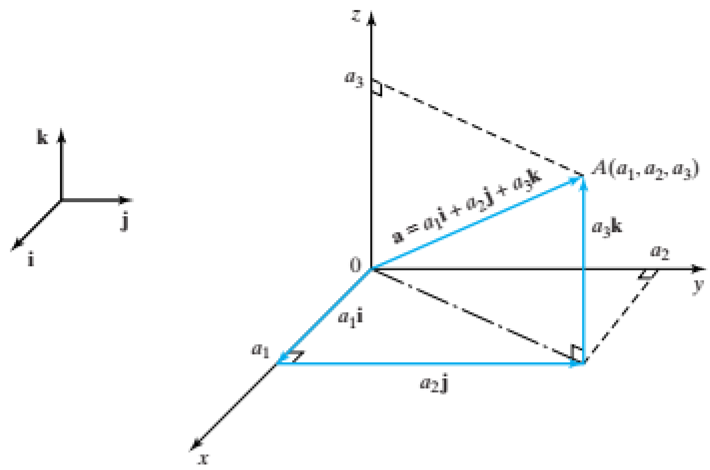


FIGURE 2.9 Vector **a** in terms of the unit vectors **i**, **j**, and **k**.

An arbitrary vector **a** can be represented by an "arrow," with its base at the origin and its tip at the point A with cartesian coordinates (a_1, a_2, a_3) where, of course, a_1, a_2, and a_3 are also the components of **a**. Consequently, scaling the unit vectors **i**, **j**, and **k** by the respective x, y, and z components a_1, a_2, and a_3 of **a**, followed by vector addition of these three vectors, shows that **a** can be written

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}, \tag{1}$$

as can be seen from Fig. 2.9. The representation of vector **a** in terms of the unit vectors **i**, **j**, and **k** in (1), and the ordered triple notation, are equivalent, so

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = (a_1, a_2, a_3). \tag{2}$$

position vector

In some applications a vector defines a point in space, so vectors of this type are called **position vectors**. The symbol **r** is normally used for a position vector, so if point P with coordinates (x, y, z) is a general point in space, as in Fig. 2.10, its

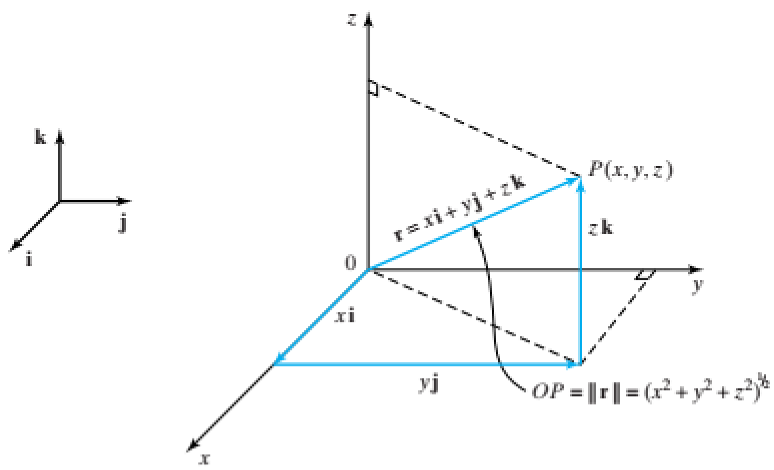


FIGURE 2.10 Position vector of a general point P in space.

position vector relative to the origin is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \tag{3}$$

and its norm (length) is

$$\|\mathbf{r}\| = (x^2 + y^2 + z^2)^{1/2}. \tag{4}$$

EXAMPLE 2.6 (a) Find the distance of point P from the origin given that its position vector is $\mathbf{r} = 2\mathbf{i} + 4\mathbf{j} - 3\mathbf{k}$. (b) If a general point P in space has position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, describe the surface defined by $\|\mathbf{r}\| = 3$ and find its cartesian equation.

Solution **(a)** As $\mathbf{r}$ is the position vector of P relative to the origin, the distance of point P from the origin is $\|\mathbf{r}\| = [2^2 + 4^2 + (-3)^2]^{1/2} = \sqrt{29}$.
(b) As $\|\mathbf{r}\| = 3$ (constant), it follows that the required surface is one for which every point lies at a distance 3 from the origin, so the surface must be a sphere of radius 3 centered on the origin. As $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is the general position vector of a point on this sphere, the result $\|\mathbf{r}\| = 3$ is equivalent to $(x^2 + y^2 + z^2)^{1/2} = 3$, so the cartesian equation of the sphere is $x^2 + y^2 + z^2 = 9$. ■

Because of the equivalence of the ordered number triple notation and the representation of vectors in terms of the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ given in (2), both systems obey the same rules governing the addition and scaling of vectors in terms of their components. Thus, the following rules apply to the combination of any two vectors $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ expressed in terms of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, and an arbitrary real number λ.

The sum $\mathbf{a} + \mathbf{b}$ is given by

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}. \tag{5}$$

The product $\lambda\,\mathbf{a}$ is given by

$$\lambda\,\mathbf{a} = \lambda a_1\,\mathbf{i} + \lambda a_2\,\mathbf{j} + \lambda a_3\,\mathbf{k}. \tag{6}$$

The norm of scaled vector $\lambda\,\mathbf{a}$ is given by

$$\begin{aligned} \|\lambda\,\mathbf{a}\| &= |\lambda| \cdot \|\mathbf{a}\| \\ &= |\lambda|\left(a_1^2 + a_2^2 + a_3^2\right)^{1/2}. \end{aligned} \tag{7}$$

EXAMPLE 2.7 If $\mathbf{a} = 5\mathbf{i} + \mathbf{j} - 3\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} - 2\mathbf{j} - 7\mathbf{k}$, find (a) $\mathbf{a} + \mathbf{b}$, (b) $\mathbf{a} - \mathbf{b}$, (c) $2\mathbf{a} + \mathbf{b}$, and (d) $|-2\mathbf{a}|$.

Solution

(a)

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= (5\mathbf{i} + \mathbf{j} - 3\mathbf{k}) + (2\mathbf{i} - 2\mathbf{j} - 7\mathbf{k}) \\ &= (5 + 2)\mathbf{i} + (1 - 2)\mathbf{j} + (-3 - 7)\mathbf{k} \\ &= 7\mathbf{i} - \mathbf{j} - 10\mathbf{k}. \end{aligned}$$

(b)

$$\begin{aligned} \mathbf{a} - \mathbf{b} &= (5\mathbf{i} + \mathbf{j} - 3\mathbf{k}) - (2\mathbf{i} - 2\mathbf{j} - 7\mathbf{k}) \\ &= (5 - 2)\mathbf{i} + (1 - (-2))\mathbf{j} + (-3 - (-7))\mathbf{k} \\ &= 3\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}. \end{aligned}$$

(c)
$$\begin{aligned}2\mathbf{a}+\mathbf{b} &= 2(5\mathbf{i}+\mathbf{j}-3\mathbf{k})+(2\mathbf{i}-2\mathbf{j}-7\mathbf{k})\\ &= (10\mathbf{i}+2\mathbf{j}-6\mathbf{k})+(2\mathbf{i}-2\mathbf{j}-7\mathbf{k})\\ &= (10+2)\mathbf{i}+(2+(-2))\mathbf{j}+(-6+(-7))\mathbf{k}\\ &= 12\mathbf{i}-13\mathbf{k}.\end{aligned}$$

(d)
$$|-2\mathbf{a}| = [(-10)^2+(-2)^2+6^2]^{1/2} = 2\sqrt{35}$$

or, equivalently,

$$|-2\mathbf{a}| = |-2|\cdot\|\mathbf{a}\| = 2\|\mathbf{a}\| = 2[5^2+1^2+(-3)^2]^{1/2} = 2\sqrt{35}.$$ ■

Finding a Unit Vector in the Direction of an Arbitrary Vector

It is often necessary to find a unit vector in the direction of an arbitrary vector $\mathbf{a} = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$. This is accomplished by dividing $\mathbf{a}$ by its norm $\|\mathbf{a}\|$, because the vector $\mathbf{a}/\|\mathbf{a}\|$ has the same sense as $\mathbf{a}$ and its norm is 1. It is convenient to use a symbol related to an arbitrary vector $\mathbf{a}$ to indicate the unit vector in its direction, so from now on such a vector will be denoted by $\hat{\mathbf{a}}$, read "**a** hat." So if $\mathbf{a} = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$,

$$\begin{aligned}\hat{\mathbf{a}} = \mathbf{a}/\|\mathbf{a}\| &= (a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})/\left(a_1^2+a_2^2+a_3^2\right)^{1/2}\\ &= (a_1/a)\mathbf{i}+(a_2/a)\mathbf{j}+(a_3/a)\mathbf{k}, \quad \text{with } a = \left(a_1^2+a_2^2+a_3^2\right)^{1/2}.\end{aligned} \tag{8}$$

As the symbols $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are used exclusively for the unit vectors in the x-, y-, and z-directions, it is not necessary to write $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$.

The relationship between $\mathbf{a}$, $\hat{\mathbf{a}}$, and $\|\mathbf{a}\|$ can be put in the useful form

$$\mathbf{a} = \|\mathbf{a}\|\hat{\mathbf{a}}, \tag{9}$$

showing that a general vector $\mathbf{a}$ can always be written as the unit vector $\hat{\mathbf{a}}$ scaled by $\|\mathbf{a}\|$. Unless otherwise stated, $\mathbf{a} \neq \mathbf{0}$.

EXAMPLE 2.8

Find a unit vector in the direction of $\mathbf{a} = 3\mathbf{i}+2\mathbf{j}+5\mathbf{k}$.

Solution As $\|\mathbf{a}\| = (3^2+2^2+5^2)^{1/2} = \sqrt{38}$, it follows that

$$\hat{\mathbf{a}} = \mathbf{a}/\|\mathbf{a}\| = (3/\sqrt{38})\mathbf{i}+(2/\sqrt{38})\mathbf{j}+(5/\sqrt{38})\mathbf{k}.$$ ■

EXAMPLE 2.9

It is known from experiments in mechanics that forces are vector quantities and so combine according to the laws of vector algebra. Use this fact to find the sum and difference of a force of 9 units in the direction of $2\mathbf{i}+\mathbf{j}-2\mathbf{k}$ and a force of 10 units in the direction of $4\mathbf{i}-3\mathbf{j}$, and determine the magnitudes of these forces.

Solution We will use the convention that a unit vector represents a force of 1 unit. Let $\mathbf{F}$ be the force of 9 units. Then as $\|2\mathbf{i}+\mathbf{j}-2\mathbf{k}\| = [2^2+1^2+(-2)^2]^{1/2} = 3$, the unit vector in the direction of $\mathbf{F}$ is

$$\hat{\mathbf{F}} = (1/3)(2\mathbf{i}+\mathbf{j}-2\mathbf{k}) = (2/3)\mathbf{i}+(1/3)\mathbf{j}-(2/3)\mathbf{k},$$

so $\mathbf{F} = 9\hat{\mathbf{F}} = 6\mathbf{i}+3\mathbf{j}-6\mathbf{k}$ units.

Similarly, let $\mathbf{G}$ be the force of 10 units. Then as $\|4\mathbf{i} - 3\mathbf{j}\| = 5$, the unit vector in the direction of $\mathbf{G}$ is

$$\hat{\mathbf{G}} = (1/5)(4\mathbf{i} - 3\mathbf{j}) = (4/5)\mathbf{i} - (3/5)\mathbf{j},$$

so $\mathbf{G} = 10\hat{\mathbf{G}} = 8\mathbf{i} - 6\mathbf{j}$ units.

Combining these results shows that $\mathbf{F} + \mathbf{G} = 14\mathbf{i} - 3\mathbf{j} - 6\mathbf{k}$ units, and $\mathbf{F} - \mathbf{G} = -2\mathbf{i} + 9\mathbf{j} - 6\mathbf{k}$ units, from which it follows that the magnitudes of the forces are given by

$$\|\mathbf{F} + \mathbf{G}\| = \sqrt{241} \text{ units and } \|\mathbf{F} - \mathbf{G}\| = 11 \text{ units.}$$ ■

Equality of vectors expressed in terms of unit vectors

As the difference of two equal and opposite vectors is the null vector $\mathbf{0}$, this shows that if $\mathbf{a} = \mathbf{b}$, where $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, then the respective components of vectors $\mathbf{a}$ and $\mathbf{b}$ must be equal, leading to the result that

$$\mathbf{a} = \mathbf{b} \text{ if, and only if, } a_1 = b_1, a_2 = b_2, \text{ and } a_3 = b_3. \tag{10}$$

Simple Geometrical Applications of Vectors

Although our use of vectors will be mainly in connection with the calculus, the following simple geometrical applications are helpful because they illustrate basic vector arguments and properties.

Although we have seen how an arbitrary vector can be expressed in terms of unit vectors associated with a cartesian coordinate system, it must be remembered that the fundamental concept of a vector and its algebra is independent of a coordinate system. Because of this, it is often possible to use the rules governing elementary vector algebra given in Theorem 2.1 to establish equations in a purely vectorial manner, without the need to appeal to any coordinate system. Once a general vector equation has been established, the representation of the vectors involved in terms of their components and the unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ can be used to convert the vector equation into the equivalent cartesian equations.

The purely vectorial approach to geometrical problems is well illustrated by finding the vector $\underline{AB}$ in terms of the position vectors of points A and B, and then using the result to find the position vector of the mid-point of $\underline{AB}$. After this, the purely vectorial derivation of a geometrical result followed by its interpretation in cartesian form will be illustrated by finding the equation of a straight line in three space dimensions.

Vector $\underline{AB}$ in terms of the position vectors of A and B

Let $\mathbf{a}$ and $\mathbf{b}$ be the position vectors of points A and B relative to an origin O, as shown in Fig. 2.11.

An application of the triangle rule for the addition of vectors gives

$$\underline{OA} + \underline{AB} = \underline{OB},$$

but $\underline{OA} = \mathbf{a}$ and $\underline{OB} = \mathbf{b}$, so

$$\mathbf{a} + \underline{AB} = \mathbf{b},$$

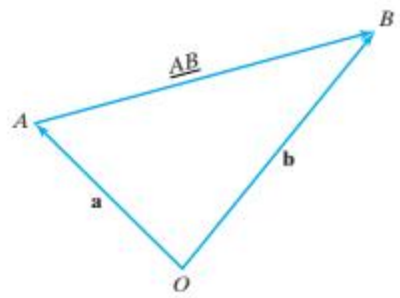


FIGURE 2.11 Vectors **a**, **b**, and $\underline{AB}$.

giving

$$\underline{AB} = \mathbf{b} - \mathbf{a}. \tag{11}$$

When expressed in words, this simple but useful result asserts that vector $\underline{AB}$ is obtained by subtracting the position vector **a** of point A from the position vector **b** of point B.

EXAMPLE 2.10

Find the position vector of the mid-point of $\underline{AB}$ if point A has position vector **a** and point B has position vector **b** relative to an origin O.

Solution Let point C, with position vector **c** relative to origin O, be the mid-point of $\underline{AB}$, as shown in Fig. 2.12.
By the triangle rule,

$$\underline{OA} + \underline{AC} = \underline{OC},$$

but $\underline{OA} = \mathbf{a}$, and from (11) $\underline{AC} = (1/2)(\mathbf{b} - \mathbf{a})$, so

$$\underline{OC} = \mathbf{a} + (1/2)(\mathbf{b} - \mathbf{a}),$$

so the required result is

$$\mathbf{c} = \underline{OC} = (1/2)(\mathbf{b} + \mathbf{a}).$$

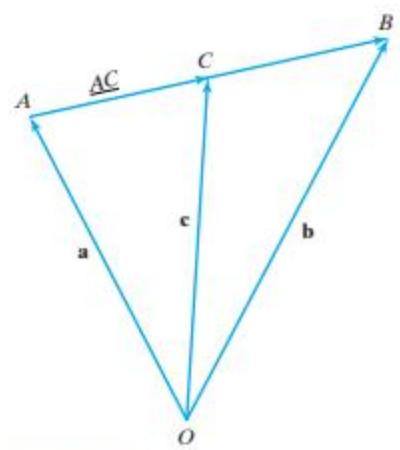


FIGURE 2.12 C is the mid-point of $\underline{AB}$.

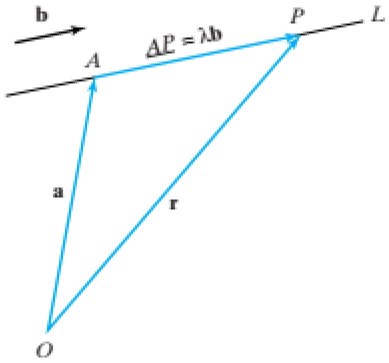


FIGURE 2.13 The straight line L.

The vector and cartesian equations of a straight line

Let line L be a straight line through point A with position vector $\mathbf{a}$ relative to an origin O, and let the line be parallel to a vector $\mathbf{b}$. If P is an arbitrary point on line L with position vector $\mathbf{r}$ relative to O, an application of the triangle rule for vector addition to the vectors shown in Fig. 2.13 gives

$$\mathbf{r} = \underline{OA} + \underline{AP}.$$

vector equation of straight line

But $\underline{OA} = \mathbf{a}$, and as $\underline{AP}$ is parallel to $\mathbf{b}$, a number λ can always be found such that $\underline{AP} = \lambda\mathbf{b}$, so the **vector equation** of line L becomes

$$\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}. \tag{12}$$

Notice that result (12) determines all points P on L if λ is taken to be a number in the interval $-\infty < \lambda < \infty$.

The cartesian equations of line L follow by setting $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ in result (12), and then using the definition of equality of vectors given in (10) to obtain the corresponding three scalar cartesian equations. Proceeding in this way we find that

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} + \lambda(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}),$$

cartesian and standard form of straight line

so equating corresponding components of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ on each side of this equation brings us to the required **cartesian equations** for L in the form

$$x_1 = a_1 + \lambda b_1, \quad x_2 = a_2 + \lambda b_2, \quad x_3 = a_3 + \lambda b_3. \tag{13}$$

An equivalent form of these equations is obtained by solving each equation for λ and equating the results to get

$$\frac{x - a_1}{b_1} = \frac{y - a_2}{b_2} = \frac{z - a_3}{b_3} = \lambda. \tag{14}$$

This is the **standard form** (also called the **canonical form**) of the cartesian equations of a straight line. It is important to notice that when written in standard form the coefficients of x, y, and z are all *unity*. Once the equation of a straight line is written in standard form, equating each numerator to zero determines the components (a_1, a_2, a_3) of a position vector of a point on the line, while the denominators in the order (b_1, b_2, b_3) determine the components of a vector parallel to the line.

EXAMPLE 2.11 A straight line L is given in the form

$$\frac{2x-3}{4} = \frac{3-y}{2} = \frac{z+1}{3}.$$

Find the position vector of a point on L and a vector parallel to L.

Solution When the equation is written in standard form it becomes

$$\frac{x-3/2}{2} = \frac{y-3}{-2} = \frac{z+1}{3} = \lambda.$$

Comparing these equations with (14) shows that $(a_1, a_2, a_3) = (3/2, 3, -1)$ and $\mathbf{b} = (b_1, b_2, b_3) = (2, -2, 3)$. So the position vector of a point on the line is $\mathbf{a} = (3/2)\mathbf{i} + 3\mathbf{j} - \mathbf{k}$, and a vector parallel to the line is $\mathbf{b} = 2\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

Neither of these results is unique, because $\mu\mathbf{b}$ is also parallel to the line for any scalar $\mu \neq 0$, and any other point on L would suffice. For example, the vector $14\mathbf{i} - 14\mathbf{j} + 21\mathbf{k}$ is also parallel to the line, while setting $\lambda = 2$ leads to the result $(a_1, a_2, a_3) = (11/2, -1, 5)$, corresponding to a different point on the same line, this time with position vector $\mathbf{a} = (11/2)\mathbf{i} - \mathbf{j} + 5\mathbf{k}$. ■

Summary

This section has introduced vectors both as geometrical quantities that can be represented by directed line segments and, using a right-handed system of cartesian axes, as ordered number triples. Definitions of the scaling, addition, and subtraction of vectors have been given, and a general vector has been defined in terms of the set of three unit vectors **i**, **j**, and **k** that lie along the orthogonal cartesian axes $O\{x, y, z\}$. Finally, the vector and cartesian equations of a straight line in space have been derived, and the standard form of the cartesian equations has been introduced from which a vector parallel to the line may be found by inspection.

EXERCISES 2.1

1. Prove Results 1, 3, and 6 of Theorem 2.1.

2. Given that $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j} + 2\mathbf{k}$, and $\mathbf{c} = 3\mathbf{i} + 4\mathbf{j} + \mathbf{k}$, find (a) $\mathbf{a} + 2\mathbf{b} - \mathbf{c}$, (b) a vector $\mathbf{d}$ such that $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = \mathbf{0}$, and (c) a vector $\mathbf{d}$ such that $\mathbf{a} - \mathbf{b} + \mathbf{c} + 3\mathbf{d} = \mathbf{0}$.

3. Given $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - 2\mathbf{j} + \mathbf{k}$, find (a) a vector $\mathbf{c}$ such that $2\mathbf{a} + \mathbf{b} + 2\mathbf{c} = \mathbf{i} + \mathbf{k}$, (b) a vector $\mathbf{c}$ such that $3\mathbf{a} - 2\mathbf{b} + \mathbf{c} = \mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

4. Given that $\mathbf{a} = 3\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j} + 5\mathbf{k}$, and $\mathbf{c} = 2\mathbf{i} + 5\mathbf{j} + 2\mathbf{k}$, find (a) $2\mathbf{a} + 3\mathbf{b} - 3\mathbf{c}$, (b) a vector $\mathbf{d}$ such that $\mathbf{a} + 3\mathbf{b} - 2\mathbf{c} + 3\mathbf{d} = \mathbf{0}$, and (c) a vector $\mathbf{d}$ such that $2\mathbf{a} - 3\mathbf{d} = \mathbf{b} + 4\mathbf{c}$.

5. Given that A and B have the respective position vectors $2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$ and $\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$, find the vector $\underline{AB}$ and a unit vector in the direction of $\underline{AB}$.

6. Given that A and B have the respective position vectors $3\mathbf{i} - \mathbf{j} + 4\mathbf{k}$ and $2\mathbf{i} + \mathbf{j} + \mathbf{k}$, find the vector $\underline{AB}$ and the position vector $\mathbf{c}$ of the mid-point of $\underline{AB}$.

7. Given that A and B have the respective position vectors $\mathbf{a}$ and $\mathbf{b}$, find the position vector of a point P on the line AB located between A and B such that

$$(\text{length } AP)/(\text{length } PB) = m/n, \quad \text{where } m, n > 0$$

are any two real numbers.

8. Find the position vector $\mathbf{r}$ of a point P on the straight line joining point A at $(1, 2, 1)$ and point B at $(3, -1, 2)$ and between A and B such that

$$(\text{length } AP)/(\text{length } PB) = 3/2.$$

9. It is known from Euclidean geometry that the medians of a triangle (lines drawn from a vertex to the mid-point of the opposite side) all meet at a single point P, and that P is two-thirds of the distance along each median from the vertex through which it passes. If the vertices A, B, and C of a triangle have the respective position vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, show that the position vector of P is $(1/3)(\mathbf{a} + \mathbf{b} + \mathbf{c})$.

10. Forces of 1, 2, and 3 units act through the origin along, and in the positive directions of, the respective x-, y-, and z-axes. Find the vector sum $\mathbf{S}$ of these forces, the magnitude $\|\mathbf{S}\|$ of the sum of the vectors, and a unit vector in the direction of $\mathbf{S}$.

11. Forces of 2, 1, and 4 units act through the origin along, and in the positive directions of, the respective x-, y-, and z-axes. Find the vector sum $\mathbf{S}$ of these forces, the magnitude $\|\mathbf{S}\|$ of the sum of the vectors, and a unit vector in the direction of $\mathbf{S}$.

12. A straight line L is given in the form

$$\frac{3x-1}{4} = \frac{2y+3}{2} = \frac{2-3z}{1}.$$

Find the position vectors of two different points on L and a unit vector parallel to L.

13. A straight line L is given in the form

$$\frac{2x+1}{3} = \frac{3y+2}{4} = \frac{2-4z}{-1}.$$

Find position vectors of two different points on L and a unit vector parallel to L.

14. Given that a straight line L_1 passes through the points $(-2, 3, 1)$ and $(1, 4, 6)$, find (a) the position vector of a point on the line and a vector parallel to it, and (b) a straight line L_2 parallel to L_1 that passes through the point $(1, 2, 1)$.

15. Given that a straight line L_1 passes through the points $(3, 2, 4)$ and $(2, 1, 6)$, find (a) the position vector of a point on the line and a vector parallel to it, and (b) a straight line L_2 parallel to L_1 that passes through the point $(-2, 1, 2)$.

16. A straight line has the vector equation $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, where $\mathbf{a} = 3\mathbf{j} + 2\mathbf{k}$, and $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$. Find the cartesian equations of the line and the coordinates of three points that lie on it.

17. A straight line passes through the point $(3, 2, -3)$ parallel to the vector $2\mathbf{i} + 3\mathbf{j} - 3\mathbf{k}$. Find the cartesian equations of the line and the coordinates of three points that lie on it.

18. In mechanics, if a point A moves with velocity $\mathbf{v}_A$ and point B moves with velocity $\mathbf{v}_B$, the velocity $\mathbf{v}_R$ of A **relative** to B (the **relative velocity** of A with respect to B) is defined as $\mathbf{v}_R = \mathbf{v}_A - \mathbf{v}_B$. Power boat A moves northeast at 20 knots and power boat B moves southeast at 30 knots. Find the velocity of boat A relative to boat B, and a unit vector in the direction of the relative velocity.

2.2 The Dot Product (Scalar Product)

A product of two vectors $\mathbf{a}$ and $\mathbf{b}$ can be formed in such a way that the result is a scalar. The result is written $\mathbf{a} \cdot \mathbf{b}$ and called the **dot product** of $\mathbf{a}$ and $\mathbf{b}$. The names **scalar product** and **inner product** are also used in place of the term *dot product.*

Dot Product

dot or scalar product

Let $\mathbf{a}$ and $\mathbf{b}$ be any two vectors that after a translation to bring their bases into coincidence are inclined to one another at an angle θ, as shown in Fig. 2.14, where $0 \le \theta \le \pi$. Then the **dot product** of $\mathbf{a}$ and $\mathbf{b}$ is defined as the number

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \cos\theta.$$

This geometrical definition of the dot product has many uses, but when working with vectors $\mathbf{a}$ and $\mathbf{b}$ that are expressed in terms of their components in the $\mathbf{i}$, $\mathbf{j}$, and

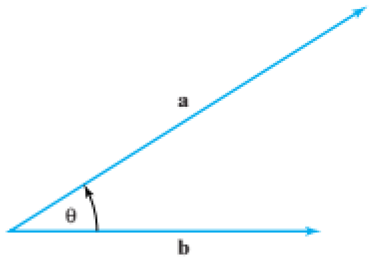


FIGURE 2.14 Vectors **a** and **b** inclined at an angle θ.

k directions, a more convenient form is needed. An equivalent definition that is easier to use is given later in (23).

properties of the dot product

Properties of the dot product

The following results, in which **a** and **b** are any two vectors and λ and μ are any two scalars, are all immediate consequences of the definition of the dot product.

The dot product is commutative

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a} \quad \text{and} \quad \lambda\mathbf{a}\cdot\mu\mathbf{b} = \mu\mathbf{a}\cdot\lambda\mathbf{b} = \lambda\mu\mathbf{a}\cdot\mathbf{b} \tag{15}$$

The dot product is distributive and linear

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c} \quad \text{and} \quad \mathbf{a}\cdot(\lambda\mathbf{b}+\mu\mathbf{c}) = \lambda\mathbf{a}\cdot\mathbf{b}+\mu\mathbf{a}\cdot\mathbf{c}. \tag{16}$$

The angle between two vectors

The angle θ between vectors **a** and **b** is given by

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\cdot\|\mathbf{b}\|}, \quad \text{with } 0 \le \theta \le \pi. \tag{17}$$

Parallel vectors ($\theta = 0$)

If vectors **a** and **b** are parallel, then

$$\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\cdot\|\mathbf{b}\| \quad \text{and, in particular,} \quad \mathbf{a}\cdot\mathbf{a} = \|\mathbf{a}\|^2. \tag{18}$$

Orthogonal vectors ($\theta = \pi/2$)

If vectors **a** and **b** are orthogonal, then

$$\mathbf{a}\cdot\mathbf{b} = 0. \tag{19}$$

Product of unit vectors

If $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are unit vectors, then

$$\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = \cos\theta, \quad \text{with } 0 \le \theta \le \pi. \tag{20}$$

An immediate consequence of properties (15), (19), and (20) is that

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1, \tag{21}$$

and

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{i} = \mathbf{i} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{j} = 0. \tag{22}$$

We now use results (21) and (22) to arrive at a simple expression for the dot product in terms of the components of **a** and **b**. To arrive at the result we set $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ and form the dot product

$$\mathbf{a} \cdot \mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (b_1\mathbf{i} + b_2\,\mathbf{j} + b_3\mathbf{k}).$$

dot product in terms of components

Expanding this product using (15) and (16) and making use of results (21) and (22) brings us to the following *alternative definition* of the **dot product** expressed in terms of the components of **a** and **b**:

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3. \tag{23}$$

Using (23) in (17) produces the following useful expression that can be used to find the angle θ between **a** and **b**:

$$\cos\theta = \frac{a_1b_1 + a_2b_2 + a_3b_3}{\left(a_1^2 + a_2^2 + a_3^2\right)^{1/2}\left(b_1^2 + b_2^2 + b_3^2\right)^{1/2}} \text{ where } 0 \le \theta \le \pi. \tag{24}$$

EXAMPLE 2.12

Find $\mathbf{a} \cdot \mathbf{b}$ and the angle between the vectors **a** and **b**, given that $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} - \mathbf{j} - 2\mathbf{k}$.

Solution $\|\mathbf{a}\| = \sqrt{14}$, $\|\mathbf{b}\| = 3$, and $\mathbf{a} \cdot \mathbf{b} = 1 \cdot 2 + 2 \cdot (-1) + 3 \cdot (-2) = -6$. Using these results in (24) gives

$$\cos\theta = -6/(3\sqrt{14}) = -2/\sqrt{14},$$

so as $0 \le \theta \le \pi$ we see that $\theta = 2.1347$ radians, or $\theta = 122.3°$. ■

projecting a vector onto a line

The projection of a vector onto the line of another vector

The projection of vector **a** onto the line of vector **b** is a scalar, and it is the *signed* length of the geometrical projection of vector **a** onto a line parallel to **b**, with the sign positive for $0 \le \theta < \pi/2$ and negative for $\pi/2 < \theta \le \pi$. This is illustrated in Fig. 2.15, from which it is seen that the signed length of the projection of **a** onto the line of vector **b** is ON, where $ON = \|\mathbf{a}\| \cos\theta$.

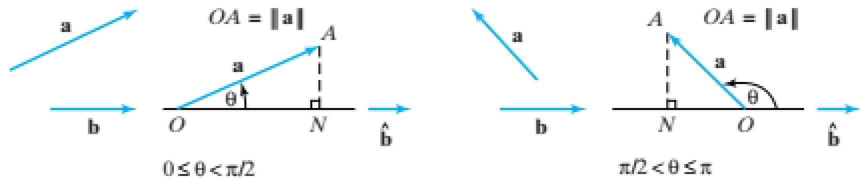


FIGURE 2.15 The projection of vector **a** onto the line of vector **b**.

If $\hat{\mathbf{b}}$ is the unit vector along **b**, then as $\mathbf{a} = \hat{\mathbf{a}}\|\mathbf{a}\|$, and $\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = \cos\theta$, the projection $ON = \|\mathbf{a}\|\cos\theta$ can be written as the dot product

$$ON = \|\mathbf{a}\|\hat{\mathbf{a}} \cdot \hat{\mathbf{b}} = \mathbf{a} \cdot \hat{\mathbf{b}} = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|} \tag{25}$$

EXAMPLE 2.13

Find the strength of the magnetic field vector $\mathbf{H} = 5\mathbf{i} + 3\mathbf{j} + 7\mathbf{k}$ in the direction of $2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, where a unit vector represents one unit of magnetic flux.

Solution We are required to find the projection of vector **H** in the direction of the vector $2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$. Setting $\mathbf{b} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, $\|\mathbf{b}\| = 3$, so $\hat{\mathbf{b}} = (1/3)(2\mathbf{i} - \mathbf{j} + 2\mathbf{k})$, so the strength of the vector **H** in the direction of **b** is

$$\mathbf{H} \cdot \hat{\mathbf{b}} = (1/3)(5\mathbf{i} + 3\mathbf{j} + 7\mathbf{k}) \cdot (2\mathbf{i} - \mathbf{j} + 2\mathbf{k}) = 7.$$

Direction cosines and direction ratios

If $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ is an arbitrary vector, the unit vector $\hat{\mathbf{a}}$ in the direction of **a** is

$$\begin{aligned}\hat{\mathbf{a}} &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})/\|\mathbf{a}\| \\ &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k})/\left(a_1^2 + a_2^2 + a_3^2\right)^{1/2}.\end{aligned} \tag{26}$$

Taking the dot product of **a** with **i**, **j**, and **k**, and setting $l = a_1/(a_1^2 + a_2^2 + a_3^2)^{1/2}$, $m = a_2/(a_1^2 + a_2^2 + a_3^2)^{1/2}$, and $n = a_3/(a_1^2 + a_2^2 + a_3^2)^{1/2}$ gives

$$l = \mathbf{i} \cdot \hat{\mathbf{a}}, \quad m = \mathbf{j} \cdot \hat{\mathbf{a}}, \quad \text{and} \quad n = \mathbf{k} \cdot \hat{\mathbf{a}},$$

so we may write

$$\hat{\mathbf{a}} = l\mathbf{i} + m\mathbf{j} + n\mathbf{k}. \tag{27}$$

The dot product $\hat{\mathbf{a}} \cdot \hat{\mathbf{a}} = l^2 + m^2 + n^2 = (a_1^2 + a_2^2 + a_3^2)/\|\mathbf{a}\|^2$, but $\|\mathbf{a}\|^2 = a_1^2 + a_2^2 + a_3^2$, so

$$l^2 + m^2 + n^2 = 1. \tag{28}$$

The number l is the cosine of the angle β_1 between **a** and the x-axis, the number m is the cosine of the angle β_2 between **a** and the y-axis, and the number n is the cosine of the angle β_3 between **a** and the z-axis, as shown in

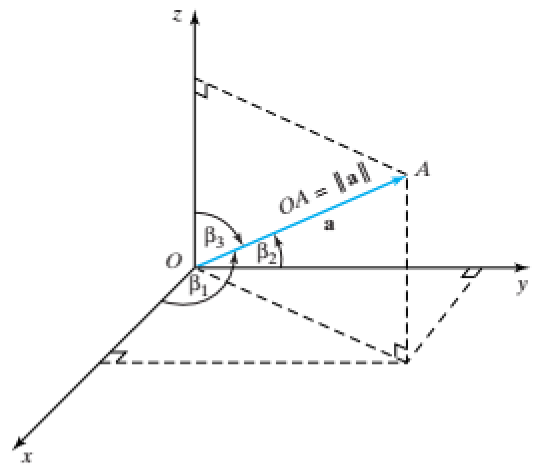


FIGURE 2.16 The angles β_1, β_2, and β_3.

direction cosines

Fig. 2.16. The numbers (l, m, n) are called the **direction cosines** of **a**, because they determine the direction of the unit vector $\hat{\mathbf{a}}$ that is parallel to **a**.

Notice that when any two of the three direction cosines l, m, and n of a vector **a** are given, the third is related to them by

$$l^2 + m^2 + n^2 = 1.$$

Because of result (27) it is always possible to write

$$\mathbf{a} = \|\mathbf{a}\|(l\mathbf{i} + m\mathbf{j} + n\mathbf{k}), \tag{29}$$

where l, m, and n are the direction cosines of **a**.

As the components a_1, a_2, and a_3 of **a** are *proportional* to the direction cosines, they are called the **direction ratios** of **a**.

direction ratios

EXAMPLE 2.14

Find the direction cosines and direction ratios of $\mathbf{a} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.

Solution As $\|\mathbf{a}\| = \sqrt{14}$, the direction cosines are $l = 3/\sqrt{14}$, $m = 1/\sqrt{14}$, and $n = -2/\sqrt{14}$. The direction ratios of **a** are 3, 1, and -2, or any nonnegative multiple of these three numbers such as $15/\sqrt{14}$, $5/\sqrt{14}$, and $-10/\sqrt{14}$. ■

The triangle inequality

The following result will be needed in the proof of the triangle inequality that is to follow. The absolute value of $\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\cdot\|\mathbf{b}\|\cos\theta$ is

$$|\mathbf{a}\cdot\mathbf{b}| = \|\mathbf{a}\|\cdot\|\mathbf{b}\||\cos\theta|,$$

but $|\cos\theta| \le 1$, so using this in the above result we obtain the **Cauchy–Schwarz inequality**,

$$|\mathbf{a}\cdot\mathbf{b}| \le \|\mathbf{a}\|\cdot\|\mathbf{b}\|. \tag{30}$$

THEOREM 2.2

The triangle inequality If $\mathbf{a}$ and $\mathbf{b}$ are any two vectors, then

$$\|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$$

Proof From (18) we have

$$\begin{aligned}\|\mathbf{a}+\mathbf{b}\|^2 &= (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) = \mathbf{a}\cdot\mathbf{a} + 2\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{b}\\ &= \|\mathbf{a}\|^2 + 2\mathbf{a}\cdot\mathbf{b} + \|\mathbf{b}\|^2,\end{aligned}$$

but $\mathbf{a}\cdot\mathbf{b} \leq |\mathbf{a}\cdot\mathbf{b}|$, so from the Cauchy–Schwarz inequality (30)

$$\begin{aligned}\|\mathbf{a}+\mathbf{b}\|^2 &\leq \|\mathbf{a}\|^2 + 2\|\mathbf{a}\|\cdot\|\mathbf{b}\| + \|\mathbf{b}\|^2\\ &= (\|\mathbf{a}\| + \|\mathbf{b}\|)^2.\end{aligned}$$

Taking the positive square root of this last result, we obtain the triangle inequality

$$\|\mathbf{a}+\mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|. \qquad \blacksquare$$

The triangle inequality will be generalized in Section 2.5, but in its present form it is the vector equivalent of the Euclidean theorem that "the sum of the lengths of any two sides of a triangle is greater than or equal to the length of the third side," and it is from this theorem that the inequality derives its name.

Equation of a Plane

When working with the vector calculus it is sometimes necessary to consider a plane that is locally tangent to a point on a surface in space so it will be useful to derive the general equation of a plane in both its vector and cartesian forms.

A plane Π can be defined by specifying a fixed point belonging to the plane and a vector $\mathbf{n}$ that is perpendicular to the plane. This follows because if $\mathbf{n}$ is perpendicular at a point on the plane, it must be perpendicular at every point on the plane. Any vector $\mathbf{n}$ that is perpendicular to a plane is called a **normal** to the plane. Clearly a normal to a plane is not unique, because a plane has two sides, so if a normal $\mathbf{n}$ is directed away from one side of the plane, the vector $-\mathbf{n}$ is a normal directed away from the other side. Both $\mathbf{n}$ and $-\mathbf{n}$ can be scaled by any nonzero number and still remain normals; consequently, if $\mathbf{n}$ is a normal to a plane, so also are all vectors of the form $\lambda\mathbf{n}$, with $\lambda \neq 0$ any real number.

Let a fixed point A on plane Π with normal $\mathbf{n}$ have position vector $\mathbf{a}$ relative to an origin O, and let P be a general point on plane Π with position vector $\mathbf{r}$ relative to O. Then, as may be seen from Fig. 2.17, the vector $\mathbf{r}-\mathbf{a}$ lies in the plane, and so is perpendicular (normal) to $\mathbf{n}$. Forming the dot product of $\mathbf{n}$ and $\mathbf{r}-\mathbf{a}$, and using (19), shows that the **vector equation** of plane Π is

vector equation of a plane

$$\mathbf{n}\cdot(\mathbf{r}-\mathbf{a}) = 0, \tag{31}$$

or, equivalently,

$$\mathbf{n}\cdot\mathbf{r} = \mathbf{n}\cdot\mathbf{a}. \tag{32}$$

cartesian equation of a plane

The **cartesian form** of this equation follows by considering a general point with coordinates (x, y, z) on plane Π, setting $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$,

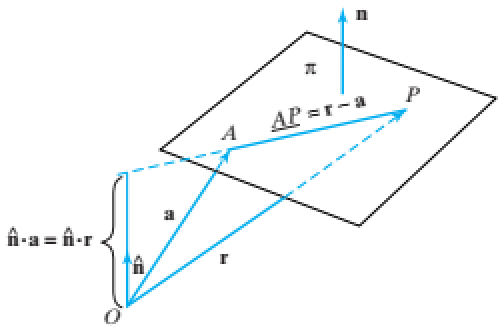


FIGURE 2.17 Plane Π with normal **n** passing through point A.

and $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$, and then substituting into (32) to get

$$(n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}) \cdot (x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = (n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}) \cdot (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}).$$

Taking the dot products and using results (21) and (22) show the cartesian equation of plane Π to be

$$n_1x + n_2y + n_3z = n_1a_1 + n_2a_2 + n_3a_3 = d, \text{ a constant.} \tag{33}$$

EXAMPLE 2.15 Find the cartesian equation of the plane through the point $(2, 5, 3)$ with normal $3\mathbf{i} + 2\mathbf{j} - 7\mathbf{k}$.

Solution Here $n_1 = 3$, $n_2 = 2$, $n_3 = -7$ and $a_1 = 2$, $a_2 = 5$, and $a_3 = 3$, so substituting into (33) shows the plane has the equation

$$3x + 2y - 7z = -5.$$

Summary

This section has introduced the dot or scalar product of two vectors in geometrical terms and, more conveniently for calculations, in terms of the components of the two vectors involved. The applications given include the important operation of projecting a vector onto the line of another vector and the derivation of the vector equation and cartesian equation of a plane.

EXERCISES 2.2

1. Find the dot products of the following pairs of vectors:
(a) $\mathbf{i} - \mathbf{j} + 3\mathbf{k}$, $2\mathbf{i} + 3\mathbf{j} + \mathbf{k}$.
(b) $2\mathbf{i} - \mathbf{j} + 4\mathbf{k}$, $-\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.
(c) $\mathbf{i} + \mathbf{j} - 3\mathbf{k}$, $2\mathbf{i} + \mathbf{j} + \mathbf{k}$.

2. Find the dot products of the following pairs of vectors:
(a) $\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$, $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.
(b) $3\mathbf{i} + \mathbf{j} + 2\mathbf{k}$, $4\mathbf{i} - 3\mathbf{j} + \mathbf{k}$.
(c) $5\mathbf{i} - 3\mathbf{j} + 3\mathbf{k}$, $2\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$.

3. Find which of the following pairs of vectors are orthogonal:
(a) $3\mathbf{i} + 2\mathbf{j} - 6\mathbf{k}$, $-9\mathbf{i} - 6\mathbf{j} + 18\mathbf{k}$.
(b) $3\mathbf{i} - \mathbf{j} + 7\mathbf{k}$, $3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$.
(c) $2\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{i} + \mathbf{j} - \mathbf{k}$.
(d) $\mathbf{i} + \mathbf{j} - 3\mathbf{k}$, $2\mathbf{i} + \mathbf{j} + \mathbf{k}$.

4. Find which, if any, of the following pairs of vectors are orthogonal:
(a) $2\mathbf{i} + \mathbf{j} + \mathbf{k}$, $8\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.
(b) $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, $2\mathbf{i} - 2\mathbf{j} - 3\mathbf{k}$.
(c) $\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$, $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$.
(d) $\mathbf{i} + \mathbf{j}$, $2\mathbf{j} + 3\mathbf{k}$.

5. Given that $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$, $\mathbf{b} = \mathbf{i} + 3\mathbf{j} + \mathbf{k}$ and $\mathbf{c} = 3\mathbf{i} + \mathbf{j} - \mathbf{k}$, find (a) $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c}$. (b) $(2\mathbf{b} - 3\mathbf{c}) \cdot \mathbf{a}$. (c) $\mathbf{a} \cdot \mathbf{a}$. (d) $\mathbf{c} \cdot (\mathbf{a} - 2\mathbf{b})$.

6. Given that $\mathbf{a} = 3\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$, and $\mathbf{c} = 5\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$, find (a) $\mathbf{b} \cdot (\mathbf{b} + (\mathbf{a} \cdot \mathbf{c})\mathbf{c})$. (b) $(\mathbf{a} + 2\mathbf{b}) \cdot (2\mathbf{b} - 3\mathbf{c})$. (c) $(\mathbf{c} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{a})\mathbf{c}$.
7. Find the angle between the following pairs of vectors:
   (a) $\mathbf{i} + \mathbf{j} + \mathbf{k}$, $2\mathbf{i} + \mathbf{j} - \mathbf{k}$.
   (b) $2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$, $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$.
   (c) $3\mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.
   (d) $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $4\mathbf{i} - 8\mathbf{j} + 16\mathbf{k}$.
8. Given $\mathbf{a} = 2\mathbf{i} - 3\mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$, and $\mathbf{c} = 3\mathbf{i} - 2\mathbf{j} - \mathbf{k}$, find the angles between the following pairs of vectors:
   (a) $\mathbf{a} + \mathbf{b}$, $\mathbf{b} - 2\mathbf{c}$. (b) $2\mathbf{a} - \mathbf{c}$, $\mathbf{a} + \mathbf{b} - \mathbf{c}$. (c) $\mathbf{b} + 3\mathbf{c}$, $\mathbf{a} - 2\mathbf{c}$.
9. Find the component of the force $\mathbf{F} = 4\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$ in the direction of the vector $\mathbf{i} + \mathbf{j} + \mathbf{k}$.
10. Find the component of the force $\mathbf{F} = 2\mathbf{i} + 5\mathbf{j} - 3\mathbf{k}$ in the direction of the vector $2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.
11. Given that $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, find (a) the projection of $\mathbf{a}$ onto the line of $\mathbf{b}$, and (b) the projection of $\mathbf{b}$ onto the line of $\mathbf{a}$.
12. Given that $\mathbf{a} = 3\mathbf{i} + 6\mathbf{j} + 9\mathbf{k}$ and $\mathbf{b} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, (a) find the projection of $\mathbf{a}$ onto the line of $\mathbf{b}$ and (b) compare the magnitude of $\mathbf{a}$ with the result found in (a) and comment on the result.
13. Find the direction cosines and corresponding angles for the following vectors:
    (a) $\mathbf{i} + \mathbf{j} + \mathbf{k}$. (b) $\mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$. (c) $4\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.
14. Find the direction cosines and corresponding angles for the following vectors:
    (a) $\mathbf{i} - \mathbf{j} - \mathbf{k}$. (b) $2\mathbf{i} + 2\mathbf{j} - 5\mathbf{k}$. (c) $-4\mathbf{j} - \mathbf{k}$.
15. Verify the triangle inequality for vectors $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$ and $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 7\mathbf{k}$.
16. Verify the triangle inequality for vectors $\mathbf{a} = 2\mathbf{i} - \mathbf{j} - 2\mathbf{k}$ and $3\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.
17. Find the equation of the plane with normal $2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ that contains the point $(1, 0, 1)$.
18. Find the equation of the plane with normal $\mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$ that contains the point $(2, -3, 4)$.
19. Given that a plane passes through the point $(2, 3, -5)$, and the vector $2\mathbf{i} + \mathbf{k}$ is normal to the plane, find the cartesian form of its equation.
20. The equation of a plane is $3x + 2y - 5z = 4$. Find a vector that is normal to the plane, and the position vector of a point on the plane.
21. Explain why if the vector equation of plane Π in (32) is divided by $\|\mathbf{n}\|$ to bring it into the form $\mathbf{r} \cdot \mathbf{n} = \mathbf{a} \cdot \mathbf{n}$, the number $|\mathbf{a} \cdot \mathbf{n}|$ is the perpendicular distance of origin O from the plane. Explain also why if $\mathbf{a} \cdot \mathbf{n} > 0$ the plane lies to the side of O toward which $\mathbf{n}$ is directed, as in Fig. 2.15, but that if $\mathbf{a} \cdot \mathbf{n} < 0$ it lies on the opposite side of O toward which $-\mathbf{n}$ is directed.
22. Use the result of Exercise 21 to find the perpendicular distance of the plane $2x - 4y - 5z = 5$ from the origin.
23. The angle between two planes is defined as the angle between their normals. Find the angle between the two planes $x + 3y + 2z = 4$ and $2x - 5y + z = 2$.
24. Find the angle between the two planes $3x + 2y - 2z = 4$ and $2x + y + 2z = 1$.
25. Let $\mathbf{a}$ and $\mathbf{b}$ be two arbitrary skew (nonparallel) vectors, and set $\mathbf{a} = \mathbf{a}_b + \mathbf{a}_p$, where $\mathbf{a}_b$ is parallel to $\mathbf{b}$ and $\mathbf{a}_p$ is perpendicular to $\mathbf{b}$ and lies in the plane of $\mathbf{a}$ and $\mathbf{b}$. Find $\mathbf{a}_b$ and $\mathbf{a}_p$ in terms of $\mathbf{a}$ and $\mathbf{b}$.
26. The **law of cosines** for a triangle with sides of length a, b, and c, in which the angle opposite the side of length c is C, takes the form
    $$c^2 = a^2 + b^2 - 2ab\cos C.$$
    Prove this by taking vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ such that $\mathbf{c} = \mathbf{a} - \mathbf{b}$ and considering the dot product $\mathbf{c} \cdot \mathbf{c} = (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b})$.
27. The work units W done by a constant force $\mathbf{F}$ when moving its point of application along a straight line L parallel to a vector $\mathbf{a}$ are defined as the product of the component of $\mathbf{F}$ in the direction of $\mathbf{a}$ and the distance d moved along line L. Express W in terms of $\mathbf{F}$, $\mathbf{a}$, and d.
28. If $\mathbf{a}$ and $\mathbf{b}$ are arbitrary vectors and λ and μ are any two scalars, prove that
    $$\|\lambda\mathbf{a} + \mu\mathbf{b}\|^2 \le \lambda^2\|\mathbf{a}\|^2 + 2\lambda\mu\mathbf{a} \cdot \mathbf{b} + \mu^2\|\mathbf{b}\|^2.$$
29. Verify the result of Exercise 28 by setting $\lambda = 2$, $\mu = -3$, $\mathbf{a} = 3\mathbf{i} + \mathbf{j} - 4\mathbf{k}$, and $\mathbf{b} = 2\mathbf{i} + 3\mathbf{j} + \mathbf{k}$.

2.3 The Cross Product

A product of two vectors $\mathbf{a}$ and $\mathbf{b}$ can be defined in such a way that the result is a vector. The result is written $\mathbf{a} \times \mathbf{b}$ and called the **cross product** of $\mathbf{a}$ and $\mathbf{b}$. The name **vector product** is also used in place of the term *cross product*.

Before defining the cross product we first formulate what is called the right-hand rule. Given any two skew vectors $\mathbf{a}$ and $\mathbf{b}$, the right-hand rule is used to determine the sense of a third vector **c** that is required to be normal to the plane containing vectors **a** and **b**.

right-hand rule

The Right-Hand Rule

Let **a** and **b** be two arbitrary skew vectors with the same base point, with **c** a vector normal to the plane containing them. If the fingers of the right hand are curled in such a way that they point from vector **a** to vector **b** through the angle θ between them, with $0 < \theta < \pi$, then when the thumb is extended away from the palm it will point in the direction of vector **c**.

When applying the right-hand rule, the order of the vectors is important. If vectors **a**, **b**, and **c** obey the right-hand rule, they will always be written in the order **a**, **b**, **c**, with the understanding that **c** is normal to the plane of **a** and **b**, with its sense determined by the right-hand rule. Figure 2.18 illustrates the right-hand rule.

An important special case of the right-hand rule has already been encountered in connection with the unit vectors **i**, **j**, and **k** that obey the rule, and because the vectors are mutually orthogonal the vectors **j**, **k**, **i** and **k**, **i**, **j** also obey the right-hand rule.

geometrical definition of a cross product

The cross product (a geometrical interpretation)

Let **a** and **b** be two arbitrary vectors, with $\hat{\mathbf{n}}$ a unit vector normal to the plane of **a** and **b** chosen so that **a**, **b**, and $\hat{\mathbf{n}}$, in this order, obey the right-hand rule. Then the **cross product** of vectors **a** and **b**, written $\mathbf{a} \times \mathbf{b}$, is defined as the vector

$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\|.\|\mathbf{b}\| \sin\theta\hat{\mathbf{n}}. \tag{34}$$

This geometrical definition of the cross product is useful in many situations, but when the vectors **a** and **b** are specified in terms of their cartesian components a different form of the definition will be needed.

The cross product can be interpreted as a *vector area*, in the sense that it can be written $\mathbf{a} \times \mathbf{b} = S\hat{\mathbf{n}}$, where $S = OA \cdot BN = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \sin\theta$ is the geometrical area

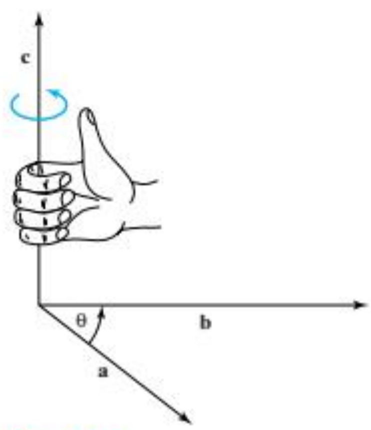


FIGURE 2.18 The right-hand rule.

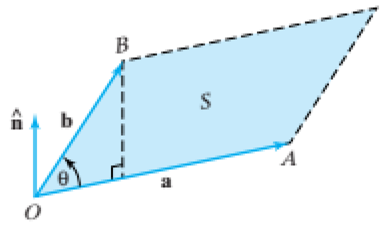


FIGURE 2.19 The cross product interpreted as the vector area of a parallelogram.

of the parallelogram in Fig. 2.19, and the unit vector $\hat{\mathbf{n}}$ is normal to the area. This shows that the geometrical area S of the vector parallelogram with sides $\mathbf{a}$ and $\mathbf{b}$ is simply the modulus of the cross product $\mathbf{a} \times \mathbf{b}$, so $S = \|\mathbf{a} \times \mathbf{b}\|$.

properties of the cross product

Properties of the cross product

The following results are consequences of the definition of the cross product.

The cross product is anticommutative

$$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a} \tag{35}$$

The cross product is associative

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}. \tag{36}$$

Parallel vectors ($\theta = 0$)

If vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel, then

$$\mathbf{a} \times \mathbf{b} = \mathbf{0}. \tag{37}$$

Orthogonal vectors ($\theta = \pi/2$)

If vectors $\mathbf{a}$ and $\mathbf{b}$ are orthogonal, then

$$\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\|.\|\mathbf{b}\|\hat{\mathbf{n}}. \tag{38}$$

Product of unit vectors

If $\mathbf{a}$ and $\mathbf{b}$ are unit vectors, then

$$\mathbf{a} \times \mathbf{b} = \sin\theta\hat{\mathbf{n}}. \tag{39}$$

An immediate consequence of properties (34), (35), and (37) is that

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}, \tag{40}$$

and

$$\mathbf{i}\times\mathbf{j}=\mathbf{k},\quad \mathbf{j}\times\mathbf{i}=-\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=\mathbf{i},\quad \mathbf{k}\times\mathbf{j}=-\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=\mathbf{j},\ \mathbf{i}\times\mathbf{k}=-\mathbf{j}. \tag{41}$$

Only results (35) and (36) require some comment, as the other results are obvious. The change of sign in (35) that makes the cross product *anticommutative* occurs because when the vectors $\mathbf{a}$ and $\mathbf{b}$ are interchanged, the right-hand rule causes the direction of $\hat{\mathbf{n}}$ to be reversed. Result (36) can be proved in several ways, but we shall postpone its proof until a different expression for the cross product has been derived.

To obtain a more convenient expression for the cross product that can be used when $\mathbf{a}$ and $\mathbf{b}$ are known in terms of their components, we proceed as follows. Let $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ and $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$, and consider the cross product $\mathbf{a}\times\mathbf{b}=(a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k})\times(b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k})$. Expanding this expression term by term is justified because of the associative property given in (36), and it leads to the result

$$\begin{aligned}\mathbf{a}\times\mathbf{b}&=a_1b_1\mathbf{i}\times\mathbf{i}+a_1b_2\mathbf{i}\times\mathbf{j}+a_1b_3\mathbf{i}\times\mathbf{k}+a_2b_1\mathbf{j}\times\mathbf{i}+a_2b_2\mathbf{j}\times\mathbf{j}\\&\quad+a_2b_3\mathbf{j}\times\mathbf{k}+a_3b_1\mathbf{k}\times\mathbf{i}+a_3b_2\mathbf{k}\times\mathbf{j}+a_3b_3\mathbf{k}\times\mathbf{k}.\end{aligned}$$

cross product in terms of components

Results (40) cause three terms on the right-hand side to vanish, and results (41) allow the remaining six terms to be collected into three groups as follows to give

$$\mathbf{a}\times\mathbf{b}=(a_2b_3-a_3b_2)\mathbf{i}-(a_1b_3-a_3b_1)\mathbf{j}+(a_1b_2-a_2b_1)\mathbf{k}. \tag{42}$$

This alternative expression for the cross product in terms of the cartesian components of vectors $\mathbf{a}$ and $\mathbf{b}$ can be further simplified by making formal use of the third-order determinant,

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix},$$

because a formal expansion in terms of elements of the first row generates result (42). We take this result as an alternative but equivalent definition of the cross product.

practical definition of a cross product using a determinant

The cross product (cartesian component form)

Let $\mathbf{a}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ and $\mathbf{b}=b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$. Then

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix}, \tag{43}$$

When expressing $\mathbf{a}\times\mathbf{b}$ as the determinant in (43), purely *formal* use was made of the method of expansion of a determinant in terms of the elements of its first row, because (43) is not a determinant in the ordinary sense as its elements are a mixture of vectors and numbers.

EXAMPLE 2.16 Given that $\mathbf{a} = 3\mathbf{i} - 2\mathbf{j} - \mathbf{k}$ and $\mathbf{b} = \mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$, find $\mathbf{a} \times \mathbf{b}$ and a unit vector $\hat{\mathbf{n}}$ normal to the plane containing $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{n}$, in this order, obey the right-hand rule.

Solution Substitution into expression (43) gives

$$\begin{aligned}\mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -2 & -1 \\ 1 & 4 & 2 \end{vmatrix} \\ &= [(-2)\cdot 2 - 4\cdot(-1)]\mathbf{i} - [3\cdot 2 - 1\cdot(-1)]\mathbf{j} + [3\cdot 4 - 1\cdot(-2)]\mathbf{k} \\ &= -7\mathbf{j} + 14\mathbf{k}.\end{aligned}$$

The required unit vector $\hat{\mathbf{n}}$ is simply the unit vector in the direction of $\mathbf{a} \times \mathbf{b}$, so

$$\begin{aligned}\hat{\mathbf{n}} &= (\mathbf{a} \times \mathbf{b})/\|\mathbf{a} \times \mathbf{b}\| = (-7\mathbf{j} + 14\mathbf{k})/(7\sqrt{5}). \\ &= (-1/\sqrt{5})\mathbf{j} + (2/\sqrt{5})\mathbf{k}.\end{aligned}$$

We now return to the proof of the associative property stated in (35) and establish it by means of result (43).

Setting $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, and $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$, we have

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ (b_1 + c_1) & (b_2 + c_2) & (b_3 + c_3) \end{vmatrix}.$$

Expanding the determinant in terms of elements of its first row and grouping terms gives

$$\begin{aligned}\mathbf{a} \times (\mathbf{b} + \mathbf{c}) &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \\ &\quad + (a_2c_3 - a_3c_2)\mathbf{i} - (a_1c_3 - a_3c_1)\mathbf{j} + (a_1c_2 - a_2c_1)\mathbf{k} \\ &= \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c},\end{aligned}$$

and the result is proved.

Summary

This section first introduced the vector or cross product of two vectors in geometrical terms and then used the result to show that the vector product is anticommutative, in the sense that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Important results involving the vector product are given in terms of the components of the two vectors that are involved. Finally, the vector product was expressed in a form that is most convenient for calculations by writing it in determinantal form, the rows of which contain the unit vectors **i**, **j**, and **k** and the components of the respective vectors.

EXERCISES 2.3

In Exercises 1 through 6 use (43) to find $\mathbf{a} \times \mathbf{b}$.

1. For $\mathbf{a} = 2\mathbf{i} - \mathbf{j} - 4\mathbf{k}$, $\mathbf{b} = 3\mathbf{i} - \mathbf{j} - \mathbf{k}$.
2. For $\mathbf{a} = -3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$.
3. For $\mathbf{a} = 7\mathbf{i} + 6\mathbf{k}$, $\mathbf{b} = 3\mathbf{j} + \mathbf{k}$.
4. For $\mathbf{a} = 3\mathbf{i} + 7\mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k}$.
5. For $\mathbf{a} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j} + \mathbf{k}$.
6. For $\mathbf{a} = 3\mathbf{i} - 2\mathbf{j} + 6\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$.

In Exercises 7 through 10 verify the equivalence of the definitions of the cross product in (34) and (43) by first using (43) to calculate $\mathbf{a} \times \mathbf{b}$, and hence $\|\mathbf{a} \times \mathbf{b}\|$ and $\hat{\mathbf{n}}$, and then calculating $\|\mathbf{a}\|$ and $\|\mathbf{b}\|$ directly, using result (17) to find $\cos\theta$ and hence $\sin\theta$, and using the results to find $\mathbf{a} \times \mathbf{b}$ from (34).

7. For $\mathbf{a} = \mathbf{i} + \mathbf{j} + 3\mathbf{k}$ and $\mathbf{b} = 3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$.

8. For $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$ and $\mathbf{b} = 4\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$.

9. For $\mathbf{a} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$ and $\mathbf{b} = 5\mathbf{i} - 2\mathbf{k}$.

10. For $\mathbf{a} = -2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$ and $\mathbf{b} = 3\mathbf{i} + \mathbf{j} + 2\mathbf{k}$.

In Exercises 11 through 14, verify by direct calculation that $(\mathbf{b} + \mathbf{c}) \times \mathbf{a} = -\mathbf{a} \times (\mathbf{b} + \mathbf{c})$.

11. $\mathbf{a} = 3\mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = \mathbf{i} - 4\mathbf{j} + \mathbf{k}$, and $\mathbf{c} = 5\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

12. $\mathbf{a} = -\mathbf{i} + 5\mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = 4\mathbf{i} + \mathbf{k}$, and $\mathbf{c} = -2\mathbf{i} - 4\mathbf{j} + 3\mathbf{k}$.

13. $\mathbf{a} = \mathbf{i} + \mathbf{k}$, $\mathbf{b} = 3\mathbf{i} - \mathbf{j} - 2\mathbf{k}$, and $\mathbf{c} = 3\mathbf{i} + \mathbf{j} + \mathbf{k}$.

14. $\mathbf{a} = 5\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j} - \mathbf{k}$, and $\mathbf{c} = 4\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.

In Exercises 15 through 18 find a unit vector normal to a plane containing the given vectors.

15. $3\mathbf{i} + \mathbf{j} + \mathbf{k}$ and $\mathbf{i} + 2\mathbf{j} + \mathbf{k}$.

16. $2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ and $2\mathbf{i} + 3\mathbf{j} + \mathbf{k}$.

17. $\mathbf{i} + \mathbf{j} + \mathbf{k}$ and $2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$.

18. $2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $3\mathbf{i} + \mathbf{j} + 4\mathbf{k}$.

19. Find a unit vector normal to a plane containing vectors $\mathbf{a} + \mathbf{b}$ and $\mathbf{a} + \mathbf{c}$, given that $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, and $\mathbf{c} = 3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$.

20. Given that $\mathbf{a} = 3\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, and $\mathbf{c} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, find (a) a vector normal to the plane containing the vectors $\mathbf{a} + (\mathbf{a} \cdot \mathbf{b})\mathbf{b}$ and $\mathbf{c}$ and, (b) explain why the normal to a plane containing the vectors $\mathbf{a}$ and $\mathbf{b}$ and the normal to a plane containing the vectors $(\mathbf{a} \cdot \mathbf{b})\mathbf{a}$ and $(\mathbf{b} \cdot \mathbf{c})\mathbf{b}$ are parallel.

In Exercises 21 through 24, find the cartesian equation of the plane that passes through the given points.

21. $(1, 3, 2)$, $(2, 0, -4)$, and $(1, 6, 11)$.

22. $(1, 4, 3)$, $(2, 0, 1)$, and $(3, 4, -6)$.

23. $(1, 2, 3)$, $(2, -4, 1)$, and $(3, 6, -1)$.

24. $(1, 0, 1)$, $(2, 5, 7)$, and $(2, 3, 9)$.

Three points with position vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ will be **collinear** (lie on a line) if the parallelogram with adjacent sides $\mathbf{a} - \mathbf{b}$ and $\mathbf{a} - \mathbf{c}$ has zero geometrical area. Use this result in Exercises 25 through 28 to determine which sets of points are collinear.

25. $(2, 2, 3)$, $(6, 1, 5)$, $(-2, 4, 3)$.

26. $(1, 2, 4)$, $(7, 0, 8)$, $(-8, 5, -2)$.

27. $(2, 3, 3)$, $(3, 7, 5)$, $(0, -5, -1)$.

28. $(1, 3, 2)$, $(4, 2, 1)$, $(1, 0, 2)$.

29. A vector $\mathbf{N}$ normal to the plane containing the skew vectors $\mathbf{a}$ and $\mathbf{b}$ can be found as follows. $\mathbf{N}$ is normal to $\mathbf{a}$ and $\mathbf{b}$, so $\mathbf{a} \cdot \mathbf{N} = 0$ and $\mathbf{b} \cdot \mathbf{N} = 0$. If a component of $\mathbf{N}$ is assigned an arbitrary nonzero value c, say, the other two components can be found from these two equations as multiples of c, and $\mathbf{N}$ will then be determined as a multiple of c. A suitable choice of c will make $\mathbf{N}$ a unit normal $\hat{\mathbf{N}}$. Apply this method to vectors $\mathbf{a}$ and $\mathbf{b}$ in Exercise 7 to find a vector $\hat{\mathbf{N}}$. Compare the result with the unit vector

$$\hat{\mathbf{n}} = (\mathbf{a} \times \mathbf{b})/\|\mathbf{a} \times \mathbf{b}\|$$

found from (43). Explain why although both $\hat{\mathbf{n}}$ and $\hat{\mathbf{N}}$ are normal to the plane containing $\mathbf{a}$ and $\mathbf{b}$ they may have opposite senses.

2.4 Linear Dependence and Independence of Vectors and Triple Products

The dot and cross products can be combined to provide a simple test that determines whether or not an arbitrary set of three vectors possesses a property of fundamental importance to the algebra of vectors. First, however, some introductory remarks are necessary.

Given a set of n vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, and a set of n constants $c_1, c_2, \ldots, c_n$, the sum

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n$$

linear combination of vectors

basis

is called a **linear combination** of the vectors. Linear combinations of the vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ were used in Section 2.1 to express *every* vector in three-dimensional space as a linear combination of these three vectors. A triad of vectors such as $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ with the property that *all* vectors in three-dimensional space can be represented as linear combinations of these three vectors is said to form a **basis** for the space.

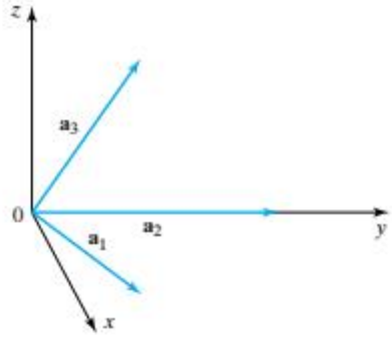


FIGURE 2.20 Nonorthogonal triad forming a basis in three-dimensional space.

It is a fundamental property of three-dimensional space that a basis for the space comprises a set of three vectors $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$, with the property that the linear combination

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3 = \mathbf{0} \tag{44}$$

linear independence and linear dependence

is *only* true when $c_1 = c_2 = c_3 = 0$. Vectors $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$ satisfying this condition are said to be **linearly independent** vectors, and a vector $\mathbf{d}$ of the form

$$\mathbf{d} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3,$$

where *not all* of c_1, c_2, and c_3 are zero, is said to be **linearly dependent** on the vectors $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$. The vectors **i**, **j**, and **k** that form a basis for three-dimensional space are linearly independent vectors, but the position vector $\mathbf{r} = 2\mathbf{i} - 3\mathbf{j} + 5\mathbf{k}$ is linearly dependent on vectors **i**, **j**, and **k**.

Clearly, vectors **i**, **j**, and **k** do not form the only basis for three-dimensional space, because any triad of linearly independent vectors $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$ will serve equally well, as, for example, the nonorthogonal set of vectors shown in Fig. 2.20.

The dot and cross products will now be combined to develop a test for linear dependence and independence based on the elementary geometrical idea of the volume of the parallelepiped shown in Fig. 2.21, three edges **a**, **b**, and **c** of which meet at the origin.

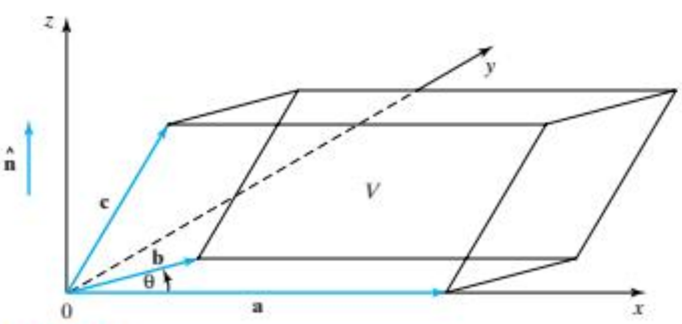


FIGURE 2.21 Volume V of a parallelepiped.

The volume V of a parallelepiped is a nonnegative number given by the product of the area of its base and its height. Suppose vectors $\mathbf{a}$ and $\mathbf{b}$ are chosen to form two sides of the base of the parallelepiped. Then the vector area of this base has already been interpreted as $\mathbf{a} \times \mathbf{b}$. The vertical height of the parallelepiped is the projection of vector $\mathbf{c}$ in the direction of the unit vector $\hat{\mathbf{n}}$ normal to the base, and so is given by $\hat{\mathbf{n}} \cdot \mathbf{c}$. Consequently, as $\mathbf{a} \times \mathbf{b} = \|\mathbf{a}\| \cdot \|\mathbf{b}\| \sin\theta \hat{\mathbf{n}}$, it follows that

$$V = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|. \tag{45}$$

The absolute value of the right-hand side of (45) has been taken because a volume must be a nonnegative quantity, whereas the dot product $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ may be of either sign.

If vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ form a basis for three-dimensional space, vector $\mathbf{c}$ cannot be linearly dependent on vectors $\mathbf{a}$ and $\mathbf{b}$, and so the parallelepiped in Fig. 2.21 with these vectors as its sides must have a nonzero volume. If, however, vectors $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are **coplanar** (all lie in the same plane), and so cannot form a basis for the space, the volume of the parallelepiped will be zero. These simple geometrical observations lead to the following test for the linear independence of three vectors in three-dimensional space.

THEOREM 2.3

a test for linear independence

Test for linear independence of vectors in three-dimensional space Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ be any three vectors. Then the vectors are linearly independent if $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \neq 0$, and they are linearly dependent if $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0$. ■

A product of the type $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ is called a **scalar triple product**. The name arises because the result is a scalar. It is also called a mixed triple product since both $\cdot$ and $\times$ appear. Three vectors are involved in this dot (scalar) product, one of which is the vector $\mathbf{a} \times \mathbf{b}$ and the other is the vector $\mathbf{c}$.

scalar triple product

Scalar triple products are easily evaluated, because taking the dot product of $\mathbf{a} \times \mathbf{b}$ in the form given in (42) with $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$ gives

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (a_2b_3 - a_3b_2)c_1 - (a_1b_3 - a_3b_1)c_2 + (a_1b_2 - a_2b_1)c_3.$$

The right-hand side of this expression is simply the value of a determinant with successive rows given by the components of $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, so we have arrived at the following convenient formula for the scalar triple product.

scalar triple product as a determinant

Scalar triple product

Let $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, and $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$. Then

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{46}$$

Interchanging any two rows in a matrix changes the sign but not the value of its determinant. Two such switches in (46) leave the value unchanged, so the dot product is commutative and so we arrive at the useful result

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}). \tag{47}$$

So, in a scalar triple product the dot and cross may be *interchanged* without altering the result.

EXAMPLE 2.17

Given the two sets of vectors (a) $\mathbf{a} = \mathbf{i} + 2\mathbf{j} - 5\mathbf{k}$, $\mathbf{b} = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$, $\mathbf{c} = \mathbf{i} + 4\mathbf{j} - 19\mathbf{k}$ and (b) $\mathbf{a} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 3\mathbf{i} + 4\mathbf{k}$, $\mathbf{c} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, find if the vectors are linearly independent or linearly dependent.

Solution We apply Theorem 2.3 to each set, using result (46) to evaluate the scalar triple products.

(a)
$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} 1 & 2 & -5 \\ 1 & 1 & 2 \\ 1 & 4 & -19 \end{vmatrix} = 0,$$

so the set of three vectors in (a) is linearly dependent. In fact this can be seen from the fact that $\mathbf{c} = 3\mathbf{a} - 2\mathbf{b}$.

(b)
$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{vmatrix} 2 & 1 & 1 \\ 3 & 0 & 4 \\ 1 & 1 & 1 \end{vmatrix} = -4 \neq 0,$$

so the set of three vectors in (b) is linearly independent. Although not required, the volume V of the parallelepiped formed by these three vectors is $V = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| = |-4| = 4$. ■

Another notation for the scalar triple product of vectors **a**, **b**, and **c** is $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$, so

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}, \tag{48}$$

or, in terms of a determinant,

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{49}$$

alternative forms of a scalar triple product

Using this definition of $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ with the row interchange property of determinants (see Section 1.7) shows that

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}], \tag{50}$$

because two row interchanges are needed to arrive at $[\mathbf{b}, \mathbf{c}, \mathbf{a}]$ from $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$, leaving the sign of the determinant unchanged, whereas two more are required to arrive at $[\mathbf{c}, \mathbf{a}, \mathbf{b}]$ from $[\mathbf{b}, \mathbf{c}, \mathbf{a}]$, again leaving the sign of the determinant unchanged.

The order of the vectors in results (46), or in the equivalent notation of (48), is easily remembered when the results are abbreviated to

$$\begin{matrix} \mathbf{a} & \mathbf{b} & \mathbf{c} \\ \mathbf{b} & \mathbf{c} & \mathbf{a} \\ \mathbf{c} & \mathbf{a} & \mathbf{b} \end{matrix}$$

In this pattern, row two follows from row one when the first letter is moved to the end position, and row three follows from row two by means of the same process. The effect of applying this process to the third row is simply to regenerate the first row. Rearrangements of this kind are called **cyclic permutations** of the three vectors.

Again making use of the row interchange property of determinants (see Section 1.7), it follows that

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = -[\mathbf{a}, \mathbf{c}, \mathbf{b}],$$

because this time only one row interchange is needed to produce the result on the right from the one on the left, so that a sign change is involved.

A different product involving the three vectors **a**, **b**, and **c** that this time generates another vector is of the form

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}),$$

and products of this type are called **vector triple products** since the results are vectors. In these products it is essential to include the brackets because, in general, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$. The most important results concerning vector triple products are given in the following theorem.

THEOREM 2.4 **Vector triple products** If **a**, **b**, and **c** are any three vectors, then

(a) $$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$$

vector triple product

and

(b) $$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}.$$

Proof The proof of the results in Theorem 2.4 both follow in similar fashion, so we only prove result (a) and leave the proof of result (b) as an exercise. We write the cross product $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ in the form of the determinant in (43), with the components of **a** in the second row and those of $\mathbf{b} \times \mathbf{c}$ (obtained from (42)) in the third row when we find that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ (b_2c_3 - b_3c_2) & (b_3c_1 - b_1c_3) & (b_1c_2 - b_2c_1) \end{vmatrix}.$$

Expanding this determinant in terms of the elements of its first row and grouping terms gives

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = {} & [(a_2c_2 + a_3c_3)b_1 - (a_2b_2 + a_3b_3)c_1]\mathbf{i} + [(a_1c_1 + a_3c_3)b_2 \\ & - (a_1b_1 + a_3b_3)c_2]\,\mathbf{j} + [(a_1c_1 + a_2c_2)b_3 - (a_1b_1 + a_2b_2)c_3]\mathbf{k}. \end{aligned}$$

As it stands, this result is not yet in the form that is required, but adding and subtracting $a_1b_1c_1$ to the coefficient of **i**, $a_2b_2c_2$ to the coefficient of **j**, and $a_3b_3c_3$ to the coefficient of **k** followed by grouping terms give

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c},$$

and the result is established. ■

EXAMPLE 2.18

Find $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ and $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$, given that $\mathbf{a} = 3\mathbf{i} + \mathbf{j} - 4\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$, and $\mathbf{c} = \mathbf{i} + 5\mathbf{j} - \mathbf{k}$.

Solution $\mathbf{a} \cdot \mathbf{b} = -5$, $\mathbf{a} \cdot \mathbf{c} = 12$, and $\mathbf{b} \cdot \mathbf{c} = 4$, so

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} = 12\mathbf{b} + 5\mathbf{c} = 29\mathbf{i} + 37\mathbf{j} + 31\mathbf{k},$$

and

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a} = 12\mathbf{b} - 4\mathbf{a} = 12\mathbf{i} + 8\mathbf{j} + 52\mathbf{k}.$$

Accounts of geometrical vectors can be found, for example, in references [2.1], [2.3], [2.6], and [1.6].

Summary

This section introduced the two fundamental concepts of linear dependence and independence of vectors. It then showed how the scalar triple product involving three vectors, that gives rise to a scalar quantity, provides a simple test for the linear dependence or independence of the vectors involved. A simple and convenient way of calculating a scalar triple product was shown to be in terms of a determinant with the elements in its rows formed by the components of the three vectors involved in the product. Finally a vector triple product was defined that gives rise to a vector quantity, and it was shown that to avoid ambiguity it is necessary to bracket a pair of vectors in such a product. A rule for the expansion of a vector triple product was derived and shown to involve a linear combination of two of the vectors multiplied by scalar products so that, for example, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

EXERCISES 2.4

In Exercises 1 through 4 use the vectors **a**, **b**, and **c** to find (a) the scalar triple product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, and (b) the volume of the parallelepiped determined by these three vectors directed away from a corner.

1. $\mathbf{a} = 2\mathbf{i} - \mathbf{j} - 3\mathbf{k}$, $\mathbf{b} = 3\mathbf{i} - 2\mathbf{k}$, $\mathbf{c} = \mathbf{i} + \mathbf{j} - 4\mathbf{k}$.
2. $\mathbf{a} = \mathbf{i} - \mathbf{j} + 2\mathbf{k}$, $\mathbf{b} = \mathbf{i} + \mathbf{j} + 3\mathbf{k}$, $\mathbf{c} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$.
3. $\mathbf{a} = -\mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$, $\mathbf{c} = -4\mathbf{i} + \mathbf{j} + 3\mathbf{k}$.
4. $\mathbf{a} = 5\mathbf{i} + 3\mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j}$, $\mathbf{c} = -2\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$.

In Exercises 5 through 10 find which sets of vectors are coplanar.

5. $\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$, $2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$, $4\mathbf{i} + 7\mathbf{j} + 8\mathbf{k}$.
6. $2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$, $\mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $4\mathbf{i} + 3\mathbf{j} + 6\mathbf{k}$.
7. $2\mathbf{i} + \mathbf{k}$, $\mathbf{i} + 4\mathbf{j} + 2\mathbf{k}$, $3\mathbf{i} + 12\mathbf{j} + 7\mathbf{k}$.
8. $\mathbf{i} + \mathbf{j} + \mathbf{k}$, $2\mathbf{i} + \mathbf{j} + 2\mathbf{k}$, $4\mathbf{i} + 3\mathbf{j} + \mathbf{k}$.
9. $2\mathbf{i} + \mathbf{j} - \mathbf{k}$, $3\mathbf{i} + \mathbf{j} + 2\mathbf{k}$, $5\mathbf{i} + \mathbf{j} + 8\mathbf{k}$.
10. $2\mathbf{i} + \mathbf{j} - \mathbf{k}$, $\mathbf{i} + 2\mathbf{j} + 2\mathbf{k}$, $5\mathbf{i} + 4\mathbf{j} + \mathbf{k}$.

In Exercises 11 through 15 use computer algebra to verify that $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] = -[\mathbf{a}, \mathbf{c}, \mathbf{b}]$.

11. $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} - \mathbf{k}$, $\mathbf{c} = 3\mathbf{i} - \mathbf{j} + \mathbf{k}$.
12. $\mathbf{a} = \mathbf{i} - \mathbf{j} - \mathbf{k}$, $\mathbf{b} = -5\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, $\mathbf{c} = 2\mathbf{i} + 3\mathbf{j} - 2\mathbf{k}$.
13. $\mathbf{a} = -3\mathbf{i} - 4\mathbf{j} + \mathbf{k}$, $\mathbf{b} = 9\mathbf{i} + 12\mathbf{j} - 3\mathbf{k}$, $\mathbf{c} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$.
14. $\mathbf{a} = 3\mathbf{i} + 4\mathbf{k}$, $\mathbf{b} = \mathbf{i} + 5\mathbf{k}$, $\mathbf{c} = 2\mathbf{j} + \mathbf{k}$.
15. Prove that if **a**, **b**, **c**, and **d** are any four vectors, and λ, μ are arbitrary scalars $[\lambda\mathbf{a} + \mu\mathbf{b}, \mathbf{c}, \mathbf{d}] = \lambda[\mathbf{a}, \mathbf{c}, \mathbf{d}] + \mu[\mathbf{b}, \mathbf{c}, \mathbf{d}]$. Use computer algebra with vectors **a**, **b**, **c**, **d** from Exercise 12 with $\mathbf{d} = 4\mathbf{c} - 2\mathbf{j} + 6\mathbf{k}$, and scalars λ, μ of your choice, to verify this result.

In Exercises 16 through 20 find (a) the cartesian equation of the plane containing the given points, and (b) a unit vector normal to the plane.

16. $(1, 2, 1)$, $(3, 1, -2)$, $(2, 1, 4)$.
17. $(2, 0, 3)$, $(0, 1, 0)$, $(2, 4, 5)$.
18. $(-1, 2, -3)$, $(2, 4, 1)$, $(3, 0, 1)$.
19. $(1, 2, 5)$, $(-2, 1, 0)$, $(0, 2, 0)$.
20. Prove result (b) of Theorem 2.4.
21. Show that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}.$$

22. The **law of sines** for a triangle with angles A, B, and C opposite sides with the respective lengths a, b, and c

takes the form

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}.$$

Prove this by considering a vector triangle with sides **a**, **b**, and **c**, where $\mathbf{c} = \mathbf{a} + \mathbf{b}$, and taking the cross product of $\mathbf{c} = \mathbf{a} + \mathbf{b}$ first with **a**, then with **b**, and finally with **c**.

In Exercises 23 through 26 use the fact that four points with position vectors **p**, **q**, **r**, and **s** will be coplanar if the vectors $\mathbf{p} - \mathbf{q}$, $\mathbf{p} - \mathbf{r}$, and $\mathbf{p} - \mathbf{s}$ are coplanar to find which sets of points all lie in a plane.

23. $(1, 1, -1), (-3, 1, 1), (-1, 2, -1), (1, 0, 0)$.

24. $(1, 2, -1), (2, 1, 1), (0, 1, 2), (1, 1, 1)$.

25. $(0, -4, 0), (2, 3, 1), (3, -4, -2), (4, -2, -2)$.

26. $(1, 2, 3), (1, 0, 1), (2, 1, 2), (4, 1, 0)$.

27. The volume of a tetrahedron is one-third of the product of the area of its base and its vertical height. Show the volume V of the tetrahedron in Fig. 2.22, in which three edges formed by the vectors **a**, **b**, and **c** are directed away from a vertex, is given by

$$V = (1/6)|\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$$

28. Let **a**, **b**, **c**, and **d** be vectors and λ, μ, ν be scalars satisfying the equation

$$\lambda(\mathbf{b} \times \mathbf{c}) + \mu(\mathbf{c} \times \mathbf{a}) + \nu(\mathbf{a} \times \mathbf{b}) + \mathbf{d} = \mathbf{0}.$$

Show that if **a**, **b**, and **c** are linearly independent, then

$$\lambda = -(\mathbf{a} \cdot \mathbf{d})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})], \quad \mu = -(\mathbf{b} \cdot \mathbf{d})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})],$$
$$\nu = -(\mathbf{c} \cdot \mathbf{d})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})].$$

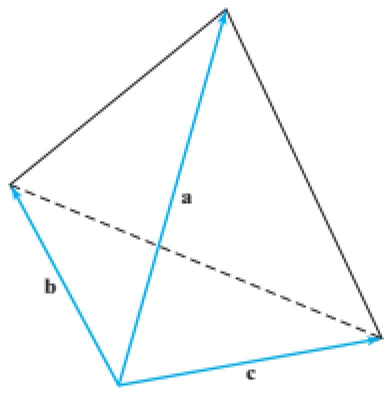


FIGURE 2.22 Tetrahedron.

29. Let **a**, **b**, **c**, and **d** be vectors and λ, μ, ν be scalars satisfying the equation

$$\lambda\mathbf{a} + \mu\mathbf{b} + \nu\mathbf{c} + \mathbf{d} = \mathbf{0}.$$

By taking the scalar products of this equation first with $\mathbf{b} \times \mathbf{c}$, then with $\mathbf{a} \times \mathbf{c}$, and finally with $\mathbf{a} \times \mathbf{b}$, show that if **a**, **b**, and **c** are linearly independent, then

$$\lambda = -\mathbf{d} \cdot (\mathbf{b} \times \mathbf{c})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})],$$
$$\mu = -\mathbf{d} \cdot (\mathbf{c} \times \mathbf{a})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})],$$
$$\nu = -\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})/[\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})].$$

30. Show that $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} - \mathbf{j} - \mathbf{k}$, and $\mathbf{c} = 4\mathbf{i} + 3\mathbf{j} + \mathbf{ik}$ are linearly independent vectors, and use them with a vector **d** of your choice to verify the results of Exercises 28 and 29.

31. Prove the **Lagrange identity**

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}).$$

2.5 *n*-Vectors and the Vector Space R^n

There are many occasions when it is convenient to generalize a vector and its associated algebra to spaces of more than three dimensions. A typical situation occurs in mechanics, where it is sometimes necessary to consider both the position and the momentum of a particle as functions of time. This leads to the study of a 6-vector, three components of which specify the particle position and three its momentum vector at a time t.

***n*-tuples**

Sets of n numbers $(x_1, x_2, \ldots, x_n)$ in a given order, that can be thought of either as n-vectors or as the coordinates of a point in n-dimensional space are called **ordered *n*-tuples** of real numbers or, simply, ***n*-tuples**.

n-vector

An n-vector

If $n \geq 2$ is an integer, and $x_1, x_2, \ldots, x_n$ are real numbers, an ***n*-vector** is an ordered n-tuple

$$(x_1, x_2, \ldots, x_n).$$

components and dimension

The numbers $x_1, x_2, \ldots, x_n$ are called the **components** of the n-vector, x_i is the ith component of the vector, and n is called the **dimension** of the space to which the n-vector belongs. For any given n, the set of all vectors with n real components is called a **real *n*-space** or, simply, an ***n*-space**, and it is denoted by the symbol R^n. A corresponding space exists when the n numbers $x_1, x_2, \ldots, x_n$ are allowed to be complex numbers, leading to a **complex *n*-space** denoted by C^n. In this notation R^3 is the three-dimensional space used in previous sections.

In R^3 the length of a vector was taken as the definition of its norm, so if $\mathbf{r} = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, then $\|\mathbf{r}\| = (x_1^2 + x_2^2 + x_3^2)^{1/2}$. A generalization of this norm to R^n leads to the following definition.

norm in R^n

The norm in R^n

The **norm** of the n-vector $(x_1, x_2, \ldots, x_n)$, denoted by $\|(x_1, x_2, \ldots, x_n)\|$ is

$$\begin{aligned} \|(x_1, x_2, \ldots, x_n)\| &= \sqrt{(x_1^2 + x_2^2 + \cdots + x_n^2)} \\ &= \left[\sum_{i=1}^{n} x_i^2\right]^{1/2}. \end{aligned} \tag{51}$$

The laws for the equality, addition, and scaling of vectors in R^3 in terms of the components of the vector generalize to R^n as follows.

algebraic rules for equality, addition, and scaling using components

Equality of n-vectors

Let $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ be two n-vectors. Then the vectors will be **equal**, written $(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_n)$, if, and only if, corresponding components are equal, so that

$$x_1 = y_1, \quad x_2 = y_2, \ldots, x_n = y_n. \tag{52}$$

Addition of n-vectors

Let $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ be any two n-vectors. Then the **sum** of these vectors, written $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n)$, is defined as the vector whose ith component is the sum of the corresponding ith components of the vectors for $i = 1, 2, \ldots, n$, so that

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n). \tag{53}$$

Scaling an n-vector

Let $(x_1, x_2, \ldots, x_n)$ be an arbitrary n-vector and λ be any scalar. Then the result of **scaling** the vector by λ, written $\lambda(x_1, x_2, \ldots, x_n)$, is defined as the vector whose ith component is λ times the ith component of the original vector, for $i = 1, 2, \ldots, n$, so that

$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n). \tag{54}$$

The **null (zero)** vector in R^n is the vector $\mathbf{0}$ in which every component is zero, so that

$$\mathbf{0} = (0, 0, \ldots, 0). \tag{55}$$

As with vectors in R^3, so also with n-vectors in R^n, it is convenient to use a single boldface symbol for a vector and the corresponding italic symbols with suffixes when it is necessary to specify the components. So we will write

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad \mathbf{y} = (y_1, y_2, \ldots, y_n).$$

The reasoning that led to the interpretation of Theorem 2.1 on the algebraic rules for the addition and scaling of vectors in R^3 leads also the following theorem for n-vectors.

THEOREM 2.5 **Algebraic rules for the addition and scaling of n-vectors in R^n** Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be arbitrary n-vectors, and let λ and μ be arbitrary real numbers. Then:

(i) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$;
(ii) $\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}$;
(iii) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$;
(iv) $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$;
(v) $(\lambda\mu)\mathbf{x} = \lambda(\mu\mathbf{x}) = \mu(\lambda\mathbf{x})$;
(vi) $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$;
(vii) $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$. ■

Because of this similarity between vectors in R^3 and in R^n, the space R^n is called a **real vector space**, though because the symbol R indicates *real* numbers this is usually abbreviated a **vector space**. Analogously, when the elements of the n-vectors are allowed to be complex, the resulting space is called the **complex vector space** C^n.

So far there would seem to be little difference between vectors in R^3 and R^n, but major differences do exist, and they are best appreciated when geometrical analogies are sought for vector operations in R^n.

dot product of n-vectors

The dot product of n-vectors

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be any two n-vectors. Then the **dot product** of these two vectors, written $\mathbf{x} \cdot \mathbf{y}$ and also called their **inner**

product, is defined as the sum of the products of corresponding components, so that

$$(x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n. \quad (56)$$

The following properties of this dot product are strictly analogous to those of the dot product in R^3 and can be deduced directly from (56).

THEOREM 2.6

Properties of the dot product in R^n Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be any three n-vectors and λ be any scalar. Then:

(i) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$;
(ii) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$;
(iii) $(\lambda\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\lambda\mathbf{y}) = \lambda(\mathbf{x} \cdot \mathbf{y})$;
(iv) $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$;
(v) $\mathbf{x} \cdot \mathbf{0} = 0$;
(vi) $\|\mathbf{x}\|^2 = 0$ if, and only if, $\mathbf{x} = \mathbf{0}$. ∎

The existence of a dot product in R^n allows the Cauchy–Schwarz and triangle inequalities to be generalized, both of which play a fundamental role in the study of vector spaces. Various forms of proof of these inequalities are possible, but the one given here has been chosen because it makes full use of the properties of the dot product listed in Theorem 2.6.

THEOREM 2.7

generalized inequalities for n-vectors

The Cauchy–Schwarz and triangle inequalities Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ be any two n-vectors. Then

(a) $$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\| \quad \textbf{(Cauchy–Schwarz inequality)},$$

and

(b) $$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \textbf{(triangle inequality)}.$$

Proof We start by proving the Cauchy–Schwarz inequality in (a). The inequality is certainly true if $\mathbf{x} \cdot \mathbf{y} = 0$, so we need only consider the case $\mathbf{x} \cdot \mathbf{y} \neq 0$. Let $\mathbf{x}$ and $\mathbf{y}$ be any two n-vectors, and λ be a scalar. Then, using properties (ii) to (iv) of Theorem 2.6,

$$\begin{aligned}\|\mathbf{x} + \lambda\mathbf{y}\|^2 &= (\mathbf{x} + \lambda\mathbf{y}) \cdot (\mathbf{x} + \lambda\mathbf{y}),\\ &= \|\mathbf{x}\|^2 + \lambda\mathbf{x} \cdot \mathbf{y} + \lambda\mathbf{y} \cdot \mathbf{x} + \lambda^2\|\mathbf{y}\|^2.\end{aligned}$$

However, by result (1) of Theorem 2.6, $\mathbf{y} \cdot \mathbf{x} = \mathbf{x} \cdot \mathbf{y}$, so

$$\|\mathbf{x} + \lambda\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\lambda\mathbf{x} \cdot \mathbf{y} + \lambda^2\|\mathbf{y}\|^2.$$

We now set $\lambda = -\|x\|^2/(\mathbf{x} \cdot \mathbf{y})$ to obtain

$$\|\mathbf{x} + \lambda\mathbf{y}\|^2 = -\|\mathbf{x}\|^2 + (\|\mathbf{x}\|^4\|\mathbf{y}\|^2)/|\mathbf{x} \cdot \mathbf{y}|^2,$$

where we have used the fact that $(\mathbf{x} \cdot \mathbf{y})^2 = |\mathbf{x} \cdot \mathbf{y}|^2$. As $\|\mathbf{x} + \lambda\mathbf{y}\|^2$ is nonnegative, this result is equivalent to

$$-\|\mathbf{x}\|^2 + (\|\mathbf{x}\|^4 \cdot \|\mathbf{y}\|^2)/|\mathbf{x} \cdot \mathbf{y}|^2 \geq 0.$$

Cancelling the nonnegative number $\|\mathbf{x}\|^2$, which leaves the inequality sign unchanged; rearranging the terms; and taking the square root of the remaining nonnegative result on each side of the inequality yields the Cauchy–Schwarz inequality

$$|\mathbf{x}\cdot\mathbf{y}| \le \|\mathbf{x}\|\cdot\|\mathbf{y}\|.$$

To prove the triangle inequality (b) we set $\lambda = 1$ and start from the result

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x}\cdot\mathbf{y} + \|\mathbf{y}\|^2.$$

As $\mathbf{x}\cdot\mathbf{y}$ may be either positive or negative, $\mathbf{x}\cdot\mathbf{y} \le |\mathbf{x}\cdot\mathbf{y}|$, so making use of the Cauchy–Schwarz inequality shows that

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 &\le \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\cdot\|\mathbf{y}\| + \|\mathbf{y}\|^2\\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.\end{aligned}$$

The *triangle inequality* follows from taking the square root of each side of this inequality, which is permitted because both are nonnegative numbers. ■

The dot product in R^3 allowed the angle between vectors to be determined and, more importantly, it provided a test for the orthogonality of vectors. These same geometrical ideas can be introduced into the vector space R^n if the Cauchy–Schwarz inequality is written in the form

$$-\|\mathbf{x}\|\cdot\|\mathbf{y}\| \le \mathbf{x}\cdot\mathbf{y} \le \|\mathbf{x}\|\cdot\|\mathbf{y}\|.$$

After division by the nonnegative number $\|\mathbf{x}\|\cdot\|\mathbf{y}\|$, this becomes

$$-1 \le \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|} \le 1.$$

This enables the angle θ between the two n-vectors $\mathbf{x}$ and $\mathbf{y}$ to be defined by the result

$$\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\cdot\|\mathbf{y}\|}.$$

orthogonality of n-vectors

On account of this result, two n-vectors $\mathbf{x}$ and $\mathbf{y}$ in R^n will be said to be **orthogonal** when $\mathbf{x}\cdot\mathbf{y} = 0$.

unit n-vector

By analogy with R^3 we will call $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ a **unit n-vector** if $\|\mathbf{x}\| = 1$. If we define the unit n-vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ as

$$\mathbf{e}_1 = (1, 0, 0, 0, \ldots, 0),\ \mathbf{e}_2 = (0, 1, 0, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, 0, 0, 0, \ldots, 1),$$

we see that

$$\mathbf{e}_i\cdot\mathbf{e}_j = \begin{cases}1 & \text{for } i = j\\ 0 & \text{for } i \ne j,\end{cases}$$

showing that the $\mathbf{e}_i$ are mutually orthogonal unit n-vectors in R^n. As a result of this the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ play the same role in R^n as the vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ in R^3. This allows the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ to be written as

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n,$$

where x_i is the ith component of $\mathbf{x}$.

Now suppose that for $n > 3$, we set

$$\mathbf{u}_1 = (1, 0, 0, 0, \ldots, 0), \quad \mathbf{u}_2 = (0, 1, 0, 0, \ldots, 0), \quad \mathbf{u}_3 = (0, 0, 1, 0, \ldots, 0),$$

and all other $\mathbf{u}_i$ identically zero, so that $\mathbf{u}_i = (0, 0, 0, 0, \ldots, 0)$ for $i = 4, 5, \ldots, n$. Then it is not difficult to see that $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3$ behave like the unit vectors **i**, **j**, and **k**, so that, in some sense the vector space R^3 is embedded in the vector space R^n with vectors in both spaces obeying the same algebraic rules for addition and scaling. This is recognized by saying that R^3 is a **subspace** of R^n.

subspaces

Subspace of R^n

A subset S of vectors in the vector space R^n is called a **subspace** of R^n if S is itself a vector space that obeys the rules for the addition and the scaling of vectors in R^n.

EXAMPLE 2.19

Find the condition that the set S of vectors of the form $(x, mx + c, 0)$, for any m and all real x forms a subspace of the vector space R^3, and give a geometrical interpretation of the result.

Solution The set S can only contain the null vector $(0, 0, 0)$ if $c = 0$, so if $c \neq 0$ the vectors in S cannot form a subspace of R^3. Now let $c = 0$, so that S contains the null vector. The vector addition law holds, because if $(x, mx, 0)$ and $(x', mx', 0)$ are vectors in S, the sum

$$(x, mx, 0) + (x', mx', 0) = (x + x', m(x + x'), 0)$$

is also a vector in S. The scaling $\lambda(x, mx, 0) = (\lambda x, m\lambda x, 0)$ also generates a vector in S, so the scaling law for vectors also holds, showing that S is a subspace of R^3 provided $c = 0$.

If the three components of vectors in S are regarded as the x-, y-, and z-components of a vector in R^3, the vectors can be interpreted as points on the straight line $y = mx$ passing through the origin and lying in the plane $z = 0$. This subspace is a one-dimensional vector space embedded in the three-dimensional vector space R^3. ■

EXAMPLE 2.20

Test the following subsets of R^n to determine if they form a subspace.

(a) S is the set of vectors $(x_1, x_1 + 1, \ldots, x_n)$ with all the x_i real numbers.

(b) S is the set of vectors $(x_1, x_2, \ldots, x_n)$ with $x_1 + x_2 + \cdots + x_n = 0$ and all the x_i are real numbers.

Solution **(a)** The set S does not contain the null vector and so cannot form a subspace of R^n. This result is sufficient to show that S is not a subspace, but to see what properties of a subspace the set S possesses we consider both the summation and scaling of vectors in S. If $(x_1, x_1 + 1, \ldots, x_n)$ and $(x_1', x_1' + 1, \ldots, x_n')$ are two vectors in S, their sum

$$(x_1, x_1 + 1, \ldots, x_n) + (x_1', x_1' + 1, \ldots, x_n') = (x_1 + x_1', x_1 + x_1' + 2, \ldots, x_n + x_n')$$

is *not* a vector in S, so the summation law is not satisfied.

The scaling condition for vectors is not satisfied, because if λ is an arbitrary scalar,

$$\lambda(x_1, x_1+1, \ldots, x_n) = (\lambda x_1, \lambda x_1 + \lambda, \ldots, \lambda x_n) \neq (a, a+1, \ldots), \quad (\lambda n_1 = a)$$

showing that scaling generates another a vector in S. We have proved that the vectors in S do *not* form a subspace of R^n.

(b) The set S does contain the null vector, because $x_1 = x_2 = \cdots = x_n = 0$ satisfies the constraint condition $x_1 + x_2 + \cdots + x_n = 0$. Both the summation law and the scaling law for vectors are easily seen to be satisfied, so this set S does form a subspace of R^n. ■

EXAMPLE 2.21 Let $C(a, b)$ be the space of all real functions of a single real variable x that are continuous for $a < x < b$, and let $S(a, b)$ be the set of all functions belonging to $C(a, b)$ that have a derivative at every point of the interval $a < x < b$. Show that $S(a, b)$ forms a subspace of $C(a, b)$.

Solution In this case a vector in the space is simply any real function of a single real variable x that is continuous in the interval $a < x < b$. The null vector corresponds to the continuous function that is identically zero in the stated interval, so as the derivative of this function is also zero, it follows that the set $S(a, b)$ must also contain the null vector. The sum of continuous functions in $a < x < b$ is a continuous function, and the sum of differentiable functions in this same interval is a differentiable function, so the summation law for vectors is satisfied. Similarly, scaling continuous functions and differentiable functions does not affect either their continuity or their differentiability, so the scaling law for vectors is also satisfied. Thus, $S(a, b)$ forms a subspace of $C(a, b)$. Think of the dimension of these spores as infinite; norm and inner product are easy to define. ■

Summary

This section generalized the concept of a three-dimensional vector to a vector with n components in R^n. It was shown that the magnitude of a vector in three space dimensions generalizes to the norm of a vector in R^n and that in terms of components, the equality, addition, and scaling of vectors in R^n follow the same pattern as with three space dimensions. The dot product was generalized and two fundamental inequalities for vectors in R^n were derived. The concept of orthogonality of vectors was generalized and the notion of a subspace of R^n was introduced.

EXERCISES 2.5

In Exercises 1 through 8 find the sum of the given pairs of vectors, their norms, and their dot product.

1. $(2, 1, 0, 2, 2), (1, -1, 2, 2, 4)$.
2. $(3, -1, -1, 2, -4), (1, 2, 0, 0, 3)$.
3. $(2, 1, -1, 2, 1), (-2, -1, 1, -2, -1)$.
4. $(3, -2, 1, 1, 2, 0, 1), (1, -1, 1, -1, 1, 0, 1)$.
5. $(3, 0, 1, 0), (0, 2, 0, 4)$.
6. $(1, -1, 2, 2, 0, 1), (2, -2, 1, 1, 1, 0)$.
7. $(-1, 2, -4, 0, 1), (2, -1, 1, 0, 2)$.
8. $(3, 1, 2, 4, 1, 1, 1), (1, 2, 3, -1, -2, 1, 3)$.

In Exercises 9 through 12 find the angle between the given pairs of n-vectors and the unit n-vector associated with each vector.

9. $(3, 1, 2, 1), (1, -1, 2, 2)$.
10. $(4, 1, 0, 2), (2, -1, 2, 1)$.
11. $(2, -2, -2, 4), (1, -1, -1, 2)$.
12. $(2, 1, -1, 1), (1, -2, 2, 2)$.

In Exercises 13 through 18 determine if the set of vectors S forms a subspace of the given vector space. Give reasons why S either is or is not a subspace.

13. S is the set of vectors of the form $(x_1, x_2, \ldots, x_n)$ in R^n, with the x_i real numbers and $x_2 = x_1^4$.
14. S is the set of vectors of the form $(x_1, x_2, \ldots, x_n)$ in R^n, with the x_i real numbers and $x_1 + 2x_2 + 3x_3 + \cdots + nx_n = 0$.
15. S is the set of vectors of the form $(x_1, x_2, \ldots, x_n)$ in R^n, with the x_i real numbers and $x_1 + x_2 + x_3 + \cdots + x_n = 2$.
16. S is the set of vectors of the form $(x_1, x_2, \ldots, x_6)$ in R^6, with the x_i real numbers and $x_1 = 0$ or $x_6 = 0$.
17. S is the set of vectors of the form $(x_1, x_2, \ldots, x_6)$ in R^6, with the x_i real numbers and $x_1 - x_2 + x_3 \cdots + x_6 = 0$.
18. S is the set of vectors of the form $(x_1, x_2, \ldots, x_5)$ in R^5, with the x_i real numbers and $x_2 < x_3$.

In Exercises 19 to 23 determine if the given set S is a subspace of the space $C[0, 1]$ of all real valued functions that are continuous on the interval $0 \le x \le 1$. Give reasons why either S is a subspace, or it is not.

19. S is the set of all polynomials of degree two.
20. S is the set of all polynomial functions.
21. S is the set of all continuous functions such that $f(0) = f(1) = 0$.
22. S is the set of all continuous functions such that $f(0) = 0$ and $f(1) = 2$.
23. S is the set of all continuous once differentiable functions such that $f(0) = 0$ and $f'(x) > 0$.
24. Prove that the set S of all vectors lying in any plane in R^3 that passes through the origin forms a subspace of R^3.
25. Explain why the set S of all vectors lying in any plane in R^3 that does not pass through the origin does not form a subspace of R^3.
26. Consider the polynomial $P(\lambda)$ defined as

$$P(\lambda) = \|\mathbf{x} + \lambda\mathbf{y}\|^2,$$

where $\mathbf{x}$ and $\mathbf{y}$ are vectors in R^n. Show, provided not both $\mathbf{x}$ and $\mathbf{y}$ are null vectors, that the graph of $P(\lambda)$ as a function of λ is nonnegative, so $P(\lambda) = 0$ cannot have real roots. Use this result to prove the Cauchy–Schwarz inequality

$$|\mathbf{x} \cdot \mathbf{y}| \le \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

27. Let $\mathbf{x}$ and $\mathbf{y}$ be vectors in R^n and λ be a scalar. Prove that

$$\|\mathbf{x} + \lambda\mathbf{y}\|^2 + \|\mathbf{x} - \lambda\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \lambda^2\|\mathbf{y}\|^2).$$

28. If $\mathbf{x}$ and $\mathbf{y}$ are orthogonal vectors in R^n, prove that the Pythagoras theorem takes the form

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

29. What conditions on the components of vectors $\mathbf{x}$ and $\mathbf{y}$ in the Cauchy–Schwarz inequality cause it to become an equality, so that

$$\sum_{i=1}^{n} x_i y_i = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \left(\sum_{i=1}^{n} y_i^2\right)^{1/2}?$$

30. Modify the method of proof used in Theorem 2.7 to prove the complex form of the Cauchy–Schwarz inequality

$$\left|\sum_{i=1}^{n} x_i y_i\right| \le \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} + \left(\sum_{i=1}^{n} |y_i|^2\right)^{1/2},$$

where the x_i and y_i are complex numbers.

2.6 Linear Independence, Basis, and Dimension

The concept of the linear independence of a set of vectors in R^3 introduced in Section 2.4 generalizes to R^n and involves a linear combination of n-vectors.

Linear combination of n-vectors

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be a set of n-vectors in R^n. Then a **linear combination** of the n-vectors is a sum of the form

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m,$$

where $c_1, c_2, \ldots, c_m$ are nonzero scalars.

An example of a linear combination of vectors in R^5 is provided by the vector sum ($m = 3, n = 5$)

$$2\mathbf{x}_1 + \mathbf{x}_2 + 3\mathbf{x}_3,$$

where $\mathbf{x}_1 = (1, 2, 3, 0, 4)$, $\mathbf{x}_2 = (2, 1, 4, 1, -3)$, and $\mathbf{x}_3 = (6, 0, 2, 2, -1)$. The vector in R^5 formed by this linear combination is

$$\begin{aligned} 2\mathbf{x}_1 + \mathbf{x}_2 + 3\mathbf{x}_3 &= 2(1, 2, 3, 0, 4) + (2, 1, 4, 1, -3) + 3(6, 0, 2, 2, -1), \\ &= (22, 5, 16, 7, 2). \end{aligned}$$

A linear combination of n-vectors is the most general way of combining n-vectors, and the definition of a linear combination of vectors contains within it the definition of the scaling of a single n-vector as a special case. This can be seen by setting $m = 1$, because this reduces the linear combination to the single scaled n-vector $c_1\mathbf{x}_1$.

linear dependence and independence of n-vectors

Linear dependence of n-vectors

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be a set of n-vectors in R^n. Then the set is said to be **linearly dependent** if, and only if, one of the n-vectors can be expressed as a linear combination of the remaining n-vectors.

An example of linear dependence in R^4 is provided by the vectors $\mathbf{x}_1 = (1, 0, 2, 5)$, $\mathbf{x}_2 = (2, 1, 2, 1)$, $\mathbf{x}_3 = (3, 2, 1, 0)$, and $\mathbf{x}_4 = (-1, -1, -1, 7)$, because

$$\mathbf{x}_4 = 2\mathbf{x}_1 - 3\mathbf{x}_2 + \mathbf{x}_3.$$

Linear independence of n-vectors

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be a set of n-vectors in R^n. Then the set is said to be **linearly independent** if, and only if, the n-vectors are not linearly dependent.

A simple example of a set of linearly independent vectors in R^4 is provided by the vectors $\mathbf{e}_1 = (1, 0, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0, 0)$, and $\mathbf{e}_3 = (0, 0, 1, 0)$. The linear independence of these 4-vectors can be seen from the fact that for no choice of c_1 and c_2 can the vector $c_1\mathbf{e}_1 + c_2\mathbf{e}_2$ be made equal to $\mathbf{e}_3$.

To make effective use of the concept of linear independence, and to understand the notion of the *basis* and *dimension* of a vector space, it is necessary to have a test for linear independence. Such a test is provided by the following theorem.

THEOREM 2.8

Linear dependence and independence Let S be a set of non-zero n-vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$, with $m \geq 2$. Then:

(a) Set S is linearly dependent if the vector equation

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m = \mathbf{0}$$

is true for some set of scalars (constants) $c_1, c_2, \ldots, c_m$ that are not all zero;

(b) Set S is linearly independent if the vector equation

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m = \mathbf{0}$$

is only true when $c_1 = c_2 = \cdots = c_m = 0$.

Proof To establish result (a) it is necessary to show that the conditions of the definition of linear dependence are satisfied. First, if the set S of n-vectors is linearly dependent, scalars $d_1, d_2, \ldots, d_m$ exist such that

$$d_1\mathbf{x}_1 + d_2\mathbf{x}_2 + \cdots + d_m\mathbf{x}_m = \mathbf{0}.$$

There is no loss of generality in assuming that $d_1 \neq 0$, because if this is not the case a renumbering of the vectors can always make this possible. Consequently,

$$\mathbf{x}_1 = (-d_2/d_1)\mathbf{x}_2 + (-d_3/d_1)\mathbf{x}_3 + \cdots + (-d_m/d_1)\mathbf{x}_m,$$

which shows, as claimed, that the set S is linearly dependent, because $\mathbf{x}_1$ is linearly dependent on $\mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_m$. A similar argument applies to show that $\mathbf{x}_r$ is linearly dependent on the remaining n-vectors in S provided $d_r \neq 0$, for $r = 2, 3, \ldots, m$.

Conversely, if one of the n-vectors in set S, say $\mathbf{x}_1$, is linearly dependent on the remaining n-vectors in the set, scalars $d_1, d_2, \ldots, d_m$ can be found such that

$$\mathbf{x}_1 = d_2\mathbf{x}_2 + \cdots + d_m\mathbf{x}_m,$$

so that

$$\mathbf{x}_1 - d_2\mathbf{x}_2 - \cdots - d_m\mathbf{x}_m = 0.$$

This result is of the form given in definition of linear dependence with $c_1 = 1$, $c_2 = -d_2, \ldots, c_m = -d_m$, not all of which constants are zero, so again the set of n-vectors in S is seen to be linearly dependent.

To establish result (b), suppose, if possible, that the set S of vectors is linearly independent, but that some scalars $d_1, d_2, \ldots, d_m$ that are not all zero can be found such that

$$d_1\mathbf{x}_1 + d_2\mathbf{x}_2 + \cdots + d_m\mathbf{x}_m = \mathbf{0}.$$

Then if $d_1 \neq 0$, say, is one of these scalars, it follows that

$$\mathbf{x}_1 = (-d_2/d_1)\mathbf{x}_2 + (-d_3/d_1)\mathbf{x}_3 + \cdots + (-d_m/d_1)\mathbf{x}_m,$$

which is impossible because this shows that, contrary to the hypothesis, $\mathbf{x}_1$ is linearly dependent on the remaining n-vectors in S. So we must have $c_1 = c_2 = \cdots = c_m = 0$. ■

A systematic and efficient computational method for the application of Theorem 2.8 to vectors in R^n will be developed in the next chapter for the three separate cases that arise, (a) $m < n$, (b) $m = n$, and (c) $m > n$. However, when n and m are small, a straightforward approach is possible, as illustrated in the next example.

EXAMPLE 2.22

Test the following sets of vectors in R^4 for linear dependence or independence.

(a) $\mathbf{x}_1 = (2, 1, 1, 0)$, $\mathbf{x}_2 = (0, 2, 0, 1)$, $\mathbf{x}_3 = (1, 1, 0, 2)$, $\mathbf{x}_4 = (0, 2, 1, 1)$.

(b) $\mathbf{x}_1 = (4, 0, 2)$, $\mathbf{x}_2 = (2, 2, 0)$, $\mathbf{x}_3 = (1, 1, 0)$, $\mathbf{x}_4 = (5, 1, 2)$.

Solution In both (a) and (b) it is necessary to consider the vector equation

$$c_1\mathbf{x}_1 + \cdots + c_m\mathbf{x}_m = \mathbf{0}.$$

If the equation is only satisfied when $c_1 = c_2 = \cdots = c_m = 0$, the set of vectors will be linearly independent, whereas if a solution can be found in which not all of the constants c_1, c_2, c_3, c_4 vanish, the set of vectors will be linearly dependent.

(a) Substituting for $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ in the preceding equation and equating corresponding components show the coefficients c_i must satisfy the following equations

$$\begin{aligned} 2c_1 + c_3 &= 0 \\ c_1 + 2c_2 + c_3 + 2c_4 &= 0 \\ c_1 + c_4 &= 0 \\ c_2 + 2c_3 + c_4 &= 0. \end{aligned}$$

The third equation shows that $c_4 = -c_1$, so the equations can be rewritten as

$$\begin{aligned} 2c_1 + c_3 &= 0 \\ -c_1 + 2c_2 + c_3 &= 0 \\ c_2 - c_1 + 2c_3 &= 0. \end{aligned}$$

Adding twice the third equation to the first equation shows that $c_3 = 0$, so $c_1 = 0$, and it then follows that $c_2 = c_3 = c_4 = 0$. This has established the linear independence of the set of vectors in (a).

(b) Proceeding in the same manner with the set of vectors in (b) leads to the following equations for the coefficients c_i:

$$\begin{aligned} 4c_1 + 2c_2 + c_3 + 5c_4 &= 0 \\ 2c_2 + c_3 + c_4 &= 0 \\ 2c_1 + 2c_4 &= 0. \end{aligned}$$

The third equation shows that $c_4 = -c_1$, so using this result in the first two equations reduces the first one to

$$-c_1 + 2c_2 + c_3 = 0$$

and the second to

$$-c_1 + 2c_2 + c_3 = 0.$$

There is only one equation connecting c_1, c_2, and c_3, and hence also c_4. This means that if c_2 and c_3 are given arbitrary values, not both of which are zero, the constants c_1 and c_4 will be determined in terms of them. Thus, a set of constants c_1, c_2, c_3, c_4 that are not all zero can be found that satisfy the vector equation, showing that the set of vectors in (b) is linearly dependent. This set of constants is not unique, but this does affect the conclusion that the set of vectors is linearly dependent, because to establish linear dependence it is sufficient that at least one such set of constants can be found. ■

Example 2.22 has shown one way in which Theorem 2.8 can be implemented for vectors in R^n, but it also illustrates the need for a systematic approach to the solution of the system of equations for the coefficients when n is large.

A trivial case of Theorem 2.8 arises when the set of vectors S contains the null vector $\mathbf{0}$, because then the set of vectors in S is always linearly dependent. This can be seen by assuming that $\mathbf{x}_1 = \mathbf{0}$, because then the vector equation in the theorem becomes

$$c_1\mathbf{0} + c_2\mathbf{x}_2 + \cdots + c_m\mathbf{x}_m = \mathbf{0}.$$

This vector equation is satisfied if $c_1 \neq 0$ (arbitrary) and $c_2 = c_3 = \cdots = c_m = 0$, so, as not all of the coefficients are zero, the set of vectors must be linearly dependent.

We conclude this introduction to the vector space R^n by defining the *span*, a *basis*, and the *dimension* of a vector space.

span of a vector space

Span of a vector space

Let the set of non-zero vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ belonging to a vector space V have the property that every vector in V can be expressed as a linear combination of these vectors. Then the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ are said to **span** the vector space V.

EXAMPLE 2.23

All vectors $\mathbf{v}$ in the (x, y)-plane are spanned by the vectors $\mathbf{i}$ and $\mathbf{j}$, because any vector $\mathbf{v} = (v_1, v_2)$ can always be written $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j}$. This is an example of vectors spanning the space R^2. ■

EXAMPLE 2.24

The vector space R^n is spanned by the unit n-vectors

$$\mathbf{e}_1 = (1, 0, 0, 0, \ldots, 0),\ \ \mathbf{e}_2 = (0, 1, 0, 0, \ldots, 0), \ldots,\ \ \mathbf{e}_n = (0, 0, 0, 0, \ldots, 1).$$ ■

EXAMPLE 2.25

The subspace R^3 of the vector space R^5 is spanned by the unit vectors

$$\mathbf{e}_1 = (1, 0, 0, 0, 0),\ \ \mathbf{e}_2 = (0, 1, 0, 0, 0),\ \ \mathbf{e}_3 = (0, 0, 1, 0, 0),$$

because all vectors $\mathbf{v} = (v_1, v_2, v_3)$ in R^3 can be written in the form of the linear combination $\mathbf{v} = v_1\mathbf{e}_1 + v_1\mathbf{e}_2 + v_3\mathbf{e}_3$. ■

basis of a vector space in R^n

Basis of a vector space

Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be vectors in R^n. Then the vectors are said to form a **basis** for the vector space R^n if:

(i) The vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent.

(ii) Every vector in R^n can be expressed as a linear combination of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

dimension of a vector space

Dimension of a vector space

The **dimension** of a vector space is the number of vectors in its basis.

EXAMPLE 2.26 A basis for the space of ordinary vectors in three dimensions is provided by the vectors **i**, **j**, and **k**, so the dimension of the space is 3. ■

EXAMPLE 2.27 A basis for R^n is provided by the n vectors

$$\mathbf{e}_1 = (1, 0, 0, 0, \dots, 0),\ \ \mathbf{e}_2 = (0, 1, 0, 0, \dots, 0), \dots,\ \ \mathbf{e}_n = (0, 0, 0, 0, \dots, 1),$$

so its dimension is n. ■

EXAMPLE 2.28 It was shown in Example 2.20 (b) that the set S of vectors $(x_1, x_2, \dots, x_n)$ with $x_1 + x_2 + \cdots + x_n = 0$ forms a subspace of R^n. The dimension of R^n is n, but the constraint condition $x_1 + x_2 + \cdots + x_n = 0$ implies that only $n - 1$ of the components $x_1, x_2, \dots, x_n$ can be specified independently, because the constraint itself determines the value of the remaining component. This in turn implies that the basis for the subspace S can only contain $n - 1$ linearly independent vectors, so S must have dimension $n - 1$. ■

More information on linear vector spaces can be found in references [2.1] and [2.5] to [2.12].

Summary

In this section the concepts of linear dependence and independence were generalized to vectors in R^n, and the span of a vector space was defined as a set of vectors in R^n with the property that every vector in R^n can be expressed as a linear combination of these vectors. Naturally in R^n, as in R^3, a set of vectors spanning the space is not unique. The smallest set of n vectors spanning a vector space is said to form a basis for the vector space, and the dimension of a vector space is the number of vectors in its basis. This corresponds to the fact that the unit vectors **i**, **j**, and **k** form a basis for the ordinary three-dimensional space R^3, because every vector in this space can be represented as a linear combination of **i**, **j**, and **k**.

EXERCISES 2.6

In Exercises 1 through 12 determine if the set of m vectors in three-dimensional space is linearly independent by solving for the scalars $c_1, c_2 \dots c_m$ in Theorem 2.8. Where appropriate, verify the result by using Theorem 2.3.

1. $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + \mathbf{k},\ \mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k},\ \mathbf{c} = 2\mathbf{i} + \mathbf{k}$.

2. $\mathbf{a} = 3\mathbf{i} - \mathbf{j} + \mathbf{k},\ \mathbf{b} = \mathbf{i} + 3\mathbf{k},\ \mathbf{c} = 5\mathbf{i} - \mathbf{j} + 7\mathbf{k}$.

3. $\mathbf{a} = 2\mathbf{i} - \mathbf{j} + \mathbf{k},\ \mathbf{b} = 3\mathbf{i} + \mathbf{j} - \mathbf{k},\ \mathbf{c} = 8\mathbf{i} + \mathbf{j} + 7\mathbf{k}$.

4. $\mathbf{a} = 3\mathbf{i} + 2\mathbf{k},\ \mathbf{b} = \mathbf{i} + \mathbf{j} + 2\mathbf{k},\ \mathbf{c} = 11\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$.

5. $\mathbf{a} = 4\mathbf{i} - \mathbf{j} + 3\mathbf{k},\ \mathbf{b} = \mathbf{i} + 4\mathbf{j} - 2\mathbf{k},\ \mathbf{c} = 3\mathbf{i} - \mathbf{j} - \mathbf{k}$.

6. $\mathbf{a} = \mathbf{i} + \mathbf{j} - \mathbf{k},\ \mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k},\ \mathbf{c} = -\mathbf{i} + \mathbf{j} + \mathbf{k}$.

7. $\mathbf{a} = \mathbf{i} + 2\mathbf{j} + \mathbf{k},\ \mathbf{b} = \mathbf{i} + 3\mathbf{j} - \mathbf{k},\ \mathbf{c} = 3\mathbf{i} + 10\mathbf{j} - 5\mathbf{k}$.

8. $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j} + \mathbf{k},\ \mathbf{b} = \mathbf{i} - 3\mathbf{j} + 2\mathbf{k},\ \mathbf{c} = \mathbf{i} + 15\mathbf{j} - 4\mathbf{k}$.

9. $\mathbf{a} = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k},\ \mathbf{b} = \mathbf{i} + \mathbf{j} + \mathbf{k}\quad (m = 2)$.

10. $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k},\ \mathbf{b} = \mathbf{i} + 2\mathbf{j} + \mathbf{k},\ \mathbf{c} = \mathbf{i} + 3\mathbf{j} + \mathbf{k},\ \mathbf{d} = \mathbf{i} - 4\mathbf{j} + \mathbf{k}\quad (m = 4)$.

11. $\mathbf{a} = \mathbf{i} - \mathbf{j} + 3\mathbf{k},\ \mathbf{b} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k},\ \mathbf{c} = \mathbf{i} + \mathbf{k},\ \mathbf{d} = 3\mathbf{i} + \mathbf{j} + \mathbf{k}\quad (m = 4)$.

12. $\mathbf{a} = \mathbf{i} + \mathbf{j},\ \mathbf{b} = \mathbf{j} + \mathbf{k},\ \mathbf{c} = \mathbf{i} - \mathbf{k}$.

In Exercises 13 through 16, determine if the set of vectors in R^4 is linearly independent by using the method of Example 2.22.

13. $(1, 3, -1, 0), (1, 2, 0, 1), (0, 1, 0, -1), (1, 1, 0, 1)$.

14. $(1, -2, 1, 2), (4, -1, 0, 2), (2, 1, -1, 1), (1, 0, 0, -1)$.

15. $(2, 1, 0, 1), (1, 0, 1, 1), (4, 1, 2, -1), (1, 0, 1, -1)$.

16. $(1, 2, 1, 1), (1, -2, 0, -1), (1, 1, 1, 2), (1, -1, 0, 0)$.

In Exercises 17 through 20, find a basis and the dimension of the given subspace S.

17. The subspace S of vectors in R^5 of the form $(x_1, x_2, x_3, x_4, x_5)$ with $x_1 = x_2$.

18. The subspace S of vectors in R^4 of the form (x_1, x_2, x_3, x_4) with $x_1 = 2x_2$.

19. The subspace S of vectors in R^5 of the form $(x_1, x_2, x_3, x_4, x_5)$ with $x_1 = x_2 = 2x_3$.

20. The subspace S of vectors in R^6 of the form $(x_1, x_2, x_3, x_4, x_5, x_6)$ with $x_1 = 2x_2$ and $x_3 = -x_4$.

21. Let $\mathbf{u} = \cos^2 x$ and $\mathbf{v} = \sin^2 x$ form a basis for a vector space V. Find which of the following can be represented in terms of $\mathbf{u}$ and $\mathbf{v}$, and so lie in V.

(a) 2. (b) $\sin 2x$. (c) 0. (d) $\cos 2x$. (e) $2 + 3x$. (f) $3 - 4\cos 2x$.

22. Given that $r \leq n$, prove that any subset S of r vectors selected from a set of n linearly independent vectors is linearly independent.

2.7 Gram–Schmidt Orthogonalization Process

A set of vectors forming a basis for a vector space is not unique, and having obtained a basis by some means, it is often useful to replace it by an equivalent set of orthogonal vectors. The **Gram–Schmidt orthogonalization process** accomplishes this by means of a sequence of simple steps that have a convenient geometrical interpretation. We now develop the Gram–Schmidt orthogonalization process for geometrical vectors in R^3, though in Section 4.2 the method will be extended to vectors in R^n to enable orthogonal matrices to be constructed from a set of eigenvectors associated with a symmetric matrix.

Let us now show how any basis for R^3, comprising three nonorthogonal linearly independent vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, can be used to construct an equivalent basis involving three linearly independent orthogonal vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$. It is essential that the vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ be linearly independent, because if not, the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ generated by the Gram–Schmidt orthogonalization process will be linearly dependent and so cannot form a basis for R^3. The derivation of the method starts by setting

$$\mathbf{u}_1 = \mathbf{a}_1,$$

where the choice of $\mathbf{a}_1$ instead of $\mathbf{a}_2$ or $\mathbf{a}_3$ is arbitrary.

The component of $\mathbf{a}_2$ in the direction of $\mathbf{u}_1$ is $\mathbf{u}_1 \cdot \mathbf{a}_2$, so the vector component of $\mathbf{a}_2$ in this direction is

$$(\mathbf{u}_1 \cdot \mathbf{a}_2)\mathbf{u}_1 = \frac{(\mathbf{u}_1 \cdot \mathbf{a}_2)\mathbf{u}_1}{\|\mathbf{u}_1\|^2},$$

and this always exists because $\|\mathbf{u}_1\|^2 > 0$. Subtracting this vector from $\mathbf{a}_2$ gives a vector $\mathbf{u}_2$ that is normal to $\mathbf{u}_1$, so

$$\mathbf{u}_2 = \mathbf{a}_2 - \frac{(\mathbf{u}_1 \cdot \mathbf{a}_2)\mathbf{u}_1}{\|\mathbf{u}_1\|^2}.$$

Similarly, to find a vector normal to both $\mathbf{u}_1$ and $\mathbf{u}_2$ involving $\mathbf{a}_3$, it is necessary to subtract from $\mathbf{a}_3$ the components of vector $\mathbf{a}_3$ in the direction of $\mathbf{u}_1$ and also in the direction of $\mathbf{u}_2$, so that

$$\mathbf{u}_3 = \mathbf{a}_3 - \frac{(\mathbf{u}_1 \cdot \mathbf{a}_3)\mathbf{u}_1}{\|\mathbf{u}_1\|^2} - \frac{(\mathbf{u}_2 \cdot \mathbf{a}_3)\mathbf{u}_2}{\|\mathbf{u}_2\|^2},$$

and this also always exists, because $\|\mathbf{u}_1\|^2 > 0$ and $\|\mathbf{u}_2\|^2 > 0$.

If an orthonormal basis is required, it is necessary to normalize the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ by dividing each by its norm.

Rule for the Gram–Schmidt orthogonalization process in R^3

A set of nonorthogonal linearly independent vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ that form a basis in R^3 can be used to generate an equivalent orthogonal basis involving the vectors, $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ by setting

$$\mathbf{u}_1 = \mathbf{a}_1, \quad \mathbf{u}_2 = \mathbf{a}_2 - \frac{(\mathbf{u}_1 \cdot \mathbf{a}_2)\mathbf{u}_1}{\|\mathbf{u}_1\|^2}, \quad \text{and}$$

$$\mathbf{u}_3 = \mathbf{a}_3 - \frac{(\mathbf{u}_1 \cdot \mathbf{a}_3)\mathbf{u}_1}{\|\mathbf{u}_1\|^2} - \frac{(\mathbf{u}_2 \cdot \mathbf{a}_3)\mathbf{u}_2}{\|\mathbf{u}_2\|^2}.$$

As already remarked, the choice of $\mathbf{a}_1$ as the vector with which to start the orthogonalization process was arbitrary, and the process could equally well have been started by setting $\mathbf{u}_1 = \mathbf{a}_2$ or $\mathbf{u}_1 = \mathbf{a}_3$. Using a different vector will produce a different set of orthogonal vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, but any basis for R^3 is equivalent to any other basis, so unless there is a practical reason for starting with a particular vector, the choice is immaterial.

EXAMPLE 2.29

Given the nonorthogonal basis $\mathbf{a}_1 = \mathbf{i} - \mathbf{j} - \mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} + \mathbf{j} + \mathbf{k}$, and $\mathbf{a}_3 = -\mathbf{i} + 2\mathbf{k}$, use the Gram–Schmidt orthogonalization process to find an equivalent orthogonal basis, and then find the corresponding orthonormal basis.

Solution Using the preceding rule we start with $\mathbf{u}_1 = \mathbf{i} - \mathbf{j} - \mathbf{k}$, and to find $\mathbf{u}_2$ we need to use the results $\mathbf{u}_1 \cdot \mathbf{a}_2 = -1$ and $\|\mathbf{u}_1\|^2 = 3$, so that

$$\mathbf{u}_2 = \mathbf{i} + \mathbf{j} + \mathbf{k} - (-1/3)(\mathbf{i} - \mathbf{j} - \mathbf{k}) = (4/3)\mathbf{i} + (2/3)\mathbf{j} + (2/3)\mathbf{k}.$$

To find $\mathbf{u}_3$ we need to use the results $\mathbf{u}_1 \cdot \mathbf{a}_3 = -3$, $\|\mathbf{u}_1\|^2 = 3$, $\mathbf{u}_2 \cdot \mathbf{a}_3 = 0$, and $\|\mathbf{u}_2\|^2 = 24/9$, so that

$$\mathbf{u}_3 = -\mathbf{i} + 2\mathbf{k} - (-3/3)(\mathbf{i} - \mathbf{j} - \mathbf{k}) = -\mathbf{j} + \mathbf{k}.$$

So the required equivalent orthogonal basis is

$$\mathbf{u}_1 = \mathbf{i} - \mathbf{j} - \mathbf{k}, \quad \mathbf{u}_2 = (4/3)\mathbf{i} + (2/3)\mathbf{j} + 2/3\mathbf{k}, \quad \text{and} \quad \mathbf{u}_3 = -\mathbf{j} + \mathbf{k}.$$

The corresponding orthonormal basis obtained by dividing each of these vectors by its norm (modulus) is

$$\hat{\mathbf{u}}_1 = (1/\sqrt{3})\mathbf{u}_1, \quad \hat{\mathbf{u}}_2 = (1/2)\sqrt{(3/2)}\mathbf{u}_2 \quad \text{and} \quad \hat{\mathbf{u}}_3 = (1/\sqrt{2})\mathbf{u}_3.$$ ■

Other accounts of the Gram–Schmidt orthogonalization process are to be found in references [2.1] and [2.7] to [2.12].

Summary

In this section it is shown how in R^3 the Gram–Schmidt orthogonalization process converts any three nonorthogonal linearly independent vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ into three orthogonal vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$. If necessary, the vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ can then be normalized in the usual manner to form an orthogonal set of unit vectors.

EXERCISES 2.7

In Exercises 1 through 6, use the given nonorthogonal basis for vectors in R^3 to find an equivalent orthogonal basis by means of the Gram–Schmidt orthogonalization process.

1. $\mathbf{a}_1 = \mathbf{i} + 2\mathbf{j} + \mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} - \mathbf{j}$, $\mathbf{a}_3 = 2\mathbf{j} - \mathbf{k}$.

2. $\mathbf{a}_1 = \mathbf{j} + 3\mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} + \mathbf{j} - \mathbf{k}$, $\mathbf{a}_3 = \mathbf{i} + 2\mathbf{k}$.

3. $\mathbf{a}_1 = 2\mathbf{i} + \mathbf{j}$, $\mathbf{a}_2 = 2\mathbf{j} + \mathbf{k}$, $\mathbf{a}_3 = \mathbf{k}$.

4. $\mathbf{a}_1 = \mathbf{i} + 3\mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{a}_3 = 2\mathbf{i} + \mathbf{j}$.

5. $\mathbf{a}_1 = -\mathbf{i} + \mathbf{k}$, $\mathbf{a}_2 = 2\mathbf{j} + \mathbf{k}$, $\mathbf{a}_3 = \mathbf{i} + \mathbf{j} + \mathbf{k}$.

6. $\mathbf{a}_1 = \mathbf{i} + \mathbf{k}$, $\mathbf{a}_2 = -\mathbf{j} + \mathbf{k}$, $\mathbf{a}_3 = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$.

In Exercises 7 and 8, find two different but equivalent sets of orthogonal vectors by arranging the same three nonorthogonal vectors in the orders indicated.

7. (a) $\mathbf{a}_1 = 3\mathbf{j} - \mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} + \mathbf{j}$, $\mathbf{a}_3 = \mathbf{i} + 2\mathbf{k}$.

(b) $\mathbf{a}_1 = \mathbf{i} + \mathbf{j}$, $\mathbf{a}_2 = 3\mathbf{j} - \mathbf{k}$, $\mathbf{a}_3 = \mathbf{i} + 2\mathbf{k}$.

8. (a) $\mathbf{a}_1 = \mathbf{j} - \mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} + \mathbf{k}$, $\mathbf{a}_3 = -\mathbf{i} - \mathbf{j} + \mathbf{k}$.

(b) $\mathbf{a}_1 = -\mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{a}_2 = \mathbf{i} + \mathbf{k}$, $\mathbf{a}_3 = \mathbf{j} - \mathbf{k}$.