In Exercises 15, 16, and 17, assume the given differential equation has a power-series solution of the form $y = \sum a_n x^n$, and determine the $n$th coefficient $a_n$.

15. $y' = \alpha y$.
16. $y'' = xy$.
17. $y'' + xy' + y = 0$.
18. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$, where $a_0 = 1$ and the remaining coefficients are determined by the identity

$$e^{-2x} = \sum_{n=0}^{\infty} \{2a_n + (n+1)a_{n+1}\}x^n .$$

Compute $a_1, a_2, a_3$, and find the sum of the series for $f(x)$.

19. Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$, where the coefficients $a_n$ are determined by the relation

$$\cos x = \sum_{n=0}^{\infty} a_n(n+2)x^n .$$

Compute $a_5, a_6$, and $f(\pi)$.

20. (a) Show that the first six terms of the binomial series for $(1-x)^{-1/2}$ are:

$$1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + \frac{63}{256}x^5 .$$

(b) Let $a_n$ denote the $n$th term of this series when $x = 1/50$, and let $r_n$ denote the remainder after $n$ terms; that is, for $n \geq 0$ let

$$r_n = a_{n+1} + a_{n+2} + a_{n+3} + \cdots .$$

Show that $0 < r_n < a_n/49$.

[*Hint:* Show that $a_{n+1} < a_n/50$, and dominate $r_n$ by a suitable geometric series.]

(c) Verify the identity

$$\sqrt{2} = \frac{7}{5}\left(1 - \frac{1}{50}\right)^{-1/2}$$

and use it to compute the first ten correct decimals of $\sqrt{2}$.

[*Hint:* Use parts (a) and (b), retain twelve decimals during the calculations, and take into account round-off errors.]

21. (a) Show that

$$\sqrt{3} = \frac{1732}{1000}\left(1 - \frac{176}{3{,}000{,}000}\right)^{-1/2}$$

(b) Proceed as suggested in Exercise 20 and compute the first fifteen correct decimals of $\sqrt{3}$.

22. Integrate the binomial series for $(1-x^2)^{-1/2}$ and thereby obtain the power-series expansion

$$\arcsin x = x + \sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \frac{x^{2n+1}}{2n+1} \qquad (|x| < 1).$$

# 12

# VECTOR ALGEBRA

## 12.1 Historical introduction

In the foregoing chapters we have presented many of the basic concepts of calculus and have illustrated their use in solving a few relatively simple geometrical and physical problems. Further applications of the calculus require a deeper knowledge of analytic geometry than has been presented so far, and therefore we turn our attention to a more detailed investigation of some fundamental geometric ideas.

As we have pointed out earlier in this book, calculus and analytic geometry were intimately related throughout their historical development. Every new discovery in one subject led to an improvement in the other. The problem of drawing tangents to curves resulted in the discovery of the derivative; that of area led to the integral; and partial derivatives were introduced to investigate curved surfaces in space. Along with these accomplishments came other parallel developments in mechanics and mathematical physics. In 1788 Lagrange published his masterpiece *Mécanique analytique* (Analytical Mechanics) which showed the great flexibility and tremendous power attained by using analytical methods in the study of mechanics. Later on, in the 19th century, the Irish mathematician William Rowan Hamilton (1805–1865) introduced his *Theory of Quaternions*, a new method and a new point of view that contributed much to the understanding of both algebra and physics. The best features of quaternion analysis and Cartesian geometry were later united, largely through the efforts of J. W. Gibbs (1839–1903) and O. Heaviside (1850–1925), and a new subject called *vector algebra* sprang into being. It was soon realized that vectors are the ideal tools for the exposition and simplification of many important ideas in geometry and physics. In this chapter we propose to discuss the elements of vector algebra. Applications to analytic geometry are given in Chapter 13. In Chapter 14 vector algebra is combined with the methods of calculus, and applications are given to both geometry and mechanics.

There are essentially three different ways to introduce vector algebra: *geometrically*, *analytically*, and *axiomatically*. In the geometric approach, vectors are represented by directed line segments, or arrows. Algebraic operations on vectors, such as addition, subtraction, and multiplication by real numbers, are defined and studied by geometric methods.

In the analytic approach, vectors and vector operations are described entirely in terms of *numbers*, called *components*. Properties of the vector operations are then deduced from corresponding properties of numbers. The analytic description of vectors arises naturally from the geometric description as soon as a coordinate system is introduced.

In the axiomatic approach, no attempt is made to describe the nature of a vector or of the algebraic operations on vectors. Instead, vectors and vector operations are thought of as *undefined concepts* of which we know nothing except that they satisfy a certain set of axioms. Such an algebraic system, with appropriate axioms, is called a *linear space* or a *linear vector space*. Examples of linear spaces occur in all branches of mathematics, and we will study many of them in Chapter 15. The algebra of directed line segments and the algebra of vectors described by components are merely two examples of linear spaces.

The study of vector algebra from the axiomatic point of view is perhaps the most mathematically satisfactory approach to use since it furnishes a description of vectors that is free of coordinate systems and free of any particular geometric representation. This study is carried out in detail in Chapter 15. In this chapter we base our treatment on the analytic approach, and we also use directed line segments to interpret many of the results geometrically. When possible, we give proofs by coordinate-free methods. Thus, this chapter serves to provide familiarity with important concrete examples of vector spaces, and it also motivates the more abstract approach in Chapter 15.

## 12.2 The vector space of $n$-tuples of real numbers

The idea of using a number to locate a point on a line was known to the ancient Greeks. In 1637 Descartes extended this idea, using a *pair* of numbers $(a_1, a_2)$ to locate a point in the plane, and a *triple* of numbers $(a_1, a_2, a_3)$ to locate a point in space. The 19th century mathematicians A. Cayley (1821–1895) and H. G. Grassmann (1809–1877) realized that there is no need to stop with three numbers. One can just as well consider a *quadruple* of numbers $(a_1, a_2, a_3, a_4)$ or, more generally, an *$n$-tuple* of real numbers

$$(a_1, a_2, \ldots, a_n)$$

for any integer $n \geq 1$. Such an $n$-tuple is called an *$n$-dimensional point* or an *$n$-dimensional vector*, the individual numbers $a_1, a_2, \ldots, a_n$ being referred to as *coordinates* or *components* of the vector. The collection of all $n$-dimensional vectors is called *the vector space of $n$-tuples*, or simply *$n$-space*. We denote this space by $V_n$.

The reader may well ask at this stage why we are interested in spaces of dimension greater than three. One answer is that many problems which involve a large number of simultaneous equations are more easily analyzed by introducing vectors in a suitable $n$-space and replacing all these equations by a single vector equation. Another advantage is that we are able to deal in one stroke with many properties common to 1-space, 2-space, 3-space, etc., that is, properties independent of the dimensionality of the space. This is in keeping with the spirit of modern mathematics which favors the development of comprehensive methods for attacking problems on a

Unfortunately, the geom

To convert $V_n$ into an algebraic system, we introduce *equality* of vectors and two vector operations called *addition* and *multiplication by scalars.* The word "scalar" is used here as a synonym for "real number."

DEFINITION. *Two vectors $A$ and $B$ in $V_n$ are called equal whenever they agree in their respective components. That is, if $A = (a_1, a_2, \ldots, a_n)$ and $B = (b_1, b_2, \ldots, b_n)$, the vector equation $A = B$ means exactly the same as the n scalar equations*

$$a_1 = b_1, \quad a_2 = b_2, \quad \ldots, \quad a_n = b_n.$$

*The sum $A + B$ is defined to be the vector obtained by adding corresponding components:*

$$A + B = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

*If $c$ is a scalar, we define $cA$ or $Ac$ to be the vector obtained by multiplying each component of $A$ by $c$:*

$$cA = (ca_1, ca_2, \ldots, ca_n).$$

From this definition it is easy to verify the following properties of these operations.

THEOREM 12.1. *Vector addition is commutative,*

$$A + B = B + A,$$

*and associative,*

$$A + (B + C) = (A + B) + C.$$

*Multiplication by scalars is associative,*

$$c(dA) = (cd)A$$

*and satisfies the two distributive laws*

$$c(A + B) = cA + cB, \qquad \textit{and} \qquad (c + d)A = cA + dA.$$

Proofs of these properties follow quickly from the definition and are left as exercises for the reader.

The vector with all components 0 is called the *zero vector* and is denoted by $O$. It has the property that $A + O = A$ for every vector $A$; in other words, $O$ is an identity element for vector addition. The vector $(-1)A$ is also denoted by $-A$ and is called the *negative* of $A$. We also write $A - B$ for $A + (-B)$ and call this the *difference* of $A$ and $B$. The equation $(A + B) - B = A$ shows that subtraction is the inverse of addition. Note that $0A = O$ and that $1A = A$.

The reader may have noticed the similarity between vectors in 2-space and complex numbers. Both are defined as ordered pairs of real numbers and both are added in exactly the same way. Thus, as far as addition is concerned, complex numbers and two-dimensional vectors are algebraically indistinguishable. They differ only when we introduce multiplication.

Multiplication of complex numbers gives the complex-number system the field properties also possessed by the real numbers. It can be shown (although the proof is difficult) that except for $n = 1$ and 2, it is not possible to introduce multiplication in $V_n$ so as to satisfy all the field properties. However, special products can be introduced in $V_n$ which do not satisfy *all* the field properties. For example, in Section 12.5 we shall discuss the *dot product* of two vectors in $V_n$. The result of this multiplication is a scalar, not a vector. Another product, called the *cross product*, is discussed in Section 13.9. This multiplication is applicable only in the space $V_3$. The result is always a vector, but the cross product is not commutative.

## 12.3 Geometric interpretation for $n \leq 3$

Although the foregoing definitions are completely divorced from geometry, vectors and vector operations have an interesting geometric interpretation for spaces of dimension three or less. We shall draw pictures in 2-space to illustrate these concepts and ask the reader to produce the corresponding visualizations for himself in 3-space and in 1-space.

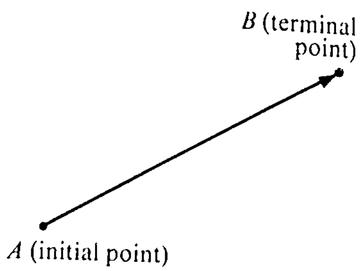


FIGURE 12.1 The geometric vector $\vec{AB}$ from $A$ to $B$.

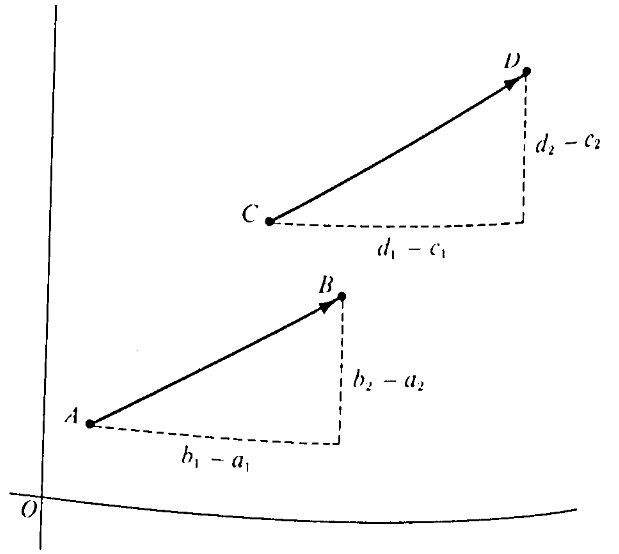


FIGURE 12.2 $\vec{AB}$ and $\vec{CD}$ are equivalent because $B - A = D - C$.

A pair of points $A$ and $B$ is called a *geometric vector* if one of the points, say $A$, is called the *initial point* and the other, $B$, the *terminal point*, or *tip*. We visualize a geometric vector as an arrow from $A$ to $B$, as shown in Figure 12.1, and denote it by the symbol $\vec{AB}$.

Geometric vectors are especially convenient for representing certain physical quantities such as force, displacement, velocity, and acceleration, which possess both magnitude and direction. The length of the arrow is a measure of the magnitude and the arrowhead indicates the required direction.

Suppose we introduce a coordinate system with origin $O$. Figure 12.2 shows two geometric vectors $\vec{AB}$ and $\vec{CD}$ with $B - A = D - C$. In terms of components, this means that we have

$$b_1 - a_1 = d_1 - c_1 \quad \text{and} \quad b_2 - a_2 = d_2 - c_2 .$$

By comparison of the congruent triangles in Figure 12.2, we see that the two arrows representing $\vec{AB}$ and $\vec{CD}$ have equal lengths, are parallel, and point in the same direction. We call such geometric vectors *equivalent*. That is, we say $\vec{AB}$ is equivalent to $\vec{CD}$ whenever

(12.1) $$B - A = D - C .$$

Note that the four points $A, B, C, D$ are vertices of a parallelogram. (See Figure 12.3.) Equation (12.1) can also be written in the form $A + D = B + C$ which tells us that *opposite vertices of the parallelogram have the same sum*. In particular, if one of the vertices, say $A$, is the origin $O$, as in Figure 12.4, the geometric vector from $O$ to the opposite vertex $D$ corresponds to the vector sum $D = B + C$. This is described by saying that vector addition corresponds geometrically to addition of geometric vectors by the *parallelogram law*. The importance of vectors in physics stems from the remarkable fact that many physical quantities (such as force, velocity, and acceleration) combine by the parallelogram law.

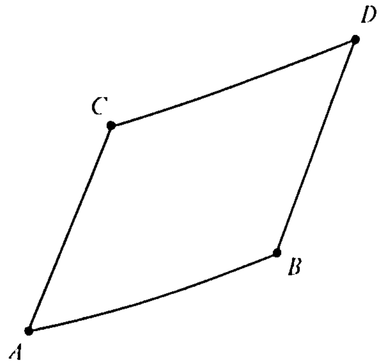


FIGURE 12.3 Opposite vertices of a parallelogram have the same sum: $A + D = B + C$.

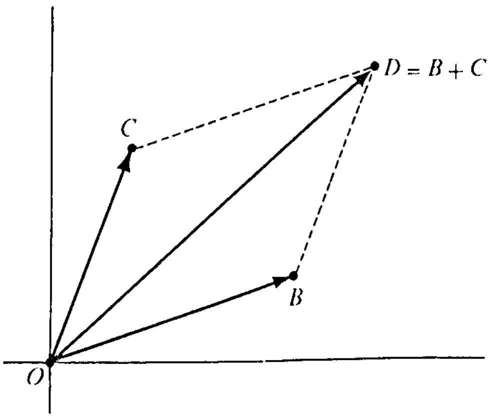


FIGURE 12.4 Vector addition interpreted geometrically by the parallelogram law.

For simplicity in notation, we shall use the same symbol to denote a point in $V_n$ (when $n \leq 3$) and the geometric vector from the origin to this point. Thus, we write $A$ instead of $\vec{OA}$, $B$ instead of $\vec{OB}$, and so on. Sometimes we also write $A$ in place of any geometric vector equivalent to $\vec{OA}$. For example, Figure 12.5 illustrates the geometric meaning of vector subtraction. Two geometric vectors are labeled as $B - A$, but these geometric vectors are equivalent. They have the same length and the same direction.

Figure 12.6 illustrates the geometric meaning of multiplication by scalars. If $B = cA$, the geometric vector $B$ has length $|c|$ times the length of $A$; it points in the same direction as $A$ if $c$ is positive, and in the opposite direction if $c$ is negative.

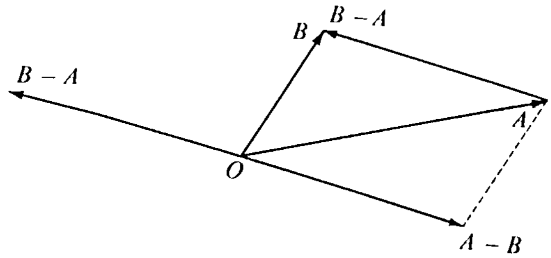


FIGURE 12.5 Geometric meaning of subtraction of vectors.

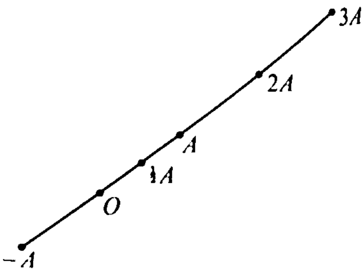


FIGURE 12.6 Multiplication of vectors by scalars.

The geometric interpretation of vectors in $V_n$ for $n \leq 3$ suggests a way to define parallelism in a general $n$-space.

DEFINITION. *Two vectors $A$ and $B$ in $V_n$ are said to have the same direction if $B = cA$ for some positive scalar $c$, and the opposite direction if $B = cA$ for some negative $c$. They are called parallel if $B = cA$ for some nonzero $c$.*

Note that this definition makes every vector have the same direction as itself—a property which we surely want. Note also that this definition ascribes the following properties to the zero vector: The zero vector is the only vector having the same direction as its negative and therefore the only vector having the opposite direction to itself. The zero vector is the only vector parallel to the zero vector.

## 12.4 Exercises

1. Let $A = (1, 3, 6)$, $B = (4, -3, 3)$, and $C = (2, 1, 5)$ be three vectors in $V_3$. Determine the components of each of the following vectors: (a) $A + B$; (b) $A - B$; (c) $A + B - C$; (d) $7A - 2B - 3C$; (e) $2A + B - 3C$.
2. Draw the geometric vectors from the origin to the points $A = (2, 1)$ and $B = (1, 3)$. On the same figure, draw the geometric vector from the origin to the point $C = A + tB$ for each of the following values of $t$: $t = \frac{1}{3}$; $t = \frac{1}{2}$; $t = \frac{3}{4}$; $t = 1$; $t = 2$; $t = -1$; $t = -2$.
3. Solve Exercise 2 if $C = tA + B$.
4. Let $A = (2, 1)$, $B = (1, 3)$, and $C = xA + yB$, where $x$ and $y$ are scalars.
   (a) Draw the geometric vector from the origin to $C$ for each of the following pairs of values of $x$ and $y$: $x = y = \frac{1}{2}$; $x = \frac{1}{4}, y = \frac{3}{4}$; $x = \frac{1}{3}, y = \frac{2}{3}$; $x = 2, y = -1$; $x = 3, y = -2$; $x = -\frac{1}{2}, y = \frac{3}{2}$; $x = -1, y = 2$.
   (b) What do you think is the set of points $C$ obtained as $x$ and $y$ run through all real numbers such that $x + y = 1$? (Just make a guess and show the locus on the figure. No proof is required.)
   (c) Make a guess for the set of all points $C$ obtained as $x$ and $y$ range independently over the intervals $0 \leq x \leq 1$, $0 \leq y \leq 1$, and make a sketch of this set.
   (d) What do you think is the set of all $C$ obtained if $x$ ranges through the interval $0 \leq x \leq 1$ and $y$ ranges through all real numbers?
   (e) What do you think is the set if $x$ and $y$ both range over all real numbers?
5. Let $A = (2, 1)$ and $B = (1, 3)$. Show that every vector $C = (c_1, c_2)$ in $V_2$ can be expressed in the form $C = xA + yB$. Express $x$ and $y$ in terms of $c_1$ and $c_2$.

6. Let $A = (1, 1, 1)$, $B = (0, 1, 1)$, and $C = (1, 1, 0)$ be three vectors in $V_3$ and let $D = xA + yB + zC$, where $x, y, z$ are scalars.
(a) Determine the components of $D$.
(b) If $D = O$, prove that $x = y = z = 0$.
(c) Find $x, y, z$ such that $D = (1, 2, 3)$.
7. Let $A = (1, 1, 1)$, $B = (0, 1, 1)$ and $C = (2, 1, 1)$ be three vectors in $V_3$, and let $D = xA + yB + zC$, where $x, y$, and $z$ are scalars.
(a) Determine the components of $D$.
(b) Find $x, y$, and $z$, not all zero, such that $D = O$.
(c) Prove that no choice of $x, y, z$ makes $D = (1, 2, 3)$.
8. Let $A = (1, 1, 1, 0)$, $B = (0, 1, 1, 1)$, $C = (1, 1, 0, 0)$ be three vectors in $V_4$, and let $D = xA + yB + zC$, where $x, y$, and $z$ are scalars.
(a) Determine the components of $D$.
(b) If $D = O$, prove that $x = y = z = 0$.
(c) Find $x, y$, and $z$ such that $D = (1, 5, 3, 4)$.
(d) Prove that no choice of $x, y, z$ makes $D = (1, 2, 3, 4)$.
9. In $V_n$, prove that two vectors parallel to the same vector are parallel to each other.
10. Given four nonzero vectors $A, B, C, D$ in $V_n$ such that $C = A + B$ and $A$ is parallel to $D$. Prove that $C$ is parallel to $D$ if and only if $B$ is parallel to $D$.
11. (a) Prove, for vectors in $V_n$, the properties of addition and multiplication by scalars given in Theorem 12.1.
(b) By drawing geometric vectors in the plane, illustrate the geometric meaning of the two distributive laws $(c + d)A = cA + dA$ and $c(A + B) = cA + cB$.
12. If a quadrilateral $OABC$ in $V_2$ is a parallelogram having $A$ and $C$ as opposite vertices, prove that $A + \frac{1}{2}(C - A) = \frac{1}{2}B$. What geometrical theorem about parallelograms can you deduce from this equation?

## 12.5 The dot product

We introduce now a new kind of multiplication called the dot product or scalar product of two vectors in $V_n$.

DEFINITION. *If $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ are two vectors in $V_n$, their dot product is denoted by $A \cdot B$ and is defined by the equation*

$$A \cdot B = \sum_{k=1}^{n} a_k b_k .$$

Thus, to compute $A \cdot B$ we multiply corresponding components of $A$ and $B$ and then add all the products. This multiplication has the following algebraic properties.

THEOREM 12.2. *For all vectors $A, B, C$ in $V_n$ and all scalars $c$, we have the following properties:*
(a) $A \cdot B = B \cdot A$ *(commutative law),*
(b) $A \cdot (B + C) = A \cdot B + A \cdot C$ *(distributive law),*
(c) $c(A \cdot B) = (cA) \cdot B = A \cdot (cB)$ *(homogeneity),*
(d) $A \cdot A > 0$ *if* $A \neq O$ *(positivity),*
(e) $A \cdot A = 0$ *if* $A = O$.

*Proof.* The first three properties are easy consequences of the definition and are left as exercises. To prove the last two, we use the relation $A \cdot A = \sum a_k^2$. Since each term is nonnegative, the sum is nonnegative. Moreover, the sum is zero if and only if each term in the sum is zero and this can happen only if $A = O$.

The dot product has an interesting geometric interpretation which will be described in Section 12.9. Before we discuss this, however, we mention an important inequality concerning dot products that is fundamental in vector algebra.

THEOREM 12.3. THE CAUCHY–SCHWARZ INEQUALITY. *If $A$ and $B$ are vectors in $V_n$, we have*

(12.2)
$$(A \cdot B)^2 \leq (A \cdot A)(B \cdot B).$$

*Moreover, the equality sign holds if and only if one of the vectors is a scalar multiple of the other.*

*Proof.* Expressing each member of (12.2) in terms of components, we obtain

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} b_k^2\right).$$

which is the inequality proved earlier in Theorem I.41.

We shall present another proof of (12.2) that makes no use of components. Such a proof is of interest because it shows that the Cauchy–Schwarz inequality is a consequence of the five properties of the dot product listed in Theorem 12.2 and does not depend on the particular definition that was used to deduce these properties.

To carry out this proof, we notice first that (12.2) holds trivially if either $A$ or $B$ is the zero vector. Therefore, we may assume that both $A$ and $B$ are nonzero. Let $C$ be the vector

$$C = xA - yB, \qquad \text{where} \quad x = B \cdot B \quad \text{and} \quad y = A \cdot B.$$

Properties (d) and (e) imply that $C \cdot C \geq 0$. When we translate this in terms of $x$ and $y$, it will yield (12.2). To express $C \cdot C$ in terms of $x$ and $y$, we use properties (a), (b) and (c) to obtain

$$C \cdot C = (xA - yB) \cdot (xA - yB) = x^2(A \cdot A) - 2xy(A \cdot B) + y^2(B \cdot B).$$

Using the definitions of $x$ and $y$ and the inequality $C \cdot C \geq 0$, we get

$$(B \cdot B)^2(A \cdot A) - 2(A \cdot B)^2(B \cdot B) + (A \cdot B)^2(B \cdot B) \geq 0.$$

Property (d) implies $B \cdot B > 0$ since $B \neq O$, so we may divide by $(B \cdot B)$ to obtain

$$(B \cdot B)(A \cdot A) - (A \cdot B)^2 \geq 0,$$

which is (12.2). This proof also shows that the equality sign holds in (12.2) if and only if $C = O$. But $C = O$ if and only if $xA = yB$. This equation holds, in turn, if and only if one of the vectors is a scalar multiple of the other.

The Cauchy-Schwarz inequality has important applications to the properties of the *length* or *norm* of a vector, a concept which we discuss next.

## 12.6 Length or norm of a vector

Figure 12.7 shows the geometric vector from the origin to a point $A = (a_1, a_2)$ in the plane. From the theorem of Pythagoras, we find that the length of $A$ is given by the formula

$$\text{length of } A = \sqrt{a_1^2 + a_2^2}.$$

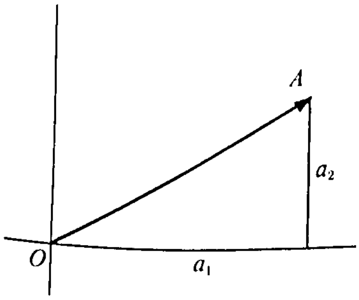


FIGURE 12.7 In $V_2$, the length of $A$ is $\sqrt{a_1^2 + a_2^2}$.

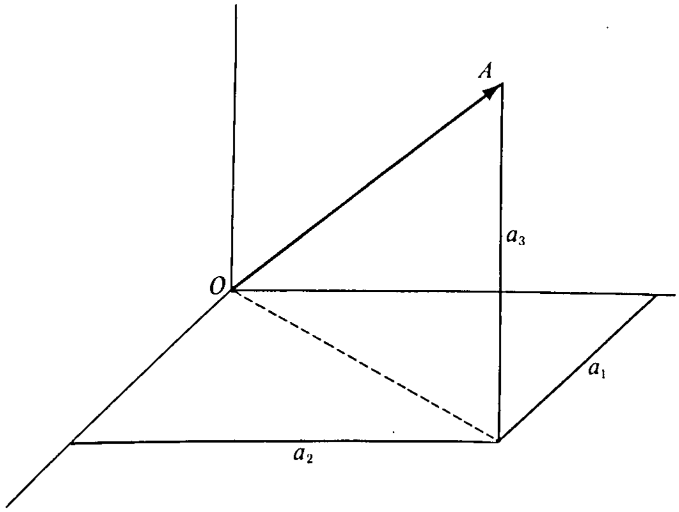


FIGURE 12.8 In $V_3$, the length of $A$ is $\sqrt{a_1^2 + a_2^2 + a_3^2}$.

A corresponding picture in 3-space is shown in Figure 12.8. Applying the theorem of Pythagoras twice, we find that the length of a geometric vector $A$ in 3-space is given by

$$\text{length of } A = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Note that in either case the length of $A$ is given by $(A \cdot A)^{1/2}$, the square root of the dot product of $A$ with itself. This formula suggests a way to introduce the concept of length in $n$-space.

DEFINITION. *If $A$ is a vector in $V_n$, its length or norm is denoted by $\|A\|$ and is defined by the equation*

$$\|A\| = (A \cdot A)^{1/2}.$$

The fundamental properties of the dot product lead to corresponding properties of norms.

THEOREM 12.4. *If $A$ is a vector in $V_n$ and if $c$ is a scalar, we have the following properties:*

(a) $\|A\| > 0$ *if* $A \neq O$ (*positivity*),
(b) $\|A\| = 0$ *if* $A = O$,
(c) $\|cA\| = |c|\,\|A\|$ (*homogeneity*).

*Proof.* Properties (a) and (b) follow at once from properties (d) and (e) of Theorem 12.2. To prove (c), we use the homogeneity property of dot products to obtain

$$\|cA\| = (cA \cdot cA)^{1/2} = (c^2 A \cdot A)^{1/2} = (c^2)^{1/2}(A \cdot A)^{1/2} = |c| \, \|A\| .$$

The Cauchy–Schwarz inequality can also be expressed in terms of norms. It states that

(12.3) $$(A \cdot B)^2 \leq \|A\|^2 \, \|B\|^2 .$$

Taking the positive square root of each member, we can also write the Cauchy–Schwarz inequality in the equivalent form

(12.4) $$|A \cdot B| \leq \|A\| \, \|B\| .$$

Now we shall use the Cauchy–Schwarz inequality to deduce the triangle inequality.

THEOREM 12.5. TRIANGLE INEQUALITY. *If $A$ and $B$ are vectors in $V_n$, we have*

$$\|A + B\| \leq \|A\| + \|B\| .$$

*Moreover, the equality sign holds if and only if $A = O$, or $B = O$, or $B = cA$ for some $c > 0$.*

*Proof.* To avoid square roots, we write the triangle inequality in the equivalent form

(12.5) $$\|A + B\|^2 \leq (\|A\| + \|B\|)^2 .$$

The left member of (12.5) is

$$\|A + B\|^2 = (A + B) \cdot (A + B) = A \cdot A + 2A \cdot B + B \cdot B = \|A\|^2 + 2A \cdot B + \|B\|^2,$$

whereas the right member is

$$(\|A\| + \|B\|)^2 = \|A\|^2 + 2\|A\| \, \|B\| + \|B\|^2 .$$

Comparing these two formulas, we see that (12.5) holds if and only if we have

(12.6) $$A \cdot B \leq \|A\| \, \|B\| .$$

But $A \cdot B \leq |A \cdot B|$ so (12.6) follows from the Cauchy–Schwarz inequality, as expressed in (12.4). This proves that the triangle inequality is a consequence of the Cauchy–Schwarz inequality.

The converse statement is also true. That is, if the triangle inequality holds then (12.6) also holds for $A$ and for $-A$, from which we obtain (12.3). If equality holds in (12.5), then $A \cdot B = \|A\| \, \|B\|$, so $B = cA$ for some scalar $c$. Hence $A \cdot B = c\|A\|^2$ and $\|A\| \, \|B\| = |c| \, \|A\|^2$. If $A \neq O$ this implies $c = |c| \geq 0$. If $B \neq O$ then $B = cA$ with $c > 0$.

The triangle inequality is illustrated geometrically in Figure 12.9. It states that the length of one side of a triangle does not exceed the sum of the lengths of the other two sides.

## 12.7 Orthogonality of vectors

In the course of the proof of the triangle inequality (Theorem 12.5), we obtained the formula

$$\|A + B\|^2 = \|A\|^2 + \|B\|^2 + 2A \cdot B \tag{12.7}$$

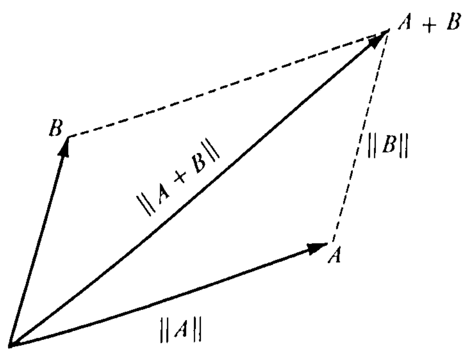


FIGURE 12.9 Geometric meaning of the triangle inequality:
$$\|A + B\| \le \|A\| + \|B\|.$$

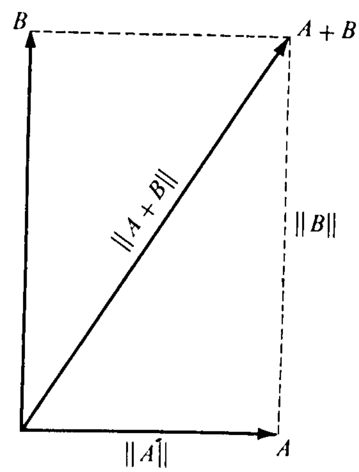


FIGURE 12.10 Two perpendicular vectors satisfy the Pythagorean identity:
$$\|A + B\|^2 = \|A\|^2 + \|B\|^2.$$

which is valid for any two vectors $A$ and $B$ in $V_n$. Figure 12.10 shows two perpendicular geometric vectors in the plane. They determine a right triangle whose legs have lengths $\|A\|$ and $\|B\|$ and whose hypotenuse has length $\|A + B\|$. The theorem of Pythagoras states that

$$\|A + B\|^2 = \|A\|^2 + \|B\|^2.$$

Comparing this with (12.7), we see that $A \cdot B = 0$. In other words, the dot product of two perpendicular vectors in the plane is zero. This property motivates the definition of perpendicularity of vectors in $V_n$.

DEFINITION. *Two vectors $A$ and $B$ in $V_n$ are called perpendicular or orthogonal if $A \cdot B = 0$.*

Equation (12.7) shows that two vectors $A$ and $B$ in $V_n$ are orthogonal if and only if $\|A + B\|^2 = \|A\|^2 + \|B\|^2$. This is called the Pythagorean identity in $V_n$.

## 12.8 Exercises

1. Let $A = (1, 2, 3, 4)$, $B = (-1, 2, -3, 0)$, and $C = (0, 1, 0, 1)$ be three vectors in $V_4$. Compute each of the following dot products:
(a) $A \cdot B$; (b) $B \cdot C$; (c) $A \cdot C$; (d) $A \cdot (B + C)$; (e) $(A - B) \cdot C$.
2. Given three vectors $A = (2, 4, -7)$, $B = (2, 6, 3)$, and $C = (3, 4, -5)$. In each of the following there is only one way to insert parentheses to obtain a meaningful expression. Insert parentheses and perform the indicated operations.
(a) $A \cdot BC$; (b) $A \cdot B + C$; (c) $A + B \cdot C$; (d) $AB \cdot C$; (e) $A/B \cdot C$.
3. Prove or disprove the following statement about vectors in $V_n$: If $A \cdot B = A \cdot C$ and $A \neq O$, then $B = C$.
4. Prove or disprove the following statement about vectors in $V_n$: If $A \cdot B = 0$ for every $B$, then $A = O$.
5. If $A = (2, 1, -1)$ and $B = (1, -1, 2)$, find a nonzero vector $C$ in $V_3$ such that $A \cdot C = B \cdot C = 0$.
6. If $A = (1, -2, 3)$ and $B = (3, 1, 2)$, find scalars $x$ and $y$ such that $C = xA + yB$ is a nonzero vector with $C \cdot B = 0$.
7. If $A = (2, -1, 2)$ and $B = (1, 2, -2)$, find two vectors $C$ and $D$ in $V_3$ satisfying all the following conditions: $A = C + D$, $B \cdot D = 0$, $C$ parallel to $B$.
8. If $A = (1, 2, 3, 4, 5)$ and $B = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5})$, find two vectors $C$ and $D$ in $V_5$ satisfying all the following conditions: $B = C + 2D$, $D \cdot A = 0$, $C$ parallel to $A$.
9. Let $A = (2, -1, 5)$, $B = (-1, -2, 3)$, and $C = (1, -1, 1)$ be three vectors in $V_3$. Calculate the norm of each of the following vectors:
(a) $A + B$; (b) $A - B$; (c) $A + B - C$; (d) $A - B + C$.
10. In each case, find a vector $B$ in $V_2$ such that $B \cdot A = 0$ and $\|B\| = \|A\|$ if:
(a) $A = (1, 1)$; (b) $A = (1, -1)$; (c) $A = (2, -3)$; (d) $A = (a, b)$.
11. Let $A = (1, -2, 3)$ and $B = (3, 1, 2)$ be two vectors in $V_3$. In each case, find a vector $C$ of length 1 parallel to:
(a) $A + B$; (b) $A - B$; (c) $A + 2B$; (d) $A - 2B$; (e) $2A - B$.
12. Let $A = (4, 1, -3)$, $B = (1, 2, 2)$, $C = (1, 2, -2)$, $D = (2, 1, 2)$, and $E = (2, -2, -1)$ be vectors in $V_3$. Determine all orthogonal pairs.
13. Find all vectors in $V_2$ that are orthogonal to $A$ and have the same length as $A$ if:
(a) $A = (1, 2)$; (b) $A = (1, -2)$; (c) $A = (2, -1)$; (d) $A = (-2, 1)$.
14. If $A = (2, -1, 1)$ and $B = (3, -4, -4)$, find a point $C$ in 3-space such that $A$, $B$, and $C$ are the vertices of a right triangle.
15. If $A = (1, -1, 2)$ and $B = (2, 1, -1)$, find a nonzero vector $C$ in $V_3$ orthogonal to $A$ and $B$.
16. Let $A = (1, 2)$ and $B = (3, 4)$ be two vectors in $V_2$. Find vectors $P$ and $Q$ in $V_2$ such that $A = P + Q$, $P$ is parallel to $B$, and $Q$ is orthogonal to $B$.
17. Solve Exercise 16 if the vectors are in $V_4$, with $A = (1, 2, 3, 4)$ and $B = (1, 1, 1, 1)$.
18. Given vectors $A = (2, -1, 1)$, $B = (1, 2, -1)$, and $C = (1, 1, -2)$ in $V_3$. Find every vector $D$ of the form $xB + yC$ which is orthogonal to $A$ and has length 1.
19. Prove that for two vectors $A$ and $B$ in $V_n$ we have the identity

$$\|A + B\|^2 - \|A - B\|^2 = 4A \cdot B,$$

and hence $A \cdot B = 0$ if and only if $\|A + B\| = \|A - B\|$. When this is interpreted geometrically in $V_2$, it states that the diagonals of a parallelogram are of equal length if and only if the parallelogram is a rectangle.
20. Prove that for any two vectors $A$ and $B$ in $V_n$ we have

$$\|A + B\|^2 + \|A - B\|^2 = 2\,\|A\|^2 + 2\,\|B\|^2.$$

What geometric theorem about the sides and diagonals of a parallelogram can you deduce from this identity?

21. The following theorem in geometry suggests a vector identity involving three vectors $A$, $B$, and $C$. Guess the identity and prove that it holds for vectors in $V_n$. This provides a proof of the theorem by vector methods.

    "The sum of the squares of the sides of any quadrilateral exceeds the sum of the squares of the diagonals by four times the square of the length of the line segment which connects the midpoints of the diagonals."
22. A vector $A$ in $V_n$ has length 6. A vector $B$ in $V_n$ has the property that for every pair of scalars $x$ and $y$ the vectors $xA + yB$ and $4yA - 9xB$ are orthogonal. Compute the length of $B$ and the length of $2A + 3B$.
23. Given two vectors $A = (1, 2, 3, 4, 5)$ and $B = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5})$ in $V_5$. Find two vectors $C$ and $D$ satisfying the following three conditions: $C$ is parallel to $A$, $D$ is orthogonal to $A$, and $B = C + D$.
24. Given two nonperpendicular vectors $A$ and $B$ in $V_n$, prove that there exist vectors $C$ and $D$ in $V_n$ satisfying the three conditions in Exercise 23 and express $C$ and $D$ in terms of $A$ and $B$.
25. Prove or disprove each of the following statements concerning vectors in $V_n$:
    (a) If $A$ is orthogonal to $B$, then $\|A + xB\| \geq \|A\|$ for all real $x$.
    (b) If $\|A + xB\| \geq \|A\|$ for all real $x$, then $A$ is orthogonal to $B$.

## 12.9 Projections. Angle between vectors in *n*-space

The dot product of two vectors in $V_2$ has an interesting geometric interpretation. Figure 12.11(a) shows two nonzero geometric vectors $A$ and $B$ making an angle $\theta$ with each other. In this example, we have $0 < \theta < \frac{1}{2}\pi$. Figure 12.11(b) shows the same vector $A$ and two perpendicular vectors whose sum is $A$. One of these, $tB$, is a scalar multiple of $B$ which we call the *projection of A along B*. In this example, $t$ is positive since $0 < \theta < \frac{1}{2}\pi$.

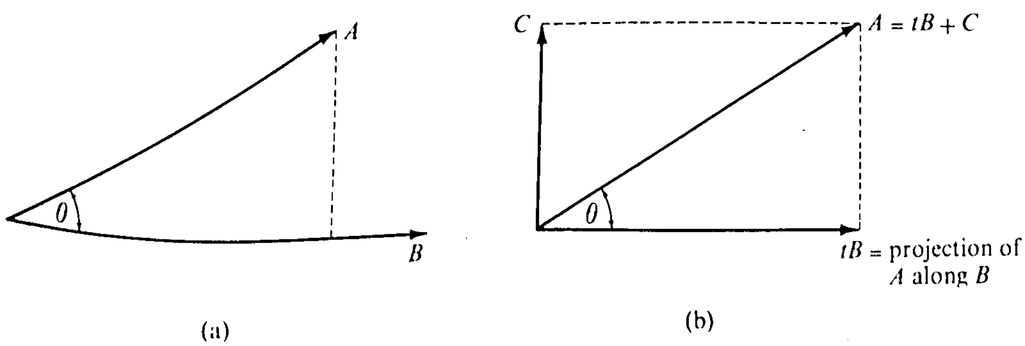


FIGURE 12.11 The vector $tB$ is the projection of $A$ along $B$.

We can use dot products to express $t$ in terms of $A$ and $B$. First we write $tB + C = A$ and then take the dot product of each member with $B$ to obtain

$$tB \cdot B + C \cdot B = A \cdot B .$$

But $C \cdot B = 0$, because $C$ was drawn perpendicular to $B$. Therefore $tB \cdot B = A \cdot B$, so we have

$$t = \frac{A \cdot B}{B \cdot B} = \frac{A \cdot B}{\|B\|^2} . \tag{12.8}$$

On the other hand, the scalar $t$ bears a simple relation to the angle $\theta$. From Figure 12.11(b), we see that

$$\cos\theta = \frac{\|tB\|}{\|A\|} = \frac{t\,\|B\|}{\|A\|}.$$

Using (12.8) in this formula, we find that

(12.9)
$$\cos\theta = \frac{A \cdot B}{\|A\|\,\|B\|}$$

or

$$A \cdot B = \|A\|\,\|B\| \cos\theta .$$

In other words, the dot product of two nonzero vectors $A$ and $B$ in $V_2$ is equal to the product of three numbers: the length of $A$, the length of $B$, and the cosine of the angle between $A$ and $B$.

Equation (12.9) suggests a way to define the concept of angle in $V_n$. The Cauchy-Schwarz inequality, as expressed in (12.4), shows that the quotient on the right of (12.9) has absolute value $\leq 1$ for any two nonzero vectors in $V_n$. In other words, we have

$$-1 \leq \frac{A \cdot B}{\|A\|\,\|B\|} \leq 1 .$$

Therefore, there is exactly one real $\theta$ in the interval $0 \leq \theta \leq \pi$ such that (12.9) holds. We define the angle between $A$ and $B$ to be this $\theta$. The foregoing discussion is summarized in the following definition.

DEFINITION. *Let $A$ and $B$ be two vectors in $V_n$, with $B \neq O$. The vector $tB$, where*

$$t = \frac{A \cdot B}{B \cdot B},$$

*is called the projection of $A$ along $B$. If both $A$ and $B$ are nonzero, the angle $\theta$ between $A$ and $B$ is defined by the equation*

$$\theta = \arccos \frac{A \cdot B}{\|A\|\,\|B\|}.$$

*Note:* The arc cosine function restricts $\theta$ to the interval $0 \leq \theta \leq \pi$. Note also that $\theta = \frac{1}{2}\pi$ when $A \cdot B = 0$.

## 12.10 The unit coordinate vectors

In Chapter 9 we learned that every complex number $(a, b)$ can be expressed in the form $a + bi$, where $i$ denotes the complex number $(0, 1)$. Similarly, every vector $(a, b)$ in $V_2$ can be expressed in the form

$$(a, b) = a(1, 0) + b(0, 1).$$

The two vectors (1, 0) and (0, 1) which multiply the components $a$ and $b$ are called *unit coordinate vectors.* We now introduce the corresponding concept in $V_n$

DEFINITION. *In* $V_n$, *the* $n$ *vectors* $E_1 = (1, 0, \ldots, 0)$, $E_2 = (0, 1, 0, \ldots, 0), \ldots, E_n = (0, 0, \ldots, 0, 1)$ *are called the unit coordinate vectors. It is understood that the kth component of* $E_k$ *is* 1 *and all other components are* 0.

The name "unit vector" comes from the fact that each vector $E_k$ has length 1. Note that these vectors are mutually orthogonal, that is, the dot product of any two distinct vectors is zero,

$$E_k \cdot E_j = 0 \qquad \text{if} \quad k \neq j.$$

THEOREM 12.6. *Every vector* $X = (x_1, \ldots, x_n)$ *in* $V_n$ *can be expressed in the form*

$$X = x_1E_1 + \cdots + x_nE_n = \sum_{k=1}^{n} x_kE_k .$$

*Moreover, this representation is unique. That is, if*

$$X = \sum_{k=1}^{n} x_kE_k \qquad \text{and} \qquad X = \sum_{k=1}^{n} y_kE_k ,$$

*then* $x_k = y_k$ *for each* $k = 1, 2, \ldots, n$.

*Proof.* The first statement follows immediately from the definition of addition and multiplication by scalars. The uniqueness property follows from the definition of vector equality.

A sum of the type $\sum c_iA_i$ is called a *linear combination* of the vectors $A_1, \ldots, A_n$. Theorem 12.6 tells us that every vector in $V_n$ can be expressed as a linear combination of the unit coordinate vectors. We describe this by saying that the unit coordinate vectors $E_1, \ldots, E_n$ *span* the space $V_n$. We also say they span $V_n$ *uniquely* because each representation of a vector as a linear combination of $E_1, \ldots, E_n$ is unique. Some collections of vectors other than $E_1, \ldots, E_n$ also span $V_n$ uniquely, and in Section 12.12 we turn to the study of such collections.

In $V_2$ the unit coordinate vectors $E_1$ and $E_2$ are often denoted, respectively, by the symbols $\boldsymbol{i}$ and $\boldsymbol{j}$ in bold-face italic type. In $V_3$ the symbols $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$ are also used in place of $E_1, E_2, E_3$. Sometimes a bar or arrow is placed over the symbol, for example, $\bar{i}$ or $\vec{i}$. The geometric meaning of Theorem 12.6 is illustrated in Figure 12.12 for $n = 3$.

When vectors are expressed as linear combinations of the unit coordinate vectors, algebraic manipulations involving vectors can be performed by treating the sums $\sum x_kE_k$ according to the usual rules of algebra. The various components can be recognized at any stage in the calculation by collecting the coefficients of the unit coordinate vectors. For example, to add two vectors, say $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$, we write

$$A = \sum_{k=1}^{n} a_kE_k , \qquad B = \sum_{k=1}^{n} b_kE_k ,$$

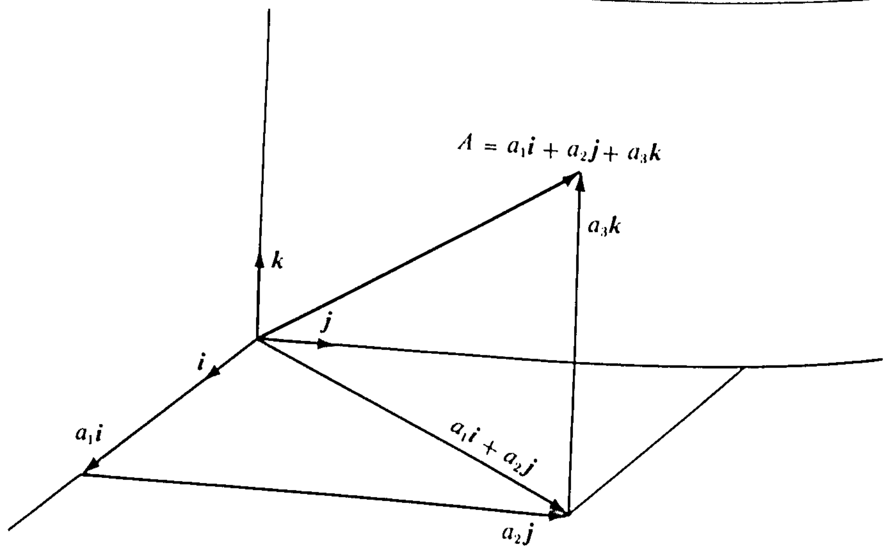


FIGURE 12.12 A vector $A$ in $V_3$ expressed as a linear combination of $i, j, k$.

and apply the linearity property of finite sums to obtain

$$A + B = \sum_{k=1}^{n} a_k E_k + \sum_{k=1}^{n} b_k E_k = \sum_{k=1}^{n} (a_k + b_k) E_k .$$

The coefficient of $E_k$ on the right is the $k$th component of the sum $A + B$.

## 12.11 Exercises

1. Determine the projection of $A$ along $B$ if $A = (1, 2, 3)$ and $B = (1, 2, 2)$.
2. Determine the projection of $A$ along $B$ if $A = (4, 3, 2, 1)$ and $B = (1, 1, 1, 1)$.
3. (a) Let $A = (6, 3, -2)$, and let $a, b, c$ denote the angles between $A$ and the unit coordinate vectors $i, j, k$, respectively. Compute $\cos a$, $\cos b$, and $\cos c$. These are called the direction cosines of $A$.
   (b) Find all vectors in $V_3$ of length 1 parallel to $A$.
4. Prove that the angle between the two vectors $A = (1, 2, 1)$ and $B = (2, 1, -1)$ is twice that between $C = (1, 4, 1)$ and $D = (2, 5, 5)$.
5. Use vector methods to determine the cosines of the angles of the triangle in 3-space whose vertices are at the points $(2, -1, 1)$, $(1, -3, -5)$, and $(3, -4, -4)$.
6. Three vectors $A, B, C$ in $V_3$ satisfy all the following properties:

$$\|A\| = \|C\| = 5, \qquad \|B\| = 1, \qquad \|A - B + C\| = \|A + B + C\| .$$

   If the angle between $A$ and $B$ is $\pi/8$, find the angle between $B$ and $C$.
7. Given three nonzero vectors $A, B, C$ in $V_n$. Assume that the angle between $A$ and $C$ is equal to the angle between $B$ and $C$. Prove that $C$ is orthogonal to the vector $\|B\| A - \|A\| B$.
8. Let $\theta$ denote the angle between the following two vectors in $V_n$: $A = (1, 1, \ldots, 1)$ and $B = (1, 2, \ldots, n)$. Find the limiting value of $\theta$ as $n \to \infty$.
9. Solve Exercise 8 if $A = (2, 4, 6, \ldots, 2n)$ and $B = (1, 3, 5, \ldots, 2n - 1)$.

10. Given vectors $A = (\cos \theta, -\sin \theta)$ and $B = (\sin \theta, \cos \theta)$ in $V_2$.
(a) Prove that $A$ and $B$ are orthogonal vectors of length 1. Make a sketch showing $A$ and $B$ when $\theta = \pi/6$.
(b) Find all vectors $(x, y)$ in $V_2$ such that $(x, y) = xA + yB$. Be sure to consider all possible values of $\theta$.
11. Use vector methods to prove that the diagonals of a rhombus are perpendicular.
12. By forming the dot product of the two vectors $(\cos a, \sin a)$ and $(\cos b, \sin b)$, deduce the trigonometric identity $\cos(a - b) = \cos a \cos b + \sin a \sin b$.
13. If $\theta$ is the angle between two nonzero vectors $A$ and $B$ in $V_n$, prove that

$$\|A - B\|^2 = \|A\|^2 + \|B\|^2 - 2\|A\|\,\|B\| \cos \theta .$$

When interpreted geometrically in $V_2$, this is the law of cosines of trigonometry.
14. Suppose that instead of defining the dot product of two vectors $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ by the formula $A \cdot B = \sum_{k=1}^{n} a_k b_k$, we used the following definition:

$$A \cdot B = \sum_{k=1}^{n} |a_k b_k| .$$

Which of the properties of Theorem 12.2 are valid with this definition? Is the Cauchy–Schwarz inequality valid with this definition?
15. Suppose that in $V_2$ we define the dot product of two vectors $A = (a_1, a_2)$ and $B = (b_1, b_2)$ by the formula

$$A \cdot B = 2a_1b_1 + a_2b_2 + a_1b_2 + a_2b_1 .$$

Prove that all the properties of Theorem 12.2 are valid with this definition of dot product. Is the Cauchy–Schwarz inequality still valid?
16. Solve Exercise 15 if the dot product of two vectors $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ in $V_3$ is defined by the formula $A \cdot B = 2a_1b_1 + a_2b_2 + a_3b_3 + a_1b_3 + a_3b_1$.
17. Suppose that instead of defining the norm of a vector $A = (a_1, \ldots, a_n)$ by the formula $(A \cdot A)^{1/2}$, we used the following definition:

$$\|A\| = \sum_{k=1}^{n} |a_k| .$$

(a) Prove that this definition of norm satisfies all the properties in Theorems 12.4 and 12.5.
(b) Use this definition in $V_2$ and describe on a figure the set of all points $(x, y)$ of norm 1.
(c) Which of the properties of Theorems 12.4 and 12.5 would hold if we used the definition

$$\|A\| = \left|\sum_{k=1}^{n} a_k\right| ?$$

18. Suppose that the norm of a vector $A = (a_1, \ldots, a_n)$ were defined by the formula

$$\|A\| = \max_{1 \le k \le n} |a_k| ,$$

where the symbol on the right means the maximum of the $n$ numbers $|a_1|, |a_2|, \ldots, |a_n|$.
(a) Which of the properties of Theorems 12.4 and 12.5 are valid with this definition?
(b) Use this definition of norm in $V_2$ and describe on a figure the set of all points $(x, y)$ of norm 1.

19. If $A = (a_1, \ldots, a_n)$ is a vector in $V_n$, define two norms as follows:

$$\|A\|_1 = \sum_{k=1}^{n} |a_k| \quad \text{and} \quad \|A\|_2 = \max_{1 \le k \le n} |a_k| .$$

Prove that $\|A\|_2 \le \|A\| \le \|A\|_1$. Interpret this inequality geometrically in the plane.

20. If $A$ and $B$ are two points in $n$-space, the distance from $A$ to $B$ is denoted by $d(A, B)$ and is defined by the equation $d(A, B) = \|A - B\|$. Prove that distance has the following properties:

(a) $d(A, B) = d(B, A)$. (b) $d(A, B) = 0$ if and only if $A = B$.
(c) $d(A, B) \le d(A, C) + d(C, B)$.

## 12.12 The linear span of a finite set of vectors

Let $S = \{A_1, \ldots, A_k\}$ be a nonempty set consisting of $k$ vectors in $V_n$, where $k$, the number of vectors, may be less than, equal to, or greater than $n$, the dimension of the space. If a vector $X$ in $V_n$ can be represented as a linear combination of $A_1, \ldots, A_k$, say

$$X = \sum_{i=1}^{k} c_i A_i ,$$

then the set $S$ is said to *span* the vector $X$.

DEFINITION. *The set of all vectors spanned by $S$ is called the linear span of $S$ and is denoted by $L(S)$.*

In other words, the linear span of $S$ is simply the set of all possible linear combinations of vectors in $S$. Note that linear combinations of vectors in $L(S)$ are again in $L(S)$. We say that $S$ *spans the whole space* $V_n$ if $L(S) = V_n$.

EXAMPLE 1. Let $S = \{A_1\}$. Then $L(S)$ consists of all scalar multiples of $A_1$.

EXAMPLE 2. Every set $S = \{A_1, \ldots, A_k\}$ spans the zero vector since $O = 0A_1 + \cdots + 0A_k$. This representation, in which all the coefficients $c_1, \ldots, c_k$ are zero, is called the *trivial representation* of the zero vector. However, there may be nontrivial linear combinations that represent $O$. For example, suppose one of the vectors in $S$ is a scalar multiple of another, say $A_2 = 2A_1$. Then we have many nontrivial representations of $O$, for example

$$O = 2tA_1 - tA_2 + 0A_3 + \cdots + 0A_k ,$$

where $t$ is any nonzero scalar.

We are especially interested in sets $S$ that span vectors in exactly one way.

DEFINITION. *A set $S = \{A_1, \ldots, A_k\}$ of vectors in $V_n$ is said to span $X$ uniquely if $S$ spans $X$ and if*

$$(12.10) \qquad X = \sum_{i=1}^{k} c_i A_i \quad \textit{and} \quad X = \sum_{i=1}^{k} d_i A_i \quad \textit{implies} \quad c_i = d_i \quad \textit{for all } i .$$

In the two sums appearing in (12.10), it is understood that the vectors $A_1, \ldots, A_k$ are written in the same order. It is also understood that the implication (12.10) is to hold for a fixed but arbitrary ordering of the vectors $A_1, \ldots, A_k$.

THEOREM 12.7. *A set S spans every vector in L(S) uniquely if and only if S spans the zero vector uniquely.*

*Proof.* If $S$ spans every vector in $L(S)$ uniquely, then it certainly spans $O$ uniquely. To prove the converse, assume $S$ spans $O$ uniquely and choose any vector $X$ in $L(S)$. Suppose $S$ spans $X$ in two ways, say

$$X = \sum_{i=1}^{k} c_i A_i \qquad \text{and} \qquad X = \sum_{i=1}^{k} d_i A_i .$$

By subtraction, we find that $O = \sum_{i=1}^{k} (c_i - d_i)A_i$. But since $S$ spans $O$ uniquely, we must have $c_i - d_i = 0$ for all $i$, so $S$ spans $X$ uniquely.

## 12.13 Linear independence

Theorem 12.7 demonstrates the importance of sets that span the zero vector uniquely. Such sets are distinguished with a special name.

DEFINITION. *A set $S = \{A_1, \ldots, A_k\}$ which spans the zero vector uniquely is said to be a linearly independent set of vectors. Otherwise, S is called linearly dependent.*

In other words, *independence* means that $S$ spans $O$ with only the trivial representation:

$$\sum_{i=1}^{k} c_i A_i = O \qquad \text{implies all } c_i = 0 .$$

*Dependence* means that $S$ spans $O$ in some nontrivial way. That is, for some choice of scalars $c_1, \ldots, c_k$, we have

$$\sum_{i=1}^{k} c_i A_i = O \qquad \text{but not all } c_i \text{ are zero} .$$

Although dependence and independence are properties of *sets* of vectors, it is common practice to also apply these terms to the vectors themselves. For example, the vectors in a linearly independent set are often called linearly independent vectors. We also agree to call the empty set linearly independent.

The following examples may serve to give further insight into the meaning of dependence and independence.

EXAMPLE 1. If a subset $T$ of a set $S$ is dependent, then $S$ itself is dependent, because if $T$ spans $O$ nontrivially, then so does $S$. This is logically equivalent to the statement that every subset of an independent set is independent.

EXAMPLE 2. The $n$ unit coordinate vectors $E_1, \ldots, E_n$ in $V_n$ span $O$ uniquely so they are linearly independent.

EXAMPLE 3. Any set containing the zero vector is dependent. For example, if $A_1 = O$, we have the nontrivial representation $O = 1A_1 + 0A_2 + \cdots + 0A_k$.

EXAMPLE 4. The set $S = \{i, j, i + j\}$ of vectors in $V_2$ is linearly dependent because we have the nontrivial representation of the zero vector

$$O = i + j + (-1)(i + j).$$

In this example the subset $T = \{i, j\}$ is linearly independent. The third vector, $i + j$, is in the linear span of $T$. The next theorem shows that if we adjoin to $i$ and $j$ any vector in the linear span of $T$, we get a dependent set.

THEOREM 12.8. *Let $S = \{A_1, \ldots, A_k\}$ be a linearly independent set of $k$ vectors in $V_n$, and let $L(S)$ be the linear span of $S$. Then, every set of $k + 1$ vectors in $L(S)$ is linearly dependent.*

*Proof.* The proof is by induction on $k$, the number of vectors in $S$. First suppose $k = 1$. Then, by hypothesis, $S$ consists of one vector, say $A_1$, where $A_1 \neq O$ since $S$ is independent. Now take any two distinct vectors $B_1$ and $B_2$ in $L(S)$. Then each is a scalar multiple of $A_1$, say $B_1 = c_1A_1$ and $B_2 = c_2A_1$, where $c_1$ and $c_2$ are not both zero. Multiplying $B_1$ by $c_2$ and $B_2$ by $c_1$ and subtracting, we find that

$$c_2B_1 - c_1B_2 = O.$$

This is a nontrivial representation of $O$ so $B_1$ and $B_2$ are dependent. This proves the theorem when $k = 1$.

Now we assume that the theorem is true for $k - 1$ and prove that it is also true for $k$. Take any set of $k + 1$ vectors in $L(S)$, say $T = \{B_1, B_2, \ldots, B_{k+1}\}$. We wish to prove that $T$ is linearly dependent. Since each $B_i$ is in $L(S)$, we may write

$$B_i = \sum_{j=1}^{k} a_{ij}A_j \qquad (12.11)$$

for each $i = 1, 2, \ldots, k + 1$. We examine all the scalars $a_{i1}$ that multiply $A_1$ and split the proof into two cases according to whether all these scalars are 0 or not.

*CASE 1. $a_{i1} = 0$ for every $i = 1, 2, \ldots, k + 1$.* In this case the sum in (12.11) does not involve $A_1$ so each $B_i$ in $T$ is in the linear span of the set $S' = \{A_2, \ldots, A_k\}$. But $S'$ is linearly independent and consists of $k - 1$ vectors. By the induction hypothesis, the theorem is true for $k - 1$ so the set $T$ is dependent. This proves the theorem in Case 1.

*CASE 2. Not all the scalars $a_{i1}$ are zero.* Let us assume that $a_{11} \neq 0$. (If necessary, we can renumber the $B$'s to achieve this.) Taking $i = 1$ in Equation (12.11) and multiplying both members by $c_i$, where $c_i = a_{i1}/a_{11}$, we get

$$c_i B_1 = a_{i1}A_1 + \sum_{j=2}^{k} c_i a_{1j} A_j .$$

From this we subtract Equation (12.11) to get

$$c_i B_1 - B_i = \sum_{j=2}^{k} (c_i a_{1j} - a_{ij})A_j ,$$

for $i = 2, \ldots, k+1$. This equation expresses each of the $k$ vectors $c_i B_1 - B_i$ as a linear combination of $k-1$ linearly independent vectors $A_2, \ldots, A_k$. By the induction hypothesis, the $k$ vectors $c_i B_1 - B_i$ must be dependent. Hence, for some choice of scalars $t_2, \ldots, t_{k+1}$, not all zero, we have

$$\sum_{i=2}^{k+1} t_i(c_i B_1 - B_i) = O ,$$

from which we find

$$\left(\sum_{i=2}^{k+1} t_i c_i\right) B_1 - \sum_{i=2}^{k+1} t_i B_i = O .$$

But this is a nontrivial linear combination of $B_1, \ldots, B_{k+1}$ which represents the zero vector, so the vectors $B_1, \ldots, B_{k+1}$ must be dependent. This completes the proof.

We show next that the concept of orthogonality is intimately related to linear independence.

DEFINITION. *A set $S = \{A_1, \ldots, A_k\}$ of vectors in $V_n$ is called an orthogonal set if $A_i \cdot A_j = 0$ whenever $i \neq j$. In other words, any two distinct vectors in an orthogonal set are perpendicular.*

THEOREM 12.9. *Any orthogonal set $S = \{A_1, \ldots, A_k\}$ of nonzero vectors in $V_n$ is linearly independent. Moreover, if $S$ spans a vector $X$, say*

$$X = \sum_{i=1}^{k} c_i A_i , \tag{12.12}$$

*then the scalar multipliers $c_1, \ldots, c_k$ are given by the formulas*

$$c_j = \frac{X \cdot A_j}{A_j \cdot A_j} \quad \text{for} \quad j = 1, 2, \ldots, k . \tag{12.13}$$

*Proof.* First we prove that $S$ is linearly independent. Assume that $\sum_{i=1}^{k} c_i A_i = O$. Taking the dot product of each member with $A_1$ and using the fact that $A_1 \cdot A_i = 0$ for each $i \neq 1$, we find $c_1(A_1 \cdot A_1) = 0$. But $(A_1 \cdot A_1) \neq 0$ since $A_1 \neq O$, so $c_1 = 0$. Repeating this argument with $A_1$ replaced by $A_j$, we find that each $c_j = 0$. Therefore $S$ spans $O$ uniquely so $S$ is linearly independent.

Now suppose that $S$ spans $X$ as in Equation (12.12). Taking the dot product of $X$ with $A_j$ as above, we find that $c_j(A_j \cdot A_j) = X \cdot A_j$ from which we obtain (12.13).

If all the vectors $A_1, \ldots, A_k$ in Theorem 12.9 have norm 1, the formula for the multipliers simplifies to

$$c_j = X \cdot A_j .$$

An orthogonal set of vectors $\{A_1, \ldots, A_k\}$, each of which has norm 1, is called an *orthonormal* set. The unit coordinate vectors $E_1, \ldots, E_n$ are an example of an orthonormal set.

## 12.14 Bases

It is natural to study sets of vectors that span every vector in $V_n$ uniquely. Such sets are called *bases* for $V_n$.

DEFINITION. *A set $S = \{A_1, \ldots, A_k\}$ of vectors in $V_n$ is called a basis for $V_n$ if $S$ spans every vector in $V_n$ uniquely. If, in addition, $S$ is orthogonal, then $S$ is called an orthogonal basis.*

Thus, a basis is a linearly independent set which spans the whole space $V_n$. The set of unit coordinate vectors is an example of a basis. This particular basis is also an orthogonal basis. Now we prove that every basis contains the same number of elements.

THEOREM 12.10. *In a given vector space $V_n$, bases have the following properties:*
(a) *Every basis contains exactly $n$ vectors.*
(b) *Any set of linearly independent vectors is a subset of some basis.*
(c) *Any set of $n$ linearly independent vectors is a basis.*

*Proof.* The unit coordinate vectors $E_1, \ldots, E_n$ form one basis for $V_n$. If we prove that any two bases contain the same number of vectors we obtain (a).

Let $S$ and $T$ be two bases, where $S$ has $k$ vectors and $T$ has $r$ vectors. If $r > k$, then $T$ contains at least $k + 1$ vectors in $L(S)$, since $L(S) = V_n$. Therefore, because of Theorem 12.8, $T$ must be linearly dependent, contradicting the assumption that $T$ is a basis. This means we cannot have $r > k$, so we must have $r \leq k$. Applying the same argument with $S$ and $T$ interchanged, we find that $k \leq r$. Hence, $k = r$ so part (a) is proved.

To prove (b), let $S = \{A_1, \ldots, A_k\}$ be any linearly independent set of vectors in $V_n$. If $L(S) = V_n$, then $S$ is a basis. If not, then there is some vector $X$ in $V_n$ which is not in $L(S)$. Adjoin this vector to $S$ and consider the new set $S' = \{A_1, \ldots, A_k, X\}$. If this set were dependent, there would be scalars $c_1, \ldots, c_{k+1}$, not all zero, such that

$$\sum_{i=1}^{k} c_i A_i + c_{k+1} X = O .$$

But $c_{k+1} \neq 0$ since $A_1, \ldots, A_k$ are independent. Hence, we could solve this equation for

$X$ and find that $X \in L(S)$, contradicting the fact that $X$ is not in $L(S)$. Therefore, the set $S'$ is linearly independent but contains $k + 1$ vectors. If $L(S') = V_n$, then $S'$ is a basis and, since $S$ is a subset of $S'$, part (b) is proved. If $S'$ is not a basis, we may argue with $S'$ as we did with $S$, getting a new set $S''$ which contains $k + 2$ vectors and is linearly independent. If $S''$ is a basis, then part (b) is proved. If not, we repeat the process. We must arrive at a basis in a finite number of steps, otherwise we would eventually obtain an independent set with $n + 1$ vectors, contradicting Theorem 12.8. Therefore part (b) is proved.

Finally, we use (a) and (b) to prove (c). Let $S$ be any linearly independent set consisting of $n$ vectors. By part (b), $S$ is a subset of some basis, say $B$. But by (a) the basis $B$ has exactly $n$ elements, so $S = B$.

## 12.15 Exercises

1. Let $\boldsymbol{i}$ and $\boldsymbol{j}$ denote the unit coordinate vectors in $V_2$. In each case find scalars $x$ and $y$ such that $x(\boldsymbol{i} - \boldsymbol{j}) + y(\boldsymbol{i} + \boldsymbol{j})$ is equal to
   (a) $\boldsymbol{i}$; (b) $\boldsymbol{j}$; (c) $3\boldsymbol{i} - 5\boldsymbol{j}$; (d) $7\boldsymbol{i} + 5\boldsymbol{j}$.
2. If $A = (1, 2)$, $B = (2, -4)$, and $C = (2, -3)$ are three vectors in $V_2$, find scalars $x$ and $y$ such that $C = xA + yB$. How many such pairs $x, y$ are there?
3. If $A = (2, -1, 1)$, $B = (1, 2, -1)$, and $C = (2, -11, 7)$ are three vectors in $V_3$, find scalars $x$ and $y$ such that $C = xA + yB$.
4. Prove that Exercise 3 has no solution if $C$ is replaced by the vector $(2, 11, 7)$.
5. Let $A$ and $B$ be two nonzero vectors in $V_n$.
   (a) If $A$ and $B$ are parallel, prove that $A$ and $B$ are linearly dependent.
   (b) If $A$ and $B$ are not parallel, prove that $A$ and $B$ are linearly independent.
6. If $(a, b)$ and $(c, d)$ are two vectors in $V_2$, prove that they are linearly independent if and only if $ad - bc \neq 0$.
7. Find all real $t$ for which the two vectors $(1 + t, 1 - t)$ and $(1 - t, 1 + t)$ in $V_2$ are linearly independent.
8. Let $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$ be the unit coordinate vectors in $V_3$. Prove that the four vectors $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}, \boldsymbol{i} + \boldsymbol{j} + \boldsymbol{k}$ are linearly dependent, but that any three of them are linearly independent.
9. Let $\boldsymbol{i}$ and $\boldsymbol{j}$ be the unit coordinate vectors in $V_2$ and let $S = \{\boldsymbol{i}, \boldsymbol{i} + \boldsymbol{j}\}$.
   (a) Prove that $S$ is linearly independent.
   (b) Prove that $\boldsymbol{j}$ is in the linear span of $S$.
   (c) Express $3\boldsymbol{i} - 4\boldsymbol{j}$ as a linear combination of $\boldsymbol{i}$ and $\boldsymbol{i} + \boldsymbol{j}$.
   (d) Prove that $L(S) = V_2$.
10. Consider the three vectors $A = \boldsymbol{i}$, $B = \boldsymbol{i} + \boldsymbol{j}$, and $C = \boldsymbol{i} + \boldsymbol{j} + 3\boldsymbol{k}$ in $V_3$.
    (a) Prove that the set $\{A, B, C\}$ is linearly independent.
    (b) Express each of $\boldsymbol{j}$ and $\boldsymbol{k}$ as a linear combination of $A$, $B$, and $C$.
    (c) Express $2\boldsymbol{i} - 3\boldsymbol{j} + 5\boldsymbol{k}$ as a linear combination of $A$, $B$, and $C$.
    (d) Prove that $\{A, B, C\}$ is a basis for $V_3$.
11. Let $A = (1, 2)$, $B = (2, -4)$, $C = (2, -3)$, and $D = (1, -2)$ be four vectors in $V_2$. Display all nonempty subsets of $\{A, B, C, D\}$ which are linearly independent.
12. Let $A = (1, 1, 1, 0)$, $B = (0, 1, 1, 1)$ and $C = (1, 1, 0, 0)$ be three vectors in $V_4$.
    (a) Determine whether $A, B, C$ are linearly dependent or independent.
    (b) Exhibit a nonzero vector $D$ such that $A, B, C, D$ are dependent.
    (c) Exhibit a vector $E$ such that $A, B, C, E$ are independent.
    (d) Having chosen $E$ in part (c), express the vector $X = (1, 2, 3, 4)$ as a linear combination of $A, B, C, E$.
13. (a) Prove that the following three vectors in $V_3$ are linearly independent: $(\sqrt{3}, 1, 0)$, $(1, \sqrt{3}, 1)$, $(0, 1, \sqrt{3})$.

(b) Prove that the following three are dependent: $(\sqrt{2}, 1, 0)$, $(1, \sqrt{2}, 1)$, $(0, 1, \sqrt{2})$.
(c) Find all real $t$ for which the following three vectors in $V_3$ are dependent: $(t, 1, 0)$, $(1, t, 1)$, $(0, 1, t)$.

14. Consider the following sets of vectors in $V_4$. In each case, find a linearly independent subset containing as many vectors as possible.
(a) $\{(1, 0, 1, 0), \quad (1, 1, 1, 1), \quad (0, 1, 0, 1), \quad (2, 0, -1, 0)\}$.
(b) $\{(1, 1, 1, 1), \quad (1, -1, 1, 1), \quad (1, -1, -1, 1), \quad (1, -1, -1, -1)\}$.
(c) $\{(1, 1, 1, 1), \quad (0, 1, 1, 1), \quad (0, 0, 1, 1), \quad (0, 0, 0, 1)\}$.
15. Given three linearly independent vectors $A, B, C$ in $V_n$. Prove or disprove each of the following statements.
(a) $A + B, B + C, A + C$ are linearly independent.
(b) $A - B, B + C, A + C$ are linearly independent.
16. (a) Prove that a set $S$ of three vectors in $V_3$ is a basis for $V_3$ if and only if its linear span $L(S)$ contains the three unit coordinate vectors $i, j$, and $k$.
(b) State and prove a generalization of part (a) for $V_n$.
17. Find two bases for $V_3$ containing the two vectors $(0, 1, 1)$ and $(1, 1, 1)$.
18. Find two bases for $V_4$ having only the two vectors $(0, 1, 1, 1)$ and $(1, 1, 1, 1)$ in common.
19. Consider the following sets of vectors in $V_3$:
$S = \{(1, 1, 1), (0, 1, 2), (1, 0, -1)\}$, $\quad T = \{(2, 1, 0), (2, 0, -2)\}$, $\quad U = \{(1, 2, 3), (1, 3, 5)\}$.
(a) Prove that $L(T) \subseteq L(S)$.
(b) Determine all inclusion relations that hold among the sets $L(S)$, $L(T)$, and $L(U)$.
20. Let $A$ and $B$ denote two finite subsets of vectors in a vector space $V_n$, and let $L(A)$ and $L(B)$ denote their linear spans. Prove each of the following statements.
(a) If $A \subseteq B$, then $L(A) \subseteq L(B)$.
(b) $L(A \cap B) \subseteq L(A) \cap L(B)$.
(c) Give an example in which $L(A \cap B) \neq L(A) \cap L(B)$.

## 12.16 The vector space $V_n(\mathbf{C})$ of $n$-tuples of complex numbers

In Section 12.2 the vector space $V_n$ was defined to be the collection of all $n$-tuples of real numbers. Equality, vector addition, and multiplication by scalars were defined in terms of the components as follows: If $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$, then

$$A = B \quad \text{means} \quad a_i = b_i \quad \text{for each } i = 1, 2, \ldots, n,$$

$$A + B = (a_1 + b_1, \ldots, a_n + b_n), \qquad cA = (ca_1, \ldots, ca_n).$$

If all the scalars $a_i$, $b_i$ and $c$ in these relations are replaced by *complex* numbers, the new algebraic system so obtained is called *complex vector space* and is denoted by $V_n(\mathbf{C})$. Here $\mathbf{C}$ is used to remind us that the scalars are complex.

Since complex numbers satisfy the same field properties as real numbers, all theorems about real vector space $V_n$ that use only the field properties of the real numbers are also valid for $V_n(\mathbf{C})$, provided all the scalars are allowed to be complex. In particular, those theorems in this chapter that involve only vector addition and multiplication by scalars are also valid for $V_n(\mathbf{C})$.

This extension is not made simply for the sake of generalization. Complex vector spaces arise naturally in the theory of linear differential equations and in modern quantum mechanics, so their study is of considerable importance. Fortunately, many of the theorems about real vector space $V_n$ carry over without change to $V_n(\mathbf{C})$. Some small changes have to be made, however, in those theorems that involve dot products. In proving that the dot product $A \cdot A$ of a nonzero vector with itself is positive, we used the fact that a sum of squares of real numbers is positive. Since a sum of squares of complex numbers can be negative, we must modify the definition of $A \cdot B$ if we wish to retain the positivity property. For $V_n(C)$, we use the following definition of dot product.

DEFINITION. *If $A = (a_1, \ldots, a_n)$ and $B = (b_1, \ldots, b_n)$ are two vectors in $V_n(C)$, we define their dot product $A \cdot B$ by the formula*

$$A \cdot B = \sum_{k=1}^{n} a_k \bar{b}_k ,$$

*where $\bar{b}_k$ is the complex conjugate of $b_k$.*

Note that this definition agrees with the one given earlier for $V_n$ because $\bar{b}_k = b_k$ when $b_k$ is real. The fundamental properties of the dot product, corresponding to those in Theorem 12.2, now take the following form.

THEOREM 12.11. *For all vectors $A$, $B$, $C$ in $V_n(C)$ and all complex scalars $c$, we have*
(a) $A \cdot B = \overline{B \cdot A}$,
(b) $A \cdot (B + C) = A \cdot B + A \cdot C$,
(c) $c(A \cdot B) = (cA) \cdot B = A \cdot (\bar{c}B)$,
(d) $A \cdot A > 0 \quad$ *if* $\quad A \neq O$,
(e) $A \cdot A = 0 \quad$ *if* $\quad A = O$.

All these properties are easy consequences of the definition and their proofs are left as exercises. The reader should note that conjugation takes place in property (a) when the order of the factors is reversed. Also, conjugation of the scalar multiplier occurs in property (c) when the scalar $c$ is moved from one side of the dot to the other.

The Cauchy–Schwarz inequality now takes the form

$$|A \cdot B|^2 \leq (A \cdot A)(B \cdot B) . \tag{12.14}$$

The proof is similar to that given for Theorem 12.3. We consider the vector $C = xA - yB$, where $x = B \cdot B$ and $y = A \cdot B$, and compute $C \cdot C$. The inequality $C \cdot C \geq 0$ leads to (12.14). Details are left as an exercise for the reader.

Since the dot product of a vector with itself is nonnegative, we can introduce the norm of a vector in $V_n(C)$ by the usual formula,

$$\|A\| = (A \cdot A)^{1/2} .$$

The fundamental properties of norms, as stated in Theorem 12.4, are also valid without change for $V_n(C)$. The triangle inequality, $\|A + B\| \leq \|A\| + \|B\|$, also holds in $V_n(C)$.

Orthogonality of vectors in $V_n(C)$ is defined by the relation $A \cdot B = 0$. As in the real case, two vectors $A$ and $B$ in $V_n(C)$ are orthogonal whenever they satisfy the Pythagorean identity, $\|A + B\|^2 = \|A\|^2 + \|B\|^2$.

The concepts of linear span, linear independence, linear dependence, and basis, are defined for $V_n(\mathbf{C})$ exactly as in the real case. Theorems 12.7 through 12.10 and their proofs are all valid without change for $V_n(\mathbf{C})$.

## 12.17 Exercises

1. Let $A = (1, i)$, $B = (i, -i)$, and $C = (2i, 1)$ be three vectors in $V_2(\mathbf{C})$. Compute each of the following dot products:
(a) $A \cdot B$; (b) $B \cdot A$; (c) $(iA) \cdot B$; (d) $A \cdot (iB)$; (e) $(iA) \cdot (iB)$;
(f) $B \cdot C$; (g) $A \cdot C$; (h) $(B + C) \cdot A$; (i) $(A - C) \cdot B$;
(j) $(A - iB) \cdot (A + iB)$.
2. If $A = (2, 1, -i)$ and $B = (i, -1, 2i)$, find a nonzero vector $C$ in $V_3(\mathbf{C})$ orthogonal to both $A$ and $B$.
3. Prove that for any two vectors $A$ and $B$ in $V_n(\mathbf{C})$, we have the identity

$$\|A + B\|^2 = \|A\|^2 + \|B\|^2 + A \cdot B + \overline{A \cdot B}.$$

4. Prove that for any two vectors $A$ and $B$ in $V_n(\mathbf{C})$, we have the identity

$$\|A + B\|^2 - \|A - B\|^2 = 2(A \cdot B + \overline{A \cdot B}).$$

5. Prove that for any two vectors $A$ and $B$ in $V_n(\mathbf{C})$, we have the identity

$$\|A + B\|^2 + \|A - B\|^2 = 2\,\|A\|^2 + 2\,\|B\|^2.$$

6. (a) Prove that for any two vectors $A$ and $B$ in $V_n(\mathbf{C})$, the sum $\overline{A \cdot B} + A \cdot B$ is real.
(b) If $A$ and $B$ are nonzero vectors in $V_n(\mathbf{C})$, prove that

$$-2 \le \frac{A \cdot B + \overline{A \cdot B}}{\|A\|\,\|B\|} \le 2.$$

7. We define the angle $\theta$ between two nonzero vectors $A$ and $B$ in $V_n(\mathbf{C})$ by the equation

$$\theta = \arccos \frac{\frac{1}{2}(A \cdot B + \overline{A \cdot B})}{\|A\|\,\|B\|}.$$

The inequality in Exercise 6 shows that there is always a unique angle $\theta$ in the closed interval $0 \le \theta \le \pi$ satisfying this equation. Prove that we have

$$\|A - B\|^2 = \|A\|^2 + \|B\|^2 - 2\,\|A\|\,\|B\| \cos \theta.$$

8. Use the definition in Exercise 7 to compute the angle between the following two vectors in $V_5(\mathbf{C})$: $A = (1, 0, i, i, i)$, and $B = (i, i, i, 0, i)$.
9. (a) Prove that the following three vectors form a basis for $V_3(\mathbf{C})$: $A = (1, 0, 0)$, $B = (0, i, 0)$, $C = (1, 1, i)$.
(b) Express the vector $(5, 2 - i, 2i)$ as a linear combination of $A$, $B$, $C$.
10. Prove that the basis of unit coordinate vectors $E_1, \ldots, E_n$ in $V_n$ is also a basis for $V_n(\mathbf{C})$.

# 13

# APPLICATIONS OF VECTOR ALGEBRA TO ANALYTIC GEOMETRY

## 13.1 Introduction

This chapter discusses applications of vector algebra to the study of lines, planes, and conic sections. In Chapter 14 vector algebra is combined with the methods of calculus, and further applications are given to the study of curves and to some problems in mechanics.

The study of geometry as a deductive system, as conceived by Euclid around 300 B.C., begins with a set of axioms or postulates which describe properties of points and lines. The concepts "point" and "line" are taken as primitive notions and remain undefined. Other concepts are defined in terms of points and lines, and theorems are systematically deduced from the axioms. Euclid listed ten axioms from which he attempted to deduce all his theorems. It has since been shown that these axioms are not adequate for the theory. For example, in the proof of his very first theorem Euclid made a tacit assumption concerning the intersection of two circles that is not covered by his axioms. Since then other lists of axioms have been formulated that do give all of Euclid's theorems. The most famous of these is a list given by the German mathematician David Hilbert (1862–1943) in his now classic *Grundlagen der Geometrie*, published in 1899. (An English translation exists: *The Foundations of Geometry*, Open Court Publishing Co., 1947.) This work, which went through seven German editions in Hilbert's lifetime, is said to have inaugurated the abstract mathematics of the twentieth century.

Hilbert starts his treatment of plane geometry with five undefined concepts: *point*, *line*, *on* (a relation holding between a point and a line), *between* (a relation between a point and a pair of points), and *congruence* (a relation between pairs of points). He then gives fifteen axioms from which he develops all of plane Euclidean geometry. His treatment of solid geometry is based on twenty-one axioms involving six undefined concepts.

The approach in analytic geometry is somewhat different. We define concepts such as point, line, on, between, etc., but we do so in terms of real numbers, which are left undefined. The resulting mathematical structure is called an *analytic model* of Euclidean geometry. In this model, properties of real numbers are used to deduce Hilbert's axioms. We shall not attempt to describe all of Hilbert's axioms. Instead, we shall merely indicate how the primitive concepts may be defined in terms of numbers and give a few proofs to illustrate the methods of analytic geometry.