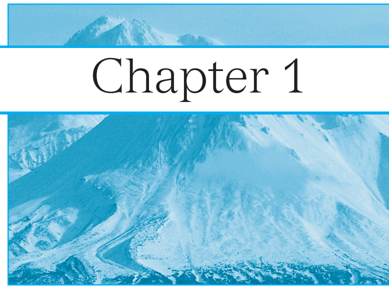


Chapter 1

Vector Analysis

1.1 Elementary Approach

In science and engineering we frequently encounter quantities that have only magnitude: mass, time, and temperature. This magnitude remains the same no matter how we orient the coordinate axes that we may use. These quantities we label scalar quantities. In contrast, many interesting physical quantities have magnitude or length and, in addition, an associated direction. Quantities with magnitude and direction are called vectors. Their length and the angle between any vectors remain unaffected by the orientation of coordinates we choose. To distinguish vectors from scalars, we identify vector quantities with boldface type (i.e., $\mathbf{V}$). Vectors are useful in solving systems of linear equations (Chapter 3). They are not only helpful in Euclidean geometry but also indispensable in classical mechanics and engineering because force, velocity, acceleration, and angular momentum are vectors. Electrodynamics is unthinkable without vector fields such as electric and magnetic fields.

Practical problems of mechanics and geometry, such as searching for the shortest distance between straight lines or parameterizing the orbit of a particle, will lead us to the differentiation of vectors and to vector analysis. Vector analysis is a powerful tool to formulate equations of motions of particles and then solve them in mechanics and engineering, or field equations of electrodynamics.

In this section, we learn to add and subtract vectors geometrically and algebraically in terms of their rectangular components.

A vector may be geometrically represented by an arrow with length proportional to the magnitude. The direction of the arrow indicates the direction of the vector, the positive sense of direction being indicated by the point. In

Figure 1.1

Triangle Law of Vector Addition

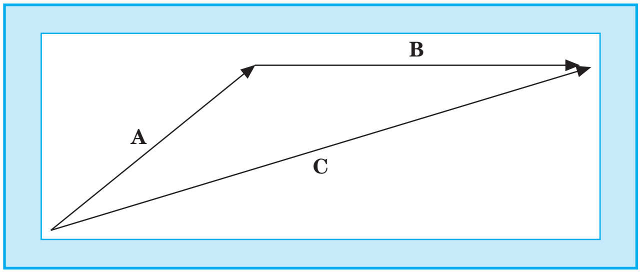
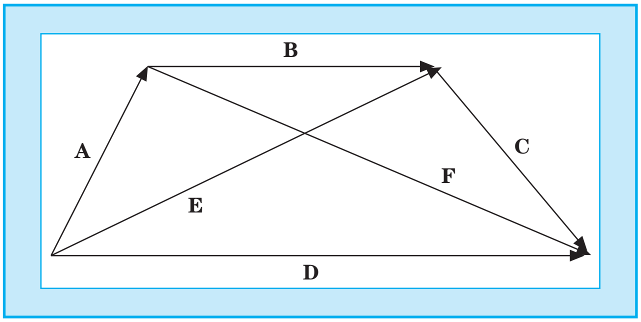


Figure 1.2

Vector Addition Is Associative



this representation, vector addition

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \tag{1.1}$$

consists of placing the rear end of vector **B** at the point of vector **A** (head to tail rule). Vector **C** is then represented by an arrow drawn from the rear of **A** to the point of **B**. This procedure, the triangle law of addition, assigns meaning to Eq. (1.1) and is illustrated in Fig. 1.1. By completing the parallelogram (sketch it), we see that

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}. \tag{1.2}$$

In words, **vector addition is commutative**.

For the sum of three vectors

$$\mathbf{D} = \mathbf{A} + \mathbf{B} + \mathbf{C},$$

illustrated in Fig. 1.2, we first add **A** and **B**:

$$\mathbf{A} + \mathbf{B} = \mathbf{E}.$$

Then this sum is added to **C**:

$$\mathbf{D} = \mathbf{E} + \mathbf{C}.$$

Alternatively, we may first add **B** and **C**:

$$\mathbf{B} + \mathbf{C} = \mathbf{F}.$$

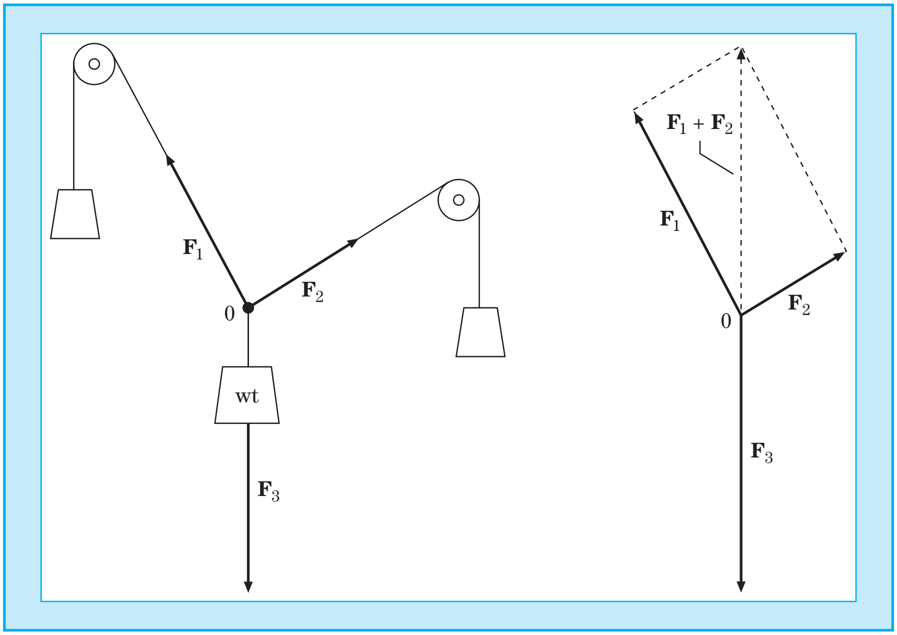


Figure 1.3

Equilibrium of Forces: $\mathbf{F}_1 + \mathbf{F}_2 = -\mathbf{F}_3$

Then

$$\mathbf{D} = \mathbf{A} + \mathbf{F}.$$

In terms of the original expression,

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$$

so that these alternative ways of summing three vectors lead to the same vector, or **vector addition is associative**.

A direct physical example of the parallelogram addition law is provided by a weight suspended by two cords in Fig. 1.3. If the junction point is in equilibrium, the vector sum of the two forces $\mathbf{F}_1$ and $\mathbf{F}_2$ must cancel the downward force of gravity, $\mathbf{F}_3$. Here, the parallelogram addition law is subject to immediate experimental verification.[1] Such a balance of forces is of immense importance for the stability of buildings, bridges, airplanes in flight, etc.

Subtraction is handled by defining the negative of a vector as a vector of the same magnitude but with reversed direction. Then

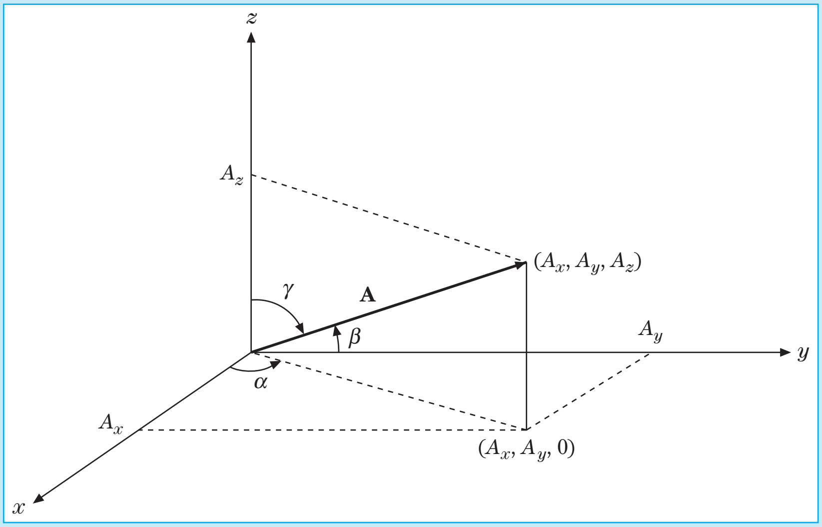
$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}).$$

The graphical representation of vector $\mathbf{A}$ by an arrow suggests using coordinates as a second possibility. Arrow $\mathbf{A}$ (Fig. 1.4), starting from the

[1]Strictly speaking, the parallelogram addition was introduced as a definition. Experiments show that forces are vector quantities that are combined by parallelogram addition, as required by the equilibrium condition of zero resultant force.

Figure 1.4

Components and Direction Cosines of A



origin,[2] terminates at the point (A_x, A_y, A_z). Thus, if we agree that the vector is to start at the origin, the positive end may be specified by giving the **rectangular** or **Cartesian** coordinates (A_x, A_y, A_z) of the arrow head.

Although **A** could have represented any vector quantity (momentum, electric field, etc.), one particularly important vector quantity, the distance from the origin to the point (x, y, z), is denoted by the special symbol **r**. We then have a choice of referring to the displacement as either the vector **r** or the collection (x, y, z), the coordinates of its end point:

$$\mathbf{r} \leftrightarrow (x, y, z). \tag{1.3}$$

Defining the magnitude r of vector **r** as its geometrical length, we find that Fig. 1.4 shows that the end point coordinates and the magnitude are related by

$$x = r\cos\alpha, \quad y = r\cos\beta, \quad z = r\cos\gamma. \tag{1.4}$$

$\cos\alpha$, $\cos\beta$, and $\cos\gamma$ are called the **direction cosines**, where α is the angle between the given vector and the positive x-axis, and so on. The (Cartesian) components A_x, A_y, and A_z can also be viewed as the **projections** of **A** on the respective axes.

Thus, any vector **A** may be resolved into its components (or projected onto the coordinate axes) to yield $A_x = A\cos\alpha$, etc., as in Eq. (1.4). We refer to the vector as a single quantity **A** or to its components (A_x, A_y, A_z). Note that the subscript x in A_x denotes the x component and not a dependence on the variable x. The choice between using **A** or its components (A_x, A_y, A_z) is

[2] We could start from any point; we choose the origin for simplicity. This freedom of shifting the origin of the coordinate system without affecting the geometry is called **translation invariance**.

essentially a choice between a **geometric or an algebraic representation**. The geometric "arrow in space" often aids in visualization. The algebraic set of components is usually more suitable for precise numerical or algebraic calculations. (This is illustrated in Examples 1.1.1–1.1.3 and also applies to Exercises 1.1.1, 1.1.3, 1.1.5, and 1.1.6.)

Vectors enter physics in two distinct forms:

- Vector **A** may represent a single force acting at a single point. The force of gravity acting at the center of gravity illustrates this form.
- Vector **A** may be defined over some extended region; that is, **A** and its components may be functions of position: $A_x = A_x(x, y, z)$, and so on.

Imagine a vector **A** attached to each point (x, y, z), whose length and direction change with position. Examples include the velocity of air around the wing of a plane in flight varying from point to point and electric and magnetic fields (made visible by iron filings). Thus, vectors defined at each point of a region are usually characterized as a **vector field**. The concept of the vector defined over a region and being a function of position will be extremely important in Section 1.2 and in later sections in which we differentiate and integrate vectors.

A **unit vector** has length 1 and may point in any direction. Coordinate unit vectors are implicit in the projection of **A** onto the coordinate axes to define its Cartesian components. Now, we define $\hat{\mathbf{x}}$ explicitly as a vector of unit magnitude pointing in the positive x-direction, $\hat{\mathbf{y}}$ as a vector of unit magnitude in the positive y-direction, and $\hat{\mathbf{z}}$ as a vector of unit magnitude in the positive z-direction. Then $\hat{\mathbf{x}}A_x$ is a vector with magnitude equal to A_x and in the positive x-direction; that is, the projection of **A** onto the x-direction, etc. By vector addition

$$\mathbf{A} = \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z, \tag{1.5}$$

which states that a vector equals the vector sum of its components or projections. Note that if **A** vanishes, all of its components must vanish individually; that is, if

$$\mathbf{A} = 0, \qquad \text{then } A_x = A_y = A_z = 0.$$

Finally, by the Pythagorean theorem, the length of vector **A** is

$$A = \left(A_x^2 + A_y^2 + A_z^2\right)^{1/2}. \tag{1.6}$$

This resolution of a vector into its components can be carried out in a variety of coordinate systems, as demonstrated in Chapter 2. Here, we restrict ourselves to Cartesian coordinates, where the unit vectors have the coordinates $\hat{\mathbf{x}} = (1, 0, 0)$, $\hat{\mathbf{y}} = (0, 1, 0)$, and $\hat{\mathbf{z}} = (0, 0, 1)$.

Equation (1.5) means that the three unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ span the real three-dimensional space: Any constant vector may be written as a linear combination of $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. Since $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are linearly independent (no one is a linear combination of the other two), they form a **basis** for the real three-dimensional space.

Complementary to the geometrical technique, algebraic addition and subtraction of vectors may now be carried out in terms of their components. For

$$\mathbf{A} = \hat{\mathbf{x}}A_x + \hat{\mathbf{y}}A_y + \hat{\mathbf{z}}A_z \text{ and } \mathbf{B} = \hat{\mathbf{x}}B_x + \hat{\mathbf{y}}B_y + \hat{\mathbf{z}}B_z,$$

$$\mathbf{A} \pm \mathbf{B} = \hat{\mathbf{x}}(A_x \pm B_x) + \hat{\mathbf{y}}(A_y \pm B_y) + \hat{\mathbf{z}}(A_z \pm B_z). \tag{1.7}$$

Biographical Data

Descartes, René. Descartes, a French mathematician and philosopher, was born in La Haye, France, in 1596 and died in Stockholm, Sweden, in 1650. Cartesius is the latinized version of his name at a time when Latin was the language of sciences, although he mainly wrote in French. He discovered his love of mathematics in the army, when he had plenty of time for research. He introduced the concept of rectangular coordinates, thereby converting geometry to algebraic equations of the coordinates of points, now called analytic geometry. Thus, he paved the way for Newton's and Leibniz's calculus. He coined the phrase "Cogito, ergo sum," which translates to "I think, therefore I am."

EXAMPLE 1.1.1

Let

$$\mathbf{A} = 6\hat{\mathbf{x}} + 4\hat{\mathbf{y}} + 3\hat{\mathbf{z}}$$

$$\mathbf{B} = 2\hat{\mathbf{x}} - 3\hat{\mathbf{y}} - 3\hat{\mathbf{z}}.$$

Then by Eq. (1.7)

$$\mathbf{A} + \mathbf{B} = (6+2)\hat{\mathbf{x}} + (4-3)\hat{\mathbf{y}} + (3-3)\hat{\mathbf{z}} = 8\hat{\mathbf{x}} + \hat{\mathbf{y}},$$

$$\mathbf{A} - \mathbf{B} = (6-2)\hat{\mathbf{x}} + (4+3)\hat{\mathbf{y}} + (3+3)\hat{\mathbf{z}} = 4\hat{\mathbf{x}} + 7\hat{\mathbf{y}} + 6\hat{\mathbf{z}}. \blacksquare$$

EXAMPLE 1.1.2

Parallelogram of Forces Find the sum of two forces **a** and **b**. To practice the geometric meaning of vector addition and subtraction, consider two forces

$$\mathbf{a} = (3, 0, 1), \quad \mathbf{b} = (4, 1, 2)$$

(in units of newtons, $1\,\mathrm{N} = 1\,\mathrm{kgm/s^2}$, in the Standard International system of units) that span a parallelogram with the diagonals forming the sum

$$\mathbf{a} + \mathbf{b} = (3+4, 1, 1+2) = (7, 1, 3) = \mathbf{b} + \mathbf{a},$$

and the difference

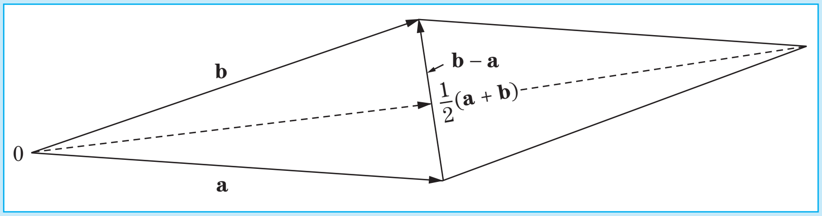
$$\mathbf{b} - \mathbf{a} = (4-3, 1, 2-1) = (1, 1, 1),$$

as shown in Fig. 1.5. The midpoint **c** is half the sum,

$$\mathbf{c} = \frac{1}{2}(\mathbf{a} + \mathbf{b}) = \left(\frac{7}{2}, \frac{1}{2}, \frac{3}{2}\right).$$

Figure 1.5

Parallelogram of Forces a and b



Alternately, to obtain the midpoint from **a**, add half of the second diagonal that points from **a** to **b**; that is,

$$\mathbf{a} + \frac{1}{2}(\mathbf{b} - \mathbf{a}) = \frac{1}{2}(\mathbf{a} + \mathbf{b}) = \mathbf{c} = \left(\frac{7}{2}, \frac{1}{2}, \frac{3}{2}\right). \quad ■$$

EXAMPLE 1.1.3

Center of Mass of Three Points at the Corners of a Triangle Consider each corner of a triangle to have a unit of mass and to be located $\mathbf{a}_i$ from the origin, where

$$\mathbf{a}_1 = (2, 0, 0),\ \mathbf{a}_2 = (4, 1, 1),\ \mathbf{a}_3 = (3, 3, 2).$$

Then, the center of mass of the triangle is

$$\frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3) = \mathbf{c} = \frac{1}{3}(2 + 4 + 3,\ 1 + 3,\ 1 + 2) = \left(3, \frac{4}{3}, 1\right).$$

THEOREM 1.1

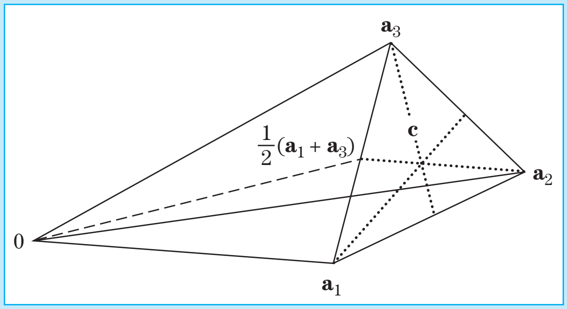
If we draw a straight line from each corner to the midpoint of the opposite side of the triangle in Fig. 1.6, these lines meet in the center, which is at a distance of two-thirds of the line length to the corner.

The three midpoints are located at the point of the vectors

$$\frac{1}{2}(\mathbf{a}_1 + \mathbf{a}_2) = \frac{1}{2}(2 + 4, 1, 1) = \left(3, \frac{1}{2}, \frac{1}{2}\right)$$
$$\frac{1}{2}(\mathbf{a}_2 + \mathbf{a}_3) = \frac{1}{2}(4 + 3, 1 + 3, 1 + 2) = \left(\frac{7}{2}, 2, \frac{3}{2}\right)$$
$$\frac{1}{2}(\mathbf{a}_3 + \mathbf{a}_1) = \frac{1}{2}(3 + 2, 3, 2) = \left(\frac{5}{2}, \frac{3}{2}, 1\right).$$

Figure 1.6

Center of a Triangle. The Dashed Line Goes from the Origin to the Midpoint of a Triangle Side, and the Dotted Lines Go from Each Corner to the Midpoint of the Opposite Triangle Side



To prove this theorem numerically or symbolically using general vectors, we start from each corner and end up in the center as follows:

$$(2, 0, 0) + \frac{2}{3}\left[\left(\frac{7}{2}, 2, \frac{3}{2}\right) - (2, 0, 0)\right] = \left(3, \frac{4}{3}, 1\right)$$

$$\mathbf{a}_1 + \frac{2}{3}\left(\frac{1}{2}(\mathbf{a}_2 + \mathbf{a}_3) - \mathbf{a}_1\right) = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3),$$

$$(4, 1, 1) + \frac{2}{3}\left[\left(\frac{5}{2}, \frac{3}{2}, 1\right) - (4, 1, 1)\right] = \left(3, \frac{4}{3}, 1\right)$$

$$\mathbf{a}_2 + \frac{2}{3}\left(\frac{1}{2}(\mathbf{a}_1 + \mathbf{a}_3) - \mathbf{a}_2\right) = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3),$$

$$(3, 3, 2) + \frac{2}{3}\left[\left(3, \frac{1}{2}, \frac{1}{2}\right) - (3, 3, 2)\right] = \left(3, \frac{4}{3}, 1\right)$$

$$\mathbf{a}_3 + \frac{2}{3}\left(\frac{1}{2}(\mathbf{a}_1 + \mathbf{a}_2) - \mathbf{a}_3\right) = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3).$$

This theorem is easy to establish using vector algebra, but it is much more tedious to prove working only with intersecting straight lines of Euclidean geometry. Thus, this example is not only useful practice but also lets us appreciate the power and versatility of the vector algebra. ■

EXAMPLE 1.1.4

Elastic Forces A movable point $\mathbf{a}$ is held elastically (denoted by springs in Fig. 1.7) by three fixed points $\mathbf{a}_i$, $i = 1, 2, 3$; that is, the force $\mathbf{F}_i = k_i(\mathbf{a}_i - \mathbf{a})$ for each i that $\mathbf{a}$ experiences is pointed to $\mathbf{a}_i$ and proportional to the distance. Let us show that the elastic forces to three points can be replaced by an elastic force to a single point.

This holds because the total force is given by

$$\mathbf{F} = \sum_i \mathbf{F}_i = \sum_i k_i\mathbf{a}_i - \mathbf{a}\sum_i k_i = \left(\sum_i k_i\right)\left(\frac{\sum_i k_i\mathbf{a}_i}{\sum_i k_i} - \mathbf{a}\right) = k_0(\mathbf{a}_0 - \mathbf{a}),$$

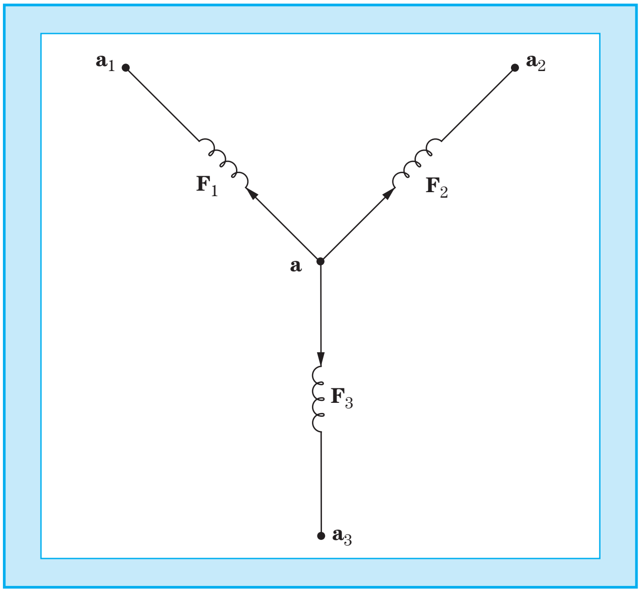


Figure 1.7

The Point a Is Held Elastically by Three Points $\mathbf{a}_i$

where $k_0 = \sum_i k_i$ and $\mathbf{a}_0 = \sum_i k_i \mathbf{a}_i / k_0$. This shows that the resulting force is equivalent to one with an effective spring constant k_0 acting from a point $\mathbf{a}_0$. Note that if all k_i's are the same, then $\mathbf{a}_0 = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3)$ is the center of mass.

Technical applications apply for bridges and buildings, for which the balance of forces is vital for stability. ■

Vectors and Vector Space Summary

An ordered triplet of real numbers (x_1, x_2, x_3) is labeled a **vector x**. The number x_n is called the nth component of vector $\mathbf{x}$. The collection of all such vectors (obeying the properties that follow) forms a three-dimensional real **vector space**, or **linear space**. We ascribe five properties to our vectors: If $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$,

1. Vector equality: $\mathbf{x} = \mathbf{y}$ means $x_i = y_i$, $i = 1, 2, 3$.
2. Vector addition: $\mathbf{x} + \mathbf{y} = \mathbf{z}$ means $x_i + y_i = z_i$, $i = 1, 2, 3$.
3. Scalar multiplication: $a\mathbf{x} = (ax_1, ax_2, ax_3)$.
4. Negative of a vector: $-\mathbf{x} = (-1)\mathbf{x} = (-x_1, -x_2, -x_3)$.
5. Null vector: There exists a null vector $\mathbf{0} = (0, 0, 0)$.

Since our vector components are numbers, the following properties also hold:

1. Addition of vectors is commutative: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
2. Addition of vectors is associative: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.

3. Scalar multiplication is distributive:

$$a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}, \qquad \text{also} \qquad (a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}.$$

4. Scalar multiplication is associative: $(ab)\mathbf{x} = a(b\mathbf{x})$.

Furthermore, the null vector $\mathbf{0}$ is unique, as is the negative of a given vector $\mathbf{x}$.

With regard to the vectors, this approach merely formalizes the component discussion of Section 1.1. The importance lies in the extensions, which will be considered later. In Chapter 3, we show that vectors form a linear space, with the transformations in the linear space described by matrices. Finally, and perhaps most important, for advanced physics the concept of vectors presented here generalizes to (i) complex quantities,[3] (ii) functions, and (iii) an infinite number of components. This leads to infinite dimensional function spaces, the Hilbert spaces, which are important in quantum mechanics. A brief introduction to function expansions and Hilbert space is provided in Chapter 9.

SUMMARY

So far, we have defined the operations of addition and subtraction of vectors guided by the use of elastic and gravitational forces in classical mechanics, set up mechanical and geometrical problems such as finding the center of mass of a system of mass points, and solved these problems using the tools of vector algebra.

Next, we address three varieties of multiplication defined on the basis of their applicability in geometry and mechanics: a scalar or inner product in Section 1.2; a vector product peculiar to three-dimensional space in Section 1.3, for which the angular momentum in mechanics is a prime example; and a direct or outer product yielding a second-rank tensor in Section 2.7. Division by a vector cannot be consistently defined.

EXERCISES

1.1.1 A jet plane is flying eastward from Kennedy Airport at a constant speed of 500 mph. There is a crosswind from the south at 50 mph. What is the resultant speed of the plane relative to the ground? Draw the velocities (using graphical software, if available).

1.1.2 A boat travels straight across a river at a speed of 5 mph when there is no current. You want to go straight across the river in that boat when there is a constant current flowing at 1 mph. At what angle do you have to steer the boat? Plot the velocities.

1.1.3 A sphere of radius a is centered at a point $\mathbf{r}_1$.

(a) Write out the algebraic equation for the sphere. Explain in words why you chose a particular form. Name theorems from geometry you may have used.

[3] The n-dimensional vector space of real n-tuples is often labeled $\mathbf{R}^n$, and the n-dimensional vector space of complex n-tuples is labeled $\mathbf{C}^n$.

(b) Write out a **vector** equation for the sphere. Identify in words what you are doing.

ANS. (a) $(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2 = a^2$.
(b) $\mathbf{r} = \mathbf{r}_1 + \mathbf{a}$ ($\mathbf{a}$ takes on all directions but has a fixed magnitude, a).

1.1.4 Show that the medians of a triangle intersect at a point. Show that this point is two-thirds of the way from any corner of the triangle to the midpoint of the opposite side. Compare a geometrical proof with one using vectors. If you use a Cartesian coordinate system, place your triangle so as to simplify the analysis as much as possible. Explain in words why you are allowed to do so.

1.1.5 The velocity of sailboat A relative to sailboat B, $\mathbf{v}_{\text{rel}}$, is defined by the equation $\mathbf{v}_{\text{rel}} = \mathbf{v}_A - \mathbf{v}_B$, where $\mathbf{v}_A$ is the velocity of A and $\mathbf{v}_B$ is the velocity of B. Determine the velocity of A relative to B if

$$\mathbf{v}_A = 30 \text{ km/hr east}$$
$$\mathbf{v}_B = 40 \text{ km/hr north.}$$

Plot the velocities (using graphical software, if available).

ANS. $\mathbf{v}_{\text{rel}} = 50$ km/hr, 53.1° south of east.

1.1.6 A sailboat sails for 1 hr at 4 km/hr (relative to the water) on a steady compass heading of 40° east of north. The sailboat is simultaneously carried along by a current. At the end of the hour the boat is 6.12 km from its starting point. The line from its starting point to its location lies 60° east of north. Find the x (easterly) and y (northerly) components of the water's velocity. Plot all velocities.

ANS. $v_{\text{east}} = 2.73$ km/hr, $v_{\text{north}} \approx 0$ km/hr.

1.1.7 A triangle is defined by the vertices of three vectors, $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, that extend from the origin. In terms of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, show that the **vector** sum of the successive sides of the triangle is zero. If software is available, plot a typical case.

1.1.8 Find the diagonal vectors of a unit cube with one corner at the origin and three adjacent sides lying along the three axes of a Cartesian coordinate system. Show that there are four diagonals with length $\sqrt{3}$. Representing these as vectors, what are their components? Show that the diagonals of the cube's surfaces have length $\sqrt{2}$. Determine their components.

1.1.9 Hubble's law: Hubble found that distant galaxies are receding with a velocity proportional to their distance (H_0 is the Hubble constant) from where we are on Earth. For the ith galaxy

$$\mathbf{v}_i = H_0 \mathbf{r}_i,$$

with our Milky Way galaxy at the origin. Show that this recession of the galaxies from us does **not** imply that we are at the center of the universe.

Specifically, take the galaxy at $\mathbf{r}_1$ as a new origin and show that Hubble's law is still obeyed.

1.2 Scalar or Dot Product

Having defined vectors, we now proceed to combine them in this section. The laws for combining vectors must be mathematically consistent. From the possibilities that are consistent we select two that are both mathematically and physically interesting. In this section, we start with the scalar product that is based on the geometric concept of projection that we used in Section 1.1 to define the Cartesian components of a vector. Also included here are some applications to particle orbits and analytic geometry that will prompt us to **differentiate vectors**, thus starting **vector analysis**.

The **projection of a vector A** onto a coordinate axis, which defines its Cartesian components in Eq. (1.5), **is a special case of the scalar product** of **A** and the coordinate unit vectors,

$$A_x = A\cos\alpha \equiv \mathbf{A}\cdot\hat{\mathbf{x}}, \quad A_y = A\cos\beta \equiv \mathbf{A}\cdot\hat{\mathbf{y}}, \quad A_z = A\cos\gamma \equiv \mathbf{A}\cdot\hat{\mathbf{z}} \tag{1.8}$$

and leads us to the general definition of the dot product. Just as the projection is linear in **A**, we want the scalar product of two vectors to be linear in **A** and **B**—that is, to obey the distributive and associative laws

$$\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) = \mathbf{A}\cdot\mathbf{B} + \mathbf{A}\cdot\mathbf{C} \tag{1.9}$$

$$\mathbf{A}\cdot(y\mathbf{B}) = (y\mathbf{A})\cdot\mathbf{B} = y\mathbf{A}\cdot\mathbf{B}, \tag{1.10}$$

where y is a real number. Now we can use the decomposition of **B** into its Cartesian components according to Eq. (1.5), $\mathbf{B} = B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}$, to construct the general scalar or dot product of the vectors **A** and **B** from the special case as

$$\begin{aligned}
\mathbf{A}\cdot\mathbf{B} &= \mathbf{A}\cdot(B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}),\\
&= B_x\mathbf{A}\cdot\hat{\mathbf{x}} + B_y\mathbf{A}\cdot\hat{\mathbf{y}} + B_z\mathbf{A}\cdot\hat{\mathbf{z}}, \qquad \text{applying Eqs. (1.9) and (1.10)}\\
&= B_xA_x + B_yA_y + B_zA_z, \qquad \text{upon substituting Eq. (1.8).}
\end{aligned}$$

Hence,

$$\mathbf{A}\cdot\mathbf{B} \equiv \sum_i A_iB_i = \sum_i B_iA_i = \mathbf{B}\cdot\mathbf{A} \tag{1.11}$$

because we are dealing with components.

If $\mathbf{A} = \mathbf{B}$ in Eq. (1.11), we recover the magnitude $A = (\sum_i A_i^2)^{1/2}$ of **A** in Eq. (1.6) from Eq. (1.11).

It is obvious from Eq. (1.11) that the scalar product treats **A** and **B** alike, is symmetric in **A** and **B**, or is commutative. Based on this observation, we can generalize Eq. (1.8) to the projection of **A** onto an arbitrary vector $\mathbf{B} \neq 0$

Figure 1.8

Scalar Product
$\mathbf{A} \cdot \mathbf{B} = AB\cos\theta$

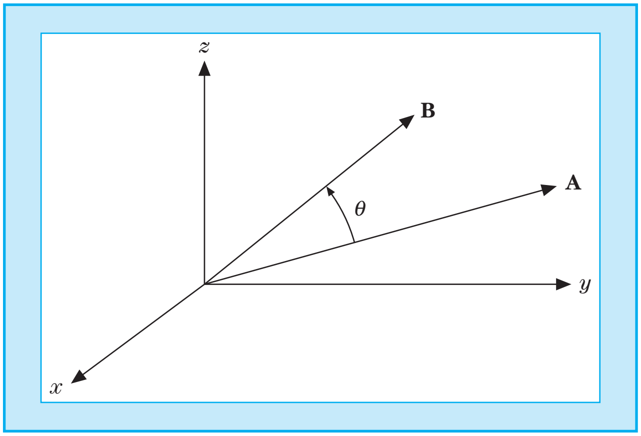
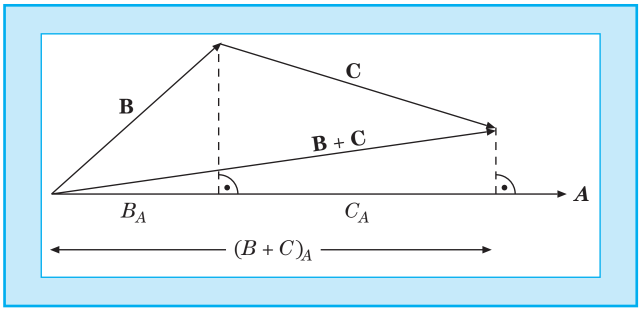


Figure 1.9

The Distributive Law
$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = AB_A + AC_A = A(\mathbf{B} + \mathbf{C})_A$
[Eq. (1.9)]



instead of the coordinate unit vectors. As a first step in this direction, we define A_B as $A_B = A\cos\theta \equiv \mathbf{A} \cdot \hat{\mathbf{B}}$, where $\hat{\mathbf{B}} = \mathbf{B}/B$ is the unit vector in the direction of $\mathbf{B}$ and θ is the angle between $\mathbf{A}$ and $\mathbf{B}$ as shown in Fig. 1.8. Similarly, we project $\mathbf{B}$ onto $\mathbf{A}$ as $B_A = B\cos\theta \equiv \mathbf{B} \cdot \hat{\mathbf{A}}$. These projections are not symmetric in $\mathbf{A}$ and $\mathbf{B}$. To make them symmetric in $\mathbf{A}$ and $\mathbf{B}$, we define

$$\mathbf{A} \cdot \mathbf{B} \equiv A_B B = AB_A = AB\cos\theta. \tag{1.12}$$

The distributive law in Eq. (1.9) is illustrated in Fig. 1.9, which states that the sum of the projections of $\mathbf{B}$ and $\mathbf{C}$ onto $\mathbf{A}$, $B_A + C_A$, is equal to the projection of $\mathbf{B} + \mathbf{C}$ onto $\mathbf{A}$, $(\mathbf{B} + \mathbf{C})_A$.

From Eqs. (1.8), (1.11), and (1.12), we infer that the coordinate unit vectors satisfy the relations

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1, \tag{1.13}$$

whereas

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0. \tag{1.14}$$

If the component definition of the dot product, Eq. (1.11), is labeled an algebraic definition, then Eq. (1.12) is a geometric definition. One of the most common applications of the scalar product in physics is in the definition of work = force · displacement · $\cos\theta$, where θ is the angle between the force and the displacement. This expression is interpreted as displacement times the projection of the force along the displacement direction—that is, the scalar product of force and displacement, $W = \mathbf{F} \cdot \mathbf{s}$.

If $\mathbf{A} \cdot \mathbf{B} = 0$ and we know that $\mathbf{A} \neq 0$ and $\mathbf{B} \neq 0$, then from Eq. (1.12) $\cos\theta = 0$ or $\theta = 90^\circ, 270^\circ$, and so on. The vectors $\mathbf{A}$ and $\mathbf{B}$ must be **perpendicular**. Alternately, we may say $\mathbf{A}$ and $\mathbf{B}$ are **orthogonal**. The unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are mutually orthogonal.

Free Motion and Other Orbits

EXAMPLE 1.2.1

Free Particle Motion To apply this notion of orthogonality in two dimensions, let us first deal with the motion of a particle free of forces along a straight line

$$\mathbf{r}(t) = (x(t), y(t)) = (-3t, 4t)$$

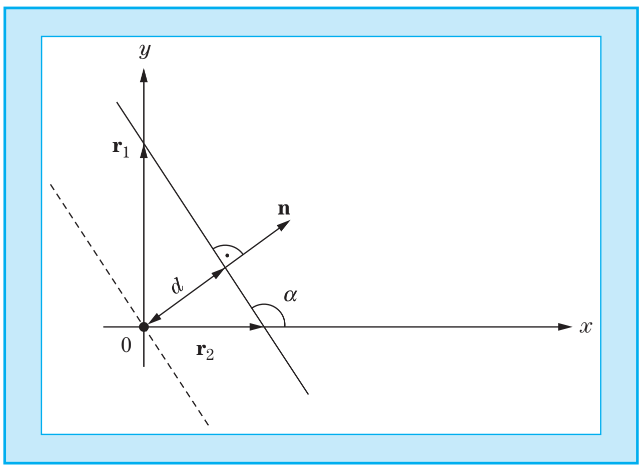
through the origin (dashed line in Fig. 1.10). The particle travels with the velocity $v_x = x/t = -3$ in the x-direction and $v_y = y/t = 4$ in the y-direction (in meters per second; e.g., 1 m/sec = 3.6 km/hr). The constant velocity $\mathbf{v} = (-3, 4)$ is characteristic of free motion according to Newton's equations.

Eliminating the time t, we find the homogeneous linear equation $4x + 3y = 0$, whose coefficient vector (4, 3) we normalize to unit length; that is, we write the linear equation as

$$\frac{4}{5}x + \frac{3}{5}y = 0 = \mathbf{n} \cdot \mathbf{r}.$$

Figure 1.10

The Dashed Line Is $\mathbf{n} \cdot \mathbf{r} = 0$ and the Solid Line Is $\mathbf{n} \cdot \mathbf{r} = d$



where $\mathbf{n} = (4/5, 3/5)$ is a constant unit vector and $\mathbf{r}$ is the coordinate vector varying in the xy-plane; that is, $\mathbf{r} = \hat{\mathbf{x}}x + \hat{\mathbf{y}}y$. The scalar product

$$\mathbf{n} \cdot \mathbf{r} = 0 \tag{1.15}$$

means that the projection onto $\mathbf{n}$ of the vector $\mathbf{r}(t)$ pointing from the origin (a point on the line) to the general point on the line is zero so that **n is the normal of the straight line**. We verify that

$$(-3t, 4t) \cdot \left(\frac{4}{5}, \frac{3}{5}\right) = \frac{t}{5}(-3 \cdot 4 + 4 \cdot 3) = 0.$$

Because the particle's velocity is a tangent vector of the line, we can also write the scalar product as $\mathbf{v} \cdot \mathbf{n} = 0$, omitting the normalization factor $t/v = t/5$.

If we throw the particle from the origin in an arbitrary direction with some velocity v, it also will travel on a line through the origin. That is, upon varying the normal unit vector the linear Eq. (1.15) defines an arbitrary straight line through the origin in the xy-plane. Notice that in three dimensions Eq. (1.15) describes a plane through the origin, and a hyperplane ($(n-1)$-dimensional subspace) in n-dimensional space.

Now we shift the line by some constant distance d along the normal direction $\mathbf{n}$ so that it passes through the point (3, 0) on the x-axis, for example. Because its tangent vector is $\mathbf{v}$, the line is parameterized as $x(t) = 3-3t$, $y(t) = 4t$. We can verify that it passes through the point $\mathbf{r}_2 = (3, 0)$ on the x-axis for $t = 0$ and $\mathbf{r}_1 = (0, 4)$ on the y-axis for $t = 1$. The particle has the same velocity and the path has the same normal. Eliminating the time as before, we find that the linear equation for the line becomes $4x + 3y = 12$, or

$$\mathbf{n} \cdot \mathbf{r} = d = \frac{12}{5}. \tag{1.16}$$

The line no longer goes through the origin (solid line in Fig. 1.10) but has the shortest distance $d = 12/5$ from the origin. If $\mathbf{r}_1 = (0, 4)$, $\mathbf{r}_2 = (3, 0)$ are our different points on that line, then $\mathbf{T} = \mathbf{r}_1 - \mathbf{r}_2 = (-3, 4) = \mathbf{v}$ is a tangent vector of the line and therefore orthogonal to the normal $\mathbf{n}$ because $\mathbf{n} \cdot \mathbf{T} = \mathbf{n} \cdot \mathbf{r}_1 - \mathbf{n} \cdot \mathbf{r}_2 = d - d = 0$ from Eq. (1.16). Then the general point on that line is parameterized by

$$\mathbf{r}(t) = \mathbf{r}_1 + t\mathbf{T} \tag{1.17}$$

because $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_1 + t\mathbf{n} \cdot \mathbf{T} = d + t \cdot 0 = d$.

Note that in general a straight line is defined by a linear relation, $\mathbf{n} \cdot \mathbf{r} = d$, and its points depend linearly on one variable t; that is, in two dimensions Eq. (1.17) represents $x = x_1 + tT_x$, $y = y_1 + tT_y$, with $\mathbf{T} = (T_x, T_y)$. The geometry of Fig. 1.10 shows that the projection of the vectors $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}$ on the normal $\mathbf{n}$ is always d—that is, the shortest distance of our line from the origin, consistent with the algebraic definition $\mathbf{n} \cdot \mathbf{r} = d$ of our line, from which we started.

Equations (1.16) and (1.17) are consistent with the conventional definition of a straight line by its constant slope (or angle α with the x-axis)

$$\tan\alpha = \frac{y - y_1}{x - x_1} \leftrightarrow (x - x_1)\sin\alpha - (y - y_1)\cos\alpha = 0, \tag{1.18}$$

where the normal $\mathbf{n} = (\sin\alpha, -\cos\alpha)$; upon comparing Eq. (1.18) with Eq. (1.16), $\mathbf{n}\cdot\mathbf{r} = d = x_1\sin\alpha - y_1\cos\alpha$.

Generalizing to three-dimensional analytic geometry, $\mathbf{n}\cdot\mathbf{r} = d$ is linear in the variables $(x, y, z) = \mathbf{r}$; that is, it represents a plane, and the unit vector $\mathbf{n} = (n_1, n_2, n_3)$ is perpendicular to the plane—it is the constant normal of the plane. If we divide the plane equation by d, the coefficients n_i/d of the coordinates x_i of the plane give the inverse lengths of the segments from the origin to the intersection of the Cartesian axes with the plane. For example, the point of the plane $6x + 3y + 2z = 6$ in Fig. 1.11 on the x-axis defined by $y = 0 = z$ is $(d/n_1 = 1, 0, 0)$ for $n_1 = 6/7$, $d = 6/7$, noting that $6^2+3^2+2^2 = 7^2$. The general point on the plane is parameterized as

$$\mathbf{r}(s, t) = \mathbf{r}_1 + s\mathbf{l}_1 + t\mathbf{l}_2,$$

where s and t are parameters, and it is constructed from three of its points $\mathbf{r}_i$, $i = 1, 2, 3$, that is, $\mathbf{r}_1 = (1, 0, 0)$, $\mathbf{r}_2 = (0, 2, 0)$, $\mathbf{r}_3 = (0, 0, 3)$ for the plane in Fig. 1.11, so that the tangent vectors $\mathbf{l}_1 = \mathbf{r}_2 - \mathbf{r}_1$, $\mathbf{l}_2 = \mathbf{r}_3 - \mathbf{r}_1$ of the plane are not parallel. All this generalizes to higher dimensions.

Geometry also tells us that two nonparallel planes $\mathbf{a}_1\cdot\mathbf{r} = d_1$, $\mathbf{a}_2\cdot\mathbf{r} = d_2$ in three-dimensional space have a line in common and three nonparallel planes a single point in general. Finding them amounts to solving linear equations, which is addressed in Section 3.1 using determinants.

Figure 1.11

The Plane
$\mathbf{6x + 3y + 2z = 6}$

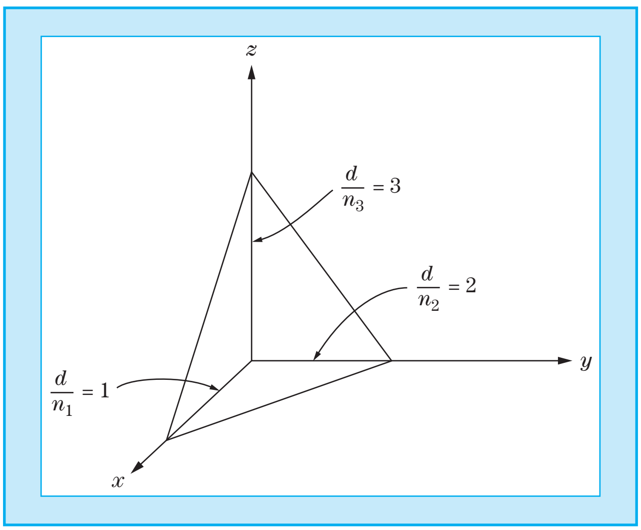
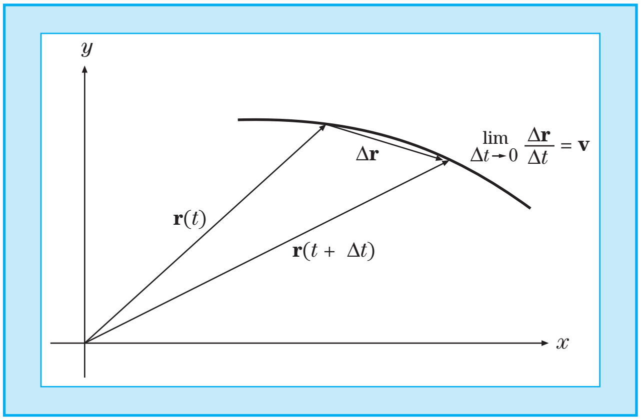


Figure 1.12

Differentiation of a Vector



More generally, the orbit of a particle or a curve in planar analytic geometry may be defined as $\mathbf{r}(t) = (x(t), y(t))$, where x and y are functions of the parameter t. In order to find the slope of a curve or the tangent to an orbit we need to differentiate vectors. Differentiating a vector function is a simple extension of differentiating scalar functions if we resolve $\mathbf{r}(t)$ into its Cartesian components. Then, for differentiation with respect to time, the linear velocity is given by

$$\frac{d\mathbf{r}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} = \mathbf{v} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \equiv (\dot{x}, \dot{y}, \dot{z})$$

because the Cartesian unit vectors are constant. Thus, differentiation of a vector always reduces directly to a vector sum of not more than three (for three-dimensional space) scalar derivatives. In other coordinate systems (see Chapter 2), the situation is more complicated because the unit vectors are no longer constant in direction. Differentiation with respect to the space coordinates is handled in the same way as differentiation with respect to time. Graphically, we have the slope of a curve, orbit, or trajectory, as shown in Fig. 1.12. ■

EXAMPLE 1.2.2

Shortest Distance of a Rocket from an Observer What is the shortest distance of a rocket traveling at a constant velocity $\mathbf{v} = (1, 2, 3)$ from an observer at $\mathbf{r}_0 = (2, 1, 3)$? The rocket is launched at **time** $t = 0$ at the point $\mathbf{r}_1 = (1, 1, 1)$.

The path of the rocket is the straight line

$$\mathbf{r} = \mathbf{r}_1 + t\mathbf{v}, \tag{1.19}$$

or, in Cartesian coordinates,

$$x(t) = 1 + t, \quad y(t) = 1 + 2t, \quad z(t) = 1 + 3t.$$

We now minimize the distance $|\mathbf{r}-\mathbf{r}_0|$ of the observer at the point $\mathbf{r}_0 = (2, 1, 3)$ from $\mathbf{r}(t)$, or equivalently $(\mathbf{r}-\mathbf{r}_0)^2 = \min$. Differentiating Eq. (1.19) with respect to t yields $\dot{\mathbf{r}} = (\dot{x}, \dot{y}, \dot{z}) = \mathbf{v}$. Setting $\frac{d}{dt}(\mathbf{r}-\mathbf{r}_0)^2 = 0$, we obtain the condition

$$2(\mathbf{r}-\mathbf{r}_0)\cdot\dot{\mathbf{r}} = 2[\mathbf{r}_1 - \mathbf{r}_0 + t\mathbf{v}]\cdot\mathbf{v} = 0.$$

Because $\dot{\mathbf{r}} = \mathbf{v}$ is the tangent vector of the line, the geometric meaning of this condition is that the **shortest distance vector through $\mathbf{r}_0$ is perpendicular to the line.** Now solving for t yields the ratio of scalar products

$$t = -\frac{(\mathbf{r}_1-\mathbf{r}_0)\cdot\mathbf{v}}{\mathbf{v}^2} = -\frac{(-1, 0, -2)\cdot(1, 2, 3)}{(1, 2, 3)\cdot(1, 2, 3)} = \frac{1+0+6}{1+4+9} = \frac{1}{2}.$$

Substituting this parameter value into Eq. (1.19) gives the point $\mathbf{r}_s = (3/2, 2, 5/2)$ on the line that is closest to $\mathbf{r}_0$. The shortest distance is $d = |\mathbf{r}_0 - \mathbf{r}_s| = |(-1/2, 1, -1/2)| = \sqrt{2/4+1} = \sqrt{3/2}$. ■

In two dimensions, $\mathbf{r}(t) = (x = a\cos t,\ y = b\sin t)$ describes an ellipse with half-axes a, b (so that $a = b$ gives a circle); for example, the orbit of a planet around the sun in a plane determined by the constant orbital angular momentum (the normal of the plane). If $\mathbf{r}_0 = (x(t_0), y(t_0)) = (x_0, y_0) = \mathbf{r}(t_0)$ is a point on our orbit, then the tangent at $\mathbf{r}_0$ has the slope $\dot{y}_0/\dot{x}_0$, where the dots denote the derivatives with respect to the time t as usual. Returning to the slope formula, imagine inverting $x = x(t)$ to find $t = t(x)$, which is substituted into $y = y(t) = y(t(x)) = f(x)$ to produce the standard form of a curve in analytic geometry. Using the **chain rule of differentiation**, the slope of $f(x)$ at x is

$$\frac{df}{dx} = f'(x) = \frac{dy(t(x))}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{\dot{y}}{\dot{x}}.$$

The tangent is a straight line and therefore depends linearly on one variable u,

$$\mathbf{r} = \mathbf{r}(t_0) + u\dot{\mathbf{r}}(t_0), \tag{1.20}$$

whereas the normal through the given point (x_0, y_0) obeys the linear equation

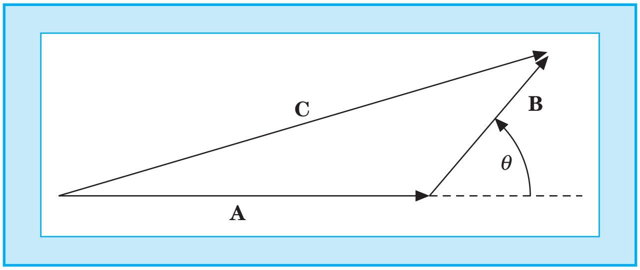
$$(x - x_0)\,\dot{x}_0 + (y - y_0)\,\dot{y}_0 = 0. \tag{1.21}$$

For the elliptical orbit mentioned previously, we check that the point $\mathbf{r}_0 = (0, b)$ for the parameter value $t = \pi/2$ lies on it. The slope at $t = \pi/2$ is zero, which we also know from geometry and because $\dot{y}_0 = b\cos t|_{t=\pi/2} = 0$, whereas $\dot{x}_0 = -a\sin t|_{\pi/2} = -a \neq 0$. The normal is the y-axis for which Eq. (1.21) yields $-ax = 0$.

A curve can also be defined implicitly by a functional relation, $F(x, y) = 0$, of the coordinates. This common case will be addressed in Section 1.5 because it involves partial derivatives.

Figure 1.13

The Law of Cosines



Law of Cosines In a similar geometrical approach, we take $\mathbf{C} = \mathbf{A} + \mathbf{B}$ and dot it into itself:

$$\mathbf{C} \cdot \mathbf{C} = (\mathbf{A} + \mathbf{B}) \cdot (\mathbf{A} + \mathbf{B}) = \mathbf{A} \cdot \mathbf{A} + \mathbf{B} \cdot \mathbf{B} + 2\mathbf{A} \cdot \mathbf{B}. \tag{1.22}$$

Since

$$\mathbf{C} \cdot \mathbf{C} = C^2, \tag{1.23}$$

the square of the magnitude of vector $\mathbf{C}$ is a scalar, we see that

$$\mathbf{A} \cdot \mathbf{B} = \frac{1}{2}(C^2 - A^2 - B^2) \tag{1.24}$$

is a scalar. Note that since the right-hand side of Eq. (1.24) is a scalar, the left-hand side $\mathbf{A} \cdot \mathbf{B}$ must also be a scalar, independent of the orientation of the coordinate system. We defer a proof that a scalar product is invariant under rotations to Section 2.6.

Equation (1.22) is another form of the **law of cosines**:

$$C^2 = A^2 + B^2 + 2AB\cos\theta. \tag{1.25}$$

Comparing Eqs. (1.22) and (1.25), we have another verification of Eq. (1.12) or, if preferred, a vector derivation of the law of cosines (Fig. 1.13). This law may also be derived from the triangle formed by the point of $\mathbf{C}$ and its line of shortest distance from the line along $\mathbf{A}$, which has the length $B\sin\theta$, whereas the projection of $\mathbf{B}$ onto $\mathbf{A}$ has length $B\cos\theta$. Applying the Pythagorean theorem to this triangle with a right angle formed by the point of $\mathbf{C}$, $\mathbf{A} + \mathbf{B} \cdot \hat{\mathbf{A}}$ and the shortest distance $B\sin\theta$ gives

$$C^2 = (A + \mathbf{B} \cdot \hat{\mathbf{A}})^2 + (B\sin\theta)^2 = A^2 + B^2(\cos^2\theta + \sin^2\theta) + 2AB\cos\theta.$$

SUMMARY

In this section, we defined the dot product as an algebraic generalization of the geometric concept of projection of vectors (their coordinates). We used it for geometrical purposes, such as finding the shortest distance of a point from a line or the cosine theorem for triangles. The geometrical meaning of the scalar product allowed us to go back and forth between the algebraic definition of a straight line as a linear equation and the parameterization of its general point $\mathbf{r}(t)$ as a linear function of time and similar steps for planes in three dimensions. We began differentiation of vectors as a tool for drawing tangents to orbits of particles, and this important step represents the start of vector analysis enlarging vector algebra.

The dot product, given by Eq. (1.11), may be generalized in two ways. The space need not be restricted to three dimensions. In n-dimensional space, Eq. (1.11) applies with the sum running from 1 to n; n may be infinity, with the sum then a convergent infinite series (see Section 5.2). The other generalization extends the concept of vector to embrace functions. The function analog of a dot or inner product is discussed in Section 9.4.

EXERCISES

1.2.1 A car is moving northward with a constant speed of 50 mph for 5 min, and then makes a 45° turn to the east and continues at 55 mph for 1 min. What is the average acceleration of the car?

1.2.2 A particle in an orbit is located at the point $\mathbf{r}$ (drawn from the origin) that terminates at and specifies the point in space (x, y, z). Find the surface swept out by the tip of $\mathbf{r}$ and draw it using graphical software if
(a) $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{a} = 0$,
(b) $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{r} = 0$.
The vector $\mathbf{a}$ is a constant (in magnitude and direction).

1.2.3 Develop a condition when two forces are parallel, with and without using their Cartesian coordinates.

1.2.4 The Newtonian equations of motion of two particles are

$$m_1 \dot{\mathbf{v}}_1 = \mathbf{F}_1^i + \mathbf{F}_1^e, \quad m_2 \dot{\mathbf{v}}_2 = \mathbf{F}_2^i + \mathbf{F}_2^e,$$

where m_i are their masses, $\mathbf{v}_i$ are their velocities, and the superscripts on the forces denote internal and external forces. What is the total force and the total external force? Write Newton's third law for the forces. Define the center of mass and derive its equation of motion. Define the relative coordinate vector of the particles and derive the relevant equation of motion. Plot a typical case at some time instant using graphical software.
Note. The resultant of all forces acting on particle 1, whose origin lies outside the system, is called external force $\mathbf{F}_1^e$; the force arising from the interaction of particle 2 with particle 1 is called the internal force $\mathbf{F}_1^i$ so that $\mathbf{F}_1^i = -\mathbf{F}_2^i$.

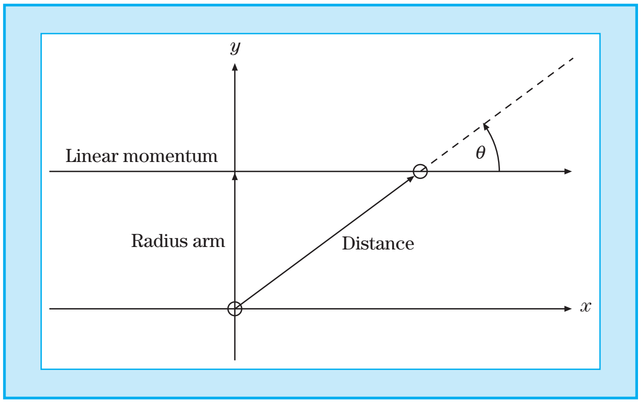
1.2.5 If $|\mathbf{A}|$, $|\mathbf{B}|$ are the magnitudes of the vectors $\mathbf{A}$, $\mathbf{B}$, show that $-|\mathbf{A}||\mathbf{B}| \le \mathbf{A} \cdot \mathbf{B} \le |\mathbf{A}||\mathbf{B}|$.

1.3 Vector or Cross Product

A second form of vector multiplication employs the sine of the included angle (denoted by θ) instead of the cosine and is called cross product. The cross product generates a vector from two vectors, in contrast with the dot product, which produces a scalar. Applications of the cross product in analytic geometry and mechanics are also discussed in this section. For instance, the orbital

Figure 1.14

Angular Momentum



angular momentum of a particle shown at the point of the distance vector in Fig. 1.14 is defined as

$$\begin{aligned}\text{Angular momentum} &= \text{radius arm} \cdot \text{ linear momentum}\\ &= \text{distance} \cdot \text{ linear momentum} \cdot \sin\theta. \end{aligned} \tag{1.26}$$

For convenience in treating problems relating to quantities such as angular momentum, torque, angular velocity, and area, we define the vector or cross product as

$$\mathbf{C} = \mathbf{A} \times \mathbf{B}, \tag{1.27}$$

with the magnitude (but not necessarily the dimensions of length)

$$C = AB \sin\theta. \tag{1.28}$$

Unlike the preceding case of the scalar product, $\mathbf{C}$ is now a vector, and we assign it a direction perpendicular to the plane of $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ form a right-handed system. If we curl the fingers of the right hand from the point of $\mathbf{A}$ to $\mathbf{B}$, then the extended thumb will point in the direction of $\mathbf{A} \times \mathbf{B}$, and these three vectors form a right-handed system. With this choice of direction, we have

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}, \qquad \text{anticommutation}. \tag{1.29}$$

In general, the cross product of two collinear vectors is zero so that

$$\hat{\mathbf{x}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{z}} = 0, \tag{1.30}$$

whereas

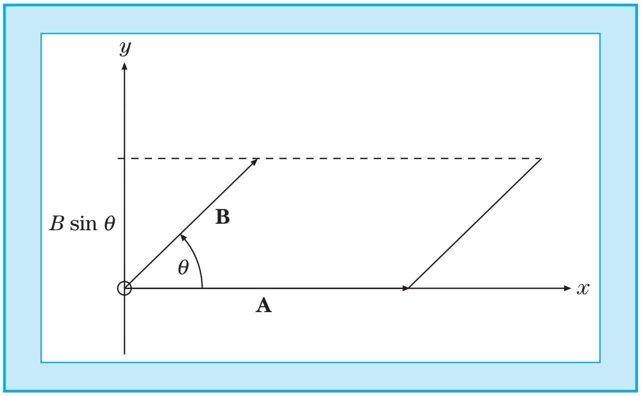
$$\begin{aligned} &\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}}, \quad \hat{\mathbf{y}} \times \hat{\mathbf{z}} = \hat{\mathbf{x}}, \quad \hat{\mathbf{z}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}},\\ &\hat{\mathbf{y}} \times \hat{\mathbf{x}} = -\hat{\mathbf{z}}, \quad \hat{\mathbf{z}} \times \hat{\mathbf{y}} = -\hat{\mathbf{x}}, \quad \hat{\mathbf{x}} \times \hat{\mathbf{z}} = -\hat{\mathbf{y}}. \end{aligned} \tag{1.31}$$

Among the examples of the cross product in mathematical physics are the relation between linear momentum $\mathbf{p}$ and angular momentum $\mathbf{L}$ (defining angular momentum),

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{1.32}$$

Figure 1.15

Parallelogram Representation of the Vector Product



and the relation between linear velocity $\mathbf{v}$ and angular velocity ω,

$$\mathbf{v} = \omega \times \mathbf{r}. \tag{1.33}$$

Vectors $\mathbf{v}$ and $\mathbf{p}$ describe properties of the particle or physical system. However, the position vector $\mathbf{r}$ is determined by the choice of the origin of the coordinates. This means that ω and $\mathbf{L}$ depend on the choice of the origin.

The familiar magnetic induction $\mathbf{B}$ occurs in the vector product force equation called Lorentz force

$$\mathbf{F}_M = q\mathbf{v} \times \mathbf{B} \qquad \text{(SI units)}, \tag{1.34}$$

where $\mathbf{v}$ is the velocity of the electric charge q, and $\mathbf{F}_M$ is the resulting magnetic force on the moving charge. The cross product has an important geometrical interpretation that we shall use in subsequent sections. In the parallelogram (Fig. 1.15) defined by $\mathbf{A}$ and $\mathbf{B}$, $B \sin\theta$ is the height if A is taken as the length of the base. Then $|\mathbf{A} \times \mathbf{B}| = AB \sin\theta$ is the **area of the parallelogram**. As a vector, $\mathbf{A} \times \mathbf{B}$ is the area of the parallelogram defined by $\mathbf{A}$ and $\mathbf{B}$, with the area vector normal to the plane of the parallelogram. This means that area (with its orientation in space) is treated as a vector.

An alternate definition of the vector product can be derived from the special case of the coordinate unit vectors in Eqs. (1.30) and (1.31) in conjunction with the linearity of the cross product in both vector arguments, in analogy with Eqs. (1.9) and (1.10) for the dot product,

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}, \tag{1.35}$$

$$(\mathbf{A} + \mathbf{B}) \times \mathbf{C} = \mathbf{A} \times \mathbf{C} + \mathbf{B} \times \mathbf{C}, \tag{1.36}$$

$$\mathbf{A} \times (y\mathbf{B}) = y\mathbf{A} \times \mathbf{B} = (y\mathbf{A}) \times \mathbf{B}, \tag{1.37}$$

where y is a number, a scalar. Using the decomposition of **A** and **B** into their Cartesian components according to Eq. (1.5), we find

$$\begin{aligned}\mathbf{A}\times\mathbf{B}\equiv\mathbf{C}=(C_x,C_y,C_z)&=(A_x\hat{\mathbf{x}}+A_y\hat{\mathbf{y}}+A_z\hat{\mathbf{z}})\times(B_x\hat{\mathbf{x}}+B_y\hat{\mathbf{y}}+B_z\hat{\mathbf{z}})\\&=(A_xB_y-A_yB_x)\hat{\mathbf{x}}\times\hat{\mathbf{y}}+(A_xB_z-A_zB_x)\hat{\mathbf{x}}\times\hat{\mathbf{z}}\\&\quad+(A_yB_z-A_zB_y)\hat{\mathbf{y}}\times\hat{\mathbf{z}},\end{aligned}$$

upon applying Eqs. (1.35) and (1.37) and substituting Eqs. (1.30) and (1.31) so that the Cartesian components of $\mathbf{A}\times\mathbf{B}$ become

$$C_x=A_yB_z-A_zB_y,\quad C_y=A_zB_x-A_xB_z,\quad C_z=A_xB_y-A_yB_x,\tag{1.38}$$

or

$$C_i=A_jB_k-A_kB_j,\quad i,j,k\ \text{ all different,}\tag{1.39}$$

and with cyclic permutation of the indices i, j, and k or $x\to y\to z\to x$ in Eq. (1.38). The vector product **C** may be represented by a determinant[4]

$$\mathbf{C}=\begin{vmatrix}\hat{\mathbf{x}}&\hat{\mathbf{y}}&\hat{\mathbf{z}}\\A_x&A_y&A_z\\B_x&B_y&B_z\end{vmatrix}\tag{1.40}$$

which, according to the expansion Eq. (3.11) of the determinant along the top row, is a shorthand form of the vector product

$$\mathbf{C}=\hat{\mathbf{x}}(A_yB_z-A_zB_y)+\hat{\mathbf{y}}(A_zB_x-A_xB_z)+\hat{\mathbf{z}}(A_xB_y-A_yB_x).$$

If Eqs. (1.27) and (1.28) are called a geometric definition of the vector product, then Eq. (1.38) is an algebraic definition.

To show the equivalence of Eqs. (1.27) and (1.28) and the component definition Eq. (1.38), let us form $\mathbf{A}\cdot\mathbf{C}$ and $\mathbf{B}\cdot\mathbf{C}$ using Eq. (1.38). We have

$$\begin{aligned}\mathbf{A}\cdot\mathbf{C}&=\mathbf{A}\cdot(\mathbf{A}\times\mathbf{B})\\&=A_x(A_yB_z-A_zB_y)+A_y(A_zB_x-A_xB_z)+A_z(A_xB_y-A_yB_x)\\&=0.\end{aligned}\tag{1.41}$$

Similarly,

$$\mathbf{B}\cdot\mathbf{C}=\mathbf{B}\cdot(\mathbf{A}\times\mathbf{B})=0.\tag{1.42}$$

Equations (1.41) and (1.42) show that **C** is perpendicular to both **A** and **B** and therefore perpendicular to the plane they determine. The positive direction is determined by considering special cases, such as the unit vectors $\hat{\mathbf{x}}\times\hat{\mathbf{y}}=\hat{\mathbf{z}}$.

[4]Determinants are discussed in detail in Section 3.1.

The magnitude is obtained from

$$
\begin{aligned}
(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{A} \times \mathbf{B}) &= A^2B^2 - (\mathbf{A} \cdot \mathbf{B})^2 \\
&= A^2B^2 - A^2B^2 \cos^2\theta \\
&= A^2B^2 \sin^2\theta,
\end{aligned} \tag{1.43}
$$

which implies Eq. (1.28). The first step in Eq. (1.43) may be verified by expanding out in component form using Eq. (1.38) for $\mathbf{A} \times \mathbf{B}$ and Eq. (1.11) for the dot product. From Eqs. (1.41)–(1.43), we see the equivalence of Eqs. (1.28) and (1.38), the two definitions of vector product.

EXAMPLE 1.3.1

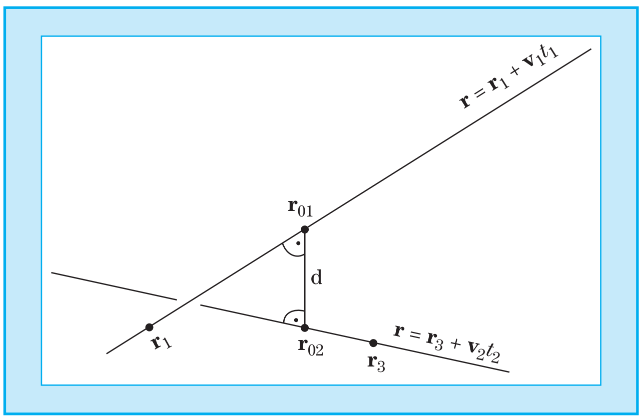
Shortest Distance between Two Rockets in Free Flight Considering Example 1.2.2 as a similar but simpler case, we remember that the shortest distance between a point and a line is measured along the normal from the line through the point. Therefore, we expect that the shortest distance between two lines is normal to both tangent vectors of the straight lines. Establishing this fact will be our first and most important step. The second step involves the projection of a vector between two points, one on each line, onto that normal to both lines. However, we also need to locate the points where the normal starts and ends. This problem we address first.

Let us take the first line from Example 1.2.2, namely $\mathbf{r} = \mathbf{r}_1 + t_1\mathbf{v}_1$ with time variable t_1 and tangent vector $\mathbf{v}_1 = \mathbf{r}_2 - \mathbf{r}_1 = (1, 2, 3)$ that goes through the points $\mathbf{r}_1 = (1, 1, 1)$ and $\mathbf{r}_2 = (2, 3, 4)$ and is shown in Fig. 1.16, along with the second line $\mathbf{r} = \mathbf{r}_3 + t_2\mathbf{v}_2$ with time variable t_2 that goes through the points $\mathbf{r}_3 = (5, 2, 1)$ and $\mathbf{r}_4 = (4, 1, 2)$, and so has the tangent vector $\mathbf{r}_4 - \mathbf{r}_3 = (-1, -1, 1) = \mathbf{v}_2$ and the parameterization

$$x = 5 - t_2, \quad y = 2 - t_2, \quad z = 1 + t_2.$$

Figure 1.16

Shortest Distance Between Two Straight Lines That Do Not Intersect



In order to find the end points $\mathbf{r}_{0k}$ of this shortest distance we minimize the distances squared $(\mathbf{r} - \mathbf{r}_{0k})^2$ to obtain the conditions

$$0 = \frac{d}{dt_1}(\mathbf{r} - \mathbf{r}_{02})^2 = \frac{d}{dt_1}(\mathbf{r}_1 - \mathbf{r}_{02} + t_1\mathbf{v}_1)^2 = 2\mathbf{v}_1 \cdot (\mathbf{r}_1 - \mathbf{r}_{02} + t_1\mathbf{v}_1),$$

$$0 = \frac{d}{dt_2}(\mathbf{r} - \mathbf{r}_{01})^2 = 2\mathbf{v}_2 \cdot (\mathbf{r}_3 - \mathbf{r}_{01} + t_2\mathbf{v}_2). \tag{1.44}$$

We can solve for $t_1 = -\mathbf{v}_1 \cdot (\mathbf{r}_1 - \mathbf{r}_{02})/\mathbf{v}_1^2$ and $t_2 = -\mathbf{v}_2 \cdot (\mathbf{r}_3 - \mathbf{r}_{01})/\mathbf{v}_2^2$ and then plug these parameter values into the line coordinates to find the points $\mathbf{r}_{0k}$ and $d = |\mathbf{r}_{01} - \mathbf{r}_{02}|$. This is straightforward but tedious. Alternatively, we can exploit the geometric meaning of Eq. (1.44) that the distance vector $\mathbf{d} = \mathbf{r}_1 + t_1\mathbf{v}_1 - \mathbf{r}_{02} = -(\mathbf{r}_3 + t_2\mathbf{v}_2 - \mathbf{r}_{01})$ is perpendicular to both tangent vectors $\mathbf{v}_k$ as shown in Fig. 1.16. Thus, the distance vector $\mathbf{d}$ is along the normal unit vector

$$\mathbf{n} = \frac{\mathbf{v}_1 \times \mathbf{v}_2}{|\mathbf{v}_1 \times \mathbf{v}_2|} = \frac{1}{\sqrt{3}\sqrt{14}}\begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ +1 & +2 & 3 \\ -1 & -1 & 1 \end{vmatrix} = \frac{1}{\sqrt{42}}(5\hat{\mathbf{x}} - 4\hat{\mathbf{y}} + \hat{\mathbf{z}}) = \frac{1}{\sqrt{42}}(5, -4, 1),$$

the **cross product** of both tangent vectors. We get the distance d by projecting the distance vector between two points $\mathbf{r}_1, \mathbf{r}_3$, one on each line, onto that normal $\mathbf{n}$—that is, $d = (\mathbf{r}_3 - \mathbf{r}_1) \cdot \mathbf{n} = \frac{1}{\sqrt{42}}(4, 1, 0) \cdot (5, -4, 1) = \frac{20-4+0}{\sqrt{42}} = \frac{16}{\sqrt{42}}$.

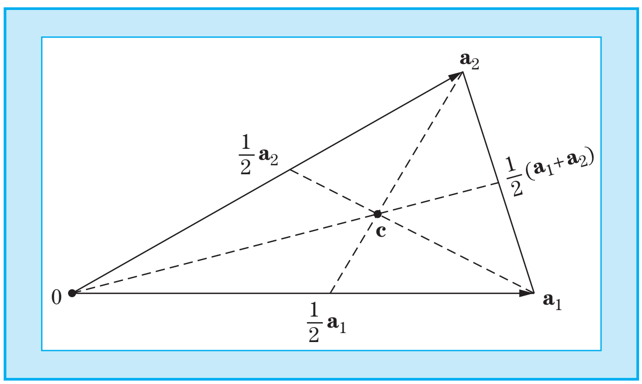
This example generalizes to the shortest distance between two orbits by examining the shortest distance beween their tangent lines. In this form, there are many applications in mechanics, space travel, and satellite orbits. ■

EXAMPLE 1.3.2

Medians of a Triangle Meet in the Center Let us consider Example 1.1.3 and Fig. 1.6 again, but now without using the 2:1 ratio of the segments from the center to the end points of each median. We put the origin of the coordinate system in one corner of the triangle, as shown in Fig. 1.17, so that the median from the origin will be given by the vector $\mathbf{m}_3 = (\mathbf{a}_1 + \mathbf{a}_2)/2$. The medians

Figure 1.17

Medians of a Triangle Meet in the Center



from the triangle corners $\mathbf{a}_1$ and $\mathbf{a}_2$ intersect at a point we call the center that is given by the vector $\mathbf{c}$ from the origin. We want to show that $\mathbf{m}_3$ and $\mathbf{c}$ are parallel (and therefore collinear), indicating that the center will also lie on the median from the origin.

From Fig. 1.17, we see that the vector $\mathbf{c}-\mathbf{a}_1$ from the corner $\mathbf{a}_1$ to the center will be parallel to $\frac{1}{2}\mathbf{a}_2 - \mathbf{a}_1$; similarly, $\mathbf{c} - \mathbf{a}_2$ will be collinear with $\frac{1}{2}\mathbf{a}_1 - \mathbf{a}_2$. We write these conditions as follows:

$$(\mathbf{c} - \mathbf{a}_1) \times \left(\frac{1}{2}\mathbf{a}_2 - \mathbf{a}_1\right) = 0, \quad (\mathbf{c} - \mathbf{a}_2) \times \left(\frac{1}{2}\mathbf{a}_1 - \mathbf{a}_2\right) = 0.$$

Expanding, and using the fact that $\mathbf{a}_1 \times \mathbf{a}_1 = 0 = \mathbf{a}_2 \times \mathbf{a}_2$, we find

$$\mathbf{c} \times \frac{1}{2}\mathbf{a}_2 - \mathbf{c} \times \mathbf{a}_1 - \frac{1}{2}(\mathbf{a}_1 \times \mathbf{a}_2) = 0, \quad \mathbf{c} \times \frac{1}{2}\mathbf{a}_1 - \mathbf{c} \times \mathbf{a}_2 - \frac{1}{2}(\mathbf{a}_2 \times \mathbf{a}_1) = 0.$$

Adding these equations, the last terms on the left-hand sides cancel, and the other terms combine to yield

$$-\frac{1}{2}\mathbf{c} \times (\mathbf{a}_1 + \mathbf{a}_2) = 0,$$

proving that $\mathbf{c}$ and $\mathbf{m}_3$ are parallel.

The center of mass (see Example 1.1.3) will be at the point $\frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2)$ and is therefore on the median from the origin. By symmetry it must be on the other medians as well, confirming both that they meet at a point and that the distance from the triangle corner to the intersection is two-thirds of the total length of the median. ■

SUMMARY

If we define a vector as an ordered triplet of numbers (or functions) as in Section 1.2, then there is no problem identifying the cross product as a vector. The cross product operation maps the two triples $\mathbf{A}$ and $\mathbf{B}$ into a third triple $\mathbf{C}$, which by definition is a vector. In Section 2.6, we shall see that the cross product also transforms like a vector.

The cross product combines two vectors antisymmetrically and involves the sine of the angle between the vectors, in contrast to their symmetric combination in the scalar product involving the cosine of their angle, and it unifies the angular momentum and velocity of mechanics with the area concept of geometry. The vector nature of the cross product is peculiar to three-dimensional space, but it can naturally be generalized to higher dimensions. The cross product occurs in many applications such as conditions for parallel forces or other vectors and the shortest distance between lines or curves more generally.

We now have two ways of multiplying vectors; a third form is discussed in Chapter 2. However, what about division by a vector? The ratio $\mathbf{B}/\mathbf{A}$ is not uniquely specified (see Exercise 3.2.21) unless $\mathbf{A}$ and $\mathbf{B}$ are also required to be parallel. Hence, division of one vector by another is meaningless.

EXERCISES

1.3.1 Prove the law of cosines starting from $\mathbf{A}^2 = (\mathbf{B} - \mathbf{C})^2$, where **A**, **B**, and **C** are the vectors collinear with the sides of a triangle. Plot the triangle and describe the theorem in words. State the analog of the law of cosines on the unit sphere (Fig. 1.18), if **A**, **B**, and **C** go from the origin to the corners of the triangle.

1.3.2 A coin with a mass of 2 g rolls on a horizontal plane at a constant velocity of 5 cm/sec. What is its kinetic energy?
Hint. Show that the radius of the coin drops out.

1.3.3 Starting with $\mathbf{C} = \mathbf{A} + \mathbf{B}$, show that $\mathbf{C} \times \mathbf{C} = 0$ leads to

$$\mathbf{A} \times \mathbf{B} = -\mathbf{B} \times \mathbf{A}.$$

1.3.4 Show that
(a) $(\mathbf{A} - \mathbf{B}) \cdot (\mathbf{A} + \mathbf{B}) = A^2 - B^2$,
(b) $(\mathbf{A} - \mathbf{B}) \times (\mathbf{A} + \mathbf{B}) = 2\mathbf{A} \times \mathbf{B}$.
The distributive laws needed here,

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$$

and

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C},$$

may be verified by expansion in Cartesian components.

1.3.5 If $\mathbf{P} = \hat{\mathbf{x}} P_x + \hat{\mathbf{y}} P_y$ and $\mathbf{Q} = \hat{\mathbf{x}} Q_x + \hat{\mathbf{y}} Q_y$ are any two nonparallel (also non-antiparallel) vectors in the xy-plane, show that $\mathbf{P} \times \mathbf{Q}$ is in the z-direction.

Figure 1.18

Spherical Triangle

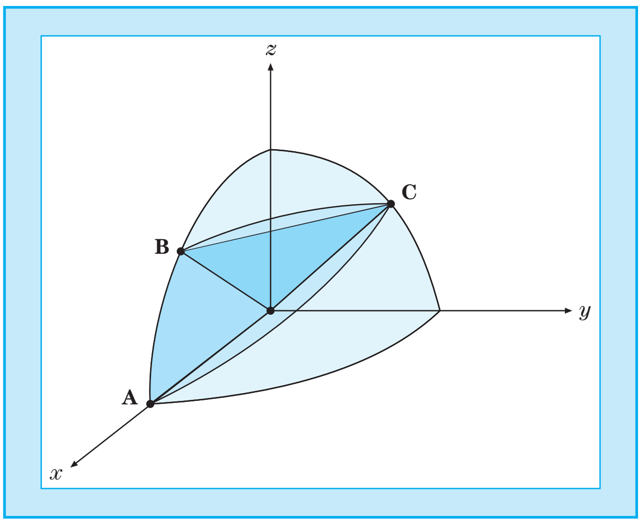
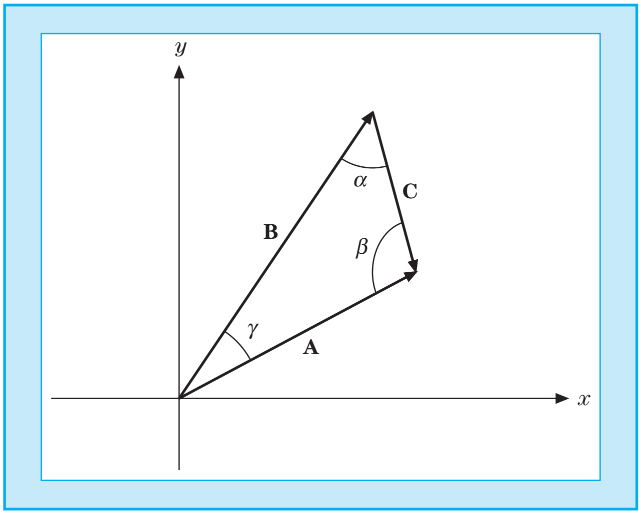


Figure 1.19

Law of Sines



1.3.6 Prove that $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{A}^2\mathbf{B}^2 - (\mathbf{A} \cdot \mathbf{B})^2$. Write the identity appropriately and describe it in geometric language. Make a plot for a typical case using graphical software.

1.3.7 Using the vectors

$$\mathbf{P} = \hat{\mathbf{x}} \cos\theta + \hat{\mathbf{y}} \sin\theta,$$

$$\mathbf{Q} = \hat{\mathbf{x}} \cos\varphi - \hat{\mathbf{y}} \sin\varphi,$$

$$\mathbf{R} = \hat{\mathbf{x}} \cos\varphi + \hat{\mathbf{y}} \sin\varphi,$$

prove the familiar trigonometric identities

$$\sin(\theta + \varphi) = \sin\theta \cos\varphi + \cos\theta \sin\varphi,$$

$$\cos(\theta + \varphi) = \cos\theta \cos\varphi - \sin\theta \sin\varphi.$$

1.3.8 If four vectors **a**, **b**, **c**, and **d** all lie in the same plane, show that

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

If graphical software is available, plot all vectors for a specific numerical case.
Hint. Consider the directions of the cross product vectors.

1.3.9 Derive the law of sines (Fig. 1.19):

$$\frac{\sin\alpha}{|\mathbf{A}|} = \frac{\sin\beta}{|\mathbf{B}|} = \frac{\sin\gamma}{|\mathbf{C}|}.$$

1.3.10 A proton of mass m, charge $+e$, and (asymptotic) momentum $p = mv$ is incident on a nucleus of charge $+Ze$ at an impact parameter b. Determine its distance of closest approach.

Hint. Consider only the Coulomb repulsion and classical mechanics, not the strong interaction and quantum mechanics.

1.3.11 Expand a vector $\mathbf{x}$ in components parallel to three linearly independent vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$.

ANS. $(\mathbf{a}\times\mathbf{b}\cdot\mathbf{c})\mathbf{x} = (\mathbf{x}\times\mathbf{b}\cdot\mathbf{c})\mathbf{a} + (\mathbf{a}\times\mathbf{x}\cdot\mathbf{c})\mathbf{b} + (\mathbf{a}\times\mathbf{b}\cdot\mathbf{x})\mathbf{c}$.

1.3.12 Let $\mathbf{F}$ be a force vector drawn from the coordinate vector $\mathbf{r}$. If $\mathbf{r}'$ goes from the origin to another point on the line through the point of $\mathbf{r}$ with tangent vector given by the force, show that the torque $\mathbf{r}'\times\mathbf{F} = \mathbf{r}\times\mathbf{F}$—that is, the torque about the origin due to the force stays the same.

1.3.13 A car drives in a horizontal circular track of radius R (to its center of mass). Find the speed at which it will overturn, if h is the height of its center of mass and d the distance between its left and right wheels.
Hint. Find the speed at which there is no vertical force on the inner wheels. (The mass of the car drops out.)

1.3.14 A force $\mathbf{F} = (3, 2, 4)$ acts at the point $(1, 4, 2)$. Find the torque about the origin. Plot the vectors using graphical software.

1.3.15 Generalize the cross product to n-dimensional space $(n = 2, 4, 5, \ldots)$ and give a geometrical interpretation of your construction. Give realistic examples in four- and higher dimensional spaces.

1.3.16 A jet plane flies due south over the north pole with a constant speed of 500 mph. Determine the angle between a plumb line hanging freely in the plane and the radius vector from the center of the earth to the plane above the north pole.
Hint. Assume that the earth's angular velocity is 2π radians in 24 hr, which is a good approximation. Why?

1.4 Triple Scalar Product and Triple Vector Product

Triple Scalar Product

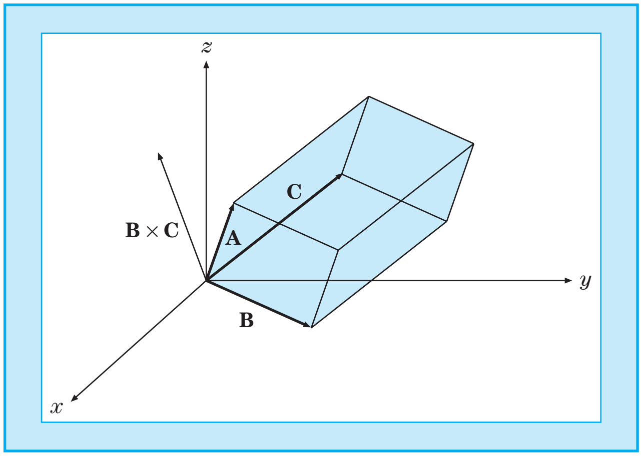
Sections 1.2 and 1.3 discussed the two types of vector multiplication. However, there are combinations of three vectors, $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$ and $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$, that occur with sufficient frequency in mechanics, electrodynamics, and analytic geometry to deserve further attention. The combination

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) \tag{1.45}$$

is known as the **triple scalar product**. $\mathbf{B}\times\mathbf{C}$ yields a vector that, dotted into $\mathbf{A}$, gives a scalar. We note that $(\mathbf{A}\cdot\mathbf{B})\times\mathbf{C}$ represents a scalar crossed into a vector, an operation that is not defined. Hence, if we agree to exclude this undefined interpretation, the parentheses may be omitted and the triple scalar product written as $\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}$.

Figure 1.20

Parallelepiped Representation of Triple Scalar Product



Using Eq. (1.38) for the cross product and Eq. (1.11) for the dot product, we obtain

$$\begin{aligned}\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} &= A_x(B_yC_z - B_zC_y) + A_y(B_zC_x - B_xC_z) + A_z(B_xC_y - B_yC_x)\\ &= \mathbf{B}\cdot\mathbf{C}\times\mathbf{A} = \mathbf{C}\cdot\mathbf{A}\times\mathbf{B} = -\mathbf{A}\cdot\mathbf{C}\times\mathbf{B}\\ &= -\mathbf{C}\cdot\mathbf{B}\times\mathbf{A} = -\mathbf{B}\cdot\mathbf{A}\times\mathbf{C}. \end{aligned} \tag{1.46}$$

The high degree of symmetry present in the component expansion should be noted. Every term contains the factors A_i, B_j, and C_k. If i, j, and k are in cyclic order (x, y, z), the sign is positive. If the order is anticyclic, the sign is negative. Furthermore, the dot and the cross may be interchanged:

$$\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} = \mathbf{A}\times\mathbf{B}\cdot\mathbf{C}. \tag{1.47}$$

A convenient representation of the component expansion of Eq. (1.46) is provided by the determinant

$$\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} = \begin{vmatrix} A_x & A_y & A_z\\ B_x & B_y & B_z\\ C_x & C_y & C_z \end{vmatrix}, \tag{1.48}$$

which follows from Eq. (1.38) by dotting $\mathbf{B}\times\mathbf{C}$ into $\mathbf{A}$. The rules for interchanging rows and columns of a determinant[5] provide an immediate verification of the permutations listed in Eq. (1.46), whereas the symmetry of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ in the determinant form suggests the relation given in Eq. (1.46). The triple products discussed in Section 1.3, which showed that $\mathbf{A}\times\mathbf{B}$ was perpendicular to both $\mathbf{A}$ and $\mathbf{B}$, were special cases of the general result [Eq. (1.46)].

The triple scalar product has a direct **geometrical interpretation** in which the three vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are interpreted as defining a **parallelepiped** (Fig. 1.20):

$$|\mathbf{B}\times\mathbf{C}| = BC\sin\theta = \text{ area of parallelogram base.} \tag{1.49}$$

[5]See Section 3.1 for a detailed discussion of the properties of determinants.

The direction, of course, is normal to the base. Dotting $\mathbf{A}$ into this means multiplying the base area by the projection of $\mathbf{A}$ onto the normal, or base times height. Therefore,

$$\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} = \text{ volume of parallelepiped defined by } \mathbf{A}, \mathbf{B}, \text{ and } \mathbf{C}. \tag{1.50}$$

Note that $\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}$ may sometimes be negative. This is not a problem, and its proper interpretation is provided in Chapter 2.

EXAMPLE 1.4.1

A Parallelepiped For

$$\mathbf{A} = \hat{\mathbf{x}} + 2\hat{\mathbf{y}} - \hat{\mathbf{z}}, \quad \mathbf{B} = \hat{\mathbf{y}} + \hat{\mathbf{z}}, \quad \mathbf{C} = \hat{\mathbf{x}} - \hat{\mathbf{y}},$$

$$\mathbf{A}\cdot\mathbf{B}\times\mathbf{C} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{vmatrix} = 4.$$

This is the volume of the parallelepiped defined by $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. ■

Recall that we already encountered a triple scalar product, namely the distance $d \sim (\mathbf{r}_3 - \mathbf{r}_1)\cdot(\mathbf{v}_1 \times \mathbf{v}_2)$ between two straight lines in Example 1.3.1.

Triple Vector Product

The second triple product of interest is $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$, which is a vector. Here, the parentheses must be retained, as is seen from a special case $(\hat{\mathbf{x}}\times\hat{\mathbf{x}})\times\hat{\mathbf{y}} = 0$, whereas $\hat{\mathbf{x}}\times(\hat{\mathbf{x}}\times\hat{\mathbf{y}}) = \hat{\mathbf{x}}\times\hat{\mathbf{z}} = -\hat{\mathbf{y}}$. Let us start with an example that illustrates a key property of the triple product.

EXAMPLE 1.4.2

A Triple Vector Product By using the three vectors given in Example 1.4.1, we obtain

$$\mathbf{B}\times\mathbf{C} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{vmatrix} = \hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}$$

and

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ 1 & 2 & -1 \\ 1 & 1 & -1 \end{vmatrix} = -\hat{\mathbf{x}} - \hat{\mathbf{z}} = -(\hat{\mathbf{y}} + \hat{\mathbf{z}}) - (\hat{\mathbf{x}} - \hat{\mathbf{y}}).$$ ■

By rewriting the result in the last line as a linear combination of $\mathbf{B}$ and $\mathbf{C}$, we notice that, taking a geometric approach, the triple product vector is perpendicular to $\mathbf{A}$ and to $\mathbf{B}\times\mathbf{C}$. The plane spanned by $\mathbf{B}$ and $\mathbf{C}$ is perpendicular to $\mathbf{B}\times\mathbf{C}$, so the triple product lies in this plane (Fig. 1.21):

$$\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = u\mathbf{B} + v\mathbf{C}, \tag{1.51}$$

where u and v are numbers. Multiplying Eq. (1.51) by $\mathbf{A}$ gives zero for the left-hand side so that $u\mathbf{A}\cdot\mathbf{B} + v\mathbf{A}\cdot\mathbf{C} = 0$. Hence, $u = w\mathbf{A}\cdot\mathbf{C}$ and $v = -w\mathbf{A}\cdot\mathbf{B}$ for

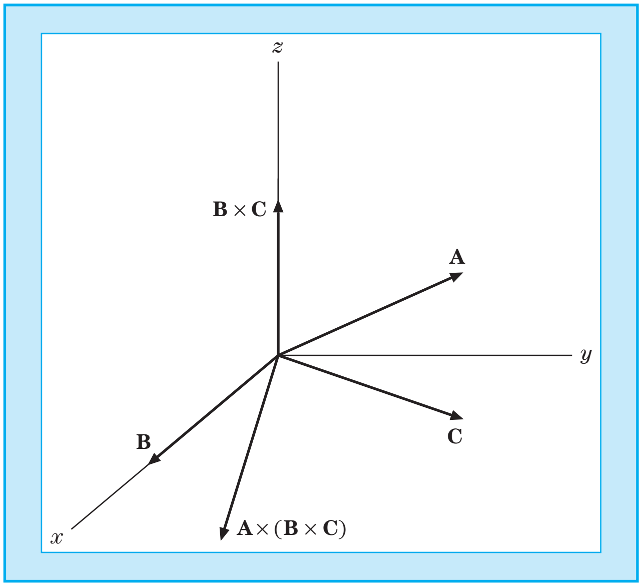


Figure 1.21

B and C Are in the *xy*-Plane. B × C Is Perpendicular to the *xy*-Plane and Is Shown Here Along the *z*-Axis. Then A × (B × C) Is Perpendicular to the *z*-Axis and Therefore Is Back in the *xy*-Plane

a suitable number w. Substituting these values into Eq. (1.50) gives

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = w[\mathbf{B}\,(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}\,(\mathbf{A} \cdot \mathbf{B})]. \tag{1.52}$$

Equation (1.51), with $w = 1$, which we now prove, is known as the ***BAC–CAB* rule**. Since Eq. (1.52) is linear in A, B, and C, w is independent of these magnitudes. That is, we only need to show that $w = 1$ for unit vectors $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$. Let us denote $\hat{\mathbf{B}} \cdot \hat{\mathbf{C}} = \cos\alpha$, $\hat{\mathbf{C}} \cdot \hat{\mathbf{A}} = \cos\beta$, $\hat{\mathbf{A}} \cdot \hat{\mathbf{B}} = \cos\gamma$, and square Eq. (1.52) to obtain

$$\begin{aligned}
[\hat{\mathbf{A}} \times (\hat{\mathbf{B}} \times \hat{\mathbf{C}})]^2 &= \hat{\mathbf{A}}^2(\hat{\mathbf{B}} \times \hat{\mathbf{C}})^2 - [\hat{\mathbf{A}} \cdot (\hat{\mathbf{B}} \times \hat{\mathbf{C}})]^2 \\
&= 1 - \cos^2\alpha - [\hat{\mathbf{A}} \cdot (\hat{\mathbf{B}} \times \hat{\mathbf{C}})]^2 \\
&= w^2[(\hat{\mathbf{A}} \cdot \hat{\mathbf{C}})^2 + (\hat{\mathbf{A}} \cdot \hat{\mathbf{B}})^2 - 2\hat{\mathbf{A}} \cdot \hat{\mathbf{B}}\,\hat{\mathbf{A}} \cdot \hat{\mathbf{C}}\,\hat{\mathbf{B}} \cdot \hat{\mathbf{C}}] \\
&= w^2(\cos^2\beta + \cos^2\gamma - 2\cos\alpha\cos\beta\cos\gamma),
\end{aligned} \tag{1.53}$$

using $(\hat{\mathbf{A}} \times \hat{\mathbf{B}})^2 = \hat{\mathbf{A}}^2\hat{\mathbf{B}}^2 - (\hat{\mathbf{A}} \cdot \hat{\mathbf{B}})^2$ repeatedly. Consequently, the (squared) volume spanned by $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$ that occurs in Eq. (1.53) can be written as

$$[\hat{\mathbf{A}} \cdot (\hat{\mathbf{B}} \times \hat{\mathbf{C}})]^2 = 1 - \cos^2\alpha - w^2(\cos^2\beta + \cos^2\gamma - 2\cos\alpha\cos\beta\cos\gamma).$$

Here, we must have $w^2 = 1$ because this volume is symmetric in α, β, γ. That is, $w = \pm 1$ and is independent of $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{C}}$. Again using the special case $\hat{\mathbf{x}} \times (\hat{\mathbf{x}} \times \hat{\mathbf{y}}) = -\hat{\mathbf{y}}$ in Eq. (1.51) finally gives $w = 1$.

An alternate and easier algebraic derivation using the Levi-Civita ε_{ijk} of Chapter 2 is the topic of Exercise 2.9.8.

Note that because vectors are independent of the coordinates, a vector equation is independent of the particular coordinate system. The coordinate system only determines the components. If the vector equation can be established in Cartesian coordinates, it is established and valid in any of the coordinate systems, as will be shown in Chapter 2. Thus, Eq. (1.52) may be verified by a direct though not very elegant method of expanding into Cartesian components (see Exercise 1.4.1).

Other, more complicated, products may be simplified by using these forms of the triple scalar and triple vector products.

SUMMARY

We have developed the geometric meaning of the triple scalar product as a volume spanned by three vectors and exhibited its component form that is directly related to a determinant whose entries are the Cartesian components of the vectors.

The main property of the triple vector product is its decomposition expressed in the *BAC–CAB* rule. It plays a role in electrodynamics, a vector field theory in which cross products abound.

EXERCISES

1.4.1 Verify the expansion of the triple vector product

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$$

by direct expansion in Cartesian coordinates.

1.4.2 Show that the first step in Eq. (1.43),

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{A} \times \mathbf{B}) = A^2B^2 - (\mathbf{A} \cdot \mathbf{B})^2,$$

is consistent with the *BAC–CAB* rule for a triple vector product.

1.4.3 The orbital angular momentum $\mathbf{L}$ of a particle is given by $\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\mathbf{r} \times \mathbf{v}$, where $\mathbf{p}$ is the linear momentum. With linear and angular velocity related by $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$, show that

$$\mathbf{L} = mr^2[\boldsymbol{\omega} - \hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \boldsymbol{\omega})],$$

where $\hat{\mathbf{r}}$ is a unit vector in the $\mathbf{r}$ direction. For $\mathbf{r} \cdot \boldsymbol{\omega} = 0$, this reduces to $\mathbf{L} = I\boldsymbol{\omega}$, with the moment of inertia I given by mr^2.

1.4.4 The kinetic energy of a single particle is given by $T = \frac{1}{2}mv^2$. For rotational motion this becomes $\frac{1}{2}m(\boldsymbol{\omega} \times \mathbf{r})^2$. Show that

$$T = \frac{1}{2}m[r^2\omega^2 - (\mathbf{r} \cdot \boldsymbol{\omega})^2].$$

For $\mathbf{r} \cdot \boldsymbol{\omega} = 0$, this reduces to $T = \frac{1}{2}I\omega^2$, with the moment of inertia I given by mr^2.

1.4.5 Show that

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0.^{6}$$

1.4.6 A vector $\mathbf{A}$ is decomposed into a radial vector $\mathbf{A}_r$ and a tangential vector $\mathbf{A}_t$. If $\hat{\mathbf{r}}$ is a unit vector in the radial direction, show that
(a) $\mathbf{A}_r = \hat{\mathbf{r}}(\mathbf{A} \cdot \hat{\mathbf{r}})$ and
(b) $\mathbf{A}_t = -\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \mathbf{A})$.

1.4.7 Prove that a necessary and sufficient condition for the three (nonvanishing) vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ to be coplanar is the vanishing of the triple scalar product

$$\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = 0.$$

1.4.8 Vector $\mathbf{D}$ is a linear combination of three noncoplanar (and nonorthogonal) vectors:

$$\mathbf{D} = a\mathbf{A} + b\mathbf{B} + c\mathbf{C}.$$

Show that the coefficients are given by a ratio of triple scalar products,

$$a = \frac{\mathbf{D} \cdot \mathbf{B} \times \mathbf{C}}{\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}}, \qquad \text{and so on.}$$

If symbolic software is available, evaluate numerically the triple scalar products and coefficients for a typical case.

1.4.9 Show that

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}).$$

1.4.10 Show that

$$(\mathbf{A} \times \mathbf{B}) \times (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{B} \times \mathbf{D})\mathbf{C} - (\mathbf{A} \cdot \mathbf{B} \times \mathbf{C})\mathbf{D}.$$

1.4.11 Given

$$\mathbf{a}' = \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \qquad \mathbf{b}' = \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \qquad \mathbf{c}' = \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}},$$

and $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c} \neq 0$, show that
(a) $\mathbf{x}' \cdot \mathbf{y}' = 0$ (if $\mathbf{x} \neq \mathbf{y}$) and $\mathbf{x}' \cdot \mathbf{y}' = 1$ (if $\mathbf{x} = \mathbf{y}$), for ($\mathbf{x}, \mathbf{y} = \mathbf{a}, \mathbf{b}, \mathbf{c}$),
(b) $\mathbf{a}' \cdot \mathbf{b}' \times \mathbf{c}' = (\mathbf{a} \cdot \mathbf{b} \times \mathbf{c})^{-1}$,
(c) $\mathbf{a} = \frac{\mathbf{b}' \times \mathbf{c}'}{\mathbf{a}' \cdot \mathbf{b}' \times \mathbf{c}'}$.

1.4.12 If $\mathbf{x}' \cdot \mathbf{y}' = 0$ if $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{x}' \cdot \mathbf{y}' = 1$ if $\mathbf{x} = \mathbf{y}$, for ($\mathbf{x}, \mathbf{y} = \mathbf{a}, \mathbf{b}, \mathbf{c}$), prove that

$$\mathbf{a}' = \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}.$$

(This is the converse of Problem 1.4.11.)

[6] This is Jacobi's identity for vector products.

1.4.13 Show that any vector $\mathbf{V}$ may be expressed in terms of the reciprocal vectors $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$ (of Problem 1.4.11) by

$$\mathbf{V} = (\mathbf{V}\cdot\mathbf{a})\,\mathbf{a}' + (\mathbf{V}\cdot\mathbf{b})\,\mathbf{b}' + (\mathbf{V}\cdot\mathbf{c})\,\mathbf{c}'.$$

1.4.14 An electric charge q_1 moving with velocity $\mathbf{v}_1$ produces a magnetic induction $\mathbf{B}$ given by

$$\mathbf{B} = \frac{\mu_0}{4\pi} q_1 \frac{\mathbf{v}_1 \times \hat{\mathbf{r}}}{r^2} \qquad \text{(SI units)},$$

where $\hat{\mathbf{r}}$ points from q_1 to the point at which $\mathbf{B}$ is measured (Biot and Savart's law).

(a) Show that the magnetic force on a second charge q_2, velocity $\mathbf{v}_2$, is given by the triple vector product

$$\mathbf{F}_2 = \frac{\mu_0}{4\pi}\frac{q_1 q_2}{r^2}\mathbf{v}_2 \times (\mathbf{v}_1 \times \hat{\mathbf{r}}).$$

(b) Write out the corresponding magnetic force $\mathbf{F}_1$ that q_2 exerts on q_1. Define your unit radial vector. How do $\mathbf{F}_1$ and $\mathbf{F}_2$ compare?

(c) Calculate $\mathbf{F}_1$ and $\mathbf{F}_2$ for the case of q_1 and q_2 moving along parallel trajectories side by side.

ANS.

(b) $\mathbf{F}_1 = -\dfrac{\mu_0}{4\pi}\dfrac{q_1 q_2}{r^2}\mathbf{v}_1 \times (\mathbf{v}_2 \times \hat{\mathbf{r}})$.

(c) $\mathbf{F}_1 = \dfrac{\mu_0}{4\pi}\dfrac{q_1 q_2}{r^2}v^2\hat{\mathbf{r}} = -\mathbf{F}_2$.

1.5 Gradient, ∇

Partial Derivatives

In this section, we deal with derivatives of functions of several variables that will lead us to the concept of directional derivative or gradient operator, which is of central importance in mechanics, electrodynamics, and engineering.

We can view a function $z = \varphi(x, y)$ of two variables geometrically as a surface over the xy-plane in three-dimensional Euclidean space because for each point (x, y) we find the z value from φ. For a fixed value y then, $z = \varphi(x, y) \equiv f(x)$ is a function of x only, viz. a curve on the intersection of the surface with the xz-plane going through y. The slope of this curve,

$$\frac{df}{dx} \equiv \frac{\partial \varphi(x, y)}{\partial x} = \lim_{h\to 0}\frac{\varphi(x+h, y) - \varphi(x, y)}{h}, \tag{1.54}$$

is the **partial derivative** of φ with respect to x defined with the understanding that the other variable y is held fixed. It is useful for drawing tangents and locating a maximum or minimum on the curve where the slope is zero. The partial derivative $\partial\varphi/\partial y$ is similarly defined holding x fixed (i.e., it is the slope of the surface in the y-direction), and so on for the higher partial derivatives.

EXAMPLE 1.5.1

Error Estimate Error estimates usually involve many partial derivatives. Let us calculate the moment of inertia of a rectangular slab of metal of length $a = 10 \pm 1$ cm, width $b = 15 \pm 2$ cm, and height $c = 5 \pm 1$ cm about an axis through its center of gravity and perpendicular to the area ab and estimate the error. The uniform density is $\rho = 5 \pm 0.1$ g/cm^3. The moment of inertia is given by

$$\begin{aligned} I &= \rho \int (x^2 + y^2)\, d\tau = \rho c \left(\int_{-a/2}^{a/2} x^2\, dx \int_{-b/2}^{b/2} dy + \int_{-a/2}^{a/2} dx \int_{-b/2}^{b/2} y^2\, dy \right) \\ &= \frac{\rho c}{3} 2 \left(b \left(\frac{a}{2}\right)^3 + a \left(\frac{b}{2}\right)^3 \right) = \frac{\rho abc}{12}(a^2 + b^2) \qquad (1.55) \\ &= \frac{1}{2} 5^6 (4 + 9) \text{ g cm}^2 = 10.15625 \times 10^{-3} \text{ kg m}^2, \end{aligned}$$

where $d\tau = c dx dy$ has been used.

The corresponding error in I derives from the errors in all variables, each being weighted by the corresponding partial derivative,

$$(\Delta I)^2 = \left(\frac{\partial I}{\partial \rho}\right)^2 (\Delta \rho)^2 + \left(\frac{\partial I}{\partial a}\right)^2 (\Delta a)^2 + \left(\frac{\partial I}{\partial b}\right)^2 (\Delta b)^2 + \left(\frac{\partial I}{\partial c}\right)^2 (\Delta c)^2,$$

where Δx is the error in the variable x, that is, $\Delta a = 1$ cm, etc. The partial derivatives

$$\begin{aligned} \frac{\partial I}{\partial \rho} &= \frac{abc}{12}(a^2 + b^2), & \frac{\partial I}{\partial a} &= \frac{\rho bc}{12}(3a^2 + b^2), \\ \frac{\partial I}{\partial b} &= \frac{\rho ac}{12}(a^2 + 3b^2), & \frac{\partial I}{\partial c} &= \frac{\rho ab}{12}(a^2 + b^2) \end{aligned} \qquad (1.56)$$

are obtained from Eq. (1.55). Substituting the numerical values of all parameters, we get

$$\begin{aligned} \frac{\partial I}{\partial \rho}\Delta \rho &= 0.203125 \times 10^{-3} \text{ kg m}^2, & \frac{\partial I}{\partial a}\Delta a &= 1.640625 \times 10^{-3} \text{ kg m}^2, \\ \frac{\partial I}{\partial b}\Delta b &= 3.2291667 \times 10^{-3} \text{ kg m}^2, & \frac{\partial I}{\partial c}\Delta c &= 2.03125 \times 10^{-3} \text{ kg m}^2. \end{aligned}$$

Squaring and summing up, we find $\Delta I = 4.1577 \times 10^{-3}$ kg m^2. This error of more than 40% of the value I is much higher than the largest error $\Delta c \sim 20\%$ of the variables on which I depends and shows how errors in several variables can add up. Thus, all decimals except the first one can be dropped safely. ■

EXAMPLE 1.5.2

Partials of a Plane Let us now take a plane $F(\mathbf{r}) = \mathbf{n} \cdot \mathbf{r} - d = 0$ that cuts the coordinate axes at $x = 1,\ y = 2,\ z = 3$ so that $n_x = d,\ 2n_y = d,\ 3n_z = d$. Because the normal $\mathbf{n}^2 = 1$, we have the constraint $d^2(1 + \frac{1}{4} + \frac{1}{9}) = 1$ so that

$d = 6/7$. Hence, the partial derivatives

$$\frac{\partial F}{\partial x} = n_x = 6/7, \quad \frac{\partial F}{\partial y} = n_y = 3/7, \quad \frac{\partial F}{\partial z} = n_z = 2/7$$

are the components of a vector $\mathbf{n}$ (the normal) for our plane $6x + 3y + 2z = 6$. This suggests the partial derivatives of any function F are a vector. ■

To provide more motivation for the vector nature of the partial derivatives, we now introduce the **total variation of a function** $F(x, y)$,

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy. \tag{1.57}$$

It consists of independent variations in the x- and y-directions. We write dF as a sum of two increments, one purely in the x- and the other in the y-direction,

$$dF(x, y) \equiv F(x + dx, y + dy) - F(x, y) = [F(x + dx, y + dy) - F(x, y + dy)] + [F(x, y + dy) - F(x, y)] = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy,$$

by adding and subtracting $F(x, y + dy)$. The mean value theorem (i.e., continuity of F) tells us that here $\partial F/\partial x$, $\partial F/\partial y$ are evaluated at some point ξ, η between x and $x + dx$, y and $y + dy$, respectively. As $dx \to 0$ and $dy \to 0$, $\xi \to x$ and $\eta \to y$. This result generalizes to three and higher dimensions. For example, for a function φ of three variables,

$$\begin{aligned} d\varphi(x, y, z) \equiv {} & [\varphi(x + dx, y + dy, z + dz) - \varphi(x, y + dy, z + dz)] \\ & + [\varphi(x, y + dy, z + dz) - \varphi(x, y, z + dz)] \\ & + [\varphi(x, y, z + dz) - \varphi(x, y, z)] \\ = {} & \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz. \end{aligned} \tag{1.58}$$

Note that if F is a scalar function, dF is also a scalar and the form of Eq. (1.57) suggests an interpretation as a scalar product of the coordinate displacement vector $d\mathbf{r} = (dx, dy)$ with the partial derivatives of F; the same holds for $d\varphi$ in three dimensions. These observations pave the way for the gradient in the next section.

As an application of the total variation, we consider the slope of an implicitly defined curve $F(x, y) = 0$, a general theorem that we postponed in Section 1.3. Because also $dF = 0$ on the curve, we find the slope of the curve

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \tag{1.59}$$

from Eq. (1.57). Compare this result with $\dot{y}/\dot{x}$ for the slope of a curve defined in terms of two functions $x(t)$, $y(t)$ of time t in Section 1.2.

Often, we are confronted with more difficult problems of finding a slope given some constraint. A case in point is the next example.

EXAMPLE 1.5.3

Extremum under a Constraint Find the points of shortest (or longest) distance from the origin on the curve $G(x, y) \equiv x^2 + xy + y^2 - 1 = 0$.

From analytic geometry we know that the points on such a quadratic form with center at the origin (there are no terms linear in x or y that would shift the center) represent a conic section. But which kind? To answer this question, note that $(x + y)^2 = x^2 + 2xy + y^2 \geq 0$ implies that $xy \geq -(x^2 + y^2)/2$ so that the quadratic form is positive definite, that is, $G(x, y) + 1 \geq 0$, and G must therefore be an ellipse (Fig. 1.22). Hence, our problem is equivalent to finding the orientation of its principal axes (see Section 3.5 for the alternative matrix diagonalization method). The square of the distance from the origin is defined by the function $F(x, y) = x^2 + y^2$, subject to the constraint that the point (x, y) lie on the ellipse defined by G. The constraint G defines $y = y(x)$. Therefore, we look for the solutions of

$$0 = \frac{dF(x, y(x))}{dx} = 2x + 2y\frac{dy}{dx}.$$

Differentiating G, we find

$$y' = -\frac{2x + y}{2y + x} \quad \text{from } 2x + y + xy' + 2yy' = 0,$$

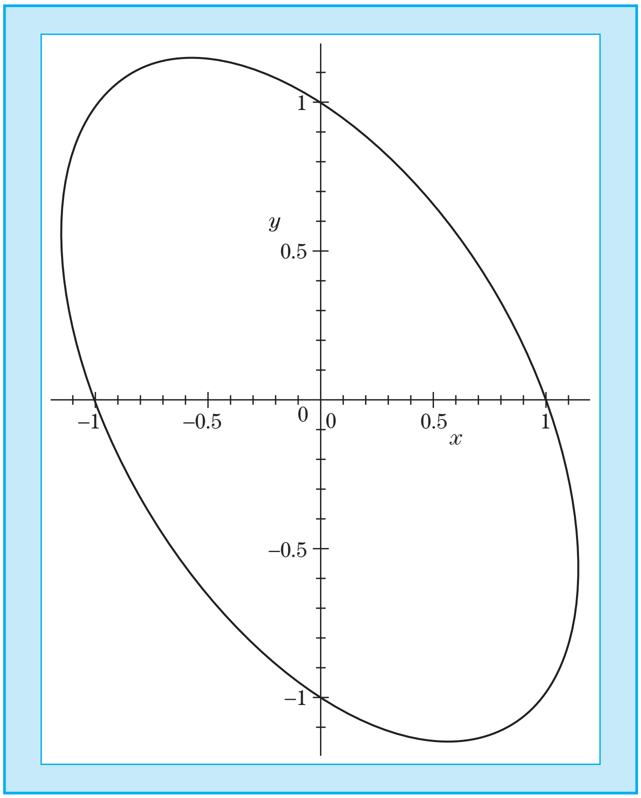


Figure 1.22

The Ellipse $x^2 + xy + y^2 = 1$

which we substitute into our min/max condition $dF/dx = 0$. This yields

$$x(2y + x) = y(2x + y), \quad \text{or} \quad y = \pm x.$$

Substituting $x = y$ into G gives the solutions $x = \pm 1/\sqrt{3}$, while $x = -y$ yields the points $x = \pm 1$ on the ellipse. Substituting $x = 1$ into G gives $y = 0$ and $y = -1$, while $x = -1$ yields $y = 0$ and $y = 1$. Although the points $(x, y) = (1, 0), (-1, 0)$ lie on the ellipse, their distance $(= 1)$ from the origin is neither shortest nor longest. However, the points $(1, -1)$, $(-1, 1)$ have the longest distance $(= \sqrt{2})$ and define the line $x + y = 0$ through the origin (at 135°) as a principal axis. The points $(1/\sqrt{3}, 1/\sqrt{3})$, $(-1/\sqrt{3}, -1/\sqrt{3})$ define the line at 45° through the origin as the second principal axis that is orthogonal to the first axis.

It is also instructive to apply the slope formula (1.59) at the intersection points of the principal axes and the ellipse, that is, $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$, $(1, -1)$. The partial derivatives there are given by $G_x \equiv \frac{\partial G}{\partial x} = 2x + y = \frac{3}{\sqrt{3}} = \sqrt{3}$ and $2 - 1 = 1$, respectively, $G_y \equiv \frac{\partial G}{\partial y} = 2y + x = \frac{3}{\sqrt{3}} = \sqrt{3}$ and $-2 + 1 = -1$, so that the slopes become $-G_x/G_y = -\frac{\sqrt{3}}{\sqrt{3}} = -1$ equal to that of the principal axis $x + y = 0$, and $-1/(-1) = 1$ equal to that of the other principal axis $x - y = 0$. ■

Although this problem was straightforward to solve, there is the more elegant **Lagrange multiplier method** for finding a maximum or minimum of a function $F(x, y)$ subject to a constraint $G(x, y) = 0$.

Introducing a Lagrange multiplier λ helps us avoid the direct (and often messy algebraic) solution for x and y as follows. Because we look for the solution of

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0, \qquad dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial y}dy = 0, \tag{1.60}$$

we can solve for the slope dy/dx from one equation and substitute that solution into the other one. Equivalently, we use the function $F + \lambda G$ of three variables x, y, λ, and solve

$$d(F + \lambda G) = \left(\frac{\partial F}{\partial x} + \lambda\frac{\partial G}{\partial x}\right)dx + \left(\frac{\partial F}{\partial y} + \lambda\frac{\partial G}{\partial y}\right)dy + \frac{\partial(F + \lambda G)}{\partial \lambda}d\lambda = 0$$

by choosing λ to satisfy $\frac{\partial F}{\partial y} + \lambda\frac{\partial G}{\partial y} = 0$, for example, and then eliminating the last term by the constraint $G = 0$ (note that F does not depend on λ) so that $\frac{\partial F}{\partial x} + \lambda\frac{\partial G}{\partial x} = 0$ follows. Including the constraint, we now have three equations for three unknowns x, y, λ, where the slope λ is not usually needed.

EXAMPLE 1.5.4

Lagrange Multiplier Method Let us illustrate the method by solving Example 1.5.3 again, this time using the Lagrange multiplier method. The x

and y partial derivative equations of the Lagrange multiplier method are given by

$$\frac{\partial F}{\partial x} + \lambda \frac{\partial G}{\partial x} \equiv 2x + \lambda(2x + y) = 0,$$
$$\frac{\partial F}{\partial y} + \lambda \frac{\partial G}{\partial y} \equiv 2y + \lambda(2y + x) = 0.$$

We find for the ratio $\xi \equiv y/x = -2(\lambda + 1)/\lambda$ and $\xi = -\lambda/2(1 + \lambda)$, that is, $\xi = 1/\xi$, or $\xi = \pm 1$, so that the principal axes are along the lines $x + y = 0$ and $x - y = 0$ through the origin. Substituting these solutions into the conic section G yields $x = 1/\sqrt{3} = y$ and $x = 1 = -y$, respectively. Contrast this simple, yet sophisticated approach with our previous lengthy solution. ■

Biographical Data

Lagrange, Joseph Louis comte de. Lagrange, a French mathematician and physicist, was born in Torino to a wealthy French-Italian family in 1736 and died in Paris in 1813. While in school, an essay on calculus by the English astronomer Halley sparked his enthusiasm for mathematics. In 1755, he became a professor in Torino. In 1766, he succeeded L. Euler (who moved to St. Petersburg to serve Catherine the Great) as director of the mathematics–physics section of the Prussian Academy of Sciences in Berlin. In 1786, he left Berlin for Paris after the death of king Frederick the Great. He was the founder of analytical mechanics. His famous book, *Mécanique Analytique*, contains not a single geometric figure.

Gradient as a Vector Operator

The **total variation** $dF(x, y)$ in Eq. (1.57) looks like a scalar product of the incremental length vector $d\mathbf{r} = (dx, dy)$ with a vector $(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y})$ of partial derivatives in two dimensions, that is, the change of F depends on the direction in which we go. For example, F could be a wave function in quantum mechanics or describe a temperature distribution in space. When we are at the peak value, the height will fall off at different rates in different directions, just like a ski slope: One side might be for beginners, whereas another has only expert runs. When we generalize this to a function $\varphi(x, y, z)$ of three variables, we obtain Eq. (1.58),

$$d\varphi = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz, \tag{1.61}$$

for the total change in the scalar function φ consisting of additive contributions of each coordinate change corresponding to a change in position

$$d\mathbf{r} = \hat{\mathbf{x}}\, dx + \hat{\mathbf{y}}\, dy + \hat{\mathbf{z}}\, dz, \tag{1.62}$$

the increment of length $d\mathbf{r}$. Algebraically, $d\varphi$ in Eq. (1.58) is a scalar product of the change in position $d\mathbf{r}$ and the **directional** change of φ. Now we are ready to recognize the three-dimensional partial derivative as a vector, which leads us to the concept of gradient. A convenient notation is

$$\nabla \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}, \tag{1.63}$$

$$\nabla \varphi = \hat{\mathbf{x}}\frac{\partial \varphi}{\partial x} + \hat{\mathbf{y}}\frac{\partial \varphi}{\partial y} + \hat{\mathbf{z}}\frac{\partial \varphi}{\partial z}, \tag{1.64}$$

so that ∇ (del) is a vector that differentiates (scalar) functions. As such, it is a **vector operator**. All the relations for ∇ can be derived from the hybrid nature of del in terms of both the partial derivatives and its vector nature.

The gradient of a scalar is extremely important in physics and engineering in expressing the relation between a force field and a potential field

$$\text{force } \mathbf{F} = -\nabla(\text{ potential } V), \tag{1.65}$$

which holds for both gravitational and electrostatic fields, among others. Note that the minus sign in Eq. (1.65) results in water flowing downhill rather than uphill. If a force can be described as in Eq. (1.65) by a single function $V(\mathbf{r})$ everywhere, we call the scalar function V its **potential**. Because the force is the directional derivative of the potential, we can find the potential, if it exists, by integrating the force along a suitable path. Because the total variation $dV = \nabla V \cdot d\mathbf{r} = -\mathbf{F} \cdot d\mathbf{r}$ is the work done against the force along the path $d\mathbf{r}$, we recognize the physical meaning of the potential (difference) as work and energy. Moreover, in a sum of path increments the intermediate points cancel,

$$\begin{aligned} &[V(\mathbf{r} + d\mathbf{r}_1 + d\mathbf{r}_2) - V(\mathbf{r} + d\mathbf{r}_1)] + [V(\mathbf{r} + d\mathbf{r}_1) - V(\mathbf{r})] \\ &\quad = V(\mathbf{r} + d\mathbf{r}_2 + d\mathbf{r}_1) - V(\mathbf{r}), \end{aligned}$$

so that the integrated work along some path from an initial point $\mathbf{r}_i$ to a final point $\mathbf{r}$ is given by the potential difference $V(\mathbf{r}) - V(\mathbf{r}_i)$ at the end points of the path. Therefore, such forces are especially simple and well behaved: They are called **conservative**. When there is loss of energy due to friction along the path or some other dissipation, the work will depend on the path and such forces cannot be conservative: No potential exists. We discuss conservative forces in more detail in Section 1.12.

EXAMPLE 1.5.5

The Gradient of a Function of r Because we often deal with **central forces** in physics and engineering, we start with the gradient of the radial distance $r = \sqrt{x^2 + y^2 + z^2}$. From r as a function of x, y, z,

$$\frac{\partial r}{\partial x} = \frac{\partial (x^2 + y^2 + z^2)^{1/2}}{\partial x} = \frac{x}{(x^2 + y^2 + z^2)^{1/2}} = \frac{x}{r},$$

etc. Now we can calculate the more general gradient of a spherically symmetric potential $f(r)$ of a central force law so that

$$\nabla f(r) = \hat{\mathbf{x}}\frac{\partial f(r)}{\partial x} + \hat{\mathbf{y}}\frac{\partial f(r)}{\partial y} + \hat{\mathbf{z}}\frac{\partial f(r)}{\partial z}, \tag{1.66}$$

where $f(r)$ depends on x through the dependence of r on x. Therefore[7],

$$\frac{\partial f(r)}{\partial x} = \frac{df(r)}{dr} \cdot \frac{\partial r}{\partial x}.$$

Therefore,

$$\frac{\partial f(r)}{\partial x} = \frac{df(r)}{dr} \cdot \frac{x}{r}. \tag{1.67}$$

Permuting coordinates ($x \to y$, $y \to z$, $z \to x$) to obtain the y and z derivatives, we get

$$\nabla f(r) = (\hat{\mathbf{x}}x + \hat{\mathbf{y}}y + \hat{\mathbf{z}}z)\frac{1}{r}\frac{df}{dr} = \frac{\mathbf{r}}{r}\frac{df}{dr} = \hat{\mathbf{r}}\frac{df}{dr}, \tag{1.68}$$

where $\hat{\mathbf{r}}$ is a unit vector ($\mathbf{r}/r$) in the **positive** radial direction. The gradient of a function of r is a vector in the (positive or negative) radial direction. ■

A Geometrical Interpretation

Example 1.5.2 illustrates the geometric meaning of the gradient of a plane: It is its normal vector. This is a special case of the general geometric meaning of the gradient of an implicitly defined surface $\varphi(\mathbf{r}) = \text{const}$. Consider P and Q to be two points on a surface $\varphi(x, y, z) = C$, a constant. If φ is a potential, the surface is an **equipotential surface**. These points are chosen so that Q is a distance $d\mathbf{r}$ from P. Then, moving from P to Q, the change in $\varphi(x, y, z)$, given by Eq. (1.58) that is now written in vector notation, must be

$$d\varphi = (\nabla\varphi) \cdot d\mathbf{r} = 0 \tag{1.69}$$

since we stay on the surface $\varphi(x, y, z) = C$. This shows that $\nabla\varphi$ is perpendicular to $d\mathbf{r}$. Since $d\mathbf{r}$ may have any direction from P **as long as it stays in the surface** $\varphi = \text{const.}$, the point Q being restricted to the surface but having arbitrary direction, $\nabla\varphi$ is seen as **normal to the surface** $\varphi = \text{const.}$ (Fig. 1.23).

If we now permit $d\mathbf{r}$ to take us from one surface $\varphi = C_1$ to an adjacent surface $\varphi = C_2$ (Fig. 1.24),

$$d\varphi = C_1 - C_2 = \Delta C = (\nabla\varphi) \cdot d\mathbf{r}. \tag{1.70}$$

For a given $d\varphi$, $|d\mathbf{r}|$ is a minimum when it is chosen parallel to $\nabla\varphi$ ($\cos\theta = 1$); for a given $|d\mathbf{r}|$, the change in the scalar function φ is maximized by choosing $d\mathbf{r}$ parallel to $\nabla\varphi$. **This identifies $\nabla\varphi$ as a vector having the direction of the maximum space rate of change of** φ, an identification that will be useful in Chapter 2 when we consider non-Cartesian coordinate systems.

[7] This is a special case of the **chain rule** generalized to partial derivatives:

$$\frac{\partial f(r,\theta,\varphi)}{\partial x} = \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial x},$$

where $\partial f/\partial\theta = \partial f/\partial\varphi = 0$, $\partial f/\partial r \to df/dr$.

Figure 1.23

The Length Increment $d\mathbf{r}$ is Required to Stay on the Surface $\varphi = C$

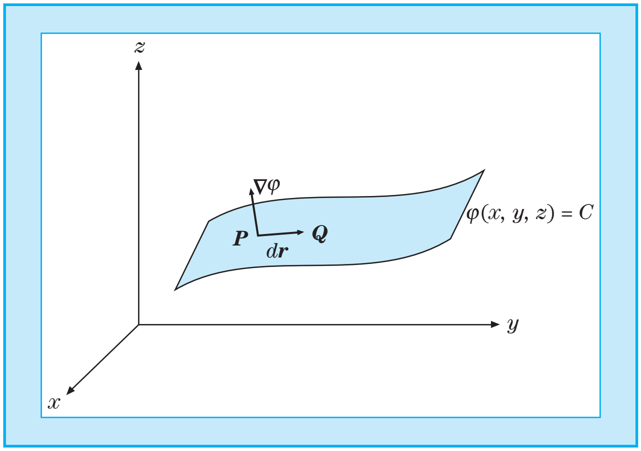
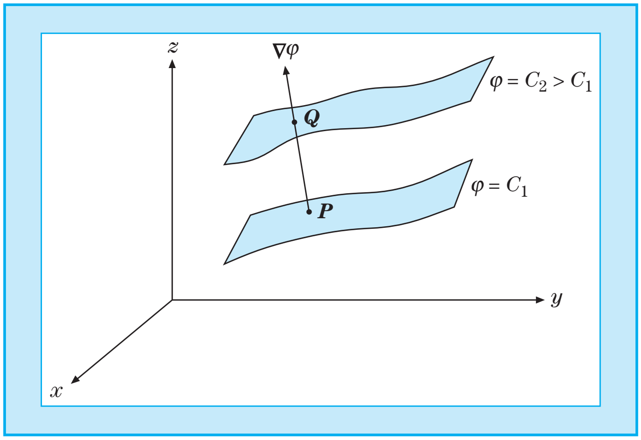


Figure 1.24

Gradient



SUMMARY

We have constructed the gradient operator as a vector of derivatives. The total variation of a function is the dot product of its gradient with the coordinate displacement vector. A conservative force is the (negative) gradient of a scalar called its potential.

EXERCISES

1.5.1 The dependence of the free fall acceleration g on geographical latitude ϕ at sea level is given by $g = g_0(1+0.0053\sin^2\phi)$. What is the southward displacement near $\phi = 30^\circ$ that changes g by 1 part in 10^8?

1.5.2 Given a vector $\mathbf{r}_{12} = \hat{\mathbf{x}}(x_1 - x_2) + \hat{\mathbf{y}}(y_1 - y_2) + \hat{\mathbf{z}}(z_1 - z_2)$, show that $\boldsymbol{\nabla}_1 r_{12}$ (gradient with respect to x_1, y_1, and z_1, of the magnitude r_{12}) is a unit vector in the direction of r_{12}. Note that a central force and a potential may depend on r_{12}.

1.5.3 If a vector function $\mathbf{F}$ depends on both space coordinates (x, y, z) and time t, show that

$$d\mathbf{F} = (d\mathbf{r} \cdot \nabla)\mathbf{F} + \frac{\partial \mathbf{F}}{\partial t}\, dt.$$

1.5.4 Show that $\nabla(uv) = v\nabla u + u\nabla v$, where u and v are differentiable scalar functions of x, y, and z (product rule).

(a) Show that a necessary and sufficient condition that $u(x, y, z)$ and $v(x, y, z)$ are related by some function $f(u, v) = 0$ is that $(\nabla u) \times (\nabla v) = 0$. Describe this geometrically. If graphical software is available, plot a typical case.

(b) If $u = u(x, y)$ and $v = v(x, y)$, show that the condition $(\nabla u) \times (\nabla v) = 0$ leads to the two-dimensional Jacobian

$$J\left(\frac{u, v}{x, y}\right) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x} = 0.$$

The functions u and v are assumed differentiable.

1.6 Divergence, ∇

In Section 1.5, ∇ was defined as a vector operator. Now, paying careful attention to both its vector and its differential properties, we let it operate on a vector. First, as a vector we dot it into a second vector to obtain

$$\nabla \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}, \tag{1.71}$$

known as the divergence of $\mathbf{V}$, which we expect to be a scalar.

EXAMPLE 1.6.1

Divergence of a Central Force Field From Eq. (1.71) we obtain for the coordinate vector with radial outward flow

$$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3. \tag{1.72}$$

Because the gravitational (or electric) force of a mass (or charge) at the origin is proportional to $\mathbf{r}$ with a radial $1/r^3$ dependence, we also consider the more general and important case of the divergence of a central force field

$$\begin{aligned}\nabla \cdot \mathbf{r} f(r) &= \frac{\partial}{\partial x}[x f(r)] + \frac{\partial}{\partial y}[y f(r)] + \frac{\partial}{\partial z}[z f(r)] \\ &= f(r)\nabla \cdot \mathbf{r} + x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 3f(r) + \frac{df}{dr}\mathbf{r} \cdot \nabla r \\ &= 3f(r) + \frac{x^2}{r}\frac{df}{dr} + \frac{y^2}{r}\frac{df}{dr} + \frac{z^2}{r}\frac{df}{dr} = 3f(r) + r\frac{df}{dr},\end{aligned} \tag{1.73}$$

using the product and chain rules of differentiation in conjunction with Example 1.5.5 and Eq. (1.71). In particular, if $f(r) = r^{n-1}$,

$$\nabla \cdot \mathbf{r} r^{n-1} = \nabla \cdot (\hat{\mathbf{r}} r^n) = 3r^{n-1} + (n-1) r^{n-1} = (n+2) r^{n-1}. \tag{1.74}$$

This divergence vanishes for $n = -2$, except at $r = 0$ (where $\hat{\mathbf{r}}/r^2$ is singular). This is relevant for the Coulomb potential

$$V(r) = A_0 = -\frac{q}{4\pi\epsilon_0 r}$$

with the electric field

$$\mathbf{E} = -\nabla V = \frac{q\hat{\mathbf{r}}}{4\pi\epsilon_0 r^2}.$$

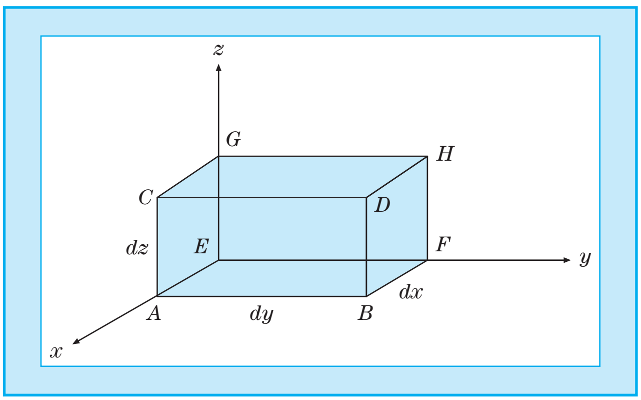
Using Eq. (1.74) we obtain the divergence $\nabla \cdot \mathbf{E} = 0$ (except at $r = 0$, where the derivatives are undefined). ■

A Physical Interpretation

To develop an understanding of the physical significance of the divergence, consider $\nabla \cdot (\rho \mathbf{v})$, with $\mathbf{v}(x, y, z)$, the velocity of a compressible fluid, and $\rho(x, y, z)$, its density at point (x, y, z). If we consider a small volume $dx\,dy\,dz$ (Fig. 1.25), the fluid flowing into this volume per unit time (positive x-direction) through the face $EFGH$ is (rate of flow in)$_{EFGH} = \rho v_x|_{x=0}\,dy\,dz$. The components of the flow ρv_y and ρv_z tangential to this face contribute nothing to the flow through this face. The rate of flow out (still positive x-direction) through face $ABCD$ is $\rho v_x|_{x=dx}\,dy\,dz$. To compare these flows and to find the net flow out, we add the change of ρv_x in the x-direction for an increment dx that

Figure 1.25

Differential Rectangular Parallelepiped (in the First or Positive Octant)



is given by its partial derivative (i.e., expand this last result in a Maclaurin series).[8] This yields

$$\begin{aligned}(\text{rate of flow out})_{ABCD} &= \rho v_x|_{x=dx}\, dy\, dz \\ &= \left[\rho v_x + \frac{\partial}{\partial x}(\rho v_x)\, dx\right]_{x=0} dy\, dz.\end{aligned}$$

Here, the derivative term is a first correction term allowing for the possibility of nonuniform density or velocity or both.[9] The zero-order term $\rho v_x|_{x=0}$ (corresponding to uniform flow) cancels out:

$$\text{Net rate of flow out}|_x = \frac{\partial}{\partial x}(\rho v_x)\, dx\, dy\, dz.$$

Equivalently, we can arrive at this result by

$$\lim_{\Delta x \to 0} \frac{\rho v_x(\Delta x, 0, 0) - \rho v_x(0, 0, 0)}{\Delta x} \equiv \left.\frac{\partial[\rho v_x(x, y, z)]}{\partial x}\right|_{(0,0,0)}.$$

Now the x-axis is not entitled to any preferred treatment. The preceding result for the two faces perpendicular to the x-axis must hold for the two faces perpendicular to the y-axis, with x replaced by y and the corresponding changes for y and z: $y \to z$, $z \to x$. This is a cyclic permutation of the coordinates. A further cyclic permutation yields the result for the remaining two faces of our parallelepiped. Adding the net rate of flow out for all three pairs of surfaces of our volume element, we have

$$\begin{aligned}\begin{matrix}\text{Net flow out}\\ \text{(per unit time)}\end{matrix} &= \left[\frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}(\rho v_y) + \frac{\partial}{\partial z}(\rho v_z)\right] dx\, dy\, dz \\ &= \nabla \cdot (\rho \mathbf{v})\, dx\, dy\, dz. \qquad (1.75)\end{aligned}$$

Therefore, the net flow of our compressible fluid out of the volume element $dx\, dy\, dz$ per unit volume per unit time is $\nabla \cdot (\rho \mathbf{v})$. Hence the name **divergence**. A direct application is in the **continuity equation**

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \qquad (1.76)$$

which states that a net flow out of the volume results in a decreased density inside the volume. Note that in Eq. (1.76), ρ is considered to be a possible function of time as well as of space: $\rho(x, y, z, t)$. The divergence appears in a wide variety of physical problems, ranging from a probability current density in quantum mechanics to neutron leakage in a nuclear reactor.

[8] A Maclaurin expansion for a single variable is given by Eq. (5.75) in Section 5.6. Here, we have the increment x of Eq. (5.75) replaced by dx. We show a partial derivative with respect to x because ρv_x may also depend on y and z.

[9] Strictly speaking, ρv_x is averaged over face $EFGH$ and the expression $\rho v_x + (\partial/\partial x)(\rho v_x)\, dx$ is similarly averaged over face $ABCD$. Using an arbitrarily small differential volume, we find that the averages reduce to the values employed here.

The combination $\nabla \cdot (f\mathbf{V})$, in which f is a scalar function and $\mathbf{V}$ a vector function, may be written as

$$\begin{aligned}\nabla \cdot (f\mathbf{V}) &= \frac{\partial}{\partial x}(fV_x) + \frac{\partial}{\partial y}(fV_y) + \frac{\partial}{\partial z}(fV_z) \\ &= \frac{\partial f}{\partial x}V_x + f\frac{\partial V_x}{\partial x} + \frac{\partial f}{\partial y}V_y + f\frac{\partial V_y}{\partial y} + \frac{\partial f}{\partial z}V_z + f\frac{\partial V_z}{\partial z} \\ &= (\nabla f)\cdot\mathbf{V} + f\nabla\cdot\mathbf{V}, \qquad (1.77)\end{aligned}$$

which is what we would expect for the derivative of a product. Notice that ∇ as a differential operator differentiates both f and $\mathbf{V}$; as a vector it is dotted into $\mathbf{V}$ (in each term).

SUMMARY

The divergence of a vector field is constructed as the dot product of the gradient with the vector field, and it locally measures its spatial outflow. In this sense, the continuity equation captures the essence of the divergence: the temporal change of the density balances the spatial outflow of the current density.

EXERCISES

1.6.1 For a particle moving in a circular orbit $\mathbf{r} = \hat{\mathbf{x}}r\cos\omega t + \hat{\mathbf{y}}r\sin\omega t$,

(a) evaluate $\mathbf{r}\times\dot{\mathbf{r}}$.

(b) Show that $\ddot{\mathbf{r}} + \omega^2\mathbf{r} = 0$.

The radius r and the angular velocity ω are constant.

ANS. (a) $\hat{\mathbf{z}}\omega r^2$. *Note*: $\dot{\mathbf{r}} = d\mathbf{r}/dt$, $\ddot{\mathbf{r}} = d^2\mathbf{r}/dt^2$.

1.6.2 Show, by differentiating components, that

(a) $\frac{d}{dt}(\mathbf{A}\cdot\mathbf{B}) = \frac{d\mathbf{A}}{dt}\cdot\mathbf{B} + \mathbf{A}\cdot\frac{d\mathbf{B}}{dt}$,

(b) $\frac{d}{dt}(\mathbf{A}\times\mathbf{B}) = \frac{d\mathbf{A}}{dt}\times\mathbf{B} + \mathbf{A}\times\frac{d\mathbf{B}}{dt}$,

in the same way as the derivative of the product of two scalar functions.

1.7 Curl, $\nabla\times$

Another possible application of the vector ∇ is to cross it into a vector field called its curl, which we discuss in this section along with its physical interpretation and applications. We obtain

$$\begin{aligned}\nabla\times\mathbf{V} &= \hat{\mathbf{x}}\left(\frac{\partial}{\partial y}V_z - \frac{\partial}{\partial z}V_y\right) + \hat{\mathbf{y}}\left(\frac{\partial}{\partial z}V_x - \frac{\partial}{\partial x}V_z\right) + \hat{\mathbf{z}}\left(\frac{\partial}{\partial x}V_y - \frac{\partial}{\partial y}V_x\right) \\ &= \begin{vmatrix}\hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z\end{vmatrix}, \qquad (1.78)\end{aligned}$$

which is called the **curl** of $\mathbf{V}$. In expanding this determinant we must consider the derivative nature of $\boldsymbol{\nabla}$. Specifically, $\mathbf{V} \times \boldsymbol{\nabla}$ is meaningless unless it acts on a function or a vector. Then it is certainly not equal, in general, to $-\boldsymbol{\nabla} \times \mathbf{V}$.[10] In the case of Eq. (1.78), the determinant must be expanded **from the top down** so that we get the derivatives as shown in the middle of Eq. (1.78). If $\boldsymbol{\nabla}$ is crossed into the product of a scalar and a vector, we can show

$$\begin{aligned}\boldsymbol{\nabla} \times (f\mathbf{V})|_x &= \left[\frac{\partial}{\partial y}(fV_z) - \frac{\partial}{\partial z}(fV_y)\right] \\ &= \left(f\frac{\partial V_z}{\partial y} + \frac{\partial f}{\partial y}V_z - f\frac{\partial V_y}{\partial z} - \frac{\partial f}{\partial z}V_y\right) \\ &= f\boldsymbol{\nabla} \times \mathbf{V}|_x + (\boldsymbol{\nabla} f) \times \mathbf{V}|_x .\end{aligned} \tag{1.79}$$

If we permute the coordinates $x \to y$, $y \to z$, $z \to x$ to pick up the y-component and then permute them a second time to pick up the z-component,

$$\boldsymbol{\nabla} \times (f\mathbf{V}) = f\boldsymbol{\nabla} \times \mathbf{V} + (\boldsymbol{\nabla} f) \times \mathbf{V}, \tag{1.80}$$

which is the vector product analog of Eq. (1.77). Again, as a differential operator, $\boldsymbol{\nabla}$ differentiates both f and $\mathbf{V}$. As a vector, it is crossed into $\mathbf{V}$ (in each term).

EXAMPLE 1.7.1

Vector Potential of a Constant B Field From electrodynamics we know that $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$, which has the general solution $\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}$, where $\mathbf{A}(\mathbf{r})$ is called the vector potential (of the magnetic induction) because $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = (\boldsymbol{\nabla} \times \boldsymbol{\nabla}) \cdot \mathbf{A} \equiv 0$ as a triple scalar product with two identical vectors. This last identity will not change if we add the gradient of some scalar function to the vector potential, which is therefore not unique.

In our case, we want to show that a vector potential is $\mathbf{A} = \frac{1}{2}(\mathbf{B} \times \mathbf{r})$.

Using the *BAC–CAB* rule in conjunction with Eq. (1.72), we find that

$$2\boldsymbol{\nabla} \times \mathbf{A} = \boldsymbol{\nabla} \times (\mathbf{B} \times \mathbf{r}) = (\boldsymbol{\nabla} \cdot \mathbf{r})\mathbf{B} - (\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{r} = 3\mathbf{B} - \mathbf{B} = 2\mathbf{B},$$

where we indicate by the ordering of the scalar product of the second term that the gradient still acts on the coordinate vector. ■

EXAMPLE 1.7.2

Curl of a Central Force As in Example 1.6.1, let us start with the curl of the coordinate vector

$$\boldsymbol{\nabla} \times \mathbf{r} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = 0. \tag{1.81}$$

[10] Note that for the quantum mechanical angular momentum **operator**, $\mathbf{L} = -i(\mathbf{r} \times \boldsymbol{\nabla})$, we find that $\mathbf{L} \times \mathbf{L} = i\mathbf{L}$. See Sections 4.3 and 4.4 for more details.

Algebraically, this comes about because each Cartesian coordinate is independent of the other two.

Now we are ready to calculate the curl of a central force $\nabla \times \mathbf{r} f(r)$, where we expect zero for the same reason. By Eq. (1.80),

$$\nabla \times \mathbf{r} f(r) = f(r)\nabla \times \mathbf{r} + [\nabla f(r)] \times \mathbf{r}. \tag{1.82}$$

Second, using $\nabla f(r) = \hat{\mathbf{r}}(df/dr)$ (Example 1.5.5), we obtain

$$\nabla \times \mathbf{r} f(r) = \frac{df}{dr}\hat{\mathbf{r}} \times \mathbf{r} = 0. \tag{1.83}$$

This vector product vanishes because $\mathbf{r} = \hat{\mathbf{r}}r$ and $\hat{\mathbf{r}} \times \hat{\mathbf{r}} = 0$.

This central force case is important in potential theory of classical mechanics and engineering (see Section 1.12). ■

To develop a better understanding of the physical significance of the curl, we consider the circulation of fluid around a differential loop in the xy-plane (Fig. 1.26).

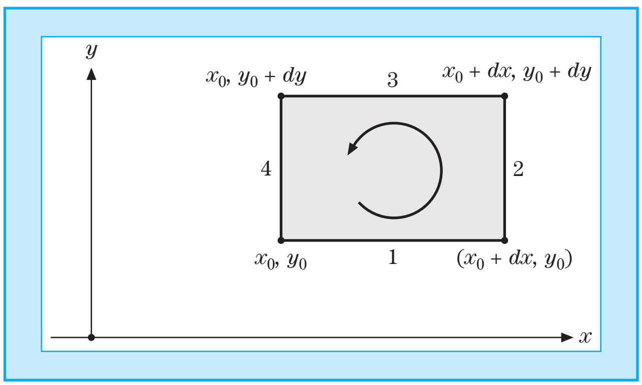
Although the circulation is technically given by a vector line integral $\int \mathbf{V} \cdot d\boldsymbol{\lambda}$, we can set up the equivalent scalar integrals here. Let us take the circulation to be

$$\begin{aligned}\text{Circulation}_{1234} &= \int_1 V_x(x, y)\, d\lambda_x + \int_2 V_y(x, y)\, d\lambda_y \\ &\quad + \int_3 V_x(x, y)\, d\lambda_x + \int_4 V_y(x, y)\, d\lambda_y. \end{aligned} \tag{1.84}$$

The numbers 1–4 refer to the numbered line segments in Fig. 1.26. In the first integral $d\lambda_x = +dx$ but in the third integral $d\lambda_x = -dx$ because the third line segment is traversed in the negative x-direction. Similarly, $d\lambda_y = +dy$ for the

Figure 1.26

Circulation Around a Differential Loop



second integral and $-dy$ for the fourth. Next, the integrands are referred to the point (x_0, y_0) with a Taylor expansion,[11] taking into account the displacement of line segment 3 from 1 and 2 from 4. For our differential line segments, this leads to

$$\begin{aligned} \text{Circulation}_{1234} &= V_x(x_0, y_0)\,dx + \left[V_y(x_0, y_0) + \frac{\partial V_y}{\partial x}\,dx\right]dy \\ &\quad + \left[V_x(x_0, y_0) + \frac{\partial V_x}{\partial y}\,dy\right](-dx) + V_y(x_0, y_0)(-dy) \\ &= \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)dx\,dy. \end{aligned} \tag{1.85}$$

Dividing by $dx\,dy$, we have

$$\text{Circulation per unit area} = \nabla \times \mathbf{V}|_z\,. \tag{1.86}$$

This is an infinitesimal case of Stokes's theorem in Section 1.11. The circulation[12] about our differential area in the xy-plane is given by the z-component of $\nabla \times \mathbf{V}$. In principle, the curl $\nabla \times \mathbf{V}$ at (x_0, y_0) could be determined by inserting a (differential) paddle wheel into the moving fluid at point (x_0, y_0). The rotation of the little paddle wheel would be a measure of the curl and its axis along the direction of $\nabla \times \mathbf{V}$, which is perpendicular to the plane of circulation.

In light of this connection of the curl with the concept of circulation, we now understand intuitively the vanishing curl of a central force in Example 1.7.2 because $\mathbf{r}$ flows radially outward from the origin with no rotation, and any scalar $f(r)$ will not affect this situation. When

$$\nabla \times \mathbf{V} = 0, \tag{1.87}$$

$\mathbf{V}$ is labeled **irrotational**. The most important physical examples of irrotational vectors are the gravitational and electrostatic forces. In each case,

$$\mathbf{V} = C\frac{\hat{\mathbf{r}}}{r^2} = C\frac{\mathbf{r}}{r^3}, \tag{1.88}$$

where C is a constant and $\hat{\mathbf{r}}$ is the unit vector in the outward radial direction. For the gravitational case, we have $C = -Gm_1m_2$, given by Newton's law of universal gravitation. If $C = q_1q_2/(4\pi\varepsilon_0)$, we have Coulomb's law of electrostatics (SI units). The force $\mathbf{V}$ given in Eq. (1.88) may be shown to be irrotational by direct expansion into Cartesian components as we did in Example 1.7.2 [Eq. (1.83)].

In Section 1.15 of Arfken and Weber's *Mathematical Methods for Physicists* (5th ed.), it is shown that a vector field may be resolved into an irrotational part and a solenoidal part (subject to conditions at infinity).

[11] $V_y(x_0 + dx, y_0) = V_y(x_0, y_0) + (\frac{\partial V_y}{\partial x})_{x_0 y_0} dx + \cdots$. The higher order terms will drop out in the limit as $dx \to 0$.

[12] In fluid dynamics, $\nabla \times \mathbf{V}$ is called the vorticity.

For waves in an elastic medium, if the displacement $\mathbf{u}$ is irrotational, $\nabla \times \mathbf{u} = 0$, plane waves (or spherical waves at large distances) become longitudinal. If $\mathbf{u}$ is solenoidal, $\nabla \cdot \mathbf{u} = 0$, then the waves become transverse. A seismic disturbance will produce a displacement that may be resolved into a solenoidal part and an irrotational part. The irrotational part yields the longitudinal P (primary) earthquake waves. The solenoidal part gives rise to the slower transverse S (secondary) waves.

Using the gradient, divergence, curl, and the *BAC–CAB* rule, we may construct or verify a large number of useful vector identities. For verification, complete expansion into Cartesian components is always a possibility. Sometimes if we use insight instead of routine shuffling of Cartesian components, the verification process can be shortened drastically.

Remember that ∇ is a vector operator, a hybrid object satisfying two sets of rules: vector rules and partial differentiation rules, including differentiation of a product.

EXAMPLE 1.7.3

Gradient of a Dot Product Verify that

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} + (\mathbf{A}\cdot\nabla)\mathbf{B} + \mathbf{B}\times(\nabla\times\mathbf{A}) + \mathbf{A}\times(\nabla\times\mathbf{B}). \qquad (1.89)$$

This particular example hinges on the recognition that $\nabla(\mathbf{A}\cdot\mathbf{B})$ is the type of term that appears in the *BAC–CAB* expansion of a triple vector product [Eq. (1.52)]. For instance,

$$\mathbf{A}\times(\nabla\times\mathbf{B}) = \nabla(\mathbf{A}\cdot\mathbf{B}) - (\mathbf{A}\cdot\nabla)\mathbf{B},$$

with the ∇ differentiating only $\mathbf{B}$, not $\mathbf{A}$. From the commutativity of factors in a scalar product we may interchange $\mathbf{A}$ and $\mathbf{B}$ and write

$$\mathbf{B}\times(\nabla\times\mathbf{A}) = \nabla(\mathbf{A}\cdot\mathbf{B}) - (\mathbf{B}\cdot\nabla)\mathbf{A},$$

now with ∇ differentiating only **A,** not $\mathbf{B}$. Adding these two equations, we obtain ∇ differentiating the product $\mathbf{A}\cdot\mathbf{B}$ and the identity [Eq. (1.89)]. This identity is used frequently in electromagnetic theory. Exercise 1.7.9 is an illustration. ■

SUMMARY

The curl is constructed as the cross product of the gradient and a vector field, and it measures the local rotational flow or circulation of the vector field. When the curl of a force field is zero, then the force is labeled conservative and derives from the gradient of a scalar, its potential. In Chapter 6, we shall see that an analytic function of a complex variable describes a two-dimensional irrotational fluid flow.

EXERCISES

1.7.1 Show that $\mathbf{u}\times\mathbf{v}$ is solenoidal if $\mathbf{u}$ and $\mathbf{v}$ are each irrotational. Start by formulating the problem in terms of mathematical equations.

1.7.2 If $\mathbf{A}$ is irrotational, show that $\mathbf{A}\times\mathbf{r}$ is solenoidal.

1.7.3 A rigid body is rotating with constant angular velocity $\boldsymbol{\omega}$. Show that the linear velocity $\mathbf{v}$ is solenoidal.

1.7.4 If a vector function $\mathbf{f}(x, y, z)$ is not irrotational but the product of f and a scalar function $g(x, y, z)$ is irrotational, show that

$$\mathbf{f} \cdot \nabla \times \mathbf{f} = 0.$$

1.7.5 Verify the vector identity

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} - \mathbf{B}(\nabla \cdot \mathbf{A}) + \mathbf{A}(\nabla \cdot \mathbf{B}).$$

Describe in words what causes the last two terms to appear in the identity beyond the *BAC–CAB* rule. If symbolic software is available, test the Cartesian components for a typical case, such as $\mathbf{A} = \mathbf{L}$, $\mathbf{B} = \mathbf{r}/r^3$.

1.7.6 As an alternative to the vector identity of Example 1.7.5, show that

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{A} \times \nabla) \times \mathbf{B} + (\mathbf{B} \times \nabla) \times \mathbf{A} + \mathbf{A}(\nabla \cdot \mathbf{B}) + \mathbf{B}(\nabla \cdot \mathbf{A}).$$

1.7.7 Verify the identity

$$\mathbf{A} \times (\nabla \times \mathbf{A}) = \frac{1}{2}\nabla(A^2) - (\mathbf{A} \cdot \nabla)\mathbf{A}.$$

Test this identity for a typical vector field, such as $\mathbf{A} \sim \mathbf{r}$ or $\mathbf{r}/r^3$.

1.7.8 If $\mathbf{A}$ and $\mathbf{B}$ are constant vectors, show that

$$\nabla(\mathbf{A} \cdot \mathbf{B} \times \mathbf{r}) = \mathbf{A} \times \mathbf{B}.$$

1.7.9 A distribution of electric currents creates a constant magnetic moment $\mathbf{m}$. The force on $\mathbf{m}$ in an external magnetic induction $\mathbf{B}$ is given by

$$\mathbf{F} = \nabla \times (\mathbf{B} \times \mathbf{m}).$$

Show that

$$\mathbf{F} = \nabla(\mathbf{m} \cdot \mathbf{B}).$$

Note. Assuming no time dependence of the fields, Maxwell's equations yield $\nabla \times \mathbf{B} = 0$. Also, $\nabla \cdot \mathbf{B} = 0$.

1.7.10 An electric dipole of moment $\mathbf{p}$ is located at the origin. The dipole creates an electric potential at $\mathbf{r}$ given by

$$\psi(\mathbf{r}) = \frac{\mathbf{p} \cdot \mathbf{r}}{4\pi\varepsilon_0 r^3}.$$

Find the electric field $\mathbf{E} = -\nabla\psi$ at $\mathbf{r}$.

1.7.11 The vector potential $\mathbf{A}$ of a magnetic dipole, dipole moment $\mathbf{m}$, is given by $\mathbf{A}(\mathbf{r}) = (\mu_0/4\pi)(\mathbf{m} \times \mathbf{r}/r^3)$. Show that the magnetic induction $\mathbf{B} = \nabla \times \mathbf{A}$ is given by

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{3\hat{\mathbf{r}}(\hat{\mathbf{r}} \cdot \mathbf{m}) - \mathbf{m}}{r^3}.$$

1.7.12 Classically, orbital angular momentum is given by $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, where $\mathbf{p}$ is the linear momentum. To go from classical mechanics to quantum mechanics, replace $\mathbf{p}$ by the operator $-i\boldsymbol{\nabla}$ (Section 14.6). Show that the quantum mechanical angular momentum operator has Cartesian components

$$L_x = -i\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$
$$L_y = -i\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$
$$L_z = -i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

(in units of $\hbar$).

1.7.13 Using the angular momentum operators previously given, show that they satisfy commutation relations of the form

$$[L_x, L_y] \equiv L_x L_y - L_y L_x = iL_z$$

and, hence,

$$\mathbf{L} \times \mathbf{L} = i\mathbf{L}.$$

These commutation relations will be taken later as the defining relations of an angular momentum operator—see Exercise 3.2.15 and the following one and Chapter 4.

1.7.14 With the commutator bracket notation $[L_x, L_y] = L_x L_y - L_y L_x$, the angular momentum vector $\mathbf{L}$ satisfies $[L_x, L_y] = iL_z$, etc., or $\mathbf{L} \times \mathbf{L} = i\mathbf{L}$. If two other vectors $\mathbf{a}$ and $\mathbf{b}$ commute with each other and with $\mathbf{L}$, that is, $[\mathbf{a}, \mathbf{b}] = [\mathbf{a}, \mathbf{L}] = [\mathbf{b}, \mathbf{L}] = 0$, show that

$$[\mathbf{a} \cdot \mathbf{L}, \mathbf{b} \cdot \mathbf{L}] = i(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{L}.$$

This vector version of the angular momentum commutation relations is an alternative to that given in Exercise 1.7.13.

1.7.15 Prove $\boldsymbol{\nabla} \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\boldsymbol{\nabla} \times \mathbf{a}) - \mathbf{a} \cdot (\boldsymbol{\nabla} \times \mathbf{b})$. Explain in words why the identity is valid.
Hint. Treat as a triple scalar product.

1.8 Successive Applications of ∇

We have now defined gradient, divergence, and curl to obtain vector, scalar, and vector quantities, respectively. Letting $\boldsymbol{\nabla}$ operate on each of these quantities, we obtain

(a) $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\varphi$ (b) $\boldsymbol{\nabla} \times \boldsymbol{\nabla}\varphi$ (c) $\boldsymbol{\nabla}\boldsymbol{\nabla} \cdot \mathbf{V}$
(d) $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \times \mathbf{V}$ (e) $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{V})$.

All five expressions involve second derivatives and all five appear in the second-order differential equations of mathematical physics, particularly in electromagnetic theory.

The first expression, $\nabla \cdot \nabla\varphi$, the divergence of the gradient, is called the Laplacian of φ. We have

$$\begin{aligned}\nabla \cdot \nabla\varphi &= \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right) \cdot \left(\hat{\mathbf{x}}\frac{\partial\varphi}{\partial x} + \hat{\mathbf{y}}\frac{\partial\varphi}{\partial y} + \hat{\mathbf{z}}\frac{\partial\varphi}{\partial z}\right) \\ &= \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}. \end{aligned} \tag{1.90}$$

When φ is the electrostatic potential, in a charge-free region we have

$$\nabla \cdot \nabla\varphi = 0, \tag{1.91}$$

which is Laplace's equation of electrostatics. Often, the combination $\nabla \cdot \nabla$ is written ∇^2, or Δ in the European literature.

Biographical Data

Laplace, Pierre Simon. Laplace, a French mathematician, physicist, and astronomer, was born in Beaumont-en-Auge in 1749 and died in Paris in 1827. He developed perturbation theory for the solar system, published a monumental treatise *Celestial Mechanics*, and applied mathematics to artillery. He made contributions of fundamental importance to hydrodynamics, differential equations and probability, the propagation of sound, and surface tension in liquids. To Napoleon's remark missing "God" in his treatise, he replied "I had no need for that hypothesis." He generally disliked giving credit to others.

EXAMPLE 1.8.1

Laplacian of a Radial Function Calculate $\nabla\cdot\nabla g(r)$. Referring to Examples 1.5.5 and 1.6.1,

$$\nabla \cdot \nabla g(r) = \nabla \cdot \hat{\mathbf{r}}\frac{dg}{dr} = \frac{2}{r}\frac{dg}{dr} + \frac{d^2g}{dr^2},$$

replacing $f(r)$ in Example 1.6.1 by $1/r \cdot dg/dr$. If $g(r) = r^n$, this reduces to

$$\nabla \cdot \nabla r^n = n(n+1)r^{n-2}.$$

This vanishes for $n = 0$ [$g(r) = $ constant] and for $n = -1$; that is, $g(r) = 1/r$ is a solution of Laplace's equation, $\nabla^2 g(r) = 0$. This is for $r \neq 0$. At the origin there is a singularity. ■

Expression (b) may be written as

$$\nabla \times \nabla\varphi = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\varphi}{\partial x} & \frac{\partial\varphi}{\partial y} & \frac{\partial\varphi}{\partial z} \end{vmatrix}.$$

By expanding the determinant, we obtain

$$\nabla \times \nabla\varphi = \hat{\mathbf{x}}\left(\frac{\partial^2\varphi}{\partial y\,\partial z} - \frac{\partial^2\varphi}{\partial z\,\partial y}\right) + \hat{\mathbf{y}}\left(\frac{\partial^2\varphi}{\partial z\,\partial x} - \frac{\partial^2\varphi}{\partial x\,\partial z}\right)$$
$$+\,\hat{\mathbf{z}}\left(\frac{\partial^2\varphi}{\partial x\,\partial y} - \frac{\partial^2\varphi}{\partial y\,\partial x}\right) = 0, \tag{1.92}$$

assuming that the order of partial differentiation may be interchanged. This is true as long as these second partial derivatives of φ are continuous functions. Then, from Eq. (1.92), the curl of a gradient is identically zero. All gradients, therefore, are irrotational. Note that the zero in Eq. (1.92) comes as a mathematical identity, independent of any physics. The zero in Eq. (1.91) is a consequence of physics.

Expression (d) is a triple scalar product that may be written as

$$\nabla \cdot \nabla \times \mathbf{V} = \begin{vmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ V_x & V_y & V_z \end{vmatrix}. \tag{1.93}$$

Again, assuming continuity so that the order of differentiation is immaterial, we obtain

$$\nabla \cdot \nabla \times \mathbf{V} = 0. \tag{1.94}$$

The divergence of a curl vanishes or all curls are solenoidal.

One of the most important cases of a vanishing divergence of a vector is

$$\nabla \cdot \mathbf{B} = 0, \tag{1.95}$$

where $\mathbf{B}$ is the magnetic induction, and Eq. (1.95) appears as one of Maxwell's equations. When a vector is solenoidal, it may be written as the curl of another vector known as its vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$. This form solves one of the four vector equations that make up Maxwell's field equations of electrodynamics. Because a vector field may be determined from its curl and divergence (Helmholtz's theorem), solving Maxwell's (often called Oersted's) equation involving the curl of $\mathbf{B}$ determines $\mathbf{A}$ and thereby $\mathbf{B}$. Similar considerations apply to the other pair of Maxwell's equations involving the divergence and curl of $\mathbf{E}$ and make plausible the fact that there are precisely four vector equations as part of Maxwell's equations.

The two remaining expressions satisfy a relation

$$\nabla \times (\nabla \times \mathbf{V}) = \nabla(\nabla \cdot \mathbf{V}) - (\nabla \cdot \nabla)\mathbf{V}. \tag{1.96}$$

This decomposition of the Laplacian $\nabla \cdot \nabla$ into a longitudinal part (the gradient) and a transverse part (the curl term) follows from Eq. (1.52), the *BAC–CAB* rule, which we rewrite so that $\mathbf{C}$ appears at the extreme right of each term. The term $(\nabla \cdot \nabla)\mathbf{V}$ was not included in our list, but it appears in the Navier–Stokes's equation and may be **defined** by Eq. (1.96). In words, this is the Laplacian (a scalar operator) acting on a vector, so it is a vector with three components in three-dimensional space. ■

EXAMPLE 1.8.2

Electromagnetic Wave Equations One important application of this vector relation [Eq. (1.96)] is in the derivation of the electromagnetic wave equation. In vacuum Maxwell's equations become

$$\nabla \cdot \mathbf{B} = 0, \tag{1.97a}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{1.97b}$$

$$\nabla \times \mathbf{B} = \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}, \tag{1.97c}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.97d}$$

where $\mathbf{E}$ is the electric field, $\mathbf{B}$ the magnetic induction, ε_0 the electric permittivity, and μ_0 the magnetic permeability (SI units), so that $\varepsilon_0\mu_0 = 1/c^2$, where c is the velocity of light. This relation has important consequences. Because ε_0, μ_0 can be measured in any frame, the velocity of light is the same in any frame.

Suppose we eliminate $\mathbf{B}$ from Eqs. (1.97c) and (1.97d). We may do this by taking the curl of both sides of Eq. (1.97d) and the time derivative of both sides of Eq. (1.97c). Since the space and time derivatives commute,

$$\frac{\partial}{\partial t} \nabla \times \mathbf{B} = \nabla \times \frac{\partial \mathbf{B}}{\partial t},$$

and we obtain

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Application of Eqs. (1.96) and (1.97b) yields

$$(\nabla \cdot \nabla)\mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \tag{1.98}$$

the electromagnetic vector wave equation. Again, if $\mathbf{E}$ is expressed in Cartesian coordinates, Eq. (1.98) separates into three scalar wave equations, each involving a scalar Laplacian.

When external electric charge and current densities are kept as driving terms in Maxwell's equations, similar wave equations are valid for the electric potential and the vector potential. To show this, we solve Eq. (1.97a) by writing $\mathbf{B} = \nabla \times \mathbf{A}$ as a curl of the vector potential. This expression is substituted into Faraday's induction law in differential form [Eq. (1.97d)] to yield $\nabla \times (\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}) = 0$. The vanishing curl implies that $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$ is a gradient and therefore can be written as $-\nabla\varphi$, where $\varphi(\mathbf{r}, t)$ is defined as the (nonstatic) electric potential. These results

$$\mathbf{B} = \nabla \times \mathbf{A}, \qquad \mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} \tag{1.99}$$

for the $\mathbf{B}$ and $\mathbf{E}$ fields solve the homogeneous Maxwell's equations.

We now show that the inhomogeneous Maxwell's equations,

$$\text{Gauss's law: } \nabla \cdot \mathbf{E} = \rho/\varepsilon_0;$$
$$\text{Oersted's law: } \nabla \times \mathbf{B} - \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J} \tag{1.100}$$

in differential form lead to wave equations for the potentials φ and $\mathbf{A}$, provided that $\nabla \cdot \mathbf{A}$ is determined by the constraint $\frac{1}{c^2}\frac{\partial\varphi}{\partial t} + \nabla \cdot \mathbf{A} = 0$. This choice of fixing the divergence of the vector potential is called the **Lorentz gauge** and serves to uncouple the partial differential equations of both potentials. This gauge constraint is not a restriction; it has no physical effect.

Substituting our electric field solution into Gauss's law yields

$$\frac{\rho}{\varepsilon_0} = \nabla \cdot \mathbf{E} = -\nabla^2\varphi - \frac{\partial}{\partial t}\nabla \cdot \mathbf{A} = -\nabla^2\varphi + \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2},$$

the wave equation for the electric potential. In the last step, we used the Lorentz gauge to replace the divergence of the vector potential by the time derivative of the electric potential and thus decouple φ from $\mathbf{A}$.

Finally, we substitute $\mathbf{B} = \nabla \times \mathbf{A}$ into Oersted's law and use Eq. (1.96), which expands ∇^2 in terms of a longitudinal (the gradient term) and a transverse component (the curl term). This yields

$$\mu_0\mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t} = \nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} = \mu_0\mathbf{J} - \frac{1}{c^2}\left(\nabla\frac{\partial\varphi}{\partial t} + \frac{\partial^2\mathbf{A}}{\partial t^2}\right),$$

where we have used the electric field solution [Eq. (1.99)] in the last step. Now we see that the Lorentz gauge condition eliminates the gradient terms so that the wave equation

$$\frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = \mu_0\mathbf{J}$$

for the vector potential remains.

Finally, looking back at Oersted's law, taking the divergence of Eq. (1.100), dropping $\nabla \cdot (\nabla \times \mathbf{B}) = 0$ and substituting Gauss's law for $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, we find $\mu_0\nabla \cdot \mathbf{J} = -\frac{1}{\epsilon_0 c^2}\frac{\partial\rho}{\partial t}$, where $\epsilon_0\mu_0 = 1/c^2$, that is, the continuity equation for the current density. This step justifies the inclusion of Maxwell's displacement current in the generalization of Oersted's law to nonstationary situations. ■

EXERCISES

1.8.1 Verify Eq. (1.96)

$$\nabla \times (\nabla \times \mathbf{V}) = \nabla(\nabla \cdot \mathbf{V}) - (\nabla \cdot \nabla)\mathbf{V}$$

by direct expansion in Cartesian coordinates. If symbolic software is available, check the identity for typical fields, such as $\mathbf{V} = \mathbf{r}, \mathbf{r}/r^3, \mathbf{a} \cdot \mathbf{r}\mathbf{b}, \mathbf{a} \times \mathbf{r}$.

1.8.2 Show that the identity

$$\nabla \times (\nabla \times \mathbf{V}) = \nabla(\nabla \cdot \mathbf{V}) - (\nabla \cdot \nabla)\mathbf{V}$$

follows from the *BAC–CAB* rule for a triple vector product. Justify any alteration of the order of factors in the *BAC* and *CAB* terms.

1.8.3 Prove that $\nabla \times (\varphi \nabla \varphi) = 0$.

1.8.4 Prove that $(\nabla u) \times (\nabla v)$ is solenoidal, where u and v are differentiable scalar functions. Start by formulating the problem as a mathematical equation.

1.8.5 φ is a scalar satisfying Laplace's equation, $\nabla^2 \varphi = 0$. Show that $\nabla \varphi$ is **both** solenoidal and irrotational.

1.8.6 With ψ a scalar function, show that

$$(\mathbf{r} \times \nabla) \cdot (\mathbf{r} \times \nabla)\psi = r^2 \nabla^2 \psi - r^2 \frac{\partial^2 \psi}{\partial r^2} - 2r \frac{\partial \psi}{\partial r}.$$

(This can actually be shown more easily in spherical polar coordinates; see Section 2.5.)

1.8.7 In the Pauli theory of the electron one encounters the expression

$$(\mathbf{p} - e\mathbf{A}) \times (\mathbf{p} - e\mathbf{A})\psi,$$

where ψ is a scalar function. $\mathbf{A}$ is the magnetic vector potential related to the magnetic induction $\mathbf{B}$ by $\mathbf{B} = \nabla \times \mathbf{A}$. Given that $\mathbf{p} = -i\nabla$, show that this expression reduces to $ie\mathbf{B}\psi$. Show that this leads to the orbital g-factor $g_L = 1$ upon writing the magnetic moment as $\boldsymbol{\mu} = g_L \mathbf{L}$ in units of Bohr magnetons. See also Example 1.7.1.

1.8.8 Show that any solution of the equation

$$\nabla \times \nabla \times \mathbf{A} - k^2 \mathbf{A} = 0$$

automatically satisfies the vector Helmholtz equation

$$\nabla^2 \mathbf{A} + k^2 \mathbf{A} = 0$$

and the solenoidal condition

$$\nabla \cdot \mathbf{A} = 0.$$

Hint. Let $\nabla \cdot$ operate on the first equation.

1.9 Vector Integration

The next step after differentiating vectors is to integrate them. Let us start with line integrals and then proceed to surface and volume integrals. In each case, the method of attack will be to reduce the vector integral to one-dimensional integrals over a coordinate interval.

Line Integrals

Using an increment of length $dr = \hat{\mathbf{x}}\,dx + \hat{\mathbf{y}}\,dy + \hat{\mathbf{z}}\,dz$, we often encounter the line integral

$$\int_C \mathbf{V} \cdot d\mathbf{r}, \tag{1.101}$$

in which the integral is over some contour C that may be open (with starting point and ending point separated) or closed (forming a loop) instead of an interval of the x-axis. The Riemann integral is defined by subdividing the curve into ever smaller segments whose number grows indefinitely. The form [Eq. (1.101)] is exactly the same as that encountered when we calculate the work done by a force that varies along the path

$$W = \int \mathbf{F} \cdot d\mathbf{r} = \int F_x(x, y, z)\,dx + \int F_y(x, y, z)\,dy + \int F_z(x, y, z)\,dz, \tag{1.102}$$

that is, a sum of conventional integrals over intervals of one variable each. In this expression, **F** is the force exerted on a particle. In general, such integrals depend on the path except for conservative forces, whose treatment we postpone to Section 1.12.

EXAMPLE 1.9.1

Path-Dependent Work The force exerted on a body is $\mathbf{F} = -\hat{\mathbf{x}}y + \hat{\mathbf{y}}x$. The problem is to calculate the work done going from the origin to the point (1, 1),

$$W = \int_{0,0}^{1,1} \mathbf{F} \cdot d\mathbf{r} = \int_{0,0}^{1,1} (-y\,dx + x\,dy). \tag{1.103}$$

Separating the two integrals, we obtain

$$W = -\int_0^1 y\,dx + \int_0^1 x\,dy. \tag{1.104}$$

The first integral cannot be evaluated until we specify the values of y as x ranges from 0 to 1. Likewise, the second integral requires x as a function of y. Consider first the path shown in Fig. 1.27. Then

$$W = -\int_0^1 0\,dx + \int_0^1 1\,dy = 1 \tag{1.105}$$

because $y = 0$ along the first segment of the path and $x = 1$ along the second. If we select the path $[x = 0, 0 \le y \le 1]$ and $[0 \le x \le 1, y = 1]$, then Eq. (1.103) gives $W = -1$. For this force, the work done depends on the choice of path. ■

EXAMPLE 1.9.2

Line Integral for Work Find the work done going around a unit circle clockwise from 0 to $-\pi$ shown in Fig. 1.28 in the xy-plane doing work against a force field given by

$$\mathbf{F} = \frac{-\hat{\mathbf{x}}y}{x^2 + y^2} + \frac{\hat{\mathbf{y}}x}{x^2 + y^2}.$$

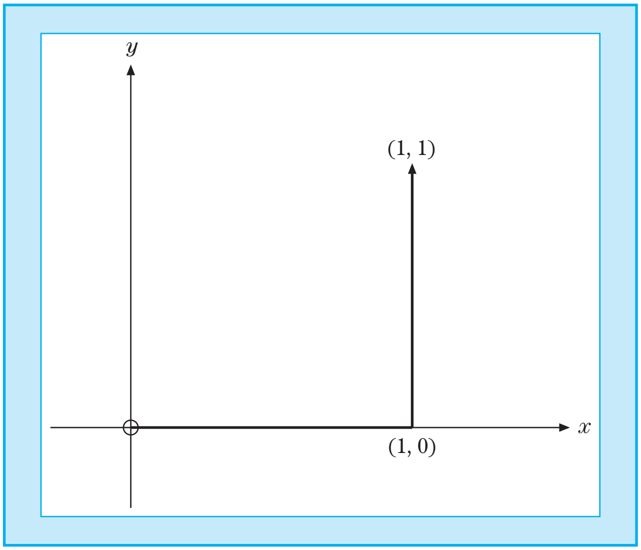


Figure 1.27

A Path of Integration

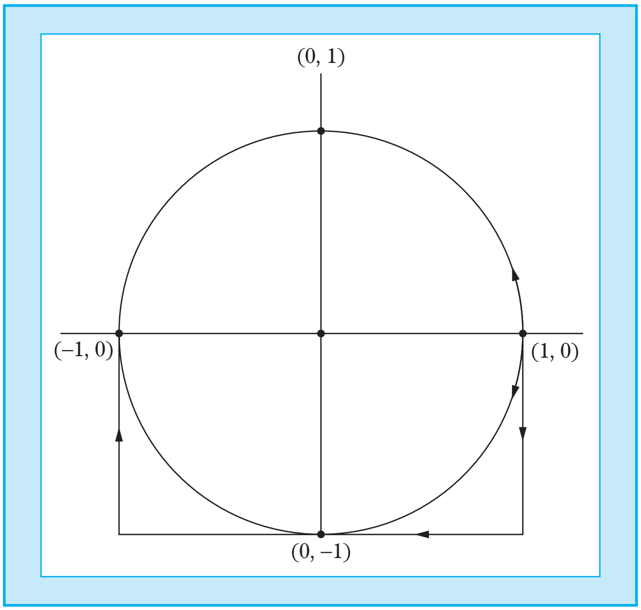


Figure 1.28

Circular and Square Integration Paths

Let us parameterize the circle C as $x = \cos\varphi$, $y = \sin\varphi$ with the polar angle φ so that $dx = -\sin\varphi d\varphi$, $dy = \cos\varphi d\varphi$. Then the force can be written as $\mathbf{F} = -\hat{\mathbf{x}}\sin\varphi + \hat{\mathbf{y}}\cos\varphi$. The work becomes

$$-\int_C \frac{xdy - ydx}{x^2 + y^2} = \int_0^{-\pi} (-\sin^2\varphi - \cos^2\varphi)\, d\varphi = \pi.$$

Here we spend energy. If we integrate anticlockwise from $\varphi = 0$ to π we find the value $-\pi$ because we are riding with the force. The work is path dependent, which is consistent with the physical interpretation that $\mathbf{F} \cdot d\mathbf{r} \sim x dy - y dx = L_z$ is proportional to the z-component of orbital angular momentum (involving circulation, as discussed in Section 1.7).

If we integrate along the square through the points $(\pm 1, 0)$, $(0, -1)$ surrounding the circle, we find for the clockwise lower half square path of Fig. 1.28

$$
\begin{aligned}
-\int \mathbf{F} \cdot d\mathbf{r} &= -\int_0^{-1} F_y \, dy|_{x=1} - \int_1^{-1} F_x \, dx|_{y=-1} - \int_{-1}^{0} F_y \, dy|_{x=-1} \\
&= \int_0^1 \frac{dy}{1+y^2} + \int_{-1}^{1} \frac{dx}{x^2 + (-1)^2} + \int_{-1}^{0} \frac{dy}{(-1)^2 + y^2} \\
&= \arctan(1) + \arctan(1) - \arctan(-1) - \arctan(-1) \\
&= 4 \cdot \frac{\pi}{4} = \pi,
\end{aligned}
$$

which is consistent with the circular path.

For the circular paths we used the $x = \cos\varphi$, $y = \sin\varphi$ parameterization, whereas for the square shape we used the standard definitions $y = f(x)$ or $x = g(y)$ of a curve, that is, $y = -1 = \text{const.}$ and $x = \pm 1 = \text{const.}$ We could have used the implicit definition $F(x, y) \equiv x^2 + y^2 - 1 = 0$ of the circle. Then the total variation

$$
dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = 2x\, dx + 2y\, dy \equiv 0
$$

so that

$$
dy = -x\, dx/y \;\text{ with }\; y = -\sqrt{1 - x^2}
$$

on our half circle. The work becomes

$$
\begin{aligned}
-\int_C \frac{x\, dy - y\, dx}{x^2 + y^2} &= \int \left(\frac{x^2}{y} + y \right) dx = \int \frac{dx}{y} = \int_1^{-1} \frac{dx}{-\sqrt{1 - x^2}} \\
&= \arcsin 1 - \arcsin(-1) = 2 \cdot \frac{\pi}{2} = \pi,
\end{aligned}
$$

in agreement with our previous results. ■

EXAMPLE 1.9.3

Gravitational Potential If a force can be described by a scalar function V_G as $\mathbf{F} = -\nabla V_G(\mathbf{r})$ [Eq. (1.65)], everywhere we call V_G its potential in mechanics and engineering. Because the total variation $dV_G = \nabla V_G \cdot d\mathbf{r} = -\mathbf{F}_G \cdot d\mathbf{r}$ is the work done against the force along the path $d\mathbf{r}$, the integrated work along any path from the initial point $\mathbf{r}_0$ to the final point $\mathbf{r}$ is given by a line integral $\int_{\mathbf{r}_0}^{\mathbf{r}} dV_G = V_G(\mathbf{r}) - V_G(\mathbf{r}_0)$, the potential difference between the end points of the path. Thus, to find the scalar potential for the gravitational force on a unit mass m_1,

$$\mathbf{F}_G = -\frac{Gm_1m_2\hat{\mathbf{r}}}{r^2} = -\frac{k\hat{\mathbf{r}}}{r^2}, \qquad \text{radially } \textbf{inward} \tag{1.106}$$

we integrate from infinity, where V_G is zero into position $\mathbf{r}$. We obtain

$$V_G(r) - V_G(\infty) = -\int_\infty^{\mathbf{r}} \mathbf{F}_G \cdot d\mathbf{r} = +\int_{\mathbf{r}}^\infty \mathbf{F}_G \cdot d\mathbf{r}. \tag{1.107}$$

By use of $\mathbf{F}_G = -\mathbf{F}_{\text{applied}}$, the potential is the work done in bringing the unit mass in from infinity. (We can define only the potential difference. Here, we arbitrarily assign infinity to be a zero of potential.) Since $\mathbf{F}_G$ is radial, we obtain a contribution to V_G only when $d\mathbf{r}$ is radial or

$$V_G(r) = -\int_r^\infty \frac{k\,dr}{r^2} = -\frac{k}{r} = -\frac{Gm_1m_2}{r}. \tag{1.108}$$

The negative sign reflects the attractive nature of gravity. ■

Surface Integrals

Surface integrals appear in the same forms as line integrals, the element of area also being a vector, $d\boldsymbol{\sigma}$.[13] Often this area element is written $\mathbf{n}\,dA$, where $\mathbf{n}$ is a unit (normal) vector to indicate the positive direction.[14] There are two conventions for choosing the positive direction. First, if the surface is a closed surface, we agree to take the outward normal as positive. Second, if the surface is an open surface, the positive normal depends on the direction in which the perimeter of the open surface is traversed. If the right-hand fingers are curled in the direction of travel around the perimeter, the positive normal is indicated by the thumb of the right hand. As an illustration, a circle in the xy-plane (Fig. 1.29) mapped out from x to y to $-x$ to $-y$ and back to x will have its positive normal parallel to the positive z-axis (for the right-handed coordinate system).

Analogous to the line integrals, Eq. (1.101), surface integrals may appear in the form

$$\int \mathbf{V} \cdot d\boldsymbol{\sigma}. \tag{1.109}$$

This surface integral $\int \mathbf{V} \cdot d\boldsymbol{\sigma}$ may be interpreted as a flow or flux through the given surface. This is really what we did in Section 1.6 to understand the significance of the concept of divergence. Note that both physically and from the dot product the tangential components of the velocity contribute nothing to the flow through the surface.

[13] Recall that in Section 1.3 the area (of a parallelogram) is represented by a cross product **vector**.
[14] Although $\mathbf{n}$ always has unit length, its direction may well be a function of position.

Figure 1.29

Right-Hand Rule for the Positive Normal

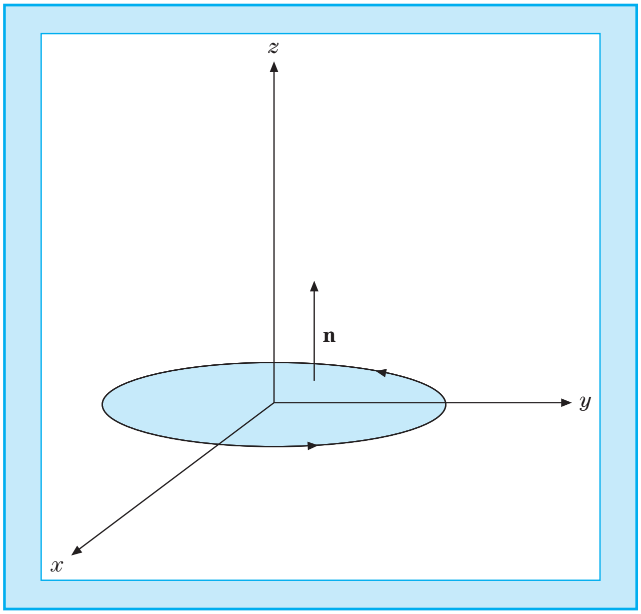
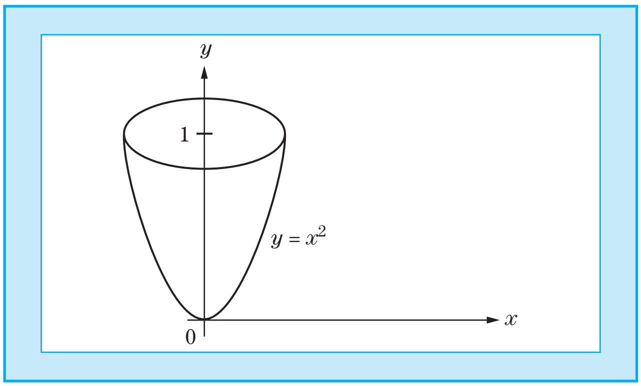


Figure 1.30

The Parabola $y = x^2$ for $0 \le y \le 1$ Rotated About the y-Axis



EXAMPLE 1.9.4

Moment of Inertia Let us determine the moment of inertia I_y of a segment of the parabola $y = x^2$ cut off by the line $y = 1$ and rotated about the y-axis (Fig. 1.30). We find

$$I_y = 2\mu \int_{x=0}^{1} \int_{y=x^2}^{1} x^2\,dx\,dy = 2\mu \int_0^1 (1 - x^2)x^2 dx = 2\mu \left(\frac{x^3}{3} - \frac{x^5}{5} \right)\Big|_0^1 = \frac{4\mu}{15}.$$

The factor of 2 originates in the reflection symmetry of $x \to -x$, and μ is the constant mass density. ■

A surface in three-dimensional space may be explicitly given as $z = f(x, y)$ or by the coordinate functions of its points

$$x = x(u, v), \qquad y = y(u, v), \qquad z = z(u, v)$$

in terms of two parameters u, v or in implicit form $F(x, y, z) = 0$. The explicit form is a special case

$$F(x, y, z) \equiv z - f(x, y)$$

of the general implicit definition of a surface. We find the area $dA = dx\,dy/n_z$ over the projection $dx\,dy$ of the surface onto the xy-plane for the latter case. Here, $n_z = \cos\gamma$ is the z-component of the normal unit vector $\mathbf{n}$ at $\mathbf{r}$ on the surface so that γ is the angle of $d\mathbf{A}$ with the xy-plane. Thus, when we project $d\mathbf{A}$ to the xy-plane, we get $dA\cos\gamma = dx\,dy$, which proves this useful formula for measuring the **area of a curved surface**. From the gradient properties we also know that $\mathbf{n} = \nabla f/\sqrt{\nabla f^2}$.

EXAMPLE 1.9.5

A Surface Integral Here we apply the general formula for surface integrals to find the area on $z = xy = f(x, y)$ cut out by the unit circle in the xy-plane shown in Fig. 1.31. We start from

$$\frac{\partial f}{\partial x} = \frac{\partial z}{\partial x} = y, \qquad \frac{\partial f}{\partial y} = \frac{\partial z}{\partial y} = x, \qquad \frac{\partial f}{\partial z} = \frac{\partial z}{\partial z} = 1,$$

which we substitute into

$$n_z = 1\Big/\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}$$

for the normal to yield the area

$$A = \int_{x=-1}^{x=1} \int_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} \sqrt{1 + x^2 + y^2}\,dx\,dy.$$

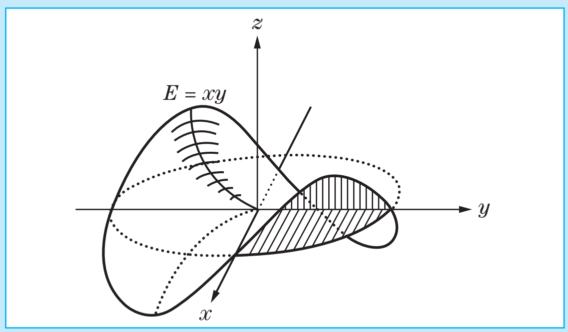


Figure 1.31

The Surface $z = xy$ Above and Below the Unit Circle $x^2 + y^2 = 1$

For the circular geometry plane polar coordinates r, φ are more appropriate, where the radial integral is evaluated by substituting $u = 1 + r^2$ in

$$A = \int_0^1 \sqrt{1+r^2}\, r\, dr \int_0^{2\pi} d\varphi = \pi \int \sqrt{u}\, du = \frac{2\pi}{3}\left|(1+r^2)^{3/2}\right|_0^1 = \frac{2\pi}{3}(2\sqrt{2}-1).$$

■

More examples of line and surface integrals are provided in Chapter 2.

Volume Integrals

Volume integrals are simpler because the volume element $d\tau$ is a scalar quantity.[15] We have

$$\int_v \mathbf{V}\, d\tau = \hat{\mathbf{x}} \int_v V_x\, d\tau + \hat{\mathbf{y}} \int_v V_y\, d\tau + \hat{\mathbf{z}} \int_v V_z\, d\tau, \tag{1.110}$$

again reducing the vector integral to a vector sum of scalar integrals.

If the vector

$$\mathbf{V} = V_\rho(\rho, \varphi, z)\hat{\boldsymbol{\rho}} + V_\varphi(\rho, \varphi, z)\hat{\boldsymbol{\varphi}} + V_z(\rho, \varphi, z)\hat{\mathbf{z}}$$

and its components are given in cylindrical coordinates $x = \rho\cos\varphi$, $y = \rho\sin\varphi$ with volume element $d\tau = \rho\, d\rho\, d\varphi\, dz$, the volume integral

$$\int_v \mathbf{V}\, d\tau = \hat{\mathbf{z}} \int_v V_z\, d\tau + \iiint (V_\rho \hat{\boldsymbol{\rho}} + V_\varphi \hat{\boldsymbol{\varphi}})\rho\, d\rho\, d\varphi\, dz$$

involves integrals over the varying unit vectors of the polar coordinates. To reduce them to scalar integrals, we need to expand the polar coordinate unit vectors in Cartesian unit vectors as follows. Dividing the plane coordinates by ρ, we find

$$\hat{\boldsymbol{\rho}} = \frac{1}{\rho}(x, y) = (\cos\varphi, \sin\varphi) = \hat{\mathbf{x}}\cos\varphi + \hat{\mathbf{y}}\sin\varphi.$$

Differentiating $\hat{\boldsymbol{\rho}}^2 = 1$, we see from $0 = \frac{d\hat{\rho}^2}{d\varphi} = 2\hat{\boldsymbol{\rho}} \cdot \frac{d\hat{\rho}}{d\varphi}$ that

$$\frac{d\hat{\boldsymbol{\rho}}}{d\varphi} = -\hat{\mathbf{x}}\sin\varphi + \hat{\mathbf{y}}\cos\varphi = \hat{\boldsymbol{\varphi}}$$

is perpendicular to $\hat{\boldsymbol{\rho}}$ and a unit vector; therefore, it is equal to $\hat{\boldsymbol{\varphi}}$. Substituting these expressions into the second integral yields the final result

$$\int_v \mathbf{V}\, d\tau = \hat{\mathbf{z}} \int_v V_z\, d\tau + \hat{\mathbf{x}} \iiint [V_\rho \cos\varphi - V_\varphi \sin\varphi]\rho\, d\rho\, d\varphi\, dz$$
$$+ \hat{\mathbf{y}} \iiint [V_\rho \sin\varphi + V_\varphi \cos\varphi]\rho\, d\rho\, d\varphi\, dz. \tag{1.111}$$

The terms in brackets are the Cartesian components V_x, V_y expressed in plane polar coordinates.

[15] Frequently, the symbols d^3r and d^3x are used to denote a volume element in coordinate (xyz or $x_1x_2x_3$) space.

In spherical polar coordinates, all of the unit vectors depend on the coordinates, none can be pulled out of the integrals, and all have to be expanded in Cartesian unit vectors. This task of rewriting Eq. (1.110) is left as an exercise.

EXAMPLE 1.9.6 **Volume of Rotated Gaussian** Rotate the Gaussian $y = \exp(-x^2)$ about the z-axis leading to $z = \exp(-x^2 - y^2)$. Then the volume in the polar (cylindrical) coordinates appropriate for the geometry is given by

$$V = \int_{r=0}^{\infty} \int_{\varphi=0}^{2\pi} \int_{z=0}^{e^{-r^2}} r\,dr\,d\varphi\,dz = 2\pi \int_0^{\infty} re^{-r^2} dr = \pi \int_0^{\infty} e^{-u} du = \pi,$$

upon substituting $\exp(-x^2 - y^2) = \exp(-r^2)$, $dx\,dy = r\,dr\,d\varphi$, $u = r^2$, and $du = 2rdr$. ■

Integral Definitions of Gradient, Divergence, and Curl

One interesting and significant application of our surface and volume integrals is their use in developing alternate definitions of our differential relations. We find

$$\nabla\varphi = \lim_{\int d\tau \to 0} \frac{\int \varphi\, d\boldsymbol{\sigma}}{\int d\tau}, \tag{1.112}$$

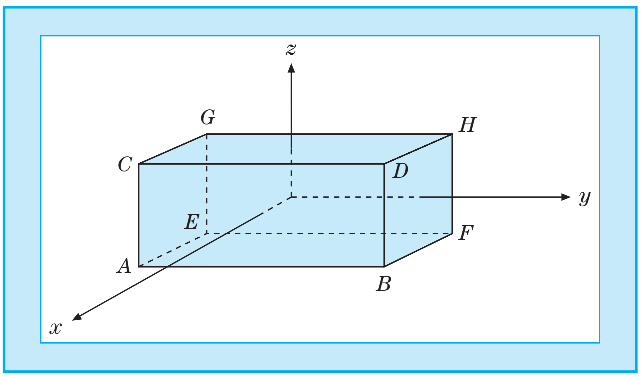
$$\nabla \cdot \mathbf{V} = \lim_{\int d\tau \to 0} \frac{\int \mathbf{V} \cdot d\boldsymbol{\sigma}}{\int d\tau}, \tag{1.113}$$

$$\nabla \times \mathbf{V} = \lim_{\int d\tau \to 0} \frac{\int d\boldsymbol{\sigma} \times \mathbf{V}}{\int d\tau}. \tag{1.114}$$

In these three equations, $\int d\tau$ is the volume of a small region of space and $d\boldsymbol{\sigma}$ is the vector area element of this volume. The identification of Eq. (1.113) as the divergence of $\mathbf{V}$ was carried out in Section 1.6. Here, we show that Eq. (1.112) is consistent with our earlier definition of $\nabla\varphi$ [Eq. (1.64)]. For simplicity, we choose $d\tau$ to be the differential volume $dx\,dy\,dz$ (Fig. 1.32). This

Figure 1.32

Differential Rectangular Parallelepiped (Origin at Center)



time, we place the origin at the geometric center of our volume element. The area integral leads to six integrals, one for each of the six faces. Remembering that $d\boldsymbol{\sigma}$ is outward, $d\boldsymbol{\sigma} \cdot \hat{\mathbf{x}} = -\,|d\boldsymbol{\sigma}|$ for surface $EFHG$, and $+\,|d\boldsymbol{\sigma}|$ for surface $ABDC$, we have

$$\begin{aligned}\int \varphi\, d\boldsymbol{\sigma} = -\hat{\mathbf{x}} \int_{EFHG} \left(\varphi - \frac{\partial\varphi}{\partial x}\frac{dx}{2}\right) dy\,dz + \hat{\mathbf{x}} \int_{ABDC} \left(\varphi + \frac{\partial\varphi}{\partial x}\frac{dx}{2}\right) dy\,dz \\ -\hat{\mathbf{y}} \int_{AEGC} \left(\varphi - \frac{\partial\varphi}{\partial y}\frac{dy}{2}\right) dx\,dz + \hat{\mathbf{y}} \int_{BFHD} \left(\varphi + \frac{\partial\varphi}{\partial y}\frac{dy}{2}\right) dx\,dz \\ -\hat{\mathbf{z}} \int_{ABFE} \left(\varphi - \frac{\partial\varphi}{\partial z}\frac{dz}{2}\right) dx\,dy + \hat{\mathbf{z}} \int_{CDHG} \left(\varphi + \frac{\partial\varphi}{\partial z}\frac{dz}{2}\right) dx\,dy.\end{aligned}$$

Using the first two terms of a Maclaurin expansion, we evaluate each integrand at the origin with a correction included to correct for the displacement ($\pm dx/2$, etc.) of the center of the face from the origin. Having chosen the total volume to be of differential size ($\int d\tau = dx\,dy\,dz$), we drop the integral signs on the right and obtain

$$\int \varphi\, d\boldsymbol{\sigma} = \left(\hat{\mathbf{x}}\frac{\partial\varphi}{\partial x} + \hat{\mathbf{y}}\frac{\partial\varphi}{\partial y} + \hat{\mathbf{z}}\frac{\partial\varphi}{\partial z}\right) dx\,dy\,dz. \tag{1.115}$$

Dividing by

$$\int d\tau = dx\,dy\,dz,$$

we verify Eq. (1.112).

This verification has been oversimplified in ignoring other correction terms beyond the first derivatives. These additional terms, which are introduced in Section 5.6 when the Taylor expansion is developed, vanish in the limit

$$\int d\tau \to 0\,(dx \to 0,\, dy \to 0,\, dz \to 0).$$

This, of course, is the reason for specifying in Eqs. (1.112)–(1.114) that this limit be taken. Verification of Eq. (1.114) follows these same lines, using a differential volume $d\tau = dx\,dy\,dz$.

EXERCISES

1.9.1 Find the potential for the electric field generated by a charge q at the origin. Normalize the potential to zero at spatial infinity.

1.9.2 Determine the gravitational field of the earth taken to be spherical and of uniform mass density. Punch out a concentric spherical cavity and show that the field is zero inside it. Show that the field is constant if the cavity is not concentric.

1.9.3 Evaluate

$$\frac{1}{3}\int_s \mathbf{r}\cdot d\boldsymbol{\sigma}$$

over the unit cube defined by the point (0, 0, 0) and the unit intercepts on the positive x-, y-, and z-axes. Note that (a) $\mathbf{r}\cdot d\boldsymbol{\sigma}$ is zero for three of the surfaces, and (b) each of the three remaining surfaces contributes the same amount to the integral.

1.9.4 Show by expansion of the surface integral that

$$\lim_{\int d\tau\to 0}\frac{\int_s d\boldsymbol{\sigma}\times\mathbf{V}}{\int d\tau} = \boldsymbol{\nabla}\times\mathbf{V}.$$

Hint. Choose the volume to be a differential volume, $dx\,dy\,dz$.

1.10 Gauss's Theorem

Here, we derive a useful relation between a surface integral of a vector and the volume integral of the divergence of that vector. Let us assume that the vector $\mathbf{V}$ and its first derivatives are continuous over the simply connected region (without holes) of interest. Then, **Gauss's theorem** states that

$$\int_S \mathbf{V}\cdot d\boldsymbol{\sigma} = \int_V \boldsymbol{\nabla}\cdot\mathbf{V}\,d\tau. \tag{1.116a}$$

In words, the surface integral of a vector over a closed surface equals the volume integral of the divergence of that vector integrated over the volume enclosed by the surface.

Imagine that volume V is subdivided into an arbitrarily large number of tiny (differential) parallelepipeds. For each parallelepiped,

$$\sum_{\text{six surfaces}} \mathbf{V}\cdot d\boldsymbol{\sigma} = \boldsymbol{\nabla}\cdot\mathbf{V}\,d\tau \tag{1.116b}$$

from the analysis of Section 1.6, Eq. (1.75), with $\rho\mathbf{v}$ replaced by $\mathbf{V}$. The summation is over the six faces of the parallelepiped. Summing over all parallelepipeds, we find that the $\mathbf{V}\cdot d\boldsymbol{\sigma}$ terms cancel (pairwise) for all **interior** faces; only the contributions of the **exterior** surfaces survive (Fig. 1.33). Analogous to the definition of a Riemann integral as the limit of a sum, we take the limit as the number of parallelepipeds approaches infinity ($\to\infty$) and the dimensions of each approach zero ($\to 0$):

$$\sum_{\text{exterior surfaces}} \mathbf{V}\underset{\downarrow}{\cdot}\, d\boldsymbol{\sigma} = \sum_{\text{volumes}} \boldsymbol{\nabla}\underset{\downarrow}{\cdot}\,\mathbf{V}\,d\tau$$

$$\int_S \mathbf{V}\cdot d\boldsymbol{\sigma} = \int_V \boldsymbol{\nabla}\cdot\mathbf{V}\,d\tau.$$

The result is Eq. (1.116a), Gauss's theorem.

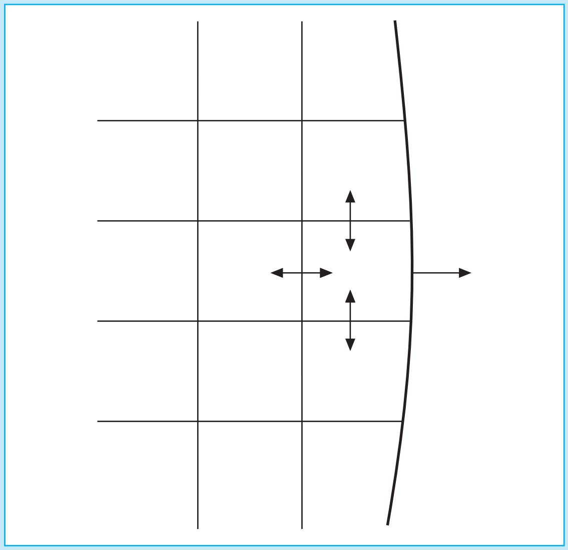


Figure 1.33

Exact Cancellation of $\mathbf{V} \cdot d\boldsymbol{\sigma}$'s on Interior Surfaces. No Cancellation on the Exterior Surface

From a physical standpoint, Eq. (1.75) has established $\nabla \cdot \mathbf{V}$ as the net outflow of field per unit volume. The volume integral then gives the total net outflow. However, the surface integral $\int \mathbf{V} \cdot d\boldsymbol{\sigma}$ is just another way of expressing this same quantity, which is the equality, Gauss's theorem.

Biographical Data

Gauss, Carl Friedrich. Gauss, a German mathematician, physicist, and astronomer, was born in Brunswick in 1777 and died in Göttingen in 1855. He was an infant prodigy in mathematics whose education was directed and financed by the Duke of Brunswick. As a teenager, he proved that regular n-polygons can be constructed in Euclidean geometry provided n is a Fermat prime number such as 3, 5, 17, and 257, a major advance in geometry since antiquity. This feat convinced him to stay in mathematics and give up the study of foreign languages. For his Ph.D., he proved the fundamental theorem of algebra, avoiding the then controversial complex numbers he had used to discover it. In his famous treatise *Disquisitiones Arithmetica* on number theory, he first proved the quadratic reciprocity theorem and originated the terse style and rigor of mathematical proofs as a series of logical steps, discarding any trace of the original heuristic ideas used in the discovery and checks of examples. Not surprisingly, he hated teaching. He is considered by many as the greatest mathematician of all times and was the last to provide major contributions to all then existing branches of mathematics. As the founder of differential geometry, he developed the intrinsic properties of surfaces, such as curvature, which later motivated B. Riemann to develop the geometry of metric spaces, the mathematical foundation of Einstein's General Relativity. In astronomy (for the orbit of the asteroid Ceres), he developed the method of least squares for fitting curves to data. In physics, he developed potential theory, and the unit of the magnetic induction is named after him in honor of his measurements and development of units in physics.

Green's Theorem

A frequently useful corollary of Gauss's theorem is a relation known as Green's theorem. If u and v are two scalar functions, we have the identities

$$\nabla \cdot (u\,\nabla v) = u\nabla \cdot \nabla v + (\nabla u) \cdot (\nabla v), \tag{1.117}$$

$$\nabla \cdot (v\,\nabla u) = v\nabla \cdot \nabla u + (\nabla v) \cdot (\nabla u), \tag{1.118}$$

which follow from the product rule of differentiation. Subtracting Eq. (1.118) from Eq. (1.117), integrating over a volume (u, v, and their derivatives, assumed continuous), and applying Eq. (1.116a) (Gauss's theorem), we obtain

$$\int_V (u\nabla \cdot \nabla v - v\nabla \cdot \nabla u)\, d\tau = \int_S (u\nabla v - v\nabla u) \cdot d\boldsymbol{\sigma}. \tag{1.119}$$

This is **Green's theorem**, which states that the antisymmetric Laplacian of a pair of functions integrated over a simply connected volume (no holes) is equivalent to the antisymmetric gradient of the pair integrated over the bounding surface. An alternate form of Green's theorem derived from Eq. (1.117) alone is

$$\int_S u\nabla v \cdot d\boldsymbol{\sigma} = \int_V u\nabla \cdot \nabla v\, d\tau + \int_V \nabla u \cdot \nabla v\, d\tau. \tag{1.120}$$

Finally, Gauss's theorem may also be extended to tensors (see Section 2.11).

Biographical Data

Green, George. Green, an English mathematician, was born in Nottingham in 1793 and died near Nottingham in 1841. He studied Laplace's papers in Cambridge and developed potential theory in electrodynamics.

EXERCISES

1.10.1 If $\mathbf{B} = \nabla \times \mathbf{A}$, show that

$$\int_S \mathbf{B} \cdot d\boldsymbol{\sigma} = 0$$

for any closed surface S. State this in words. If symbolic software is available, check this for a typical vector potential and specific surfaces, such as a sphere or cube.

1.10.2 Over some volume V, let ψ be a solution of Laplace's equation (with the derivatives appearing there continuous). Prove that the integral over any closed surface in V of the normal derivative of ψ ($\partial\psi/\partial n$, or $\nabla\psi \cdot \mathbf{n}$) will be zero.

1.10.3 In analogy to the integral definition of gradient, divergence, and curl of Section 1.10, show that

$$\nabla^2\varphi = \lim_{\int d\tau \to 0} \frac{\int \nabla\varphi \cdot d\boldsymbol{\sigma}}{\int d\tau}.$$

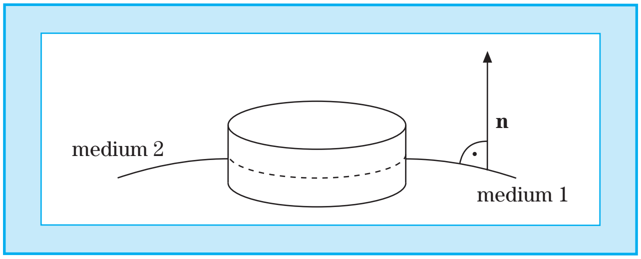
1.10.4 The electric displacement vector $\mathbf{D}$ satisfies the Maxwell equation $\nabla \cdot \mathbf{D} = \rho$, where ρ is the charge density (per unit volume). At the boundary between two media there is a surface charge density σ (per unit area). Show that a boundary condition for $\mathbf{D}$ is

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \sigma,$$

where $\mathbf{n}$ is a unit vector normal to the surface and out of medium 1.
Hint. Consider a thin pillbox as shown in Fig. 1.34.

Figure 1.34

Pillbox



1.10.5 From Eq. (1.77) and Example 1.6.1, with $\mathbf{V}$ the electric field $\mathbf{E}$ and f the electrostatic potential φ, show that

$$\int \rho\varphi\, d\tau = \varepsilon_0 \int E^2\, d\tau.$$

This corresponds to a three-dimensional integration by parts.
Hint. $\mathbf{E} = -\nabla\varphi$, $\nabla \cdot \mathbf{E} = \rho/\varepsilon_0$. You may assume that φ vanishes at large r at least as fast as r^{-1}.

1.10.6 The creation of a **localized** system of steady electric currents (current density $\mathbf{J}$) and magnetic fields may be shown to require an amount of work

$$W = \frac{1}{2}\int \mathbf{H} \cdot \mathbf{B}\, d\tau.$$

Transform this into

$$W = \frac{1}{2}\int \mathbf{J} \cdot \mathbf{A}\, d\tau,$$

where $\mathbf{A}$ is the magnetic vector potential, $\nabla \times \mathbf{A} = \mathbf{B}$.

Hint. In Maxwell's equations, take the displacement current term $\partial \mathbf{D}/\partial t = 0$ and explain why using Ohm's law. If the fields and currents are localized, a bounding surface may be taken far enough out so that the integrals of the fields and currents over the surface yield zero.

1.11 Stokes's Theorem

Gauss's theorem relates the volume integral of a derivative of a function to an integral of the function over the closed surface bounding the volume. Here, we consider an analogous relation between the surface integral of a derivative of a function and the line integral of the function, the path of integration being the perimeter bounding the surface.

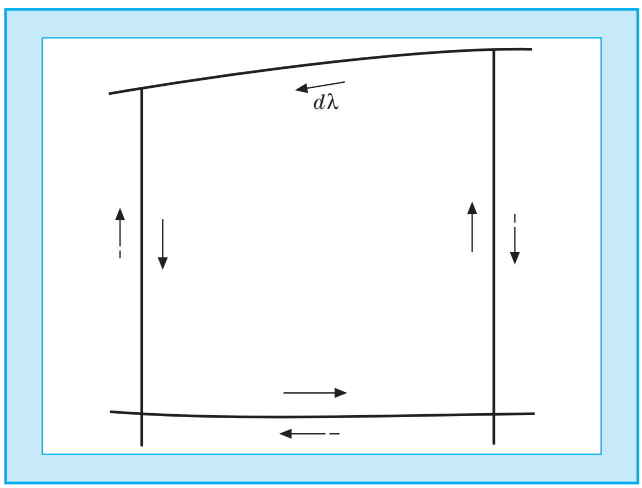
Let us take the surface and subdivide it into a network of arbitrarily small rectangles. In Section 1.7, we showed that the circulation about such a differential rectangle (in the xy-plane) is $\nabla \times \mathbf{V}|_z \, dx\,dy$. From Eq. (1.85) applied to **one** differential rectangle,

$$\sum_{\text{four sides}} \mathbf{V} \cdot d\boldsymbol{\lambda} = \nabla \times \mathbf{V} \cdot d\boldsymbol{\sigma}. \qquad (1.121)$$

We sum over all the little rectangles as in the definition of a Riemann integral. The surface contributions [right-hand side of Eq. (1.121)] are added together. The line integrals [left-hand side of Eq. (1.121)] of all **interior** line segments cancel identically. Only the line integral around the perimeter survives (Fig. 1.35). Taking the usual limit as the number of rectangles approaches

Figure 1.35

Exact Cancellation on Interior Paths; No Cancellation on the Exterior Path



infinity while $dx \to 0$, $dy \to 0$, we have

$$\sum_{\substack{\text{exterior line}\\\text{segments}}} \mathbf{V}\cdot d\boldsymbol{\lambda} = \sum_{\text{rectangles}} \nabla\times\mathbf{V}\cdot d\boldsymbol{\sigma}$$

$$\downarrow \qquad\qquad \downarrow$$

$$\oint \mathbf{V}\cdot d\boldsymbol{\lambda} = \int_S \nabla\times\mathbf{V}\cdot d\boldsymbol{\sigma}. \tag{1.122}$$

This is Stokes's theorem. The surface integral on the right is over the surface bounded by the perimeter or contour for the line integral on the left. The direction of the vector representing the area is out of the paper plane toward the reader if the direction of traversal around the contour for the line integral is in the positive mathematical sense as shown in Fig. 1.35.

This demonstration of Stokes's theorem is limited by the fact that we used a Maclaurin expansion of $\mathbf{V}(x, y, z)$ in establishing Eq. (1.85) in Section 1.7. Actually, we need only demand that the curl of $\mathbf{V}(x, y, z)$ exists and that it be integrable over the surface. Stokes's theorem obviously applies to an open, simply connected surface. It is possible to consider a closed surface as a limiting case of an open surface with the opening (and therefore the perimeter) shrinking to zero. This is the point of Exercise 1.11.4.

As a special case of Stokes's theorem, consider the curl of a two-dimensional vector field $\mathbf{V} = (V_1(x, y), V_2(x, y), 0)$. The curl $\nabla\times\mathbf{V} = (0, 0, \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y})$ so

$$\int_S \nabla\times\mathbf{V}\cdot\hat{\mathbf{z}}\,dx\,dy = \int_S \left(\frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y}\right) dx\,dy = \int_C \mathbf{V}\cdot d\mathbf{r} = \int_C (V_1 dx + V_2 dy),$$

where the curve C is the boundary of the simply connected surface S that is integrated in the positive mathematical sense (anticlockwise). This relation is sometimes also called Green's theorem. In Chapter 6, we shall use it to prove Cauchy's theorem for analytic functions.

EXAMPLE 1.11.1

Area as a Line Integral For the two-dimensional Stokes's theorem, we first choose $\mathbf{V} = x\hat{\mathbf{y}}$, which gives the area $S = \int_S dx\,dy = \int_C x\,dy$, and for $\mathbf{V} = (y\hat{\mathbf{x}})$ we get similarly $S = \int_S dx\,dy = -\int_C y\,dx$. Adding both results gives the area

$$S = \frac{1}{2}\int_C (x\,dy - y\,dx). \qquad ■$$

We can use Stokes's theorem to derive Oersted's and Faraday's laws from two of Maxwell's equations and vice versa, thus recognizing that the former are an integrated form of the latter.

EXAMPLE 1.11.2

Oersted's and Faraday's Laws Consider the magnetic field generated by a long wire that carries a stationary current I (Fig. 1.36). Starting from Maxwell's differential law $\nabla\times\mathbf{H} = \mathbf{J}$ [Eq. (1.97c); with Maxwell's displacement current $\partial\mathbf{D}/\partial t = 0$ for a stationary current case by Ohm's law], we integrate over a closed area S perpendicular to and surrounding the wire and apply Stokes's

Figure 1.36

Oersted's Law for a Long Wire Carrying a Current

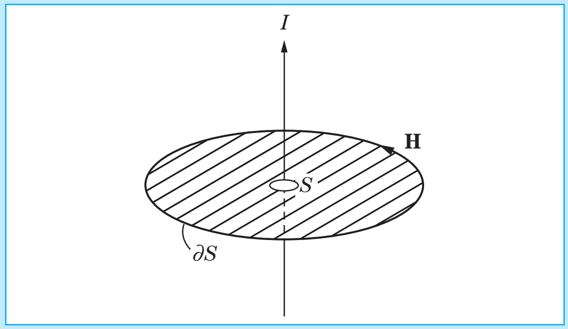
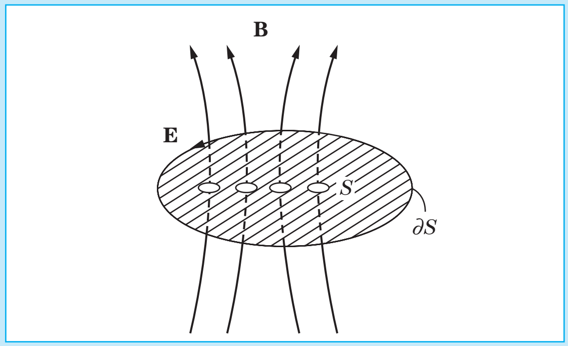


Figure 1.37

Faraday's Induction Law Across a Magnetic Induction Field



theorem to get

$$I = \int_S \mathbf{J} \cdot d\boldsymbol{\sigma} = \int_S (\boldsymbol{\nabla} \times \mathbf{H}) \cdot d\boldsymbol{\sigma} = \oint_{\partial S} \mathbf{H} \cdot d\mathbf{r},$$

which is Oersted's law. Here, the line integral is along ∂S, the closed curve surrounding the cross section area S.

Similarly, we can integrate Maxwell's equation for $\boldsymbol{\nabla} \times \mathbf{E}$ [Eq. (1.97d)] to yield Faraday's induction law. Imagine moving a closed loop (∂S) of wire (of area S) across a magnetic induction field $\mathbf{B}$ (Fig. 1.37). At a fixed moment of time we integrate Maxwell's equation and use Stokes's theorem, yielding

$$\int_{\partial S} \mathbf{E} \cdot d\mathbf{r} = \int_S (\boldsymbol{\nabla} \times \mathbf{E}) \cdot d\boldsymbol{\sigma} = -\frac{d}{dt} \int_S \mathbf{B} \cdot d\boldsymbol{\sigma} = -\frac{d\Phi}{dt},$$

which is Faraday's law. The line integral on the left-hand side represents the voltage induced in the wire loop, whereas the right-hand side is the change with time of the magnetic flux Φ through the moving surface S of the wire. ■

SUMMARY

Both Stokes's and Gauss's theorems are of tremendous importance in a wide variety of problems involving vector calculus in electrodynamics, where they allow us to derive the local form of Maxwell's differential equations from the global (integral) form of the experimental laws. An indication of their power and versatility may be obtained from the exercises in Sections 1.10 and 1.11 and the development of potential theory in Section 1.12.

Biographical Data

Stokes, Sir George Gabriel. Stokes, a British mathematician and physicist, was born in Skreen, Ireland, in 1819 and died in Cambridge in 1903. Son of a clergyman, his talent for mathematics was already evident in school. In 1849, he became Lucasian professor at Cambridge, the chair Isaac Newton once held and currently held by S. Hawking. In 1885, he became president of the Royal Society. He is known for the theory of viscous fluids, with practical applications to the motion of ships in water. He demonstrated his vision by hailing Joule's work early on and recognizing X-rays as electromagnetic radiation. He received the Rumford and Copley medals of the Royal Society and served as a member of Parliament for Cambridge University in 1887–1892.

EXERCISES

1.11.1 The calculation of the magnetic moment of a current loop leads to the line integral

$$\oint \mathbf{r} \times d\mathbf{r}.$$

(a) Integrate around the perimeter of a current loop (in the xy-plane) and show that the scalar magnitude of this line integral is twice the area of the enclosed surface.

(b) The perimeter of an ellipse is described by $\mathbf{r} = \hat{\mathbf{x}}a\cos\theta + \hat{\mathbf{y}}b\sin\theta$. From part (a), show that the area of the ellipse is πab.

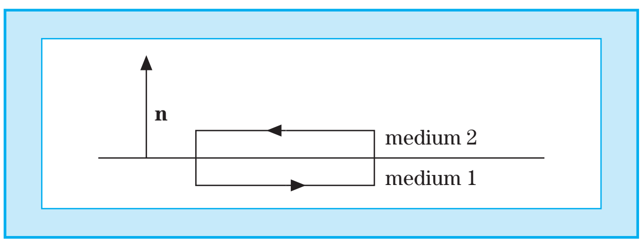
1.11.2 In steady state, the magnetic field $\mathbf{H}$ satisfies the Maxwell equation $\nabla \times \mathbf{H} = \mathbf{J}$, where $\mathbf{J}$ is the current density (per square meter). At the boundary between two media there is a surface current density $\mathbf{K}$ (perimeter). Show that a boundary condition on $\mathbf{H}$ is

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K},$$

where $\mathbf{n}$ is a unit vector normal to the surface and out of medium 1.
Hint. Consider a narrow loop perpendicular to the interface as shown in Fig. 1.38.

Figure 1.38

Loop Contour



1.11.3 A magnetic induction $\mathbf{B}$ is generated by electric current in a ring of radius R. Show that the **magnitude** of the vector potential $\mathbf{A}$ ($\mathbf{B} = \nabla \times \mathbf{A}$) at the ring is

$$|\mathbf{A}| = \frac{\Phi}{2\pi R},$$

where Φ is the total magnetic flux passing through the ring.
Note. $\mathbf{A}$ is tangential to the ring.

1.11.4 Prove that

$$\int_S \nabla \times \mathbf{V} \cdot d\boldsymbol{\sigma} = 0$$

if S is a closed surface.

1.11.5 Prove that

$$\oint u\nabla v \cdot d\boldsymbol{\lambda} = -\oint v\nabla u \cdot d\boldsymbol{\lambda}.$$

1.11.6 Prove that

$$\oint u\nabla v \cdot d\boldsymbol{\lambda} = \int_S (\nabla u) \times (\nabla v) \cdot d\boldsymbol{\sigma}.$$

1.12 Potential Theory

Scalar Potential

This section formulates the conditions under which a force field $\mathbf{F}$ is conservative. From a mathematical standpoint, it is a practice session of typical applications of Gauss's and Stokes's theorems in physics.

If a force in a given **simply connected region** of space V (i.e., no holes in it) can be expressed as the negative gradient of a scalar function φ,

$$\mathbf{F} = -\nabla\varphi, \tag{1.123}$$

we call φ a scalar potential that describes the force by one function instead of three, which is a significant simplification. A scalar potential is only determined up to an additive constant, which can be used to adjust its value at infinity (usually zero) or at some other point. The force $\mathbf{F}$ appearing as the negative gradient of a single-valued scalar potential is labeled a **conservative** force. We want to know when a scalar potential function exists. To answer this question, we establish two other relations as equivalent to Eq. (1.123):

$$\nabla \times \mathbf{F} = 0 \tag{1.124}$$

and

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0, \tag{1.125}$$

for every closed path in our simply connected region V. We proceed to show that each of these three equations implies the other two. Let us start with

$$\mathbf{F} = -\boldsymbol{\nabla}\varphi. \tag{1.126}$$

Then

$$\boldsymbol{\nabla} \times \mathbf{F} = -\boldsymbol{\nabla} \times \boldsymbol{\nabla}\varphi = 0 \tag{1.127}$$

by Eq. (1.92), or Eq. (1.123) implies Eq. (1.124). Turning to the line integral, we have

$$\oint \mathbf{F} \cdot d\mathbf{r} = -\oint \boldsymbol{\nabla}\varphi \cdot d\mathbf{r} = -\oint d\varphi \tag{1.128}$$

using Eq. (1.58). Now $d\varphi$ integrates to give φ. Because we have specified a closed loop, the end points coincide and we get zero for every closed path in our region S for which Eq. (1.123) holds. It is important to note the restriction that the potential be single-valued and that Eq. (1.123) hold for **all** points in S. This derivation may also apply to a scalar magnetic potential as long as no net current is encircled. As soon as we choose a path in space that encircles a net current, the scalar magnetic potential ceases to be single-valued and our analysis no longer applies because V is no longer simply connected.

Continuing this demonstration of equivalence, let us assume that Eq. (1.125) holds. If $\oint \mathbf{F} \cdot d\mathbf{r} = 0$ for all paths in S, the value of the integral joining two distinct points A and B is independent of the path (Fig. 1.39). Our premise is that

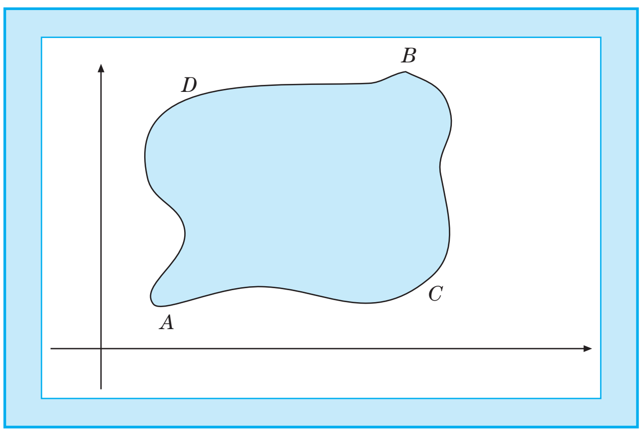
$$\oint_{ACBDA} \mathbf{F} \cdot d\mathbf{r} = 0. \tag{1.129}$$

Therefore,

$$\oint_{ACB} \mathbf{F} \cdot d\mathbf{r} = -\int_{BDA} \mathbf{F} \cdot d\mathbf{r} = \int_{ADB} \mathbf{F} \cdot d\mathbf{r}, \tag{1.130}$$

Figure 1.39

Possible Paths for Doing Work



reversing the sign by reversing the direction of integration. Physically, this means that the work done in going from A to B is independent of the path and that the work done in going around a closed path is zero. This is the reason for labeling such a force conservative: Energy is conserved.

With the result shown in Eq. (1.130), we have the work done dependent only on the end points A and B. That is,

$$\text{Work done by force} = \int_A^B \mathbf{F} \cdot d\mathbf{r} = \varphi(A) - \varphi(B). \tag{1.131}$$

Equation (1.131) defines a scalar potential (strictly speaking, the difference in potential between points A and B) and provides a means of calculating the potential. If point B is taken as a variable such as (x, y, z), then differentiation with respect to x, y, and z will recover Eq. (1.123).

The choice of sign on the right-hand side is arbitrary. The choice here is made to achieve agreement with Eq. (1.123) and to ensure that water will run downhill rather than uphill. For points A and B separated by a length $d\mathbf{r}$, Eq. (1.131) becomes

$$\mathbf{F} \cdot d\mathbf{r} = -d\varphi = -\nabla\varphi \cdot d\mathbf{r}. \tag{1.132}$$

This may be rewritten

$$(\mathbf{F} + \nabla\varphi) \cdot d\mathbf{r} = 0, \tag{1.133}$$

and since $d\mathbf{r} \neq 0$ is arbitrary, Eq. (1.126) must follow. If

$$\oint \mathbf{F} \cdot d\mathbf{r} = 0, \tag{1.134}$$

we may obtain Eq. (1.123) by using Stokes's theorem [Eq. (1.122)]:

$$\oint \mathbf{F} \cdot d\mathbf{r} = \int \nabla \times \mathbf{F} \cdot d\boldsymbol{\sigma}. \tag{1.135}$$

If we take the path of integration to be the perimeter of an arbitrary differential area $d\boldsymbol{\sigma}$, the integrand in the surface integral must vanish. Hence, Eq. (1.125) implies Eq. (1.123).

Finally, if $\nabla \times \mathbf{F} = 0$, we need only reverse our statement of Stokes's theorem [Eq. (1.135)] to derive Eq. (1.125). Then, by Eqs. (1.131)–(1.133) the initial statement $\mathbf{F} = -\nabla\varphi$ is derived. The triple equivalence is illustrated in Fig. 1.40.

SUMMARY

A single-valued scalar potential function φ exists if and only if $\mathbf{F}$ is irrotational so that the work done around every closed loop is zero. The gravitational and electrostatic force fields given by Eq. (1.88) are irrotational and therefore conservative. Gravitational and electrostatic scalar potentials exist. Now, by calculating the work done [Eq. (1.131)], we proceed to determine three potentials (Fig. 1.41).

Figure 1.40

Equivalent Formulations of a Conservative Force

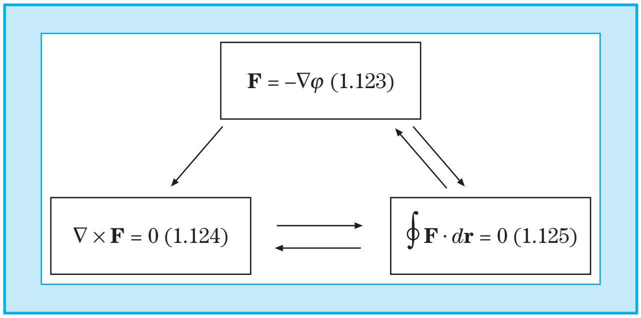
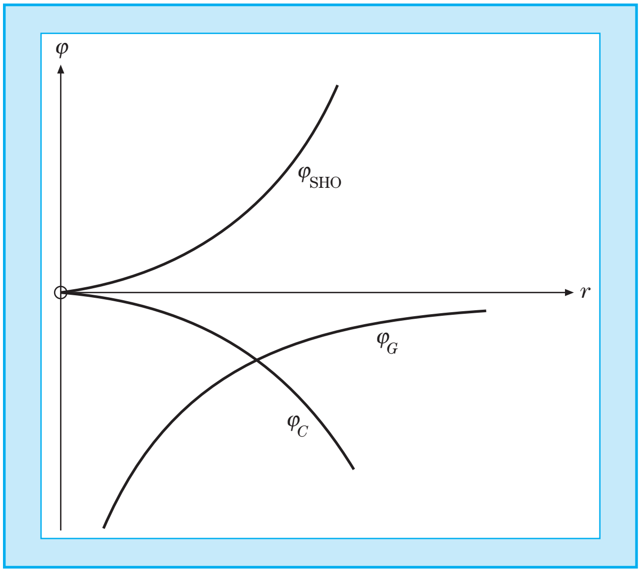


Figure 1.41

Potential Energy Versus Distance (Gravitational, Centrifugal, and Simple Harmonic Oscillator)



EXAMPLE 1.12.1

Centrifugal Potential Calculate the scalar potential for the **centrifugal** force per unit mass, $\mathbf{F}_C = \omega^2\mathbf{r}$, radially **outward**. Physically, the centrifugal force is what you feel when on a merry-go-round. Proceeding as in Example 1.9.3, but integrating from the origin outward and taking $\varphi_C(0) = 0$, we have

$$\varphi_C(r) = -\int_0^r \mathbf{F}_C \cdot d\mathbf{r} = -\frac{\omega^2 r^2}{2}.$$

If we reverse signs, taking $\mathbf{F}_{\text{SHO}} = -k\mathbf{r}$, we obtain $\varphi_{\text{SHO}} = \frac{1}{2}kr^2$, the simple harmonic oscillator potential.

The gravitational, centrifugal, and simple harmonic oscillator potentials are shown in Fig. 1.41. Clearly, the simple harmonic oscillator yields stability and describes a restoring force. The centrifugal potential describes an unstable situation. ■

SUMMARY

When a vector **B** is solenoidal, a vector potential **A** exists such that $\mathbf{B} = \nabla \times \mathbf{A}$. **A** is undetermined to within an additive gradient of a scalar function. This is similar to the arbitrary zero of a potential, due to an additive constant of the scalar potential.

In many problems, the magnetic vector potential **A** will be obtained from the current distribution that produces the magnetic induction **B**. This means solving Poisson's (vector) equation (see Exercise 1.13.4).

EXERCISES

1.12.1 The usual problem in classical mechanics is to calculate the motion of a particle given the potential. For a uniform density (ρ_0), nonrotating massive sphere, Gauss's law (Section 1.10) leads to a gravitational force on a unit mass m_0 at a point r_0 produced by the attraction of the mass at $r \le r_0$. The mass at $r > r_0$ contributes nothing to the force.

(a) Show that $\mathbf{F}/m_0 = -(4\pi G\rho_0/3)\mathbf{r}$, $0 \le r \le a$, where a is the radius of the sphere.

(b) Find the corresponding gravitational potential, $0 \le r \le a$.

(c) Imagine a vertical hole running completely through the center of the earth and out to the far side. Neglecting the rotation of the earth and assuming a uniform density $\rho_0 = 5.5\,\mathrm{g/cm^3}$, calculate the nature of the motion of a particle dropped into the hole. What is its period?

Note. $\mathbf{F} \propto \mathbf{r}$ is actually a very poor approximation. Because of varying density, the approximation $\mathbf{F} =$ constant, along the outer half of a radial line, and $\mathbf{F} \propto \mathbf{r}$, along the inner half, is much closer.

1.12.2 The origin of the Cartesian coordinates is at the earth's center. The moon is on the z-axis, a fixed distance R away (center-to-center distance). The tidal force exerted by the moon on a particle at the earth's surface (point x, y, z) is given by

$$F_x = -GMm\frac{x}{R^3}, \quad F_y = -GMm\frac{y}{R^3}, \quad F_z = +2GMm\frac{z}{R^3}.$$

Find the potential that yields this tidal force.

ANS. $-\dfrac{GMm}{R^3}\left(z^2 - \dfrac{1}{2}x^2 - \dfrac{1}{2}y^2\right)$

In terms of the Legendre polynomials of Chapter 11, this becomes

$$-\frac{GMm}{R^3}r^2 P_2(\cos\theta).$$

1.12.3 Vector **B** is formed by the product of two gradients

$$\mathbf{B} = (\nabla u) \times (\nabla v),$$

where u and v are scalar functions.

(a) Show that $\mathbf{B}$ is solenoidal.
(b) Show that

$$\mathbf{A} = \frac{1}{2}(u\,\nabla v - v\,\nabla u)$$

is a vector potential for $\mathbf{B}$ in that

$$\mathbf{B} = \nabla \times \mathbf{A}.$$

1.12.4 The magnetic induction $\mathbf{B}$ is related to the magnetic vector potential $\mathbf{A}$ by $\mathbf{B} = \nabla \times \mathbf{A}$. By Stokes's theorem,

$$\int \mathbf{B} \cdot d\boldsymbol{\sigma} = \oint \mathbf{A} \cdot d\mathbf{r}.$$

Show that each side of this equation is invariant under the **gauge transformation**, $\mathbf{A} \to \mathbf{A} + \nabla\Lambda$, where Λ is an arbitrary scalar function.
Note. Take the function Λ to be single-valued.

1.12.5 With $\mathbf{E}$ as the electric field and $\mathbf{A}$ as the magnetic vector potential, show that $[\mathbf{E} + \partial\mathbf{A}/\partial t]$ is irrotational and that we may therefore write

$$\mathbf{E} = -\nabla\varphi - \frac{\partial\mathbf{A}}{\partial t}.$$

1.12.6 The total force on a charge q moving with velocity $\mathbf{v}$ is

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

Using the scalar and vector potentials, show that

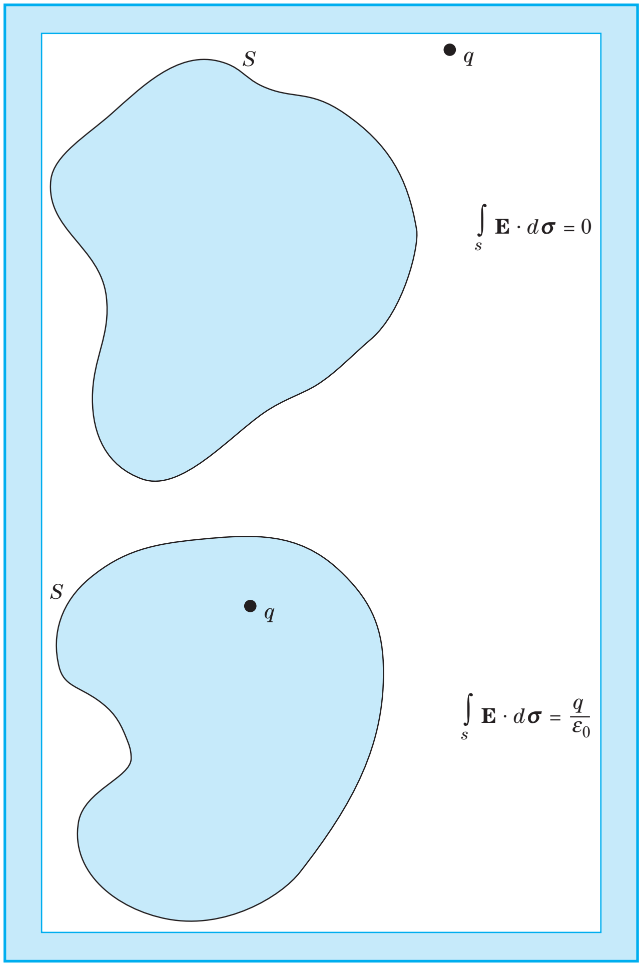
$$\mathbf{F} = q\left[-\nabla\varphi - \frac{d\mathbf{A}}{dt} + \nabla(\mathbf{A} \cdot \mathbf{v})\right].$$

Note that we now have a total time derivative of $\mathbf{A}$ in place of the partial derivative of Exercise 1.12.5.

1.12.7 A planet of mass m moves on a circular orbit of radius r around a star in an attractive gravitational potential $V = kr^n$. Find the conditions on the exponent n for the orbit to be stable.
Note. You can set $k = -GmM$, where M is the mass of the star, and use classical mechanics. Einstein's General Relativity gives $n = -1$, whereas in Newton's gravitation the Kepler laws are needed in addition to determining that $n = -1$.

Figure 1.42

Gauss's Law



1.13 Gauss's Law and Poisson's Equation

Gauss's Law

Consider a point electric charge q at the origin of our coordinate system. This produces an electric field $\mathbf{E}$[16] given by

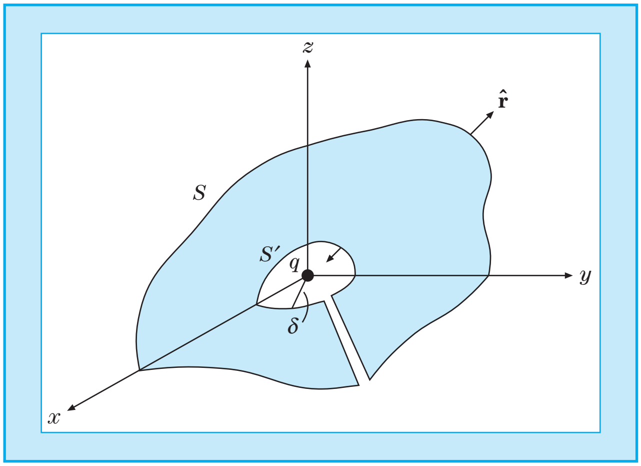
$$\mathbf{E} = \frac{q\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2}. \tag{1.136}$$

We now derive Gauss's law, which states that the surface integral in Fig. 1.42 is q/ε_0 if the closed surface S includes the origin (where q is located) and zero

[16] The electric field $\mathbf{E}$ is defined as the force per unit charge on a small stationary test charge q_t: $\mathbf{E} = \mathbf{F}/q_t$. From Coulomb's law, the force on q_t due to q is $\mathbf{F} = (qq_t/4\pi\varepsilon_0)(\hat{\mathbf{r}}/r^2)$. When we divide by q_t, Eq. (1.136) follows.

Figure 1.43

Exclusion of the Origin



if the surface does not include the origin. The surface S is any closed surface; it need not be spherical.

Using Gauss's theorem [Eq. (1.116a)] (and neglecting the scale factor $q/4\pi\varepsilon_0$), we obtain

$$\int_S \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}}{r^2} = \int_V \boldsymbol{\nabla}\cdot\left(\frac{\hat{\mathbf{r}}}{r^2}\right) d\tau = 0 \tag{1.137}$$

by Example 1.6.1, provided the surface S does not include the origin, where the integrands are not defined. This proves the second part of Gauss's law.

The first part, in which the surface S must include the origin, may be handled by surrounding the origin with a small sphere S' of radius δ (Fig. 1.43). So that there will be no question as to what is inside and what is outside, imagine the volume outside the outer surface S and the volume inside surface $S'(r < \delta)$ connected by a small hole. This joins surfaces S and S', combining them into one **single, simply connected closed surface**. Because the radius of the imaginary hole may be made vanishingly small, there is no additional contribution to the surface integral. The inner surface is deliberately chosen to be spherical so that we will be able to integrate over it. Gauss's theorem now applies to the volume between S and S' without any difficulty. We have

$$\int_S \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}}{r^2} + \int_{S'} \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}'}{\delta^2} = 0. \tag{1.138}$$

We may evaluate the second integral for $d\boldsymbol{\sigma}' = -\hat{\mathbf{r}}\delta^2\, d\Omega$, in which $d\Omega$ is an element of solid angle. The minus sign appears because we agreed in Section 1.9 to have the positive normal $\hat{\mathbf{r}}'$ **outward** from the volume. In this case, the outward $\hat{\mathbf{r}}'$ is in the negative radial direction, $\hat{\mathbf{r}}' = -\hat{\mathbf{r}}$. By integrating over all angles, we have

$$\int_{S'} \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}'}{\delta^2} = -\int_{S'} \frac{\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}\delta^2 d\Omega}{\delta^2} = -4\pi, \tag{1.139}$$

independent of the radius δ. With the constants from Eq. (1.136), this results in

$$\int_S \mathbf{E} \cdot d\boldsymbol{\sigma} = \frac{q}{4\pi\varepsilon_0} 4\pi = \frac{q}{\varepsilon_0}, \tag{1.140}$$

completing the proof of Gauss's law. Notice that although the surface S may be spherical, it **need not** be spherical.

Going a bit further, we consider a distributed charge so that

$$q = \int_V \rho \, d\tau. \tag{1.141}$$

Equation (1.140) still applies, with q now interpreted as the total distributed charge enclosed by surface S:

$$\int_S \mathbf{E} \cdot d\boldsymbol{\sigma} = \int_V \frac{\rho}{\varepsilon_0} d\tau. \tag{1.142}$$

Using Gauss's theorem, we have

$$\int_V \boldsymbol{\nabla} \cdot \mathbf{E} \, d\tau = \int_V \frac{\rho}{\varepsilon_0} d\tau. \tag{1.143}$$

Since our volume is completely arbitrary, the integrands must be equal or

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \tag{1.144}$$

one of Maxwell's equations. If we reverse the argument, Gauss's law follows immediately from Maxwell's equation by integration.

Poisson's Equation

Replacing $\mathbf{E}$ by $-\boldsymbol{\nabla}\varphi$, Eq. (1.144) becomes

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\varphi = -\frac{\rho}{\varepsilon_0}, \tag{1.145}$$

which is Poisson's equation. We know a solution,

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')d\tau'}{|\mathbf{r} - \mathbf{r}'|},$$

from generalizing a sum of Coulomb potentials for discrete charges in electrostatics to a continuous charge distribution.

For the condition $\rho = 0$ this reduces to an even more famous equation, the **Laplace equation**.

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\varphi = 0. \tag{1.146}$$

We encounter Laplace's equation frequently in discussing various curved coordinate systems (Chapter 2) and the special functions of mathematical physics that appear as its solutions in Chapter 11.

From direct comparison of the Coulomb electrostatic force law and Newton's law of universal gravitation,

$$\mathbf{F}_E = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r^2} \hat{\mathbf{r}}, \qquad \mathbf{F}_G = -G \frac{m_1 m_2}{r^2} \hat{\mathbf{r}}.$$

All of the potential theory of this section therefore applies equally well to gravitational potentials. For example, the gravitational Poisson equation is

$$\nabla \cdot \nabla \varphi = +4\pi G \rho, \tag{1.147}$$

with ρ now a mass density.

> ***Biographical Data***
> **Poisson, Siméon Denis.** Poisson, a French mathematician, was born in Pithiviers, France in 1781 and died in Paris in 1840. He studied mathematics at the Ecole Polytechnique under Laplace and Lagrange, whom he so impressed with his talent that he became professor there in 1802. He contributed to their celestial mechanics, Fourier's heat theory, and probability theory, among others.

EXERCISES

1.13.1 Develop Gauss's law for the two-dimensional case in which

$$\varphi = -q \frac{\ln \rho}{2\pi \varepsilon_0}, \quad \mathbf{E} = -\nabla \varphi = q \frac{\hat{\boldsymbol{\rho}}}{2\pi \varepsilon_0 \rho},$$

where q is the charge at the origin or the line charge per unit length if the two-dimensional system is a unit thickness slice of a three-dimensional (circular cylindrical) system. The variable ρ is measured radially outward from the line charge. $\hat{\boldsymbol{\rho}}$ is the corresponding unit vector (see Section 2.2). If graphical software is available, draw the potential and field for the $q/2\pi\varepsilon_0 = 1$ case.

1.13.2 (a) Show that Gauss's law follows from Maxwell's equation

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$$

by integrating over a closed surface. Here, ρ is the charge density.

(b) Assuming that the electric field of a point charge q is spherically symmetric, show that Gauss's law implies the Coulomb inverse square expression

$$\mathbf{E} = \frac{q\hat{\mathbf{r}}}{4\pi \varepsilon_0 r^2}.$$

1.13.3 Show that the value of the electrostatic potential φ at any point P is equal to the average of the potential over any spherical surface centered on P. There are no electric charges on or within the sphere.
Hint. Use Green's theorem [Eq. (1.119)], with $u^{-1} = r$, the distance from P, and $v = \varphi$.

1.13.4 Using Maxwell's equations, show that for a system (steady current) the magnetic vector potential $\mathbf{A}$ satisfies a vector Poisson equation

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J},$$

provided we require $\nabla \cdot \mathbf{A} = 0$ in Coulomb gauge.

1.14 Dirac Delta Function

From Example 1.6.1 and the development of Gauss's law in Section 1.13,

$$\int \nabla \cdot \nabla \left(\frac{1}{r}\right) d\tau = -\int \nabla \cdot \left(\frac{\hat{\mathbf{r}}}{r^2}\right) d\tau = \begin{cases} -4\pi \\ 0 \end{cases}, \tag{1.148}$$

depending on whether the integration includes the origin $\mathbf{r} = 0$ or not. This result may be conveniently expressed by introducing the Dirac delta function,

$$\nabla^2 \left(\frac{1}{r}\right) = -4\pi\,\delta\,(\mathbf{r}) = -4\pi\,\delta(x)\delta(y)\delta(z). \tag{1.149}$$

This Dirac delta function is **defined** by its assigned properties

$$\delta\,(x) = 0, \qquad x \neq 0 \tag{1.150}$$

$$f\,(0) = \int_{-\infty}^{\infty} f\,(x)\,\delta(x)\,dx, \tag{1.151}$$

where $f(x)$ is any well-behaved function and the integration includes the origin. As a special case of Eq. (1.151),

$$\int_{-\infty}^{\infty} \delta(x)\,dx = 1. \tag{1.152}$$

From Eq. (1.151), $\delta(x)$ must be an infinitely high, infinitely thin spike at $x = 0$, as in the description of an impulsive force or the charge density for a point charge.[17] The problem is that **no such function exists** in the usual sense of function. However, the crucial property in Eq. (1.151) can be developed rigorously as the limit of a **sequence** of functions, a distribution. For example, the delta function may be approximated by the sequences of functions in n for $n \to \infty$ [Eqs. (1.153)–(1.156) and Figs. 1.44–1.47]:

$$\delta_n(x) = \begin{cases} 0, & x < -\frac{1}{2n} \\ n, & -\frac{1}{2n} < x < \frac{1}{2n} \\ 0, & x > \frac{1}{2n} \end{cases} \tag{1.153}$$

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2x^2) \tag{1.154}$$

$$\delta_n(x) = \frac{n}{\pi} \cdot \frac{1}{1 + n^2x^2} \tag{1.155}$$

$$\delta_n(x) = \frac{\sin nx}{\pi x} = \frac{1}{2\pi} \int_{-n}^{n} e^{ixt}\,dt. \tag{1.156}$$

[17] The delta function is frequently invoked to describe very short-range forces such as nuclear forces. It also appears in the normalization of continuum wave functions of quantum mechanics.

Figure 1.44

δ Sequence Function Eq. (1.153)

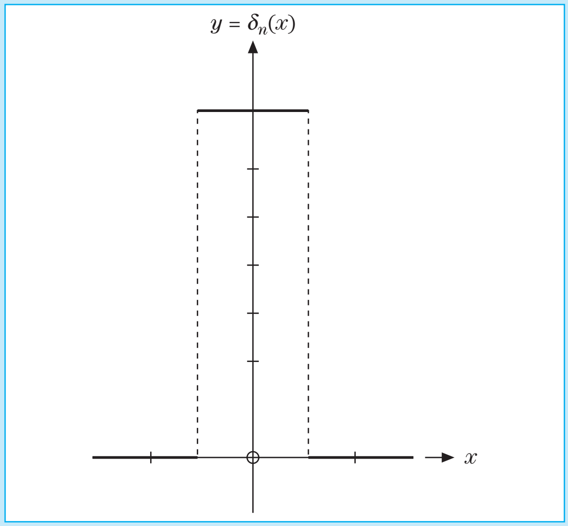
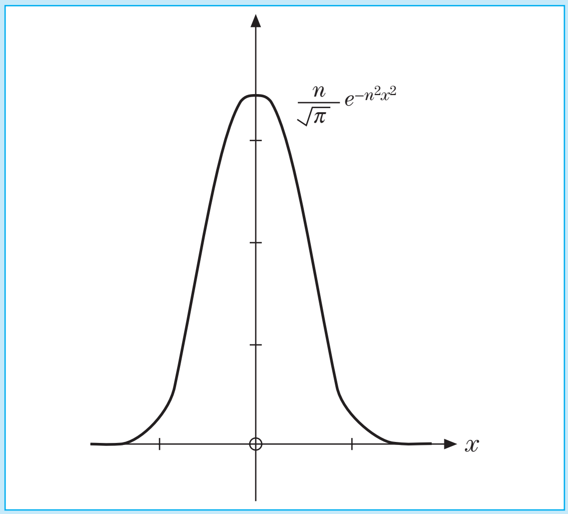


Figure 1.45

δ Sequence Function Eq. (1.154)



EXAMPLE 1.14.1

Let us evaluate $\int_{-\pi}^{\pi} \cos x \delta(x) dx = \cos 0 = 1$ using the sequence of Eq. (1.153). We find

$$\int_{-1/2n}^{1/2n} n \cos x \, dx = n \sin x|_{-1/2n}^{1/2n} = n \left(\sin\left(\frac{1}{2n}\right) - \sin\left(-\frac{1}{2n}\right)\right)$$

$$= 2n \sin \frac{1}{2n} = 2n \left(\frac{1}{2n} + O(1/n^3) \right) \to 1 \text{ for } n \to \infty.$$

Figure 1.46

δ Sequence Function Eq. (1.155)

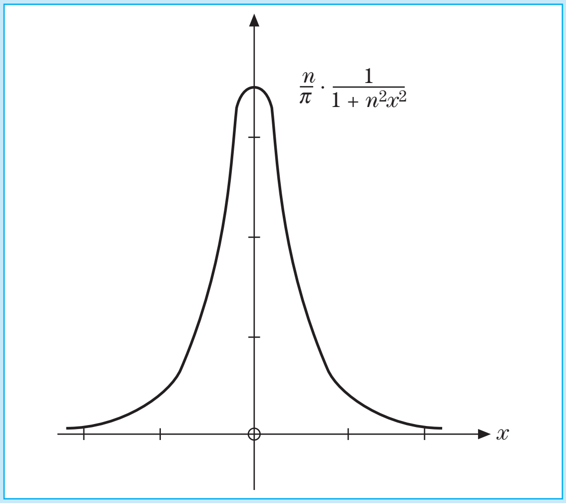
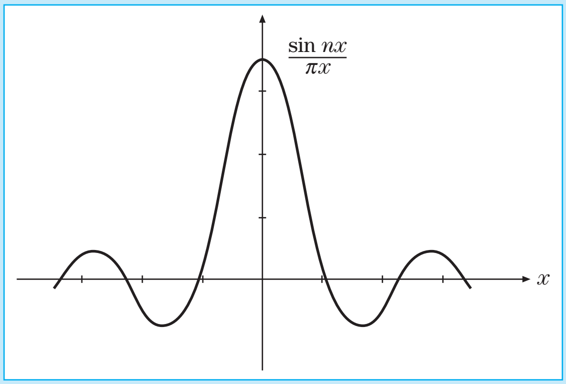


Figure 1.47

δ Sequence Function Eq. (1.156)



Notice how the integration limits change in the first step. Similarly, $\int_{-\pi}^{\pi} \sin x\delta(x) \cdot dx = \sin 0 = 0$. We could have used Eq. (1.155) instead,

$$\frac{n}{\pi}\int_{-\pi}^{\pi}\frac{\cos x\,dx}{1+n^2x^2} = \frac{n}{\pi}\int_{-\pi}^{\pi}\frac{1-x^2/2+\cdots}{1+n^2x^2}dx = \frac{n}{\pi}\int_{-\pi}^{\pi}\frac{dx}{1+n^2x^2}$$
$$= \frac{1}{\pi}\int_{-n\pi}^{n\pi}\frac{dy}{1+y^2} = \frac{1}{\pi}[\arctan(n\pi) - \arctan(-n\pi)]$$
$$= \frac{2}{\pi}\arctan(n\pi) \to \frac{2}{\pi}\frac{\pi}{2} = 1, \quad \text{for } n \to \infty,$$

by keeping just the first term of the power expansion of $\cos x$. Again, we could have changed the integration limits to $\pm\pi/n$ in the first step for all terms with positive powers of x because the denominator is so large, except close to $x = 0$ for large n. This explains why the higher order terms of the $\cos x$ power series do not contribute. ■

These approximations have varying degrees of usefulness. Equation (1.153) is useful in providing a simple derivation of the integral property [Eq. (1.151)]. Equation (1.154) is convenient to differentiate. Its derivatives lead to the Hermite polynomials. Equation (1.156) is particularly useful in Fourier analysis and in its applications to quantum mechanics. In the theory of Fourier series, Eq. (1.156) often appears (modified) as the Dirichlet kernel:

$$\delta_n(x) = \frac{1}{2\pi}\,\frac{\sin[(n+\frac{1}{2})x]}{\sin(\frac{1}{2}x)}. \tag{1.157}$$

In using these approximations in Eq. (1.151) and later, we assume that $f(x)$ is integrable—it offers no problems at large x.

For most physical purposes such approximations are quite adequate. From a mathematical standpoint, the situation is still unsatisfactory: The limits

$$\lim_{n\to\infty} \delta_n(x)$$

do not exist.

A way out of this difficulty is provided by the theory of distributions. Recognizing that Eq. (1.151) is the fundamental property, we focus our attention on it rather than on $\delta(x)$. Equations (1.153)–(1.156), with $n = 1, 2, 3, \ldots,$ may be interpreted as **sequences** of normalized functions:

$$\int_{-\infty}^{\infty} \delta_n(x)\,dx = 1. \tag{1.158}$$

The sequence of integrals has the limit

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x) f(x)\,dx = f(0). \tag{1.159}$$

Note that Eq. (1.158) is the limit of a sequence of integrals. Again, the limit of $\delta_n(x)$, $n \to \infty$, does not exist. [The limits for all four forms of $\delta_n(x)$ diverge at $x = 0$.]

We may treat $\delta(x)$ consistently in the form

$$\int_{-\infty}^{\infty} \delta(x) f(x)\,dx = \lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x) f(x)\,dx. \tag{1.160}$$

$\delta(x)$ is labeled a distribution (not a function) defined by the sequences $\delta_n(x)$ as indicated in Eq. (1.158). We might emphasize that the integral on the left-hand side of Eq. (1.160) is not a Riemann integral.[18] It is a limit.

[18]It can be treated as a Stieltjes integral if desired. $\delta(x)\,dx$ is replaced by $du(x)$, where $u(x)$ is the Heaviside step function.

This distribution $\delta(x)$ is only one of an infinity of possible distributions, but it is the one we are interested in because of Eq. (1.151).

From these sequences of functions, we see that Dirac's delta function must be even in x, $\delta(-x) = \delta(x)$.

Let us now consider a detailed application of the Dirac delta function to a single charge and illustrate the singularity of the electric field at the origin.

EXAMPLE 1.14.2

Total Charge inside a Sphere Consider the total electric flux $\oint \mathbf{E} \cdot d\boldsymbol{\sigma}$ out of a sphere of radius R around the origin surrounding n charges e_j located at the points $\mathbf{r}_j$ with $r_j < R$ (i.e., inside the sphere). The electric field strength $\mathbf{E} = -\nabla\varphi(\mathbf{r})$, where the potential

$$\varphi = \sum_{j=1}^{n} \frac{e_j}{|\mathbf{r} - \mathbf{r}_j|} = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r'$$

is the sum of the Coulomb potentials generated by each charge and the total charge density is $\rho(\mathbf{r}) = \sum_j e_j \delta(\mathbf{r} - \mathbf{r}_j)$. The delta function is used here as an abbreviation of a pointlike density. Now we use Gauss's theorem for

$$\oint \mathbf{E} \cdot d\boldsymbol{\sigma} = -\oint \nabla\varphi \cdot d\boldsymbol{\sigma} = -\int \nabla^2 \varphi d\tau = \int \frac{\rho(\mathbf{r})}{\varepsilon_0} d\tau = \frac{\sum_j e_j}{\varepsilon_0}$$

in conjunction with the differential form of Gauss's law $\nabla \cdot \mathbf{E} = -\rho/\varepsilon_0$ and $\sum_j e_j \int \delta(\mathbf{r} - \mathbf{r}_j) d\tau = \sum_j e_j$. ■

The integral property [Eq. (1.151)] is useful in cases in which the argument of the delta function is a function $g(x)$ with **simple zeros** on the real axis, which leads to the rules

$$\delta(ax) = \frac{1}{a}\delta(x), \qquad a > 0, \tag{1.161}$$

$$\delta(g(x)) = \sum_{\substack{a, \\ g(a)=0, \\ g'(a)\neq 0}} \frac{\delta(x-a)}{|g'(a)|}. \tag{1.162}$$

To obtain Eq. (1.161) we change the integration variable in

$$\int_{-\infty}^{\infty} f(x)\delta(ax)dx = \frac{1}{a}\int_{-\infty}^{\infty} f\left(\frac{y}{a}\right)\delta(y)dy = \frac{1}{a} f(0)$$

and apply Eq. (1.151). To prove Eq. (1.162), we decompose the integral

$$\int_{-\infty}^{\infty} f(x)\delta(g(x))dx = \sum_a \int_{a-\varepsilon}^{a+\varepsilon} f(x)\delta((x-a)g'(a))dx \tag{1.163}$$

into a sum of integrals over small intervals containing the first-order zeros of $g(x)$. In these intervals, $g(x) \approx g(a) + (x-a)g'(a) = (x-a)g'(a)$. Using Eq. (1.161) on the right-hand side of Eq. (1.163), we obtain the integral of Eq. (1.162).

EXAMPLE 1.14.3

Evaluate $I \equiv \int_{-\infty}^{\infty} f(x)\delta(x^2-2)dx$ Because the zeros of the argument of the delta function, $x^2 = 2$, are $x = \pm\sqrt{2}$, we can write the integral as a sum of two contributions:

$$
\begin{aligned}
I &= \int_{\sqrt{2}-\epsilon}^{\sqrt{2}+\epsilon} \delta(x-\sqrt{2}) \frac{f(x)dx}{\frac{d(x^2-2)}{dx}|_{x=\sqrt{2}}} dx + \int_{-\sqrt{2}-\epsilon}^{-\sqrt{2}+\epsilon} \delta(x+\sqrt{2}) \frac{f(x)dx}{\frac{d(x^2-2)}{dx}|_{x=-\sqrt{2}}} \\
&= \int_{\sqrt{2}-\epsilon}^{\sqrt{2}+\epsilon} \delta(x-\sqrt{2}) \frac{f(x)dx}{2\sqrt{2}} + \int_{-\sqrt{2}-\epsilon}^{-\sqrt{2}+\epsilon} \delta(x+\sqrt{2}) \frac{f(x)dx}{2\sqrt{2}} \\
&= \frac{f(\sqrt{2}) + f(-\sqrt{2})}{2\sqrt{2}}. \quad \blacksquare
\end{aligned}
$$

This example is good training for the following one.

EXAMPLE 1.14.4

Phase Space In the scattering theory of relativistic particles using Feynman diagrams, we encounter the following integral over energy of the scattered particle (we set the velocity of light $c = 1$):

$$
\begin{aligned}
\int d^4p\delta(p^2-m^2)f(p) &\equiv \int d^3p \int dp_0 \delta\left(p_0^2 - \mathbf{p}^2 - m^2\right) f(p) \\
&= \int_{E>0} \frac{d^3pf(E,\mathbf{p})}{2\sqrt{m^2+\mathbf{p}^2}} - \int_{E<0} \frac{d^3pf(E,\mathbf{p})}{2\sqrt{m^2+\mathbf{p}^2}},
\end{aligned}
$$

where we have used Eq. (1.162) at the zeros $E = \pm\sqrt{m^2+\mathbf{p}^2}$ of the argument of the delta function. The physical meaning of $\delta(p^2-m^2)$ is that the particle of mass m and four-momentum $p^\mu = (p_0, \mathbf{p})$ is on its mass shell because $p^2 = m^2$ is equivalent to $E = \pm\sqrt{m^2+\mathbf{p}^2}$. Thus, the on-mass-shell volume element in momentum space is the Lorentz invariant $\frac{d^3p}{2E}$, in contrast to the nonrelativistic d^3p of momentum space. The fact that a negative energy occurs is a peculiarity of relativistic kinematics that is related to the antiparticle. ■

Using integration by parts we can also **define the derivative** $\delta'(x)$ of the Dirac delta function by the relation

$$\int_{-\infty}^{\infty} f(x)\delta'(x-x')\,dx = -\int_{-\infty}^{\infty} f'(x)\delta(x-x')\,dx = -f'(x'). \tag{1.164}$$

It should be understood that our Dirac delta function has significance only as part of an integrand. Thus, the Dirac delta function is often regarded as a linear operator: $\delta(x-x_0)$ operates on $f(x)$ and yields $f(x_0)$:

$$\mathcal{L}(x_0)f(x) \equiv \int_{-\infty}^{\infty} \delta(x-x_0)f(x)\,dx = f(x_0). \tag{1.165}$$

It may also be classified as a linear mapping or simply as a generalized function. Shifting our singularity to the point $x = x'$, we write the Dirac delta function as $\delta(x - x')$. Equation (1.151) becomes

$$\int_{-\infty}^{\infty} f(x)\delta(x - x')\,dx = f(x'). \tag{1.166}$$

As a description of a singularity at $x = x'$, the Dirac delta function may be written as $\delta(x - x')$ or as $\delta(x' - x)$. Expanding to three dimensions and using spherical polar coordinates, we obtain

$$\begin{aligned} f(0) &= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} f(\mathbf{r})\delta\,(\mathbf{r})\,r^2\,dr\sin\theta\,d\theta\,d\varphi \\ &= \iiint_{-\infty}^{\infty} f(x,y,z)\delta\,(x)\,\delta\,(y)\,\delta\,(z)\;dx\,dy\,dz, \end{aligned}$$

$$\int_0^{\infty}\frac{\delta(r)}{r^2}r^2dr\int_{-1}^{1}\delta(\cos\theta)\,d\cos\theta\int_0^{2\pi}\delta(\varphi)\,d\varphi = 1, \tag{1.167}$$

where each one-dimensional integral is equal to unity. This corresponds to a singularity (or source) at the origin. Again, if our source is at $\mathbf{r} = \mathbf{r}_1$, Eq. (1.167) generalizes to

$$\iiint f(\mathbf{r}_2)\delta\,(\mathbf{r}_2 - \mathbf{r}_1)\,r_2^2dr_2\sin\theta_2\,d\theta_2\,d\varphi_2 = f(\mathbf{r}_1), \tag{1.168}$$

where

$$\int_0^{\infty}\frac{\delta(r_2 - r_1)}{r_2^2}r_2^2dr_2\int_{-1}^{1}\delta(\cos\theta_2 - \cos\theta_1)d\cos\theta_2\int_0^{2\pi}\delta(\varphi_2 - \varphi_1)d\varphi_2 = 1.$$

SUMMARY

We use $\delta\,(x)$ frequently and call it the Dirac delta function—for historical reasons.[19] Remember that it is not really a function. It is essentially a shorthand notation, defined implicitly as the limit of integrals in a sequence, $\delta_n\,(x)$, according to Eq. (1.160).

Biographical Data

Dirac, Paul Adrien Maurice. Dirac, an English physicist, was born in Bristol in 1902 and died in Bristol in 1984. He obtained a degree in electrical engineering at Bristol and obtained his Ph.D. in mathematical physics in 1926 at Cambridge. By 1932, he was Lucasian professor, like Stokes, the chair Newton once held. In the 1920s, he advanced quantum mechanics, became one of the founders of quantum field theory, and, in 1928, discovered his relativistic equation for the electron that predicted antiparticles for which he was awarded the Nobel prize in 1933.

[19] Dirac introduced the delta function to quantum mechanics. Actually, the delta function can be traced back to Kirchhoff, 1882. For further details, see M. Jammer (1966). *The Conceptual Development of Quantum Mechanics*, p. 301. McGraw-Hill, New York.

EXERCISES

1.14.1 Let

$$\delta_n(x) = \begin{cases} 0, & x < -\frac{1}{2n} \\ n, & -\frac{1}{2n} < x < \frac{1}{2n}, \\ 0, & \frac{1}{2n} < x. \end{cases}$$

Show that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\delta_n(x)\, dx = f(0),$$

assuming that $f(x)$ is continuous at $x = 0$.

1.14.2 Verify that the sequence $\delta_n(x)$, based on the function

$$\delta_n(x) = \begin{cases} 0, & x < 0, \\ ne^{-nx} & x > 0, \end{cases}$$

is a delta sequence [satisfying Eq. (1.159)]. Note that the singularity is at $+0$, the positive side of the origin.
Hint. Replace the upper limit (∞) by c/n, where c is large but finite, and use the mean value theorem of integral calculus.

1.14.3 For

$$\delta_n(x) = \frac{n}{\pi} \cdot \frac{1}{1 + n^2x^2},$$

[Eq. (1.155)], show that

$$\int_{-\infty}^{\infty} \delta_n(x)\, dx = 1.$$

1.14.4 Demonstrate that $\delta_n = \sin nx/\pi x$ is a delta distribution by showing that

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x)\frac{\sin nx}{\pi x}\, dx = f(0).$$

Assume that $f(x)$ is continuous at $x = 0$ and vanishes as $x \to \pm\infty$.
Hint. Replace x by y/n and take $\lim n \to \infty$ **before** integrating.

1.14.5 Fejer's method of summing series is associated with the function

$$\delta_n(t) = \frac{1}{2\pi n}\left[\frac{\sin(nt/2)}{\sin(t/2)}\right]^2.$$

Show that $\delta_n(t)$ is a delta distribution in the sense that

$$\lim_{n\to\infty} \frac{1}{2\pi n} \int_{-\infty}^{\infty} f(t)\left[\frac{\sin(nt/2)}{\sin(t/2)}\right]^2 dt = f(0).$$

1.14.6 Using the Gaussian delta sequence (δ_n), Eq. (1.154), show that

$$x\frac{d}{dx}\delta(x) = -\delta(x),$$

treating $\delta(x)$ and its derivative as in Eq. (1.151).

1.14.7 Show that

$$\int_{-\infty}^{\infty} \delta'(x) f(x)\, dx = -f'(0).$$

Assume that $f'(x)$ is continuous at $x = 0$.

1.14.8 Prove that

$$\delta(f(x)) = \left| \frac{df(x)}{dx} \right|^{-1} \delta(x - x_0),$$

where x_0 is chosen so that $f(x_0) = 0$ with $df/dx \neq 0$; that is, $f(x)$ has a simple zero at x_0.
Hint. Use $\delta(f)\, df = \delta(x)\, dx$ after explaining why this holds.

1.14.9 Show that in spherical polar coordinates $(r, \cos\theta, \varphi)$ the delta function $\delta(\mathbf{r}_1 - \mathbf{r}_2)$ becomes

$$\frac{1}{r_1^2} \delta(r_1 - r_2) \delta(\cos\theta_1 - \cos\theta_2) \delta(\varphi_1 - \varphi_2).$$

1.14.10 For the finite interval $(-\pi, \pi)$ expand the Dirac delta function $\delta(x - t)$ in a series of sines and cosines: $\sin nx$, $\cos nx$, $n = 0, 1, 2, \ldots$. Note that although these functions are orthogonal, they are not normalized to unity.

1.14.11 In the interval $(-\pi, \pi)$, $\delta_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2)$.

(a) Expand $\delta_n(x)$ as a Fourier cosine series.
(b) Show that your Fourier series agrees with a Fourier expansion of $\delta(x)$ in the limit as $n \to \infty$.
(c) Confirm the delta function nature of your Fourier series by showing that for any $f(x)$ that is finite in the interval $[-\pi, \pi]$ and continuous at $x = 0$,

$$\int_{-\pi}^{\pi} f(x) [\text{Fourier expansion of } \delta_\infty(x)]\, dx = f(0).$$

1.14.12 (a) Expand $\delta_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2 x^2)$ in the interval $(-\infty, \infty)$ as a Fourier integral.
(b) Expand $\delta_n(x) = n \exp(-nx)$ as a Laplace transform.

1.14.13 We may define a sequence

$$\delta_n(x) = \begin{cases} n, & |x| < 1/2n, \\ 0, & |x| > 1/2n. \end{cases}$$

[Eq. (1.153)]. Express $\delta_n(x)$ as a Fourier integral (via the Fourier integral theorem, inverse transform, etc.). Finally, show that we may write

$$\delta(x) = \lim_{n \to \infty} \delta_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx}\, dk.$$

1.14.14 Using the sequence

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} \exp(-n^2x^2),$$

show that

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx}\, dk.$$

Note. Remember that $\delta(x)$ is defined in terms of its behavior as part of an integrand, especially Eq. (1.159).

1.14.15 Derive sine and cosine representations of $\delta(t - x)$.

$$ANS.\quad \frac{2}{\pi}\int_0^{\infty} \sin\omega t \sin\omega x\, d\omega, \quad \frac{2}{\pi}\int_0^{\infty} \cos\omega t \cos\omega x\, d\omega.$$

Additional Reading

Borisenko, A. I., and Taropov, I. E. (1968). *Vector and Tensor Analysis with Applications*. Prentice-Hall, Englewood Cliffs, NJ. Reprinted, Dover, New York (1980).

Davis, H. F., and Snider, A. D. (1995). *Introduction to Vector Analysis*, 7th ed. Allyn & Bacon, Boston.

Kellogg, O. D. (1953). *Foundations of Potential Theory*. Dover, New York. Originally published 1929. The classic text on potential theory.

Lewis, P. E., and Ward, J. P. (1989). *Vector Analysis for Engineers and Scientists*. Addison-Wesley, Reading, MA.

Marion, J. B. (1965). *Principles of Vector Analysis*. Academic Press, New York. A moderately advanced presentation of vector analysis oriented toward tensor analysis. Rotations and other transformations are described with the appropriate matrices.

Spiegel, M. R. (1989). *Vector Analysis*. McGraw-Hill, New York.

Tai, C.-T. (1996). *Generalized Vector and Dyadic Analysis*. Oxford Univ. Press, Oxford.

Wrede, R. C. (1963). *Introduction to Vector and Tensor Analysis*. Wiley, New York. Reprinted, Dover, New York (1972). Fine historical introduction; excellent discussion of differentiation of vectors and applications to mechanics.