CHAPTER 3

VECTOR ANALYSIS

The introductory section on vectors, Section 1.7, identified some basic properties that are universal, in the sense that they occur in a similar fashion in spaces of different dimension. In summary, these properties are (1) vectors can be represented as linear forms, with operations that include addition and multiplication by a scalar, (2) vectors have a commutative and distributive dot product operation that associates a scalar with a pair of vectors and depends on their relative orientations and hence is independent of the coordinate system, and (3) vectors can be decomposed into components that can be identified as projections onto the coordinate directions. In Section 2.2 we found that the components of vectors could be identified as the elements of a **column vector** and that the scalar product of two vectors corresponded to the matrix multiplication of the transpose of one (the transposition makes it a **row vector**) with the column vector of the other.

The current chapter builds on these ideas, mainly in ways that are specific to three-dimensional (3-D) physical space, by (1) introducing a quantity called a **vector cross product** to permit the use of vectors to represent rotational phenomena and volumes in 3-D space, (2) studying the transformational properties of vectors when the coordinate system used to describe them is rotated or subjected to a reflection operation, (3) developing mathematical methods for treating vectors that are defined over a spatial region (vector **fields**), with particular attention to quantities that depend on the spatial variation of the vector field, including vector differential operators and integrals of vector quantities, and (4) extending vector concepts to curvilinear coordinate systems, which are very useful when the symmetry of the coordinate system corresponds to a symmetry of the problem under study (an example is the use of spherical polar coordinates for systems with spherical symmetry).

A key idea of the present chapter is that a quantity that is properly called a **vector** must have the transformation properties that preserve its essential features under coordinate transformation; there exist quantities with direction and magnitude that do not transform appropriately and hence are not vectors. This study of transformation properties will, in a subsequent chapter, ultimately enable us to generalize to related quantities such as tensors.

Mathematical Methods for Physicists. DOI: 10.1016/B978-0-12-384654-9.00003-7

Finally, we note that the methods developed in this chapter have direct application in electromagnetic theory as well as in mechanics, and these connections are explored through the study of examples.

3.1 Review of Basic Properties

In Section 1.7 we established the following properties of vectors:

1. Vectors satisfy an addition law that corresponds to successive displacements that can be represented by arrows in the underlying space. Vector addition is commutative and associative: $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ and $(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$.
2. A vector $\mathbf{A}$ can be multiplied by a scalar k; if $k>0$ the result will be a vector in the direction of $\mathbf{A}$ but with its length multiplied by k; if $k<0$ the result will be in the direction opposite to $\mathbf{A}$ but with its length mutiplied by $|k|$.
3. The vector $\mathbf{A}-\mathbf{B}$ is interpreted as $\mathbf{A}+(-1)\mathbf{B}$, so vector polynomials, e.g., $\mathbf{A}-2\mathbf{B}+3\mathbf{C}$, are well-defined.
4. A vector of unit length in the coordinate direction x_i is denoted $\hat{\mathbf{e}}_i$. An arbitrary vector $\mathbf{A}$ can be written as a sum of vectors along the coordinate directions, as
$$\mathbf{A}=A_1\hat{\mathbf{e}}_1+A_2\hat{\mathbf{e}}_2+\cdots .$$
The A_i are called the components of A, and the operations in Properties 1 to 3 correspond to the component formulas
$$\mathbf{G}=\mathbf{A}-2\mathbf{B}+3\mathbf{C} \quad\Longrightarrow\quad G_i=A_i-2B_i+3C_i, \quad \text{(each i)}.$$
5. The magnitude or length of a vector $\mathbf{A}$, denoted $|\mathbf{A}|$ or A, is given in terms of its components as
$$|\mathbf{A}|=\left(A_1^2+A_2^2+\cdots\right)^{1/2}.$$
6. The dot product of two vectors is given by the formula
$$\mathbf{A}\cdot\mathbf{B}=A_1B_1+A_2B_2+\cdots ;$$
consequences are
$$|\mathbf{A}|^2=\mathbf{A}\cdot\mathbf{A}, \quad \mathbf{A}\cdot\mathbf{B}=|\mathbf{A}|\,|\mathbf{B}|\cos\theta,$$
where θ is the angle between $\mathbf{A}$ and $\mathbf{B}$.
7. If two vectors are perpendicular to each other, their dot product vanishes and they are termed **orthogonal**. The unit vectors of a Cartesian coordinate system are orthogonal:
$$\hat{\mathbf{e}}_i\cdot\hat{\mathbf{e}}_\mathbf{j}=\delta_{ij}, \tag{3.1}$$
where δ_{ij} is the Kronecker delta, Eq. (1.164).
8. The projection of a vector in any direction has an algebraic magnitude given by its dot product with a unit vector in that direction. In particular, the projection of $\mathbf{A}$ on the $\hat{\mathbf{e}}_i$ direction is $A_i\hat{\mathbf{e}}_i$, with
$$A_i=\hat{\mathbf{e}}_i\cdot\mathbf{A}.$$

9. The components of $\mathbf{A}$ in $\mathbb{R}^3$ are related to its direction cosines (cosines of the angles that $\mathbf{A}$ makes with the coordinate axes) by the formulas

$$A_x = A\cos\alpha, \quad A_y = A\cos\beta, \quad A_z = A\cos\gamma,$$

and $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

In Section 2.2 we noted that matrices consisting of a single column could be used to represent vectors. In particular, we found, illustrating for the 3-D space $\mathbb{R}^3$, the following properties.

10. A vector $\mathbf{A}$ can be represented by a single-column matrix $\mathbf{a}$ whose elements are the components of $\mathbf{A}$, as in

$$\mathbf{A} \quad \Longrightarrow \quad \mathbf{a} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}.$$

The rows (i.e., individual elements A_i) of $\mathbf{a}$ are the coefficients of the individual members of the **basis** used to represent $\mathbf{A}$, so the element A_i is associated with the basis unit vector $\hat{\mathbf{e}}_i$.

11. The vector operations of addition and multiplication by a scalar correspond exactly to the operations of the same names applied to the single-column matrices representing the vectors, as illustrated here:

$$\mathbf{G} = \mathbf{A} - 2\mathbf{B} + 3\mathbf{C} \quad \Longrightarrow \quad \begin{pmatrix} G_1 \\ G_2 \\ G_3 \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} - 2\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} + 3\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix}$$

$$= \begin{pmatrix} A_1 - 2B_1 + 3C_1 \\ A_2 - 2B_2 + 3C_2 \\ A_3 - 2B_3 + 3C_3 \end{pmatrix}, \quad \text{or} \quad \mathbf{g} = \mathbf{a} - 2\mathbf{b} + 3\mathbf{c}.$$

It is therefore appropriate to call these single-column matrices **column vectors**.

12. The transpose of the matrix representing a vector $\mathbf{A}$ is a single-row matrix, called a **row vector**:

$$\mathbf{a}^T = (A_1 \;\; A_2 \;\; A_3).$$

The operations illustrated in Property 11 also apply to row vectors.

13. The dot product $\mathbf{A}\cdot\mathbf{B}$ can be evaluated as $\mathbf{a}^T\mathbf{b}$, or alternatively, because $\mathbf{a}$ and $\mathbf{b}$ are real, as $\mathbf{a}^\dagger\mathbf{b}$. Moreover, $\mathbf{a}^T\mathbf{b} = \mathbf{b}^T\mathbf{a}$.

$$\mathbf{A}\cdot\mathbf{B} = \mathbf{a}^T\mathbf{b} = (A_1 \;\; A_2 \;\; A_3)\begin{pmatrix} B_1 \\ B_2 \\ B_3 \end{pmatrix} = A_1B_1 + A_2B_2 + A_3B_3.$$

3.2 VECTORS IN 3-D SPACE

We now proceed to develop additional properties for vectors, most of which are applicable only for vectors in 3-D space.

Vector or Cross Product

A number of quantities in physics are related to angular motion or the torque required to cause angular acceleration. For example, **angular momentum** about a point is defined as having a magnitude equal to the distance r from the point times the component of the linear momentum $\mathbf{p}$ perpendicular to $\mathbf{r}$—the component of $\mathbf{p}$ causing angular motion (see Fig. 3.1). The direction assigned to the angular momentum is that perpendicular to both $\mathbf{r}$ and $\mathbf{p}$, and corresponds to the axis about which angular motion is taking place. The mathematical construction needed to describe angular momentum is the **cross product**, defined as

$$\mathbf{C} = \mathbf{A} \times \mathbf{B} = (AB \sin\theta)\hat{\mathbf{e}}_c. \tag{3.2}$$

Note that $\mathbf{C}$, the result of the cross product, is stated to be a vector, with a magnitude that is the product of the magnitudes of $\mathbf{A}$, $\mathbf{B}$ and the sine of the angle $\theta \leq \pi$ between $\mathbf{A}$ and $\mathbf{B}$. The direction of $\mathbf{C}$, i.e., that of $\hat{\mathbf{e}}_c$, is perpendicular to the plane of $\mathbf{A}$ and $\mathbf{B}$, such that $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ form a right-handed system.[1] This causes $\mathbf{C}$ to be aligned with the rotational axis, with a sign that indicates the sense of the rotation.

From Fig. 3.2, we also see that $\mathbf{A} \times \mathbf{B}$ has a magnitude equal to the area of the parallelogram formed by $\mathbf{A}$ and $\mathbf{B}$, and with a direction **normal** to the parallelogram.

Other places the cross product is encountered include the formulas

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r} \quad \text{and} \quad \mathbf{F}_M = q\mathbf{v} \times \mathbf{B}.$$

The first of these equations is the relation between linear velocity $\mathbf{v}$ and and angular velocity $\boldsymbol{\omega}$, and the second equation gives the force $\mathbf{F}_M$ on a particle of charge q and velocity $\mathbf{v}$ in the magnetic induction field $\mathbf{B}$ (in SI units).

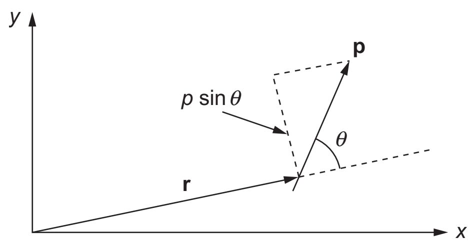

FIGURE 3.1 Angular momentum about the origin, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. $\mathbf{L}$ has magnitude $rp\sin\theta$ and is directed out of the plane of the paper.

[1] The inherent ambiguity in this statement can be resolved by the following anthropomorphic prescription: Point the right hand in the direction $\mathbf{A}$, and then bend the fingers through the **smaller** of the two angles that can cause the fingers to point in the direction $\mathbf{B}$; the thumb will then point in the direction of $\mathbf{C}$.

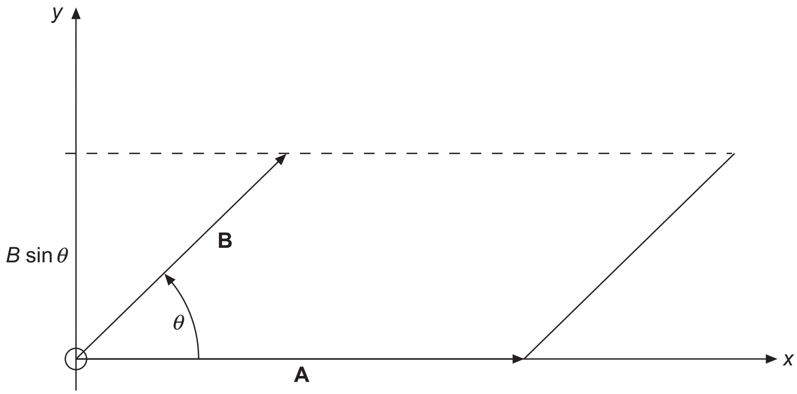

FIGURE 3.2 Parallelogram of $\mathbf{A} \times \mathbf{B}$.

We can get our right hands out of the analysis by compiling some algebraic properties of the cross product. If the roles of **A** and **B** are reversed, the cross product changes sign, so

$$\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B} \quad \text{(anticommutation).} \tag{3.3}$$

The cross product also obeys the distributive laws

$$\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}, \quad k(\mathbf{A} \times \mathbf{B}) = (k\mathbf{A}) \times \mathbf{B}, \tag{3.4}$$

and when applied to unit vectors in the coordinate directions, we get

$$\hat{\mathbf{e}}_i \times \hat{\mathbf{e}}_j = \sum_k \varepsilon_{ijk} \hat{\mathbf{e}}_k. \tag{3.5}$$

Here ε_{ijk} is the Levi-Civita symbol defined in Eq. (2.8); Eq. (3.5) therefore indicates, for example, that $\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_x = 0$, $\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_y = \hat{\mathbf{e}}_z$, but $\hat{\mathbf{e}}_y \times \hat{\mathbf{e}}_x = -\hat{\mathbf{e}}_z$.

Using Eq. (3.5) and writing **A** and **B** in component form, we can expand $\mathbf{A} \times \mathbf{B}$ to obtain

$$\begin{aligned}
\mathbf{C} = \mathbf{A} \times \mathbf{B} &= (A_x \hat{\mathbf{e}}_x + A_y \hat{\mathbf{e}}_y + A_z \hat{\mathbf{e}}_z) \times (B_x \hat{\mathbf{e}}_x + B_y \hat{\mathbf{e}}_y + B_z \hat{\mathbf{e}}_z) \\
&= (A_x B_y - A_y B_x)(\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_y) + (A_x B_z - A_z B_x)(\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_z) \\
&\quad + (A_y B_z - A_z B_y)(\hat{\mathbf{e}}_y \times \hat{\mathbf{e}}_z) \\
&= (A_x B_y - A_y B_x)\hat{\mathbf{e}}_z + (A_x B_z - A_z B_x)(-\hat{\mathbf{e}}_y) + (A_y B_z - A_z B_y)\hat{\mathbf{e}}_x.
\end{aligned} \tag{3.6}$$

The components of **C** are important enough to be displayed prominently:

$$C_x = A_y B_z - A_z B_y, \quad C_y = A_z B_x - A_x B_z, \quad C_z = A_x B_y - A_y B_x, \tag{3.7}$$

equivalent to

$$C_i = \sum_{jk} \varepsilon_{ijk} A_j B_k. \tag{3.8}$$

Yet another way of expressing the cross product is to write it as a determinant. It is straightforward to verify that Eqs. (3.7) are reproduced by the determinantal equation

$$\mathbf{C} = \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \tag{3.9}$$

when the determinant is expanded in minors of its top row. The anticommutation of the cross product now clearly follows if the rows for the components of **A** and **B** are interchanged.

We need to reconcile the geometric form of the cross product, Eq. (3.2), with the algebraic form in Eq. (3.6). We can confirm the magnitude of $\mathbf{A} \times \mathbf{B}$ by evaluating (from the component form of **C**)

$$\begin{aligned}(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{A} \times \mathbf{B}) &= A^2B^2 - (\mathbf{A} \cdot \mathbf{B})^2 = A^2B^2 - A^2B^2\cos^2\theta \\ &= A^2B^2\sin^2\theta. \end{aligned} \tag{3.10}$$

The first step in Eq. (3.10) can be verified by expanding its left-hand side in component form, then collecting the result into the terms constituting the central member of the first line of the equation.

To confirm the direction of $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, we can check that $\mathbf{A} \cdot \mathbf{C} = \mathbf{B} \cdot \mathbf{C} = 0$, showing that **C** (in component form) is perpendicular to both **A** and **B**. We illustrate for $\mathbf{A} \cdot \mathbf{C}$:

$$\mathbf{A} \cdot \mathbf{C} = A_x(A_yB_z - A_zB_y) + A_y(A_zB_x - A_xB_z) + A_z(A_xB_y - A_yB_x) = 0. \tag{3.11}$$

To verify the sign of **C**, it suffices to check special cases (e.g., $\mathbf{A} = \hat{\mathbf{e}}_x$, $\mathbf{B} = \hat{\mathbf{e}}_y$, or $A_x = B_y = 1$, all other components zero).

Next, we observe that it is obvious from Eq. (3.2) that if $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ in a given coordinate system, then that equation will also be satisfied if we rotate the coordinates, even though the individual components of all three vectors will thereby be changed. In other words, the cross product, like the dot product, is a rotationally invariant relationship.

Finally, note that the cross product is a quantity specifically defined for 3-D space. It is possible to make analogous definitions for spaces of other dimensionality, but they do not share the interpretation or utility of the cross product in $\mathbb{R}^3$.

Scalar Triple Product

While the various vector operations can be combined in many ways, there are two combinations involving three operands that are of particular importance. We call attention first to the **scalar triple product**, of the form $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$. Taking $(\mathbf{B} \times \mathbf{C})$ in the determinantal form, Eq. (3.9), one can see that taking the dot product with **A** will cause the unit vector $\hat{\mathbf{e}}_x$ to be replaced by A_x, with corresponding replacements to $\hat{\mathbf{e}}_y$ and $\hat{\mathbf{e}}_z$. The overall result is

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}. \tag{3.12}$$

We can draw a number of conclusions from this highly symmetric determinantal form. To start, we see that the determinant contains no vector quantities, so it must evaluate

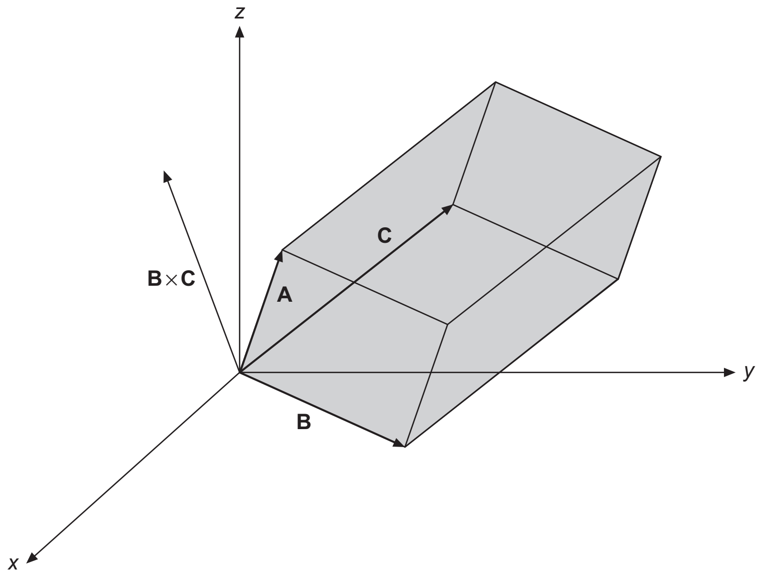

FIGURE 3.3 $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ parallelepiped.

to an ordinary number. Because the left-hand side of Eq. (3.12) is a rotational invariant, the number represented by the determinant must also be rotationally invariant, and can therefore be identified as a scalar. Since we can permute the rows of the determinant (with a sign change for an odd permutation, and with no sign change for an even permutation), we can permute the vectors **A**, **B**, and **C** to obtain

$$\mathbf{A} \cdot \mathbf{B} \times \mathbf{C} = \mathbf{B} \cdot \mathbf{C} \times \mathbf{A} = \mathbf{C} \cdot \mathbf{A} \times \mathbf{B} = -\mathbf{A} \cdot \mathbf{C} \times \mathbf{B}, \quad \text{etc.} \tag{3.13}$$

Here we have followed common practice and dropped the parentheses surrounding the cross product, on the basis that they must be understood to be present in order for the expressions to have meaning. Finally, noting that $\mathbf{B} \times \mathbf{C}$ has a magnitude equal to the area of the **BC** parallelog ram and a direction perpendicular to it, and that the dot product with **A** will multiply that area by the projection of **A** on $\mathbf{B} \times \mathbf{C}$, we see that the scalar triple product gives us (±) the volume of the parallelepiped defined by **A**, **B**, and **C**; see Fig. 3.3.

Example 3.2.1 Reciprocal Lattice

Let **a**, **b**, and **c** (not necessarily mutually perpendicular) represent the vectors that define a crystal lattice. The displacements from one lattice point to another may then be written

$$\mathbf{R} = n_a \mathbf{a} + n_b \mathbf{b} + n_c \mathbf{c}, \tag{3.14}$$

with n_a, n_b, and n_c taking integral values. In the band theory of solids,[2] it is useful to introduce what is called a **reciprocal lattice** $\mathbf{a}'$, $\mathbf{b}'$, $\mathbf{c}'$ such that

$$\mathbf{a} \cdot \mathbf{a}' = \mathbf{b} \cdot \mathbf{b}' = \mathbf{c} \cdot \mathbf{c}' = 1, \tag{3.15}$$

and with

$$\mathbf{a} \cdot \mathbf{b}' = \mathbf{a} \cdot \mathbf{c}' = \mathbf{b} \cdot \mathbf{a}' = \mathbf{b} \cdot \mathbf{c}' = \mathbf{c} \cdot \mathbf{a}' = \mathbf{c} \cdot \mathbf{b}' = 0. \tag{3.16}$$

The reciprocal-lattice vectors are easily constructed by calling on the fact that for any $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{u} \times \mathbf{v}$ is perpendicular to both $\mathbf{u}$ and $\mathbf{v}$; we have

$$\mathbf{a}' = \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \quad \mathbf{b}' = \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}, \quad \mathbf{c}' = \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}}. \tag{3.17}$$

The scalar triple product causes these expressions to satisfy the scale condition of Eq. (3.15). ■

Vector Triple Product

The other triple product of importance is the **vector triple product**, of the form $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$. Here the parentheses are essential since, for example, $(\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_x) \times \hat{\mathbf{e}}_y = 0$, while $\hat{\mathbf{e}}_x \times (\hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_y) = \hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_z = -\hat{\mathbf{e}}_y$. Our interest is in reducing this triple product to a simpler form; the result we seek is

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \tag{3.18}$$

Equation (3.18), which for convenience we will sometimes refer to as the BAC–CAB rule, can be proved by inserting components for all vectors and evaluating all the products, but it is instructive to proceed in a more elegant fashion. Using the formula for the cross product in terms of the Levi-Civita symbol, Eq. (3.8), we write

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \sum_i \hat{\mathbf{e}}_i \sum_{jk} \varepsilon_{ijk} A_j \left(\sum_{pq} \varepsilon_{kpq} B_p C_q \right) \\ &= \sum_{ij} \sum_{pq} \hat{\mathbf{e}}_i A_j B_p C_q \sum_k \varepsilon_{ijk} \varepsilon_{kpq}. \end{aligned} \tag{3.19}$$

The summation over k of the product of Levi-Civita symbols reduces, as shown in Exercise 2.1.9, to $\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}$; we are left with

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \sum_{ij} \hat{\mathbf{e}}_i A_j (B_i C_j - B_j C_i) = \sum_i \hat{\mathbf{e}}_i \left(B_i \sum_j A_j C_j - C_i \sum_j A_j B_j \right),$$

which is equivalent to Eq. (3.18).

[2] It is often chosen to require $\mathbf{a} \cdot \mathbf{a}'$, etc., to be 2π rather than unity, because when Bloch states for a crystal (labeled by $\mathbf{k}$) are set up, a constituent atomic function in cell $\mathbf{R}$ enters with coefficient $\exp(i\mathbf{k} \cdot \mathbf{R})$, and if $\mathbf{k}$ is changed by a reciprocal lattice step (in, say, the $\mathbf{a}'$ direction), the coefficient becomes $\exp(i[\mathbf{k} + \mathbf{a}'] \cdot \mathbf{R})$, which reduces to $\exp(2\pi i n_a)\exp(i\mathbf{k} \cdot \mathbf{R})$ and therefore, because $\exp(2\pi i n_a) = 1$, to its original value. Thus, the reciprocal lattice identifies the periodicity in $\mathbf{k}$. The unit cell of the $\mathbf{k}$ vectors is called the **Brillouin zone**

Exercises

3.2.1 If $\mathbf{P} = \hat{\mathbf{e}}_x P_x + \hat{\mathbf{e}}_y P_y$ and $\mathbf{Q} = \hat{\mathbf{e}}_x Q_x + \hat{\mathbf{e}}_y Q_y$ are any two nonparallel (also nonantiparallel) vectors in the xy-plane, show that $\mathbf{P} \times \mathbf{Q}$ is in the z-direction.

3.2.2 Prove that $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{A} \times \mathbf{B}) = (AB)^2 - (\mathbf{A} \cdot \mathbf{B})^2$.

3.2.3 Using the vectors

$$\mathbf{P} = \hat{\mathbf{e}}_x \cos\theta + \hat{\mathbf{e}}_y \sin\theta,$$

$$\mathbf{Q} = \hat{\mathbf{e}}_x \cos\varphi - \hat{\mathbf{e}}_y \sin\varphi,$$

$$\mathbf{R} = \hat{\mathbf{e}}_x \cos\varphi + \hat{\mathbf{e}}_y \sin\varphi,$$

prove the familiar trigonometric identities

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

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

3.2.4 (a) Find a vector $\mathbf{A}$ that is perpendicular to

$$\mathbf{U} = 2\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y - \hat{\mathbf{e}}_z,$$

$$\mathbf{V} = \hat{\mathbf{e}}_x - \hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z.$$

(b) What is $\mathbf{A}$ if, in addition to this requirement, we demand that it have unit magnitude?

3.2.5 If four vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$ all lie in the same plane, show that

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

Hint. Consider the directions of the cross-product vectors.

3.2.6 Derive the law of sines (see Fig. 3.4):

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

3.2.7 The magnetic induction $\mathbf{B}$ is **defined** by the Lorentz force equation,

$$\mathbf{F} = q(\mathbf{v} \times \mathbf{B}).$$

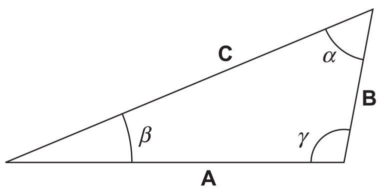

FIGURE 3.4 Plane triangle.

Carrying out three experiments, we find that if

$$\mathbf{v} = \hat{\mathbf{e}}_x, \quad \frac{\mathbf{F}}{q} = 2\hat{\mathbf{e}}_z - 4\hat{\mathbf{e}}_y,$$

$$\mathbf{v} = \hat{\mathbf{e}}_y, \quad \frac{\mathbf{F}}{q} = 4\hat{\mathbf{e}}_x - \hat{\mathbf{e}}_z,$$

$$\mathbf{v} = \hat{\mathbf{e}}_z, \quad \frac{\mathbf{F}}{q} = \hat{\mathbf{e}}_y - 2\hat{\mathbf{e}}_x.$$

From the results of these three separate experiments calculate the magnetic induction **B**.

3.2.8 You are given the three vectors **A**, **B**, and **C**,

$$\mathbf{A} = \hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y,$$

$$\mathbf{B} = \hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z,$$

$$\mathbf{C} = \hat{\mathbf{e}}_x - \hat{\mathbf{e}}_z.$$

(a) Compute the scalar triple product, $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$. Noting that $\mathbf{A} = \mathbf{B} + \mathbf{C}$, give a geometric interpretation of your result for the scalar triple product.

(b) Compute $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$.

3.2.9 Prove Jacobi's identity for vector products:

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

3.2.10 A vector **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) $\mathbf{A}_t = -\hat{\mathbf{r}} \times (\hat{\mathbf{r}} \times \mathbf{A})$.

3.2.11 Prove that a necessary and sufficient condition for the three (nonvanishing) vectors **A**, **B**, and **C** to be coplanar is the vanishing of the scalar triple product

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

3.2.12 Three vectors **A**, **B**, and **C** are given by

$$\mathbf{A} = 3\hat{\mathbf{e}}_x - 2\hat{\mathbf{e}}_y + 2\hat{\mathbf{z}},$$

$$\mathbf{B} = 6\hat{\mathbf{e}}_x + 4\hat{\mathbf{e}}_y - 2\hat{\mathbf{z}},$$

$$\mathbf{C} = -3\hat{\mathbf{e}}_x - 2\hat{\mathbf{e}}_y - 4\hat{\mathbf{z}}.$$

Compute the values of $\mathbf{A} \cdot \mathbf{B} \times \mathbf{C}$ and $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$, $\mathbf{C} \times (\mathbf{A} \times \mathbf{B})$ and $\mathbf{B} \times (\mathbf{C} \times \mathbf{A})$.

3.2.13 Show that

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

3.2.14 Show that

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

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

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

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

(a) Show that the magnetic force exerted by q_1 on a second charge q_2, velocity $\mathbf{v}_2$, is given by the vector triple product

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

(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) $\mathbf{F}_1 = -\frac{\mu_0}{4\pi} \frac{q_1 q_2}{r^2} \mathbf{v}_1 \times \big(\mathbf{v}_2 \times \hat{\mathbf{r}}\big).$
In general, there is no simple relation between $\mathbf{F}_1$ and $\mathbf{F}_2$. Specifically, Newton's third law, $\mathbf{F}_1 = -\mathbf{F}_2$, does not hold.

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

3.3 Coordinate Transformations

As indicated in the chapter introduction, an object classified as a vector must have specific transformation properties under rotation of the coordinate system; in particular, the components of a vector must transform in a way that describes the same object in the rotated system.

Rotations

Considering initially $\mathbb{R}^2$, and a rotation of the coordinate axes as shown in Fig. 3.5, we wish to find how the components A_x and A_y of a vector $\mathbf{A}$ in the unrotated system are related to A'_x and A'_y, its components in the rotated coordinate system. Perhaps the easiest way to answer this question is by first asking how the unit vectors $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$ are represented in the new coordinates, after which we can perform vector addition on the new incarnations of $A_x\hat{\mathbf{e}}_x$ and $A_y\hat{\mathbf{e}}_y$.

From the right-hand part of Fig. 3.5, we see that

$$\hat{\mathbf{e}}_x = \cos\varphi\,\hat{\mathbf{e}}'_x - \sin\varphi\,\hat{\mathbf{e}}'_y, \quad \text{and} \quad \hat{\mathbf{e}}_y = \sin\varphi\,\hat{\mathbf{e}}'_x + \cos\varphi\,\hat{\mathbf{e}}'_y, \tag{3.20}$$

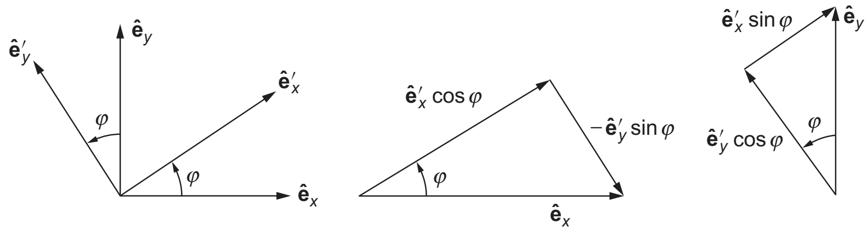

FIGURE 3.5 Left: Rotation of two-dimensional (2-D) coordinate axes through angle φ. Center and right: Decomposition of $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$ into their components in the rotated system.

so the **unchanged** vector **A** now takes the **changed** form

$$\begin{aligned}\mathbf{A} = A_x\hat{\mathbf{e}}_x + A_y\hat{\mathbf{e}}_y &= A_x(\cos\varphi\,\hat{\mathbf{e}}'_x - \sin\varphi\,\hat{\mathbf{e}}'_y) + A_y(\sin\varphi\,\hat{\mathbf{e}}'_x + \cos\varphi\,\hat{\mathbf{e}}'_y)\\ &= (A_x\cos\varphi + A_y\sin\varphi)\hat{\mathbf{e}}'_x + (-A_x\sin\varphi + A_y\cos\varphi)\hat{\mathbf{e}}'_y.\end{aligned}\tag{3.21}$$

If we write the vector **A** in the rotated (primed) coordinate system as

$$\mathbf{A} = A'_x\hat{\mathbf{e}}'_x + A'_y\hat{\mathbf{e}}'_y,$$

we then have

$$A'_x = A_x\cos\varphi + A_y\sin\varphi, \quad A'_y = -A_x\sin\varphi + A_y\cos\varphi, \tag{3.22}$$

which is equivalent to the matrix equation

$$\mathbf{A}' = \begin{pmatrix} A'_x \\ A'_y \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} A_x \\ A_y \end{pmatrix}. \tag{3.23}$$

Suppose now that we start from **A** as given by its components in the rotated system, (A'_x, A'_y), and rotate the coordinate system back to its original orientation. This will entail a rotaton in the amount $-\varphi$, and corresponds to the matrix equation

$$\begin{pmatrix} A_x \\ A_y \end{pmatrix} = \begin{pmatrix} \cos(-\varphi) & \sin(-\varphi) \\ -\sin(-\varphi) & \cos(-\varphi) \end{pmatrix}\begin{pmatrix} A'_x \\ A'_y \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} A'_x \\ A'_y \end{pmatrix}. \tag{3.24}$$

Assigning the 2×2 matrices in Eqs. (3.23) and (3.24) the respective names S and S', we see that these two equations are equivalent to $\mathbf{A}' = \mathsf{S}\mathbf{A}$ and $\mathbf{A} = \mathsf{S}'\mathbf{A}'$, with

$$\mathsf{S} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \quad \text{and} \quad \mathsf{S}' = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{3.25}$$

Now, applying S to **A** and then S' to $\mathsf{S}\mathbf{A}$ (corresponding to first rotating the coordinate system an amount $+\varphi$ and then an amount $-\varphi$), we recover **A**, or

$$\mathbf{A} = \mathsf{S}'\mathsf{S}\mathbf{A}.$$

Since this result must be valid for any **A**, we conclude that $\mathsf{S}' = \mathsf{S}^{-1}$. We also see that $\mathsf{S}' = \mathsf{S}^T$. We can check that $\mathsf{S}S' = \mathbf{1}$ by matrix multiplication:

$$\mathsf{S}\mathsf{S}' = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since S is real, the fact that $\mathsf{S}^{-1} = \mathsf{S}^T$ means that it is **orthogonal**. In summary, we have found that the transformation connecting $\mathbf{A}$ and $\mathbf{A}'$ (the same vector, but represented in the rotated coordinate system) is

$$\mathbf{A}' = \mathsf{S}\mathbf{A}, \tag{3.26}$$

with S an orthogonal matrix.

Orthogonal Transformations

It was no accident that the transformation describing a rotation in $\mathbb{R}^2$ was **orthogonal**, by which we mean that the matrix effecting the transformation was an orthogonal matrix.

An instructive way of writing the transformation S is, returning to Eq. (3.20), to rewrite those equations as

$$\hat{\mathbf{e}}_x = (\hat{\mathbf{e}}'_x \cdot \hat{\mathbf{e}}_x)\hat{\mathbf{e}}'_x + (\hat{\mathbf{e}}'_y \cdot \hat{\mathbf{e}}_x)\hat{\mathbf{e}}'_y, \quad \hat{\mathbf{e}}_y = (\hat{\mathbf{e}}'_x \cdot \hat{\mathbf{e}}_y)\hat{\mathbf{e}}'_x + (\hat{\mathbf{e}}'_y \cdot \hat{\mathbf{e}}_y)\hat{\mathbf{e}}'_y. \tag{3.27}$$

This corresponds to writing $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$ as the sum of their projections on the orthogonal vectors $\hat{\mathbf{e}}'_x$ and $\hat{\mathbf{e}}'_y$. Now we can rewrite S as

$$\mathsf{S} = \begin{pmatrix} \hat{\mathbf{e}}'_x \cdot \hat{\mathbf{e}}_x & \hat{\mathbf{e}}'_x \cdot \hat{\mathbf{e}}_y \\ \hat{\mathbf{e}}'_y \cdot \hat{\mathbf{e}}_x & \hat{\mathbf{e}}'_y \cdot \hat{\mathbf{e}}_y \end{pmatrix}. \tag{3.28}$$

This means that each row of S contains the components (in the unprimed coordinates) of a unit vector (either $\hat{\mathbf{e}}'_x$ or $\hat{\mathbf{e}}'_y$) that is orthogonal to the vector whose components are in the other row. In turn, this means that the dot products of different row vectors will be zero, while the dot product of any row vector with itself (because it is a unit vector) will be unity. That is the deeper significance of an **orthogonal** matrix S; the $\mu\nu$ element of $\mathsf{S}\mathsf{S}^T$ is the dot product formed from the μth row of S and the νth column of S^T (which is the same as the νth row of S). Since these row vectors are orthogonal, we will get zero if $\mu \neq \nu$, and because they are unit vectors, we will get unity if $\mu = \nu$. In other words, $\mathsf{S}\mathsf{S}^T$ will be a unit matrix.

Before leaving Eq. (3.28), note that its columns also have a simple interpretation: Each contains the components (in the primed coordinates) of one of the unit vectors of the unprimed set. Thus the dot product formed from two different **columns** of S will vanish, while the dot product of any column with itself will be unity. This corresponds to the fact that, for an orthogonal matrix, we also have $\mathsf{S}^T\mathsf{S} = \mathbf{1}$.

Summarizing part of the above,

> *The transformation from one orthogonal Cartesian coordinate system to another Cartesian system is described by an* **orthogonal** *matrix.*

In Chapter 2 we found that an orthogonal matrix must have a determinant that is real and of magnitude unity, i.e., ± 1. However, for rotations in ordinary space the value of the determinant will always be $+1$. One way to understand this is to consider the fact that any rotation can be built up from a large number of small rotations, and that the determinant must vary continuously as the amount of rotation is changed. The identity rotation (i.e., no rotation at all) has determinant $+1$. Since no value close to $+1$ except $+1$ itself is a permitted value for the determinant, rotations cannot change the value of the determinant.

Reflections

Another possibility for changing a coordinate system is to subject it to a reflection operation. For simplicity, consider first the **inversion** operation, in which the sign of each coordinate is reversed. In $\mathbb{R}^3$, the transformation matrix S will be the 3×3 analog of Eq. (3.28), and the transformation under discussion is to set $\hat{\mathbf{e}}'_\mu = -\hat{\mathbf{e}}_\mu$, with $\mu = x,\ y,$ and z. This will lead to

$$\mathsf{S} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

which clearly results in $\det \mathsf{S} = -1$. The change in sign of the determinant corresponds to the change from a right-handed to a left-handed coordinate system (which obviously cannot be accomplished by a rotation). Reflection about a plane (as in the image produced by a plane mirror) also changes the sign of the determinant and the handedness of the coordinate system; for example, reflection in the xy-plane changes the sign of $\hat{\mathbf{e}}_z$, leaving the other two unit vectors unchanged; the transformation matrix S for this transformation is

$$\mathsf{S} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Its determinant is also -1.

The formulas for vector addition, multiplication by a scalar, and the dot product are unaffected by a reflection transformation of the coordinates, but this is not true of the cross product. To see this, look at the formula for any one of the components of $\mathbf{A} \times \mathbf{B}$, and how it would change under inversion (where the same, unchanged vectors in physical space now have sign changes to all their components):

$$C_x: \quad A_y B_z - A_z B_y \quad \longrightarrow \quad (-A_y)(-B_z) - (-A_z)(-B_y) = A_y B_z - A_z B_y.$$

Note that this formula says that the sign of C_x should not change, even though it must in order to describe the unchanged physical situation. The conclusion is that our transformation law fails for the result of a cross-product operation. However, the mathematics can be salvaged if we classify $\mathbf{B} \times \mathbf{C}$ as a different type of quantity than $\mathbf{B}$ and $\mathbf{C}$. Many texts on vector analysis call vectors whose components change sign under coordinate reflection **polar vectors**, and those whose components do not then change sign **axial vectors**. The term **axial** doubtless arises from the fact that cross products frequently describe phenomena associated with rotation about the axis defined by the axial vector. Nowadays, it is becoming more usual to call **polar vectors** just **vectors**, because we want that term to describe objects that obey for all S the transformation law

$$\mathbf{A}' = \mathsf{S}\mathbf{A} \quad \text{(vectors)}, \tag{3.29}$$

(and specifically without a restriction to S whose determinants are $+1$). Axial vectors, for which the vector transformation law fails for coordinate reflections, are then referred to as **pseudovectors**, and their transformation law can be expressed in the somewhat more complicated form

$$\mathbf{C}' = \det(\mathsf{S})\, \mathsf{S}\mathbf{C} \quad \text{(pseudovectors)}. \tag{3.30}$$

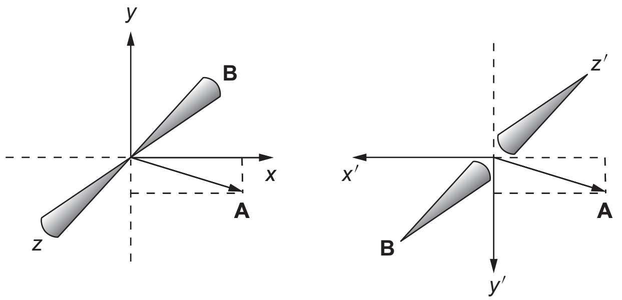

FIGURE 3.6 Inversion (right) of original coordinates (left) and the effect on a vector **A** and a pseudovector **B**.

The effect of an inversion operation on a coordinate system and on a vector and a pseudovector are shown in Fig. 3.6.

Since vectors and pseudovectors have different transformation laws, it is in general without physical meaning to add them together.[3] It is also usually meaningless to equate quantities of different transformational properties: in $\mathbf{A} = \mathbf{B}$, both quantities must be either vectors or pseudovectors.

Pseudovectors, of course, enter into more complicated expressions, of which an example is the scalar triple product $\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}$. Under coordinate reflection, the components of $\mathbf{B}\times\mathbf{C}$ do not change (as observed earlier), but those of $\mathbf{A}$ are reversed, with the result that $\mathbf{A}\cdot\mathbf{B}\times\mathbf{C}$ changes sign. We therefore need to reclassify it as a **pseudoscalar**. On the other hand, the vector triple product, $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$, which contains two cross products, evaluates, as shown in Eq. (3.18), to an expression containing only legitimate scalars and (polar) vectors. It is therefore proper to identify $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$ as a vector. These cases illustrate the general principle that a product with an odd number of pseudo quantities is "pseudo," while those with even numbers of pseudo quantities are not.

Successive Operations

One can carry out a succession of coordinate rotations and/or reflections by applying the relevant orthogonal transformations. In fact, we already did this in our introductory discussion for $\mathbb{R}^2$ where we applied a rotation and then its inverse. In general, if R and R' refer to such operations, the application to $\mathbf{A}$ of R followed by the application of R' corresponds to

$$\mathbf{A}' = \mathsf{S}(R')\mathsf{S}(R)\mathbf{A}, \tag{3.31}$$

and the overall result of the two transformations can be identified as a single transformation whose matrix $\mathsf{S}(R'R)$ is the matrix product $\mathsf{S}(R')\mathsf{S}(R)$.

[3]The big exception to this is in beta-decay weak interactions. Here the universe distinguishes between right- and left-handed systems, and we add polar and axial vector interactions.

Two points should be noted:

1. The operations take place in right-to-left order: The rightmost operator is the one applied to the original **A**; that to its left then applies to the result of the first operation, etc.
2. The combined operation $R'R$ is a transformation between two orthogonal coordinate systems and therefore can be described by an orthogonal matrix: The product of two orthogonal matrices is orthogonal.

Exercises

3.3.1 A rotation $\varphi_1 + \varphi_2$ about the z-axis is carried out as two successive rotations φ_1 and φ_2, each about the z-axis. Use the matrix representation of the rotations to derive the trigonometric identities

$$\cos(\varphi_1 + \varphi_2) = \cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2,$$

$$\sin(\varphi_1 + \varphi_2) = \sin\varphi_1 \cos\varphi_2 + \cos\varphi_1 \sin\varphi_2.$$

3.3.2 A corner reflector is formed by three mutually perpendicular reflecting surfaces. Show that a ray of light incident upon the corner reflector (striking all three surfaces) is reflected back along a line parallel to the line of incidence.
Hint. Consider the effect of a reflection on the components of a vector describing the direction of the light ray.

3.3.3 Let **x** and **y** be column vectors. Under an orthogonal transformation S, they become $\mathbf{x}' = \mathsf{S}\mathbf{x}$ and $\mathbf{y}' = \mathsf{S}\mathbf{y}$. Show that $(\mathbf{x}')^T\mathbf{y}' = \mathbf{x}^T\mathbf{y}$, a result equivalent to the invariance of the dot product under a rotational transformation.

3.3.4 Given the orthogonal transformation matrix S and vectors **a** and **b**,

$$\mathsf{S} = \begin{pmatrix} 0.80 & 0.60 & 0.00 \\ -0.48 & 0.64 & 0.60 \\ 0.36 & -0.48 & 0.80 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix},$$

(a) Calculate $\det(\mathsf{S})$.
(b) Verify that $\mathbf{a} \cdot \mathbf{b}$ is invariant under application of S to **a** and **b**.
(c) Determine what happens to $\mathbf{a} \times \mathbf{b}$ under application of S to **a** and **b**. Is this what is expected?

3.3.5 Using **a** and **b** as defined in Exercise 3.3.4, but with

$$\mathsf{S} = \begin{pmatrix} 0.60 & 0.00 & 0.80 \\ -0.64 & -0.60 & 0.48 \\ -0.48 & 0.80 & 0.36 \end{pmatrix} \quad \text{and} \quad \mathbf{c} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix},$$

(a) Calculate $\det(\mathsf{S})$.

Apply S to **a**, **b**, and **c**, and determine what happens to

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

(c) $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$,
(d) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$.
(e) Classify the expressions in (b) through (d) as scalar, vector, pseudovector, or pseudoscalar.

3.4 Rotations in $\mathbb{R}^3$

Because of its practical importance, we discuss now in some detail the treatment of rotations in $\mathbb{R}^3$. An obvious starting point, based on our experience in $\mathbb{R}^2$, would be to write the 3×3 matrix S of Eq. (3.28), with rows that describe the orientations of a rotated (primed) set of unit vectors in terms of the original (unprimed) unit vectors:

$$\mathsf{S} = \begin{pmatrix} \hat{\mathbf{e}}'_1 \cdot \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}'_1 \cdot \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}'_1 \cdot \hat{\mathbf{e}}_3 \\ \hat{\mathbf{e}}'_2 \cdot \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}'_2 \cdot \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}'_2 \cdot \hat{\mathbf{e}}_3 \\ \hat{\mathbf{e}}'_3 \cdot \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}'_3 \cdot \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}'_3 \cdot \hat{\mathbf{e}}_3 \end{pmatrix} \tag{3.32}$$

We have switched the coordinate labels from x, y, z to 1, 2, 3 for convenience in some of the formulas that use Eq. (3.32). It is useful to make one observation about the elements of S, namely $s_{\mu\nu} = \hat{\mathbf{e}}'_\mu \cdot \hat{\mathbf{e}}_\nu$. This dot product is the projection of $\hat{\mathbf{e}}'_\mu$ onto the $\hat{\mathbf{e}}_\nu$ direction, and is therefore the change in x_ν that is produced by a unit change in x'_μ. Since the relation between the coordinates is linear, we can identify $\hat{\mathbf{e}}'_\mu \cdot \hat{\mathbf{e}}_\nu$ as $\partial x_\nu / \partial x'_\mu$, so our transformation matrix S can be written in the alternate form

$$\mathsf{S} = \begin{pmatrix} \partial x_1/\partial x'_1 & \partial x_2/\partial x'_1 & \partial x_3/\partial x'_1 \\ \partial x_1/\partial x'_2 & \partial x_2/\partial x'_2 & \partial x_3/\partial x'_2 \\ \partial x_1/\partial x'_3 & \partial x_2/\partial x'_3 & \partial x_3/\partial x'_3 \end{pmatrix}. \tag{3.33}$$

The argument we made to evaluate $\hat{\mathbf{e}}'_\mu \cdot \hat{\mathbf{e}}_\nu$ could as easily have been made with the roles of the two unit vectors reversed, yielding instead of $\partial x_\nu / \partial x'_\mu$ the derivative $\partial x'_\mu / \partial x_\nu$. We then have what at first may seem to be a surprising result:

$$\frac{\partial x_\nu}{\partial x'_\mu} = \frac{\partial x'_\mu}{\partial x_\nu}. \tag{3.34}$$

A superficial look at this equation suggests that its two sides would be reciprocals. The problem is that we have not been notationally careful enough to avoid ambiguity: the derivative on the left-hand side is to be taken with the other x' coordinates fixed, while that on the right-hand side is with the other unprimed coordinates fixed. In fact, the equality in Eq. (3.34) is needed to make S an orthogonal matrix.

We note in passing that the observation that the coordinates are related linearly restricts the current discussion to Cartesian coordinate systems. Curvilinear coordinates are treated later.

Neither Eq. (3.32) nor Eq. (3.33) makes obvious the possibility of relations among the elements of S. In $\mathbb{R}^2$, we found that all the elements of S depended on a single variable, the rotation angle. In $\mathbb{R}^3$, the number of independent variables needed to specify a general rotation is three: Two parameters (usually angles) are needed to specify the direction of $\hat{\mathbf{e}}'_3$; then one angle is needed to specify the direction of $\hat{\mathbf{e}}'_1$ in the plane perpendicular to $\hat{\mathbf{e}}'_3$;

at this point the orientation of $\hat{\mathbf{e}}'_2$ is completely determined. Therefore, of the nine elements of S, only three are in fact independent. The usual parameters used to specify $\mathbb{R}^3$ rotations are the **Euler angles**.[4] It is useful to have S given explicitly in terms of them, as the Lagrangian formulation of mechanics requires the use of a set of *independent* variables.

The Euler angles describe an $\mathbb{R}^3$ rotation in three steps, the first two of which have the effect of fixing the orientation of the new $\hat{\mathbf{e}}_3$ axis (the polar direction in spherical coordinates), while the third Euler angle indicates the amount of subsequent rotation about that axis. The first two steps do more than identify a new polar direction; they describe rotations that cause the realignment. As a result, we can obtain the matrix representations of these (and the third rotation), and apply them sequentially (i.e., as a matrix product) to obtain the overall effect of the rotation.

The three steps describing rotation of the coordinate axes are the following (also illustrated in Fig. 3.7):

1. The coordinates are rotated about the $\hat{\mathbf{e}}_3$ axis counterclockwise (as viewed from positive $\hat{\mathbf{e}}_3$) through an angle α in the range $0 \le \alpha < 2\pi$, into new axes denoted $\hat{\mathbf{e}}'_1$, $\hat{\mathbf{e}}'_2$, $\hat{\mathbf{e}}'_3$. (The polar direction is not changed; the $\hat{\mathbf{e}}_3$ and $\hat{\mathbf{e}}'_3$ axes coincide.)
2. The coordinates are rotated about the $\hat{\mathbf{e}}'_2$ axis counterclockwise (as viewed from positive $\hat{\mathbf{e}}'_2$) through an angle β in the range $0 \le \beta \le \pi$, into new axes denoted $\hat{\mathbf{e}}''_1$, $\hat{\mathbf{e}}''_2$, $\hat{\mathbf{e}}''_3$. (This tilts the polar direction toward the $\hat{\mathbf{e}}'_1$ direction, but leaves $\hat{\mathbf{e}}'_2$ unchanged.)
3. The coordinates are now rotated about the $\hat{\mathbf{e}}''_3$ axis counterclockwise (as viewed from positive $\hat{\mathbf{e}}''_3$) through an angle γ in the range $0 \le \gamma < 2\pi$, into the final axes, denoted $\hat{\mathbf{e}}'''_1$, $\hat{\mathbf{e}}'''_2$, $\hat{\mathbf{e}}'''_3$. (This rotation leaves the polar direction, $\hat{\mathbf{e}}''_3$, unchanged.)

In terms of the usual spherical polar coordinates (r, θ, φ), the final polar axis is at the orientation $\theta = \beta$, $\varphi = \alpha$. The final orientations of the other axes depend on all three Euler angles.

We now need the transformation matrices. The first rotation causes $\hat{\mathbf{e}}'_1$ and $\hat{\mathbf{e}}'_2$ to remain in the xy-plane, and has in its first two rows and columns exactly the same form

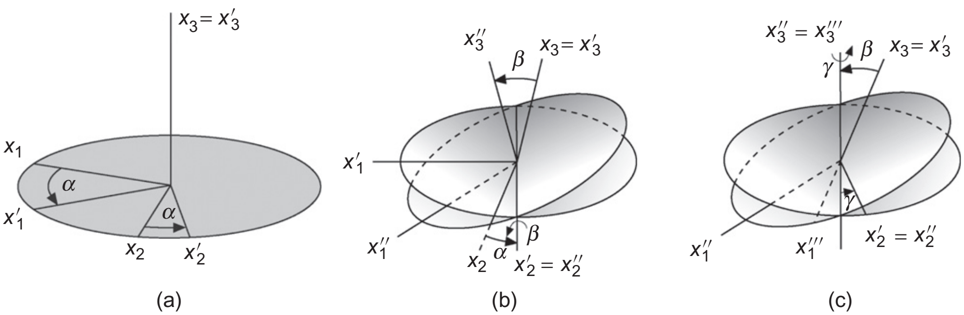

FIGURE 3.7 Euler angle rotations: (a) about $\hat{\mathbf{e}}_3$ through angle α; (b) about $\hat{\mathbf{e}}'_2$ through angle β; (c) about $\hat{\mathbf{e}}''_3$ through angle γ.

[4] There are almost as many definitions of the Euler angles as there are authors. Here we follow the choice generally made by workers in the area of group theory and the quantum theory of angular momentum.

as S in Eq. (3.25):

$$\mathsf{S}_1(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.35}$$

The third row and column of S_1 indicate that this rotation leaves unchanged the $\hat{\mathbf{e}}_3$ component of any vector on which it operates. The second rotation (applied to the coordinate system as it exists **after** the first rotation) is in the $\hat{\mathbf{e}}_3'\hat{\mathbf{e}}_1'$ plane; note that the signs of $\sin\beta$ have to be consistent with a cyclic permutation of the axis numbering:

$$\mathsf{S}_2(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}.$$

The third rotation is like the first, but with rotation amount γ:

$$\mathsf{S}_3(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The total rotation is described by the triple matrix product

$$\mathsf{S}(\alpha,\beta,\gamma) = \mathsf{S}_3(\gamma)\mathsf{S}_2(\beta)\mathsf{S}_1(\alpha). \tag{3.36}$$

Note the order: $\mathsf{S}_1(\alpha)$ operates first, then $\mathsf{S}_2(\beta)$, and finally $\mathsf{S}_3(\gamma)$. Direct multiplication gives

$$\mathsf{S}(\alpha,\beta,\gamma) = \begin{pmatrix} \cos\gamma\cos\beta\cos\alpha - \sin\gamma\sin\alpha & \cos\gamma\cos\beta\sin\alpha + \sin\gamma\cos\alpha & -\cos\gamma\sin\beta \\ -\sin\gamma\cos\beta\cos\alpha - \cos\gamma\sin\alpha & -\sin\gamma\cos\beta\sin\alpha + \cos\gamma\cos\alpha & \sin\gamma\sin\beta \\ \sin\beta\cos\alpha & \sin\beta\sin\alpha & \cos\beta \end{pmatrix}. \tag{3.37}$$

In case they are wanted, note that the elements s_{ij} in Eq. (3.37) give the explicit forms of the dot products $\hat{\mathbf{e}}_i''' \cdot \hat{\mathbf{e}}_j$ (and therefore also the partial derivatives $\partial x_i/\partial x_j'''$).

Note that each of S_1, S_2, and S_3 are orthogonal, with determinant $+1$, so that the overall S will also be orthogonal with determinant $+1$.

Example 3.4.1 An $\mathbb{R}^3$ Rotation

Consider a vector originally with components $(2, -1, 3)$. We want its components in a coordinate system reached by Euler angle rotations $\alpha = \beta = \gamma = \pi/2$. Evaluating $\mathsf{S}(\alpha,\beta,\gamma)$:

$$\mathsf{S}(\alpha,\beta,\gamma) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

A partial check on this value of S is obtained by verifying that $\det(\mathsf{S}) = +1$.

Then, in the new coordinates, our vector has components

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}.$$

The reader should check this result by visualizing the rotations involved. ■

Exercises

3.4.1 Another set of Euler rotations in common use is

(1) a rotation about the x_3-axis through an angle φ, counterclockwise,
(2) a rotation about the x_1'-axis through an angle θ, counterclockwise,
(3) a rotation about the x_3''-axis through an angle ψ, counterclockwise.

If

$$\begin{aligned} \alpha &= \varphi - \pi/2 \\ \beta &= \theta \\ \gamma &= \psi + \pi/2 \end{aligned} \qquad \text{or} \qquad \begin{aligned} \varphi &= \alpha + \pi/2 \\ \theta &= \beta \\ \psi &= \gamma - \pi/2, \end{aligned}$$

show that the final systems are identical.

3.4.2 Suppose the Earth is moved (rotated) so that the north pole goes to 30° north, 20° west (original latitude and longitude system) and the 10° west meridian points due south (also in the original system).

(a) What are the Euler angles describing this rotation?
(b) Find the corresponding direction cosines.

$$\textit{ANS.} \quad (b) \quad \mathsf{S} = \begin{pmatrix} 0.9551 & -0.2552 & -0.1504 \\ 0.0052 & 0.5221 & -0.8529 \\ 0.2962 & 0.8138 & 0.5000 \end{pmatrix}.$$

3.4.3 Verify that the Euler angle rotation matrix, Eq. (3.37), is invariant under the transformation

$$\alpha \to \alpha + \pi, \quad \beta \to -\beta, \quad \gamma \to \gamma - \pi.$$

3.4.4 Show that the Euler angle rotation matrix $\mathsf{S}(\alpha, \beta, \gamma)$ satisfies the following relations:

(a) $\mathsf{S}^{-1}(\alpha, \beta, \gamma) = \tilde{\mathsf{S}}(\alpha, \beta, \gamma)$,
(b) $\mathsf{S}^{-1}(\alpha, \beta, \gamma) = \mathsf{S}(-\gamma, -\beta, -\alpha)$.

3.4.5 The coordinate system (x, y, z) is rotated through an angle $\mathbf{\Phi}$ counterclockwise about an axis defined by the unit vector $\hat{\mathbf{n}}$ into system (x', y', z'). In terms of the new coordinates the radius vector becomes

$$\mathbf{r}' = \mathbf{r}\cos\boldsymbol{\Phi} + \mathbf{r}\times\mathbf{n}\sin\boldsymbol{\Phi} + \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{r})(1-\cos\boldsymbol{\Phi}).$$

(a) Derive this expression from geometric considerations.

(b) Show that it reduces as expected for $\hat{\mathbf{n}} = \hat{\mathbf{e}}_z$. The answer, in matrix form, appears in Eq. (3.35).

(c) Verify that $r'^2 = r^2$.

3.5 Differential Vector Operators

We move now to the important situation in which a vector is associated with each point in space, and therefore has a value (its set of components) that depends on the coordinates specifying its position. A typical example in physics is the electric field $\mathbf{E}(x, y, z)$, which describes the direction and magnitude of the electric force if a unit "test charge" was placed at x, y, z. The term **field** refers to a quantity that has values at all points of a region; if the quantity is a vector, its distribution is described as a **vector field**. While we already have a standard name for a simple algebraic quantity which is assigned a value at all points of a spatial region (it is called a **function**), in physics contexts it may also be referred to as a **scalar field**.

Physicists need to be able to characterize the rate at which the values of vectors (and also scalars) change with position, and this is most effectively done by introducing differential vector operator concepts. It turns out that there are a large number of relations between these differential operators, and it is our current objective to identify such relations and learn how to use them.

Gradient, ∇

Our first differential operator is that known as the **gradient**, which characterizes the change of a scalar quantity, here φ, with position. Working in $\mathbb{R}^3$, and labeling the coordinates x_1, x_2, x_3, we write $\varphi(\mathbf{r})$ as the value of φ at the point $\mathbf{r} = x_1\hat{\mathbf{e}}_1 + x_2\hat{\mathbf{e}}_2 + x_3\hat{\mathbf{e}}_3$, and consider the effect of small changes dx_1, dx_2, dx_3, respectively, in x_1, x_2, and x_3. This situation corresponds to that discussed in Section 1.9, where we introduced **partial derivatives** to describe how a function of several variables (there x, y, and z) changes its value when these variables are changed by respective amounts dx, dy, and dz. The equation governing this process is Eq. (1.141).

To first order in the differentials dx_i, φ in our present problem changes by an amount

$$d\varphi = \left(\frac{\partial\varphi}{\partial x_1}\right)dx_1 + \left(\frac{\partial\varphi}{\partial x_2}\right)dx_2 + \left(\frac{\partial\varphi}{\partial x_3}\right)dx_3, \tag{3.38}$$

which is of the form corresponding to the dot product of

$$\boldsymbol{\nabla}\varphi = \begin{pmatrix}\partial\varphi/\partial x_1\\ \partial\varphi/\partial x_2\\ \partial\varphi/\partial x_3\end{pmatrix} \quad \text{and} \quad d\mathbf{r} = \begin{pmatrix}dx_1\\ dx_2\\ dx_3\end{pmatrix}.$$

These quantities can also be written

$$\nabla\varphi = \left(\frac{\partial\varphi}{\partial x_1}\right)\hat{\mathbf{e}}_1 + \left(\frac{\partial\varphi}{\partial x_2}\right)\hat{\mathbf{e}}_2 + \left(\frac{\partial\varphi}{\partial x_3}\right)\hat{\mathbf{e}}_3, \tag{3.39}$$

$$d\mathbf{r} = dx_1\hat{\mathbf{e}}_1 + dx_2\hat{\mathbf{e}}_2 + dx_3\hat{\mathbf{e}}_3, \tag{3.40}$$

in terms of which we have

$$d\varphi = (\nabla\varphi)\cdot d\mathbf{r}. \tag{3.41}$$

We have given the 3×1 matrix of derivatives the name $\nabla\varphi$ (often referred to in speech as "del phi" or "grad phi"); we give the differential of position its customary name $d\mathbf{r}$.

The notation of Eqs. (3.39) and (3.41) is really only appropriate if $\nabla\varphi$ is actually a vector, because the utility of the present approach depends on our ability to use it in coordinate systems of arbitrary orientation. To prove that $\nabla\varphi$ is a vector, we must show that it transforms under rotation of the coordinate system according to

$$(\nabla\varphi)' = \mathsf{S}\,(\nabla\varphi). \tag{3.42}$$

Taking S in the form given in Eq. (3.33), we examine $\mathsf{S}(\nabla\varphi)$. We have

$$\mathsf{S}(\nabla\varphi) = \begin{pmatrix} \partial x_1/\partial x_1' & \partial x_2/\partial x_1' & \partial x_3/\partial x_1' \\ \partial x_1/\partial x_2' & \partial x_2/\partial x_2' & \partial x_3/\partial x_2' \\ \partial x_1/\partial x_3' & \partial x_2/\partial x_3' & \partial x_3/\partial x_3' \end{pmatrix} \begin{pmatrix} \partial\varphi/\partial x_1 \\ \partial\varphi/\partial x_2 \\ \partial\varphi/\partial x_3 \end{pmatrix}$$

$$= \begin{pmatrix} \displaystyle\sum_{\nu=1}^{3} \frac{\partial x_\nu}{\partial x_1'}\frac{\partial\varphi}{\partial x_\nu} \\ \displaystyle\sum_{\nu=1}^{3} \frac{\partial x_\nu}{\partial x_2'}\frac{\partial\varphi}{\partial x_\nu} \\ \displaystyle\sum_{\nu=1}^{3} \frac{\partial x_\nu}{\partial x_3'}\frac{\partial\varphi}{\partial x_\nu} \end{pmatrix}. \tag{3.43}$$

Each of the elements in the final expression in Eq. (3.43) is a chain-rule expression for $\partial\varphi/\partial x_\mu'$, $\mu = 1, 2, 3$, showing that the transformation did produce $(\nabla\varphi)'$, the representation of $\nabla\varphi$ in the rotated coordinates.

Having now established the legitimacy of the form $\nabla\varphi$, we proceed to give ∇ a life of its own. We therefore define (calling the coordinates x, y, z)

$$\nabla = \hat{\mathbf{e}}_x\frac{\partial}{\partial x} + \hat{\mathbf{e}}_y\frac{\partial}{\partial y} + \hat{\mathbf{e}}_z\frac{\partial}{\partial z}. \tag{3.44}$$

We note that ∇ is a **vector differential operator**, capable of operating on a scalar (such as φ) to produce a vector as the result of the operation. Because a differential operator only operates on what is to its right, we have to be careful to maintain the correct order in expressions involving ∇, and we have to use parentheses when necessary to avoid ambiguity as to what is to be differentiated.

The gradient of a scalar is extremely important in physics and engineering, as it expresses the relation between a force field $\mathbf{F}(\mathbf{r})$ experienced by an object at $\mathbf{r}$ and the related potential $V(\mathbf{r})$,

$$\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r}). \tag{3.45}$$

The minus sign in Eq. (3.45) is important; it causes the force exerted by the field to be in a direction that lowers the potential. We consider later (in Section 3.9) the conditions that must be satisfied if a potential corresponding to a given force can exit.

The gradient has a simple geometric interpretation. From Eq. (3.41), we see that, if $d\mathbf{r}$ is constrained to have a fixed magnitude, the direction of $d\mathbf{r}$ that maximizes $d\varphi$ will be when $\nabla\varphi$ and $d\mathbf{r}$ are collinear. So, the direction of most rapid increase in φ is the gradient direction, and the magnitude of the gradient is the directional derivative of φ in that direction. We now see that $-\nabla V$, in Eq. (3.45), is the direction of most rapid *decrease* in V, and is the direction of the force associated with the potential V.

Example 3.5.1 GRADIENT OF r^n

As a first step toward computation of ∇r^n, let's look at the even simpler ∇r. We begin by writing $r = (x^2 + y^2 + z^2)^{1/2}$, from which we get

$$\frac{\partial r}{\partial x} = \frac{x}{(x^2 + y^2 + z^2)^{1/2}} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r}. \tag{3.46}$$

From these formulas we construct

$$\nabla r = \frac{x}{r}\hat{\mathbf{e}}_x + \frac{y}{r}\hat{\mathbf{e}}_y + \frac{z}{r}\hat{\mathbf{e}}_z = \frac{1}{r}(x\hat{\mathbf{e}}_x + y\hat{\mathbf{e}}_y + z\hat{\mathbf{e}}_z) = \frac{\mathbf{r}}{r}. \tag{3.47}$$

The result is a unit vector in the direction of $\mathbf{r}$, denoted $\hat{\mathbf{r}}$. For future reference, we note that

$$\hat{\mathbf{r}} = \frac{x}{r}\hat{\mathbf{e}}_x + \frac{y}{r}\hat{\mathbf{e}}_y + \frac{z}{r}\hat{\mathbf{e}}_z \tag{3.48}$$

and that Eq. (3.47) takes the form

$$\nabla r = \hat{\mathbf{r}}. \tag{3.49}$$

The geometry of $\mathbf{r}$ and $\hat{\mathbf{r}}$ is illustrated in Fig. 3.8.

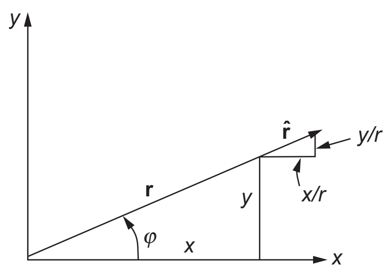

FIGURE 3.8 Unit vector $\hat{\mathbf{r}}$ (in xy-plane).

Continuing now to ∇r^n, we have

$$\frac{\partial r^n}{\partial x} = nr^{n-1}\frac{\partial r}{\partial x},$$

with corresponding results for the y and z derivatives. We get

$$\nabla r^n = nr^{n-1}\nabla r = nr^{n-1}\hat{\mathbf{r}}. \tag{3.50}$$

■

Example 3.5.2 Coulomb's Law

In electrostatics, it is well known that a point charge produces a potential proportional to $1/r$, where r is the distance from the charge. To check that this is consistent with the Coulomb force law, we compute

$$\mathbf{F} = -\nabla\left(\frac{1}{r}\right).$$

This is a case of Eq. (3.50) with $n = -1$, and we get the expected result

$$\mathbf{F} = \frac{1}{r^2}\hat{\mathbf{r}}.$$

■

Example 3.5.3 General Radial Potential

Another situation of frequent occurrence is that the potential may be a function only of the radial distance from the origin, i.e., $\varphi = f(r)$. We then calculate

$$\frac{\partial \varphi}{\partial x} = \frac{df(r)}{dr}\frac{\partial r}{\partial x}, \quad \text{etc.},$$

which leads, invoking Eq. (3.49), to

$$\nabla\varphi = \frac{df(r)}{dr}\nabla\mathbf{r} = \frac{df(r)}{dr}\hat{\mathbf{r}}. \tag{3.51}$$

This result is in accord with intuition; the direction of maximum increase in φ must be radial, and numerically equal to $d\varphi/dr$. ■

Divergence, $\nabla\cdot$

The **divergence** of a vector **A** is defined as the operation

$$\nabla\cdot\mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \tag{3.52}$$

The above formula is exactly what one might expect given both the vector and differential-operator character of ∇.

After looking at some examples of the calculation of the divergence, we will discuss its physical significance.

Example 3.5.4 Divergence of Coordinate Vector

Calculate $\nabla \cdot \mathbf{r}$:

$$\nabla \cdot \mathbf{r} = \left(\hat{\mathbf{e}}_x \frac{\partial}{\partial x} + \hat{\mathbf{e}}_y \frac{\partial}{\partial y} + \hat{\mathbf{e}}_z \frac{\partial}{\partial z} \right) \cdot \left(\hat{\mathbf{e}}_x x + \hat{\mathbf{e}}_y y + \hat{\mathbf{e}}_z z \right)$$
$$= \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z},$$

which reduces to $\nabla \cdot \mathbf{r} = 3$. ■

Example 3.5.5 Divergence of Central Force Field

Consider next $\nabla \cdot f(r)\hat{\mathbf{r}}$. Using Eq. (3.48), we write

$$\nabla \cdot f(r)\hat{\mathbf{r}} = \left(\hat{\mathbf{e}}_x \frac{\partial}{\partial x} + \hat{\mathbf{e}}_y \frac{\partial}{\partial y} + \hat{\mathbf{e}}_z \frac{\partial}{\partial z} \right) \cdot \left(\frac{xf(r)}{r}\hat{\mathbf{e}}_x + \frac{yf(r)}{r}\hat{\mathbf{e}}_y + \frac{zf(r)}{r}\hat{\mathbf{e}}_z \right).$$
$$= \frac{\partial}{\partial x}\left(\frac{xf(r)}{r} \right) + \frac{\partial}{\partial y}\left(\frac{yf(r)}{r} \right) + \frac{\partial}{\partial z}\left(\frac{zf(r)}{r} \right).$$

Using

$$\frac{\partial}{\partial x}\left(\frac{xf(r)}{r} \right) = \frac{f(r)}{r} - \frac{xf(r)}{r^2}\frac{\partial r}{\partial x} + \frac{x}{r}\frac{df(r)}{dr}\frac{\partial r}{\partial x} = f(r)\left[\frac{1}{r} - \frac{x^2}{r^3} \right] + \frac{x^2}{r^2}\frac{df(r)}{dr}$$

and corresponding formulas for the y and z derivatives, we obtain after simplification

$$\nabla \cdot f(r)\hat{\mathbf{r}} = 2\frac{f(r)}{r} + \frac{df(r)}{dr}. \tag{3.53}$$

In the special case $f(r) = r^n$, Eq. (3.53) reduces to

$$\nabla \cdot r^n \hat{\mathbf{r}} = (n+2)r^{n-1}. \tag{3.54}$$

For $n = 1$, this reduces to the result of Example 3.5.4. For $n = -2$, corresponding to the Coulomb field, the divergence vanishes, except at $r = 0$, where the differentiations we performed are not defined. ■

If a vector field represents the flow of some quantity that is distributed in space, its divergence provides information as to the accumulation or depletion of that quantity at the point at which the divergence is evaluated. To gain a clearer picture of the concept, let us suppose that a vector field $\mathbf{v}(\mathbf{r})$ represents the velocity of a fluid[5] at the spatial points $\mathbf{r}$, and that $\rho(\mathbf{r})$ represents the fluid density at $\mathbf{r}$ at a given time t. Then the direction and magnitude of the flow rate at any point will be given by the product $\rho(\mathbf{r})\mathbf{v}(\mathbf{r})$.

Our objective is to calculate the net rate of change of the fluid density in a volume element at the point $\mathbf{r}$. To do so, we set up a parallelepiped of dimensions dx, dy, dz centered at $\mathbf{r}$ and with sides parallel to the xy, xz, and yz planes. See Fig. 3.9. To first order (infinitesimal $d\mathbf{r}$ and dt), the density of fluid exiting the parallelepiped per unit time

[5]It may be helpful to think of the fluid as a collection of molecules, so the number per unit volume (the density) at any point is affected by the flow in and out of a volume element at the point.

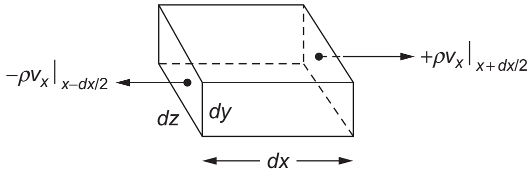

FIGURE 3.9 Outward flow of $\rho\mathbf{v}$ from a volume element in the $\pm x$ directions. The quantities $\pm\rho v_x$ must be multiplied by $dy\,dz$ to represent the total flux through the bounding surfaces at $x \pm dx/2$.

through the yz face located at $x - (dx/2)$ will be

$$\text{Flow out, face at } x - \frac{dx}{2}: \quad -(\rho v_x)\Big|_{(x-dx/2,y,z)} dy\,dz.$$

Note that only the velocity component v_x is relevant here. The other components of $\mathbf{v}$ will not cause motion through a yz face of the parallelepiped. Also, note the following: $dy\,dz$ is the area of the yz face; the average of ρv_x over the face is to first order its value at $(x-dx/2, y, z)$, as indicated, and the amount of fluid leaving per unit time can be identified as that in a column of area $dy\,dz$ and height v_x. Finally, keep in mind that **outward** flow corresponds to that in the $-x$ direction, explaining the presence of the minus sign.

We next compute the outward flow through the yz planar face at $x + dx/2$. The result is

$$\text{Flow out, face at } x + \frac{dx}{2}: \quad +(\rho v_x)\Big|_{(x+dx/2,y,z)} dy\,dz.$$

Combining these, we have for both yz faces

$$\left(-(\rho v_x)\Big|_{x-dx/2} + (\rho v_x)\Big|_{x+dx/2}\right) dy\,dz = \left(\frac{\partial(\rho v_x)}{\partial x}\right) dx\,dy\,dz.$$

Note that in combining terms at $x - dx/2$ and $x + dx/2$ we used the partial derivative notation, because all the quantities appearing here are also functions of y and z. Finally, adding corresponding contributions from the other four faces of the parallelepiped, we reach

$$\begin{aligned}\begin{matrix}\text{Net flow out}\\ \text{per unit time}\end{matrix} &= \left[\frac{\partial}{\partial x}(\rho v_x) + \frac{\partial}{\partial y}(\rho v_y) + \frac{\partial}{\partial z}(\rho v_z)\right] dx\,dy\,dz \\ &= \nabla\cdot(\rho\mathbf{v})\,dx\,dy\,dz. \end{aligned} \tag{3.55}$$

We now see that the name **divergence** is aptly chosen. As shown in Eq. (3.55), the divergence of the vector $\rho\mathbf{v}$ represents the net outflow per unit volume, per unit time. If the physical problem being described is one in which fluid (molecules) are neither created or destroyed, we will also have an **equation of continuity**, of the form

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0. \tag{3.56}$$

This equation quantifies the obvious statement that a net outflow from a volume element results in a smaller density inside the volume.

When a vector quantity is divergenceless (has zero divergence) in a spatial region, we can interpret it as describing a steady-state "fluid-conserving" flow (**flux**) within that region

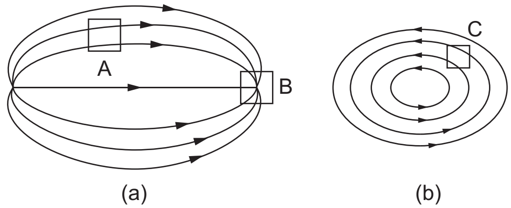

FIGURE 3.10 Flow diagrams: (a) with source and sink; (b) solenoidal. The divergence vanishes at volume elements A and C, but is negative at B.

(even if the vector field does not represent material that is moving). This is a situation that arises frequently in physics, applying in general to the magnetic field, and, in charge-free regions, also to the electric field. If we draw a diagram with lines that follow the flow paths, the lines (depending on the context) may be called **stream lines** or **lines of force**. Within a region of zero divergence, these lines must exit any volume element they enter; they cannot terminate there. However, lines will begin at points of positive divergence (sources) and end at points where the divergence is negative (sinks). Possible patterns for a vector field are shown in Fig. 3.10.

If the divergence of a vector field is zero everywhere, its lines of force will consist entirely of closed loops, as in Fig. 3.10(b); such vector fields are termed **solenoidal**. For emphasis, we write

$$\nabla \cdot \mathbf{B} = 0 \text{ everywhere} \quad \longrightarrow \quad \mathbf{B} \text{ is solenoidal.} \tag{3.57}$$

Curl, $\nabla \times$

Another possible operation with the vector operator ∇ is to take its cross product with a vector. Using the established formula for the cross product, and being careful to write the derivatives to the left of the vector on which they are to act, we obtain

$$\begin{aligned}\nabla \times \mathbf{V} &= \hat{\mathbf{e}}_x \left(\frac{\partial}{\partial y} V_z - \frac{\partial}{\partial z} V_y\right) + \hat{\mathbf{e}}_y \left(\frac{\partial}{\partial z} V_x - \frac{\partial}{\partial x} V_z\right) + \hat{\mathbf{e}}_z \left(\frac{\partial}{\partial x} V_y - \frac{\partial}{\partial y} V_x\right) \\ &= \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ V_x & V_y & V_z \end{vmatrix}.\end{aligned} \tag{3.58}$$

This vector operation is called the **curl** of $\mathbf{V}$. Note that when the determinant in Eq. (3.58) is evaluated, it must be expanded in a way that causes the derivatives in the second row to be applied to the functions in the third row (and not to anything in the top row); we will encounter this situation repeatedly, and will identify the evaluation as being **from the top down**.

Example 3.5.6 CURL OF A CENTRAL FORCE FIELD

Calculate $\nabla \times [f(r)\hat{\mathbf{r}}]$. Writing

$$\hat{\mathbf{r}} = \frac{x}{r}\hat{\mathbf{e}}_x + \frac{y}{r}\hat{\mathbf{e}}_y + \frac{z}{r}\hat{\mathbf{e}}_z,$$

and remembering that $\partial r/\partial y = y/r$ and $\partial r/\partial z = z/r$, the x-component of the result is found to be

$$\left[\nabla \times [f(r)\hat{\mathbf{r}}]\right]_x = \frac{\partial}{\partial y}\frac{zf(r)}{r} - \frac{\partial}{\partial z}\frac{yf(r)}{r}$$
$$= z\left(\frac{d}{dr}\frac{f(r)}{r}\right)\frac{\partial r}{\partial y} - y\left(\frac{d}{dr}\frac{f(r)}{r}\right)\frac{\partial r}{\partial z}$$
$$= z\left(\frac{d}{dr}\frac{f(r)}{r}\right)\frac{y}{r} - y\left(\frac{d}{dr}\frac{f(r)}{r}\right)\frac{z}{r} = 0.$$

By symmetry, the other components are also zero, yielding the final result

$$\nabla \times [f(r)\hat{\mathbf{r}}] = 0. \tag{3.59}$$

■

Example 3.5.7 A Nonzero Curl

Calculate $\mathbf{F} = \nabla \times (-y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y)$, which is of the form $\nabla \times \mathbf{b}$, where $b_x = -y$, $b_y = x$, $b_z = 0$. We have

$$F_x = \frac{\partial b_z}{\partial y} - \frac{\partial b_y}{\partial z} = 0, \quad F_y = \frac{\partial b_x}{\partial z} - \frac{\partial v_z}{\partial x} = 0, \quad F_z = \frac{\partial b_y}{\partial x} - \frac{\partial b_x}{\partial y} = 2,$$

so $\mathbf{F} = 2\hat{\mathbf{e}}_z$. ■

The results of these two examples can be better understood from a geometric interpretation of the curl operator. We proceed as follows: Given a vector field $\mathbf{B}$, consider the line integral $\oint \mathbf{B} \cdot d\mathbf{s}$ for a small closed path. The circle through the integral sign is a signal that the path is closed. For simplicity in the computations, we take a rectangular path in the xy-plane, centered at a point (x_0, y_0), of dimensions $\Delta x \times \Delta y$, as shown in Fig. 3.11. We will traverse this path in the counterclockwise direction, passing through the four segments labeled 1 through 4 in the figure. Since everywhere in this discussion $z = 0$, we do not show it explicitly.

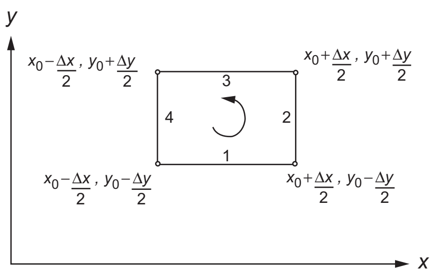

FIGURE 3.11 Path for computing circulation at (x_0, y_0).

Segment 1 of the path contributes to the integral

$$\text{Segment 1} = \int_{x_0-\Delta x/2}^{x_0+\Delta x/2} B_x(x, y_0 - \Delta y/2)\,dx \approx B_x(x_0, y_0 - \Delta y/2)\Delta x,$$

where the approximation, replacing B_x by its value at the middle of the segment, is good to first order. In a similar fashion, we have

$$\text{Segment 2} = \int_{y_0-\Delta y/2}^{y_0+\Delta y/2} B_y(x_0 + \Delta x/2, y)\,dy \approx B_y(x_0 + \Delta x/2, y_0)\Delta y,$$

$$\text{Segment 3} = \int_{x_0+\Delta x/2}^{x_0-\Delta x/2} B_x(x, y_0 + \Delta y/2)\,dx \approx -B_x(x_0, y_0 + \Delta y/2)\Delta x,$$

$$\text{Segment 4} = \int_{y_0+\Delta y/2}^{y_0-\Delta y/2} B_y(x_0 - \Delta x/2, y)\,dy \approx -B_y(x_0 - \Delta x/2, y_0)\Delta y.$$

Note that because the paths of segments 3 and 4 are in the direction of decrease in the value of the integration variable, we obtain minus signs in the contributions of these segments. Combining the contributions of Segments 1 and 3, and those of Segments 2 and 4, we have

$$\text{Segments } 1+3 = \big(B_x(x_0, y_0 - \Delta y/2) - B_x(x_0, y_0 + \Delta y/2)\big)\Delta x \approx -\frac{\partial B_x}{\partial y}\Delta y \Delta x,$$

$$\text{Segments } 2+4 = \big(B_y(x_0 + \Delta x/2, y_0) - B_y(x_0 - \Delta x/2, y_0)\big)\Delta y \approx +\frac{\partial B_y}{\partial x}\Delta x\, \Delta y.$$

Combining these contributions to obtain the value of the entire line integral, we have

$$\oint \mathbf{B}\cdot d\mathbf{s} \approx \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)\Delta x \Delta y \approx [\nabla \times \mathbf{B}]_z \Delta x \Delta y. \tag{3.60}$$

The thing to note is that a nonzero closed-loop line integral of **B** corresponds to a nonzero value of the component of $\nabla \times \mathbf{B}$ normal to the loop. In the limit of a small loop, the line integral will have a value proportional to the loop area; the value of the line integral per unit area is called the **circulation** (in fluid dynamics, it is also known as the **vorticity**). A nonzero circulation corresponds to a pattern of stream lines that form closed loops. Obviously, to form a closed loop, a stream line must curl; hence the name of the $\nabla\times$ operator.

Returning now to Example 3.5.6, we have a situation in which the lines of force must be entirely radial; there is no possibility to form closed loops. Accordingly, we found this example to have a zero curl. But, looking next at Example 3.5.7, we have a situation in which the stream lines of $-y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y$ form counterclockwise circles about the origin, and the curl is nonzero.

We close the discussion by noting that a vector whose curl is zero everywhere is termed **irrotational**. This property is in a sense the opposite of solenoidal, and deserves a parallel degree of emphasis:

$$\nabla \times \mathbf{B} = 0 \text{ everywhere} \quad \longrightarrow \quad \mathbf{B} \text{ is irrotational.} \tag{3.61}$$

Exercises

3.5.1 If $S(x, y, z) = \left(x^2 + y^2 + z^2\right)^{-3/2}$, find

(a) ∇S at the point $(1, 2, 3)$,
(b) the magnitude of the gradient of S, $|\nabla S|$ at $(1, 2, 3)$, and
(c) the direction cosines of ∇S at $(1, 2, 3)$.

3.5.2 (a) Find a unit vector perpendicular to the surface

$$x^2 + y^2 + z^2 = 3$$

at the point $(1, 1, 1)$.
(b) Derive the equation of the plane tangent to the surface at $(1, 1, 1)$.

ANS. (a) $\left(\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z\right)/\sqrt{3}$, (b) $x + y + z = 3$.

3.5.3 Given a vector $\mathbf{r}_{12} = \hat{\mathbf{e}}_x(x_1 - x_2) + \hat{\mathbf{e}}_y(y_1 - y_2) + \hat{\mathbf{e}}_z(z_1 - z_2)$, show that $\nabla_1 r_{12}$ (gradient with respect to x_1, y_1, and z_1 of the magnitude r_{12}) is a unit vector in the direction of $\mathbf{r}_{12}$.

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

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

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

3.5.6 For a particle moving in a circular orbit $\mathbf{r} = \hat{\mathbf{e}}_x r \cos\omega t + \hat{\mathbf{e}}_y r \sin\omega t$:

(a) Evaluate $\mathbf{r} \times \dot{\mathbf{r}}$, with $\dot{\mathbf{r}} = d\mathbf{r}/dt = \mathbf{v}$.
(b) Show that $\ddot{\mathbf{r}} + \omega^2 \mathbf{r} = 0$ with $\ddot{\mathbf{r}} = d\mathbf{v}/dt$.

Hint. The radius r and the angular velocity ω are constant.

ANS. (a) $\hat{\mathbf{e}}_z \omega r^2$.

3.5.7 Vector $\mathbf{A}$ satisfies the vector transformation law, Eq. (3.26). Show directly that its time derivative $d\mathbf{A}/dt$ also satisfies Eq. (3.26) and is therefore a vector.

3.5.8 Show, by differentiating components, that

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

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

just like the derivative of the product of two algebraic functions.

3.5.9 Prove $\nabla\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\nabla\times\mathbf{a}) - \mathbf{a}\cdot(\nabla\times\mathbf{b})$.

Hint. Treat as a scalar triple product.

3.5.10 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, $\mathbf{p}$ is replaced (in units with $\hbar = 1$) by the operator $-i\nabla$. Show that the quantum mechanical angular momentum operator has Cartesian components

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

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

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

3.5.11 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.

3.5.12 With the aid of the results of Exercise 3.5.11, show that if two 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

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

3.5.13 Prove that the stream lines of $\mathbf{b}$ in of Example 3.5.7 are counterclockwise circles.

3.6 Differential Vector Operators: Further Properties

Successive Applications of ∇

Interesting results are obtained when we operate with ∇ on the differential vector operator forms we have already introduced. The possible results include the following:

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

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

Laplacian

The first of these expressions, $\nabla \cdot \nabla\varphi$, the divergence of the gradient, is named the Laplacian of φ. We have

$$\nabla \cdot \nabla\varphi = \left(\hat{\mathbf{e}}_x \frac{\partial}{\partial x} + \hat{\mathbf{e}}_y \frac{\partial}{\partial y} + \hat{\mathbf{e}}_z \frac{\partial}{\partial z}\right) \cdot \left(\hat{\mathbf{e}}_x \frac{\partial \varphi}{\partial x} + \hat{\mathbf{e}}_y \frac{\partial \varphi}{\partial y} + \hat{\mathbf{e}}_z \frac{\partial \varphi}{\partial z}\right)$$
$$= \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2}. \tag{3.62}$$

When φ is the electrostatic potential, we have

$$\nabla \cdot \nabla\varphi = 0 \tag{3.63}$$

at points where the charge density vanishes, which is Laplace's equation of electrostatics. Often the combination $\nabla \cdot \nabla$ is written ∇^2, or Δ in the older European literature.

Example 3.6.1 Laplacian of a Central Field Potential

Calculate $\nabla^2\varphi(r)$. Using Eq. (3.51) to evaluate $\nabla\varphi$ and then Eq. (3.53) for the divergence, we have

$$\nabla^2\varphi(r) = \nabla \cdot \nabla\varphi(r) = \nabla \cdot \frac{d\varphi(r)}{dr}\hat{\mathbf{e}}_r = \frac{2}{r}\frac{d\varphi(r)}{dr} + \frac{d^2\varphi(r)}{dr^2}.$$

We get a term in addition to $d^2\varphi/dr^2$ because $\hat{\mathbf{e}}_r$ has a direction that depends on $\mathbf{r}$.

In the special case $\varphi(r) = r^n$, this reduces to

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

This vanishes for $n = 0$ (φ =constant) and for $n = -1$ (Coulomb potential). For $n = -1$, our derivation fails for $\mathbf{r} = 0$, where the derivatives are undefined. ■

Irrotational and Solenoidal Vector Fields

Expression (b), the second of our five forms involving two ∇ operators, may be written as a determinant:

$$\nabla \times \nabla\varphi = \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial\varphi/\partial x & \partial\varphi/\partial y & \partial\varphi/\partial z \end{vmatrix} = \begin{vmatrix} \hat{\mathbf{e}}_x & \hat{\mathbf{e}}_y & \hat{\mathbf{e}}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \end{vmatrix} \varphi = 0.$$

Because the determinant is to be evaluated from the top down, it is meaningful to move φ outside and to its right, leaving a determinant with two identical rows and yielding the indicated value of zero. We are thereby actually assuming that the order of the partial differentiations can be reversed, which is true so long as these second derivatives of φ are continuous.

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

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{V})=\begin{vmatrix}\partial/\partial x & \partial/\partial y & \partial/\partial z\\ \partial/\partial x & \partial/\partial y & \partial/\partial z\\ V_x & V_y & V_z\end{vmatrix}=0.$$

This determinant also has two identical rows and yields zero if $\mathbf{V}$ has sufficient continuity.

These two vanishing results tell us that any gradient has a vanishing curl and is therefore **irrotational**, and that any curl has a vanishing divergence, and is therefore **solenoidal**. These properties are of such importance that we set them out here in display form:

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\varphi=0,\quad \text{all } \varphi, \tag{3.64}$$

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{V})=0,\quad \text{all } \mathbf{V}. \tag{3.65}$$

Maxwell's Equations

The unification of electric and magnetic phenomena that is encapsulated in Maxwell's equations provides an excellent example of the use of differential vector operators. In SI units, these equations take the form

$$\boldsymbol{\nabla}\cdot\mathbf{B}=0, \tag{3.66}$$

$$\boldsymbol{\nabla}\cdot\mathbf{E}=\frac{\rho}{\varepsilon_0}, \tag{3.67}$$

$$\boldsymbol{\nabla}\times\mathbf{B}=\varepsilon_0\mu_0\frac{\partial\mathbf{E}}{\partial t}+\mu_0\mathbf{J}, \tag{3.68}$$

$$\boldsymbol{\nabla}\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t}. \tag{3.69}$$

Here $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic induction field, ρ is the charge density, $\mathbf{J}$ is the current density, ε_0 is the electric permittivity, and μ_0 is the magnetic permeability, so $\varepsilon_0\mu_0=1/c^2$, where c is the velocity of light.

Vector Laplacian

Expressions (c) and (e) in the list at the beginning of this section satisfy the relation

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{V})=\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{V})-\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\mathbf{V}. \tag{3.70}$$

The term $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\mathbf{V}$, which is called the **vector Laplacian** and sometimes written $\boldsymbol{\nabla}^2\mathbf{V}$, has prior to this point not been defined; Eq. (3.70) (solved for $\boldsymbol{\nabla}^2\mathbf{V}$) can be taken to be its definition. In Cartesian coordinates, $\boldsymbol{\nabla}^2\mathbf{V}$ is a vector whose i component is $\boldsymbol{\nabla}^2 V_i$, and that fact can be confirmed either by direct component expansion or by applying the BAC–CAB rule, Eq. (3.18), with care always to place $\mathbf{V}$ so that the differential operators act on it. While Eq. (3.70) is general, $\boldsymbol{\nabla}^2\mathbf{V}$ separates into Laplacians for the components of $\mathbf{V}$ only in Cartesian coordinates.

Example 3.6.2 Electromagnetic Wave Equation

Even in vacuum, Maxwell's equations can describe electromagnetic waves. To derive an electromagnetic wave equation, we start by taking the time derivative of Eq. (3.68) for the case $\mathbf{J} = 0$, and the curl of Eq. (3.69). We then have

$$\frac{\partial}{\partial t}\nabla \times \mathbf{B} = \epsilon_0\mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2},$$

$$\nabla \times (\nabla \times \mathbf{E}) = -\frac{\partial}{\partial t}\nabla \times \mathbf{B} = -\epsilon_0\mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

We now have an equation that involves only $\mathbf{E}$; it can be brought to a more convenient form by applying Eq. (3.70), dropping the first term on the right of that equation because, in vacuum, $\nabla \cdot \mathbf{E} = 0$. The result is the vector electromagnetic wave equation for $\mathbf{E}$,

$$\nabla^2 \mathbf{E} = \epsilon_0\mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{3.71}$$

Equation (3.71) separates into three scalar wave equations, each involving the (scalar) Laplacian. There is a separate equation for each Cartesian component of $\mathbf{E}$. ■

Miscellaneous Vector Identities

Our introduction of differential vector operators is now formally complete, but we present two further examples to illustrate how the relationships between these operators can be manipulated to obtain useful vector identities.

Example 3.6.3 Divergence and Curl of a Product

First, simplify $\nabla \cdot (f\mathbf{V})$, where f and $\mathbf{V}$ are, respectively, scalar and vector functions. Working with the components,

$$\begin{aligned}
\nabla \cdot (f\mathbf{V}) &= \frac{\partial}{\partial x}(fV_x) + \frac{\partial}{\partial y}(fV_y) + \frac{\partial}{\partial z}(fV_z) \\
&= \frac{\partial f}{\partial x}V_x + f\frac{\partial V_x}{\partial x} + \frac{\partial f}{\partial y}V_y + f\frac{\partial V_y}{\partial y} + \frac{\partial f}{\partial z}V_z + f\frac{\partial V_z}{\partial z} \\
&= (\nabla f)\cdot \mathbf{V} + f\nabla \cdot \mathbf{V}.
\end{aligned} \tag{3.72}$$

Now simplify $\nabla \times (f\mathbf{V})$. Consider the x-component:

$$\frac{\partial}{\partial y}(fV_z) - \frac{\partial}{\partial z}(fV_y) = f\left[\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right] + \left[\frac{\partial f}{\partial y}V_z - \frac{\partial f}{\partial z}V_y\right].$$

This is the x-component of $f(\nabla \times \mathbf{V}) + (\nabla f) \times \mathbf{V}$, so we have

$$\nabla \times (f\mathbf{V}) = f(\nabla \times \mathbf{V}) + (\nabla f) \times \mathbf{V}. \tag{3.73}$$

■

Example 3.6.4 GRADIENT OF A DOT PRODUCT

Verify that

$$\nabla(\mathbf{A}\cdot\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} + (\mathbf{A}\cdot\nabla)\mathbf{B} + \mathbf{B}\times(\nabla\times\mathbf{A}) + \mathbf{A}\times(\nabla\times\mathbf{B}). \tag{3.74}$$

This problem is easier to solve if we recognize that $\nabla(\mathbf{A}\cdot\mathbf{B})$ is a type of term that appears in the BAC–CAB expansion of a vector triple product, Eq. (3.18). From that equation, we have

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

where we placed $\mathbf{B}$ at the end of the final term because ∇ must act on it. We write ∇_B to indicate an operation our notation is not really equipped to handle. In this term, ∇ acts only on $\mathbf{B}$, because $\mathbf{A}$ appeared to its left on the left-hand side of the equation. Interchanging the roles of $\mathbf{A}$ and $\mathbf{B}$, we also have

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

where ∇_A acts only on $\mathbf{A}$. Adding these two equations together, noting that $\nabla_B + \nabla_A$ is simply an unrestricted ∇, we recover Eq. (3.74). ■

Exercises

3.6.1 Show that $\mathbf{u}\times\mathbf{v}$ is solenoidal if $\mathbf{u}$ and $\mathbf{v}$ are each irrotational.

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

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

3.6.4 If a vector function $\mathbf{V}(x, y, z)$ is not irrotational, show that if there exists a scalar function $g(x, y, z)$ such that $g\mathbf{V}$ is irrotational, then

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

3.6.5 Verify the vector identity

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

3.6.6 As an alternative to the vector identity of Example 3.6.4 show that

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

3.6.7 Verify the identity

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

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

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

3.6.9 Verify Eq. (3.70),

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

by direct expansion in Cartesian coordinates.

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

3.6.11 You are given that the curl of **F** equals the curl of **G**. Show that **F** and **G** may differ by (a) a constant and (b) a gradient of a scalar function.

3.6.12 The Navier-Stokes equation of hydrodynamics contains a nonlinear term of the form $(\mathbf{v} \cdot \nabla)\mathbf{v}$. Show that the curl of this term may be written as $-\nabla \times [\mathbf{v} \times (\nabla \times \mathbf{v})]$.

3.6.13 Prove that $(\nabla u) \times (\nabla v)$ is solenoidal, where u and v are differentiable scalar functions.

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

3.6.15 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.

3.6.16 The theory of heat conduction leads to an equation

$$\nabla^2 \Psi = k\,|\nabla \Phi|^2,$$

where Φ is a potential satisfying Laplace's equation: $\nabla^2 \Phi = 0$. Show that a solution of this equation is $\Psi = k\Phi^2/2$.

3.6.17 Given the three matrices

$$\mathsf{M}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \mathsf{M}_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix},$$

and

$$\mathsf{M}_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

show that the matrix-vector equation

$$\left(\mathsf{M} \cdot \nabla + \mathbf{1}_3 \frac{1}{c}\frac{\partial}{\partial t}\right) \boldsymbol{\psi} = 0$$

reproduces Maxwell's equations in vacuum. Here $\boldsymbol{\psi}$ is a column vector with components $\psi_j = B_j - iE_j/c$, $j = x, y, z$. Note that $\varepsilon_0\mu_0 = 1/c^2$ and that $\mathbf{1}_3$ is the 3×3 unit matrix.

3.6.18 Using the Pauli matrices $\boldsymbol{\sigma}_i$ of Eq. (2.28), show that

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = (\mathbf{a} \cdot \mathbf{b})\mathbf{1}_2 + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}).$$

Here

$$\boldsymbol{\sigma} \equiv \hat{\mathbf{e}}_x\boldsymbol{\sigma}_1 + \hat{\mathbf{e}}_y\boldsymbol{\sigma}_2 + \hat{\mathbf{e}}_z\boldsymbol{\sigma}_3,$$

a and **b** are ordinary vectors, and $\mathbf{1}_2$ is the 2×2 unit matrix.

3.7 Vector Integration

In physics, vectors occur in line, surface, and volume integrals. At least in principle, these integrals can be decomposed into scalar integrals involving the vector components; there are some useful general observations to make at this time.

Line Integrals

Possible forms for line integrals include the following:

$$\int_C \varphi d\mathbf{r}, \quad \int_C \mathbf{F} \cdot d\mathbf{r}, \quad \int_C \mathbf{V} \times d\mathbf{r}. \tag{3.75}$$

In each of these the integral is over some path C that may be open (with starting and endpoints distinct) or closed (forming a loop). Inserting the form of $d\mathbf{r}$, the first of these integrals reduces immediately to

$$\int_C \varphi d\mathbf{r} = \hat{\mathbf{e}}_x \int_C \varphi(x, y, z)\, dx + \hat{\mathbf{e}}_y \int_C \varphi(x, y, z)\, dy + \hat{\mathbf{e}}_z \int_C \varphi(x, y, z)\, dz. \tag{3.76}$$

The unit vectors need not remain within the integral beause they are constant in both magnitude and direction.

The integrals in Eq. (3.76) are one-dimensional scalar integrals. Note, however, that the integral over x cannot be evaluated unless y and z are known in terms of x; similar observations apply for the integrals over y and z. This means that the path C must be specified. Unless φ has special properties, the value of the integral will depend on the path.

The other integrals in Eq. (3.75) can be handled similarly. For the second integral, which is of common occurrence, being that which evaluates the work associated with displacement on the path C, we have:

$$W = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_C F_x(x, y, z)\, dx + \int_C F_y(x, y, z)\, dy + \int_C F_z(x, y, z)\, dz. \tag{3.77}$$

Example 3.7.1 LINE INTEGRALS

We consider two integrals in 2-D space:

$$I_C = \int_C \varphi(x, y) d\mathbf{r}, \quad \text{with } \varphi(x, y) = 1,$$

$$J_C = \int_C \mathbf{F}(x, y) \cdot d\mathbf{r}, \quad \text{with } \mathbf{F}(x, y) = -y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y.$$

We perform integrations in the xy-plane from (0,0) to (1,1) by the two different paths shown in Fig. 3.12:

Path C_1 is $(0, 0) \to (1, 0) \to (1, 1)$,
Path C_2 is the straight line $(0, 0) \to (1, 1)$.

For the first segment of C_1, x ranges from 0 to 1 while y is fixed at zero. For the second segment, y ranges from 0 to 1 while $x = 1$. Thus,

$$I_{C_1} = \hat{\mathbf{e}}_x \int_0^1 dx\varphi(x, 0) + \hat{\mathbf{e}}_y \int_0^1 dy\varphi(1, y) = \hat{\mathbf{e}}_x \int_0^1 dx + \hat{\mathbf{e}}_y \int_0^1 dy = \hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y,$$

$$J_{C_1} = \int_0^1 dx\, F_x(x, 0) + \int_0^1 dy\, F_y(1, y) = \int_0^1 = \int_0^1 dx(0) + \int_0^1 dy(1) = 1.$$

On Path 2, both dx and dy range from 0 to 1, with $x = y$ at all points of the path. Thus,

$$I_{C_2} = \hat{\mathbf{e}}_x \int_0^1 dx\, \varphi(x, x) + \hat{\mathbf{e}}_y \int_0^1 dy\, \varphi(y, y) = \hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y,$$

$$J_{C_2} = \int_0^1 dx F_x(x, x) + \int_0^1 dy F_y(y, y) = \int_0^1 dx(-x) + \int_0^1 dy(y) = -\frac{1}{2} + \frac{1}{2} = 0.$$

We see that integral I is independent of the path from (0,0) to (1,1), a nearly trivial special case, while the integral J is not. ■

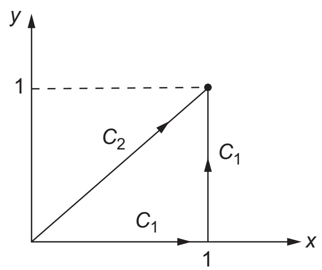

FIGURE 3.12 Line integration paths.

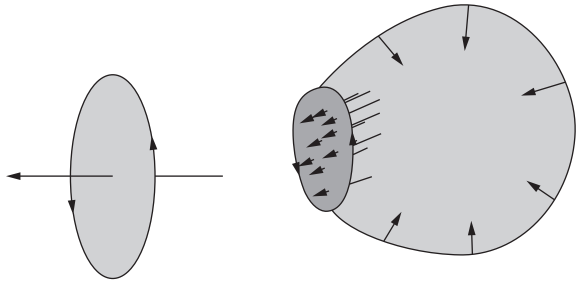

FIGURE 3.13 Positive normal directions: left, disk; right, spherical surface with hole.

Surface Integrals

Surface integrals appear in the same forms as line integrals, the element of area being a vector, $d\boldsymbol{\sigma}$, normal to the surface:

$$\int \varphi d\boldsymbol{\sigma}, \quad \int \mathbf{V}\cdot d\boldsymbol{\sigma}, \quad \int \mathbf{V}\times d\boldsymbol{\sigma}.$$

Often $d\boldsymbol{\sigma}$ is written $\hat{\mathbf{n}}\,dA$, where $\hat{\mathbf{n}}$ is a unit vector indicating the normal direction. There are two conventions for choosing the positive direction. First, if the surface is closed (has no boundary), we agree to take the outward normal as positive. Second, for an open surface, the positive normal depends on the direction in which the perimeter of the surface is traversed. Starting from an arbitrary point on the perimeter, we define a vector $\mathbf{u}$ to be in the direction of travel along the perimeter, and define a second vector $\mathbf{v}$ at our perimeter point but tangent to and lying on the surface. We then take $\mathbf{u}\times\mathbf{v}$ as the positive normal direction. This corresponds to a right-hand rule, and is illustrated in Fig. 3.13. It is necessary to define the orientation carefully so as to deal with cases such as that of Fig. 3.13, right.

The dot-product form is by far the most commonly encountered surface integral, as it corresponds to a flow or flux through the given surface.

Example 3.7.2 A SURFACE INTEGRAL

Consider a surface integral of the form $I=\int_S \mathbf{B}\cdot d\boldsymbol{\sigma}$ over the surface of a tetrahedron whose vertices are at the origin and at the points (1,0,0), (0,1,0), and (0,0,1), with $\mathbf{B}=(x+1)\hat{\mathbf{e}}_x+y\hat{\mathbf{e}}_y-z\hat{\mathbf{e}}_z$. See Fig. 3.14.

The surface consists of four triangles, which can be identified and their contributions evaluated, as follows:

1. On the xy-plane ($z=0$), vertices at $(x,y)=$ (0,0), (1,0), and (0,1); direction of outward normal is $-\hat{\mathbf{e}}_z$, so $d\boldsymbol{\sigma}=-\hat{\mathbf{e}}_z dA$ ($dA=$ element of area on this triangle). Here, $\mathbf{B}=(x+1)\hat{\mathbf{e}}_x+y\hat{\mathbf{e}}_y$, and $\mathbf{B}\cdot d\boldsymbol{\sigma}=0$. So there is no contribution to I.
2. On the xz plane ($y=0$), vertices at $(x,z)=$ (0,0), (1,0), and (0,1); direction of outward normal is $-\hat{\mathbf{e}}_y$, so $d\boldsymbol{\sigma}=-\hat{\mathbf{e}}_y dA$. On this triangle, $\mathbf{B}=(x+1)\hat{\mathbf{e}}_x-z\hat{\mathbf{e}}_z$, Again, $\mathbf{B}\cdot d\boldsymbol{\sigma}=0$. There is no contribution to I.
3. On the yz plane ($x=0$), vertices at $(y,z)=$ (0,0), (1,0), and (0,1); direction of outward normal is $-\hat{\mathbf{e}}_x$, so $d\boldsymbol{\sigma}=-\hat{\mathbf{e}}_x dA$. Here, $\mathbf{B}=\hat{\mathbf{e}}_x+y\hat{\mathbf{e}}_y-z\hat{\mathbf{e}}_z$, and

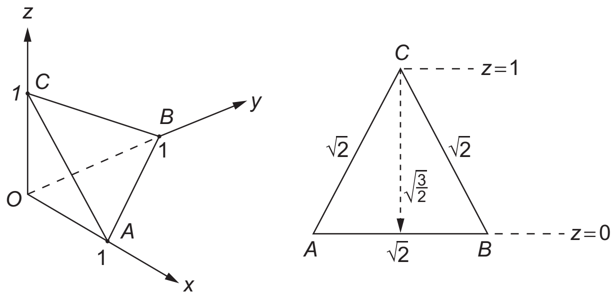

FIGURE 3.14 Tetrahedron, and detail of the oblique face.

$\mathbf{B} \cdot d\boldsymbol{\sigma} = (-1)dA$; the contribution to I is -1 times the area of the triangle (=1/2), or $I_3 = -1/2$.

4. Obliquely oriented, vertices at $(x, y, z) = (1,0,0), (0,1,0), (0,0,1)$; direction of outward normal is $\hat{\mathbf{n}} = (\hat{\mathbf{e}}_x + \hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z)/\sqrt{3}$, and $d\boldsymbol{\sigma} = \hat{\mathbf{n}} dA$. Using also $\mathbf{B} = (x+1)\hat{\mathbf{e}}_x + y\hat{\mathbf{e}}_y - z\hat{\mathbf{e}}_z$, this contribution to I becomes

$$I_4 = \int\limits_{\Delta_4} \frac{x+1+y-z}{\sqrt{3}} dA = \int\limits_{\Delta_4} \frac{2(1-z)}{\sqrt{3}} dA,$$

where we have used the fact that on this triangle, $x + y + z = 1$.

To complete the evaluation, we note that the geometry of the triangle is as shown in Fig. 3.14, that the width of the triangle at height z is $\sqrt{2}\,(1-z)$, and a change dz in z produces a displacement $\sqrt{3/2}dz$ on the triangle. I_4 therefore can be written

$$I_4 = \int\limits_0^1 2(1-z)^2 dz = \frac{2}{3}.$$

Combining the nonzero contributions I_3 and I_4, we obtain the final result

$$I = -\frac{1}{2} + \frac{2}{3} = \frac{1}{6}.$$

■

Volume Integrals

Volume integrals are somewhat simpler, because the volume element $d\tau$ is a scalar quantity. Sometimes $d\tau$ is written d^3r, or d^3x when the coordinates were designated (x_1, x_2, x_3). In the literature, the form $d\mathbf{r}$ is frequently encountered, but in contexts that usually reveal that it is a synonym for $d\tau$, and not a vector quantity. The volume integrals under consideration here are of the form

$$\int \mathbf{V}\, d\tau = \hat{\mathbf{e}}_x \int V_x\, d\tau + \hat{\mathbf{e}}_y \int V_y\, d\tau + \hat{\mathbf{e}}_z \int V_z\, d\tau.$$

The integral reduces to a vector sum of scalar integrals.

Some volume integrals contain vector quantities in combinations that are actually scalar. Often these can be rearranged by applying techniques such as integration by parts.

Example 3.7.3 INTEGRATION BY PARTS

Consider an integral over all space of the form $\int \mathbf{A}(\mathbf{r})\boldsymbol{\nabla}\cdot f(\mathbf{r})d^3r$ in the frequently occurring special case in which either f or $\mathbf{A}$ vanish sufficiently strongly at infinity. Expanding the integrand into components,

$$\begin{aligned}\int \mathbf{A}(\mathbf{r})\cdot\boldsymbol{\nabla} f(\mathbf{r})d^3r &= \iint dy\,dz\left[A_x f\big|_{x=-\infty}^{\infty} - \int f\frac{\partial A_x}{\partial x}\,dx\right] + \cdots \\ &= -\iiint f\frac{\partial A_x}{\partial x}\,dx\,dy\,dz - \iiint f\frac{\partial A_y}{\partial y}dx\,dy\,dz - \iiint f\frac{\partial A_z}{\partial z}dx\,dy\,dz \\ &= -\int f(r)\boldsymbol{\nabla}\cdot\mathbf{A}(\mathbf{r})d^3r. \qquad (3.78)\end{aligned}$$

For example, if $\mathbf{A} = e^{ikz}\hat{\mathbf{p}}$ describes a photon with a constant polarization vector in the direction $\hat{\mathbf{p}}$ and $\psi(\mathbf{r})$ is a bound-state wave function (so it vanishes at infinity), then

$$\int e^{ikz}\hat{\mathbf{p}}\cdot\boldsymbol{\nabla}\psi(\mathbf{r})d^3r = -(\hat{\mathbf{p}}\cdot\hat{\mathbf{e}}_z)\int \psi(\mathbf{r})\frac{de^{ikz}}{dz}d^3r = -ik(\hat{\mathbf{p}}\cdot\hat{\mathbf{e}}_z)\int \psi(\mathbf{r})e^{ikz}d^3r.$$

Only the z-component of the gradient contributes to the integral.

Analogous rearrangements (assuming the integrated terms vanish at infinity) include

$$\int f(\mathbf{r})\boldsymbol{\nabla}\cdot\mathbf{A}(\mathbf{r})d^3r = -\int \mathbf{A}(\mathbf{r})\cdot\boldsymbol{\nabla} f(\mathbf{r})d^3r, \qquad (3.79)$$

$$\int \mathbf{C}(\mathbf{r})\cdot\big(\boldsymbol{\nabla}\times\mathbf{A}(\mathbf{r})\big)d^3r = \int \mathbf{A}(\mathbf{r})\cdot\big(\boldsymbol{\nabla}\times\mathbf{C}(\mathbf{r})\big)d^3r. \qquad (3.80)$$

In the cross-product example, the sign change from the integration by parts combines with the signs from the cross product to give the result shown. ■

Exercises

3.7.1 The origin and the three vectors **A**, **B**, and **C** (all of which start at the origin) define a tetrahedron. Taking the outward direction as positive, calculate the total vector area of the four tetrahedral surfaces.

3.7.2 Find the work $\oint \mathbf{F}\cdot d\mathbf{r}$ done moving on a unit circle in the xy-plane, doing work **against** a force field given by

$$\mathbf{F} = \frac{-\hat{\mathbf{e}}_x y}{x^2+y^2} + \frac{\hat{\mathbf{e}}_y x}{x^2+y^2}:$$

(a) Counterclockwise from 0 to π,
(b) Clockwise from 0 to $-\pi$.

Note that the work done depends on the path.

3.7.3 Calculate the work you do in going from point $(1, 1)$ to point $(3, 3)$. The force **you exert** is given by

$$\mathbf{F} = \hat{\mathbf{e}}_\mathbf{x}(x - y) + \hat{\mathbf{e}}_\mathbf{y}(x + y).$$

Specify clearly the path you choose. Note that this force field is nonconservative.

3.7.4 Evaluate $\oint \mathbf{r} \cdot d\mathbf{r}$ for a closed path of your choosing.

3.7.5 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 $\mathbf{r} \cdot d\boldsymbol{\sigma}$ is zero for three of the surfaces and that each of the three remaining surfaces contributes the same amount to the integral.

3.8 Integral Theorems

The formulas in this section relate a volume integration to a surface integral on its boundary (Gauss' theorem), or relate a surface integral to the line defining its perimeter (Stokes' theorem). These formulas are important tools in vector analysis, particularly when the functions involved are known to vanish on the boundary surface or perimeter.

Gauss' 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 a vector **A** and its first derivatives are continuous over a **simply connected** region of $\mathbb{R}^3$ (regions that contain holes, like a donut, are not simply connected). Then Gauss' theorem states that

$$\oint_{\partial V} \mathbf{A} \cdot d\boldsymbol{\sigma} = \int_V \boldsymbol{\nabla} \cdot \mathbf{A}\, d\tau. \tag{3.81}$$

Here the notations V and ∂V respectively denote a volume of interest and the closed surface that bounds it. The circle on the surface integral is an additional indication that the surface is closed.

To prove the theorem, consider the volume V to be subdivided into an arbitrary large number of tiny (differential) parallelepipeds, and look at the behavior of $\boldsymbol{\nabla} \cdot \mathbf{A}$ for each. See Fig. 3.15. For any given parallelepiped, this quantity is a measure of the net outward flow (of whatever **A** describes) through its boundary. If that boundary is interior (i.e., is shared by another parallelepiped), outflow from one parallelepiped is inflow to its neighbor; in a summation of all the outflows, all the contributions of interior boundaries cancel. Thus, the sum of all the outflows in the volume will just be the sum of those through the exterior boundary. In the limit of infinite subdivision, these sums become integrals: The left-hand side of Eq. (3.81) becomes the total outflow to the exterior, while its right-hand side is the sum of the outflows of the differential elements (the parallelepipeds).

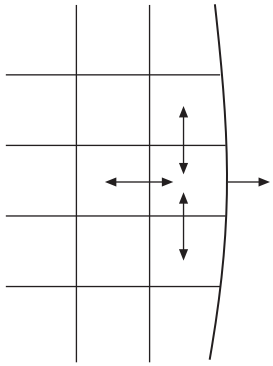

FIGURE 3.15 Subdivision for Gauss' theorem.

A simple alternate explanation of Gauss' theorem is that the volume integral sums the outflows $\nabla \cdot \mathbf{A}$ from all elements of the volume; the surface integral computes the same thing, by directly summing the flow through all elements of the boundary.

If the region of interest is the complete $\mathbb{R}^3$, and the volume integral converges, the surface integral in Eq. (3.81) must vanish, giving the useful result

$$\int \nabla \cdot \mathbf{A}\, d\tau = 0, \quad \text{integration over } \mathbb{R}^3 \text{ and convergent.} \tag{3.82}$$

Example 3.8.1 TETRAHEDRON

We check Gauss' theorem for a vector $\mathbf{B} = (x+1)\hat{\mathbf{e}}_x + y\hat{\mathbf{e}}_y - z\hat{\mathbf{e}}_z$, comparing

$$\int_V \nabla \cdot \mathbf{B}\, d\tau \quad \text{vs.} \quad \int_{\partial V} \mathbf{B} \cdot d\boldsymbol{\sigma},$$

where V is the tetrahedron of Example 3.7.2. In that example we computed the surface integral needed here, obtaining the value 1/6. For the integral over V, we take the divergence, obtaining $\nabla \cdot \mathbf{B} = 1$. The volume integral therefore reduces to the volume of the tetrahedron that, with base of area 1/2 and height 1, has volume $1/3 \times 1/2 \times 1 = 1/6$. This instance of Gauss' theorem is confirmed. ■

Green's Theorem

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

$$\nabla \cdot (u\nabla v) = u\nabla^2 v + (\nabla u)\cdot(\nabla v), \tag{3.83}$$

$$\nabla \cdot (u\nabla v) = u\nabla^2 v + (\nabla u)\cdot(\nabla v). \tag{3.84}$$

Subtracting Eq. (3.84) from Eq. (3.83), integrating over a volume V on which u, v, and their derivatives are continuous, and applying Gauss' theorem, Eq. (3.81), we obtain

$$\int_V (u\nabla^2 v - v\nabla^2 u)d\tau = \oint_{\partial V} (u\nabla v - v\nabla u) \cdot d\boldsymbol{\sigma}. \tag{3.85}$$

This is Green's theorem. An alternate form of Green's theorem, obtained from Eq. (3.83) alone, is

$$\oint_{\partial V} u\nabla v \cdot d\boldsymbol{\sigma} = \int_V u\nabla^2 v\, d\tau + \int_V \nabla_u \cdot \nabla v\, d\tau. \tag{3.86}$$

While the results already obtained are by far the most important forms of Gauss' theorem, volume integrals involving the gradient or the curl may also appear. To derive these, we consider a vector of the form

$$\mathbf{B}(x, y, z) = B(x, y, z)\mathbf{a}, \tag{3.87}$$

in which $\mathbf{a}$ is a vector with constant magnitude and constant but arbitrary direction. Then Eq. (3.81) becomes, applying Eq. (3.72),

$$\mathbf{a} \cdot \oint_{\partial V} B\, d\boldsymbol{\sigma} = \int_V \nabla \cdot (B\mathbf{a})d\tau = \mathbf{a} \int_V \nabla B\, d\tau.$$

This may be rewritten

$$\mathbf{a} \cdot \left[\oint_{\partial V} B\, d\boldsymbol{\sigma} - \int_V \nabla B\, d\tau \right] = 0. \tag{3.88}$$

Since the direction of $\mathbf{a}$ is arbitrary, Eq. (3.88) cannot always be satisfied unless the quantity in the square brackets evaluates to zero.[6] The result is

$$\oint_{\partial V} B\, d\boldsymbol{\sigma} = \int_V \nabla B\, d\tau. \tag{3.89}$$

In a similar manner, using $\mathbf{B} = \mathbf{a} \times \mathbf{P}$ in which $\mathbf{a}$ is a constant vector, we may show

$$\oint_{\partial V} d\boldsymbol{\sigma} \times \mathbf{P} = \int_V \nabla \times \mathbf{P}\, d\tau. \tag{3.90}$$

These last two forms of Gauss' theorem are used in the vector form of Kirchoff diffraction theory.

[6]This exploitation of the **arbitrary** nature of a part of a problem is a valuable and widely used technique.

Stokes' Theorem

Stokes' theorem is the analog of Gauss' theorem that relates a surface integral of a derivative of a function to the line integral of the function, with 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 Eq. (3.60) we saw that the circulation of a vector **B** about such a differential rectangles (in the xy-plane) is $\nabla \times \mathbf{B}\big|_z \hat{\mathbf{e}}_z \, dx \, dy$. Identifying $dx \, dy \, \hat{\mathbf{e}}_z$ as the element of area $d\boldsymbol{\sigma}$, Eq. (3.60) generalizes to

$$\sum_{\text{four sides}} \mathbf{B} \cdot d\mathbf{r} = \nabla \times \mathbf{B} \cdot d\boldsymbol{\sigma}. \tag{3.91}$$

We now sum over all the little rectangles; the surface contributions, from the right-hand side of Eq. (3.91), are added together. The line integrals (left-hand side) of all **interior** line segments cancel identically. See Fig. 3.16. Only the line integral around the perimeter survives. Taking the limit as the number of rectangles approaches infinity, we have

$$\oint_{\partial S} \mathbf{B} \cdot d\mathbf{r} = \int_S \nabla \times \mathbf{B} \cdot d\boldsymbol{\sigma}. \tag{3.92}$$

Here ∂S is the perimeter of S. This is Stokes' theorem. Note that both the sign of the line integral and the direction of $d\boldsymbol{\sigma}$ depend on the direction the perimeter is traversed, so consistent results will always be obtained. For the area and the line-integral direction shown in Fig. 3.16, the direction of $\boldsymbol{\sigma}$ for the shaded rectangle will be **out** of the plane of the paper.

Finally, consider what happens if we apply Stokes' theorem to a closed surface. Since it has no perimeter, the line integral vanishes, so

$$\int_S \nabla \times B \cdot d\boldsymbol{\sigma} = 0, \quad \text{for } S \text{ a closed surface.} \tag{3.93}$$

As with Gauss' theorem, we can derive additional relations connecting surface integrals with line integrals on their perimeter. Using the arbitrary-vector technique employed to

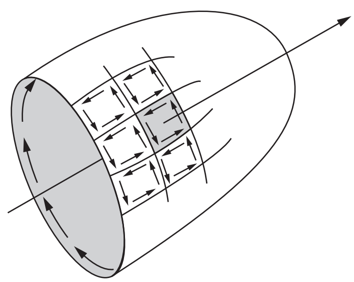

FIGURE 3.16 Direction of normal for the shaded rectangle when perimeter of the surface is traversed as indicated.

reach Eqs. (3.89) and (3.90), we can obtain

$$\int_S d\boldsymbol{\sigma} \times \nabla\varphi = \oint_{\partial S} \varphi d\mathbf{r}, \tag{3.94}$$

$$\int_S (d\boldsymbol{\sigma} \times \nabla) \times \mathbf{P} = \oint_{\partial S} d\mathbf{r} \times \mathbf{P}. \tag{3.95}$$

Example 3.8.2 OERSTED'S AND FARADAY'S LAWS

Consider the magnetic field generated by a long wire that carries a time-independent current I (meaning that $\partial\mathbf{E}/\partial t = \partial\mathbf{B}/\partial t = 0$). The relevant Maxwell equation, Eq. (3.68), then takes the form $\nabla \times \mathbf{B} = \mu_0\mathbf{J}$. Integrating this equation over a disk S perpendicular to and surrounding the wire (see Fig. 3.17), we have

$$I = \int_S \mathbf{J} \cdot d\boldsymbol{\sigma} = \frac{1}{\mu_0} \int_S (\nabla \times \mathbf{B}) \cdot d\boldsymbol{\sigma}.$$

Now we apply Stokes' theorem, obtaining the result $I = (1/\mu_0)\oint_{\partial S} \mathbf{B} \cdot d\mathbf{r}$, which is Oersted's law.

Similarly, we can integrate Maxwell's equation for $\nabla \times \mathbf{E}$, Eq. (3.69). Imagine moving a closed loop (∂S) of wire (of area S) across a magnetic induction field B. We have

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

where Φ is the magnetic flux through the area S. By Stokes' theorem, we have

$$\int_{\partial S} \mathbf{E} \cdot d\mathbf{r} = -\frac{d\Phi}{dt}.$$

This is Faraday's law. The line integral represents the voltage induced in the wire loop; it is equal in magnitude to the rate of change of the magnetic flux through the loop. There is no sign ambiguity; if the direction of ∂S is reversed, that causes a reversal of the direction of $d\boldsymbol{\sigma}$ and thereby of Φ. ■

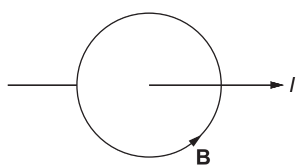

FIGURE 3.17 Direction of **B** given by Oersted's law.

Exercises

3.8.1 Using Gauss' theorem, prove that

$$\oint_S d\boldsymbol{\sigma} = 0$$

if $S = \partial V$ is a closed surface.

3.8.2 Show that

$$\frac{1}{3}\oint_S \mathbf{r}\cdot d\boldsymbol{\sigma} = V,$$

where V is the volume enclosed by the closed surface $S = \partial V$.
Note. This is a generalization of Exercise 3.7.5.

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

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

for any closed surface S.

3.8.4 From Eq. (3.72), with $\mathbf{V}$ the electric field $\mathbf{E}$ and f the electrostatic potential φ, show that, for integration over all space,

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

This corresponds to a 3-D 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}.

3.8.5 A particular steady-state electric current distribution is localized in space. Choosing a bounding surface far enough out so that the current density $\mathbf{J}$ is zero everywhere on the surface, show that

$$\int \mathbf{J} d\tau = 0.$$

Hint. Take one component of $\mathbf{J}$ at a time. With $\nabla\cdot\mathbf{J} = 0$, show that $\mathbf{J}_i = \nabla\cdot(x_i\mathbf{J})$ and apply Gauss' theorem.

3.8.6 Given a vector $\mathbf{t} = -\hat{\mathbf{e}}_\mathbf{x} y + \hat{\mathbf{e}}_\mathbf{y} x$, show, with the help of Stokes' theorem, that the integral of $\mathbf{t}$ around a continuous closed curve in the xy-plane satisfies

$$\frac{1}{2}\oint \mathbf{t}\cdot d\boldsymbol{\lambda} = \frac{1}{2}\oint (x\,dy - y\,dx) = A,$$

where A is the area enclosed by the curve.

3.8.7 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 $\mathbf{r} = \hat{\mathbf{e}}_\mathbf{x} a \cos\theta + \hat{\mathbf{e}}_\mathbf{y} b \sin\theta$. From part (a) show that the area of the ellipse is πab.

3.8.8 Evaluate $\oint \mathbf{r} \times d\mathbf{r}$ by using the alternate form of Stokes' theorem given by Eq. (3.95):

$$\int_S (d\boldsymbol{\sigma} \times \nabla) \times \mathbf{P} = \oint d\boldsymbol{\lambda} \times \mathbf{P}.$$

Take the loop to be entirely in the xy-plane.

3.8.9 Prove that

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

3.8.10 Prove that

$$\oint u\nabla v \cdot d\boldsymbol{\lambda} = \int_S (\nabla u) \times (\nabla v) \cdot d\boldsymbol{\sigma}.$$

3.8.11 Prove that

$$\oint_{\partial V} d\boldsymbol{\sigma} \times \mathbf{P} = \int_V \nabla \times \mathbf{P} d\tau.$$

3.8.12 Prove that

$$\int_S d\boldsymbol{\sigma} \times \nabla\varphi = \oint_{\partial S} \varphi d\mathbf{r}.$$

3.8.13 Prove that

$$\int_S (d\boldsymbol{\sigma} \times \nabla) \times \mathbf{P} = \oint_{\partial S} d\mathbf{r} \times \mathbf{P}.$$

3.9 Potential Theory

Much of physics, particularly electromagnetic theory, can be treated more simply by introducing **potentials** from which forces can be derived. This section deals with the definition and use of such potentials.

Scalar Potential

If, over a given simply connected region of space (one with no holes), a force can be expressed as the negative gradient of a scalar function φ,

$$\mathbf{F} = -\nabla\varphi, \tag{3.96}$$

we call φ a **scalar potential**, and we benefit from the feature that the force can be described in terms of one function instead of three. Since the force is a derivative of the scalar potential, the 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 reference point. We want to know what conditions $\mathbf{F}$ must satisfy in order for a scalar potential to exist.

First, consider the result of computing the work done against a force given by $-\nabla\varphi$ when an object subject to the force is moved from a point A to a point B. This is a line integral of the form

$$-\int_A^B \mathbf{F}\cdot d\mathbf{r} = \int_A^B \nabla\varphi\cdot d\mathbf{r}. \tag{3.97}$$

But, as pointed out in Eq. (3.41), $\nabla\varphi\cdot d\mathbf{r} = d\varphi$, so the integral is in fact independent of the path, depending only on the endpoints A and B. So we have

$$-\int_A^B \mathbf{F}\cdot d\mathbf{r} = \varphi(\mathbf{r}_B) - \varphi(\mathbf{r}_A), \tag{3.98}$$

which also means that if A and B are the same point, forming a closed loop,

$$\oint \mathbf{F}\cdot d\mathbf{r} = 0. \tag{3.99}$$

We conclude that a force (on an object) described by a scalar potential is a **conservative force**, meaning that the work needed to move the object between any two points is independent of the path taken, and that $\varphi(\mathbf{r})$ is the work needed to move to the point $\mathbf{r}$ from a reference point where the potential has been assigned the value zero.

Another property of a force given by a scalar potential is that

$$\nabla\times\mathbf{F} = -\nabla\times\nabla\varphi = 0 \tag{3.100}$$

as prescribed by Eq. (3.64). This observation is consistent with the notion that the lines of force of a conservative $\mathbf{F}$ cannot form closed loops.

The three conditions, Eqs. (3.96), (3.99), and (3.100), are all equivalent. If we take Eq. (3.99) for a differential loop, its left side and that of Eq. (3.100) must, according to Stokes' theorem, be equal. We already showed both these equations followed from Eq. (3.96). To complete the establishment of full equivalence, we need only to derive Eq. (3.96) from Eq. (3.99). Going backward to Eq. (3.97), we rewrite it as

$$\int_A^B (\mathbf{F} + \nabla\varphi)\cdot d\mathbf{r} = 0,$$

which must be satisfied for all A and B. This means its integrand must be identically zero, thereby recovering Eq. (3.96).

Example 3.9.1 GRAVITATIONAL POTENTIAL

We have previously, in Example 3.5.2, illustrated the generation of a force from a scalar potential. To perform the reverse process, we must integrate. Let us find the scalar potential for the gravitational force

$$\mathbf{F}_G = -\frac{Gm_1m_2\hat{\mathbf{r}}}{r^2} = -\frac{k\hat{\mathbf{r}}}{r^2},$$

radially **inward**. Setting the zero of scalar potential at infinity, we obtain by integrating (radially) from infinity to position $\mathbf{r}$,

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

The minus sign in the central member of this equation arises because we are calculating the work done **against** the gravitational force. Evaluating the integral,

$$\varphi_G(r) = -\int_{r}^{\infty} \frac{k\,dr}{r^2} = -\frac{k}{r} = -\frac{Gm_1m_2}{r}.$$

The final negative sign corresponds to the fact that gravity is an attractive force. ■

Vector Potential

In some branches of physics, especially electrodynamics, it is convenient to introduce a **vector potential A** such that a (force) field **B** is given by

$$\mathbf{B} = \nabla \times \mathbf{A}. \tag{3.101}$$

An obvious reason for introducing **A** is that it causes **B** to be solenoidal; if **B** is the magnetic induction field, this property is required by Maxwell's equations. Here we want to develop a converse, namely to show that when **B** is solenoidal, a vector potential **A** exists. We demonstrate the existence of **A** by actually writing it.

Our construction is

$$\mathbf{A} = \hat{\mathbf{e}}_y \int_{x_0}^{x} B_z(x, y, z)\,dx + \hat{\mathbf{e}}_z \left[\int_{y_0}^{y} B_x(x_0, y, z)\,dy - \int_{x_0}^{x} B_y(x, y, z)\,dx \right]. \tag{3.102}$$

Checking the y- and z-components of $\nabla \times \mathbf{A}$ first, noting that $A_x = 0$,

$$(\nabla \times \mathbf{A})_y = -\frac{\partial A_z}{\partial x} = +\frac{\partial}{\partial x}\int_{x_0}^{x} B_y(x, y, z)\,dx = B_y,$$

$$(\nabla \times \mathbf{A})_z = +\frac{\partial A_y}{\partial x} == \frac{\partial}{\partial x}\int_{x_0}^{x} B_z(x, y, z)\,dx = B_z.$$

The x-component of $\nabla \times \mathbf{A}$ is a bit more complicated. We have

$$\begin{aligned}(\nabla \times \mathbf{A})_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\\ &= \frac{\partial}{\partial y}\left[\int_{y_0}^{y} B_x(x_0, y, z)\,dy - \int_{x_0}^{x} B_y(x, y, z)\,dx\right] - \frac{\partial}{\partial z}\int_{x_0}^{x} B_z(x, y, z)\,dx\\ &= B_x(x_0, y, z) - \int_{x_0}^{x}\left[\frac{\partial B_y(x, y, z)}{\partial y} + \frac{\partial B_z(x, y, z)}{\partial z}\right]dx.\end{aligned}$$

To go further, we must use the fact that $\mathbf{B}$ is solenoidal, which means $\nabla \cdot \mathbf{B} = 0$. We can therefore make the replacement

$$\frac{\partial B_y(x, y, z)}{\partial y} + \frac{\partial B_z(x, y, z)}{\partial z} = -\frac{\partial B_x(x, y, z)}{\partial x},$$

after which the x integration becomes trivial, yielding

$$+\int_{x_0}^{x} \frac{\partial B_x(x, y, z)}{\partial x}\,dx = B_x(x, y, z) - B_x(x_0, y, z),$$

leading to the desired final result $(\nabla \times \mathbf{A})_x = B_x$.

While we have shown that there exists a vector potential $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \mathbf{B}$ subject only to the condition that $\mathbf{B}$ be solenoidal, we have in no way established that $\mathbf{A}$ is unique. In fact, $\mathbf{A}$ is far from unique, as we can add to it not only an arbitrary constant, but also the gradient of **any** scalar function, $\nabla\varphi$, without affecting $\mathbf{B}$ at all. Moreover, our verification of $\mathbf{A}$ was independent of the values of x_0 and y_0, so these can be assigned arbitrarily without affecting $\mathbf{B}$. In addition, we can derive another formula for $\mathbf{A}$ in which the roles of x and y are interchanged:

$$\mathbf{A} = -\hat{\mathbf{e}}_x \int_{y_0}^{y} B_z(x, y, z)\,dy - \hat{\mathbf{e}}_z\left[\int_{x_0}^{x} B_y(x, y_0, z)\,dx - \int_{y_0}^{y} B_x(x, y, z)\,dy\right]. \qquad (3.103)$$

Example 3.9.2 MAGNETIC VECTOR POTENTIAL

We consider the construction of the vector potential for a constant magnetic induction field

$$\mathbf{B} = B_z \hat{\mathbf{e}}_z. \tag{3.104}$$

Using Eq. (3.102), we have (choosing the arbitrary value of x_0 to be zero)

$$\mathbf{A} = \hat{\mathbf{e}}_y \int_0^x B_z \, dx = \hat{\mathbf{e}}_y x B_z. \tag{3.105}$$

Alternatively, we could use Eq. (3.103) for $\mathbf{A}$, leading to

$$\mathbf{A}' = -\hat{\mathbf{e}}_x y B_z. \tag{3.106}$$

Neither of these is the form for $\mathbf{A}$ found in many elementary texts, which for $\mathbf{B}$ from Eq. (3.104) is

$$\mathbf{A}'' = \frac{1}{2}(\mathbf{B} \times \mathbf{r}) = \frac{B_z}{2}(x\hat{\mathbf{e}}_y - y\hat{\mathbf{e}}_x). \tag{3.107}$$

These disparate forms can be reconciled if we use the freedom to add to $\mathbf{A}$ any expression of the form $\nabla\varphi$. Taking $\varphi = Cxy$, the quantity that can be added to $\mathbf{A}$ will be of the form

$$\nabla\varphi = C(y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y).$$

We now see that

$$\mathbf{A} - \frac{B_z}{2}(y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y) = \mathbf{A}' + \frac{B_z}{2}(y\hat{\mathbf{e}}_x + x\hat{\mathbf{e}}_y) = \mathbf{A}'',$$

showing that all these formulas predict the same value of $\mathbf{B}$. ■

Example 3.9.3 POTENTIALS IN ELECTROMAGNETISM

If we introduce suitably defined scalar and vector potentials φ and $\mathbf{A}$ into Maxwell's equations, we can obtain equations giving these potentials in terms of the sources of the electromagnetic field (charges and currents). We start with $\mathbf{B} = \nabla \times \mathbf{A}$, thereby assuring satisfaction of the Maxwell's equation $\nabla \cdot \mathbf{B} = 0$. Substitution into the equation for $\nabla \times \mathbf{E}$ yields

$$\nabla \times \mathbf{E} = -\nabla \times \frac{\partial \mathbf{A}}{\partial t} \quad \longrightarrow \quad \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0,$$

showing that $\mathbf{E} + \partial\mathbf{A}/\partial t$ is a gradient and can be written as $-\nabla\varphi$, thereby defining φ. This preserves the notion of an electrostatic potential in the absence of time dependence, and means that $\mathbf{A}$ and φ have now been defined to give

$$\mathbf{B} = \nabla \times \mathbf{A}, \quad \mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}. \tag{3.108}$$

At this point $\mathbf{A}$ is still arbitrary to the extent of adding any gradient, which is equivalent to making an arbitrary choice of $\nabla \cdot \mathbf{A}$. A convenient choice is to require

$$\frac{1}{c^2}\frac{\partial\varphi}{\partial t} + \nabla \cdot \mathbf{A} = 0. \tag{3.109}$$

This **gauge condition** is called the **Lorentz gauge**, and transformations of $\mathbf{A}$ and φ to satisfy it or any other legitimate gauge condition are called **gauge transformations**. The invariance of electromagnetic theory under gauge transformation is an important precursor of contemporary directions in fundamental physical theory.

From Maxwell's equation for $\nabla \cdot \mathbf{E}$ and the Lorentz gauge condition, we get

$$\frac{\rho}{\varepsilon_0} = \nabla \cdot \mathbf{E} = -\nabla^2 \mathbf{E} - \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = -\nabla^2 \varphi + \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2}, \tag{3.110}$$

showing that the Lorentz gauge permitted us to decouple $\mathbf{A}$ and φ to the extent that we have an equation for φ in terms only of the charge density ρ; neither $\mathbf{A}$ nor the current density $\mathbf{J}$ enters this equation.

Finally, from the equation for $\nabla \times \mathbf{B}$, we obtain

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J}. \tag{3.111}$$

Proof of this formula is the subject of Exercise 3.9.11. ■

Gauss' Law

Consider a point charge q at the origin of our coordinate system. It produces an electric field $\mathbf{E}$, given by

$$\mathbf{E} = \frac{q\hat{\mathbf{r}}}{4\pi\varepsilon_0 r^2}. \tag{3.112}$$

Gauss' law states that for an arbitrary volume V,

$$\oint_{\partial V} \mathbf{E} \cdot d\boldsymbol{\sigma} = \begin{cases} \dfrac{q}{\varepsilon_0} & \text{if } \partial V \text{ encloses } q, \\ 0 & \text{if } \partial V \text{ does not enclose } q. \end{cases} \tag{3.113}$$

The case that ∂V does not enclose q is easily handled. From Eq. (3.54), the r^{-2} central force $\mathbf{E}$ is divergenceless everywhere except at $r = 0$, and for this case, throughout the entire volume V. Thus, we have, invoking Gauss' theorem, Eq. (3.81),

$$\int_V \nabla \cdot \mathbf{E} = 0 \quad \longrightarrow \quad \mathbf{E} \cdot d\boldsymbol{\sigma} = 0.$$

If q is within the volume V, we must be more devious. We surround $\mathbf{r} = 0$ by a small spherical hole (of radius δ), with a surface we designate S', and connect the hole with the boundary of V via a small tube, thereby creating a simply connected region V' to which Gauss' theorem will apply. See Fig. 3.18. We now consider $\oint \mathbf{E} \cdot d\boldsymbol{\sigma}$ on the surface of this modified volume. The contribution from the connecting tube will become negligible in the limit that it shrinks toward zero cross section, as $\mathbf{E}$ is finite everywhere on the tube's surface. The integral over the modified ∂V will thus be that of the original ∂V (over the outer boundary, which we designate S), plus that of the inner spherical surface (S').

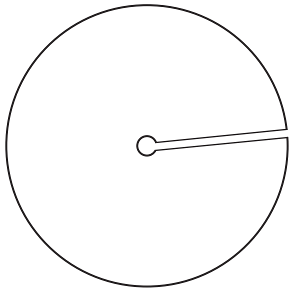

FIGURE 3.18 Making a multiply connected region simply connected.

But note that the "outward" direction for S' is toward smaller r, so $d\boldsymbol{\sigma}' = -\hat{\mathbf{r}}\,dA$. Because the modified volume contains no charge, we have

$$\oint_{\partial V'} \mathbf{E}\cdot d\boldsymbol{\sigma} = \oint_{S} \mathbf{E}\cdot d\boldsymbol{\sigma} + \frac{q}{4\pi\varepsilon_0}\oint_{S'} \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}'}{\delta^2} = 0, \tag{3.114}$$

where we have inserted the explicit form of $\mathbf{E}$ in the S' integral. Because S' is a sphere of radius δ, this integral can be evaluated. Writing $d\Omega$ as the element of solid angle, so $dA = \delta^2 d\Omega$,

$$\oint_{S'} \frac{\hat{\mathbf{r}}\cdot d\boldsymbol{\sigma}'}{\delta^2} = \int \frac{\hat{\mathbf{r}}}{\delta^2}\cdot(-\hat{\mathbf{r}}\,\delta^2\,d\Omega) = -\int d\Omega = -4\pi,$$

independent of the value of δ. Returning now to Eq. (3.114), it can be rearranged into

$$\oint_{S} \mathbf{E}\cdot d\boldsymbol{\sigma} = -\frac{q}{4\pi\varepsilon_0}(-4\pi) = +\frac{q}{\varepsilon_0},$$

the result needed to confirm the second case of Gauss' law, Eq. (3.113).

Because the equations of electrostatics are linear, Gauss' law can be extended to collections of charges, or even to continuous charge distributions. In that case, q can be replaced by $\int_V \rho\,d\tau$, and Gauss' law becomes

$$\int_{\partial V} \mathbf{E}\cdot d\boldsymbol{\sigma} = \int_V \frac{\rho}{\varepsilon_0}\,d\tau. \tag{3.115}$$

If we apply Gauss' theorem to the left side of Eq. (3.115), we have

$$\int_V \nabla\cdot\mathbf{E}\,d\tau = \int_V \frac{\rho}{\varepsilon_0}\,d\tau.$$

Since our volume is completely arbitrary, the integrands of this equation must be equal, so

$$\nabla\cdot\mathbf{E} = \frac{\rho}{\varepsilon_0}. \tag{3.116}$$

We thus see that Gauss' law is the integral form of one of Maxwell's equations.

Poisson's Equation

If we return to Eq. (3.116) and, assuming a situation independent of time, write $\mathbf{E} = -\nabla\varphi$, we obtain

$$\nabla^2\varphi = -\frac{\rho}{\varepsilon_0}. \tag{3.117}$$

This equation, applicable to electrostatics,[7] is called Poisson's equation. If, in addition, $\rho = 0$, we have an even more famous equation,

$$\nabla^2\varphi = 0, \tag{3.118}$$

Laplace's equation.

To make Poisson's equation apply to a point charge q, we need to replace ρ by a concentration of charge that is localized at a point and adds up to q. The Dirac delta function is what we need for this purpose. Thus, for a point charge q at the origin, we write

$$\nabla^2\varphi = -\frac{q}{\varepsilon_0}\,\delta(\mathbf{r}), \qquad (\text{charge } q \text{ at } \mathbf{r} = 0). \tag{3.119}$$

If we rewrite this equation, inserting the point-charge potential for φ, we have

$$\frac{q}{4\pi\varepsilon_0}\nabla^2\left(\frac{1}{r}\right) = -\frac{q}{\varepsilon_0}\,\delta(\mathbf{r}),$$

which reduces to

$$\nabla^2\left(\frac{1}{r}\right) = -4\pi\,\delta(\mathbf{r}). \tag{3.120}$$

This equation circumvents the problem that the derivatives of $1/r$ do not exist at $\mathbf{r} = 0$, and gives appropriate and correct results for systems containing point charges. Like the definition of the delta function itself, Eq. (3.120) is only meaningful when inserted into an integral. It is an important result that is used repeatedly in physics, often in the form

$$\nabla_1^2\left(\frac{1}{r_{12}}\right) = -4\pi\,\delta(\mathbf{r}_1 - \mathbf{r}_2). \tag{3.121}$$

Here $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$, and the subscript in ∇_1 indicates that the derivatives apply to $\mathbf{r}_1$.

Helmholtz's Theorem

We now turn to two theorems that are of great formal importance, in that they establish conditions for the existence and uniqueness of solutions to time-independent problems in electromagnetic theory. The first of these theorems is:

> *A vector field is uniquely specified by giving its divergence and its curl within a simply connected region and its normal component on the boundary.*

[7] For general time dependence, see Eq. (3.110).

Note that both for this theorem and the next (Helmholtz's theorem), even if there are points in the simply connected region where the divergence or the curl is only defined in terms of delta functions, these points are not to be removed from the region.

Let **P** be a vector field satisfying the conditions

$$\nabla \cdot \mathbf{P} = s, \qquad \nabla \times \mathbf{P} = \mathbf{c}, \tag{3.122}$$

where s may be interpreted as a given source (charge) density and **c** as a given circulation (current) density. Assuming that the normal component P_n on the boundary is also given, we want to show that **P** is unique.

We proceed by assuming the existence of a second vector, $\mathbf{P}'$, which satisfies Eq. (3.122) and has the same value of P_n. We form $\mathbf{Q} = \mathbf{P} - \mathbf{P}'$, which must have $\nabla \cdot \mathbf{Q}$, $\nabla \times \mathbf{Q}$, and Q_n all identically zero. Because **Q** is irrotational, there must exist a potential φ such that $\mathbf{Q} = -\nabla\varphi$, and because $\nabla \cdot \mathbf{Q} = 0$, we also have

$$\nabla^2\varphi = 0.$$

Now we draw on Green's theorem in the form given in Eq. (3.86), letting u and v each equal φ. Because $Q_n = 0$ on the boundary, Green's theorem reduces to

$$\int_V (\nabla\varphi) \cdot (\nabla\varphi)\, d\tau = \int_V \mathbf{Q} \cdot \mathbf{Q}\, d\tau = 0.$$

This equation can only be satisfied if **Q** is identically zero, showing that $\mathbf{P}' = \mathbf{P}$, thereby proving the theorem.

The second theorem we shall prove, Helmholtz's theorem, is

> *A vector* **P** *with both source and circulation densities vanishing at infinity may be written as the sum of two parts, one of which is irrotational, the other of which is solenoidal.*

Helmholtz's theorem will clearly be satisfied if **P** can be written in the form

$$\mathbf{P} = -\nabla\varphi + \nabla \times \mathbf{A}, \tag{3.123}$$

since $-\nabla\varphi$ is irrotational, while $\nabla \times \mathbf{A}$ is solenoidal. Because **P** is known, so are also s and **c**, defined as

$$s = \nabla \cdot \mathbf{P}, \quad \mathbf{c} = \nabla \times \mathbf{P}.$$

We proceed by exhibiting expressions for φ and **A** that enable the recovery of s and **c**. Because the region here under study is simply connected and the vector involved vanishes at infinity (so that the first theorem of this subsection applies), having the correct s and **c** guarantees that we have properly reproduced **P**.

The formulas proposed for φ and **A** are the following, written in terms of the spatial variable $\mathbf{r}_1$:

$$\varphi(\mathbf{r}_1) = \frac{1}{4\pi} \int \frac{s(\mathbf{r}_2)}{r_{12}} d\tau_2, \tag{3.124}$$

$$\mathbf{A}(\mathbf{r}_1) = \frac{1}{4\pi} \int \frac{\mathbf{c}(\mathbf{r}_2)}{r_{12}} d\tau_2. \tag{3.125}$$

Here $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$.

If Eq. (3.123) is to be satisfied with the proposed values of φ and $\mathbf{A}$, it is necessary that

$$\nabla \cdot \mathbf{P} = -\nabla \cdot \nabla\varphi + \nabla \cdot (\nabla \times \mathbf{A}) = -\nabla^2\varphi = s,$$

$$\nabla \times \mathbf{P} = -\nabla \times \nabla\varphi + \nabla \times (\nabla \times \mathbf{A}) = \nabla \times (\nabla \times \mathbf{A}) = \mathbf{c}.$$

To check that $-\nabla^2\varphi = s$, we examine

$$-\nabla_1^2\varphi(\mathbf{r}_1) = -\frac{1}{4\pi}\int \nabla_1^2\left(\frac{1}{r_{12}}\right)s(\mathbf{r}_2)d\tau_2$$

$$= -\frac{1}{4\pi}\int \big[-4\pi\delta(\mathbf{r}_1 - \mathbf{r}_2)\big]s(\mathbf{r}_2)d\tau_2 = s(\mathbf{r}_1). \tag{3.126}$$

We have written ∇_1 to make clear that it operates on $\mathbf{r}_1$ and not $\mathbf{r}_2$, and we have used the delta-function property given in Eq. (3.121). So s has been recovered.

We now check that $\nabla \times (\nabla \times \mathbf{A}) = \mathbf{c}$. We start by using Eq. (3.70) to convert this condition to a more easily utilized form:

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

Taking $\mathbf{r}_1$ as the free variable, we look first at

$$\nabla_1\big(\nabla_1 \cdot \mathbf{A}(\mathbf{r}_1)\big) = \frac{1}{4\pi}\nabla_1\int \nabla_1 \cdot \left(\frac{\mathbf{c}(\mathbf{r}_2)}{r_{12}}\right)d\tau_2$$

$$= \frac{1}{4\pi}\nabla_1\int \mathbf{c}(\mathbf{r}_2)\cdot\nabla_1\left(\frac{1}{r_{12}}\right)d\tau_2$$

$$= \frac{1}{4\pi}\nabla_1\int \mathbf{c}(\mathbf{r}_2)\cdot\left[-\nabla_2\left(\frac{1}{r_{12}}\right)\right]d\tau_2.$$

To reach the second line of this equation, we used Eq. (3.72) for the special case that the vector in that equation is not a function of the variable being differentiated. Then, to obtain the third line, we note that because the ∇_1 within the integral acts on a function of $\mathbf{r}_1 - \mathbf{r}_2$, we can change ∇_1 into ∇_2 and introduce a sign change.

Now we integrate by parts, as in Example 3.7.3, reaching

$$\nabla_1\big[\nabla_1 \cdot \mathbf{A}(\mathbf{r}_1)\big] = \frac{1}{4\pi}\nabla_1\int \big(\nabla_2 \cdot \mathbf{c}(\mathbf{r}_2)\big)\left(\frac{1}{r_{12}}\right)d\tau_2.$$

At last we have the result we need: $\nabla_2 \cdot \mathbf{c}(\mathbf{r}_2)$ vanishes, because $\mathbf{c}$ is a curl, so the entire $\nabla(\nabla \cdot \mathbf{A})$ term is zero and may be dropped. This reduces the condition we are checking to $-\nabla^2\mathbf{A} = \mathbf{c}$.

The quantity $-\nabla^2\mathbf{A}$ is a vector Laplacian and we may individually evaluate its Cartesian components. For component j,

$$-\nabla_1^2 A_j(\mathbf{r}_1) = -\frac{1}{4\pi}\int c_j(\mathbf{r}_2)\nabla_1^2\left(\frac{1}{r_{12}}\right)d\tau_2$$

$$= -\frac{1}{4\pi}\int c_j(\mathbf{r}_2)\big[-4\pi\delta(\mathbf{r}_1 - \mathbf{r}_2)\big]d\tau_2 = c_j(\mathbf{r}_1).$$

This completes the proof of Helmholtz's theorem.

Helmholtz's theorem legitimizes the division of the quantities appearing in electromagnetic theory into an irrotational vector field $\mathbf{E}$ and a solenoidal vector field $\mathbf{B}$, together with their respective representations using scalar and vector potentials. As we have seen in numerous examples, the source s is identified as the charge density (divided by ε_0) and the circulation $\mathbf{c}$ is the current density (multiplied by μ_0).

Exercises

3.9.1 If a force $\mathbf{F}$ is given by

$$\mathbf{F} = (x^2 + y^2 + z^2)^n (\hat{\mathbf{e}}_\mathbf{x} x + \hat{\mathbf{e}}_\mathbf{y} y + \hat{\mathbf{e}}_\mathbf{z} z),$$

find

(a) $\nabla \cdot \mathbf{F}$.
(b) $\nabla \times \mathbf{F}$.
(c) A scalar potential $\varphi(x, y, z)$ so that $\mathbf{F} = -\nabla\varphi$.
(d) For what value of the exponent n does the scalar potential diverge at both the origin and infinity?

ANS. (a) $(2n+3)r^{2n}$ (b) 0
(c) $-r^{2n+2}/(2n+2),\ n \neq -1$ (d) $n = -1,\ \varphi = -\ln r$.

3.9.2 A sphere of radius a is uniformly charged (throughout its volume). Construct the electrostatic potential $\varphi(r)$ for $0 \le r < \infty$.

3.9.3 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. $-\dfrac{GMm}{R^3}\left(z^2 - \dfrac{1}{2}x^2 - \dfrac{1}{2}y^2\right).$

3.9.4 A long, straight wire carrying a current I produces a magnetic induction $\mathbf{B}$ with components

$$\mathbf{B} = \frac{\mu_0 I}{2\pi}\left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right).$$

Find a magnetic vector potential $\mathbf{A}$.

ANS. $\mathbf{A} = -\hat{\mathbf{z}}(\mu_0 I/4\pi)\ln(x^2+y^2)$. (This solution is not unique.)

3.9.5 If

$$\mathbf{B} = \frac{\hat{\mathbf{r}}}{r^2} = \left(\frac{x}{r^3}, \frac{y}{r^3}, \frac{z}{r^3}\right),$$

find a vector $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \mathbf{B}$.

ANS. One possible solution is $\mathbf{A} = \dfrac{\hat{\mathbf{e}}_x yz}{r(x^2+y^2)} - \dfrac{\hat{\mathbf{e}}_y xz}{r(x^2+y^2)}$.

3.9.6 Show that the pair of equations

$$\mathbf{A} = \frac{1}{2}(\mathbf{B} \times \mathbf{r}), \quad \mathbf{B} = \nabla \times \mathbf{A},$$

is satisfied by any constant magnetic induction **B**.

3.9.7 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 **B** is solenoidal.

(b) Show that

$$\mathbf{A} = \frac{1}{2}(u \nabla v - v \nabla u)$$

is a vector potential for **B**, in that

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

3.9.8 The magnetic induction **B** is related to the magnetic vector potential **A** by $\mathbf{B} = \nabla \times \mathbf{A}$. By Stokes' 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**, $\mathbf{A} \to \mathbf{A} + \nabla\varphi$.
Note. Take the function φ to be single-valued.

3.9.9 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, provided that there are no electric charges on or within the sphere.
Hint. Use Green's theorem, Eq. (3.85), with $u = r^{-1}$, the distance from P, and $v = \varphi$. Equation (3.120) will also be useful.

3.9.10 Using Maxwell's equations, show that for a system (steady current) the magnetic vector potential **A** satisfies a vector Poisson equation,

$$\nabla^2 \mathbf{A} = -\mu \mathbf{J},$$

provided we require $\nabla \cdot \mathbf{A} = 0$.

3.9.11 Derive, assuming the Lorentz gauge, Eq. (3.109):

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

Hint. Eq. (3.70) will be helpful.

3.9.12 Prove that an arbitrary solenoidal vector $\mathbf{B}$ can be described as $\mathbf{B} = \nabla \times \mathbf{A}$, with

$$\mathbf{A} = -\hat{\mathbf{e}}_x \int_{y_0}^{y} B_z(x, y, z)\, dy - \hat{\mathbf{e}}_z \left[\int_{x_0}^{x} B_y(x, y_0, z)\, dx - \int_{y_0}^{y} B_x(x, y, z)\, dy \right].$$

3.10 CURVILINEAR COORDINATES

Up to this point we have treated vectors essentially entirely in Cartesian coordinates; when $\mathbf{r}$ or a function of it was encountered, we wrote $\mathbf{r}$ as $\sqrt{x^2 + y^2 + z^2}$, so that Cartesian coordinates could continue to be used. Such an approach ignores the simplifications that can result if one uses a coordinate system that is appropriate to the symmetry of a problem. Central force problems are frequently easiest to deal with in spherical polar coordinates. Problems involving geometrical elements such as straight wires may be best handled in cylindrical coordinates. Yet other coordinate systems (of use too infrequent to be described here) may be appropriate for other problems.

Naturally, there is a price that must be paid for the use of a non-Cartesian coordinate system. Vector operators become different in form, and their specific forms may be position-dependent. We proceed here to examine these questions and derive the necessary formulas.

Orthogonal Coordinates in $\mathbb{R}^3$

In Cartesian coordinates the point (x_0, y_0, z_0) can be identified as the intersection of three planes: (1) the plane $x = x_0$ (a surface of constant x), (2) the plane $y = y_0$ (constant y), and (3) the plane $z = z_0$ (constant z). A change in x corresponds to a displacement **normal** to the surface of constant x; similar remarks apply to changes in y or z. The planes of constant coordinate value are mutually perpendicular, and have the obvious feature that the normal to any given one of them is in the same direction, no matter where on the plane it is constructed (a plane of constant x has a normal that is, of course, everywhere in the direction of $\hat{\mathbf{e}}_x$).

Consider now, as an example of a curvilinear coordinate system, spherical polar coordinates (see Fig. 3.19). A point $\mathbf{r}$ is identified by r (distance from the origin), θ (angle of $\mathbf{r}$ relative to the polar axis, which is conventionally in the z direction), and φ (dihedral angle between the zx plane and the plane containing $\hat{\mathbf{e}}_z$ and $\mathbf{r}$). The point $\mathbf{r}$ is therefore at the intersection of (1) a sphere of radius r, (2) a cone of opening angle θ, and (3) a half-plane through equatorial angle φ. This example provides several observations: (1) general

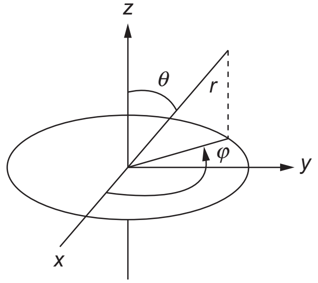

FIGURE 3.19 Spherical polar coordinates.

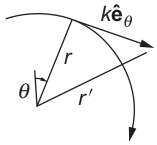

FIGURE 3.20 Effect of a "large" displacement in the direction $\hat{\mathbf{e}}_\theta$. Note that $r' \neq r$.

coordinates need not be lengths, (2) a surface of constant coordinate value may have a normal whose direction depends on position, (3) surfaces with different constant values of the same coordinate need not be parallel, and therefore also (4) changes in the value of a coordinate may move $\mathbf{r}$ in both an amount and a direction that depends on position.

It is convenient to define unit vectors $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$, $\hat{\mathbf{e}}_\varphi$ in the directions of the normals to the surfaces, respectively, of constant r, θ, and φ. The spherical polar coordinate system has the feature that these unit vectors are mutually perpendicular, meaning that, for example, $\hat{\mathbf{e}}_\theta$ will be tangent to both the constant-r and constant-φ surfaces, so that a small displacement in the $\hat{\mathbf{e}}_\theta$ direction will not change the values of either the r or the φ coordinate. The reason for the restriction to "small" displacements is that the directions of the normals are position-dependent; a "large" displacement in the $\hat{\mathbf{e}}_\theta$ direction would change r (see Fig. 3.20). If the coordinate unit vectors are mutually perpendicular, the coordinate system is said to be **orthogonal**.

If we have a vector field $\mathbf{V}$ (so we associate a value of $\mathbf{V}$ with each point in a region of $\mathbb{R}^3$), we can write $\mathbf{V}(\mathbf{r})$ in terms of the orthogonal set of unit vectors that are defined for the point $\mathbf{r}$; symbolically, the result is

$$\mathbf{V}(\mathbf{r}) = V_r\,\hat{\mathbf{e}}_r + V_\theta\,\hat{\mathbf{e}}_\theta + V_\varphi\,\hat{\mathbf{e}}_\varphi.$$

It is important to realize that the unit vectors $\hat{\mathbf{e}}_i$ have directions that depend on the value of $\mathbf{r}$. If we have another vector field $\mathbf{W}(\mathbf{r})$ for the **same** point $\mathbf{r}$, we can perform **algebraic** processes[8] on $\mathbf{V}$ and $\mathbf{W}$ by the same rules as for Cartesian coordinates. For example, **at the point r**,

$$\mathbf{V}\cdot\mathbf{W} = V_r W_r + V_\theta W_\theta + V_\varphi W_\varphi.$$

However, if $\mathbf{V}$ and $\mathbf{W}$ are not associated with the same $\mathbf{r}$, we cannot carry out such operations in this way, and it is important to realize that

$$\mathbf{r} \neq r\hat{\mathbf{e}}_r + \theta\hat{\mathbf{e}}_\theta + \varphi\hat{\mathbf{e}}_\varphi.$$

Summarizing, the component formulas for $\mathbf{V}$ or $\mathbf{W}$ describe component decompositions applicable to the point at which the vector is specified; an attempt to decompose $\mathbf{r}$ as illustrated above is incorrect because it uses fixed unit-vector orientations where they do not apply.

Dealing for the moment with an arbitrary curvilinear system, with coordinates labeled (q_1, q_2, q_3), we consider how changes in the q_i are related to changes in the Cartesian coordinates. Since x can be thought of as a function of the q_i, namely $x(q_1, q_2, q_3)$, we have

$$dx = \frac{\partial x}{\partial q_1}dq_1 + \frac{\partial x}{\partial q_2}dq_2 + \frac{\partial x}{\partial q_3}dq_3, \tag{3.127}$$

with similar formulas for dy and dz.

[8] Addition, multiplication by a scalar, dot and cross products (but not application of differential or integral operators).

We next form a measure of the differential displacement, $d\mathbf{r}$, associated with changes dq_i. We actually examine

$$(dr)^2 = (dx)^2 + (dy)^2 + (dz)^2.$$

Taking the square of Eq. (3.127), we get

$$(dx)^2 = \sum_{ij} \frac{\partial x}{\partial q_i} \frac{\partial x}{\partial q_j} dq_i \, dq_j$$

and similar expressions for $(dy)^2$ and $(dz)^2$. Combining these and collecting terms with the same $dq_i \, dq_j$, we reach the result

$$(dr)^2 = \sum_{ij} g_{ij} \, dq_i \, dq_j, \tag{3.128}$$

where

$$g_{ij}(q_1, q_2, q_3) = \frac{\partial x}{\partial q_i} \frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i} \frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i} \frac{\partial z}{\partial q_j}. \tag{3.129}$$

Spaces with a measure of distance given by Eq. (3.128) are called **metric** or **Riemannian**.

Equation (3.129) can be interpreted as the dot product of a vector in the dq_i direction, of components $(\partial x/\partial q_i, \ \partial y/\partial q_i, \ \partial z/\partial q_i)$, with a similar vector in the dq_j direction. If the q_i coordinates are perpendicular, the coefficients g_{ij} will vanish when $i \neq j$.

Since it is our objective to discuss orthogonal coordinate systems, we specialize Eqs. (3.128) and (3.129) to

$$(dr)^2 = (h_1 \, dq_1)^2 + (h_2 \, dq_2)^2 + (h_3 \, dq_3)^2, \tag{3.130}$$

$$h_i^2 = \left(\frac{\partial x}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2 + \left(\frac{\partial y}{\partial q_i}\right)^2. \tag{3.131}$$

If we consider Eq. (3.130) for a case $dq_2 = dq_3 = 0$, we see that we can identify $h_1 dq_1$ as dr_1, meaning that the element of displacement in the q_1 direction is $h_1 dq_1$. Thus, in general,

$$dr_i = h_i dq_i, \quad \text{or} \quad \frac{\partial \mathbf{r}}{\partial q_i} = h_i \, \hat{\mathbf{e}}_i. \tag{3.132}$$

Here $\hat{\mathbf{e}}_i$ is a unit vector in the q_i direction, and the overall $d\mathbf{r}$ takes the form

$$d\mathbf{r} = h_1 dq_1 \, \hat{\mathbf{e}}_1 + h_2 dq_2 \, \hat{\mathbf{e}}_2 + h_3 dq_3 \, \hat{\mathbf{e}}_3. \tag{3.133}$$

Note that h_i may be position-dependent and must have the dimension needed to cause $h_i dq_i$ to be a length.

Integrals in Curvilinear Coordinates

Given the scale factors h_i for a set of coordinates, either because they have been tabulated or because we have evaluated them via Eq. (3.131), we can use them to set up formulas for integration in the curvilinear coordinates. Line integrals will take the form

$$\int_C \mathbf{V} \cdot d\mathbf{r} = \sum_i \int_C V_i h_i dq_i. \tag{3.134}$$

Surface integrals take the same form as in Cartesian coordinates, with the exception that instead of expressions like $dx\,dy$ we have $(h_1 dq_1)(h_2 dq_2) = h_1 h_2\, dq_1\, dq_2$ etc. This means that

$$\int_S \mathbf{V} \cdot d\boldsymbol{\sigma} = \int_S V_1 h_2 h_3\, dq_2 dq_3 + \int_S V_2 h_3 h_1\, dq_3 dq_1 + \int_S V_3 h_1 h_2\, dq_1 dq_2. \tag{3.135}$$

The element of volume in orthogonal curvilinear coordinates is

$$d\tau = h_1 h_2 h_3\, dq_1 dq_2 dq_3, \tag{3.136}$$

so volume integrals take the form

$$\int_V \varphi(q_1, q_2, q_3) h_1 h_2 h_3 dq_1 dq_2 dq_3, \tag{3.137}$$

or the analogous expression with φ replaced by a vector $\mathbf{V}(q_1, q_2, q_3)$.

Differential Operators in Curvilinear Coordinates

We continue with a restriction to orthogonal coordinate systems.

Gradient—Because our curvilinear coordinates are orthogonal, the gradient takes the same form as for Cartesian coordinates, providing we use the differential displacements $dr_i = h_i\, dq_i$ in the formula. Thus, we have

$$\nabla\varphi(q_1, q_2, q_3) = \hat{\mathbf{e}}_1 \frac{1}{h_1}\frac{\partial\varphi}{\partial q_1} + \hat{\mathbf{e}}_2 \frac{1}{h_2}\frac{\partial\varphi}{\partial q_2} + \hat{\mathbf{e}}_3 \frac{1}{h_3}\frac{\partial\varphi}{\partial q_3}, \tag{3.138}$$

this corresponds to writing ∇ as

$$\nabla = \hat{\mathbf{e}}_1 \frac{1}{h_1}\frac{\partial}{\partial q_1} + \hat{\mathbf{e}}_2 \frac{1}{h_2}\frac{\partial}{\partial q_2} + \hat{\mathbf{e}}_3 \frac{1}{h_3}\frac{\partial}{\partial q_3}. \tag{3.139}$$

Divergence—This operator must have the same meaning as in Cartesian coordinates, so $\nabla \cdot \mathbf{V}$ must give the net outward flux of $\mathbf{V}$ per unit volume at the point of evaluation. The key difference from the Cartesian case is that an element of volume will no longer be a parallelepiped, as the scale factors h_i are in general functions of position. See Fig. 3.21. To compute the net outflow of $\mathbf{V}$ in the q_1 direction from a volume element defined by

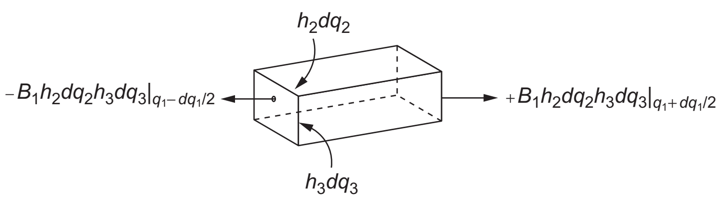

FIGURE 3.21 Outflow of B_1 in the q_1 direction from a curvilinear volume element.

dq_1, dq_2, dq_3 and centered at (q_1, q_2, q_3), we must form

$$\text{Net } q_1 \text{ outflow} = -V_1 h_2 h_3 \, dq_2 dq_3 \Big|_{q_1 - dq_1/2, q_2, q_3} + V_1 h_2 h_3 \, dq_2 dq_3 \Big|_{q_1 + dq_1/2, q_2, q_3}. \tag{3.140}$$

Note that not only V_1, but also $h_2 h_3$ must be evaluated at the displaced values of q_1; this product may have different values at $q_1 + dq_1/2$ and $q_1 - dq_1/2$. Rewriting Eq. (3.140) in terms of a derivative with respect to q_1, we have

$$\text{Net } q_1 \text{ outflow} = \frac{\partial}{\partial q_1}(V_1 h_2 h_3) dq_1 dq_2 dq_3.$$

Combining this with the q_2 and q_3 outflows and dividing by the differential volume $h_1 h_2 h_3 \, dq_1 dq_2 dq_3$, we get the formula

$$\nabla \cdot \mathbf{V}(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}(V_1 h_2 h_3) + \frac{\partial}{\partial q_2}(V_2 h_3 h_1) + \frac{\partial}{\partial q_3}(V_3 h_1 h_2) \right]. \tag{3.141}$$

Laplacian—From the formulas for the gradient and divergence, we can form the Laplacian in curvilinear coordinates:

$$\nabla^2 \varphi(q_1, q_2, q_3) = \nabla \cdot \nabla \varphi =$$
$$\frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial \varphi}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial \varphi}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial \varphi}{\partial q_3} \right) \right]. \tag{3.142}$$

Note that the Laplacian contains no cross derivatives, such as $\partial^2/\partial q_1 \partial q_2$. They do not appear because the coordinate system is orthogonal.

Curl—In the same spirit as our treatment of the divergence, we calculate the circulation around an element of area in the $q_1 q_2$ plane, and therefore associated with a vector in the q_3 direction. Referring to Fig. 3.22, the line integral $\oint \mathbf{B} \cdot d\mathbf{r}$ consists of four segment

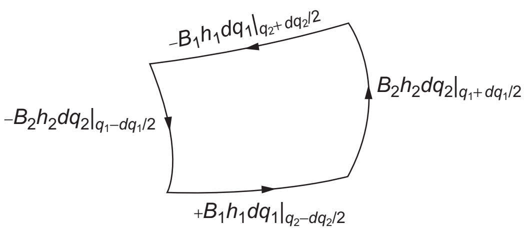

FIGURE 3.22 Circulation $\oint \mathbf{B} \cdot d\mathbf{r}$ around curvilinear element of area on a surface of constant q_3.

contributions, which to first order are

$$\text{Segment } 1 = (h_1 B_1)\Big|_{q_1, q_2 - dq_2/2, q_3} dq_1,$$

$$\text{Segment } 2 = (h_2 B_2)\Big|_{q_1 + dq_1/2, q_2, q_3} dq_2,$$

$$\text{Segment } 3 = -(h_1 B_1)\Big|_{q_1, q_2 + dq_2/2, q_3} dq_1,$$

$$\text{Segment } 4 = -(h_2 B_2)\Big|_{q_1 - dq_1/2, q_2, q_3} dq_2.$$

Keeping in mind that the h_i are functions of position, and that the loop has area $h_1 h_2\, dq_1 dq_2$, these contributions combine into a circulation per unit area

$$(\nabla \times \mathbf{B})_3 = \frac{1}{h_1 h_2}\left[-\frac{\partial}{\partial q_2}(h_1 B_1) + \frac{\partial}{\partial q_1}(h_2 B_2)\right].$$

The generalization of this result to arbitrary orientation of the circulation loop can be brought to the determinantal form

$$\nabla \times \mathbf{B} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} \hat{\mathbf{e}}_1 h_1 & \hat{\mathbf{e}}_2 h_2 & \hat{\mathbf{e}}_3 h_3 \\ \dfrac{\partial}{\partial q_1} & \dfrac{\partial}{\partial q_2} & \dfrac{\partial}{\partial q_3} \\ h_1 B_1 & h_2 B_2 & h_3 B_3 \end{vmatrix}. \tag{3.143}$$

Just as for Cartesian coordinates, this determinant is to be evaluated from the top down, so that the derivatives will act on its bottom row.

Circular Cylindrical Coordinates

Although there are at least 11 coordinate systems that are appropriate for use in solving physics problems, the evolution of computers and efficient programming techniques have greatly reduced the need for most of these coordinate systems, with the result that the discussion in this book is limited to (1) Cartesian coordinates, (2) spherical polar coordinates (treated in the next subsection), and (3) circular cylindrical coordinates, which we discuss here. Specifications and details of other coordinate systems will be found in the first two editions of this work and in Additional Readings at the end of this chapter (Morse and Feshbach, Margenau and Murphy).

In the circular cylindrical coordinate system the three curvilinear coordinates are labeled (ρ, φ, z). We use ρ for the perpendicular distance from the z-axis because we reserve r for the distance from the origin. The ranges of ρ, φ, and z are

$$0 \le \rho < \infty, \quad 0 \le \varphi < 2\pi, \quad -\infty < z < \infty.$$

For $\rho = 0$, φ is not well defined. The coordinate surfaces, shown in Fig. 3.23, follow:

1. Right circular cylinders having the z-axis as a common axis,

$$\rho = \left(x^2 + y^2\right)^{1/2} = \text{constant}.$$

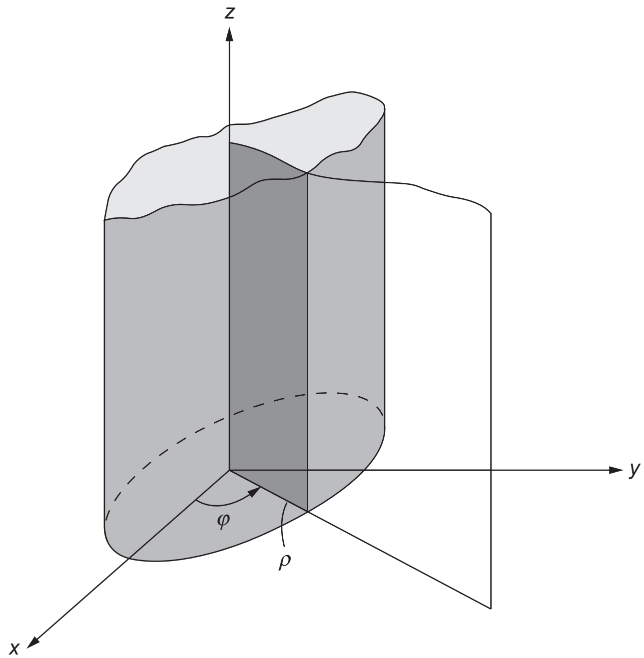

FIGURE 3.23 Cylindrical coordinates ρ, φ, z.

2. Half-planes through the z-axis, at an angle φ measured from the x direction,

$$\varphi = \tan^{-1}\left(\frac{y}{x}\right) = \text{constant}.$$

The arctangent is double valued on the range of φ, and the correct value of φ must be determined by the individual signs of x and y.

3. Planes parallel to the xy-plane, as in the Cartesian system,

$$z = \text{constant}.$$

Inverting the preceding equations, we can obtain

$$x = \rho\cos\varphi, \quad y = \rho\sin\varphi, \quad z = z. \tag{3.144}$$

This is essentially a 2-D curvilinear system with a Cartesian z-axis added on to form a 3-D system.

The coordinate vector $\mathbf{r}$ and a general vector $\mathbf{V}$ are expressed as

$$\mathbf{r} = \rho\,\hat{\mathbf{e}}_\rho + z\,\hat{\mathbf{e}}_z, \quad \mathbf{V} = V_\rho\,\hat{\mathbf{e}}_\rho + V_\varphi\,\hat{\mathbf{e}}_\varphi + V_z\,\hat{\mathbf{e}}_z.$$

From Eq. (3.131), the scale factors for these coordinates are

$$h_\rho = 1, \quad h_\varphi = \rho, \quad h_z = 1, \tag{3.145}$$

so the elements of displacement, area, and volume are

$$\begin{aligned} d\mathbf{r} &= \hat{\mathbf{e}}_\rho\, d\rho + \rho\, \hat{\mathbf{e}}_\varphi\, d\varphi + \hat{\mathbf{e}}_z\, dz, \\ d\boldsymbol{\sigma} &= \rho\, \hat{\mathbf{e}}_\rho\, d\varphi\, dz + \hat{\mathbf{e}}_\varphi d\rho\, dz + \rho\, \hat{\mathbf{e}}_z\, d\rho\, d\varphi, \\ d\tau &= \rho\, d\rho\, d\varphi\, dz. \end{aligned} \tag{3.146}$$

It is perhaps worth emphasizing that the unit vectors $\hat{\mathbf{e}}_\rho$ and $\hat{\mathbf{e}}_\varphi$ have directions that vary with φ; if expressions containing these unit vectors are differentiated with respect to φ, the derivatives of these unit vectors must be included in the computations.

Example 3.10.1 Kepler's Area Law for Planetary Motion

One of Kepler's laws states that the radius vector of a planet, relative to an origin at the sun, sweeps out equal areas in equal time. It is instructive to derive this relationship using cylindrical coordinates. For simplicity we consider a planet of unit mass and motion in the plane $z = 0$.

The gravitational force $\mathbf{F}$ is of the form $f(r)\,\hat{\mathbf{e}}_r$, and hence the torque about the origin, $\mathbf{r} \times \mathbf{F}$, vanishes, so angular momentum $\mathbf{L} = \mathbf{r} \times d\mathbf{r}/dt$ is conserved. To evaluate $d\mathbf{r}/dt$, we start from $d\mathbf{r}$ as given in Eq. (3.146), writing

$$\frac{d\mathbf{r}}{dt} = \hat{\mathbf{e}}_\rho \dot{\rho} + \hat{\mathbf{e}}_\varphi \rho \dot{\varphi},$$

where we have used the dot notation (invented by Newton) to indicate time derivatives. We now form

$$\mathbf{L} = \rho\, \hat{\mathbf{e}}_\rho \times \left(\hat{\mathbf{e}}_\rho \dot{\rho} + \hat{\mathbf{e}}_\varphi \rho \dot{\varphi}\right) = \rho^2 \dot{\varphi} \hat{\mathbf{e}}_z.$$

We conclude that $\rho^2\, \dot{\varphi}$ is constant. Making the identification $\rho^2\, \dot{\varphi} = 2\, dA/dt$, where A is the area swept out, we confirm Kepler's law. ■

Continuing now to the vector differential operators, using Eqs. (3.138), (3.141), (3.142), and (3.143), we have

$$\nabla\psi(\rho,\varphi,z) = \hat{\mathbf{e}}_\rho \frac{\partial\psi}{\partial b\rho} + \hat{\mathbf{e}}_\varphi \frac{1}{\rho}\frac{\partial\psi}{\partial\varphi} + \hat{\mathbf{e}}_z \frac{\partial\psi}{\partial z}, \tag{3.147}$$

$$\nabla\cdot\mathbf{V} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho V_\rho) + \frac{1}{\rho}\frac{\partial V_\varphi}{\partial\varphi} + \frac{\partial V_z}{\partial z}, \tag{3.148}$$

$$\nabla^2\psi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{\partial^2\psi}{\partial z^2}, \tag{3.149}$$

$$\nabla\times V = \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\hat{\mathbf{e}}_\varphi & \hat{\mathbf{e}}_z \\ \dfrac{\partial}{\partial\rho} & \dfrac{\partial}{\partial\varphi} & \dfrac{\partial}{\partial z} \\ V_\rho & \rho V_\varphi & V_z \end{vmatrix}. \tag{3.150}$$

Finally, for problems such as circular wave guides and cylindrical cavity resonators, one needs the vector Laplacian $\nabla^2\mathbf{V}$. From Eq. (3.70), its components in cylindrical coordinates can be shown to be

$$\begin{aligned}
\nabla^2\mathbf{V}\Big|_\rho &= \nabla^2 V_\rho - \frac{1}{\rho^2}V_\rho - \frac{2}{\rho^2}\frac{\partial V_\varphi}{\partial\varphi}, \\
\nabla^2\mathbf{V}\Big|_\varphi &= \nabla^2 V_\varphi - \frac{1}{\rho^2}V_\varphi + \frac{2}{\rho^2}\frac{\partial V_\rho}{\partial\varphi}, \\
\nabla^2\mathbf{V}\Big|_z &= \nabla^2 V_z.
\end{aligned} \tag{3.151}$$

Example 3.10.2 A Navier-Stokes Term

The Navier-Stokes equations of hydrodynamics contain a nonlinear term

$$\nabla\times\left[\mathbf{v}\times(\nabla\times\mathbf{v})\right],$$

where $\mathbf{v}$ is the fluid velocity. For fluid flowing through a cylindrical pipe in the z direction,

$$\mathbf{v} = \hat{\mathbf{e}}_z v(\rho).$$

From Eq. (3.150),

$$\nabla\times\mathbf{v} = \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\hat{\mathbf{e}}_\varphi & \hat{\mathbf{e}}_z \\ \dfrac{\partial}{\partial\rho} & \dfrac{\partial}{\partial\varphi} & \dfrac{\partial}{\partial z} \\ 0 & 0 & v(\rho) \end{vmatrix} = -\hat{\mathbf{e}}_\varphi\frac{\partial v}{\partial\rho},$$

$$\mathbf{v}\times(\nabla\times\mathbf{v}) = \begin{vmatrix} \hat{\mathbf{e}}_\rho & \hat{\mathbf{e}}_\varphi & \hat{\mathbf{e}}_z \\ 0 & 0 & v \\ 0 & -\dfrac{\partial v}{\partial\rho} & 0 \end{vmatrix} = \hat{\mathbf{e}}_\rho\, v(\rho)\,\frac{\partial v}{\partial\rho}.$$

Finally,

$$\nabla\times\big(\mathbf{v}\times(\nabla\times\mathbf{v})\big) = \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\hat{\mathbf{e}}_\varphi & \hat{\mathbf{e}}_z \\ \dfrac{\partial}{\partial\rho} & \dfrac{\partial}{\partial\varphi} & \dfrac{\partial}{\partial z} \\ v\dfrac{\partial v}{\partial\rho} & 0 & 0 \end{vmatrix} = 0.$$

For this particular case, the nonlinear term vanishes. ■

Spherical Polar Coordinates

Spherical polar coordinates were introduced as an initial example of a curvilinear coordinate system, and were illustrated in Fig. 3.19. We reiterate: The coordinates are labeled (r,θ,φ). Their ranges are

$$0\le r<\infty, \quad 0\le\theta\le\pi, \quad 0\le\varphi<2\pi.$$

For $r = 0$, neither θ nor φ is well defined. Additionally, φ is ill-defined for $\theta = 0$ and $\theta = \pi$. The coordinate surfaces follow:

1. Concentric spheres centered at the origin,

$$r = \left(x^2 + y^2 + z^2\right)^{1/2} = \text{constant}.$$

2. Right circular cones centered on the z (polar) axis with vertices at the origin,

$$\theta = \arccos \frac{z}{r} = \text{constant}.$$

3. Half-planes through the z (polar) axis, at an angle φ measured from the x direction,

$$\varphi = \arctan \frac{y}{x} = \text{constant}.$$

The arctangent is double valued on the range of φ, and the correct value of φ must be determined by the individual signs of x and y.

Inverting the preceding equations, we can obtain

$$x = r \, \sin\theta \cos\varphi, \quad y = r \, \sin\theta \sin\varphi, \quad z = r \, \cos\theta. \tag{3.152}$$

The coordinate vector $\mathbf{r}$ and a general vector $\mathbf{V}$ are expressed as

$$\mathbf{r} = r \, \hat{\mathbf{e}}_r, \quad \mathbf{V} = V_r \, \hat{\mathbf{e}}_r + V_\theta \, \hat{\mathbf{e}}_\theta + V_\varphi \, \hat{\mathbf{e}}_\varphi.$$

From Eq. (3.131), the scale factors for these coordinates are

$$h_r = 1, \quad h_\theta = r, \quad h_\varphi = r \, \sin\theta, \tag{3.153}$$

so the elements of displacement, area, and volume are

$$\begin{aligned} d\mathbf{r} &= \hat{\mathbf{e}}_r \, dr + r \, \hat{\mathbf{e}}_\theta \, d\theta + r \, \sin\theta \, \hat{\mathbf{e}}_\varphi \, d\varphi, \\ d\boldsymbol{\sigma} &= r^2 \sin\theta \, \hat{\mathbf{e}}_r \, d\theta \, d\varphi + r \, \sin\theta \, \hat{\mathbf{e}}_\theta \, dr \, d\varphi + r \, \hat{\mathbf{e}}_\varphi \, dr \, d\theta, \\ d\tau &= r^2 \sin\theta \, d\rho \, d\theta \, d\varphi. \end{aligned} \tag{3.154}$$

Frequently one encounters a need to perform a surface integration over the angles, in which case the angular dependence of $d\boldsymbol{\sigma}$ reduces to

$$d\Omega = \sin\theta \, d\theta \, d\varphi, \tag{3.155}$$

where $d\Omega$ is called an element of solid angle, and has the property that its integral over all angles has the value

$$\int d\Omega = 4\pi.$$

Note that for spherical polar coordinates, all three of the unit vectors have directions that depend on position, and this fact must be taken into account when expressions containing the unit vectors are differentiated.

The vector differential operators may now be evaluated, using Eqs. (3.138), (3.141), (3.142), and (3.143):

$$\nabla\psi(r,\theta,\varphi) = \hat{\mathbf{e}}_r \frac{\partial\psi}{\partial r} + \hat{\mathbf{e}}_\theta \frac{1}{r}\frac{\partial\psi}{\partial\theta} + \hat{\mathbf{e}}_\varphi \frac{1}{r\sin\theta}\frac{\partial\psi}{\partial\varphi}, \tag{3.156}$$

$$\nabla\cdot\mathbf{V} = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}(r^2 V_r) + r\frac{\partial}{\partial\theta}(\sin\theta\, V_\theta) + r\frac{\partial V_\varphi}{\partial\varphi}\right], \tag{3.157}$$

$$\nabla^2\psi = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin\theta}\frac{\partial^2\psi}{\partial\varphi^2}\right], \tag{3.158}$$

$$\nabla\times V = \frac{1}{r^2\sin\theta}\begin{vmatrix} \hat{\mathbf{e}}_r & r\,\hat{\mathbf{e}}_\theta & r\sin\theta\,\hat{\mathbf{e}}_\varphi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial\varphi} \\ V_r & rV_\theta & r\sin\theta\, V_\varphi \end{vmatrix}. \tag{3.159}$$

Finally, again using Eq. (3.70), the components of the vector Laplacian $\nabla^2\mathbf{V}$ in spherical polar coordinates can be shown to be

$$\begin{aligned}
\nabla^2\mathbf{V}\Big|_r &= \nabla^2 V_r - \frac{2}{r^2}V_r - \frac{2}{r^2}\cot\theta\, V_\theta - \frac{2}{r^2}\frac{\partial V_\theta}{\partial\theta} - \frac{2}{r^2\sin\theta}\frac{\partial V_\varphi}{\partial\varphi},\\
\nabla^2\mathbf{V}\Big|_\theta &= \nabla^2 V_\theta - \frac{1}{r^2\sin^2\theta}V_\theta + \frac{2}{r^2}\frac{\partial V_r}{\partial\theta} - \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial V_\varphi}{\partial\varphi},\\
\nabla^2\mathbf{V}\Big|_\varphi &= \nabla^2 V_\varphi - \frac{1}{r^2\sin^2\theta}V_\varphi + \frac{2}{r^2\sin\theta}\frac{\partial V_r}{\partial\varphi} + \frac{2\cos\theta}{r^2\sin^2\theta}\frac{\partial V_\varphi}{\partial\varphi}.
\end{aligned} \tag{3.160}$$

Example 3.10.3 ∇, $\nabla\cdot$, $\nabla\times$ for a Central Force

We can now easily derive some of the results previously obtained more laboriously in Cartesian coordinates:

From Eq. (3.156),

$$\nabla f(r) = \hat{\mathbf{e}}_r\frac{df}{dr}, \qquad \nabla r^n = \hat{\mathbf{e}}_r n r^{n-1}. \tag{3.161}$$

Specializing to the Coulomb potential of a point charge at the origin, $V = Ze/(4\pi\varepsilon_0 r)$, so the electric field has the expected value $\mathbf{E} = -\nabla V = (Ze/4\pi\varepsilon_0 r^2)\hat{\mathbf{e}}_r$.

Taking next the divergence of a radial function, we have from Eq. (3.157),

$$\nabla\cdot\big(\hat{\mathbf{e}}_r\, f(r)\big) = \frac{2}{r}f(r) + \frac{df}{dr}, \qquad \nabla\cdot(\hat{\mathbf{e}}_r\, r^n) = (n+2)r^{n-1}. \tag{3.162}$$

Specializing the above to the Coulomb force ($n = -2$), we have (except for $r = 0$) $\nabla\cdot r^{-2} = 0$, which is consistent with Gauss' law.

Continuing now to the Laplacian, from Eq. (3.158) we have

$$\nabla^2 f(r) = \frac{2}{r}\frac{df}{dr} + \frac{d^2 f}{dr^2}, \qquad \nabla^2 r^n = n(n+1)r^{n-2}, \tag{3.163}$$

in contrast to the ordinary second derivative of r^n involving $n-1$.

Finally, from Eq. (3.159),

$$\nabla \times \big(\hat{\mathbf{e}}_r f(r)\big) = 0, \tag{3.164}$$

which confirms that central forces are irrotational. ■

Example 3.10.4 MAGNETIC VECTOR POTENTIAL

A single current loop in the xy-plane has a vector potential $\mathbf{A}$ that is a function only of r and θ, is entirely in the $\hat{\mathbf{e}}_\varphi$ direction and is related to the current density $\mathbf{J}$ by the equation

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B} = \nabla \times \big[\nabla \times \hat{\mathbf{e}}_\varphi A_\varphi(r,\theta)\big].$$

In spherical polar coordinates this reduces to

$$\begin{aligned}\mu_0 \mathbf{J} &= \nabla \times \frac{1}{r^2 \sin\theta}\begin{vmatrix} \hat{\mathbf{e}}_r & r\hat{\mathbf{e}}_\theta & r\sin\theta\,\hat{\mathbf{e}}_\varphi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} & \dfrac{\partial}{\partial \varphi} \\ 0 & 0 & r\sin\theta A_\varphi \end{vmatrix} \\ &= \nabla \times \frac{1}{r^2 \sin\theta}\left[\hat{\mathbf{e}}_r \frac{\partial}{\partial\theta}(r\sin\theta A_\varphi) - r\hat{\mathbf{e}}_\theta \frac{\partial}{\partial r}(r\sin\theta A_\varphi)\right].\end{aligned}$$

Taking the curl a second time, we obtain

$$\mu_0 \mathbf{J} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \hat{\mathbf{e}}_r & r\hat{\mathbf{e}}_\theta & r\sin\theta\,\hat{\mathbf{e}}_\varphi \\ \dfrac{\partial}{\partial r} & \dfrac{\partial}{\partial \theta} & \dfrac{\partial}{\partial \varphi} \\ \dfrac{1}{r\sin\theta}\dfrac{\partial}{\partial\theta}(\sin\theta A_\varphi) & -\dfrac{1}{r}\dfrac{\partial}{\partial r}(rA_\varphi) & 0 \end{vmatrix}.$$

Expanding this determinant from the top down, we reach

$$\mu_0 \mathbf{J} = -\hat{\mathbf{e}}_\varphi \left[\frac{\partial^2 A_\varphi}{\partial r^2} + \frac{2}{r}\frac{\partial A_\varphi}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial A_\varphi}{\partial\theta}\right) - \frac{1}{r^2\sin^2\theta}A_\varphi\right]. \tag{3.165}$$

Note that we get, in addition to $\nabla^2 A_\varphi$, one more term: $-A_\varphi/r^2\sin^2\theta$. ■

Example 3.10.5 STOKES' THEOREM

As a final example, let's compute $\oint \mathbf{B}\cdot d\mathbf{r}$ for a closed loop, comparing the result with integrals $\int(\nabla \times \mathbf{B})\cdot d\boldsymbol{\sigma}$ for two different surfaces having the same perimeter. We use spherical polar coordinates, taking $\mathbf{B} = e^{-r}\hat{\mathbf{e}}_\varphi$.

The loop will be a unit circle about the origin in the xy-plane; the line integral about it will be taken in a counterclockwise sense as viewed from positive z, so the normal to the surfaces it bounds will pass through the xy-plane in the direction of positive z. The surfaces we consider are (1) a circular disk bounded by the loop, and (3) a hemisphere bounded by the loop, with its surface in the region $z < 0$. See Fig. 3.24.

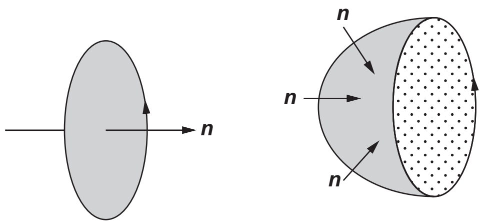

FIGURE 3.24 Surfaces for Example 3.10.5: (left) S_1, disk; (right) S_2, hemisphere.

For the line integral, $d\mathbf{r} = r\sin\theta\,\hat{\mathbf{e}}_\varphi\,d\varphi$, which reduces to $d\mathbf{r} = \hat{\mathbf{e}}_\varphi\,d\varphi$ since $\theta = \pi/2$ and $r = 1$ on the entire loop. We then have

$$\oint \mathbf{B}\cdot d\mathbf{r} = \int_{\varphi=0}^{2\pi} e^{-1}\,\hat{\mathbf{e}}_\varphi\cdot\hat{\mathbf{e}}_\varphi\,d\varphi = \frac{2\pi}{e}.$$

For the surface integrals, we need $\nabla\times\mathbf{B}$:

$$\nabla\times\mathbf{B} = \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial\theta}(r\sin\theta\,e^{-r})\hat{\mathbf{e}}_r - r\frac{\partial}{\partial r}(r\sin\theta\,e^{-r})\hat{\mathbf{e}}_\theta\right]$$
$$= \frac{e^{-r}\cos\theta}{r\sin\theta}\,\hat{\mathbf{e}}_r - (1-r)e^{-r}\,\hat{\mathbf{e}}_\theta.$$

Taking first the disk, at all points of which $\theta = \pi/2$, with integration range $0 \le r \le 1$, and $0 \le \varphi < 2\pi$, we note that $d\boldsymbol{\sigma} = -\hat{\mathbf{e}}_\theta\,r\sin\theta\,dr\,d\varphi = -\hat{\mathbf{e}}_\theta\,r\,dr\,d\varphi$. The minus sign arises because the positive normal is in the direction of **decreasing** θ. Then,

$$\int_{S_1} -(\nabla\times\mathbf{B})\cdot\hat{\mathbf{e}}_\theta\,r\,dr\,d\varphi = \int_0^{2\pi} d\varphi\int_0^1 dr\,(1-r)\,e^{-r} = \frac{2\pi}{e}.$$

For the hemisphere, defined by $r = 1$, $\pi/2 \le \theta < \pi$, and $0 \le \varphi < 2\pi$, we have $d\boldsymbol{\sigma} = -\hat{\mathbf{e}}_r\,r^2\sin\theta\,d\theta\,d\varphi = -\hat{\mathbf{e}}_r\sin\theta\,d\theta\,d\varphi$ (the normal is in the direction of decreasing r), and

$$\int_{S_2} -(\nabla\times\mathbf{B})\cdot\hat{\mathbf{e}}_r\,\sin\theta\,d\theta\,d\varphi = -\int_{\pi/2}^{\pi} d\theta e^{-1}\cos\theta\int_0^{2\pi} d\varphi = \frac{2\pi}{e}.$$

The results for both surfaces agree with that from the line integral of their common perimeter. Because $\nabla\times\mathbf{B}$ is solenoidal, all the flux that passes through the disk in the xy-plane must continue through the hemispherical surface, and for that matter, through **any** surface with the same perimeter. That is why Stokes' theorem is indifferent to features of the surface other than its perimeter. ■

Rotation and Reflection in Spherical Coordinates

It is infrequent that rotational coordinate transformations need be applied in curvilinear coordinate systems, and they usually arise only in contexts that are compatible with the symmetry of the coordinate system. We limit the current discussion to rotations (and reflections) in spherical polar coordinates.

Rotation—Suppose a coordinate rotation identified by Euler angles (α, β, γ) converts the coordinates of a point from (r, θ, φ) to (r, θ', φ'). It is obvious that r retains its original value. Two questions arise: (1) How are θ' and φ' related to θ and φ? and (2) How do the components of a vector $\mathbf{A}$, namely $(A_r,\ A_\theta,\ A_\varphi)$, transform?

It is simplest to proceed, as we did for Cartesian coordinates, by analyzing the three consecutive rotations implied by the Euler angles. The first rotation, by an angle α about the z-axis, leaves θ unchanged, and converts φ into $\varphi - \alpha$. However, it causes no change in any of the components of $\mathbf{A}$.

The second rotation, which inclines the polar direction by an angle β toward the (new) x-axis, does change the values of both θ and φ and, in addition, changes the directions of $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$. Referring to Fig. 3.25, we see that these two unit vectors are subjected to a rotation χ in the plane tangent to the sphere of constant r, thereby yielding new unit vectors $\hat{\mathbf{e}}'_\theta$ and $\hat{\mathbf{e}}'_\varphi$ such that

$$\hat{\mathbf{e}}_\theta = \cos\chi\,\hat{\mathbf{e}}'_\theta - \sin\chi\,\hat{\mathbf{e}}'_\varphi, \quad \hat{\mathbf{e}}_\varphi = \sin\chi\,\hat{\mathbf{e}}'_\theta + \cos\chi\,\hat{\mathbf{e}}'_\varphi.$$

This transformation corresponds to

$$\mathsf{S}_2 = \begin{pmatrix} \cos\chi & \sin\chi \\ -\sin\chi & \cos\chi \end{pmatrix}.$$

Carrying out the spherical trigonometry corresponding to Fig. 3.25, we have the new coordinates

$$\cos\theta' = cos\beta\cos\theta + \sin\beta\sin\theta\cos(\varphi-\alpha), \quad \cos\varphi' = \frac{\cos\beta\cos\theta' - \cos\theta}{\sin\beta\sin\theta'}, \tag{3.166}$$

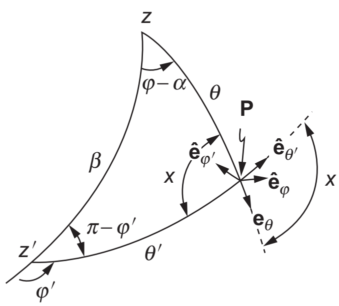

FIGURE 3.25 Rotation and unit vectors in spherical polar coordinates, shown on a sphere of radius r. The original polar direction is marked z; it is moved to the direction z', at an inclination given by the Euler angle β. The unit vectors $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$ at the point $\mathbf{P}$ are thereby rotated through the angle χ.

and

$$\cos\chi = \frac{\cos\beta - \cos\theta\cos\theta'}{\sin\theta\sin\theta'}. \tag{3.167}$$

The third rotation, by an angle γ about the new z-axis, leaves the components of $\mathbf{A}$ unchanged but requires the replacement of φ' by $\varphi' - \gamma$.

Summarizing,

$$\begin{pmatrix} A'_r \\ A'_\theta \\ A'_\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\chi & \sin\chi \\ 0 & -\sin\chi & \cos\chi \end{pmatrix} \begin{pmatrix} A_r \\ A_\theta \\ A_\varphi \end{pmatrix}. \tag{3.168}$$

This equation specifies the components of $\mathbf{A}$ in the rotated coordinates at the point $(r, \theta', \varphi' - \gamma)$ in terms of the original components at the same physical point, (r, θ, φ).

Reflection—Inversion of the coordinate system reverses the sign of each Cartesian coordinate. Taking the angle φ as that which moves the new $+x$ coordinate toward the new $+y$ coordinate, the system (which was originally right-handed) now becomes left-handed. The coordinates (r, θ, φ) of a (fixed) point become, in the new system, $(r, \pi - \theta, \pi + \varphi)$. The unit vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\varphi$ are invariant under inversion, but $\hat{\mathbf{e}}_\theta$ changes sign, so

$$\begin{pmatrix} A'_r \\ A'_\theta \\ A'_\varphi \end{pmatrix} = \begin{pmatrix} A_r \\ -A_\theta \\ A_\varphi \end{pmatrix}, \quad \text{coordinate inversion.} \tag{3.169}$$

Exercises

3.10.1 The u-, v-, z-coordinate system frequently used in electrostatics and in hydrodynamics is defined by

$$xy = u, \quad x^2 - y^2 = v, \quad z = z.$$

This u-, v-, z-system is orthogonal.

(a) In words, describe briefly the nature of each of the three families of coordinate surfaces.

(b) Sketch the system in the xy-plane showing the intersections of surfaces of constant u and surfaces of constant v with the xy-plane.

(c) Indicate the directions of the unit vectors $\hat{\mathbf{e}}_u$ and $\hat{\mathbf{e}}_v$ in all four quadrants.

(d) Finally, is this u-, v-, z-system right-handed ($\hat{\mathbf{e}}_u \times \hat{\mathbf{e}}_v = +\hat{\mathbf{e}}_z$) or left-handed ($\hat{\mathbf{e}}_u \times \hat{\mathbf{e}}_v = -\hat{\mathbf{e}}_z$)?

3.10.2 The elliptic cylindrical coordinate system consists of three families of surfaces:

$$(1)\quad \frac{x^2}{a^2\cosh^2 u} + \frac{y^2}{a^2\sinh^2 u} = 1; \qquad (2)\quad \frac{x^2}{a^2\cos^2 v} - \frac{y^2}{a^2\sin^2 v} = 1; \qquad (3)\quad z = z.$$

Sketch the coordinate surfaces $u =$ constant and $v =$ constant as they intersect the first quadrant of the xy-plane. Show the unit vectors $\hat{\mathbf{e}}_u$ and $\hat{\mathbf{e}}_v$. The range of u is $0 \le u < \infty$. The range of v is $0 \le v \le 2\pi$.

3.10.3 Develop arguments to show that dot and cross products (not involving ∇) in orthogonal curvilinear coordinates in $\mathbb{R}^3$ proceed, as in Cartesian coordinates, **with no involvement of scale factors**.

3.10.4 With $\hat{\mathbf{e}}_1$ a unit vector in the direction of increasing q_1, show that

(a) $$\nabla \cdot \hat{\mathbf{e}}_1 = \frac{1}{h_1 h_2 h_3} \frac{\partial(h_2 h_3)}{\partial q_1}$$

(b) $$\nabla \times \hat{\mathbf{e}}_1 = \frac{1}{h_1}\left[\hat{\mathbf{e}}_2 \frac{1}{h_3}\frac{\partial h_1}{\partial q_3} - \hat{\mathbf{e}}_3 \frac{1}{h_2}\frac{\partial h_1}{\partial q_2}\right].$$

Note that even though $\hat{\mathbf{e}}_1$ is a unit vector, its divergence and curl **do not necessarily vanish**.

3.10.5 Show that a set of orthogonal unit vectors $\hat{\mathbf{e}}_i$ may be defined by

$$\hat{\mathbf{e}}_i = \frac{1}{h_i}\frac{\partial \mathbf{r}}{\partial q_i}.$$

In particular, show that $\hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_i = 1$ leads to an expression for h_i in agreement with Eq. (3.131).
The above equation for $\hat{\mathbf{e}}_i$ may be taken as a starting point for deriving

$$\frac{\partial \hat{\mathbf{e}}_i}{\partial q_j} = \hat{\mathbf{e}}_j \frac{1}{h_i}\frac{\partial h_j}{\partial q_i}, \quad i \neq j$$

and

$$\frac{\partial \hat{\mathbf{e}}_i}{\partial q_i} = -\sum_{j \neq i} \hat{\mathbf{e}}_j \frac{1}{h_j}\frac{\partial h_i}{\partial q_j}.$$

3.10.6 Resolve the circular cylindrical unit vectors into their Cartesian components (see Fig. 3.23).

ANS. $$\hat{\mathbf{e}}_\rho = \hat{\mathbf{e}}_x \cos\varphi + \hat{\mathbf{e}}_y \sin\varphi,$$
$$\hat{\mathbf{e}}_\varphi = -\hat{\mathbf{e}}_x \sin\varphi + \hat{\mathbf{e}}_y \cos\varphi,$$
$$\hat{\mathbf{e}}_z = \hat{\mathbf{e}}_z.$$

3.10.7 Resolve the Cartesian unit vectors into their circular cylindrical components (see Fig. 3.23).

ANS. $$\hat{\mathbf{e}}_x = \hat{\mathbf{e}}_\rho \cos\varphi - \hat{\mathbf{e}}_\varphi \sin\varphi,$$
$$\hat{\mathbf{e}}_y = \hat{\mathbf{e}}_\rho \sin\varphi + \hat{\mathbf{e}}_\varphi \cos\varphi,$$
$$\hat{\mathbf{e}}_z = \hat{\mathbf{e}}_z.$$

3.10.8 From the results of Exercise 3.10.6, show that

$$\frac{\partial \hat{\mathbf{e}}_\rho}{\partial \varphi} = \hat{\mathbf{e}}_\varphi, \quad \frac{\partial \hat{\mathbf{e}}_\varphi}{\partial \varphi} = -\hat{\mathbf{e}}_\rho$$

and that all other first derivatives of the circular cylindrical unit vectors with respect to the circular cylindrical coordinates vanish.

3.10.9 Compare $\nabla \cdot \mathbf{V}$ as given for cylindrical coordinates in Eq. (3.148) with the result of its computation by applying to $\mathbf{V}$ the operator

$$\nabla = \hat{\mathbf{e}}_\rho \frac{\partial}{\partial \rho} + \hat{\mathbf{e}}_\varphi \frac{1}{\rho}\frac{\partial}{\partial \varphi} + \hat{\mathbf{e}}_z \frac{\partial}{\partial z}$$

Note that ∇ acts both on the unit vectors and on the components of $\mathbf{V}$.

3.10.10 (a) Show that $\mathbf{r} = \hat{\mathbf{e}}_\rho \rho + \hat{\mathbf{e}}_z z$.

(b) Working entirely in circular cylindrical coordinates, show that

$$\nabla \cdot \mathbf{r} = 3 \quad \text{and} \quad \nabla \times \mathbf{r} = 0.$$

3.10.11 (a) Show that the parity operation (reflection through the origin) on a point (ρ, φ, z) relative to **fixed** x-, y-, z-axes consists of the transformation

$$\rho \to \rho, \quad \varphi \to \varphi \pm \pi, \quad z \to -z.$$

(b) Show that $\hat{\mathbf{e}}_\rho$ and $\hat{\mathbf{e}}_\varphi$ have odd parity (reversal of direction) and that $\hat{\mathbf{e}}_z$ has even parity.

Note. The Cartesian unit vectors $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, and $\hat{\mathbf{e}}_z$ remain constant.

3.10.12 A rigid body is rotating about a fixed axis with a constant angular velocity $\boldsymbol{\omega}$. Take $\boldsymbol{\omega}$ to lie along the z-axis. Express the position vector $\mathbf{r}$ in circular cylindrical coordinates and using circular cylindrical coordinates,

(a) calculate $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$,

(b) calculate $\nabla \times \mathbf{v}$.

ANS. (a) $\mathbf{v} = \hat{\mathbf{e}}_\varphi \omega \rho$

(b) $\nabla \times \mathbf{v} = 2\boldsymbol{\omega}$.

3.10.13 Find the circular cylindrical components of the velocity and acceleration of a moving particle,

$$v_\rho = \dot{\rho}, \quad a_\rho = \ddot{\rho} - \rho\dot{\varphi}^2,$$
$$v_\varphi = \rho\dot{\varphi}, \quad a_\varphi = \rho\ddot{\varphi} + 2\dot{\rho}\dot{\varphi},$$
$$v_z = \dot{z}, \quad a_z = \ddot{z}.$$

Hint.

$$\begin{aligned}\mathbf{r}(t) &= \hat{\mathbf{e}}_\rho(t)\rho(t) + \hat{\mathbf{e}}_z z(t)\\ &= [\hat{\mathbf{e}}_x \cos\varphi(t) + \hat{\mathbf{e}}_y \sin\varphi(t)]\rho(t) + \hat{\mathbf{e}}_z z(t).\end{aligned}$$

Note. $\dot{\rho} = d\rho/dt$, $\ddot{\rho} = d^2\rho/dt^2$, and so on.

3.10.14 In right circular cylindrical coordinates, a particular vector function is given by

$$\mathbf{V}(\rho, \varphi) = \hat{\mathbf{e}}_\rho V_\rho(\rho, \varphi) + \hat{\mathbf{e}}_\varphi V_\varphi(\rho, \varphi).$$

Show that $\nabla \times \mathbf{V}$ has only a z-component. Note that this result will hold for any vector confined to a surface $q_3 =$ constant as long as the products $h_1 V_1$ and $h_2 V_2$ are each independent of q_3.

3.10.15 A conducting wire along the z-axis carries a current I. The resulting magnetic vector potential is given by

$$\mathbf{A} = \hat{\mathbf{e}}_z \frac{\mu I}{2\pi} \ln\left(\frac{1}{\rho}\right).$$

Show that the magnetic induction $\mathbf{B}$ is given by

$$\mathbf{B} = \hat{\mathbf{e}}_\varphi \frac{\mu I}{2\pi\rho}.$$

3.10.16 A force is described by

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

(a) Express $\mathbf{F}$ in circular cylindrical coordinates.
Operating entirely in circular cylindrical coordinates for (b) and (c),
(b) Calculate the curl of $\mathbf{F}$ and
(c) Calculate the work done by $\mathbf{F}$ in encircling the unit circle once counter-clockwise.
(d) How do you reconcile the results of (b) and (c)?

3.10.17 A calculation of the magnetohydrodynamic pinch effect involves the evaluation of $(\mathbf{B}\cdot\nabla)\mathbf{B}$. If the magnetic induction $\mathbf{B}$ is taken to be $\mathbf{B} = \hat{\mathbf{e}}_\varphi B_\varphi(\rho)$, show that

$$(\mathbf{B}\cdot\nabla)\mathbf{B} = -\hat{\mathbf{e}}_\rho B_\varphi^2/\rho.$$

3.10.18 Express the spherical polar unit vectors in terms of Cartesian unit vectors.

ANS.
$$\hat{\mathbf{e}}_r = \hat{\mathbf{e}}_x \sin\theta\cos\varphi + \hat{\mathbf{e}}_y \sin\theta\sin\varphi + \hat{\mathbf{e}}_z\cos\theta,$$
$$\hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_x \cos\theta\cos\varphi + \hat{\mathbf{e}}_y \cos\theta\sin\varphi - \hat{\mathbf{e}}_z\sin\theta,$$
$$\hat{\mathbf{e}}_\varphi = -\hat{\mathbf{e}}_x \sin\varphi + \hat{\mathbf{e}}_y\cos\varphi.$$

3.10.19 Resolve the Cartesian unit vectors into their spherical polar components:

$$\hat{\mathbf{e}}_x = \hat{\mathbf{e}}_r \sin\theta\cos\varphi + \hat{\mathbf{e}}_\theta\cos\theta\cos\varphi - \hat{\mathbf{e}}_\varphi\sin\varphi,$$

$$\hat{\mathbf{e}}_y = \hat{\mathbf{e}}_r \sin\theta\sin\varphi + \hat{\mathbf{e}}_\theta\cos\theta\sin\varphi + \hat{\mathbf{e}}_\varphi\cos\varphi,$$

$$\hat{\mathbf{e}}_z = \hat{\mathbf{e}}_r\cos\theta - \hat{\mathbf{e}}_\theta\sin\theta.$$

3.10.20 (a) Explain why it is not possible to relate a column vector $\mathbf{r}$ (with components x, y, z) to another column vector $\mathbf{r}'$ (with components r, θ, φ), via a matrix equation of the form $\mathbf{r}' = \mathsf{B}\mathbf{r}$.
(b) One can write a matrix equation relating the Cartesian components of a vector to its components in spherical polar coordinates. Find the transformation matrix and determine whether it is orthogonal.

3.10.21 Find the transformation matrix that converts the components of a vector in spherical polar coordinates into its components in circular cylindrical coordinates. Then find the matrix of the inverse transformation.

3.10.22 (a) From the results of Exercise 3.10.18, calculate the partial derivatives of $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\theta$, and $\hat{\mathbf{e}}_\varphi$ with respect to r, θ, and φ.

(b) With ∇ given by

$$\hat{\mathbf{e}}_r \frac{\partial}{\partial r} + \hat{\mathbf{e}}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\mathbf{e}}_\varphi \frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi}$$

(greatest space rate of change), use the results of part (a) to calculate $\nabla \cdot \nabla \psi$. This is an alternate derivation of the Laplacian.

Note. The derivatives of the left-hand ∇ operate on the unit vectors of the right-hand ∇ **before** the dot product is evaluated.

3.10.23 A rigid body is rotating about a fixed axis with a constant angular velocity $\boldsymbol{\omega}$. Take $\boldsymbol{\omega}$ to be along the z-axis. Using spherical polar coordinates,

(a) calculate $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$.
(b) calculate $\nabla \times \mathbf{v}$.

ANS. (a) $\mathbf{v} = \hat{\mathbf{e}}_\varphi \omega r \sin\theta$.
(b) $\nabla \times \mathbf{v} = 2\boldsymbol{\omega}$.

3.10.24 A certain vector $\mathbf{V}$ has no radial component. Its curl has no tangential components. What does this imply about the radial dependence of the tangential components of $\mathbf{V}$?

3.10.25 Modern physics lays great stress on the property of parity (whether a quantity remains invariant or changes sign under an inversion of the coordinate system). In Cartesian coordinates this means $x \to -x$, $y \to -y$, and $z \to -z$.

(a) Show that the inversion (reflection through the origin) of a point (r, θ, φ) relative to **fixed** x-, y-, z-axes consists of the transformation

$$r \to r, \quad \theta \to \pi - \theta, \quad \varphi \to \varphi \pm \pi.$$

(b) Show that $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\varphi$ have odd parity (reversal of direction) and that $\hat{\mathbf{e}}_\theta$ has even parity.

3.10.26 With $\mathbf{A}$ any vector,

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

(a) Verify this result in Cartesian coordinates.
(b) Verify this result using spherical polar coordinates. Equation (3.156) provides ∇.

3.10.27 Find the spherical coordinate components of the velocity and acceleration of a moving particle:

$$v_r = \dot{r}, \qquad a_r = \ddot{r} - r\dot{\theta}^2 - r\sin^2\theta\,\dot{\varphi}^2,$$

$$v_\theta = r\dot{\theta}, \qquad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\sin\theta\cos\theta\,\dot{\varphi}^2,$$

$$v_\varphi = r\sin\theta\,\dot{\varphi}, \qquad a_\varphi = r\sin\theta\,\ddot{\varphi} + 2\dot{r}\sin\theta\,\dot{\varphi} + 2r\cos\theta\,\dot{\theta}\dot{\varphi}.$$

Hint. $\mathbf{r}(t) = \hat{\mathbf{e}}_r(t)r(t)$

$$= [\hat{\mathbf{e}}_x \sin\theta(t)\cos\varphi(t) + \hat{\mathbf{e}}_y \sin\theta(t)\sin\varphi(t) + \hat{\mathbf{e}}_z \cos\theta(t)]r(t).$$

Note. The dot in $\dot{r}, \dot{\theta}, \dot{\varphi}$ means time derivative: $\dot{r} = dr/dt, \dot{\theta} = d\theta/dt$, $\dot{\varphi} = d\varphi/dt$.

3.10.28 Express $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$ in spherical polar coordinates.

ANS.
$$\frac{\partial}{\partial x} = \sin\theta\cos\varphi\frac{\partial}{\partial r} + \cos\theta\cos\varphi\frac{1}{r}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi},$$
$$\frac{\partial}{\partial y} = \sin\theta\sin\varphi\frac{\partial}{\partial r} + \cos\theta\sin\varphi\frac{1}{r}\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi},$$
$$\frac{\partial}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \sin\theta\frac{1}{r}\frac{\partial}{\partial\theta}.$$

Hint. Equate ∇_{xyz} and $\nabla_{r\theta\varphi}$.

3.10.29 Using results from Exercise 3.10.28, show that

$$-i\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) = -i\frac{\partial}{\partial\varphi}.$$

This is the quantum mechanical operator corresponding to the z-component of orbital angular momentum.

3.10.30 With the quantum mechanical orbital angular momentum operator defined as $\mathbf{L} = -i(\mathbf{r}\times\nabla)$, show that

(a) $L_x + iL_y = e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\frac{\partial}{\partial\varphi}\right)$,

(b) $L_x - iL_y = -e^{-i\varphi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\frac{\partial}{\partial\varphi}\right)$.

3.10.31 Verify that $\mathbf{L}\times\mathbf{L} = i\mathbf{L}$ in spherical polar coordinates. $\mathbf{L} = -i(\mathbf{r}\times\nabla)$, the quantum mechanical orbital angular momentum operator.
Written in component form, this relation is

$$L_yL_z - L_zL_y = iL_x, \quad L_zL_x - L_xL_z = -L_y, \quad L_xL_y - L_yL_x = iL_z.$$

Using the commutator notation, $[A, B] = AB - BA$, and the definition of the Levi-Civita symbol ε_{ijk}, the above can also be written

$$[L_i, L_j] = i\,\varepsilon_{ijk}\,L_k,$$

where i, j, k are x, y, z in any order.
Hint. Use spherical polar coordinates for $\mathbf{L}$ but Cartesian components for the cross product.

3.10.32 (a) Using Eq. (3.156) show that

$$\mathbf{L} = -i(\mathbf{r} \times \nabla) = i\left(\hat{\mathbf{e}}_\theta \frac{1}{\sin\theta}\frac{\partial}{\partial\varphi} - \hat{\mathbf{e}}_\varphi \frac{\partial}{\partial\theta}\right).$$

(b) Resolving $\hat{\mathbf{e}}_\theta$ and $\hat{\mathbf{e}}_\varphi$ into Cartesian components, determine L_x, L_y, and L_z in terms of θ, φ, and their derivatives.

(c) From $L^2 = L_x^2 + L_y^2 + L_z^2$ show that

$$\begin{aligned}\mathbf{L}^2 &= -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\\ &= -r^2\nabla^2 + \frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right).\end{aligned}$$

3.10.33 With $\mathbf{L} = -i\mathbf{r} \times \nabla$, verify the operator identities

(a) $\nabla = \hat{\mathbf{e}}_r \dfrac{\partial}{\partial r} - i\dfrac{\mathbf{r} \times \mathbf{L}}{r^2}$,

(b) $\mathbf{r}\nabla^2 - \nabla\left(1 + r\dfrac{\partial}{\partial r}\right) = i\nabla \times \mathbf{L}$.

3.10.34 Show that the following three forms (spherical coordinates) of $\nabla^2\psi(r)$ are equivalent:

(a) $\dfrac{1}{r^2}\dfrac{d}{dr}\left[r^2\dfrac{d\psi(r)}{dr}\right]$; (b) $\dfrac{1}{r}\dfrac{d^2}{dr^2}[r\psi(r)]$; (c) $\dfrac{d^2\psi(r)}{dr^2} + \dfrac{2}{r}\dfrac{d\psi(r)}{dr}$.

The second form is particularly convenient in establishing a correspondence between spherical polar and Cartesian descriptions of a problem.

3.10.35 A certain force field is given in spherical polar coordinates by

$$\mathbf{F} = \hat{\mathbf{e}}_r \frac{2P\cos\theta}{r^3} + \hat{\mathbf{e}}_\theta \frac{P}{r^3}\sin\theta, \quad r \geq P/2.$$

(a) Examine $\nabla \times \mathbf{F}$ to see if a potential exists.

(b) Calculate $\oint \mathbf{F} \cdot d\mathbf{r}$ for a unit circle in the plane $\theta = \pi/2$. What does this indicate about the force being conservative or nonconservative?

(c) If you believe that $\mathbf{F}$ may be described by $\mathbf{F} = -\nabla\psi$, find ψ. Otherwise simply state that no acceptable potential exists.

3.10.36 (a) Show that $\mathbf{A} = -\hat{\mathbf{e}}_\varphi \cot\theta/r$ is a solution of $\nabla \times \mathbf{A} = \hat{\mathbf{e}}_r/r^2$.

(b) Show that this spherical polar coordinate solution agrees with the solution given for Exercise 3.9.5:

$$\mathbf{A} = \hat{\mathbf{e}}_x \frac{yz}{r(x^2+y^2)} - \hat{\mathbf{e}}_y \frac{xz}{r(x^2+y^2)}.$$

Note that the solution diverges for $\theta = 0, \pi$ corresponding to $x, y = 0$.

(c) Finally, show that $\mathbf{A} = -\hat{\mathbf{e}}_\theta \varphi \sin\theta/r$ is a solution. Note that although this solution does not diverge ($r \neq 0$), it is no longer single-valued for all possible azimuth angles.

3.10.37 An electric dipole of moment **p** is located at the origin. The dipole creates an electric potential at **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 **r**.

Additional Readings

Borisenko, A. I., and I. E. Tarpov, *Vector and Tensor Analysis with Applications*. Englewood Cliffs, NJ: Prentice-Hall (1968), reprinting, Dover (1980).

Davis, H. F., and A. D. Snider, *Introduction to Vector Analysis*, 7th ed. Boston: Allyn & Bacon (1995).

Kellogg, O. D., *Foundations of Potential Theory*. Berlin: Springer (1929), reprinted, Dover (1953). The classic text on potential theory.

Lewis, P. E., and J. P. Ward, *Vector Analysis for Engineers and Scientists*. Reading, MA: Addison-Wesley (1989).

Margenau, H., and G. M. Murphy, *The Mathematics of Physics and Chemistry*, 2nd ed. Princeton NJ: Van Nostrand (1956). Chapter 5 covers curvilinear coordinates and 13 specific coordinate systems.

Marion, J. B., *Principles of Vector Analysis*. New York: Academic Press (1965). A moderately advanced presentation of vector analysis oriented toward tensor analysis. Rotations and other transformations are described with the appropriate matrices.

Morse, P. M., and H. Feshbach, *Methods of Theoretical Physics*. New York: McGraw-Hill (1953). Chapter 5 includes a description of several different coordinate systems. Note that Morse and Feshbach are not above using left-handed coordinate systems even for Cartesian coordinates. Elsewhere in this excellent (and difficult) book there are many examples of the use of the various coordinate systems in solving physical problems. Eleven additional fascinating but seldom-encountered orthogonal coordinate systems are discussed in the second (1970) edition of *Mathematical Methods for Physicists*.

Spiegel, M. R., *Vector Analysis*. New York: McGraw-Hill (1989).

Tai, C.-T., *Generalized Vector and Dyadic Analysis*. Oxford: Oxford University Press (1966).

Wrede, R. C., *Introduction to Vector and Tensor Analysis*. New York: Wiley (1963), reprinting, Dover (1972). Fine historical introduction. Excellent discussion of differentiation of vectors and applications to mechanics.