CHAPTER

1 Vector Analysis

1.1 ■ VECTOR ALGEBRA

1.1.1 ■ Vector Operations

If you walk 4 miles due north and then 3 miles due east (Fig. 1.1), you will have gone a total of 7 miles, but you're *not* 7 miles from where you set out—you're only 5. We need an arithmetic to describe quantities like this, which evidently do not add in the ordinary way. The reason they don't, of course, is that **displacements** (straight line segments going from one point to another) have *direction* as well as *magnitude* (length), and it is essential to take both into account when you combine them. Such objects are called **vectors**: velocity, acceleration, force and momentum are other examples. By contrast, quantities that have magnitude but no direction are called **scalars**: examples include mass, charge, density, and temperature.

I shall use **boldface** (**A**, **B**, and so on) for vectors and ordinary type for scalars. The magnitude of a vector **A** is written $|\mathbf{A}|$ or, more simply, A. In diagrams, vectors are denoted by arrows: the length of the arrow is proportional to the magnitude of the vector, and the arrowhead indicates its direction. *Minus* **A** ($-\mathbf{A}$) is a vector with the same magnitude as **A** but of opposite direction (Fig. 1.2). Note that vectors have magnitude and direction but *not location:* a displacement of 4 miles due north from Washington is represented by the same vector as a displacement 4 miles north from Baltimore (neglecting, of course, the curvature of the earth). On a diagram, therefore, you can slide the arrow around at will, as long as you don't change its length or direction.

We define four vector operations: addition and three kinds of multiplication.

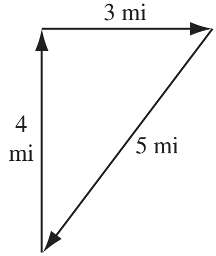


FIGURE 1.1

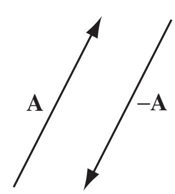


FIGURE 1.2

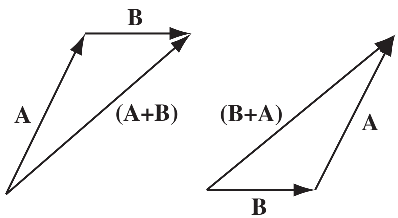


FIGURE 1.3

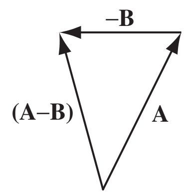


FIGURE 1.4

(i) Addition of two vectors. Place the tail of **B** at the head of **A**; the sum, $\mathbf{A}+\mathbf{B}$, is the vector from the tail of **A** to the head of **B** (Fig. 1.3). (This rule generalizes the obvious procedure for combining two displacements.) Addition is *commutative:*

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A};$$

3 miles east followed by 4 miles north gets you to the same place as 4 miles north followed by 3 miles east. Addition is also *associative:*

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

To subtract a vector, add its opposite (Fig. 1.4):

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

(ii) Multiplication by a scalar. Multiplication of a vector by a positive scalar a multiplies the *magnitude* but leaves the direction unchanged (Fig. 1.5). (If a is negative, the direction is reversed.) Scalar multiplication is *distributive:*

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

(iii) Dot product of two vectors. The dot product of two vectors is defined by

$$\mathbf{A}\cdot\mathbf{B}\equiv AB\cos\theta, \tag{1.1}$$

where θ is the angle they form when placed tail-to-tail (Fig. 1.6). Note that $\mathbf{A}\cdot\mathbf{B}$ is itself a *scalar* (hence the alternative name **scalar product**). The dot product is *commutative,*

$$\mathbf{A}\cdot\mathbf{B}=\mathbf{B}\cdot\mathbf{A},$$

and *distributive,*

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

Geometrically, $\mathbf{A}\cdot\mathbf{B}$ is the product of A times the projection of **B** along **A** (or the product of B times the projection of **A** along **B**). If the two vectors are parallel, then $\mathbf{A}\cdot\mathbf{B}=AB$. In particular, for any vector **A**,

$$\mathbf{A}\cdot\mathbf{A}=A^2. \tag{1.3}$$

If **A** and **B** are perpendicular, then $\mathbf{A}\cdot\mathbf{B}=0$.

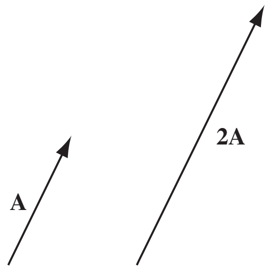


FIGURE 1.5

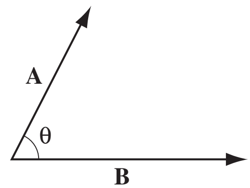


FIGURE 1.6

Example 1.1. Let **C** = **A** − **B** (Fig. 1.7), and calculate the dot product of **C** with itself.

Solution

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

or

$$C^2 = A^2 + B^2 - 2AB\cos\theta.$$

This is the **law of cosines**.

(iv) Cross product of two vectors. The cross product of two vectors is defined by

$$\mathbf{A}\times\mathbf{B} \equiv AB\sin\theta\,\hat{\mathbf{n}}, \tag{1.4}$$

where $\hat{\mathbf{n}}$ is a **unit vector** (vector of magnitude 1) pointing perpendicular to the plane of **A** and **B**. (I shall use a hat (ˆ) to denote unit vectors.) Of course, there are *two* directions perpendicular to any plane: "in" and "out." The ambiguity is resolved by the **right-hand rule**: let your fingers point in the direction of the first vector and curl around (via the smaller angle) toward the second; then your thumb indicates the direction of $\hat{\mathbf{n}}$. (In Fig. 1.8, $\mathbf{A}\times\mathbf{B}$ points *into* the page; $\mathbf{B}\times\mathbf{A}$ points *out* of the page.) Note that $\mathbf{A}\times\mathbf{B}$ is itself a *vector* (hence the alternative name **vector product**). The cross product is *distributive,*

$$\mathbf{A}\times(\mathbf{B}+\mathbf{C}) = (\mathbf{A}\times\mathbf{B}) + (\mathbf{A}\times\mathbf{C}), \tag{1.5}$$

but *not commutative.* In fact,

$$(\mathbf{B}\times\mathbf{A}) = -(\mathbf{A}\times\mathbf{B}). \tag{1.6}$$

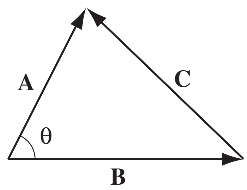


FIGURE 1.7

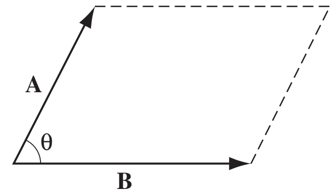


FIGURE 1.8

Geometrically, $|\mathbf{A} \times \mathbf{B}|$ is the area of the parallelogram generated by **A** and **B** (Fig. 1.8). If two vectors are parallel, their cross product is zero. In particular,

$$\mathbf{A} \times \mathbf{A} = \mathbf{0}$$

for any vector **A**. (Here **0** is the **zero vector**, with magnitude 0.)

Problem 1.1 Using the definitions in Eqs. 1.1 and 1.4, and appropriate diagrams, show that the dot product and cross product are distributive,

a) when the three vectors are coplanar;

! b) in the general case.

Problem 1.2 Is the cross product associative?

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} \stackrel{?}{=} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}).$$

If so, *prove* it; if not, provide a counterexample (the simpler the better).

1.1.2 ■ Vector Algebra: Component Form

In the previous section, I defined the four vector operations (addition, scalar multiplication, dot product, and cross product) in "abstract" form—that is, without reference to any particular coordinate system. In practice, it is often easier to set up Cartesian coordinates x, y, z and work with vector **components**. Let $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ be unit vectors parallel to the x, y, and z axes, respectively (Fig. 1.9(a)). An arbitrary vector **A** can be expanded in terms of these **basis vectors** (Fig. 1.9(b)):

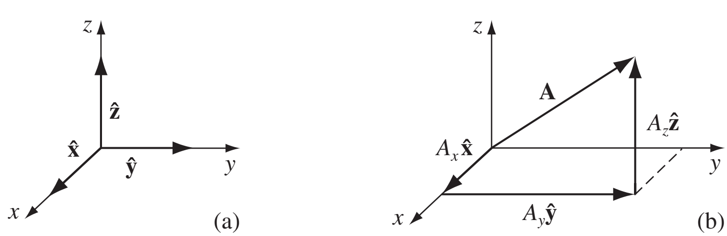


FIGURE 1.9

$$\mathbf{A} = A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}}.$$

The numbers A_x, A_y, and A_z, are the "components" of $\mathbf{A}$; geometrically, they are the projections of $\mathbf{A}$ along the three coordinate axes ($A_x = \mathbf{A} \cdot \hat{\mathbf{x}}$, $A_y = \mathbf{A} \cdot \hat{\mathbf{y}}$, $A_z = \mathbf{A} \cdot \hat{\mathbf{z}}$). We can now reformulate each of the four vector operations as a rule for manipulating components:

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

Rule (i): *To add vectors, add like components.*

$$a\mathbf{A} = (aA_x)\hat{\mathbf{x}} + (aA_y)\hat{\mathbf{y}} + (aA_z)\hat{\mathbf{z}}. \tag{1.8}$$

Rule (ii): *To multiply by a scalar, multiply each component.*

Because $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are mutually perpendicular unit vectors,

$$\hat{\mathbf{x}} \cdot \hat{\mathbf{x}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} = 1; \quad \hat{\mathbf{x}} \cdot \hat{\mathbf{y}} = \hat{\mathbf{x}} \cdot \hat{\mathbf{z}} = \hat{\mathbf{y}} \cdot \hat{\mathbf{z}} = 0. \tag{1.9}$$

Accordingly,

$$\begin{aligned}\mathbf{A} \cdot \mathbf{B} &= (A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}}) \cdot (B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}) \\ &= A_xB_x + A_yB_y + A_zB_z.\end{aligned} \tag{1.10}$$

Rule (iii): *To calculate the dot product, multiply like components, and add.* In particular,

$$\mathbf{A} \cdot \mathbf{A} = A_x^2 + A_y^2 + A_z^2,$$

so

$$A = \sqrt{A_x^2 + A_y^2 + A_z^2}. \tag{1.11}$$

(This is, if you like, the three-dimensional generalization of the Pythagorean theorem.)

Similarly,[1]

$$\begin{aligned}\hat{\mathbf{x}} \times \hat{\mathbf{x}} &= \quad \hat{\mathbf{y}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}} \times \hat{\mathbf{z}} = \mathbf{0}, \\ \hat{\mathbf{x}} \times \hat{\mathbf{y}} &= -\hat{\mathbf{y}} \times \hat{\mathbf{x}} = \hat{\mathbf{z}}, \\ \hat{\mathbf{y}} \times \hat{\mathbf{z}} &= -\hat{\mathbf{z}} \times \hat{\mathbf{y}} = \hat{\mathbf{x}}, \\ \hat{\mathbf{z}} \times \hat{\mathbf{x}} &= -\hat{\mathbf{x}} \times \hat{\mathbf{z}} = \hat{\mathbf{y}}.\end{aligned} \tag{1.12}$$

[1]These signs pertain to a *right-handed* coordinate system (x-axis out of the page, y-axis to the right, z-axis up, or any rotated version thereof). In a *left-handed* system (z-axis down), the signs would be reversed: $\hat{\mathbf{x}} \times \hat{\mathbf{y}} = -\hat{\mathbf{z}}$, and so on. We shall use right-handed systems exclusively.

Therefore,

$$\mathbf{A} \times \mathbf{B} = (A_x\hat{\mathbf{x}} + A_y\hat{\mathbf{y}} + A_z\hat{\mathbf{z}}) \times (B_x\hat{\mathbf{x}} + B_y\hat{\mathbf{y}} + B_z\hat{\mathbf{z}}) \tag{1.13}$$
$$= (A_yB_z - A_zB_y)\hat{\mathbf{x}} + (A_zB_x - A_xB_z)\hat{\mathbf{y}} + (A_xB_y - A_yB_x)\hat{\mathbf{z}}.$$

This cumbersome expression can be written more neatly as a determinant:

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \tag{1.14}$$

Rule (iv): *To calculate the cross product, form the determinant whose first row is* $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$, *whose second row is* **A** *(in component form), and whose third row is* **B**.

Example 1.2. Find the angle between the face diagonals of a cube.

Solution

We might as well use a cube of side 1, and place it as shown in Fig. 1.10, with one corner at the origin. The face diagonals **A** and **B** are

$$\mathbf{A} = 1\,\hat{\mathbf{x}} + 0\,\hat{\mathbf{y}} + 1\,\hat{\mathbf{z}}; \qquad \mathbf{B} = 0\,\hat{\mathbf{x}} + 1\,\hat{\mathbf{y}} + 1\,\hat{\mathbf{z}}.$$

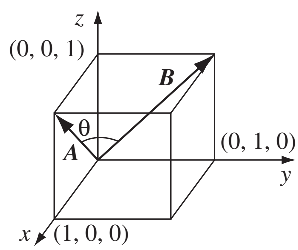


FIGURE 1.10

So, in component form,

$$\mathbf{A} \cdot \mathbf{B} = 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 = 1.$$

On the other hand, in "abstract" form,

$$\mathbf{A} \cdot \mathbf{B} = AB\cos\theta = \sqrt{2}\sqrt{2}\cos\theta = 2\cos\theta.$$

Therefore,

$$\cos\theta = 1/2, \quad \text{or} \quad \theta = 60^\circ.$$

Of course, you can get the answer more easily by drawing in a diagonal across the top of the cube, completing the equilateral triangle. But in cases where the geometry is not so simple, this device of comparing the abstract and component forms of the dot product can be a very efficient means of finding angles.

Problem 1.3 Find the angle between the body diagonals of a cube.

Problem 1.4 Use the cross product to find the components of the unit vector $\hat{\mathbf{n}}$ perpendicular to the shaded plane in Fig. 1.11.

1.1.3 ■ Triple Products

Since the cross product of two vectors is itself a vector, it can be dotted or crossed with a third vector to form a *triple* product.

(i) Scalar triple product: $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$. Geometrically, $|\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})|$ is the volume of the parallelepiped generated by **A**, **B**, and **C**, since $|\mathbf{B} \times \mathbf{C}|$ is the area of the base, and $|A \cos\theta|$ is the altitude (Fig. 1.12). Evidently,

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}), \tag{1.15}$$

for they all correspond to the same figure. Note that "alphabetical" order is preserved—in view of Eq. 1.6, the "nonalphabetical" triple products,

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

have the opposite sign. In component form,

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}. \tag{1.16}$$

Note that the dot and cross can be interchanged:

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

(this follows immediately from Eq. 1.15); however, the placement of the parentheses is critical: $(\mathbf{A} \cdot \mathbf{B}) \times \mathbf{C}$ is a meaningless expression—you can't make a cross product from a *scalar* and a vector.

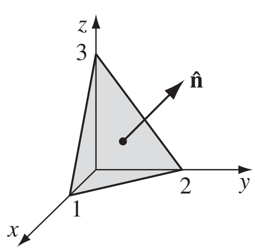


FIGURE 1.11

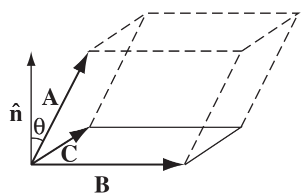


FIGURE 1.12

(ii) Vector triple product: $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$. The vector triple product can be simplified by the so-called **BAC-CAB** rule:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}). \tag{1.17}$$

Notice that

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

is an entirely different vector (cross-products are not associative). All *higher* vector products can be similarly reduced, often by repeated application of Eq. 1.17, so it is never necessary for an expression to contain more than one cross product in any term. For instance,

$$\begin{aligned}(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C});\\ \mathbf{A} \times [\mathbf{B} \times (\mathbf{C} \times \mathbf{D})] &= \mathbf{B}[\mathbf{A} \cdot (\mathbf{C} \times \mathbf{D})] - (\mathbf{A} \cdot \mathbf{B})(\mathbf{C} \times \mathbf{D}).\end{aligned} \tag{1.18}$$

Problem 1.5 Prove the **BAC-CAB** rule by writing out both sides in component form.

Problem 1.6 Prove that

$$[\mathbf{A} \times (\mathbf{B} \times \mathbf{C})] + [\mathbf{B} \times (\mathbf{C} \times \mathbf{A})] + [\mathbf{C} \times (\mathbf{A} \times \mathbf{B})] = \mathbf{0}.$$

Under what conditions does $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \times \mathbf{C}$?

1.1.4 ■ Position, Displacement, and Separation Vectors

The location of a point in three dimensions can be described by listing its Cartesian coordinates (x, y, z). The vector to that point from the origin ($\mathcal{O}$) is called the **position vector** (Fig. 1.13):

$$\mathbf{r} \equiv x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + z\,\hat{\mathbf{z}}. \tag{1.19}$$

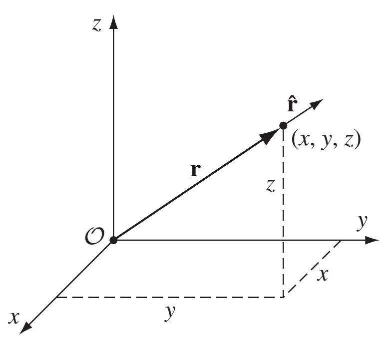


FIGURE 1.13

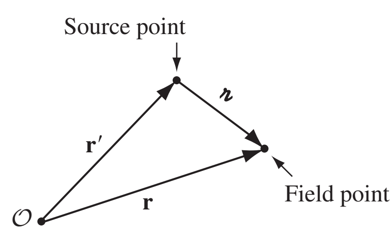


FIGURE 1.14

I will reserve the letter **r** for this purpose, throughout the book. Its magnitude,

$$r = \sqrt{x^2 + y^2 + z^2}, \tag{1.20}$$

is the distance from the origin, and

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} = \frac{x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + z\,\hat{\mathbf{z}}}{\sqrt{x^2 + y^2 + z^2}} \tag{1.21}$$

is a unit vector pointing radially outward. The **infinitesimal displacement vector**, from (x, y, z) to $(x + dx, y + dy, z + dz)$, is

$$d\mathbf{l} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}}. \tag{1.22}$$

(We could call this $d\mathbf{r}$, since that's what it *is*, but it is useful to have a special notation for infinitesimal displacements.)

In electrodynamics, one frequently encounters problems involving *two* points—typically, a **source point**, $\mathbf{r}'$, where an electric charge is located, and a **field point**, $\mathbf{r}$, at which you are calculating the electric or magnetic field (Fig. 1.14). It pays to adopt right from the start some short-hand notation for the **separation vector** from the source point to the field point. I shall use for this purpose the script letter $\boldsymbol{\mathscr{r}}$:

$$\boldsymbol{\mathscr{r}} \equiv \mathbf{r} - \mathbf{r}'. \tag{1.23}$$

Its magnitude is

$$\mathscr{r} = |\mathbf{r} - \mathbf{r}'|, \tag{1.24}$$

and a unit vector in the direction from $\mathbf{r}'$ to $\mathbf{r}$ is

$$\hat{\boldsymbol{\mathscr{r}}} = \frac{\boldsymbol{\mathscr{r}}}{\mathscr{r}} = \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}. \tag{1.25}$$

In Cartesian coordinates,

$$\boldsymbol{\mathscr{r}} = (x - x')\hat{\mathbf{x}} + (y - y')\hat{\mathbf{y}} + (z - z')\hat{\mathbf{z}}, \tag{1.26}$$

$$\mathscr{r} = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}, \tag{1.27}$$

$$\hat{\boldsymbol{\mathscr{r}}} = \frac{(x - x')\hat{\mathbf{x}} + (y - y')\hat{\mathbf{y}} + (z - z')\hat{\mathbf{z}}}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}} \tag{1.28}$$

(from which you can appreciate the economy of the script-$\mathscr{r}$ notation).

Problem 1.7 Find the separation vector $\boldsymbol{\mathscr{r}}$ from the source point (2,8,7) to the field point (4,6,8). Determine its magnitude ($\mathscr{r}$), and construct the unit vector $\hat{\boldsymbol{\mathscr{r}}}$.

1.1.5 ■ How Vectors Transform[2]

The definition of a vector as "a quantity with a magnitude and direction" is not altogether satisfactory: What precisely does "direction" *mean*? This may seem a pedantic question, but we shall soon encounter a species of derivative that *looks* rather like a vector, and we'll want to know for sure whether it *is* one.

You might be inclined to say that a vector is anything that has three components that combine properly under addition. Well, how about this: We have a barrel of fruit that contains N_x pears, N_y apples, and N_z bananas. Is $\mathbf{N} = N_x\hat{\mathbf{x}} + N_y\hat{\mathbf{y}} + N_z\hat{\mathbf{z}}$ a vector? It has three components, and when you add another barrel with M_x pears, M_y apples, and M_z bananas the result is $(N_x + M_x)$ pears, $(N_y + M_y)$ apples, $(N_z + M_z)$ bananas. So it does *add* like a vector. Yet it's obviously *not* a vector, in the physicist's sense of the word, because it doesn't really have a direction. What exactly is wrong with it?

The answer is that $\mathbf{N}$ *does not transform properly when you change coordinates.* The coordinate frame we use to describe positions in space is of course entirely arbitrary, but there is a specific geometrical transformation law for converting vector components from one frame to another. Suppose, for instance, the $\overline{x}, \overline{y}, \overline{z}$ system is rotated by angle ϕ, relative to x, y, z, about the common $x = \overline{x}$ axes. From Fig. 1.15,

$$A_y = A\cos\theta, \qquad A_z = A\sin\theta,$$

while

$$\begin{aligned}
\overline{A}_y &= A\cos\overline{\theta} = A\cos(\theta - \phi) = A(\cos\theta\cos\phi + \sin\theta\sin\phi)\\
&= \cos\phi A_y + \sin\phi A_z,\\
\overline{A}_z &= A\sin\overline{\theta} = A\sin(\theta - \phi) = A(\sin\theta\cos\phi - \cos\theta\sin\phi)\\
&= -\sin\phi A_y + \cos\phi A_z.
\end{aligned}$$

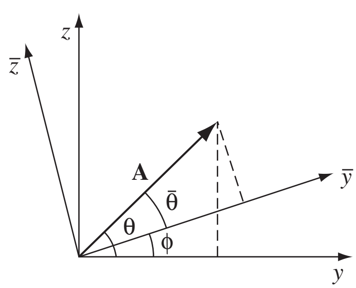


FIGURE 1.15

[2]This section can be skipped without loss of continuity.

We might express this conclusion in matrix notation:

$$\begin{pmatrix} \overline{A}_y \\ \overline{A}_z \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} A_y \\ A_z \end{pmatrix}. \tag{1.29}$$

More generally, for rotation about an *arbitrary* axis in three dimensions, the transformation law takes the form

$$\begin{pmatrix} \overline{A}_x \\ \overline{A}_y \\ \overline{A}_z \end{pmatrix} = \begin{pmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}, \tag{1.30}$$

or, more compactly,

$$\overline{A}_i = \sum_{j=1}^{3} R_{ij} A_j, \tag{1.31}$$

where the index 1 stands for x, 2 for y, and 3 for z. The elements of the matrix R can be ascertained, for a given rotation, by the same sort of trigonometric arguments as we used for a rotation about the x axis.

Now: *Do* the components of **N** transform in this way? Of *course* not—it doesn't matter what coordinates you use to represent positions in space; there are still just as many apples in the barrel. You can't convert a pear into a banana by choosing a different set of axes, but you *can* turn A_x into $\overline{A}_y$. Formally, then, a *vector is any set of three components that transforms in the same manner as a displacement when you change coordinates.* As always, displacement is the *model* for the behavior of all vectors.[3]

By the way, a (second-rank) **tensor** is a quantity with *nine* components, T_{xx}, T_{xy}, T_{xz}, $T_{yx}, \ldots, T_{zz}$, which transform with *two* factors of R:

$$\begin{aligned} \overline{T}_{xx} = {} & R_{xx}(R_{xx}T_{xx} + R_{xy}T_{xy} + R_{xz}T_{xz}) \\ & + R_{xy}(R_{xx}T_{yx} + R_{xy}T_{yy} + R_{xz}T_{yz}) \\ & + R_{xz}(R_{xx}T_{zx} + R_{xy}T_{zy} + R_{xz}T_{zz}), \ldots \end{aligned}$$

or, more compactly,

$$\overline{T}_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} R_{ik}R_{jl}T_{kl}. \tag{1.32}$$

[3] If you're a mathematician you might want to contemplate generalized vector spaces in which the "axes" have nothing to do with direction and the basis vectors are no longer $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ (indeed, there may be more than three dimensions). This is the subject of **linear algebra**. But for our purposes all vectors live in ordinary 3-space (or, in Chapter 12, in 4-dimensional space-time.)

In general, an nth-rank tensor has n indices and 3^n components, and transforms with n factors of R. In this hierarchy, a vector is a tensor of rank 1, and a scalar is a tensor of rank zero.[4]

Problem 1.8

(a) Prove that the two-dimensional rotation matrix (Eq. 1.29) preserves dot products. (That is, show that $\overline{A}_y \overline{B}_y + \overline{A}_z \overline{B}_z = A_y B_y + A_z B_z$.)

(b) What constraints must the elements (R_{ij}) of the three-dimensional rotation matrix (Eq. 1.30) satisfy, in order to preserve the length of **A** (for all vectors **A**)?

Problem 1.9 Find the transformation matrix R that describes a rotation by 120° about an axis from the origin through the point (1, 1, 1). The rotation is clockwise as you look down the axis toward the origin.

Problem 1.10

(a) How do the components of a vector[5] transform under a **translation** of coordinates ($\overline{x} = x$, $\overline{y} = y - a$, $\overline{z} = z$, Fig. 1.16a)?

(b) How do the components of a vector transform under an **inversion** of coordinates ($\overline{x} = -x$, $\overline{y} = -y$, $\overline{z} = -z$, Fig. 1.16b)?

(c) How do the components of a cross product (Eq. 1.13) transform under inversion? [The cross-product of two vectors is properly called a **pseudovector** because of this "anomalous" behavior.] Is the cross product of two pseudovectors a vector, or a pseudovector? Name two pseudovector quantities in classical mechanics.

(d) How does the scalar triple product of three vectors transform under inversions? (Such an object is called a **pseudoscalar.**)

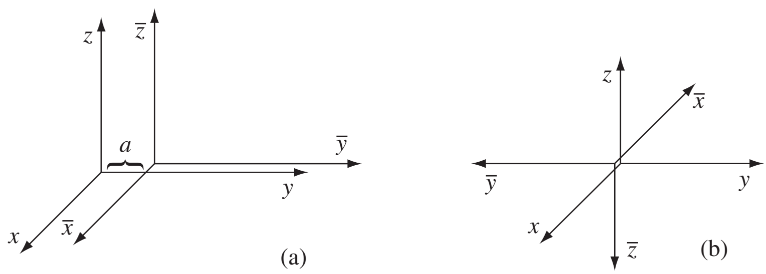


FIGURE 1.16

[4]A scalar does not change when you change coordinates. In particular, the components of a vector are *not* scalars, but the magnitude is.

[5]*Beware:* The vector **r** (Eq. 1.19) goes from a specific point in space (the origin, $\mathcal{O}$) to the point $P = (x, y, z)$. Under translations the *new* origin ($\bar{\mathcal{O}}$) is at a different location, and the arrow from $\bar{\mathcal{O}}$ to P is a completely different vector. The original vector **r** still goes from $\mathcal{O}$ to P, regardless of the coordinates used to label these points.

1.2 ■ DIFFERENTIAL CALCULUS

1.2.1 ■ "Ordinary" Derivatives

Suppose we have a function of one variable: $f(x)$. *Question:* What does the derivative, df/dx, do for us? *Answer:* It tells us how rapidly the function $f(x)$ varies when we change the argument x by a tiny amount, dx:

$$df = \left(\frac{df}{dx}\right) dx. \tag{1.33}$$

In words: If we increment x by an infinitesimal amount dx, then f changes by an amount df; the derivative is the proportionality factor. For example, in Fig. 1.17(a), the function varies slowly with x, and the derivative is correspondingly small. In Fig. 1.17(b), f increases rapidly with x, and the derivative is large, as you move away from $x = 0$.

Geometrical Interpretation: The derivative df/dx is the *slope* of the graph of f versus x.

1.2.2 ■ Gradient

Suppose, now, that we have a function of *three* variables—say, the temperature $T(x, y, z)$ in this room. (Start out in one corner, and set up a system of axes; then for each point (x, y, z) in the room, T gives the temperature at that spot.) We want to generalize the notion of "derivative" to functions like T, which depend not on *one* but on *three* variables.

A derivative is supposed to tell us how fast the function varies, if we move a little distance. But this time the situation is more complicated, because it depends on what *direction* we move: If we go straight up, then the temperature will probably increase fairly rapidly, but if we move horizontally, it may not change much at all. In fact, the question "How fast does T vary?" has an infinite number of answers, one for each direction we might choose to explore.

Fortunately, the problem is not as bad as it looks. A theorem on partial derivatives states that

$$dT = \left(\frac{\partial T}{\partial x}\right) dx + \left(\frac{\partial T}{\partial y}\right) dy + \left(\frac{\partial T}{\partial z}\right) dz. \tag{1.34}$$

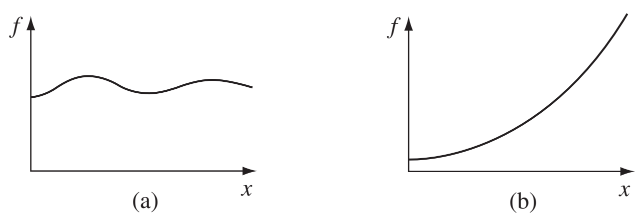


FIGURE 1.17

This tells us how T changes when we alter all three variables by the infinitesimal amounts dx, dy, dz. Notice that we do *not* require an infinite number of derivatives—*three* will suffice: the *partial* derivatives along each of the three coordinate directions.

Equation 1.34 is reminiscent of a dot product:

$$dT = \left(\frac{\partial T}{\partial x}\hat{\mathbf{x}} + \frac{\partial T}{\partial y}\hat{\mathbf{y}} + \frac{\partial T}{\partial z}\hat{\mathbf{z}}\right) \cdot (dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}})$$
$$= (\nabla T) \cdot (d\mathbf{l}), \tag{1.35}$$

where

$$\nabla T \equiv \frac{\partial T}{\partial x}\hat{\mathbf{x}} + \frac{\partial T}{\partial y}\hat{\mathbf{y}} + \frac{\partial T}{\partial z}\hat{\mathbf{z}} \tag{1.36}$$

is the **gradient** of T. Note that ∇T is a *vector* quantity, with three components; it is the generalized derivative we have been looking for. Equation 1.35 is the three-dimensional version of Eq. 1.33.

Geometrical Interpretation of the Gradient: Like any vector, the gradient has *magnitude* and *direction*. To determine its geometrical meaning, let's rewrite the dot product (Eq. 1.35) using Eq. 1.1:

$$dT = \nabla T \cdot d\mathbf{l} = |\nabla T||d\mathbf{l}|\cos\theta, \tag{1.37}$$

where θ is the angle between ∇T and $d\mathbf{l}$. Now, if we *fix* the *magnitude* $|d\mathbf{l}|$ and search around in various *directions* (that is, vary θ), the *maximum* change in T evidentally occurs when $\theta = 0$ (for then $\cos\theta = 1$). That is, for a fixed distance $|d\mathbf{l}|$, dT is greatest when I move in the *same direction* as ∇T. Thus:

> *The gradient ∇T points in the direction of maximum increase of the function T.*

Moreover:

> *The magnitude $|\nabla T|$ gives the slope (rate of increase) along this maximal direction.*

Imagine you are standing on a hillside. Look all around you, and find the direction of steepest ascent. That is the *direction* of the gradient. Now measure the *slope* in that direction (rise over run). That is the *magnitude* of the gradient. (Here the function we're talking about is the height of the hill, and the coordinates it depends on are positions—latitude and longitude, say. This function depends on only *two* variables, not *three*, but the geometrical meaning of the gradient is easier to grasp in two dimensions.) Notice from Eq. 1.37 that the direction of maximum *descent* is opposite to the direction of maximum *ascent*, while at right angles ($\theta = 90°$) the slope is zero (the gradient is perpendicular to the contour lines). You can conceive of surfaces that do not have these properties, but they always have "kinks" in them, and correspond to nondifferentiable functions.

What would it mean for the gradient to vanish? If $\nabla T = \mathbf{0}$ at (x, y, z), then $dT = 0$ for small displacements about the point (x, y, z). This is, then, a **stationary point** of the function $T(x, y, z)$. It could be a maximum (a summit), a minimum (a valley), a saddle point (a pass), or a "shoulder." This is analogous to the situation for functions of *one* variable, where a vanishing derivative signals a maximum, a minimum, or an inflection. In particular, if you want to locate the extrema of a function of three variables, set its gradient equal to zero.

Example 1.3. Find the gradient of $r = \sqrt{x^2 + y^2 + z^2}$ (the magnitude of the position vector).

Solution

$$\begin{aligned} \nabla r &= \frac{\partial r}{\partial x}\,\hat{\mathbf{x}} + \frac{\partial r}{\partial y}\,\hat{\mathbf{y}} + \frac{\partial r}{\partial z}\,\hat{\mathbf{z}} \\ &= \frac{1}{2}\frac{2x}{\sqrt{x^2+y^2+z^2}}\,\hat{\mathbf{x}} + \frac{1}{2}\frac{2y}{\sqrt{x^2+y^2+z^2}}\,\hat{\mathbf{y}} + \frac{1}{2}\frac{2z}{\sqrt{x^2+y^2+z^2}}\,\hat{\mathbf{z}} \\ &= \frac{x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + z\,\hat{\mathbf{z}}}{\sqrt{x^2+y^2+z^2}} = \frac{\mathbf{r}}{r} = \hat{\mathbf{r}}. \end{aligned}$$

Does this make sense? Well, it says that the distance from the origin increases most rapidly in the radial direction, and that its *rate* of increase in that direction is 1. . . just what you'd expect.

Problem 1.11 Find the gradients of the following functions:

(a) $f(x, y, z) = x^2 + y^3 + z^4$.

(b) $f(x, y, z) = x^2 y^3 z^4$.

(c) $f(x, y, z) = e^x \sin(y) \ln(z)$.

Problem 1.12 The height of a certain hill (in feet) is given by

$$h(x, y) = 10(2xy - 3x^2 - 4y^2 - 18x + 28y + 12),$$

where y is the distance (in miles) north, x the distance east of South Hadley.

(a) Where is the top of the hill located?

(b) How high is the hill?

(c) How steep is the slope (in feet per mile) at a point 1 mile north and one mile east of South Hadley? In what direction is the slope steepest, at that point?

• **Problem 1.13** Let $\boldsymbol{\mathscr{r}}$ be the separation vector from a fixed point (x', y', z') to the point (x, y, z), and let $\mathscr{r}$ be its length. Show that

(a) $\nabla(\mathscr{r}^2) = 2\boldsymbol{\mathscr{r}}$.

(b) $\nabla(1/\mathscr{r}) = -\hat{\boldsymbol{\mathscr{r}}}/\mathscr{r}^2$.

(c) What is the *general* formula for $\nabla(\mathscr{r}^n)$?

! **Problem 1.14** Suppose that f is a function of two variables (y and z) only. Show that the gradient $\nabla f = (\partial f/\partial y)\hat{\mathbf{y}} + (\partial f/\partial z)\hat{\mathbf{z}}$ transforms as a vector under rotations, Eq. 1.29. [*Hint:* $(\partial f/\partial \overline{y}) = (\partial f/\partial y)(\partial y/\partial \overline{y}) + (\partial f/\partial z)(\partial z/\partial \overline{y})$, and the analogous formula for $\partial f/\partial \overline{z}$. We know that $\overline{y} = y\cos\phi + z\sin\phi$ and $\overline{z} = -y\sin\phi + z\cos\phi$; "solve" these equations for y and z (as functions of $\overline{y}$ and $\overline{z}$), and compute the needed derivatives $\partial y/\partial \overline{y}$, $\partial z/\partial \overline{y}$, etc.]

1.2.3 ■ The Del Operator

The gradient has the formal appearance of a vector, ∇, "multiplying" a scalar T:

$$\nabla T = \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right)T. \tag{1.38}$$

(For once, I write the unit vectors to the *left,* just so no one will think this means $\partial\hat{\mathbf{x}}/\partial x$, and so on—which would be zero, since $\hat{\mathbf{x}}$ is constant.) The term in parentheses is called **del**:

$$\boxed{\nabla = \hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}.} \tag{1.39}$$

Of course, del is *not* a vector, in the usual sense. Indeed, it doesn't mean much until we provide it with a function to act upon. Furthermore, it does not "multiply" T; rather, it is an instruction to *differentiate* what follows. To be precise, then, we say that ∇ is a **vector operator** that *acts upon* T, not a vector that multiplies T.

With this qualification, though, ∇ mimics the behavior of an ordinary vector in virtually every way; almost anything that can be done with other vectors can also be done with ∇, if we merely translate "multiply" by "act upon." So by all means take the vector appearance of ∇ seriously: it is a marvelous piece of notational simplification, as you will appreciate if you ever consult Maxwell's original work on electromagnetism, written without the benefit of ∇.

Now, an ordinary vector **A** can multiply in three ways:

1. By a scalar a : $\mathbf{A}a$;
2. By a vector **B**, via the dot product: $\mathbf{A} \cdot \mathbf{B}$;
3. By a vector **B** via the cross product: $\mathbf{A} \times \mathbf{B}$.

Correspondingly, there are three ways the operator ∇ can act:

1. On a scalar function T : ∇T (the gradient);
2. On a vector function **v**, via the dot product: $\nabla \cdot \mathbf{v}$ (the **divergence**);
3. On a vector function **v**, via the cross product: $\nabla \times \mathbf{v}$ (the **curl**).

We have already discussed the gradient. In the following sections we examine the other two vector derivatives: divergence and curl.

1.2.4 ■ The Divergence

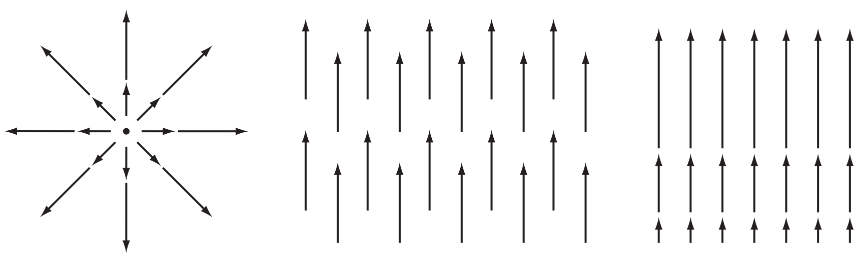
From the definition of ∇ we construct the divergence:

$$\nabla \cdot \mathbf{v} = \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right) \cdot (v_x\hat{\mathbf{x}} + v_y\hat{\mathbf{y}} + v_z\hat{\mathbf{z}})$$
$$= \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}. \tag{1.40}$$

Observe that the divergence of a vector function[6] $\mathbf{v}$ is itself a *scalar* $\nabla \cdot \mathbf{v}$.

Geometrical Interpretation: The name **divergence** is well chosen, for $\nabla \cdot \mathbf{v}$ is a measure of how much the vector $\mathbf{v}$ spreads out (diverges) from the point in question. For example, the vector function in Fig. 1.18a has a large (positive) divergence (if the arrows pointed *in*, it would be a *negative* divergence), the function in Fig. 1.18b has zero divergence, and the function in Fig. 1.18c again has a positive divergence. (Please understand that $\mathbf{v}$ here is a *function*—there's a different vector associated with every point in space. In the diagrams, of course, I can only draw the arrows at a few representative locations.)

Imagine standing at the edge of a pond. Sprinkle some sawdust or pine needles on the surface. If the material spreads out, then you dropped it at a point of positive divergence; if it collects together, you dropped it at a point of negative divergence. (The vector function $\mathbf{v}$ in this model is the velocity of the water at the surface—this is a *two*-dimensional example, but it helps give one a "feel" for what the divergence means. A point of positive divergence is a source, or "faucet"; a point of negative divergence is a sink, or "drain.")



(a) (b) (c)

FIGURE 1.18

[6]A vector function $\mathbf{v}(x, y, z) = v_x(x, y, z)\,\hat{\mathbf{x}} + v_y(x, y, z)\,\hat{\mathbf{y}} + v_z(x, y, z)\,\hat{\mathbf{z}}$ is really *three* functions—one for each component. There's no such thing as the divergence of a scalar.

Example 1.4. Suppose the functions in Fig. 1.18 are $\mathbf{v}_a = \mathbf{r} = x\,\hat{\mathbf{x}} + y\,\hat{\mathbf{y}} + z\,\hat{\mathbf{z}}$, $\mathbf{v}_b = \hat{\mathbf{z}}$, and $\mathbf{v}_c = z\,\hat{\mathbf{z}}$. Calculate their divergences.

Solution

$$\nabla \cdot \mathbf{v}_a = \frac{\partial}{\partial x}(x) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(z) = 1 + 1 + 1 = 3.$$

As anticipated, this function has a positive divergence.

$$\nabla \cdot \mathbf{v}_b = \frac{\partial}{\partial x}(0) + \frac{\partial}{\partial y}(0) + \frac{\partial}{\partial z}(1) = 0 + 0 + 0 = 0,$$

as expected.

$$\nabla \cdot \mathbf{v}_c = \frac{\partial}{\partial x}(0) + \frac{\partial}{\partial y}(0) + \frac{\partial}{\partial z}(z) = 0 + 0 + 1 = 1.$$

Problem 1.15 Calculate the divergence of the following vector functions:

(a) $\mathbf{v}_a = x^2\,\hat{\mathbf{x}} + 3xz^2\,\hat{\mathbf{y}} - 2xz\,\hat{\mathbf{z}}$.

(b) $\mathbf{v}_b = xy\,\hat{\mathbf{x}} + 2yz\,\hat{\mathbf{y}} + 3zx\,\hat{\mathbf{z}}$.

(c) $\mathbf{v}_c = y^2\,\hat{\mathbf{x}} + (2xy + z^2)\,\hat{\mathbf{y}} + 2yz\,\hat{\mathbf{z}}$.

• **Problem 1.16** Sketch the vector function

$$\mathbf{v} = \frac{\hat{\mathbf{r}}}{r^2},$$

and compute its divergence. The answer may surprise you... can you explain it?

! **Problem 1.17** In two dimensions, show that the divergence transforms as a scalar under rotations. [*Hint:* Use Eq. 1.29 to determine $\bar{v}_y$ and $\bar{v}_z$, and the method of Prob. 1.14 to calculate the derivatives. Your aim is to show that $\partial \bar{v}_y/\partial \bar{y} + \partial \bar{v}_z/\partial \bar{z} = \partial v_y/\partial y + \partial v_z/\partial z$.]

1.2.5 ■ The Curl

From the definition of ∇ we construct the curl:

$$\nabla \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ v_x & v_y & v_z \end{vmatrix}$$

$$= \hat{\mathbf{x}}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right) + \hat{\mathbf{y}}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) + \hat{\mathbf{z}}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right). \quad (1.41)$$

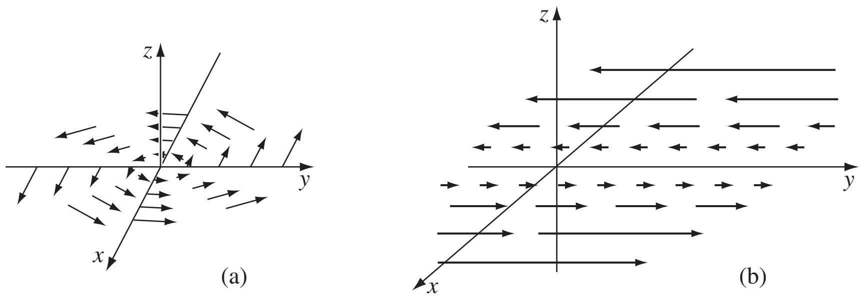


FIGURE 1.19

Notice that the curl of a vector function[7] **v** is, like any cross product, a *vector*.

Geometrical Interpretation: The name **curl** is also well chosen, for $\nabla \times \mathbf{v}$ is a measure of how much the vector **v** swirls around the point in question. Thus the three functions in Fig. 1.18 all have zero curl (as you can easily check for yourself), whereas the functions in Fig. 1.19 have a substantial curl, pointing in the z direction, as the natural right-hand rule would suggest. Imagine (again) you are standing at the edge of a pond. Float a small paddlewheel (a cork with toothpicks pointing out radially would do); if it starts to rotate, then you placed it at a point of nonzero *curl.* A whirlpool would be a region of large curl.

Example 1.5. Suppose the function sketched in Fig. 1.19a is $\mathbf{v}_a = -y\hat{\mathbf{x}} + x\hat{\mathbf{y}}$, and that in Fig. 1.19b is $\mathbf{v}_b = x\hat{\mathbf{y}}$. Calculate their curls.

Solution

$$\nabla \times \mathbf{v}_a = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ -y & x & 0 \end{vmatrix} = 2\hat{\mathbf{z}},$$

and

$$\nabla \times \mathbf{v}_b = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ 0 & x & 0 \end{vmatrix} = \hat{\mathbf{z}}.$$

As expected, these curls point in the $+z$ direction. (Incidentally, they both have zero divergence, as you might guess from the pictures: nothing is "spreading out"... it just "swirls around.")

---

[7] There's no such thing as the curl of a scalar.

Problem 1.18 Calculate the curls of the vector functions in Prob. 1.15.

Problem 1.19 Draw a circle in the xy plane. At a few representative points draw the vector $\mathbf{v}$ tangent to the circle, pointing in the clockwise direction. By comparing adjacent vectors, determine the *sign* of $\partial v_x/\partial y$ and $\partial v_y/\partial x$. According to Eq. 1.41, then, what is the direction of $\nabla \times \mathbf{v}$? Explain how this example illustrates the geometrical interpretation of the curl.

Problem 1.20 Construct a vector function that has zero divergence and zero curl everywhere. (A *constant* will do the job, of course, but make it something a little more interesting than that!)

1.2.6 ■ Product Rules

The calculation of ordinary derivatives is facilitated by a number of rules, such as the sum rule:

$$\frac{d}{dx}(f+g) = \frac{df}{dx} + \frac{dg}{dx},$$

the rule for multiplying by a constant:

$$\frac{d}{dx}(kf) = k\frac{df}{dx},$$

the product rule:

$$\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx},$$

and the quotient rule:

$$\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{g\dfrac{df}{dx} - f\dfrac{dg}{dx}}{g^2}.$$

Similar relations hold for the vector derivatives. Thus,

$$\boldsymbol{\nabla}(f+g) = \boldsymbol{\nabla} f + \boldsymbol{\nabla} g, \qquad \boldsymbol{\nabla}\cdot(\mathbf{A}+\mathbf{B}) = (\boldsymbol{\nabla}\cdot\mathbf{A}) + (\boldsymbol{\nabla}\cdot\mathbf{B}),$$

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

and

$$\boldsymbol{\nabla}(kf) = k\boldsymbol{\nabla} f, \qquad \boldsymbol{\nabla}\cdot(k\mathbf{A}) = k(\boldsymbol{\nabla}\cdot\mathbf{A}), \qquad \boldsymbol{\nabla}\times(k\mathbf{A}) = k(\boldsymbol{\nabla}\times\mathbf{A}),$$

as you can check for yourself. The product rules are not quite so simple. There are two ways to construct a scalar as the product of two functions:

$$fg \quad \text{(product of two scalar functions)},$$
$$\mathbf{A}\cdot\mathbf{B} \quad \text{(dot product of two vector functions)},$$

and two ways to make a vector:

$$
\begin{aligned}
f\mathbf{A} &\quad \text{(scalar times vector)},\\
\mathbf{A}\times\mathbf{B} &\quad \text{(cross product of two vectors)}.
\end{aligned}
$$

Accordingly, there are *six* product rules, two for gradients:

(i) $$\nabla(fg) = f\nabla g + g\nabla f,$$

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

two for divergences:

(iii) $$\nabla\cdot(f\mathbf{A}) = f(\nabla\cdot\mathbf{A}) + \mathbf{A}\cdot(\nabla f),$$

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

and two for curls:

(v) $$\nabla\times(f\mathbf{A}) = f(\nabla\times\mathbf{A}) - \mathbf{A}\times(\nabla f),$$

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

You will be using these product rules so frequently that I have put them inside the front cover for easy reference. The proofs come straight from the product rule for ordinary derivatives. For instance,

$$
\begin{aligned}
\nabla\cdot(f\mathbf{A}) &= \frac{\partial}{\partial x}(fA_x) + \frac{\partial}{\partial y}(fA_y) + \frac{\partial}{\partial z}(fA_z)\\
&= \left(\frac{\partial f}{\partial x}A_x + f\frac{\partial A_x}{\partial x}\right) + \left(\frac{\partial f}{\partial y}A_y + f\frac{\partial A_y}{\partial y}\right) + \left(\frac{\partial f}{\partial z}A_z + f\frac{\partial A_z}{\partial z}\right)\\
&= (\nabla f)\cdot\mathbf{A} + f(\nabla\cdot\mathbf{A}).
\end{aligned}
$$

It is also possible to formulate three quotient rules:

$$
\begin{aligned}
\nabla\left(\frac{f}{g}\right) &= \frac{g\nabla f - f\nabla g}{g^2},\\
\nabla\cdot\left(\frac{\mathbf{A}}{g}\right) &= \frac{g(\nabla\cdot\mathbf{A}) - \mathbf{A}\cdot(\nabla g)}{g^2},\\
\nabla\times\left(\frac{\mathbf{A}}{g}\right) &= \frac{g(\nabla\times\mathbf{A}) + \mathbf{A}\times(\nabla g)}{g^2}.
\end{aligned}
$$

However, since these can be obtained quickly from the corresponding product rules, there is no point in listing them separately.

Problem 1.21 Prove product rules (i), (iv), and (v).

Problem 1.22

(a) If $\mathbf{A}$ and $\mathbf{B}$ are two vector functions, what does the expression $(\mathbf{A} \cdot \boldsymbol{\nabla})\mathbf{B}$ mean? (That is, what are its x, y, and z components, in terms of the Cartesian components of $\mathbf{A}$, $\mathbf{B}$, and $\boldsymbol{\nabla}$?)

(b) Compute $(\hat{\mathbf{r}} \cdot \boldsymbol{\nabla})\hat{\mathbf{r}}$, where $\hat{\mathbf{r}}$ is the unit vector defined in Eq. 1.21.

(c) For the functions in Prob. 1.15, evaluate $(\mathbf{v}_a \cdot \boldsymbol{\nabla})\mathbf{v}_b$.

Problem 1.23 (For masochists only.) Prove product rules (ii) and (vi). Refer to Prob. 1.22 for the definition of $(\mathbf{A} \cdot \boldsymbol{\nabla})\mathbf{B}$.

Problem 1.24 Derive the three quotient rules.

Problem 1.25

(a) Check product rule (iv) (by calculating each term separately) for the functions

$$\mathbf{A} = x\,\hat{\mathbf{x}} + 2y\,\hat{\mathbf{y}} + 3z\,\hat{\mathbf{z}}; \qquad \mathbf{B} = 3y\,\hat{\mathbf{x}} - 2x\,\hat{\mathbf{y}}.$$

(b) Do the same for product rule (ii).

(c) Do the same for rule (vi).

1.2.7 ■ Second Derivatives

The gradient, the divergence, and the curl are the only first derivatives we can make with $\boldsymbol{\nabla}$; by applying $\boldsymbol{\nabla}$ *twice*, we can construct five species of *second* derivatives. The gradient $\boldsymbol{\nabla} T$ is a *vector*, so we can take the *divergence* and *curl* of it:

(1) Divergence of gradient: $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} T)$.

(2) Curl of gradient: $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} T)$.

The divergence $\boldsymbol{\nabla} \cdot \mathbf{v}$ is a *scalar*—all we can do is take its *gradient:*

(3) Gradient of divergence: $\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v})$.

The curl $\boldsymbol{\nabla} \times \mathbf{v}$ is a *vector*, so we can take its *divergence* and *curl:*

(4) Divergence of curl: $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v})$.

(5) Curl of curl: $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v})$.

This exhausts the possibilities, and in fact not all of them give anything new. Let's consider them one at a time:

(1)
$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} T) = \left(\hat{\mathbf{x}}\frac{\partial}{\partial x} + \hat{\mathbf{y}}\frac{\partial}{\partial y} + \hat{\mathbf{z}}\frac{\partial}{\partial z}\right) \cdot \left(\frac{\partial T}{\partial x}\hat{\mathbf{x}} + \frac{\partial T}{\partial y}\hat{\mathbf{y}} + \frac{\partial T}{\partial z}\hat{\mathbf{z}}\right)$$
$$= \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}. \tag{1.42}$$

This object, which we write as $\nabla^2 T$ for short, is called the **Laplacian** of T; we shall be studying it in great detail later on. Notice that the Laplacian of a *scalar* T is a *scalar*. Occasionally, we shall speak of the Laplacian of a *vector,* $\nabla^2\mathbf{v}$. By this we mean a *vector* quantity whose x-component is the Laplacian of v_x, and so on:[8]

$$\nabla^2\mathbf{v} \equiv (\nabla^2 v_x)\hat{\mathbf{x}} + (\nabla^2 v_y)\hat{\mathbf{y}} + (\nabla^2 v_z)\hat{\mathbf{z}}. \tag{1.43}$$

This is nothing more than a convenient extension of the meaning of ∇^2.

(2) The curl of a gradient is always zero:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} T) = \mathbf{0}. \tag{1.44}$$

This is an important fact, which we shall use repeatedly; you can easily prove it from the definition of $\boldsymbol{\nabla}$, Eq. 1.39. *Beware*: You might think Eq. 1.44 is "obviously" true—isn't it just $(\boldsymbol{\nabla} \times \boldsymbol{\nabla})T$, and isn't the cross product of *any* vector (in this case, $\boldsymbol{\nabla}$) with itself always zero? This reasoning is suggestive, but not quite conclusive, since $\boldsymbol{\nabla}$ is an *operator* and does not "multiply" in the usual way. The proof of Eq. 1.44, in fact, hinges on the equality of cross derivatives:

$$\frac{\partial}{\partial x}\left(\frac{\partial T}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial T}{\partial x}\right). \tag{1.45}$$

If you think I'm being fussy, test your intuition on this one:

$$(\boldsymbol{\nabla} T) \times (\boldsymbol{\nabla} S).$$

Is *that* always zero? (It *would* be, of course, if you replaced the $\boldsymbol{\nabla}$'s by an ordinary vector.)

(3) $\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v})$ seldom occurs in physical applications, and it has not been given any special name of its own—it's just **the gradient of the divergence.** Notice that $\boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v})$ is *not* the same as the Laplacian of a vector: $\nabla^2\mathbf{v} = (\boldsymbol{\nabla} \cdot \boldsymbol{\nabla})\mathbf{v} \neq \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v})$.

(4) The divergence of a curl, like the curl of a gradient, is always zero:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{v}) = 0. \tag{1.46}$$

You can prove this for yourself. (Again, there is a fraudulent short-cut proof, using the vector identity $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}$.)

(5) As you can check from the definition of $\boldsymbol{\nabla}$:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{v}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{v}) - \nabla^2\mathbf{v}. \tag{1.47}$$

So curl-of-curl gives nothing new; the first term is just number (3), and the second is the Laplacian (of a vector). (In fact, Eq. 1.47 is often used to *define* the

[8]In curvilinear coordinates, where the unit vectors themselves depend on position, they too must be differentiated (see Sect. 1.4.1).

Laplacian of a vector, in preference to Eq. 1.43, which makes explicit reference to Cartesian coordinates.)

Really, then, there are just two kinds of second derivatives: the Laplacian (which is of fundamental importance) and the gradient-of-divergence (which we seldom encounter). We could go through a similar ritual to work out *third* derivatives, but fortunately second derivatives suffice for practically all physical applications.

A final word on vector differential calculus: It *all* flows from the operator $\boldsymbol{\nabla}$, and from taking seriously its vectorial character. Even if you remembered *only* the definition of $\boldsymbol{\nabla}$, you could easily reconstruct all the rest.

Problem 1.26 Calculate the Laplacian of the following functions:

(a) $T_a = x^2 + 2xy + 3z + 4$.

(b) $T_b = \sin x \sin y \sin z$.

(c) $T_c = e^{-5x} \sin 4y \cos 3z$.

(d) $\mathbf{v} = x^2\,\hat{\mathbf{x}} + 3xz^2\,\hat{\mathbf{y}} - 2xz\,\hat{\mathbf{z}}$.

Problem 1.27 Prove that the divergence of a curl is always zero. *Check* it for function $\mathbf{v}_a$ in Prob. 1.15.

Problem 1.28 Prove that the curl of a gradient is always zero. *Check* it for function (b) in Prob. 1.11.

1.3 ■ INTEGRAL CALCULUS

1.3.1 ■ Line, Surface, and Volume Integrals

In electrodynamics, we encounter several different kinds of integrals, among which the most important are **line** (or **path**) **integrals**, **surface integrals** (or **flux**), and **volume integrals**.

(a) **Line Integrals.** A line integral is an expression of the form

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{v} \cdot d\mathbf{l}, \tag{1.48}$$

where $\mathbf{v}$ is a vector function, $d\mathbf{l}$ is the infinitesimal displacement vector (Eq. 1.22), and the integral is to be carried out along a prescribed path $\mathcal{P}$ from point $\mathbf{a}$ to point $\mathbf{b}$ (Fig. 1.20). If the path in question forms a closed loop (that is, if $\mathbf{b} = \mathbf{a}$), I shall put a circle on the integral sign:

$$\oint \mathbf{v} \cdot d\mathbf{l}. \tag{1.49}$$

At each point on the path, we take the dot product of $\mathbf{v}$ (evaluated at that point) with the displacement $d\mathbf{l}$ to the next point on the path. To a physicist, the most familiar example of a line integral is the work done by a force $\mathbf{F}$: $W = \int \mathbf{F} \cdot d\mathbf{l}$.

Ordinarily, the value of a line integral depends critically on the path taken from $\mathbf{a}$ to $\mathbf{b}$, but there is an important special class of vector functions for which the line