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

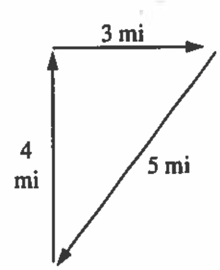

Figure 1.1

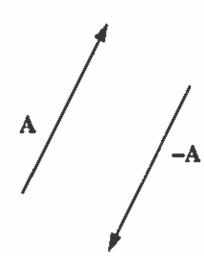

Figure 1.2

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.

(i) Addition of two vectors. Place the tail of **B** at the head of **A**; the sum, **A** + **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 (Fig. 1.4), add its opposite:

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

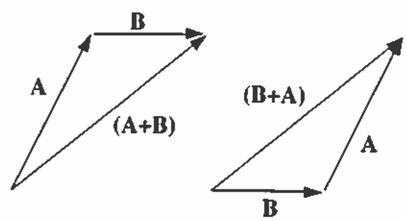

Figure 1.3

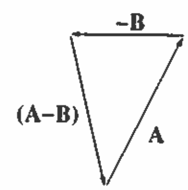

Figure 1.4

(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,*

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

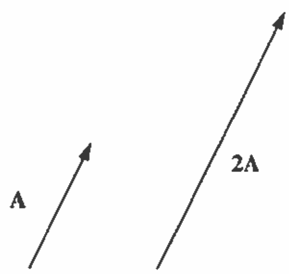

Figure 1.5

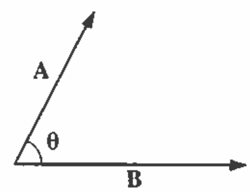

Figure 1.6

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

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

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

Example 1.1

Let $\mathbf{C} = \mathbf{A} - \mathbf{B}$ (Fig. 1.7), and calculate the dot product of $\mathbf{C}$ 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 length 1) pointing perpendicular to the plane of $\mathbf{A}$ and $\mathbf{B}$. (I shall use a hat (ˆ) to designate 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*,

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

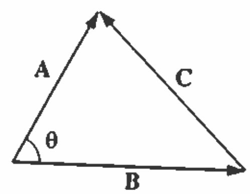

Figure 1.7

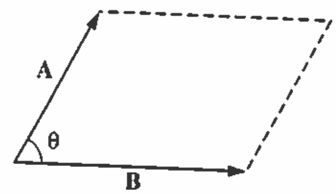

Figure 1.8

but *not commutative*. In fact,

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

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,

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

for any vector **A**.

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.

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)):

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

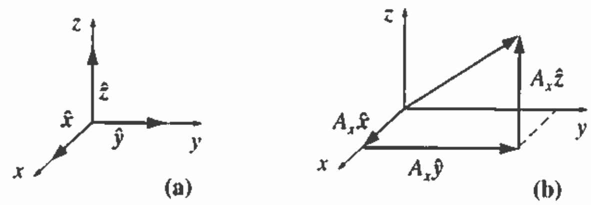

Figure 1.9

The numbers A_x, A_y, and A_z, are called **components** of **A**; geometrically, they are the projections of **A** along the three coordinate axes. We can now reformulate each of the four vector operations as a rule for manipulating components:

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

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

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

(ii) Rule: *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,

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

(iii) Rule: *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.) Note that the dot product of **A** with any *unit* vector is the component of **A** along that direction (thus $\mathbf{A}\cdot\hat{\mathbf{x}} = A_x$, $\mathbf{A}\cdot\hat{\mathbf{y}} = A_y$, and $\mathbf{A}\cdot\hat{\mathbf{z}} = A_z$).

Similarly,[1]

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

Therefore,

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

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

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

(iv) Rule: *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

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

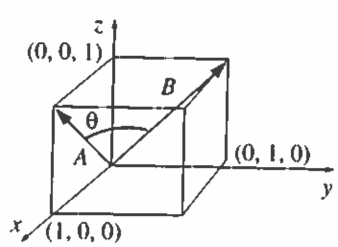

Figure 1.10

[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 are reversed: $\hat{\mathbf{x}}\times\hat{\mathbf{y}} = -\hat{\mathbf{z}}$, and so on. We shall use right-handed systems exclusively.

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 plane shown 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,

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

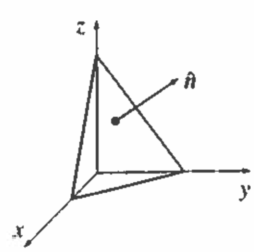

Figure 1.11

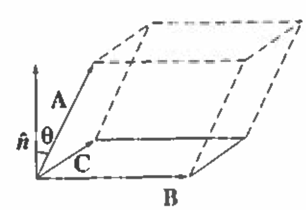

Figure 1.12

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.

(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. Incidentally, 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

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

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 (Fig. 1.13) is called the **position vector**:

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

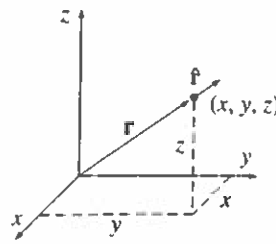

Figure 1.13

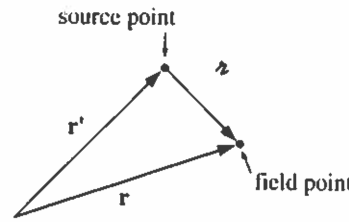

Figure 1.14

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 reserve a special letter 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 begin to appreciate the advantage 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

The definition of a vector as "a quantity with a magnitude and direction" is not altogether satisfactory: What precisely does "direction" *mean*?[2] This may seem a pedantic question, but we shall shortly 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 θ, 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

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

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

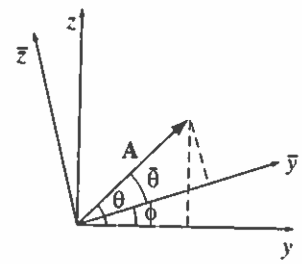

Figure 1.15

We might express this conclusion in matrix notation:

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

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

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

or, more compactly,

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

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 geometrical 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 is still the same number of 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 $\bar{A}_y$. 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.

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 transforms with *two* factors of R:

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

or, more compactly,

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

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.

Problem 1.8

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

(b) What constraints must the elements (R_{ij}) of the three-dimensional rotation matrix (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 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 does the cross product (1.13) of two vectors 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**.)

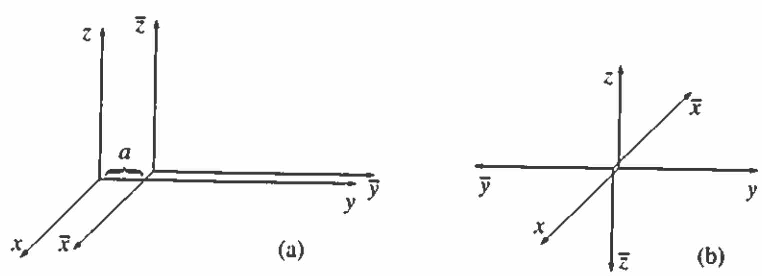

Figure 1.16

1.2 Differential Calculus

1.2.1 "Ordinary" Derivatives

Question: Suppose we have a function of one variable: $f(x)$. 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 change x by an 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.

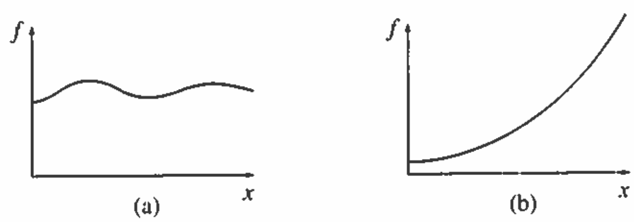

Figure 1.17

1.2.2 Gradient

Suppose, now, that we have a function of *three* variables—say, the temperature $T(x, y, z)$ in a 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.

Now 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

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

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:

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

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. ∇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 (1.35) in abstract form:

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

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 coutour 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 $\nabla T = 0$ 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:* $(\partial f/\partial \bar{y}) = (\partial f/\partial y)(\partial y/\partial \bar{y}) + (\partial f/\partial z)(\partial z/\partial \bar{y})$, and the analogous formula for $\partial f/\partial \bar{z}$. We know that $\bar{y} = y\cos\phi + z\sin\phi$ and $\bar{z} = -y\sin\phi + z\cos\phi$; "solve" these equations for y and z (as functions of $\bar{y}$ and $\bar{z}$), and compute the needed derivatives $\partial y/\partial \bar{y}$, $\partial z/\partial \bar{y}$, etc.]

1.2.3 The 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 is without specific meaning 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 should 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. Multiply a scalar a : $\mathbf{A}a$;
2. Multiply another vector **B**, via the dot product: $\mathbf{A}\cdot\mathbf{B}$;
3. Multiply another vector 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 $\mathbf{v}$ is itself a *scalar* $\nabla \cdot \mathbf{v}$. (You can't have the divergence of a scalar: that's meaningless.)

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 large *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,

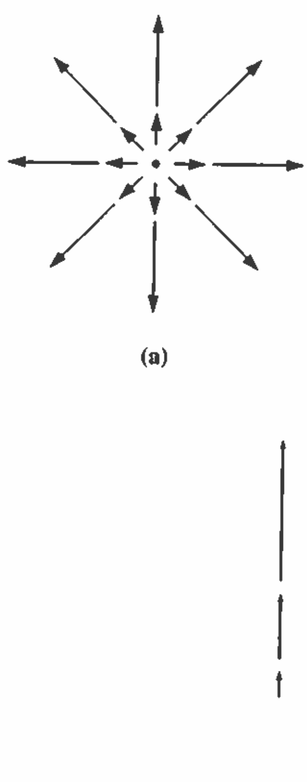

Figure 1.18

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 **v** in this model is the velocity of the water—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.")

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:

$$\nabla \times \mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ v_x & v_y & v_z \end{vmatrix}$$
$$= \hat{\mathbf{x}}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right) + \hat{\mathbf{y}}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) + \hat{\mathbf{z}}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right). \quad (1.41)$$

Notice that the curl of a vector function $\mathbf{v}$ is, like any cross product, a *vector*. (You cannot have the curl of a scalar; that's meaningless.)

Geometrical Interpretation: The name **curl** is also well chosen, for $\nabla \times \mathbf{v}$ is a measure of how much the vector $\mathbf{v}$ "curls 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.

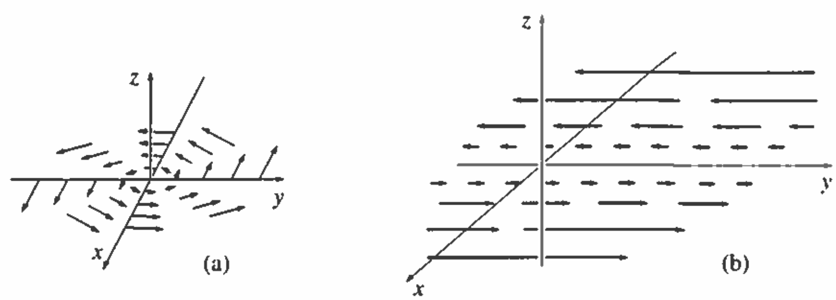

Figure 1.19

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 "curls around.")

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

Problem 1.19 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 general 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:

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

Similar relations hold for the vector derivatives. Thus,

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

and

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

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

and two ways to make a vector:

$f\mathbf{A}$ (scalar times vector),
$\mathbf{A}\times\mathbf{B}$ (cross product of two vectors).

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 on the inside 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, I haven't bothered to put them on the inside front cover.

Problem 1.20 Prove product rules (i), (iv), and (v).

Problem 1.21

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

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

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

Problem 1.22 (For masochists only.) Prove product rules (ii) and (vi). Refer to Prob. 1.21 for the definition of $(\mathbf{A}\cdot\nabla)\mathbf{B}$.

Problem 1.23 Derive the three quotient rules.

Problem 1.24

(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) 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 ∇; by applying ∇ *twice* we can construct five species of *second* derivatives. The gradient ∇T is a *vector*, so we can take the *divergence* and *curl* of it:

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

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

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

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

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

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

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

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

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

This object, which we write $\nabla^2 T$ 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:[3]

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

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

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

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

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

$$(\nabla T) \times (\nabla S).$$

Is *that* always zero? (It *would* be, of course, if you replaced the ∇'s by an ordinary vector.)

(3) $\nabla(\nabla \cdot \mathbf{v})$ for some reason 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 $\nabla(\nabla \cdot \mathbf{v})$ is *not* the same as the Laplacian of a vector: $\nabla^2 \mathbf{v} = (\nabla \cdot \nabla)\mathbf{v} \neq \nabla(\nabla \cdot \mathbf{v})$.

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

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

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 ∇:

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

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 Laplacian of a vector, in preference to Eq. 1.43, which makes specific 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).

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

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 ∇, and from taking seriously its vector character. Even if you remembered *only* the definition of ∇, you should be able, in principle, to reconstruct all the rest.

Problem 1.25 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 1.26 Prove that the divergence of a curl is always zero. *Check* it for function $\mathbf{v}_a$ in Prob. 1.15.

Problem 1.27 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

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

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 particular path taken from $\mathbf{a}$ to $\mathbf{b}$, but there is an important special class of vector functions for which the line integral is *independent* of the path, and is determined entirely by the end points. It will be our business in due course to characterize this special class of vectors. (A *force* that has this property is called **conservative**.)

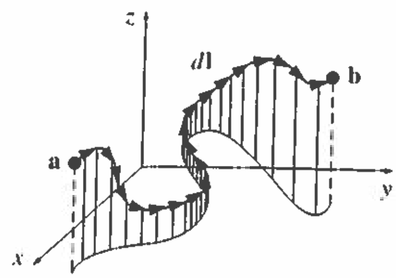

Figure 1.20

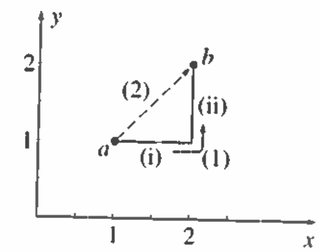

Figure 1.21

Example 1.6

Calculate the line integral of the function $\mathbf{v} = y^2\,\hat{\mathbf{x}} + 2x(y+1)\,\hat{\mathbf{y}}$ from the point $\mathbf{a} = (1, 1, 0)$ to the point $\mathbf{b} = (2, 2, 0)$, along the paths (1) and (2) in Fig. 1.21. What is $\oint \mathbf{v} \cdot d\mathbf{l}$ for the loop that goes from $\mathbf{a}$ to $\mathbf{b}$ along (1) and returns to $\mathbf{a}$ along (2)?

Solution: As always, $d\mathbf{l} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}}$. Path (1) consists of two parts. Along the "horizontal" segment $dy = dz = 0$, so

(i) $d\mathbf{l} = dx\,\hat{\mathbf{x}}$, $y = 1$, $\mathbf{v} \cdot d\mathbf{l} = y^2\,dx = dx$, so $\int \mathbf{v} \cdot d\mathbf{l} = \int_1^2 dx = 1$.

On the "vertical" stretch $dx = dz = 0$, so

(ii) $d\mathbf{l} = dy\,\hat{\mathbf{y}}$, $x = 2$, $\mathbf{v} \cdot d\mathbf{l} = 2x(y+1)\,dy = 4(y+1)\,dy$, so

$$\int \mathbf{v} \cdot d\mathbf{l} = 4\int_1^2 (y+1)\,dy = 10.$$

By path (1), then,

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{v} \cdot d\mathbf{l} = 1 + 10 = 11.$$

Meanwhile, on path (2) $x = y$, $dx = dy$, and $dz = 0$, so

$$d\mathbf{l} = dx\,\hat{\mathbf{x}} + dx\,\hat{\mathbf{y}}, \quad \mathbf{v} \cdot d\mathbf{l} = x^2\,dx + 2x(x+1)\,dx = (3x^2 + 2x)\,dx,$$

so

$$\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{v} \cdot d\mathbf{l} = \int_1^2 (3x^2 + 2x)\,dx = (x^3 + x^2)\Big|_1^2 = 10.$$

(The strategy here is to get everything in terms of one variable; I could just as well have eliminated x in favor of y.)

For the loop that goes *out* (1) and *back* (2), then,

$$\oint \mathbf{v} \cdot d\mathbf{l} = 11 - 10 = 1.$$

(b) **Surface Integrals.** A surface integral is an expression of the form

$$\int_S \mathbf{v} \cdot d\mathbf{a}, \tag{1.50}$$

where $\mathbf{v}$ is again some vector function, and $d\mathbf{a}$ is an infinitesimal patch of area, with direction perpendicular to the surface (Fig. 1.22). There are, of course, *two* directions perpendicular to any surface, so the *sign* of a surface integral is intrinsically ambiguous. If the surface is *closed* (forming a "balloon"), in which case I shall again put a circle on the integral sign

$$\oint \mathbf{v} \cdot d\mathbf{a},$$

then tradition dictates that "outward" is positive, but for open surfaces it's arbitrary. If $\mathbf{v}$ describes the flow of a fluid (mass per unit area per unit time), then $\int \mathbf{v} \cdot d\mathbf{a}$ represents the total mass per unit time passing through the surface—hence the alternative name, "flux."

Ordinarily, the value of a surface integral depends on the particular surface chosen, but there is a special class of vector functions for which it is *independent* of the surface, and is determined entirely by the boundary line. We shall soon be in a position to characterize this special class.

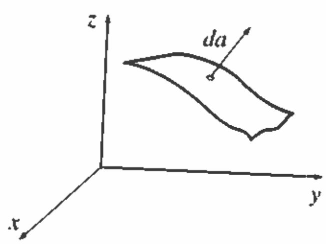

Figure 1.22

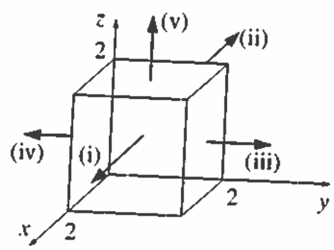

Figure 1.23

Example 1.7

Calculate the surface integral of $\mathbf{v} = 2xz\,\hat{\mathbf{x}} + (x+2)\,\hat{\mathbf{y}} + y(z^2-3)\,\hat{\mathbf{z}}$ over five sides (excluding the bottom) of the cubical box (side 2) in Fig. 1.23. Let "upward and outward" be the positive direction, as indicated by the arrows.

Solution: Taking the sides one at a time:

(i) $x = 2$, $d\mathbf{a} = dy\,dz\,\hat{\mathbf{x}}$, $\mathbf{v} \cdot d\mathbf{a} = 2xz\,dy\,dz = 4z\,dy\,dz$, so

$$\int \mathbf{v} \cdot d\mathbf{a} = 4 \int_0^2 dy \int_0^2 z\,dz = 16.$$

(ii) $x = 0$, $d\mathbf{a} = -dy\,dz\,\hat{\mathbf{x}}$, $\mathbf{v} \cdot d\mathbf{a} = -2xz\,dy\,dz = 0$, so

$$\int \mathbf{v} \cdot d\mathbf{a} = 0.$$

(iii) $y = 2,\ d\mathbf{a} = dx\,dz\,\hat{\mathbf{y}},\ \mathbf{v}\cdot d\mathbf{a} = (x+2)\,dx\,dz$, so

$$\int \mathbf{v}\cdot d\mathbf{a} = \int_0^2 (x+2)\,dx \int_0^2 dz = 12.$$

(iv) $y = 0,\ d\mathbf{a} = -dx\,dz\,\hat{\mathbf{y}},\ \mathbf{v}\cdot d\mathbf{a} = -(x+2)\,dx\,dz$, so

$$\int \mathbf{v}\cdot d\mathbf{a} = -\int_0^2 (x+2)\,dx \int_0^2 dz = -12.$$

(v) $z = 2,\ d\mathbf{a} = dx\,dy\,\hat{\mathbf{z}},\ \mathbf{v}\cdot d\mathbf{a} = y(z^2-3)\,dx\,dy = y\,dx\,dy$, so

$$\int \mathbf{v}\cdot d\mathbf{a} = \int_0^2 dx \int_0^2 y\,dy = 4.$$

Evidently the *total* flux is

$$\int_{\text{surface}} \mathbf{v}\cdot d\mathbf{a} = 16 + 0 + 12 - 12 + 4 = 20.$$

(c) **Volume Integrals.** A volume integral is an expression of the form

$$\int_{\mathcal{V}} T\,d\tau, \tag{1.51}$$

where T is a scalar function and $d\tau$ is an infinitesimal volume element. In Cartesian coordinates,

$$d\tau = dx\,dy\,dz. \tag{1.52}$$

For example, if T is the density of a substance (which might vary from point to point), then the volume integral would give the total mass. Occasionally we shall encounter volume integrals of *vector* functions:

$$\int \mathbf{v}\,d\tau = \int (v_x\,\hat{\mathbf{x}} + v_y\,\hat{\mathbf{y}} + v_z\,\hat{\mathbf{z}})d\tau = \hat{\mathbf{x}}\int v_x d\tau + \hat{\mathbf{y}}\int v_y d\tau + \hat{\mathbf{z}}\int v_z d\tau; \tag{1.53}$$

because the unit vectors are constants, they come outside the integral.

Example 1.8

Calculate the volume integral of $T = xyz^2$ over the prism in Fig. 1.24.

Solution: You can do the three integrals in any order. Let's do x first: it runs from 0 to $(1-y)$; then y (it goes from 0 to 1); and finally z (0 to 3):

$$\int T\,d\tau = \int_0^3 z^2 \left\{\int_0^1 y\left[\int_0^{1-y} x\,dx\right]dy\right\}dz =$$

$$\frac{1}{2}\int_0^3 z^2\,dz \int_0^1 (1-y)^2 y\,dy = \tfrac{1}{2}(9)(\tfrac{1}{12}) = \tfrac{3}{8}.$$

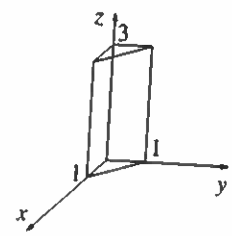

Figure 1.24

Problem 1.28 Calculate the line integral of the function $\mathbf{v} = x^2\,\hat{\mathbf{x}} + 2yz\,\hat{\mathbf{y}} + y^2\,\hat{\mathbf{z}}$ from the origin to the point (1,1,1) by three different routes:

(a) $(0,0,0) \to (1,0,0) \to (1,1,0) \to (1,1,1)$;

(b) $(0,0,0) \to (0,0,1) \to (0,1,1) \to (1,1,1)$;

(c) The direct straight line.

(d) What is the line integral around the closed loop that goes *out* along path (a) and *back* along path (b)?

Problem 1.29 Calculate the surface integral of the function in Ex. 1.7, over the *bottom* of the box. For consistency, let "upward" be the positive direction. Does the surface integral depend only on the boundary line for this function? What is the total flux over the *closed* surface of the box (*including* the bottom)? [*Note:* For the *closed* surface the positive direction is "outward," and hence "down," for the bottom face.]

Problem 1.30 Calculate the volume integral of the function $T = z^2$ over the tetrahedron with corners at (0,0,0), (1,0,0), (0,1,0), and (0,0,1).

1.3.2 The Fundamental Theorem of Calculus

Suppose $f(x)$ is a function of one variable. The **fundamental theorem of calculus** states:

$$\int_a^b \frac{df}{dx}dx = f(b) - f(a). \tag{1.54}$$

In case this doesn't look familiar, let's write it another way:

$$\int_a^b F(x)\,dx = f(b) - f(a),$$

where $df/dx = F(x)$. The fundamental theorem tells you how to integrate $F(x)$: you think up a function $f(x)$ whose *derivative* is equal to F.

Geometrical Interpretation: According to Eq. 1.33, $df = (df/dx)dx$ is the infinitesimal change in f when you go from (x) to $(x + dx)$. The fundamental theorem (1.54) says that if you chop the interval from a to b (Fig. 1.25) into many tiny pieces, dx, and add up the increments df from each little piece, the result is (not surprisingly) equal to the total change in f: $f(b) - f(a)$. In other words, there are two ways to determine the total change in the function: *either* subtract the values at the ends *or* go step-by-step, adding up all the tiny increments as you go. You'll get the same answer either way.

Notice the basic format of the fundamental theorem: *the integral of a derivative over an interval is given by the value of the function at the end points (boundaries).* In vector calculus there are three species of derivative (gradient, divergence, and curl), and each has its own "fundamental theorem," with essentially the same format. I don't plan to prove these theorems here; rather, I shall explain what they *mean*, and try to make them *plausible*. Proofs are given in Appendix A.

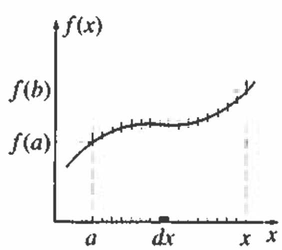

Figure 1.25

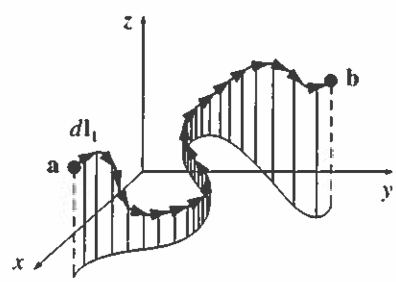

Figure 1.26

1.3.3 The Fundamental Theorem for Gradients

Suppose we have a scalar function of three variables $T(x, y, z)$. Starting at point $\mathbf{a}$, we move a small distance $d\mathbf{l}_1$ (Fig. 1.26). According to Eq. 1.37, the function T will change by an amount

$$dT = (\nabla T) \cdot d\mathbf{l}_1.$$

Now we move a little further, by an additional small displacement $d\mathbf{l}_2$; the incremental change in T will be $(\nabla T) \cdot d\mathbf{l}_2$. In this manner, proceeding by infinitesimal steps, we make the journey to point $\mathbf{b}$. At each step we compute the gradient of T (at that point) and dot it into the displacement $d\mathbf{l}$... this gives us the change in T. Evidently the *total* change in T in going from $\mathbf{a}$ to $\mathbf{b}$ *along the path selected* is

$$\boxed{\int_{\mathbf{a}\atop\mathcal{P}}^{\mathbf{b}} (\nabla T) \cdot d\mathbf{l} = T(\mathbf{b}) - T(\mathbf{a}).} \tag{1.55}$$

This is called the **fundamental theorem for gradients**; like the "ordinary" fundamental theorem, it says that the integral (here a *line* integral) of a derivative (here the *gradient*) is given by the value of the function at the boundaries (**a** and **b**).

Geometrical Interpretation: Suppose you wanted to determine the height of the Eiffel Tower. You could climb the stairs, using a ruler to measure the rise at each step, and adding them all up (that's the left side of Eq. 1.55), or you could place altimeters at the top and the bottom, and subtract the two readings (that's the right side); you should get the same answer either way (that's the fundamental theorem).

Incidentally, as we found in Ex. 1.6, line integrals ordinarily depend on the *path* taken from **a** to **b**. But the *right* side of Eq. 1.55 makes no reference to the path—only to the end points. Evidently, *gradients* have the special property that their line integrals are path independent:

Corollary 1: $\int_{\mathbf{a}}^{\mathbf{b}}(\nabla T)\cdot d\mathbf{l}$ is independent of path taken from **a** to **b**.

Corollary 2: $\oint(\nabla T)\cdot d\mathbf{l}=0$, since the beginning and end points are identical, and hence $T(\mathbf{b})-T(\mathbf{a})=0$.

Example 1.9

Let $T = xy^2$, and take point **a** to be the origin $(0, 0, 0)$ and **b** the point $(2, 1, 0)$. Check the fundamental theorem for gradients.

Solution: Although the integral is independent of path, we must *pick* a specific path in order to evaluate it. Let's go out along the x axis (step i) and then up (step ii) (Fig. 1.27). As always, $d\mathbf{l} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}}$; $\nabla T = y^2\,\hat{\mathbf{x}} + 2xy\,\hat{\mathbf{y}}$.

(i) $y = 0$; $d\mathbf{l} = dx\,\hat{\mathbf{x}}$, $\nabla T\cdot d\mathbf{l} = y^2\,dx = 0$, so

$$\int_{\mathrm{i}} \nabla T\cdot d\mathbf{l} = 0.$$

(ii) $x = 2$; $d\mathbf{l} = dy\,\hat{\mathbf{y}}$, $\nabla T\cdot d\mathbf{l} = 2xy\,dy = 4y\,dy$, so

$$\int_{\mathrm{ii}} \nabla T\cdot d\mathbf{l} = \int_0^1 4y\,dy = 2y^2\Big|_0^1 = 2.$$

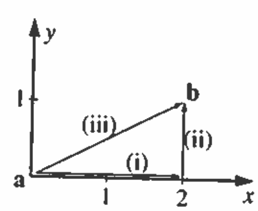

Figure 1.27

Evidently the total line integral is 2. Is this consistent with the fundamental theorem? Yes: $T(\mathbf{b}) - T(\mathbf{a}) = 2 - 0 = 2$.

Now, just to convince you that the answer is independent of path, let me calculate the same integral along path iii (the straight line from **a** to **b**):

(iii) $y = \frac{1}{2}x,\ dy = \frac{1}{2}\,dx,\ \nabla T \cdot d\mathbf{l} = y^2\,dx + 2xy\,dy = \frac{3}{4}x^2\,dx$, so

$$\int_{\text{iii}} \nabla T \cdot d\mathbf{l} = \int_0^2 \tfrac{3}{4}x^2\,dx = \tfrac{1}{4}x^3\Big|_0^2 = 2.$$

Problem 1.31 Check the fundamental theorem for gradients, using $T = x^2 + 4xy + 2yz^3$, the points $\mathbf{a} = (0, 0, 0)$, $\mathbf{b} = (1, 1, 1)$, and the three paths in Fig. 1.28:

(a) $(0, 0, 0) \to (1, 0, 0) \to (1, 1, 0) \to (1, 1, 1)$;

(b) $(0, 0, 0) \to (0, 0, 1) \to (0, 1, 1) \to (1, 1, 1)$;

(c) the parabolic path $z = x^2$; $y = x$.

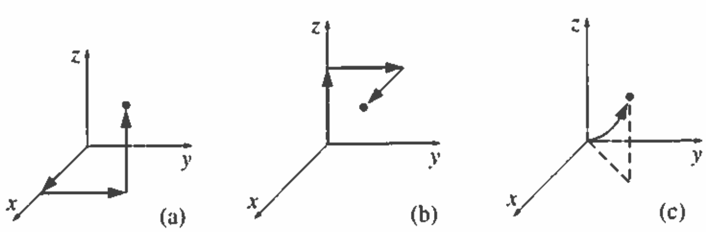

Figure 1.28

1.3.4 The Fundamental Theorem for Divergences

The fundamental theorem for divergences states that:

$$\boxed{\int_{\mathcal{V}} (\nabla \cdot \mathbf{v})\,d\tau = \oint_{\mathcal{S}} \mathbf{v} \cdot d\mathbf{a}.} \tag{1.56}$$

In honor, I suppose of its great importance, this theorem has at least three special names: **Gauss's theorem, Green's theorem**, or, simply, the **divergence theorem**. Like the other "fundamental theorems," it says that the *integral* of a *derivative* (in this case the *divergence*) over a *region* (in this case a *volume*) is equal to the value of the function at the *boundary* (in this case the *surface* that bounds the volume). Notice that the boundary term is itself an integral (specifically, a surface integral). This is reasonable: the "boundary" of a *line* is just two end points, but the boundary of a *volume* is a (closed) surface.

Geometrical Interpretation: If $\mathbf{v}$ represents the flow of an incompressible fluid, then the *flux* of $\mathbf{v}$ (the right side of Eq. 1.56) is the total amount of fluid passing out through the surface, per unit time. Now, the divergence measures the "spreading out" of the vectors from a point—a place of high divergence is like a "faucet," pouring out liquid. If we have lots of faucets in a region filled with incompressible fluid, an equal amount of liquid will be forced out through the boundaries of the region. In fact, there are *two* ways we could determine how much is being produced: (a) we could count up all the faucets, recording how much each puts out, or (b) we could go around the boundary, measuring the flow at each point, and add it all up. You get the same answer either way:

$$\int (\text{faucets within the volume}) = \oint (\text{flow out through the surface}).$$

This, in essence, is what the divergence theorem says.

Example 1.10

Check the divergence theorem using the function

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

and the unit cube situated at the origin (Fig. 1.29).

Solution: In this case

$$\nabla \cdot \mathbf{v} = 2(x + y),$$

and

$$\int_{\mathcal{V}} 2(x + y)\,d\tau = 2\int_0^1\int_0^1\int_0^1 (x + y)\,dx\,dy\,dz,$$

$$\int_0^1 (x + y)\,dx = \tfrac{1}{2} + y, \quad \int_0^1 (\tfrac{1}{2} + y)\,dy = 1, \quad \int_0^1 1\,dz = 1.$$

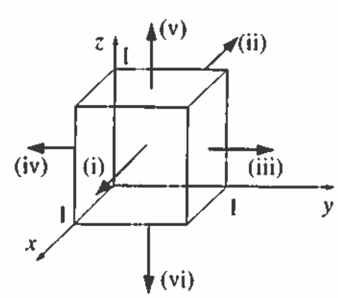

Figure 1.29

Evidently,

$$\int_{\mathcal{V}} \nabla \cdot \mathbf{v}\, d\tau = 2.$$

So much for the left side of the divergence theorem. To evaluate the surface integral we must consider separately the six sides of the cube:

(i) $$\int \mathbf{v} \cdot d\mathbf{a} = \int_0^1 \int_0^1 y^2 dy\, dz = \tfrac{1}{3}.$$

(ii) $$\int \mathbf{v} \cdot d\mathbf{a} = -\int_0^1 \int_0^1 y^2\, dy\, dz = -\tfrac{1}{3}.$$

(iii) $$\int \mathbf{v} \cdot d\mathbf{a} = \int_0^1 \int_0^1 (2x + z^2)\, dx\, dz = \tfrac{4}{3}.$$

(iv) $$\int \mathbf{v} \cdot d\mathbf{a} = -\int_0^1 \int_0^1 z^2\, dx\, dz = -\tfrac{1}{3}.$$

(v) $$\int \mathbf{v} \cdot d\mathbf{a} = \int_0^1 \int_0^1 2y\, dx\, dy = 1.$$

(vi) $$\int \mathbf{v} \cdot d\mathbf{a} = -\int_0^1 \int_0^1 0\, dx\, dy = 0.$$

So the total flux is:

$$\oint_{\mathcal{S}} \mathbf{v} \cdot d\mathbf{a} = \tfrac{1}{3} - \tfrac{1}{3} + \tfrac{4}{3} - \tfrac{1}{3} + 1 + 0 = 2,$$

as expected.

Problem 1.32 Test the divergence theorem for the function $\mathbf{v} = (xy)\,\hat{\mathbf{x}} + (2yz)\,\hat{\mathbf{y}} + (3zx)\,\hat{\mathbf{z}}$. Take as your volume the cube shown in Fig. 1.30, with sides of length 2.

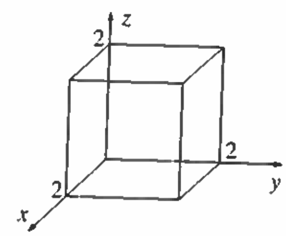

Figure 1.30

1.3.5 The Fundamental Theorem for Curls

The fundamental theorem for curls, which goes by the special name of **Stokes' theorem**, states that

$$\boxed{\int_{\mathcal{S}} (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = \oint_{\mathcal{P}} \mathbf{v} \cdot d\mathbf{l}.} \tag{1.57}$$

As always, the *integral* of a *derivative* (here, the *curl*) over a *region* (here, a patch of *surface*) is equal to the value of the function at the *boundary* (here, the perimeter of the patch). As in the case of the divergence theorem, the boundary term is itself an integral—specifically, a closed line integral.

Geometrical Interpretation: Recall that the curl measures the "twist" of the vectors **v**; a region of high curl is a whirlpool—if you put a tiny paddle wheel there, it will rotate. Now, the *integral* of the curl over some surface (or, more precisely, the *flux* of the curl *through* that surface) represents the "total amount of swirl," and we can determine that swirl just as well by going around the *edge* and finding how much the flow is following the boundary (Fig. 1.31). You may find this a rather forced interpretation of Stokes' theorem, but it's a helpful mnemonic, if nothing else.

You might have noticed an apparent ambiguity in Stokes' theorem: concerning the boundary line integral, which *way* are we supposed to go around (clockwise or counter-clockwise)? If we go the "wrong" way we'll pick up an overall sign error. The answer is that it doesn't *matter* which way you go as long as you are *consistent*, for there is a compensating sign ambiguity in the surface integral: Which way does $d\mathbf{a}$ point? For a *closed* surface (as in the divergence theorem) $d\mathbf{a}$ points in the direction of the *outward* normal; but for an *open* surface, which way is "out?" Consistency in Stokes' theorem (as in all such matters) is given by the right-hand rule: If your fingers point in the direction of the line integral, then your thumb fixes the direction of $d\mathbf{a}$ (Fig. 1.32).

Now, there are plenty of surfaces (infinitely many) that share any given boundary line. Twist a paper clip into a loop and dip it in soapy water. The soap film constitutes a surface, with the wire loop as its boundary. If you blow on it, the soap film will expand, making a larger surface, with the same boundary. Ordinarily, a flux integral depends critically on what surface you integrate over, but evidently this is *not* the case with curls. For Stokes'

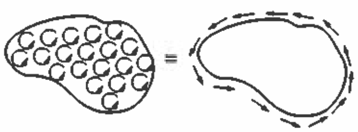

Figure 1.31

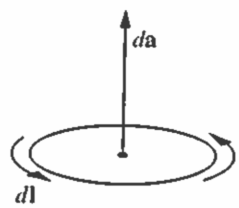

Figure 1.32

theorem says that $\int(\nabla \times \mathbf{v}) \cdot d\mathbf{a}$ is equal to the line integral of $\mathbf{v}$ around the boundary, and the latter makes no reference to the specific surface you choose.

Corollary 1: $\int(\nabla \times \mathbf{v}) \cdot d\mathbf{a}$ depends only on the boundary line, not on the particular surface used.

Corollary 2: $\oint(\nabla \times \mathbf{v}) \cdot d\mathbf{a} = 0$ for any closed surface, since the boundary line, like the mouth of a balloon, shrinks down to a point, and hence the right side of Eq. 1.57 vanishes.

These corollaries are analogous to those for the gradient theorem. We shall develop the parallel further in due course.

Example 1.11

Suppose $\mathbf{v} = (2xz + 3y^2)\hat{\mathbf{y}} + (4yz^2)\hat{\mathbf{z}}$. Check Stokes' theorem for the square surface shown in Fig. 1.33.

Solution: Here

$$\nabla \times \mathbf{v} = (4z^2 - 2x)\,\hat{\mathbf{x}} + 2z\,\hat{\mathbf{z}} \quad \text{and} \quad d\mathbf{a} = dy\,dz\,\hat{\mathbf{x}}.$$

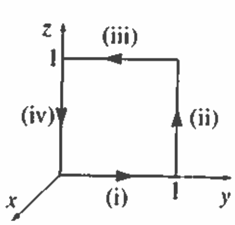

Figure 1.33

(In saying that $d\mathbf{a}$ points in the x direction, we are committing ourselves to a counterclockwise line integral. We could as well write $d\mathbf{a} = -dy\,dz\,\hat{\mathbf{x}}$, but then we would be obliged to go clockwise.) Since $x = 0$ for this surface,

$$\int (\nabla \times \mathbf{v}) \cdot d\mathbf{a} = \int_0^1 \int_0^1 4z^2\, dy\, dz = \tfrac{4}{3}.$$

Now, what about the line integral? We must break this up into four segments:

(i) $x = 0, \quad z = 0, \quad \mathbf{v} \cdot d\mathbf{l} = 3y^2\, dy, \quad \int \mathbf{v} \cdot d\mathbf{l} = \int_0^1 3y^2\, dy = 1,$

(ii) $x = 0, \quad y = 1, \quad \mathbf{v} \cdot d\mathbf{l} = 4z^2\, dz, \quad \int \mathbf{v} \cdot d\mathbf{l} = \int_0^1 4z^2\, dz = \frac{4}{3},$

(iii) $x = 0, \quad z = 1, \quad \mathbf{v} \cdot d\mathbf{l} = 3y^2\, dy, \quad \int \mathbf{v} \cdot d\mathbf{l} = \int_1^0 3y^2\, dy = -1,$

(iv) $x = 0, \quad y = 0, \quad \mathbf{v} \cdot d\mathbf{l} = 0, \quad \int \mathbf{v} \cdot d\mathbf{l} = \int_1^0 0\, dz = 0.$

So

$$\oint \mathbf{v} \cdot d\mathbf{l} = 1 + \tfrac{4}{3} - 1 + 0 = \tfrac{4}{3}.$$

It checks.

A point of strategy: notice how I handled step (iii). There is a temptation to write $d\mathbf{l} = -dy\,\hat{\mathbf{y}}$ here, since the path goes to the left. You can get away with this, if you insist, by running the integral from $0 \to 1$. Personally, I prefer to say $d\mathbf{l} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}}$ *always* (never any minus signs) and let the limits of the integral take care of the direction.

Problem 1.33 Test Stokes' theorem for the function $\mathbf{v} = (xy)\,\hat{\mathbf{x}} + (2yz)\,\hat{\mathbf{y}} + (3zx)\,\hat{\mathbf{z}}$, using the triangular shaded area of Fig. 1.34.

Problem 1.34 Check Corollary 1 by using the same function and boundary line as in Ex. 1.11, but integrating over the five sides of the cube in Fig. 1.35. The back of the cube is open.

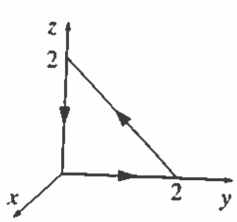

Figure 1.34

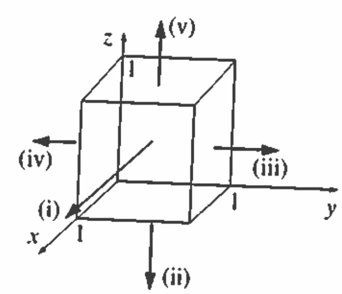

Figure 1.35

1.3.6 Integration by Parts

The technique known (awkwardly) as **integration by parts** exploits the product rule for derivatives:

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

Integrating both sides, and invoking the fundamental theorem:

$$\int_a^b \frac{d}{dx}(fg)\,dx = fg\Big|_a^b = \int_a^b f\left(\frac{dg}{dx}\right)dx + \int_a^b g\left(\frac{df}{dx}\right)dx,$$

or

$$\int_a^b f\left(\frac{dg}{dx}\right)dx = -\int_a^b g\left(\frac{df}{dx}\right)dx + fg\Big|_a^b. \tag{1.58}$$

That's integration by parts. It pertains to the situation in which you are called upon to integrate the product of one function (f) and the *derivative* of another (g); it says you can *transfer the derivative from g to f*, at the cost of a minus sign and a boundary term.

Example 1.12

Evaluate the integral

$$\int_0^\infty xe^{-x}\,dx.$$

Solution: The exponential can be expressed as a derivative:

$$e^{-x} = \frac{d}{dx}\left(-e^{-x}\right);$$

in this case, then, $f(x) = x$, $g(x) = -e^{-x}$, and $df/dx = 1$, so

$$\int_0^\infty xe^{-x}\,dx = \int_0^\infty e^{-x}\,dx - xe^{-x}\Big|_0^\infty = -e^{-x}\Big|_0^\infty = 1.$$

We can exploit the product rules of vector calculus, together with the appropriate fundamental theorems, in exactly the same way. For example, integrating

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

over a volume, and invoking the divergence theorem, yields

$$\int \nabla\cdot(f\mathbf{A})\,d\tau = \int f(\nabla\cdot\mathbf{A})\,d\tau + \int \mathbf{A}\cdot(\nabla f)\,d\tau = \oint f\mathbf{A}\cdot d\mathbf{a},$$

or

$$\int_{\mathcal{V}} f(\nabla\cdot\mathbf{A})\,d\tau = -\int_{\mathcal{V}} \mathbf{A}\cdot(\nabla f)\,d\tau + \oint_{\mathcal{S}} f\mathbf{A}\cdot d\mathbf{a}. \tag{1.59}$$

Here again the integrand is the product of one function (f) and the derivative (in this case the *divergence*) of another ($\mathbf{A}$), and integration by parts licenses us to transfer the derivative from $\mathbf{A}$ to f (where it becomes a *gradient*), at the cost of a minus sign and a boundary term (in this case a surface integral).

You might wonder how often one is likely to encounter an integral involving the product of one function and the derivative of another; the answer is *surprisingly* often, and integration by parts turns out to be one of the most powerful tools in vector calculus.

Problem 1.35

(a) Show that

$$\int_{\mathcal{S}} f(\nabla \times \mathbf{A}) \cdot d\mathbf{a} = \int_{\mathcal{S}} [\mathbf{A} \times (\nabla f)] \cdot d\mathbf{a} + \oint_{\mathcal{P}} f\mathbf{A} \cdot d\mathbf{l}. \tag{1.60}$$

(b) Show that

$$\int_{\mathcal{V}} \mathbf{B} \cdot (\nabla \times \mathbf{A})\, d\tau = \int_{\mathcal{V}} \mathbf{A} \times (\nabla \times \mathbf{B})\, d\tau + \oint_{\mathcal{S}} (\mathbf{A} \times \mathbf{B}) \cdot d\mathbf{a}. \tag{1.61}$$

1.4 Curvilinear Coordinates

1.4.1 Spherical Polar Coordinates

The spherical polar coordinates (r, θ, ϕ) of a point P are defined in Fig. 1.36; r is the distance from the origin (the magnitude of the position vector), θ (the angle down from the z axis) is called the **polar angle**, and ϕ (the angle around from the x axis) is the **azimuthal angle**. Their relation to Cartesian coordinates (x, y, z) can be read from the figure:

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \tag{1.62}$$

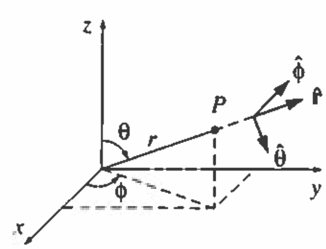

Figure 1.36

Figure 1.36 also shows three unit vectors, $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$, pointing in the direction of increase of the corresponding coordinates. They constitute an orthogonal (mutually perpendicular) basis set (just like $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$), and any vector **A** can be expressed in terms of them in the usual way:

$$\mathbf{A} = A_r\,\hat{\mathbf{r}} + A_\theta\,\hat{\boldsymbol{\theta}} + A_\phi\,\hat{\boldsymbol{\phi}}. \tag{1.63}$$

A_r, A_θ, and A_ϕ are the radial, polar, and azimuthal components of **A**. In terms of the Cartesian unit vectors,

$$\left.\begin{aligned} \hat{\mathbf{r}} &= \sin\theta\cos\phi\,\hat{\mathbf{x}} + \sin\theta\sin\phi\,\hat{\mathbf{y}} + \cos\theta\,\hat{\mathbf{z}}, \\ \hat{\boldsymbol{\theta}} &= \cos\theta\cos\phi\,\hat{\mathbf{x}} + \cos\theta\sin\phi\,\hat{\mathbf{y}} - \sin\theta\,\hat{\mathbf{z}}, \\ \hat{\boldsymbol{\phi}} &= -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}}, \end{aligned}\right\} \tag{1.64}$$

as you can easily check for yourself (Prob. 1.37). I have put these formulas inside the back cover, for easy reference.

But there is a poisonous snake lurking here that I'd better warn you about: $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ are associated with a *particular point P*, and they *change direction* as P moves around. For example, $\hat{\mathbf{r}}$ always points radially outward, but "radially outward" can be the x direction, the y direction, or any other direction, depending on where you are. In Fig. 1.37, $\mathbf{A} = \hat{\mathbf{y}}$ and $\mathbf{B} = -\hat{\mathbf{y}}$, and yet *both* of them would be written as $\hat{\mathbf{r}}$ in spherical coordinates. One could take account of this by explicitly indicating the point of reference: $\hat{\mathbf{r}}(\theta, \phi), \hat{\boldsymbol{\theta}}(\theta, \phi), \hat{\boldsymbol{\phi}}(\theta, \phi)$, but this would be cumbersome, and as long as you are alert to the problem I don't think it will cause difficulties.[4] In particular, do not naïvely combine the spherical components of vectors associated with different points (in Fig. 1.37, $\mathbf{A} + \mathbf{B} = 0$, not $2\hat{\mathbf{r}}$, and $\mathbf{A} \cdot \mathbf{B} = -1$, not $+1$). Beware of differentiating a vector that is expressed in spherical coordinates, since the unit vectors themselves are functions of position ($\partial\hat{\mathbf{r}}/\partial\theta = \hat{\boldsymbol{\theta}}$, for example). And do not take $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ outside an integral, as we did with $\hat{\mathbf{x}}, \hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ in Eq. 1.53. In general, if you're uncertain about the validity of an operation, reexpress the problem in Cartesian coordinates, where this difficulty does not arise.

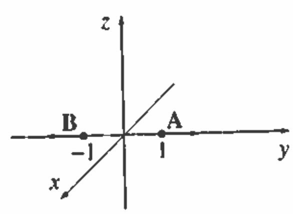

Figure 1.37

[4] I claimed on the very first page that vectors have no location, and I'll stand by that. The vectors themselves live "out there," completely independent of our choice of coordinates. But the *notation* we use to represent them *does* depend on the point in question, in curvilinear coordinates.

An infinitesimal displacement in the $\hat{\mathbf{r}}$ direction is simply dr (Fig. 1.38a), just as an infinitesimal element of length in the x direction is dx:

$$dl_r = dr. \tag{1.65}$$

On the other hand, an infinitesimal element of length in the $\hat{\boldsymbol{\theta}}$ direction (Fig. 1.38b) is *not* just $d\theta$ (that's an *angle*—it doesn't even have the right *units* for a length), but rather $r\,d\theta$:

$$dl_\theta = r\,d\theta. \tag{1.66}$$

Similarly, an infinitesimal element of length in the $\hat{\boldsymbol{\phi}}$ direction (Fig. 1.38c) is $r\sin\theta\,d\phi$:

$$dl_\phi = r\sin\theta\,d\phi. \tag{1.67}$$

Thus, the general infinitesimal displacement $d\mathbf{l}$ is

$$d\mathbf{l} = dr\,\hat{\mathbf{r}} + r\,d\theta\,\hat{\boldsymbol{\theta}} + r\sin\theta\,d\phi\,\hat{\boldsymbol{\phi}}. \tag{1.68}$$

This plays the role (in line integrals, for example) that $d\mathbf{l} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} + dz\,\hat{\mathbf{z}}$ played in Cartesian coordinates.

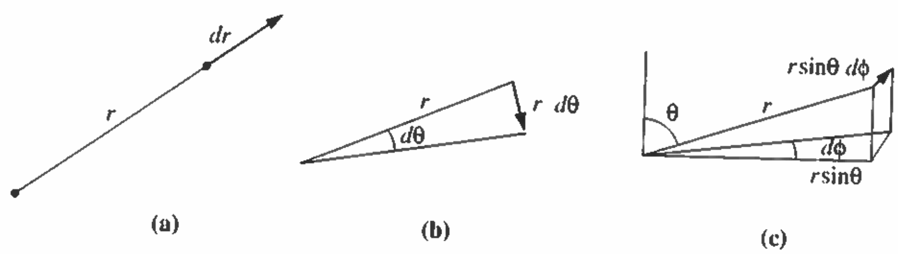

Figure 1.38

The infinitesimal volume element $d\tau$, in spherical coordinates, is the product of the three infinitesimal displacements:

$$d\tau = dl_r\,dl_\theta\,dl_\phi = r^2\sin\theta\,dr\,d\theta\,d\phi. \tag{1.69}$$

I cannot give you a general expression for *surface* elements $d\mathbf{a}$, since these depend on the orientation of the surface. You simply have to analyze the geometry for any given case (this goes for Cartesian and curvilinear coordinates alike). If you are integrating over the surface of a sphere, for instance, then r is constant, whereas θ and ϕ change (Fig. 1.39), so

$$d\mathbf{a}_1 = dl_\theta\,dl_\phi\,\hat{\mathbf{r}} = r^2\sin\theta\,d\theta\,d\phi\,\hat{\mathbf{r}}.$$

On the other hand, if the surface lies in the xy plane, say, so that θ is constant (to wit: $\pi/2$) while r and ϕ vary, then

$$d\mathbf{a}_2 = dl_r\,dl_\phi\,\hat{\boldsymbol{\theta}} = r\,dr\,d\phi\,\hat{\boldsymbol{\theta}}.$$

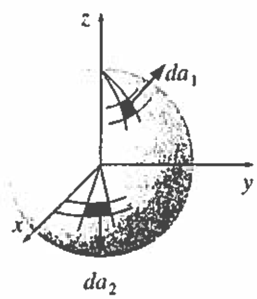

Figure 1.39

Notice, finally, that r ranges from 0 to ∞, ϕ from 0 to 2π, and θ from 0 to π (*not* 2π—that would count every point twice).[5]

Example 1.13

Find the volume of a sphere of radius R.

Solution:

$$\begin{aligned} V &= \int d\tau = \int_{r=0}^{R}\int_{\theta=0}^{\pi}\int_{\phi=0}^{2\pi} r^2 \sin\theta \, dr\, d\theta\, d\phi \\ &= \left(\int_0^R r^2\,dr\right)\left(\int_0^\pi \sin\theta\,d\theta\right)\left(\int_0^{2\pi} d\phi\right) \\ &= \left(\frac{R^3}{3}\right)(2)(2\pi) = \tfrac{4}{3}\pi R^3. \end{aligned}$$

(Not a big surprise.)

So far we have talked only about the *geometry* of spherical coordinates. Now I would like to "translate" the vector derivatives (gradient, divergence, curl, and Laplacian) into r, θ, ϕ notation. In principle this is entirely straightforward: in the case of the gradient,

$$\nabla T = \frac{\partial T}{\partial x}\hat{\mathbf{x}} + \frac{\partial T}{\partial y}\hat{\mathbf{y}} + \frac{\partial T}{\partial z}\hat{\mathbf{z}},$$

for instance, we would first use the chain rule to reexpress the partials:

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial r}\left(\frac{\partial r}{\partial x}\right) + \frac{\partial T}{\partial \theta}\left(\frac{\partial \theta}{\partial x}\right) + \frac{\partial T}{\partial \phi}\left(\frac{\partial \phi}{\partial x}\right).$$

[5] Alternatively, you could run ϕ from 0 to π (the "eastern hemisphere") and cover the "western hemisphere" by extending θ from π up to 2π. But this is very bad notation, since, among other things, $\sin\theta$ will then run negative, and you'll have to put absolute value signs around that term in volume and surface elements (area and volume being intrinsically positive quantities).

The terms in parentheses could be worked out from Eq. 1.62—or rather, the *inverse* of those equations (Prob. 1.36). Then we'd do the same for $\partial T/\partial y$ and $\partial T/\partial z$. Finally, we'd substitute in the formulas for $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ in terms of $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$ (Prob. 1.37). It would take an hour to figure out the gradient in spherical coordinates by this brute-force method. I suppose this is how it was first done, but there is a much more efficient indirect approach, explained in Appendix A, which has the extra advantage of treating all coordinate systems at once. I described the "straightforward" method only to show you that there is nothing subtle or mysterious about transforming to spherical coordinates: you're expressing the *same quantity* (gradient, divergence, or whatever) in different notation, that's all.

Here, then, are the vector derivatives in spherical coordinates:

Gradient:

$$\nabla T = \frac{\partial T}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial T}{\partial \theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial T}{\partial \phi}\hat{\boldsymbol{\phi}}. \tag{1.70}$$

Divergence:

$$\nabla\cdot\mathbf{v} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 v_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\, v_\theta) + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial\phi}. \tag{1.71}$$

Curl:

$$\begin{aligned}\nabla\times\mathbf{v} &= \frac{1}{r\sin\theta}\left[\frac{\partial}{\partial\theta}(\sin\theta\, v_\phi) - \frac{\partial v_\theta}{\partial\phi}\right]\hat{\mathbf{r}} + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{\partial}{\partial r}(r v_\phi)\right]\hat{\boldsymbol{\theta}} \\ &+ \frac{1}{r}\left[\frac{\partial}{\partial r}(r v_\theta) - \frac{\partial v_r}{\partial\theta}\right]\hat{\boldsymbol{\phi}}.\end{aligned} \tag{1.72}$$

Laplacian:

$$\nabla^2 T = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial T}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial T}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 T}{\partial\phi^2}. \tag{1.73}$$

For reference, these formulas are listed inside the front cover.

Problem 1.36 Find formulas for r, θ, ϕ in terms of x, y, z (the inverse, in other words, of Eq. 1.62).

• **Problem 1.37** Express the unit vectors $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$ in terms of $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ (that is, derive Eq. 1.64). Check your answers several ways ($\hat{\mathbf{r}}\cdot\hat{\mathbf{r}} \stackrel{?}{=} 1$, $\hat{\boldsymbol{\theta}}\cdot\hat{\boldsymbol{\phi}} \stackrel{?}{=} 0$, $\hat{\mathbf{r}}\times\hat{\boldsymbol{\theta}} \stackrel{?}{=} \hat{\boldsymbol{\phi}}, \ldots$). Also work out the inverse formulas, giving $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ in terms of $\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$ (and θ, ϕ).

• **Problem 1.38**

(a) Check the divergence theorem for the function $\mathbf{v}_1 = r^2\hat{\mathbf{r}}$, using as your volume the sphere of radius R, centered at the origin.

b) Do the same for $\mathbf{v}_2 = (1/r^2)\hat{\mathbf{r}}$. (If the answer surprises you, look back at Prob. 1.16.)

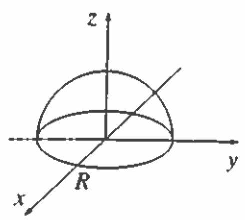

Figure 1.40

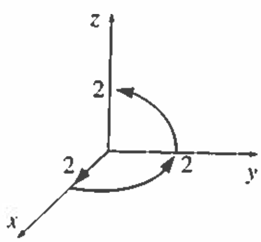

Figure 1.41

Problem 1.39 Compute the divergence of the function

$$\mathbf{v} = (r\cos\theta)\,\hat{\mathbf{r}} + (r\sin\theta)\,\hat{\boldsymbol{\theta}} + (r\sin\theta\cos\phi)\,\hat{\boldsymbol{\phi}}.$$

Check the divergence theorem for this function, using as your volume the inverted hemispherical bowl of radius R, resting on the xy plane and centered at the origin (Fig. 1.40).

Problem 1.40 Compute the gradient and Laplacian of the function $T = r(\cos\theta + \sin\theta\cos\phi)$. Check the Laplacian by converting T to Cartesian coordinates and using Eq. 1.42. Test the gradient theorem for this function, using the path shown in Fig. 1.41, from (0, 0, 0) to (0, 0, 2).

1.4.2 Cylindrical Coordinates

The cylindrical coordinates (s, ϕ, z) of a point P are defined in Fig. 1.42. Notice that ϕ has the same meaning as in spherical coordinates, and z is the same as Cartesian; s is the distance to P *from the z axis*, whereas the spherical coordinate r is the distance from the *origin*. The relation to Cartesian coordinates is

$$x = s\cos\phi, \qquad y = s\sin\phi, \qquad z = z. \tag{1.74}$$

The unit vectors (Prob. 1.41) are

$$\left.\begin{array}{lcl} \hat{\mathbf{s}} & = & \cos\phi\,\hat{\mathbf{x}} + \sin\phi\,\hat{\mathbf{y}}, \\ \hat{\boldsymbol{\phi}} & = & -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}}, \\ \hat{\mathbf{z}} & = & \hat{\mathbf{z}}. \end{array}\right\} \tag{1.75}$$

The infinitesimal displacements are

$$dl_s = ds, \qquad dl_\phi = s\,d\phi, \qquad dl_z = dz, \tag{1.76}$$

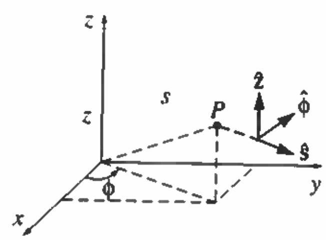

Figure 1.42

so

$$d\mathbf{l} = ds\,\hat{\mathbf{s}} + s\,d\phi\,\hat{\boldsymbol{\phi}} + dz\,\hat{\mathbf{z}}, \tag{1.77}$$

and the volume element is

$$d\tau = s\,ds\,d\phi\,dz. \tag{1.78}$$

The range of s is $0 \to \infty$, ϕ goes from $0 \to 2\pi$, and z from $-\infty$ to ∞.

The vector derivatives in cylindrical coordinates are:

Gradient:

$$\nabla T = \frac{\partial T}{\partial s}\,\hat{\mathbf{s}} + \frac{1}{s}\frac{\partial T}{\partial \phi}\,\hat{\boldsymbol{\phi}} + \frac{\partial T}{\partial z}\,\hat{\mathbf{z}}. \tag{1.79}$$

Divergence:

$$\nabla \cdot \mathbf{v} = \frac{1}{s}\frac{\partial}{\partial s}(s v_s) + \frac{1}{s}\frac{\partial v_\phi}{\partial \phi} + \frac{\partial v_z}{\partial z}. \tag{1.80}$$

Curl:

$$\nabla \times \mathbf{v} = \left(\frac{1}{s}\frac{\partial v_z}{\partial \phi} - \frac{\partial v_\phi}{\partial z}\right)\hat{\mathbf{s}} + \left(\frac{\partial v_s}{\partial z} - \frac{\partial v_z}{\partial s}\right)\hat{\boldsymbol{\phi}} + \frac{1}{s}\left[\frac{\partial}{\partial s}(s v_\phi) - \frac{\partial v_s}{\partial \phi}\right]\hat{\mathbf{z}}. \tag{1.81}$$

Laplacian:

$$\nabla^2 T = \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial T}{\partial s}\right) + \frac{1}{s^2}\frac{\partial^2 T}{\partial \phi^2} + \frac{\partial^2 T}{\partial z^2}. \tag{1.82}$$

These formulas are also listed inside the front cover.

Problem 1.41 Express the cylindrical unit vectors $\hat{\mathbf{s}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{z}}$ in terms of $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ (that is, derive Eq. 1.75). "Invert" your formulas to get $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ in terms of $\hat{\mathbf{s}}, \hat{\boldsymbol{\phi}}, \hat{\mathbf{z}}$ (and ϕ).

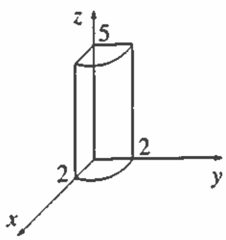

Figure 1.43

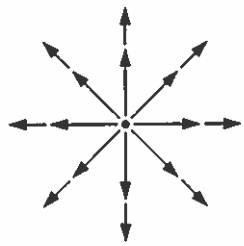

Figure 1.44

Problem 1.42

(a) Find the divergence of the function

$$\mathbf{v} = s(2+\sin^2\phi)\,\hat{\mathbf{s}} + s\sin\phi\cos\phi\,\hat{\boldsymbol{\phi}} + 3z\,\hat{\mathbf{z}}.$$

(b) Test the divergence theorem for this function, using the quarter-cylinder (radius 2, height 5) shown in Fig. 1.43.

(c) Find the curl of v.

1.5 The Dirac Delta Function

1.5.1 The Divergence of $\hat{\mathbf{r}}/r^2$

Consider the vector function

$$\mathbf{v} = \frac{1}{r^2}\,\hat{\mathbf{r}}. \tag{1.83}$$

At every location, v is directed radially outward (Fig. 1.44); if ever there was a function that ought to have a large positive divergence, this is it. And yet, when you actually *calculate* the divergence (using Eq. 1.71), you get precisely *zero*:

$$\nabla\cdot\mathbf{v} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{1}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}(1) = 0. \tag{1.84}$$

(You will have encountered this paradox already, if you worked Prob. 1.16.) The plot thickens if you apply the divergence theorem to this function. Suppose we integrate over a sphere of radius R, centered at the origin (Prob. 1.38b); the surface integral is

$$\begin{aligned}\oint \mathbf{v}\cdot d\mathbf{a} &= \int\left(\frac{1}{R^2}\hat{\mathbf{r}}\right)\cdot(R^2\sin\theta\,d\theta\,d\phi\,\hat{\mathbf{r}})\\ &= \left(\int_0^{\pi}\sin\theta\,d\theta\right)\left(\int_0^{2\pi}d\phi\right) = 4\pi.\end{aligned} \tag{1.85}$$