1
Vector Algebra

1.1 Vectors and scalars

This book is concerned with the mathematical description of physical quantities. These physical quantities include vectors and scalars, which are defined below.

1.1.1 Definition of a vector and a scalar

A *vector* is a physical quantity which has both magnitude and direction. There are many examples of such quantities, including velocity, force and electric field. A *scalar* is a physical quantity which has magnitude only. Examples of scalars include mass, temperature and pressure.

In this book, vectors will be written in bold italic type (for example, $\boldsymbol{u}$ is a vector) while scalar quantities will be written in plain italic type (for example, a is a scalar). There are two other commonly used ways of denoting vectors which are more convenient when writing by hand: an arrow over the symbol ($\vec{u}$) or a line under the symbol ($\underline{u}$).

Vectors can be represented diagrammatically by a line with an arrow at the end, as shown in Figure 1.1. The length of the line shows the magnitude of the vector and the arrow indicates its direction. If the vector has magnitude one, it is said to be a *unit vector*. Two vectors are said to be equal if they have the same magnitude and the same direction.

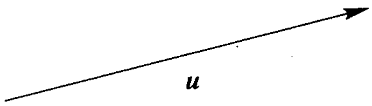


Fig. 1.1. Representation of a vector.

Example 1.1

Classify the following quantities according to whether they are vectors or scalars: energy, electric charge, electric current.

Energy and electric charge are scalars since there is no direction associated with them. Electric current is a vector because it flows in a particular direction.

1.1.2 Addition of vectors

Two vector quantities can be added together by the 'triangle rule' as shown in Figure 1.2. The vector $\boldsymbol{a}+\boldsymbol{b}$ is obtained by drawing the vector $\boldsymbol{a}$ and then drawing the vector $\boldsymbol{b}$ starting from the arrow at the end of $\boldsymbol{a}$.

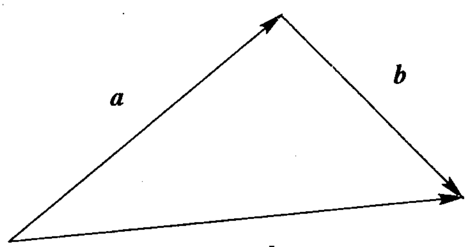


Fig. 1.2. Addition of vectors.

The vector $-\boldsymbol{a}$ is defined as the vector with magnitude equal to that of $\boldsymbol{a}$ but pointing in the opposite direction.

By adding $\boldsymbol{a}$ and $-\boldsymbol{a}$ we obtain the *zero vector*, $\mathbf{0}$. This has magnitude zero and so does not have a direction; nevertheless it is sensible to regard $\mathbf{0}$ as a vector.

1.1.3 Components of a vector

Vectors are often written using a Cartesian coordinate system with axes x, y, z. Such a system is usually assumed to be *right-handed*, which means that a screw rotated from the x-axis to the y-axis would move in the direction of the z-axis. Alternatively, if the thumb of the right hand points in the x direction and the first finger in the y direction, then the second finger points in the z direction.

Suppose that a vector $\boldsymbol{a}$ is drawn in a Cartesian coordinate system and extends from the point $(x_1,\ y_1,\ z_1)$ to the point $(x_2,\ y_2,\ z_2)$, as shown in Figure 1.3. Then the *components* of the vector are defined to be the three numbers $a_1 = x_2 - x_1$, $a_2 = y_2 - y_1$ and $a_3 = z_2 - z_1$. The vector can then be written in the form $\boldsymbol{a} = (a_1,\ a_2,\ a_3)$.

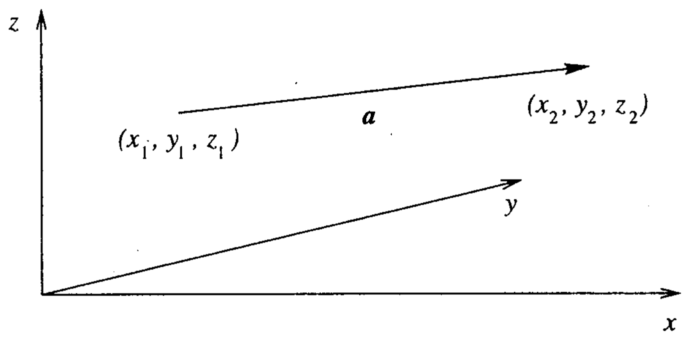


Fig. 1.3. The components of the vector $\boldsymbol{a}$ are $(x_2 - x_1, y_2 - y_1, z_2 - z_1)$.

By introducing three unit vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, which point along the coordinate axes x, y and z respectively, the vector can also be written in the form $\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3$. Using this form, the sum of the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is $\boldsymbol{a}+\boldsymbol{b} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3 + b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + b_3\boldsymbol{e}_3 = (a_1+b_1)\boldsymbol{e}_1 + (a_2 + b_2)\boldsymbol{e}_2 + (a_3+b_3)\boldsymbol{e}_3$. It follows that vectors can be added simply by adding their components, so that the vector equation $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$ is equivalent to the three equations $c_1 = a_1 + b_1$, $c_2 = a_2 + b_2$, $c_3 = a_3 + b_3$.

The magnitude of the vector is written $|\boldsymbol{a}|$. It can be deduced from Pythagoras's theorem that the magnitude of the vector can be written in terms of its components as $|\boldsymbol{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$.

The position of a point in space (x, y, z) defines a vector which points from the origin of the coordinate system to the point (x, y, z). This vector is called the *position vector* of the point, and is usually denoted by the symbol $\boldsymbol{r}$, with components given by $\boldsymbol{r} = (x, y, z)$.

Example 1.2

The vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are defined by $\boldsymbol{a} = (1, 1, 1)$, $\boldsymbol{b} = (1, 2, 2)$. Find the magnitudes of $\boldsymbol{a}$ and $\boldsymbol{b}$, and find the vectors $\boldsymbol{a} + \boldsymbol{b}$ and $\boldsymbol{a} - \boldsymbol{b}$.

The magnitude of the vector $\boldsymbol{a}$ is $|\boldsymbol{a}| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$. The magnitude of $\boldsymbol{b}$ is $|\boldsymbol{b}| = \sqrt{1^2 + 2^2 + 2^2} = 3$. The vector $\boldsymbol{a}+\boldsymbol{b}$ is $(1, 1, 1) + (1, 2, 2) = (2, 3, 3)$ and $\boldsymbol{a} - \boldsymbol{b} = (0, -1, -1)$.

1.2 Dot product

The *dot product* or *scalar product* of two vectors is a scalar quantity. It is written $\boldsymbol{a} \cdot \boldsymbol{b}$ and is defined as the product of the magnitudes of the two vectors and the cosine of the angle between them:

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta. \tag{1.1}$$

A number of properties of the dot product follow from this definition:

- The dot product is commutative, i.e. $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$.
- If the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular (orthogonal) then $\boldsymbol{a} \cdot \boldsymbol{b} = 0$.
- Conversely, if $\boldsymbol{a} \cdot \boldsymbol{b} = 0$ then either the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular or one of the vectors is the zero vector.
- $\boldsymbol{a} \cdot \boldsymbol{a} = |\boldsymbol{a}|^2$.
- Since the quantity $|\boldsymbol{b}| \cos\theta$ represents the component of the vector $\boldsymbol{b}$ in the direction of the vector $\boldsymbol{a}$, the scalar $\boldsymbol{a} \cdot \boldsymbol{b}$ can be thought of as the magnitude of $\boldsymbol{a}$ multiplied by the component of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$ (see Figure 1.4).
- The dot product is distributive over addition, i.e. $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$. This follows geometrically from the fact that the component of $\boldsymbol{b} + \boldsymbol{c}$ in the direction of $\boldsymbol{a}$ is the same as the component of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$ plus the component of $\boldsymbol{c}$ in the direction of $\boldsymbol{a}$ (see Figure 1.5).

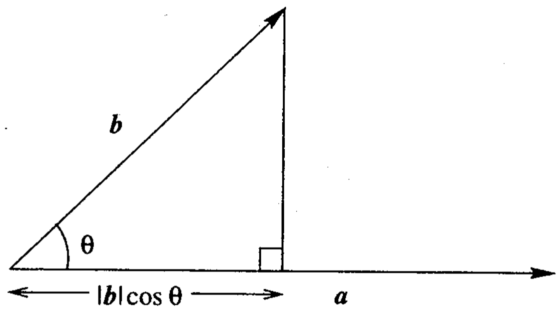


Fig. 1.4. The component of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$ is $|\boldsymbol{b}| \cos\theta$.

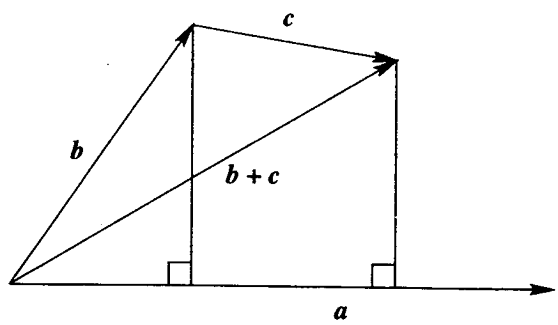


Fig. 1.5. Geometrical demonstration that the dot product is distributive over addition.

A formula for the dot product $\boldsymbol{a} \cdot \boldsymbol{b}$ in terms of the components of the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can be derived from the above properties. Considering first the unit vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, it follows from the fact that these vectors have magnitude 1 and are orthogonal to each other that

$$\boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = 1,\ \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = 1,\ \boldsymbol{e}_3 \cdot \boldsymbol{e}_3 = 1,\ \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = 0,\ \boldsymbol{e}_2 \cdot \boldsymbol{e}_3 = 0,\ \boldsymbol{e}_3 \cdot \boldsymbol{e}_1 = 0.$$

The dot product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is therefore

$$\begin{aligned} \boldsymbol{a} \cdot \boldsymbol{b} &= (a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3) \cdot (b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + b_3\boldsymbol{e}_3) \\ &= a_1b_1\boldsymbol{e}_1 \cdot \boldsymbol{e}_1 + a_2b_2\boldsymbol{e}_2 \cdot \boldsymbol{e}_2 + a_3b_3\boldsymbol{e}_3 \cdot \boldsymbol{e}_3 \\ &= a_1b_1 + a_2b_2 + a_3b_3. \qquad (1.2) \end{aligned}$$

Example 1.3

Find the dot product of the vectors $(1,1,2)$ and $(2,3,2)$.
$(1,1,2) \cdot (2,3,2) = 1 \times 2 + 1 \times 3 + 2 \times 2 = 9$.

Example 1.4

For what value of c are the vectors $(c,1,1)$ and $(-1,2,0)$ perpendicular?

They are perpendicular when their dot product is zero. The dot product is $-c + 2 + 0$ so the vectors are perpendicular if $c = 2$.

Example 1.5

Show that a triangle inscribed in a circle is right-angled if one of the sides of the triangle is a diameter of the circle.

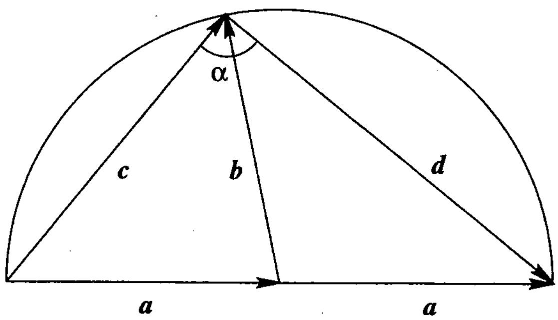


Fig. 1.6. Geometrical construction to show that α is a right angle.

Introduce two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ as shown in Figure 1.6. Since these two vectors are both along radii of the circle they are of equal magnitude. The two sides $\boldsymbol{c}$ and $\boldsymbol{d}$ of the triangle are then given by $\boldsymbol{c} = \boldsymbol{a}+\boldsymbol{b}$ and $\boldsymbol{d} = \boldsymbol{a}-\boldsymbol{b}$. The dot product of these two vectors is $\boldsymbol{c}\cdot\boldsymbol{d} = (\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b}) = |\boldsymbol{a}|^2-\boldsymbol{a}\cdot\boldsymbol{b}+\boldsymbol{b}\cdot\boldsymbol{a}-|\boldsymbol{b}|^2 = 0$. Since the dot product is zero the vectors are perpendicular, so the angle α is a right angle. This is just one of many geometrical results that can be obtained using vector methods.

1.2.1 Applications of the dot product

Work done against a force

Suppose that a constant force $\boldsymbol{F}$ acts on a body and that the body is moved a distance $\boldsymbol{d}$. Then the work done against the force is given by the magnitude of the force times the distance moved in the direction opposite to the force; this is simply $-\boldsymbol{F}\cdot\boldsymbol{d}$ (Figure 1.7).

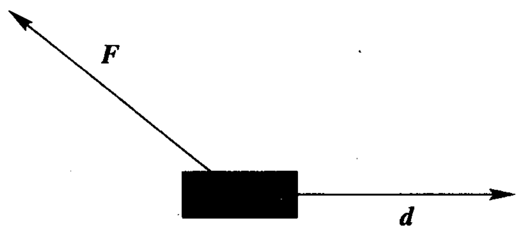


Fig. 1.7. The work done against a force $\boldsymbol{F}$ when an object is moved a distance $\boldsymbol{d}$ is $-\boldsymbol{F}\cdot\boldsymbol{d}$.

Equation of a plane

Consider a two-dimensional plane in three-dimensional space (Figure 1.8). Let $\boldsymbol{r}$ be the position vector of any point in the plane, and let $\boldsymbol{a}$ be a vector perpendicular to the plane. The condition for a point with position vector $\boldsymbol{r}$ to lie in the plane is that the component of $\boldsymbol{r}$ in the direction of $\boldsymbol{a}$ is equal to the perpendicular distance p from the origin to the plane. The general form of the equation of a plane is therefore

$$\boldsymbol{r}\cdot\boldsymbol{a} = \text{constant}.$$

An alternative way to write this is in terms of components. Writing $\boldsymbol{r} = (x, y, z)$ and $\boldsymbol{a} = (a_1, a_2, a_3)$, the equation of a plane becomes

$$a_1x + a_2y + a_3z = \text{constant}. \tag{1.3}$$

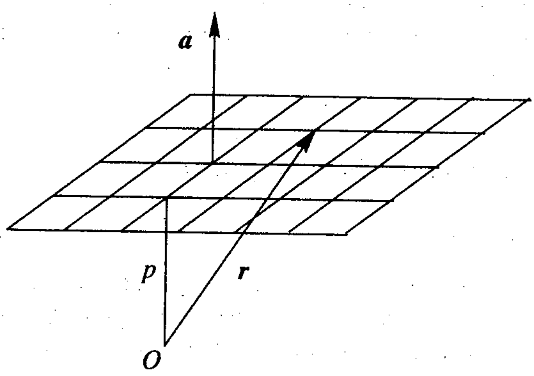


Fig. 1.8. The equation of a plane is $\boldsymbol{r} \cdot \boldsymbol{a} =$ constant.

EXERCISES

1.1 Classify the following quantities according to whether they are vectors or scalars: density, magnetic field strength, power, momentum, angular momentum, acceleration.

1.2 If $\boldsymbol{a} = (2, 0, 3)$ and $\boldsymbol{b} = (1, 0, -1)$, find $|\boldsymbol{a}|$, $|\boldsymbol{b}|$, $\boldsymbol{a} + \boldsymbol{b}$, $\boldsymbol{a} - \boldsymbol{b}$ and $\boldsymbol{a} \cdot \boldsymbol{b}$. What is the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$?

1.3 If $\boldsymbol{u} = (1, 2, 2)$ and $\boldsymbol{v} = (-6, 2, 3)$, find the component of $\boldsymbol{u}$ in the direction of $\boldsymbol{v}$ and the component of $\boldsymbol{v}$ in the direction of $\boldsymbol{u}$.

1.4 Find the equation of the plane that is perpendicular to the vector $(1, 1, -1)$ and passes through the point $x = 1$, $y = 2$, $z = 1$.

1.5 Use vector methods to show that the diagonals of a rhombus are perpendicular.

1.6 What is the angle between any two diagonals of a cube?

1.7 Use vectors to show that for any triangle, the three lines drawn from each vertex to the midpoint of the opposite side all pass through the same point.

1.3 Cross product

The *cross product* or *vector product* of two vectors is a vector quantity, written $\boldsymbol{a} \times \boldsymbol{b}$. Since it is a vector, its definition must specify both its magnitude and direction. The magnitude of $\boldsymbol{a} \times \boldsymbol{b}$ is $|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$, where θ is the angle between the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The direction of $\boldsymbol{a} \times \boldsymbol{b}$ is perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$ in a right-handed sense, i.e. a right-handed screw rotated from $\boldsymbol{a}$ towards $\boldsymbol{b}$ moves in the direction of $\boldsymbol{a} \times \boldsymbol{b}$ (Figure 1.9). We may therefore write $\boldsymbol{a} \times \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta\, \boldsymbol{u}$, where $\boldsymbol{u}$ is a unit vector perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$ in a right-handed sense.

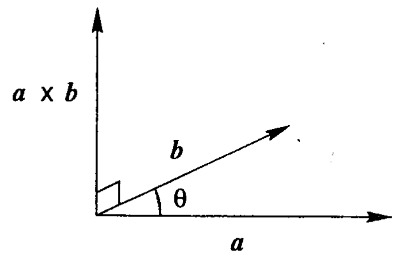


Fig. 1.9. The cross product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$, in a right-handed sense.

The cross product has the following properties:

- The cross product is *not* commutative. Because of the right-hand rule, $\boldsymbol{a} \times \boldsymbol{b}$ and $\boldsymbol{b} \times \boldsymbol{a}$ point in opposite directions, so $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$.
- If the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are parallel then $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{0}$.
- $\boldsymbol{a} \times \boldsymbol{a} = \boldsymbol{0}$.
- The magnitude of the cross product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is the area of the parallelogram made by the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ (Figure 1.10). Similarly the area of the triangle made by $\boldsymbol{a}$ and $\boldsymbol{b}$ is $|\boldsymbol{a} \times \boldsymbol{b}|/2$.
- The cross product of $\boldsymbol{a}$ and $\boldsymbol{b}$ only depends on the component of $\boldsymbol{b}$ perpendicular to $\boldsymbol{a}$. This is apparent from Figure 1.10 since the component of $\boldsymbol{b}$ perpendicular to $\boldsymbol{a}$ is $|\boldsymbol{b}| \sin\theta$.
- The cross product is distributive over addition, i.e. $\boldsymbol{a} \times (\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$. This is demonstrated geometrically in Figure 1.11, where the vector $\boldsymbol{a}$ points into the page. The vectors $\boldsymbol{b}$, $\boldsymbol{c}$ and $\boldsymbol{b}+\boldsymbol{c}$ do not necessarily lie in the page, but from the previous point the cross products of these vectors with $\boldsymbol{a}$ only depend on their projections onto the page. The effect of taking the cross

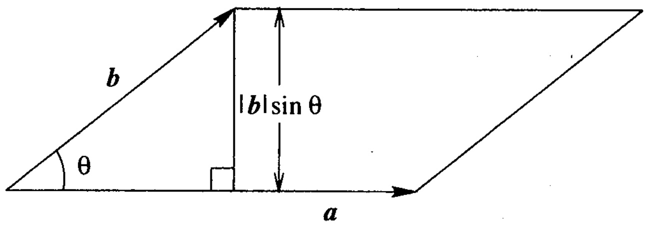


Fig. 1.10. The area of the parallelogram is the length of its base, $|\boldsymbol{a}|$, multiplied by its height, $|\boldsymbol{b}|\sin\theta$.

product with $\boldsymbol{a}$ on any vector is to project it onto the page, rotate through $\pi/2$ clockwise and then multiply by $|\boldsymbol{a}|$. Thus the triangle made by the vectors $\boldsymbol{b}$, $\boldsymbol{c}$ and $\boldsymbol{b}+\boldsymbol{c}$ becomes rotated and scaled as in Figure 1.11 but remains a triangle.

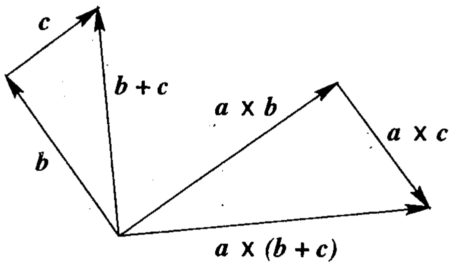


Fig. 1.11. Geometrical demonstration that the cross product is distributive over addition. The vector $\boldsymbol{a}$ points into the page.

A formula for the cross product $\boldsymbol{a}\times\boldsymbol{b}$ in terms of the components of the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can be derived in a similar manner to that carried out for the dot product. Consider first $\boldsymbol{e}_1 \times \boldsymbol{e}_2$. Since these two vectors have magnitude 1 and are perpendicular, $\sin\theta = 1$ and the magnitude of $\boldsymbol{e}_1 \times \boldsymbol{e}_2$ is 1. The direction of $\boldsymbol{e}_1 \times \boldsymbol{e}_2$ is perpendicular to both $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ in a right-handed sense, so $\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3$.

It follows that the unit vectors $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ obey

$$\boldsymbol{e}_1\times\boldsymbol{e}_1=\boldsymbol{0},\ \boldsymbol{e}_2\times\boldsymbol{e}_2=\boldsymbol{0},\ \boldsymbol{e}_3\times\boldsymbol{e}_3=\boldsymbol{0},\ \boldsymbol{e}_1\times\boldsymbol{e}_2=\boldsymbol{e}_3,\ \boldsymbol{e}_2\times\boldsymbol{e}_3=\boldsymbol{e}_1,\ \boldsymbol{e}_3\times\boldsymbol{e}_1=\boldsymbol{e}_2.$$

The cross product of $\boldsymbol{a}$ and $\boldsymbol{b}$ is therefore

$$\begin{aligned}
\boldsymbol{a} \times \boldsymbol{b} &= (a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3) \times (b_1\boldsymbol{e}_1 + b_2\boldsymbol{e}_2 + b_3\boldsymbol{e}_3) \\
&= a_1b_2\,\boldsymbol{e}_1 \times \boldsymbol{e}_2 + a_1b_3\,\boldsymbol{e}_1 \times \boldsymbol{e}_3 + a_2b_1\,\boldsymbol{e}_2 \times \boldsymbol{e}_1 \\
&\quad + a_2b_3\,\boldsymbol{e}_2 \times \boldsymbol{e}_3 + a_3b_1\,\boldsymbol{e}_3 \times \boldsymbol{e}_1 + a_3b_2\,\boldsymbol{e}_3 \times \boldsymbol{e}_2 \\
&= (a_2b_3 - a_3b_2)\boldsymbol{e}_1 + (a_3b_1 - a_1b_3)\boldsymbol{e}_2 + (a_1b_2 - a_2b_1)\boldsymbol{e}_3. \qquad (1.4)
\end{aligned}$$

This can also be written as the determinant of a 3×3 matrix as follows:

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Example 1.6

Find the cross product of the vectors $(1,3,0)$ and $(2,-1,1)$.

$(1,3,0) \times (2,-1,1) = (3-0, 0-1, -1-6) = (3,-1,-7)$.

Example 1.7

Find a unit vector which is perpendicular to both $(1,0,1)$ and $(0,1,1)$.

A perpendicular vector is $(1,0,1) \times (0,1,1) = (-1,-1,1)$. To make this a unit vector we must divide by its magnitude, which is $\sqrt{3}$, so the unit vector perpendicular to $(1,0,1)$ and $(0,1,1)$ is $(-1,-1,1)/\sqrt{3}$.

Example 1.8

What is the area of the triangle which has its vertices at the points $P = (1,1,1)$, $Q = (2,3,3)$ and $R = (4,1,2)$?

First construct two vectors that make up two sides of the triangle. The vector from P to Q is $\boldsymbol{a} = (1,2,2)$ and the vector from P to R is $\boldsymbol{b} = (3,0,1)$. The cross product of these vectors is $\boldsymbol{a} \times \boldsymbol{b} = (2,5,-6)$. The area of the triangle is then $|\boldsymbol{a} \times \boldsymbol{b}|/2 = \sqrt{65}/2 \approx 4.03$.

1.3.1 Applications of the cross product

Solid body rotation

Suppose that a solid body is rotating steadily about an axis. What is the velocity vector of a point within the body?

Consider a body rotating with angular velocity Ω (this means that in a time t the body rotates through an angle Ωt radians). Since there is a rotation axis, a vector $\boldsymbol{\Omega}$ can be defined, with magnitude $|\boldsymbol{\Omega}| = \Omega$ and directed along the rotation axis. Since this vector could point in either direction, the following form of the right-hand rule is used to define the direction of $\boldsymbol{\Omega}$: a screw rotating in the same direction as the body moves in the direction of $\boldsymbol{\Omega}$. Alternatively, if the fingers of the right hand point in the direction of the rotation, the thumb of the right hand points in the direction of $\boldsymbol{\Omega}$. This means that for a body which is rotating to the right, $\boldsymbol{\Omega}$ points upwards (Figure 1.12).

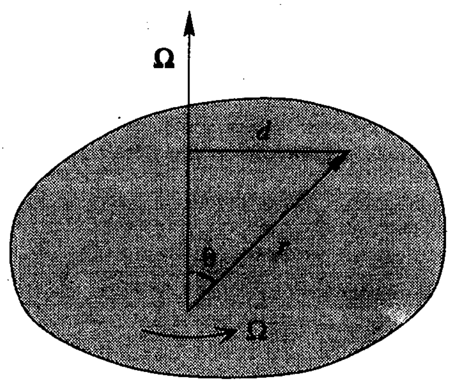


Fig. 1.12. Motion of a rotating body.

Now consider the motion of a point at a position vector $\boldsymbol{r}$, which makes an angle θ with the rotation axis. The speed at which this point moves is Ωd, where d is the perpendicular distance from the point to the rotation axis. Since $d = |\boldsymbol{r}| \sin\theta$ (Figure 1.12), the speed of motion is $v = \Omega|\boldsymbol{r}|\sin\theta$. Note that this is equal to $|\boldsymbol{\Omega} \times \boldsymbol{r}|$. Now consider the direction of the motion. In Figure 1.12, where both $\boldsymbol{\Omega}$ and $\boldsymbol{r}$ lie in the plane of the page, the direction of motion is into the page, perpendicular to both $\boldsymbol{\Omega}$ and $\boldsymbol{r}$ and so in the direction of $\boldsymbol{\Omega} \times \boldsymbol{r}$. Therefore the velocity vector of the point at $\boldsymbol{r}$ is

$$\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{r}, \tag{1.5}$$

since this vector has both the correct magnitude and the correct direction.

Equation of a straight line

The equation of a straight line can be written in terms of the cross product as follows. Suppose that $\boldsymbol{a}$ is the position vector of a particular fixed point on the line, and that $\boldsymbol{u}$ is a vector pointing along the line (Figure 1.13). Then any point $\boldsymbol{r}$ on the line can be reached from the origin by travelling first along the vector $\boldsymbol{a}$ onto the line and then some multiple of the vector $\boldsymbol{u}$ along the line:

$$\boldsymbol{r} = \boldsymbol{a} + \lambda\boldsymbol{u}, \tag{1.6}$$

where λ is a parameter. This is referred to as the *parametric* form of the equation of a line.

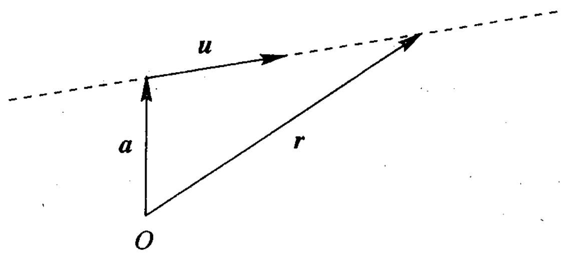


Fig. 1.13. The equation of a line is $\boldsymbol{r} = \boldsymbol{a} + \lambda\boldsymbol{u}$.

To obtain a form of (1.6) that does not involve the parameter λ, the term involving the vector $\boldsymbol{u}$ must be eliminated. This can be done by taking the cross product of (1.6) with $\boldsymbol{u}$. This gives $\boldsymbol{r} \times \boldsymbol{u} = \boldsymbol{a} \times \boldsymbol{u}$. Since the vector $\boldsymbol{a} \times \boldsymbol{u}$ is a constant, it can be relabelled $\boldsymbol{b}$, giving the second form for the equation of a straight line:

$$\boldsymbol{r} \times \boldsymbol{u} = \boldsymbol{b}. \tag{1.7}$$

Physical applications of the cross product

There are many physical quantities that are defined in terms of the cross product. These include the following:

- A particle of mass m has position vector $\boldsymbol{r}$ and is moving with velocity $\boldsymbol{v}$. Its angular momentum about the origin is $\boldsymbol{h} = m\,\boldsymbol{r} \times \boldsymbol{v}$.
- A particle of mass m moves with velocity $\boldsymbol{u}$ in a frame which is rotating with angular velocity $\boldsymbol{\Omega}$. Due to the rotation, the particle experiences a sideways force called the Coriolis force, $\boldsymbol{F} = 2m\,\boldsymbol{u} \times \boldsymbol{\Omega}$. Since the Earth is rotating, this force influences motion on the surface of the Earth. The effect deflects particles to the right in the northern hemisphere and is strongest for motions on large scales such as ocean currents and weather systems.
- A particle with electric charge q moves with velocity $\boldsymbol{v}$ in the presence of a magnetic field $\boldsymbol{B}$. This results in a force, called the Lorentz force, equal to $q\,\boldsymbol{v} \times \boldsymbol{B}$. This is the force which is responsible for the operation of an electric motor.

1.4 Scalar triple product

The *scalar triple product* of three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ is defined to be $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c})$. In fact the brackets here are unnecessary: $(\boldsymbol{a} \cdot \boldsymbol{b}) \times \boldsymbol{c}$ is meaningless since $(\boldsymbol{a} \cdot \boldsymbol{b})$ is a scalar and so cannot be crossed with the vector $\boldsymbol{c}$. Therefore the expression $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c}$ is well defined.

The formula for the scalar triple product in terms of the components of the three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ can be obtained using the formula for the cross product (1.4):

$$\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1. \tag{1.8}$$

The scalar triple product has a number of properties, listed below. The first four follow directly from (1.8).

- The dot and the cross can be interchanged:

$$\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c}.$$

- The vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ can be permuted cyclically:

$$\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{b} \cdot \boldsymbol{c} \times \boldsymbol{a} = \boldsymbol{c} \cdot \boldsymbol{a} \times \boldsymbol{b}.$$

- The scalar triple product can be written in the form of a determinant:

$$\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

- If any two of the vectors are equal, the scalar triple product is zero.
- Geometrically, the magnitude of the scalar triple product is the volume of the three-dimensional object known as a parallelepiped formed by the three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ (Figure 1.14). This can be shown as follows. The area of the parallelogram forming the base is $|\boldsymbol{b} \times \boldsymbol{c}|$. The height is the vertical component of $\boldsymbol{a}$, which is the magnitude of the component of $\boldsymbol{a}$ in the direction of $\boldsymbol{b} \times \boldsymbol{c}$. This is $|\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c}|/|\boldsymbol{b} \times \boldsymbol{c}|$, so the volume is the area of the base multiplied by the height, which is $|\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c}|$. Similarly, the volume of the tetrahedron made by the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ is $|\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c}|/6$.

The scalar triple product of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ is often written $[\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$. This notation highlights the fact that the dot and the cross can be interchanged.

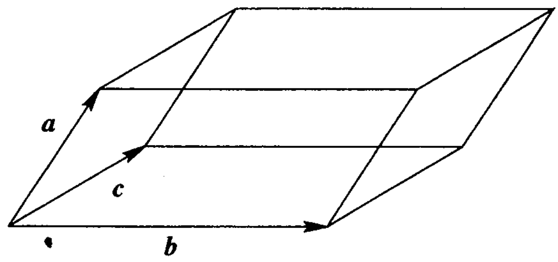


Fig. 1.14. The volume of the object formed by the three vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ is $|\boldsymbol{a}\cdot\boldsymbol{b}\times\boldsymbol{c}|$.

Example 1.9

Find the scalar triple product of the vectors (1,2,1), (0,1,1) and (2,1,0).

First find the vector $(0,1,1) \times (2,1,0) = (-1,2,-2)$. Now dot this with $(1,2,1)$, giving the answer 1.

Example 1.10

Show that if three vectors lie in a plane, then their scalar triple product is zero.

If $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ lie in a plane, then the vector $\boldsymbol{b} \times \boldsymbol{c}$ is perpendicular to the plane and hence perpendicular to $\boldsymbol{a}$. Since the dot product of perpendicular vectors is always zero, it follows that $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = 0$.

Example 1.11

A particle with mass m and electric charge q moves in a uniform magnetic field $\boldsymbol{B}$. Given that the force $\boldsymbol{F}$ on the particle is $\boldsymbol{F} = q\,\boldsymbol{v} \times \boldsymbol{B}$, where $\boldsymbol{v}$ is the velocity of the particle, show that the particle moves at constant speed.

The equation of motion of the particle is written using Newton's second law, force equals mass times acceleration. The acceleration of the particle is the rate of change of the velocity, written $\dot{\boldsymbol{v}}$, so the equation of motion is

$$q\,\boldsymbol{v} \times \boldsymbol{B} = m\,\dot{\boldsymbol{v}}.$$

Now taking the dot product of both sides of this equation with $\boldsymbol{v}$, the scalar triple product on the left-hand side gives zero since two of the vectors are equal. Hence

$$0 = m\,\dot{\boldsymbol{v}} \cdot \boldsymbol{v} = m\frac{d}{dt}(\boldsymbol{v} \cdot \boldsymbol{v})/2 = m\frac{d}{dt}\left(|\boldsymbol{v}|^2\right)/2,$$

so the speed of the particle, $|\boldsymbol{v}|$, does not change with time.

1.5 Vector triple product

The *vector triple product* of three vectors is $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$. The brackets are important here, since $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) \neq (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$. Since only cross products are involved, the result is a vector. An alternative expression for $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ can be obtained by writing out the components. Since

$$\boldsymbol{b} \times \boldsymbol{c} = (b_2c_3 - b_3c_2)\boldsymbol{e}_1 + (b_3c_1 - b_1c_3)\boldsymbol{e}_2 + (b_1c_2 - b_2c_1)\boldsymbol{e}_3,$$

the first component of $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ is

$$\begin{aligned} [\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})]_1 &= a_2(b_1c_2 - b_2c_1) - a_3(b_3c_1 - b_1c_3) \\ &= b_1(a_2c_2 + a_3c_3) - c_1(a_2b_2 + a_3b_3). \end{aligned}$$

By adding and subtracting the quantity $a_1b_1c_1$, this can be written

$$\begin{aligned} [\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})]_1 &= b_1(a_1c_1 + a_2c_2 + a_3c_3) - c_1(a_1b_1 + a_2b_2 + a_3b_3) \\ &= b_1\boldsymbol{a} \cdot \boldsymbol{c} - c_1\boldsymbol{a} \cdot \boldsymbol{b}. \end{aligned}$$

Similar equations hold for the second and third components, so the vector triple product can be expanded as

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}. \tag{1.9}$$

From this result it also follows that

$$(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = -\boldsymbol{c} \times (\boldsymbol{a} \times \boldsymbol{b}) = -(\boldsymbol{c} \cdot \boldsymbol{b})\boldsymbol{a} + (\boldsymbol{c} \cdot \boldsymbol{a})\boldsymbol{b}. \tag{1.10}$$

Example 1.12

Under what conditions are $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c})$ and $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}$ equal?

By comparing (1.9) with (1.10), the two are equal if $-(\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c} = -(\boldsymbol{c} \cdot \boldsymbol{b})\boldsymbol{a}$. This can alternatively be written $\boldsymbol{b} \times (\boldsymbol{a} \times \boldsymbol{c}) = \boldsymbol{0}$.

Example 1.13

Find an alternative expression for $(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d})$.

Since the dot and cross can be interchanged in a scalar triple product,

$$\begin{aligned} (\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{c} \times \boldsymbol{d}) &= \boldsymbol{a} \cdot (\boldsymbol{b} \times (\boldsymbol{c} \times \boldsymbol{d})) \\ &= \boldsymbol{a} \cdot ((\boldsymbol{b} \cdot \boldsymbol{d})\boldsymbol{c} - (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{d}) \\ &= (\boldsymbol{a} \cdot \boldsymbol{c})(\boldsymbol{b} \cdot \boldsymbol{d}) - (\boldsymbol{a} \cdot \boldsymbol{d})(\boldsymbol{b} \cdot \boldsymbol{c}). \end{aligned}$$

1.6 Scalar fields and vector fields

A scalar or vector quantity is said be a *field* if it is a function of position. An example of a scalar field is the temperature inside a room; in general the temperature has a different value at different points in space, so the temperature T is a function of position. This is indicated by writing $T(\boldsymbol{r})$, where $\boldsymbol{r}$ is the position vector of a point in space, $\boldsymbol{r} = (x, y, z)$. Other examples of scalar fields include pressure and density. An example of a vector field is the velocity of the air within a room.

In general, a scalar field T is three-dimensional, i.e. it depends on all three coordinates, $T = T(x, y, z)$. Such fields are difficult to visualise. However, if the scalar field only depends on two coordinates, $T = T(x, y)$, then it can be visualised by sketching a contour plot. To do this, the line $T(x, y) =$ constant is plotted for different values of the constant. For example, consider the scalar field $T(x, y) = x^2 + y^2$. The contour lines are the lines $x^2 + y^2 =$ constant, which are concentric circles centred at the origin, as shown in Figure 1.15(a).

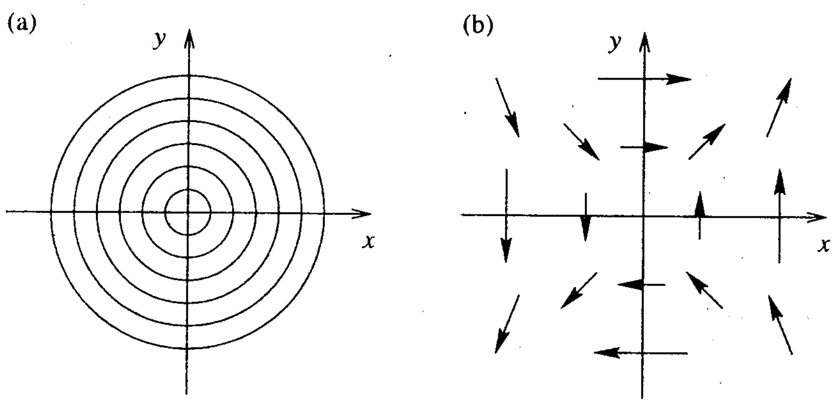


Fig. 1.15. (a) Contours of the scalar field $T(x, y) = x^2 + y^2$. (b) The vector field $\boldsymbol{u}(x, y) = (y, x)$.

Vector fields in two dimensions can also be visualised by a sketch. In this case the simplest procedure is to evaluate the vector field at a sequence of points and draw vectors indicating the magnitude and direction of the vector field at each point. An example of this procedure is the drawing of wind speeds and directions on weather maps. For example, consider the vector field $\boldsymbol{u}(x, y) =$

(y, x). At the point $(1,0)$, $\boldsymbol{u} = (0,1)$, so at this point a vector of magnitude 1 pointing in the y direction is drawn. Similarly, at $(0,1)$, $\boldsymbol{u} = (1,0)$ and at $(1,1)$, $\boldsymbol{u} = (1,1)$. By considering a few additional points, a sketch of the vector field can be built up (Figure 1.15(b)).

Summary of Chapter 1

- A *vector* is a physical quantity with magnitude and direction.
- A *scalar* is a physical quantity with magnitude only.
- In Cartesian coordinates a vector can be written in terms of its *components* as either $\boldsymbol{a} = (a_1, a_2, a_3)$ or $\boldsymbol{a} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + a_3\boldsymbol{e}_3$, where $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are unit vectors along the x-, y- and z-axes respectively.
- The *magnitude* of the vector $\boldsymbol{a}$ is $|\boldsymbol{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}$.
- The *dot product* or *scalar product* of $\boldsymbol{a}$ and $\boldsymbol{b}$ is a scalar,

  $$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

  This can also be thought of as $|\boldsymbol{a}|$ multiplied by the component of $\boldsymbol{b}$ in the direction of $\boldsymbol{a}$. Applications of the dot product include the work done when moving an object acted on by a force and the equation of a plane.
- The *cross product* or *vector product* of $\boldsymbol{a}$ and $\boldsymbol{b}$ is a vector, $\boldsymbol{a} \times \boldsymbol{b}$, with magnitude $|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$, perpendicular to $\boldsymbol{a}$ and $\boldsymbol{b}$ in a right-handed sense. In component form,

  $$\boldsymbol{a} \times \boldsymbol{b} = (a_2 b_3 - a_3 b_2)\boldsymbol{e}_1 + (a_3 b_1 - a_1 b_3)\boldsymbol{e}_2 + (a_1 b_2 - a_2 b_1)\boldsymbol{e}_3.$$

  The magnitude of $\boldsymbol{a} \times \boldsymbol{b}$ is $|\boldsymbol{a}|$ multiplied by the component of $\boldsymbol{b}$ perpendicular to $\boldsymbol{a}$, which is the area of the parallelogram made by $\boldsymbol{a}$ and $\boldsymbol{b}$. Applications of the cross product include the equation of a straight line and the rotation of a rigid body.
- The *scalar triple product* is $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c} = \boldsymbol{b} \cdot \boldsymbol{c} \times \boldsymbol{a}$.
- The *vector triple product* is $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$.
- A scalar or vector quantity is a *field* if it is a function of position.

EXERCISES

1.8 Find the equation of the straight line which passes through the points $(1, 1, 1)$ and $(2, 3, 5)$, (a) in parametric form; (b) in cross product form.

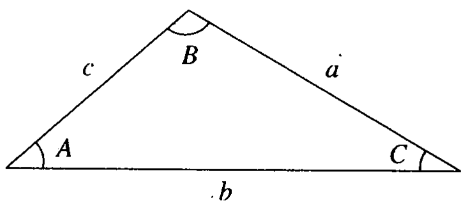
1.9 Using vector methods, prove the sine rule,

$$\frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c} \tag{1.11}$$

and the cosine rule,

$$c^2 = a^2 + b^2 - 2ab\cos C \tag{1.12}$$

for the triangle with angles A, B, C and sides a, b, c in the figure below.



1.10 (a) Show that the set of vectors and the operation of vector addition form a group. (The set of objects $a, b, c, \ldots$ and the operation $\star$ form a group if the following four conditions are satisfied: (i) for any two elements a and b, $a \star b$ is in the set; (ii) $(a \star b) \star c = a \star (b \star c)$; (iii) there is an element I obeying $a \star I = I \star a = a$; (iv) each element a has an inverse a^{-1} such that $a \star a^{-1} = a^{-1} \star a = I$.)
(b) Do the set of vectors and the dot product form a group?
(c) Do the set of vectors and the cross product form a group?

1.11 Simplify the following expressions:
(a) $|\boldsymbol{a} \times \boldsymbol{b}|^2 + (\boldsymbol{a} \cdot \boldsymbol{b})^2$;
(b) $\boldsymbol{a} \times (\boldsymbol{b} \times (\boldsymbol{a} \times \boldsymbol{b}))$;
(c) $(\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{b} - \boldsymbol{c}) \times (\boldsymbol{c} - \boldsymbol{a})$;
(d) $(\boldsymbol{a} \times \boldsymbol{b}) \cdot (\boldsymbol{b} \times \boldsymbol{c}) \times (\boldsymbol{c} \times \boldsymbol{a})$.

1.12 The vector $\boldsymbol{x}$ obeys the two equations $\boldsymbol{x} \cdot \boldsymbol{a} = 1$ and $\boldsymbol{x} \times \boldsymbol{a} = \boldsymbol{b}$, where $\boldsymbol{a}$ and $\boldsymbol{b}$ are constant vectors. Solve these equations to find an expression for $\boldsymbol{x}$ in terms of $\boldsymbol{a}$ and $\boldsymbol{b}$. Give a geometrical interpretation of this question.

1.13 Find the equation of the line on which the two planes $\boldsymbol{r} \cdot \boldsymbol{a} = 1$ and $\boldsymbol{r} \cdot \boldsymbol{b} = 1$ meet.

1.14 (a) Express the vector $\boldsymbol{a} \times \boldsymbol{b}$ in the form $\alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c}$, assuming that the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are not coplanar.

(b) Hence find an expression for $(\boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c})^2$ that does not involve any cross products.

(c) Hence find the volume of a tetrahedron made from four equilateral triangles with sides of length 1.

1.15 A particle of mass m at position $\boldsymbol{r}$ and moving with velocity $\boldsymbol{v}$ is subject to a force $\boldsymbol{F}$ directed towards the origin, $\boldsymbol{F} = -f(r)\boldsymbol{r}$. Show that the angular momentum vector $\boldsymbol{h} = m\boldsymbol{r} \times \boldsymbol{v}$ is constant.

1.16 Sketch the scalar field $T(x, y) = x^2 - y$.

1.17 Sketch the vector field $\boldsymbol{u}(x, y) = (x + y, -x)$.

2

Line, Surface and Volume Integrals

2.1 Applications and methods of integration

This chapter is concerned with extending the concept of integration to vector quantities and to three dimensions. Before embarking on these more complicated types of integration, however, it is useful to review the concept of integration and some standard techniques for evaluating integrals. It is important that the reader is familiar with these methods, since this will be assumed in the following sections.

2.1.1 Examples of the use of integration

Example 2.1

A rod of length a has a mass per unit length $\rho(x)$ that varies along the length of the rod according to the formula $\rho(x) = 1 + x$. What is the total mass of the rod?

Consider dividing the rod into N small sections, each of length dx_i. The mass of each section is $\rho(x_i)dx_i$. The total mass M of the rod is the sum of the masses of all these sections,

$$M = \sum_{i=1}^{N} \rho(x_i)\, dx_i.$$

The integral of $\rho(x)$ is defined to be the limit of this sum as $N \to \infty$:

$$\int_0^a \rho(x)\, dx = \lim_{N\to\infty} \sum_{i=1}^{N} \rho(x_i) dx_i.$$

The total mass M is therefore

$$M = \int_0^a 1 + x\, dx = \left[x + x^2/2\right]_0^a = a + a^2/2.$$

Example 2.2

A vehicle starts from rest and accelerates uniformly up to a speed of 10 m/s over a time of 20 s. What is the total distance travelled during this time?

The vehicle starts from rest and reaches a speed of 10 m/s after 20 s, so its speed at a time t is $v(t) = t/2$ m/s. In a small time interval dt the distance travelled is $v(t)\, dt = t/2\, dt$. The total distance S travelled in the total time of 20 s is therefore

$$S = \int_0^{20} t/2\, dt = \left[t^2/4\right]_0^{20} = 100\,\mathrm{m}.$$

2.1.2 Integration by substitution

In this method for the evaluation of integrals, a complicated integral is transformed to a simpler one by a substitution or change of variable. In some cases the choice of the change of variable is easy to find, but in others it can be difficult to spot the most sensible substitution. Often, there is more than one possible substitution. Three examples of the application of this method are given below.

Example 2.3

Evaluate $\int x/\sqrt{1-x}\; dx$.

Here, the difficulty is caused by the $\sqrt{1-x}$ in the denominator. This suggests that the appropriate substitution is $u = 1 - x$, so $x = 1 - u$ and $dx = -du$. The integral becomes $\int -(1-u)/\sqrt{u}\, du = \int -1/\sqrt{u} + \sqrt{u}\, du = -2u^{1/2} + 2u^{3/2}/3 + c$, where c is an arbitrary constant of integration. The result can be expressed in terms of the original variable x as $-2\sqrt{1-x}\,(2+x)/3 + c$.

Example 2.4

Evaluate $\int \sqrt{1-x^2}\, dx$.

For integrals involving the quantity $\sqrt{1-x^2}$, the appropriate substitution is $x = \sin\theta$ (or $x = \cos\theta$, which would do equally well). With this choice, $\sqrt{1-x^2}$ becomes $\cos\theta$ and $dx = \cos\theta\, d\theta$. The integral then simplifies to $\int \cos^2\theta\, d\theta$. Integrals of this type, which occur very frequently, are evaluated using the trigonometric formula $\cos^2\theta = (1+\cos 2\theta)/2$, so the value of the integral is $(2\theta + \sin 2\theta)/4 + c$. In terms of x this result can be written

$$\int \sqrt{1-x^2}\, dx = \left(\sin^{-1} x + x\sqrt{1-x^2}\right)/2 + c.$$

Example 2.5

Evaluate $\int_0^1 x^2\sqrt{1-x^2}\, dx$.

Again the substitution $x = \sin\theta$ is used. Since the limits of integration are given, these can also be expressed in terms of the new variable θ. When $x = 0$, $\theta = 0$ and when $x = 1$, $\theta = \pi/2$, so the integral becomes $\int_0^{\pi/2} \sin^2\theta\cos^2\theta\, d\theta$. This can be simplified to $\int_0^{\pi/2} 1/4\,\sin^2 2\theta\, d\theta = \int_0^{\pi/2} 1/8\,(1-\cos 4\theta)\, d\theta = \pi/16$.

2.1.3 Integration by parts

Integration by parts is an important and useful technique, used when an integral involves a product of two terms. The integration by parts formula is derived from the product rule for differentiation. Given two functions of x, $u(x)$ and $v(x)$, the rule for the derivative of their product is

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}. \tag{2.1}$$

Integrating this expression and rearranging the terms gives the integration by parts formula:

$$\int u\,\frac{dv}{dx}\,dx = u\,v - \int v\,\frac{du}{dx}\,dx. \tag{2.2}$$

As with the case of integration by substitution, some experience is helpful in determining whether this formula will be useful in evaluating an integral, and exactly how to split the integral into the two parts. In general, it is best to choose u to be a function which becomes simpler when differentiated. The following two examples illustrate the use of the method of integration by parts.

Example 2.6

Evaluate $\int x \sin x \, dx$.

In this example we choose $u = x$, $dv/dx = \sin x$, so $v = -\cos x$. Applying the formula (2.2) gives

$$\int x \sin x \, dx = -x \cos x + \int \cos x \, dx = -x \cos x + \sin x + c.$$

Note that it is essential to make the right choice for u and v. If we had chosen $u = \sin x$, $dv/dx = x$ then the resulting integral would have involved $x^2 \cos x$ which is more complicated than the integral we started with.

Example 2.7

Evaluate $\int \exp ax \cos x \, dx$.

For this case two applications of (2.2) are necessary. Choosing $u = \exp ax$, $dv/dx = \cos x$,

$$\int \exp ax \cos x \, dx = \exp ax \sin x - \int a \exp ax \sin x \, dx + c.$$

Now apply (2.2) to the integral on the r.h.s. with $u = a \exp ax$, $dv/dx = \sin x$:

$$\int \exp ax \cos x \, dx = \exp ax \sin x + a \exp ax \cos x - \int a^2 \exp ax \cos x \, dx + c.$$

At this stage it may appear that no progress has been made, since the original integral has reappeared on the right-hand side. However, by rearranging the terms,

$$(1 + a^2) \int \exp ax \cos x \, dx = \exp ax \sin x + a \exp ax \cos x + c$$

and so the value of the original integral is

$$\int \exp ax \cos x \, dx = (\exp ax \sin x + a \exp ax \cos x + c)/(1 + a^2).$$

In this case the choice of u and v does not matter: the result can also be obtained by choosing $u = \cos x$, $dv/dx = \exp ax$, provided that a similar choice, with u chosen to be the trigonometric term and dv/dx chosen to be the exponential term, is made for the second application of (2.2).

2.2 Line integrals

2.2.1 Introductory example: work done against a force

As an introductory example of a line integral, consider a particle moving along a curved path C through space. The particle is acted on by a force $\boldsymbol{F}(\boldsymbol{r})$, which is a vector field. What is the total amount of work done as the particle moves along the curve C?

To answer this question, first divide the curve C into a large number of small pieces. Consider the work done when the particle moves from the position $\boldsymbol{r}$ to $\boldsymbol{r} + \boldsymbol{dr}$ (Figure 2.1). On this small section of the curve C, the work done is $-\boldsymbol{F} \cdot \boldsymbol{dr}$. The total amount of work done W as the particle moves along C is therefore the sum of the contributions from all the small segments of the curve,

$$W = \sum_{i=1}^{N} -\boldsymbol{F}_i \cdot \boldsymbol{dr}_i. \tag{2.3}$$

The *line integral* of $\boldsymbol{F}$ along the curve C is defined by

$$\int_C \boldsymbol{F} \cdot \boldsymbol{dr} = \lim_{N\to\infty} \sum_{i=1}^{N} \boldsymbol{F}_i \cdot \boldsymbol{dr}_i. \tag{2.4}$$

The vector $\boldsymbol{dr}$ is often referred to as a *line element*.

Note that the direction of the integral along the curve C must be specified. If the direction of the curve is reversed, all the line elements $\boldsymbol{dr}$ are reversed and so the value of the integral is multiplied by -1.

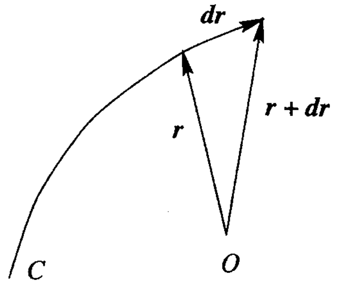


Fig. 2.1. A small section of the curve C is represented by the line element $\boldsymbol{dr}$.

2.2.2 Evaluation of line integrals

Line integrals are evaluated by using a parameter, t for example, together with a formula giving the value of the position vector $\boldsymbol{r}$ in terms of t. This can be regarded as an example of integration by substitution, since

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int \boldsymbol{F} \cdot \frac{d\boldsymbol{r}}{dt}\, dt. \tag{2.5}$$

For example, suppose that the curve C is given in terms of t by

$$x = t, \qquad y = t, \qquad z = 2t^2, \tag{2.6}$$

and t lies in the range $0 \leq t \leq 1$. Then as t varies between 0 and 1, the position vector $\boldsymbol{r} = (x, y, z)$ moves along a curve C in space connecting the points $(0,0,0)$ and $(1,1,2)$. Suppose that the vector field $\boldsymbol{F}$ is given by $\boldsymbol{F} = (y, x, z)$. To evaluate the line integral, both $\boldsymbol{F}$ and $d\boldsymbol{r}/dt$ must be written in terms of t. Substituting (2.6) into the expression for $\boldsymbol{F}$ gives $\boldsymbol{F} = (t, t, 2t^2)$, and

$$\frac{d\boldsymbol{r}}{dt} = \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) = (1, 1, 4t).$$

The line integral can now be evaluated:

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_0^1 (t, t, 2t^2) \cdot (1, 1, 4t)\, dt = \int_0^1 2t + 8t^3\, dt = 3.$$

In this first example, the parametric form of the curve C was given. If the curve C is given in a different form, then a parametric form must be constructed so that the line integral can be evaluated. For example, suppose now that $\boldsymbol{F} = (y, x, z)$ as before, but C is the straight line connecting the origin to the point $(1,2,3)$. The way in which the curve C is parametrised is not unique, so we can make the arbitrary choice $x = t$. Since x varies between 0 and 1 along the line, this is also the range for t. The end point of C is $(1,2,3)$, so y and z must be given by $y = 2t$, $z = 3t$, and so $d\boldsymbol{r} = (1,2,3)\, dt$. The value of the integral is therefore

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_0^1 (2t, t, 3t) \cdot (1, 2, 3)\, dt = \int_0^1 13t\, dt = 6.5.$$

Line integrals sometimes occur over curves that are closed, i.e. when the starting point and end point of the curve are equal. In this case the integral is written using the symbol $\oint$, which indicates that the integral is along a closed curve. For example, consider the integral of $\boldsymbol{F} = (y, x, z)$ around the closed curve given by $x = \cos\theta$, $y = \sin\theta$, $z = 0$, where $0 \leq \theta \leq 2\pi$. Here, as θ varies, the curve C describes a circle in the x, y plane. The line element $d\boldsymbol{r}$ is expressed in terms of the parameter θ as $\boldsymbol{dr} = (dx, dy, dz) = (-\sin\theta, \cos\theta, 0)\, d\theta$, so the value of the line integral is

$$\begin{aligned}\oint_C \boldsymbol{F}\cdot\boldsymbol{dr} &= \int_0^{2\pi} (\sin\theta, \cos\theta, 0)\cdot(-\sin\theta, \cos\theta, 0)\, d\theta \\ &= \int_0^{2\pi} -\sin^2\theta + \cos^2\theta\, d\theta \\ &= \int_0^{2\pi} \cos 2\theta\, d\theta = \left[1/2\,\sin 2\theta\right]_0^{2\pi} = 0.\end{aligned}$$

The line integral of a vector field $\boldsymbol{F}$ around a closed curve C is often called the *circulation* of $\boldsymbol{F}$ around C.

Example 2.8

Evaluate the line integral of the vector field $\boldsymbol{u} = (xy, z^2, x)$ along the curve given by $x = 1+t$, $y = 0$, $z = t^2$, $0 \le t \le 3$.

First write $\boldsymbol{u}$ and $\boldsymbol{dr}$ in terms of t: $\boldsymbol{u} = (0, t^4, 1+t)$, $\boldsymbol{dr} = (1, 0, 2t)\, dt$. The value of the integral is therefore

$$\int_C \boldsymbol{F}\cdot\boldsymbol{dr} = \int_0^3 (0, t^4, 1+t)\cdot(1, 0, 2t)\, dt = \int_0^3 2t + 2t^2\, dt = \left[t^2 + 2t^3/3\right]_0^3 = 27.$$

Example 2.9

Find the line integral of $\boldsymbol{F} = (y, -x, 0)$ along the curve consisting of the two straight line segments (a) $y = 1$, $0 \le x \le 1$, (b) $x = 1$, $1 \le y \le 2$.

Here, the contributions from the two line segments must be taken separately. On section (a), using x as the parameter, we have

$$\int_0^1 (1, -x, 0)\cdot(dx, 0, 0) = 1.$$

Similarly on section (b) we have

$$\int_1^2 (y, -1, 0)\cdot(0, dy, 0) = -1.$$

Therefore the total value of the integral is 0.

Example 2.10

Find the circulation of the vector $\boldsymbol{F} = (y, -x, 0)$ around the unit circle, $x^2 + y^2 = 1$, $z = 0$, taken in an anticlockwise direction.

The circle is written in terms of a parameter θ as $x = \cos\theta$, $y = \sin\theta$, $z = 0$, $0 \le \theta \le 2\pi$. The value of the integral is

$$\oint_C \boldsymbol{F}\cdot\boldsymbol{dr} = \int_0^{2\pi} (\sin\theta, -\cos\theta, 0)\cdot(-\sin\theta, \cos\theta, 0)\, d\theta = \int_0^{2\pi} -1\, d\theta = -2\pi.$$

2.2.3 Conservative vector fields

A vector field $\boldsymbol{F}$ is said to be *conservative* if it has the property that the line integral of $\boldsymbol{F}$ around any closed curve C is zero:

$$\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = 0. \tag{2.7}$$

An equivalent definition is that $\boldsymbol{F}$ is conservative if the line integral of $\boldsymbol{F}$ along a curve only depends on the endpoints of the curve, not on the path taken by the curve,

$$\int_{C_1} \boldsymbol{F} \cdot \boldsymbol{dr} = \int_{C_2} \boldsymbol{F} \cdot \boldsymbol{dr} \tag{2.8}$$

where C_1 and C_2 are any two curves that have the same endpoints but different paths (Figure 2.2).

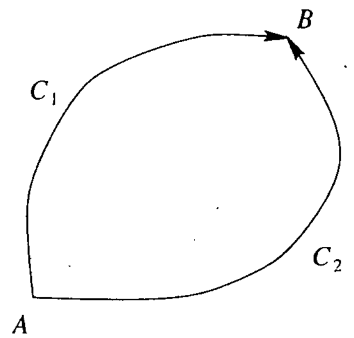


Fig. 2.2. A vector field is conservative if the line integrals along two different curves C_1 and C_2 from A to B are equal.

The equivalence of these two definitions can be demonstrated as follows. Consider two curves C_1 and C_2 that start from the point A and end at the point B (Figure 2.2). Let C be the closed curve that starts from the point A, follows the curve C_1 to the point B and then follows the curve C_2 in the reverse direction to return to A. Then

$$\oint_C \boldsymbol{F} \cdot \boldsymbol{dr} = \int_{C_1} \boldsymbol{F} \cdot \boldsymbol{dr} - \int_{C_2} \boldsymbol{F} \cdot \boldsymbol{dr} \tag{2.9}$$

since the effect of the reversed direction of the integral along C_2 is to change the sign of the integral. From this equation it follows that if the integral around C is zero, then the integrals along C_1 and C_2 are equal, and similarly, if the integrals along C_1 and C_2 are equal, then the integral around C is zero.

Conservative vector fields are of great importance, since many physical examples of vector fields are conservative. Consider for example the Earth's gravitational field, $\boldsymbol{g}$. A particle of mass m experiences a force $m\boldsymbol{g}$, so the work done in moving the particle along a path C from point A to point B is just minus the line integral of $m\boldsymbol{g}$ along C. However, we know physically that the work done only depends on the position of the points A and B – in fact the work done is simply the difference in the potential energy of the particle at A and B. Equivalently, if the particle is moved around but returns to its starting point, the total work done is zero. Therefore, the Earth's gravitational field is an example of a conservative vector field.

Example 2.11

By considering the line integral of $\boldsymbol{F} = (y, x^2 - x, 0)$ around the square in the x, y plane connecting the four points $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$, show that $\boldsymbol{F}$ cannot be a conservative vector field.

This line integral consists of four parts (Figure 2.3). On the first section

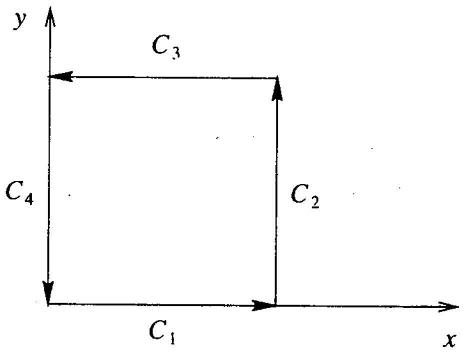


Fig. 2.3. The line integral around the square is split into four straight sections.

C_1, from $(0,0)$ to $(1,0)$, $\boldsymbol{dr} = (dx, 0, 0)$ and $y = 0$ so $\boldsymbol{F} = (0, x^2 - x, 0)$ and $\boldsymbol{F} \cdot \boldsymbol{dr} = 0$. On the second section C_2, $\boldsymbol{dr} = (0, dy, 0)$ and $x = 1$ so $\boldsymbol{F} = (y, 0, 0)$ and again $\boldsymbol{F} \cdot \boldsymbol{dr} = 0$. On the third section C_3, $\boldsymbol{dr} = (dx, 0, 0)$ and $y = 1$ so $\boldsymbol{F} = (1, x^2 - x, 0)$ and the contribution to the line integral is

$$\int_1^0 1\, dx = -1.$$

Finally, on the fourth section C_4 of the square $\boldsymbol{F} \cdot \boldsymbol{dr} = 0$ again, so the total value of the integral is -1. Since the integral around the closed circuit is non-zero, $\boldsymbol{F}$ cannot be conservative.

2.2.4 Other forms of line integrals

The line integral $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$ is the most important type of line integral, but there are two other forms of the line integral which can occur. These are

$$\int_C \phi \, d\boldsymbol{r} \qquad \text{and} \qquad \int_C \boldsymbol{F} \times d\boldsymbol{r}$$

where ϕ is a scalar field and $\boldsymbol{F}$ is a vector field. Note that in each of these cases the result of the integral is a vector quantity. These integrals can be evaluated using a parameter, as in the following examples.

Example 2.12

Evaluate the line integral

$$\int_C x + y^2 \, d\boldsymbol{r},$$

where C is the parabola $y = x^2$ in the plane $z = 0$ connecting the points $(0,0,0)$ and $(1,1,0)$.

The curve can be written in terms of a parameter t as $x = t$, $y = t^2$, $z = 0$, $0 \le t \le 1$, so $d\boldsymbol{r} = (1, 2t, 0)\, dt$. The value of the integral is therefore

$$\begin{aligned}
\int_C x + y^2 \, d\boldsymbol{r} &= \int_0^1 (t + t^4)(1, 2t, 0)\, dt \\
&= \boldsymbol{e}_1 \left(\int_0^1 t + t^4 \, dt \right) + \boldsymbol{e}_2 \left(\int_0^1 2t^2 + 2t^5 \, dt \right) = 0.7\boldsymbol{e}_1 + \boldsymbol{e}_2.
\end{aligned}$$

Example 2.13

Evaluate the line integral

$$\int_C \boldsymbol{F} \times d\boldsymbol{r},$$

where $\boldsymbol{F}$ is the vector field $(y, x, 0)$ and C is the curve $y = \sin x$, $z = 0$, between $x = 0$ and $x = \pi$.

The curve can be written as $x = t$, $y = \sin t$, $z = 0$, $0 \le t \le \pi$. Then $\boldsymbol{F} = (\sin t, t, 0)$ and $d\boldsymbol{r} = (1, \cos t, 0)\, dt$, so $\boldsymbol{F} \times d\boldsymbol{r} = (0, 0, \sin t \cos t - t)\, dt$ and the integral is

$$\int_C \boldsymbol{F} \times d\boldsymbol{r} = \boldsymbol{e}_3 \int_0^\pi \sin t \cos t - t \, dt = 1/2 \left[\sin^2 t - t^2 \right]_0^\pi \boldsymbol{e}_3 = -\pi^2/2\, \boldsymbol{e}_3.$$

EXERCISES

2.1 Evaluate the line integral

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} \quad \text{,where} \quad \boldsymbol{F} = (5z^2, 2x, x + 2y) \tag{2.10}$$

and the curve C is given by $x = t$, $y = t^2$, $z = t^2$, $0 \leq t \leq 1$.

2.2 Evaluate the line integral of the same vector field $\boldsymbol{F}$ given in (2.10) along the straight line joining the points $(0, 0, 0)$ and $(1, 1, 1)$. Is $\boldsymbol{F}$ a conservative vector field?

2.3 Find the line integral of the vector field $\boldsymbol{u} = (y^2, x, z)$ along the curve given by $z = y = e^x$ from $x = 0$ to $x = 1$.

2.4 Find the line integral $\oint_C \boldsymbol{r} \times d\boldsymbol{r}$ where the curve C is the ellipse $x^2/a^2 + y^2/b^2 = 1$ taken in an anticlockwise direction. What do you notice about the magnitude of the answer?

2.3 Surface integrals

2.3.1 Introductory example: flow through a pipe

Suppose that fluid flows with velocity $\boldsymbol{u}$ through a pipe. What is the total volume of fluid passing through the pipe per unit time (Figure 2.4)? This

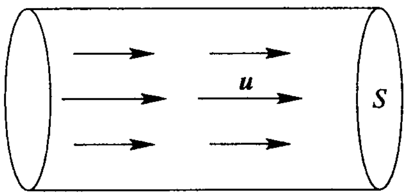


Fig. 2.4. Fluid flows with velocity $\boldsymbol{u}$ along a pipe. The rate at which it crosses the surface S at the end of the pipe is an example of a surface integral.

volume flow rate is often called the *flux* of fluid through the pipe, or the flux of fluid across the surface S that forms the end of the pipe. We will consider this question in several different cases, beginning with the simplest case and progressing to more complicated examples.

Suppose first that the pipe has cross-sectional area A and that the velocity $\boldsymbol{u}$, which in general can be a function of space and time, is a constant and is directed parallel to the walls of the pipe, with speed $|\boldsymbol{u}| = U_0$. In this case, the fluid moves along the pipe as if it were a solid block. In a time t, the fluid moves a distance $U_0 t$, so a 'block' of fluid of volume $U_0 t A$ emerges from the end of the pipe. The flow rate Q, or flux, of fluid through the pipe is therefore this volume divided by the time t, giving $Q = U_0 A$.

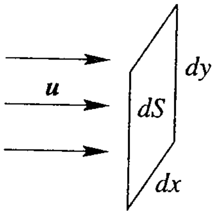


Fig. 2.5. Enlargement of a small surface element dS forming part of the surface S.

Now suppose that the flow is again directed parallel to the walls of the pipe but that the speed of the flow depends on the position within the pipe, so $|\boldsymbol{u}| = U_0(x, y)$, and that the pipe has a square cross-section with the walls at $x = 0, 1$ and $y = 0, 1$. Now consider a small *surface element* with area dS on the surface S, which is a small rectangle with sides of length dx and dy located at the point (x, y) on the surface S, so that $dS = dx\,dy$ (Figure 2.5). Following the argument of the previous paragraph, the flux dQ of fluid across this surface element dS is $dQ = U_0(x, y)\,dS = U_0(x, y)\,dx\,dy$. To calculate the total flux Q across the surface S, we need to add up the contributions from all the small surface elements dS. This sum of contributions becomes an integral, but since the surface is two-dimensional the resulting integral is a surface integral or double integral, representing integration in both the x and y directions:

$$Q = \iint_S U_0(x, y)\,dS = \iint_S U_0(x, y)\,dx\,dy. \qquad (2.11)$$

In the above example, the fluid flow direction is perpendicular to the surface S. Consider now the case where the vector field $\boldsymbol{u}$ and the surface S are both arbitrary. In general, S may represent a curved surface. Again we consider a small surface element dS (Figure 2.6) and compute the flux of $\boldsymbol{u}$ across dS. Now if $\boldsymbol{u}$ is perpendicular to dS, this flux is just $|\boldsymbol{u}|dS$, but if $\boldsymbol{u}$ is not perpendicular

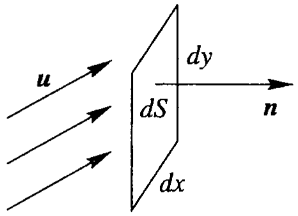


Fig. 2.6. The unit normal vector $\boldsymbol{n}$ is perpendicular to the surface element dS.

to dS, only the component of $\boldsymbol{u}$ perpendicular to dS contributes to the flux across dS. To extract this component it is necessary to introduce a *normal vector* $\boldsymbol{n}$ to the surface dS, with the properties that $\boldsymbol{n}$ is perpendicular to dS and that $\boldsymbol{n}$ is a unit vector, $|\boldsymbol{n}| = 1$. The component of $\boldsymbol{u}$ perpendicular to dS is then the component of $\boldsymbol{u}$ in the direction of $\boldsymbol{n}$, which is just $\boldsymbol{u} \cdot \boldsymbol{n}$ (see Section 1.2), so the flux across the surface element dS is $\boldsymbol{u} \cdot \boldsymbol{n}\, dS$. The total flux across the surface S is given by the surface integral

$$Q = \iint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dS = \iint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dx\, dy.$$

Note that the direction of the normal vector $\boldsymbol{n}$ has not been uniquely specified: the vector $\boldsymbol{n}$ could have been chosen to point in the opposite direction in Figure 2.6, and this would change the sign of the answer. Therefore the direction of $\boldsymbol{n}$ must be specified when a surface integral is written down.

Surface integrals often occur over surfaces which are closed. In this case, the normal to the surface which points outward is used (Figure 2.7). To indicate that the surface is closed, a circle is sometimes drawn through the integral, in a similar manner to the notation used for line integrals around closed curves:

$$\oiint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dS.$$

2.3.2 Evaluation of surface integrals

Surface integrals can be evaluated by carrying out two successive integrations. Consider first the evaluation of (2.11) where S is the square surface given by $0 \le x \le 1$, $0 \le y \le 1$. The surface integral is then

$$Q = \iint_S U_0(x, y)\, dS = \int_0^1 \int_0^1 U_0(x, y)\, dx\, dy.$$

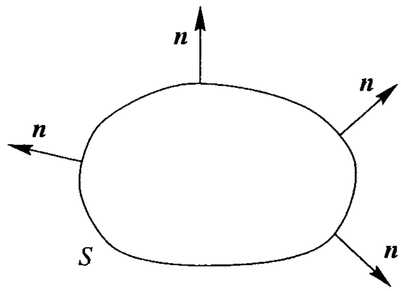


Fig. 2.7. For a closed surface, the convention is that the normal points outward.

There is some potential ambiguity here, since it is not immediately clear which of the integral signs refers to the x integration and which to the y integration. The convention adopted is that the integrals are 'nested', so that the first integral sign represents the y integral and the second one represents the x integral. The double integral is then interpreted as

$$Q = \int_0^1 \left(\int_0^1 U_0(x,y)\,dx \right) dy.$$

It is also assumed that the inner integral, the x integral in the above equation, is to be evaluated first. For example, suppose that $U_0(x,y) = (x - x^2)(y - y^2)$. The inner, x, integral is evaluated first, and within this inner integral y is regarded as a constant. Carrying out the first integral gives

$$\begin{aligned} Q &= \int_0^1 \int_0^1 (x - x^2)(y - y^2)\,dx\,dy \\ &= \int_0^1 \left[x^2/2 - x^3/3\right]_0^1 (y - y^2)\,dy \\ &= \int_0^1 1/6\,(y - y^2)\,dy. \end{aligned}$$

The double integral has now been reduced to a single integral which can be evaluated in the usual way:

$$Q = \int_0^1 1/6\,(y - y^2)\,dy = 1/6\,\left[y^2/2 - y^3/3\right]_0^1 = 1/36.$$

Consider now the case where the surface is circular. In the original example of pipe flow, this corresponds to a cylindrical pipe. If the surface is circular, it is best to use polar coordinates to evaluate the integral. Polar coordinates (r, θ) are related to Cartesian coordinates (x, y) by $x = r\cos\theta$, $y = r\sin\theta$. The angle

θ is measured in radians and ranges from 0 to 2π. The area element dS can be computed by considering a small angle $d\theta$ and a small change in radius dr as shown in Figure 2.8. If dr and $d\theta$ are both small then the corresponding area

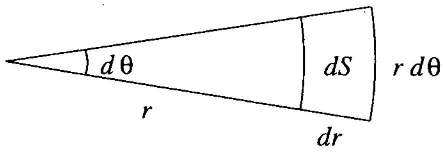


Fig. 2.8. The area element dS in polar coordinates.

element dS is almost rectangular, with length dr in the r direction and $r\,d\theta$ in the θ direction, so $dS = r\,d\theta\,dr$. Suppose now that the radius of the surface S is 1 and that $U_0 = 1 - r^2$. The value of the surface integral is then

$$Q = \iint_S U_0\,dS = \int_0^1 \int_0^{2\pi} (1 - r^2) r\,d\theta\,dr.$$

The inner, θ, integral is carried out first, with r temporarily regarded as a constant. Since there is no dependence on θ in the integral, this inner integral just gives a factor of 2π, so

$$Q = \int_0^1 2\pi(1 - r^2) r\,dr = 2\pi \left[r^2/2 - r^4/4\right]_0^1 = \pi/2.$$

Finally, consider the case where the surface S is curved. The surface can be written in terms of two parameters, v and w, so that a position vector $\boldsymbol{r}$ lying in the surface is written $\boldsymbol{r} = \boldsymbol{r}(v, w)$. Now consider a small change in the value of v, to $v + dv$. The vector $\boldsymbol{r}(v + dv, w)$ also lies on the surface, so the difference between these two vectors, $\boldsymbol{r}(v + dv, w) - \boldsymbol{r}(v, w) = (\partial \boldsymbol{r}/\partial v)\,dv$ must be a vector lying in, or tangent to, the surface, and similarly for the vector $(\partial \boldsymbol{r}/\partial w)\,dw$. To evaluate the surface integral we need an expressions for $\boldsymbol{n}\,dS$, but we know from Section 1.3 that this is simply the cross product of the two vectors, since the cross product of two vectors gives a vector perpendicular to both and with a magnitude equal to the area of the parallelogram created by the two vectors. Therefore the surface integral can be written as

$$\iint_S \boldsymbol{u} \cdot \boldsymbol{n}\,dS = \iint_S \boldsymbol{u} \cdot \frac{\partial \boldsymbol{r}}{\partial v} \times \frac{\partial \boldsymbol{r}}{\partial w}\,dv\,dw.$$

For example, consider the integral of the vector field $\boldsymbol{u} = (x, z, -y)$ over the curved surface of the cylinder $x^2 + y^2 = 1$ lying between $z = 0$ and $z = 1$. The

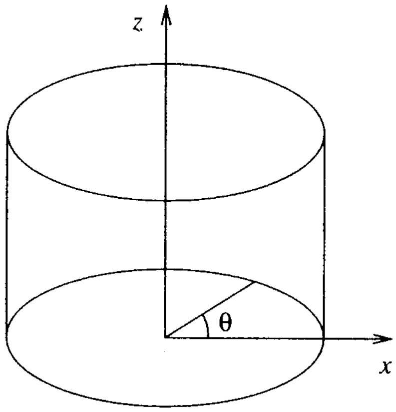


Fig. 2.9. A point on the cylindrical surface $x^2 + y^2 = 1$ can be denoted by the parameters z and θ.

two parameters describing the surface are the height z and the angle θ around the cylinder (Figure 2.9). In terms of these parameters the position vector is $\boldsymbol{r} = (x, y, z) = (\cos\theta, \sin\theta, z)$, so

$$\frac{\partial \boldsymbol{r}}{\partial \theta} = (-\sin\theta, \cos\theta, 0), \qquad \frac{\partial \boldsymbol{r}}{\partial z} = (0, 0, 1)$$

$$\text{and} \quad \frac{\partial \boldsymbol{r}}{\partial \theta} \times \frac{\partial \boldsymbol{r}}{\partial z} = (\cos\theta, \sin\theta, 0).$$

The value of the integral is therefore

$$\begin{aligned} \iint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dS &= \int_0^1 \int_0^{2\pi} (\cos\theta, z, -\sin\theta) \cdot (\cos\theta, \sin\theta, 0)\, d\theta\, dz \\ &= \int_0^1 \int_0^{2\pi} \cos^2\theta + z \sin\theta\, d\theta\, dz \\ &= \int_0^1 \pi\, dz = \pi. \end{aligned}$$

Example 2.14

Evaluate the surface integral of $\boldsymbol{u} = (y, x^2, z^2)$, over the surface S, where S is the triangular surface on $x = 0$ with $y \geq 0$, $z \geq 0$, $y + z \leq 1$, with the normal $\boldsymbol{n}$ directed in the positive x direction.

In this example $\boldsymbol{n} = (1, 0, 0)$ and so $\boldsymbol{u} \cdot \boldsymbol{n} = y$. When the surface is not rectangular, care must be taken when setting the limits of the integration. If we choose to do the z integral first, then for any given value of y, the range

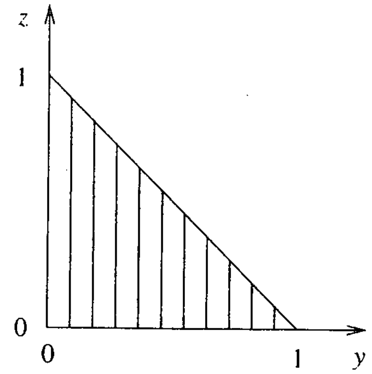


Fig. 2.10. For each value of y, z runs from the line $z = 0$ up to the line $y + z = 1$.

of values for z is $0 \leq z \leq 1 - y$ (Figure 2.10). The outer y integral then has limits of 0 and 1. This corresponds to covering the triangular area with vertical strips. The value of the integral is

$$\int_0^1 \int_0^{1-y} y \, dz \, dy = \int_0^1 [yz]_0^{1-y} \, dy = \int_0^1 y - y^2 \, dy = 1/6.$$

The integral could also have been evaluated by doing the y integral first, in which case the limits for y are $0 \leq y \leq 1 - z$ and the limits for z are $0 \leq z \leq 1$. This ordering corresponds to covering the region of integration in Figure 2.10 with horizontal strips. The value of the integral is the same:

$$\int_0^1 \int_0^{1-z} y \, dy \, dz = \int_0^1 [y^2/2]_0^{1-z} \, dz = \int_0^1 (1-z)^2/2 \, dz = 1/6.$$

Example 2.15

Find the surface integral of $\boldsymbol{u} = \boldsymbol{r}$ over the part of the paraboloid $z = 1 - x^2 - y^2$ with $z > 0$, with the normal pointing upwards.

Since the surface is curved, a description of the surface in terms of two parameters is needed. Using simply x and y, a point on the surface is $(x, y, 1 - x^2 - y^2)$ and the two tangent vectors in the surface, obtained by differentiating with respect to x and y, are $(1, 0, -2x)$ and $(0, 1, -2y)$. Taking the cross product of these two vectors, the quantity $\boldsymbol{n} \, dS$ is $(2x, 2y, 1) \, dx \, dy$. Note that this has a positive z component, so is directed upwards as required. Taking the dot product with $\boldsymbol{u}$ gives $\boldsymbol{u} \cdot \boldsymbol{n} \, dS = 2x^2 + 2y^2 + z \, dx \, dy = 1 + x^2 + y^2 \, dx \, dy$. The limits on the integral are determined in a similar manner to Example 2.14. The edge of the surface is given as $z = 0$, which is the circle $x^2 + y^2 = 1$. Choosing

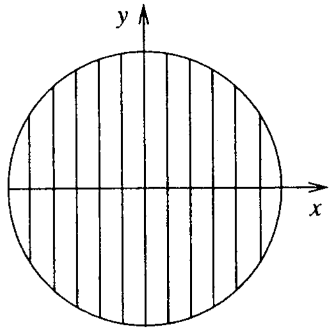


Fig. 2.11. For any value of x, y ranges from the lower half of the circle $x^2 + y^2 = 1$ to the upper half of the circle.

to do the y integral first, the y integration is carried out with x fixed, so the range of values for y is $-\sqrt{1-x^2} < y < \sqrt{1-x^2}$ (Figure 2.11). The range of values for x is $-1 < x < 1$. The value of the integral is

$$\begin{aligned}
\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 + x^2 + y^2 \, dy \, dx &= \int_{-1}^{1} \left[y + x^2 y + y^3/3\right]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dx \\
&= \int_{-1}^{1} (8/3 + 4/3\,x^2)\sqrt{1-x^2}\, dx \\
&= 4\pi/3 + \pi/6 = 3\pi/2,
\end{aligned}$$

where the final integral has been evaluated using the results of Section 2.1.2.

2.3.3 Other forms of surface integrals

The surface integral of $\boldsymbol{u} \cdot \boldsymbol{n}$ is the most important type of surface integral. However, as in the case of line integrals, other types of surface integrals can be defined, for example

$$\iint_S f \, dS, \qquad \iint_S f \, \boldsymbol{n} \, dS \qquad \text{and} \qquad \iint_S \boldsymbol{v} \times \boldsymbol{n} \, dS,$$

where f is a scalar field and $\boldsymbol{v}$ is a vector field. These integrals are evaluated using the methods of the previous section.

Example 2.16

If S is the entire x, y plane, evaluate the integral

$$I = \iint_S e^{-x^2-y^2}\, dS,$$

by transforming the integral into polar coordinates.

In polar coordinates (r, θ), $x^2 + y^2 = r^2$ and $dS = r\, d\theta\, dr$. The ranges of the variables to cover the whole plane are $0 \leq r < \infty$ and $0 \leq \theta < 2\pi$, so

$$I = \int_0^\infty \int_0^{2\pi} e^{-r^2} r\, d\theta\, dr = \int_0^\infty 2\pi e^{-r^2} r\, dr = \pi \left[-e^{-r^2}\right]_0^\infty = \pi.$$

This answer can be used to show that $\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}$, a result which cannot be obtained by standard methods of integration.

2.4 Volume integrals

2.4.1 Introductory example: mass of an object with variable density

Suppose that an object of volume V has a density ρ. If ρ is a constant, the mass M of the object is simply $M = \rho V$. Now suppose that the object has a density which is a function of position, $\rho = \rho(\boldsymbol{r})$. What is the total mass of the object?

The argument proceeds in a similar manner to the examples of line and surface integrals. The volume V is divided into N small pieces with volumes δV_i, $i = 1, \ldots, N$, which are called *volume elements*. Within each of the volume elements the density is approximately constant (assuming that ρ is a continuous function of position) and so the mass M_i of the volume element at position $\boldsymbol{r}_i$ is $M_i = \rho(\boldsymbol{r}_i)\, \delta V_i$. The total mass of the object is the sum of the masses of all the volume elements,

$$M = \sum_{i=1}^{N} \rho(\boldsymbol{r}_i)\, \delta V_i.$$

The *volume integral* of ρ over the volume V is defined to be the limit of this sum as $N \to \infty$:

$$\iiint_V \rho\, dV = \lim_{N\to\infty} \sum_{i=1}^{N} \rho(\boldsymbol{r}_i)\, \delta V_i. \tag{2.12}$$

Volume integrals can also be used to compute the volumes of objects, in which case $\rho = 1$ in the above example. Note that volume integrals usually occur as integrals of scalar quantities. However, the volume integral of a vector field $\boldsymbol{u}$,

$$\iiint_V \boldsymbol{u}\, dV$$

can be defined in a similar way.

2.4.2 Evaluation of volume integrals

Volume integrals are evaluated by carrying out three successive integrals. The same rule for the evaluation of the triple integral applies as for double integrals: the inner integral is evaluated first. The main difficulty in this process is in determining the correct limits for the integrals when the shape of the object is complicated. It is often helpful to sketch the region of integration in order to find the limits on the integrals. Also useful is the rule that in general, the limits on an integral can depend only on the variables of integrals that lie outside that integral. For example, if the integrals are evaluated in the order x, y, z then the limits on the y integral may depend on z but not on x.

Example 2.17

A cube $0 \le x, y, z, \le 1$ has a variable density given by $\rho = 1 + x + y + z$. What is the total mass of the cube?

The total mass is

$$\begin{aligned}
M &= \iiint_V \rho\, dV \\
&= \int_0^1 \int_0^1 \int_0^1 1 + x + y + z\, dx\, dy\, dz \\
&= \int_0^1 \int_0^1 \left[x + x^2/2 + xy + xz\right]_0^1 dy\, dz \\
&= \int_0^1 \int_0^1 (3/2 + y + z)\, dy\, dz \\
&= \int_0^1 \left[3y/2 + y^2/2 + yz\right]_0^1 dz \\
&= \int_0^1 (2 + z)\, dz \\
&= \left[2z + z^2/2\right]_0^1 = 5/2.
\end{aligned}$$

Note that in this example the integrals were carried out in the order x, y, z, but any other choice of ordering is equally valid.

Example 2.18

Find the volume of the tetrahedron with vertices at $(0,0,0)$, $(a,0,0)$, $(0,b,0)$ and $(0,0,c)$.

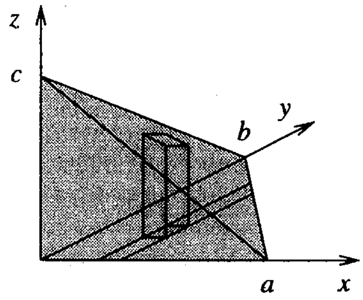


Fig. 2.12. For any given values of x and y, z ranges from the plane $z=0$ up to the plane $x/a+y/b+z/c=1$. This is indicated by the vertical column.

A sketch of the tetrahedron is shown in Figure 2.12. The faces of the tetrahedron are the planes $x=0$, $y=0$, $z=0$ and the plane which passes through the three points $(a,0,0)$, $(0,b,0)$ and $(0,0,c)$. The equation of this plane is $x/a+y/b+z/c=1$, which can be deduced from the general formula for the equation of a plane (1.3). Suppose that we choose to do the z integral first. This integral is carried out for fixed values of x and y, so the range of z is from the plane $z=0$ to the plane $z=c(1-x/a-y/b)$. Choosing to do the y integral next, y ranges from 0 to the line that passes through $(a,0,0)$ and $(0,b,0)$, which is $y=b(1-x/a)$. Finally the range of x is from 0 to a. The volume V is therefore

$$\begin{aligned}
V &= \int_0^a \int_0^{b(1-x/a)} \int_0^{c(1-x/a-y/b)} dz\,dy\,dx \\
&= \int_0^a \int_0^{b(1-x/a)} c(1-x/a-y/b)\,dy\,dx \\
&= c\int_0^a \left[y(1-x/a)-y^2/2b\right]_0^{b(1-x/a)} dx \\
&= \frac{cb}{2}\int_0^a (1-x/a)^2\,dx = \frac{abc}{6}.
\end{aligned}$$

Note that this result is consistent with the formula for the volume of a tetrahedron given in terms of the scalar triple product in Section 1.4.

Summary of Chapter 2

- There are several different types of integral, but each one should be interpreted as the limit of a sum.
- A *line integral*, written

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r},$$

represents the sum of the elements $\boldsymbol{F} \cdot d\boldsymbol{r}$ along the curve C. Applications include the total amount of work done when a particle moves in the presence of a force that is a function of position.
- Line integrals are evaluated by writing the vector $\boldsymbol{F}$ and the curve C in terms of a parameter, t.
- If the line integral of $\boldsymbol{F}$ around any closed curve is zero, $\boldsymbol{F}$ is said to be *conservative*.
- The *surface integral*,

$$\iint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dS,$$

represents the flux of $\boldsymbol{u}$ across the surface S; this can be thought of as the volume of fluid flowing with velocity $\boldsymbol{u}$ across the surface S per unit time. The *normal vector* $\boldsymbol{n}$ is a unit vector that is perpendicular to the surface S.
- Surface integrals are evaluated by carrying out two successive integrations.
- The *volume integral*,

$$\iiint_V \rho\, dV,$$

represents the sum of $\rho\, dV$ over all the volume elements dV contained within V. If ρ is the density, the volume integral gives the total mass of the object with volume V.
- Volume integrals are evaluated by carrying out three successive integrations. Care must be taken over setting the limits of the integrals and over the order in which the three integrals are evaluated.
- In both surface and volume integrals, the inner integral is evaluated first.
- There are other forms of line, surface and volume integrals, but the forms displayed above are the most commonly occurring.

EXERCISES

2.5 Evaluate the surface integral of $\boldsymbol{u} = (xy, x, x + y)$ over the surface S defined by $z = 0$ with $0 \le x \le 1$, $0 \le y \le 2$, with the normal $\boldsymbol{n}$ directed in the positive z direction.

2.6 Find the surface integral of $\boldsymbol{u} = \boldsymbol{r}$ over the surface of the unit cube $0 \le x, y, z \le 1$, with $\boldsymbol{n}$ pointing outward.

2.7 The surface S is defined to be that part of the plane $z = 0$ lying between the curves $y = x^2$ and $x = y^2$. Find the surface integral of $\boldsymbol{u} \cdot \boldsymbol{n}$ over S where $\boldsymbol{u} = (z, xy, x^2)$ and $\boldsymbol{n} = (0, 0, 1)$.

2.8 Find the surface integral of $\boldsymbol{u} \cdot \boldsymbol{n}$ over S where S is the part of the surface $z = x + y^2$ with $z < 0$ and $x > -1$, $\boldsymbol{u}$ is the vector field $\boldsymbol{u} = (2y + x, -1, 0)$ and $\boldsymbol{n}$ has a negative z component.

2.9 Find the volume integral of the scalar field $\phi = x^2 + y^2 + z^2$ over the region V specified by $0 \le x \le 1$, $1 \le y \le 2$, $0 \le z \le 3$.

2.10 Find the volume of the section of the cylinder $x^2 + y^2 = 1$ that lies between the planes $z = x + 1$ and $z = -x - 1$.

2.11 A circular pond with radius 1 m and a maximum depth of 1 m has the shape of a paraboloid, so that its depth z is $z = 1 - x^2 - y^2$. What is the total volume of the pond? How does this compare with the case where the pond has the same radius and depth but has the shape of a hemisphere?

3
Gradient, Divergence and Curl

3.1 Partial differentiation and Taylor series

This chapter introduces important concepts concerning the differentiation of scalar and vector quantities in three dimensions. These concepts form the core of the subject of vector calculus. In this preliminary section, the methods of partial differentiation and Taylor series are reviewed.

3.1.1 Partial differentiation

Consider a scalar quantity f which is a function of three variables, so $f = f(x, y, z)$. Then the ***partial derivative*** of f with respect to x is defined to be the derivative of f with respect to x, regarding y and z as constants. To indicate that f is a function of more than one variable, the partial derivative is written using a curly d, ∂. More formally, the definition of the partial derivative is

$$\frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x + \delta x, y, z) - f(x, y, z)}{\delta x}. \tag{3.1}$$

Second derivatives, such as $\partial^2 f/\partial x^2$ and mixed derivatives, such as $\partial^2 f/\partial y \partial x$ can also be defined: the mixed derivative means that f is differentiated with respect to x regarding y as a constant, and then differentiated with respect to y regarding x as a constant. An important property of this mixed derivative, or cross-derivative, is that the order of the two derivatives does not matter, i.e.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}, \tag{3.2}$$

provided that these second partial derivatives exist and are continuous.

The main applications of partial differentiation are in the following sections. One additional application is in finding the maximum or minimum of a function of more than one variable: when a function is at a maximum or a minimum, all of its partial derivatives are zero.

Example 3.1

The function $f(x, y, z)$ is defined by $f = x^2 + xy \sin z - yz$. Find the partial derivatives of f with respect to x, y and z and verify the result (3.2) that the order of partial differentiation does not matter.

Differentiating f with respect to x, y and z in turn gives the three partial derivatives

$$\frac{\partial f}{\partial x} = 2x + y \sin z, \qquad \frac{\partial f}{\partial y} = x \sin z - z, \qquad \frac{\partial f}{\partial z} = xy \cos z - y.$$

Differentiating $\partial f/\partial y$ with respect to x gives

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \sin z,$$

and similarly, differentiating $\partial f/\partial x$ with respect to y gives

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \sin z.$$

In the same way it can be confirmed that the ordering of the cross-derivatives in x and z or in y and z does not matter.

Example 3.2

A rectangular box has height a, length b and breadth c, and is open at the top. If the volume of the box is fixed, deduce how a, b and c should be related to minimise the surface area of the box.

Let the fixed volume of the box be V, so $abc = V$. Two sides of the box have area ab, two have area ac and the base has area bc, so the total surface area is $A = 2ab + 2ac + bc$. Using the constraint $abc = V$, a can be eliminated so that $A = 2V/c + 2V/b + bc$. At a maximum or a minimum, the partial derivatives of A with respect to b and c must both be zero. This gives the two simultaneous equations

$$-2V/b^2 + c = 0, \qquad -2V/c^2 + b = 0.$$

From the first equation, $c = 2V/b^2$ and substituting this into the second equation gives the solution for b, $b^3 = 2V$ or $b = (2V)^{1/3}$, and similarly $c = (2V)^{1/3}$. The condition $abc = V$ gives $a = V/((2V)^{2/3}) = (2V)^{1/3}/2$. Therefore in the arrangement that minimises the surface area, $b = c$ and $a = b/2$, so the height of the box is half its length.

Note that we have not shown that this is a minimum and not a maximum. However, common sense suggests that since the area would be very large if the box were tall and thin, the solution found probably does represent a minimum.

3.1.2 Taylor series in more than one variable

The Taylor series for an infinitely differentiable function $f(x)$ of a single variable is

$$\begin{aligned} f(x) &= f(a) + (x-a)\frac{df}{dx}(a) + \frac{(x-a)^2}{2!}\frac{d^2 f}{dx^2}(a) + \dots \\ &= \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!}\frac{d^n f}{dx^n}(a). \end{aligned} \tag{3.3}$$

This can also be written as

$$\delta f = \delta x \frac{df}{dx} + \frac{(\delta x)^2}{2!}\frac{d^2 f}{dx^2} + \dots, \tag{3.4}$$

where $\delta x = (x - a)$ is a small perturbation and $\delta f = f(x) - f(a)$ is the corresponding perturbation in the value of the function.

Taylor series can also be constructed for functions of more than one variable. For a function $f(x, y)$ of two independent variables, the analogous formula is

$$\delta f = \delta x \frac{\partial f}{\partial x} + \delta y \frac{\partial f}{\partial y} + \frac{(\delta x)^2}{2!}\frac{\partial^2 f}{\partial x^2} + \frac{(\delta y)^2}{2!}\frac{\partial^2 f}{\partial y^2} + \delta x \delta y \frac{\partial^2 f}{\partial x \partial y} + \dots. \tag{3.5}$$

In the following sections we will make use of the Taylor series for a function $f(x, y, z)$ of three variables, but in all cases only the linear terms, that is, only those that only involve a single power of δx, δy or δz, will be needed:

$$\delta f = \delta x \frac{\partial f}{\partial x} + \delta y \frac{\partial f}{\partial y} + \delta z \frac{\partial f}{\partial z} + \dots. \tag{3.6}$$

Taylor series can be useful for approximating functions, as in the following example.

Example 3.3

Find an approximate value for the function $f(x, y, z) = 2x + (1 + y)\sin z$ at the point $x = 0.1$, $y = 0.2$, $z = 0.3$.

The function f takes the value 0 at the point $(0, 0, 0)$. Near to this point, the function can be approximated by its Taylor series expansion. To do this, the three partial derivatives of f evaluated at $(0, 0, 0)$ are required. These are

$$\frac{\partial f}{\partial x} = 2, \qquad \frac{\partial f}{\partial y} = \sin z = 0, \qquad \frac{\partial f}{\partial z} = (1 + y)\cos z = 1.$$

Hence the Taylor expansion (3.6) is

$$\delta f = 2\delta x + \delta z + \dots,$$

which at the point $(0.1, 0.2, 0.3)$ gives the approximate value $f \approx 0.5$ (the correct value to four decimal places is 0.5546).

3.2 Gradient of a scalar field

In Section 1.6 the concept of a scalar field was introduced as a scalar quantity which is a function of position in space. A scalar field f can be visualised in terms of the ***level surfaces*** or ***isosurfaces*** on which f is constant. The ***gradient*** of the scalar field f is a vector field, with a direction that is perpendicular to the level surfaces, pointing in the direction of increasing f, with a magnitude equal to the rate of change of f in this direction (Figure 3.1).

The gradient of a scalar field f can be written as $\operatorname{grad} f$, but the gradient is so important that a special symbol for grad, $\boldsymbol{\nabla}$, is used, so $\operatorname{grad} f = \boldsymbol{\nabla} f$. This symbol is sometimes referred to as 'del' or 'nabla'.

The gradient of f can also be defined in a Cartesian coordinate system in terms of the partial derivatives of f:

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial x}\boldsymbol{e}_1 + \frac{\partial f}{\partial y}\boldsymbol{e}_2 + \frac{\partial f}{\partial z}\boldsymbol{e}_3. \tag{3.7}$$

We will now show that these two definitions are equivalent, by showing that the vector $\boldsymbol{\nabla} f$ defined in (3.7) satisfies the two conditions of being perpendicular to the level surfaces and with magnitude equal to the rate of change of f in this direction.

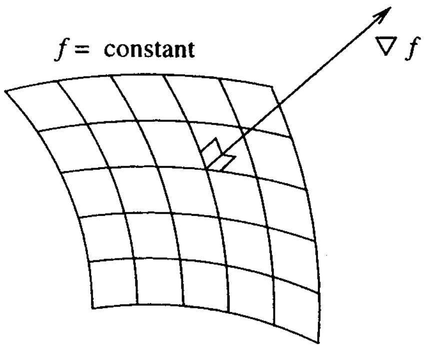


Fig. 3.1. The gradient of f is a vector perpendicular to the surface f = constant.

Consider an infinitesimal change in position in space from $\boldsymbol{r}$ to $\boldsymbol{r}+\boldsymbol{dr}$. This results in a small change in the value of the scalar field f from f to $f + df$, where, from (3.6),

$$\begin{aligned} df &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \cdot (dx, dy, dz) \\ &= \boldsymbol{\nabla} f \cdot \boldsymbol{dr}. \end{aligned} \tag{3.8}$$

Now suppose that $\boldsymbol{dr}$ lies in the surface f = constant (Figure 3.2). In this case the change in the value of f must be zero, so we have $df = \boldsymbol{\nabla} f \cdot \boldsymbol{dr} = 0$. Now in general $\boldsymbol{\nabla} f \neq \boldsymbol{0}$ and $\boldsymbol{dr} \neq \boldsymbol{0}$, so the two vectors $\boldsymbol{\nabla} f$ and $\boldsymbol{dr}$ must be perpendicular. Since $\boldsymbol{dr}$ is in the level surface f = constant, the vector $\boldsymbol{\nabla} f$ must be perpendicular to the level surface.

So the vector $\boldsymbol{\nabla} f$ defined by (3.7) has the correct direction; it remains to be shown that it has the correct magnitude. This is achieved by using (3.8) with $\boldsymbol{dr} = \boldsymbol{n}\,ds$, where $\boldsymbol{n}$ is the unit normal to the level surface and s is a distance measured along the normal. In this case, $df = \boldsymbol{\nabla} f \cdot \boldsymbol{n}\,ds = |\boldsymbol{\nabla} f|\,ds$, since $\boldsymbol{\nabla} f$ and $\boldsymbol{n}$ are parallel and $|\boldsymbol{n}| = 1$. Hence the magnitude of $\boldsymbol{\nabla} f$ is

$$|\boldsymbol{\nabla} f| = \frac{df}{ds}, \tag{3.9}$$

which is the rate of change of f with position along the normal.

From (3.8) it follows that $\boldsymbol{\nabla} f$ can be used to find the rate of change of f in any direction. To find the rate of change of f in the direction of the unit vector $\boldsymbol{u}$, set $\boldsymbol{dr} = \boldsymbol{u}\,ds$ where ds is the distance along $\boldsymbol{u}$. Then $df = \boldsymbol{\nabla} f \cdot \boldsymbol{u}\,ds$ and so

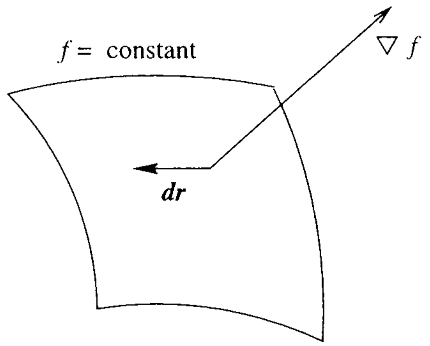


Fig. 3.2. Grad f is perpendicular to any vector $\boldsymbol{dr}$ lying in the surface f = constant.

$$\frac{df}{ds} = \boldsymbol{\nabla} f \cdot \boldsymbol{u}. \tag{3.10}$$

This is the rate of change of f in the direction of the unit vector $\boldsymbol{u}$, and is called the *directional derivative* of f. This can also be written as

$$\frac{df}{ds} = |\boldsymbol{\nabla} f| \cos\theta, \tag{3.11}$$

where θ is the angle between $\boldsymbol{\nabla} f$ and the unit vector $\boldsymbol{u}$. Since $-1 \leq \cos\theta \leq 1$, it follows that the magnitude of $\boldsymbol{\nabla} f$ is equal to the maximum rate of change of f with position.

The symbol $\boldsymbol{\nabla}$ can be interpreted as a vector differential operator,

$$\boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right), \tag{3.12}$$

where the term *operator* means that $\boldsymbol{\nabla}$ only has a meaning when it acts on some other quantity.

The gradient has many important applications. These include finding normals to surfaces and obtaining the rates of change of functions in any direction, as in the following examples.

Example 3.4

Find the unit normal $\boldsymbol{n}$ to the surface $x^2 + y^2 - z = 0$ at the point $(1, 1, 2)$.

Define $f(x, y, z) = x^2 + y^2 - z = 0$, so the surface is $f = 0$. Then $\boldsymbol{\nabla} f = (2x, 2y, -1)$. At the point $(1, 1, 2)$, $\boldsymbol{\nabla} f = (2, 2, -1)$. This is a vector normal to the surface. To find the unit normal we need to divide by the magnitude, which is $(2^2 + 2^2 + 1^2)^{1/2} = 3$ so $\boldsymbol{n} = \boldsymbol{\nabla} f / |\boldsymbol{\nabla} f| = (2/3, 2/3, -1/3)$. Note that the unit normal is not uniquely defined: the vector $-\boldsymbol{n} = (-2/3, -2/3, 1/3)$ is also a unit normal to the surface.

Example 3.5

Find the directional derivative of the scalar field $f = 2x+y+z^2$ in the direction of the vector (1,1,1), and evaluate this at the origin.

The gradient of f is $\nabla f = (2, 1, 2z)$. To find the directional derivative, we must take the dot product with the unit vector in the direction of $(1, 1, 1)$ which is $\boldsymbol{u} = (1, 1, 1)/\sqrt{3}$. The directional derivative is then $\nabla f \cdot \boldsymbol{u} = (3 + 2z)/\sqrt{3}$. At the origin, $x = y = z = 0$, the directional derivative takes the value $\sqrt{3}$.

3.2.1 Gradients, conservative fields and potentials

There is a very important link between the gradient of a scalar field and the concept of a conservative vector field defined in Section 2.2.3. Recall that a conservative vector field is one in which the line integral along a curve connecting two points does not depend on the path taken. The connection between gradients and conservative fields is given by the following theorem.

Theorem 3.1

Suppose that a vector field $\boldsymbol{F}$ is related to a scalar field ϕ by $\boldsymbol{F} = \nabla\phi$ and $\nabla\phi$ exists everywhere in some region D. Then $\boldsymbol{F}$ is conservative within D. Conversely, if $\boldsymbol{F}$ is conservative, then $\boldsymbol{F}$ can be written as the gradient of a scalar field, $\boldsymbol{F} = \nabla\phi$.

Proof

Suppose that $\boldsymbol{F} = \nabla\phi$. Then the line integral of $\boldsymbol{F}$ along a curve C connecting two points A and B is

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_C \nabla\phi \cdot d\boldsymbol{r}.$$

Using (3.8) this can be written as

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_C d\phi = [\phi]_A^B = \phi(B) - \phi(A),$$

where the line integral has been evaluated simply using ϕ as the parameter. Since this result only depends on the end points of C, $\boldsymbol{F}$ is conservative.

Conversely, suppose that $\boldsymbol{F}$ is conservative. Then a scalar field $\phi(\boldsymbol{r})$ can be defined as the line integral of $\boldsymbol{F}$ from the origin to the point $\boldsymbol{r}$:

$$\phi(\boldsymbol{r}) = \int_{\mathbf{0}}^{\boldsymbol{r}} \boldsymbol{F} \cdot d\boldsymbol{r}. \tag{3.13}$$

Since $\boldsymbol{F}$ is conservative, the value of ϕ does not depend on the path taken from $\mathbf{0}$ to $\boldsymbol{r}$, so ϕ is well defined. From the definition of an integral it then follows that an infinitesimal change in ϕ is given by

$$d\phi = \boldsymbol{F} \cdot \boldsymbol{dr}.$$

Comparing this with (3.8) shows that $\boldsymbol{F} \cdot \boldsymbol{dr} = \boldsymbol{\nabla}\phi \cdot \boldsymbol{dr}$. This must be true for any choice of $\boldsymbol{dr}$ and so $\boldsymbol{F} = \boldsymbol{\nabla}\phi$. □

If a vector field $\boldsymbol{F}$ is conservative, the corresponding scalar field ϕ which obeys $\boldsymbol{F} = \boldsymbol{\nabla}\phi$ is called the *potential* for $\boldsymbol{F}$. Note that the potential is not unique, since an arbitrary constant can be added to ϕ without affecting $\boldsymbol{\nabla}\phi$. This arbitrary constant corresponds to the arbitrary choice of the origin for the lower limit in the integral in the definition (3.13).

Example 3.6

Show that the vector field $\boldsymbol{F} = (2x + y, x, 2z)$ is conservative.

$\boldsymbol{F}$ is conservative if it can be written as the gradient of a scalar field ϕ. This gives the three equations

$$\frac{\partial \phi}{\partial x} = 2x + y, \qquad \frac{\partial \phi}{\partial y} = x, \qquad \frac{\partial \phi}{\partial z} = 2z.$$

Integrating the first of these equations with respect to x gives $\phi = x^2 + xy + h(y, z)$ where h is an arbitrary function of y and z, analogous to a constant of integration. The second equation forces the partial derivative of h with respect to y to be zero, so that h only depends on z. The third equation yields $dh/dz = 2z$, so $h(z) = z^2 + c$, where c is any constant. Therefore all three equations are satisfied by the potential function

$$\phi = x^2 + xy + z^2$$

and $\boldsymbol{F}$ is a conservative vector field.

3.2.2 Physical applications of the gradient

The gradient of a scalar field appears in many physical contexts. Two examples are given below.

- Let p denote the pressure within a gas. Then there is a force $\boldsymbol{F}$ acting on any volume element δV due to the pressure gradient, given by $\boldsymbol{F} = -\boldsymbol{\nabla}p\,\delta V$.
- A material has a constant thermal conductivity K and a variable temperature $T(\boldsymbol{r})$. Because of the temperature variation, heat flows from the hot regions to the cold regions. The heat flux $\boldsymbol{q}$ is a vector quantity, $\boldsymbol{q} = -K\boldsymbol{\nabla}T$.

EXERCISES

3.1 Find the gradient of the scalar field $f = xyz$, and evaluate it at the point (1,2,3). Hence find the directional derivative of f at this point in the direction of the vector $(1,1,0)$.

3.2 Find the unit normal to the surface $y = x + z^3$ at the point $(1,2,1)$.

3.3 Show that the gradient of the scalar field $\phi = r = |\boldsymbol{r}|$ is $\boldsymbol{r}/r$ and interpret this result geometrically.

3.4 Find the angle between the surfaces of the sphere $x^2 + y^2 + z^2 = 2$ and the cylinder $x^2 + y^2 = 1$ at a point where they intersect.

3.5 Find the gradient of the scalar field $f = yx^2 + y^3 - y$ and hence find the minima and maxima of f. Sketch the contours $f =$ constant and the vector field ∇f.

3.6 If $\boldsymbol{a}$ is a constant vector, find the gradient of $f = \boldsymbol{a} \cdot \boldsymbol{r}$ and interpret this result geometrically.

3.7 Determine whether or not the vector field $\boldsymbol{F} = (\sin y, x, 0)$ is conservative.

3.8 Consider the vector field $\boldsymbol{F} = (y/(x^2 + y^2), -x/(x^2 + y^2), 0)$. Show that $\boldsymbol{F}$ can be written as the gradient of a potential ϕ. Show also that the line integral of $\boldsymbol{F}$ around the unit circle $x^2 + y^2 = 1$ is non-zero. Explain why this result does not contradict Theorem 3.1.

3.3 Divergence of a vector field

This section introduces the first of two ways of differentiating a vector field, the divergence. The second way of differentiating a vector field, the curl, is defined in Section 3.4. Each of these quantities is defined in terms of an integral.

The *divergence* of a vector field $\boldsymbol{u}$ is a scalar field. Its value at a point P is defined by

$$\operatorname{div} \boldsymbol{u} = \lim_{\delta V \to 0} \frac{1}{\delta V} \oiint_{\delta S} \boldsymbol{u} \cdot \boldsymbol{n} \, dS, \tag{3.14}$$

where δV is a small volume enclosing P with surface δS and $\boldsymbol{n}$ is the outward pointing normal to δS. Physically, this corresponds to the amount of flux of the vector field $\boldsymbol{u}$ out of δV divided by the volume δV (Figure 3.3).

As in the case of the gradient, this physical definition leads to an equivalent definition in terms of the components of $\boldsymbol{u} = (u_1, u_2, u_3)$ in a Cartesian

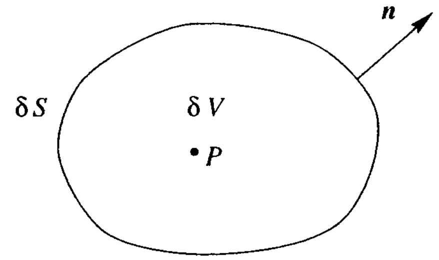


Fig. 3.3. Definition of the divergence. The small volume δV has surface δS and outward normal $\boldsymbol{n}$.

coordinate system. To derive this alternative form, take the volume δV to be a small rectangular box with sides of length δx, δy and δz and centred on a point (x, y, z) (Figure 3.4). It is assumed that the components of $\boldsymbol{u}$ have continuous partial derivatives.

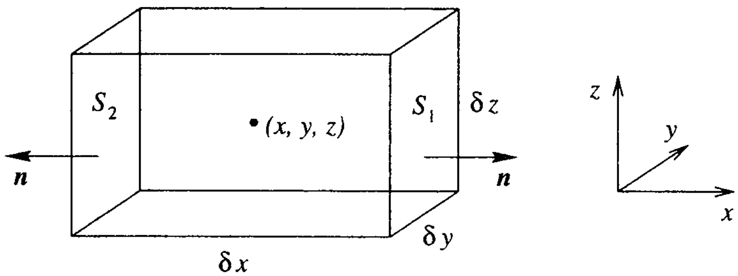


Fig. 3.4. Rectangular box used to obtain an expression for div $\boldsymbol{u}$ in Cartesian coordinates.

Since the rectangular box has six faces, there are six different contributions to the surface integral in (3.14). Consider first the contribution from the face labelled S_1 in Figure 3.4. This face is perpendicular to the x-axis, so the unit outward normal is $(1, 0, 0)$ and hence $\boldsymbol{u} \cdot \boldsymbol{n} = u_1$. The centre of face S_1 is at the point $(x + \delta x/2, y, z)$ and the area of the face is $\delta y\, \delta z$, so the contribution to the surface integral from this face is

$$\iint_{S_1} \boldsymbol{u} \cdot \boldsymbol{n}\, dS \approx u_1(x + \delta x/2, y, z)\, \delta y\, \delta z,$$

where we have used the fact that since the surface is small, the surface integral can be approximated as the value of $\boldsymbol{u} \cdot \boldsymbol{n}$ multiplied by the area of the surface.

A similar argument can be used to approximate the contribution to the surface integral from S_2, which is located at $(x - \delta x/2, y, z)$. The unit outward normal for S_2 is $(-1, 0, 0)$ so $\boldsymbol{u} \cdot \boldsymbol{n} = -u_1$ and the contribution to the surface integral is

$$\iint_{S_2} \boldsymbol{u} \cdot \boldsymbol{n}\, dS \approx -u_1(x - \delta x/2, y, z)\, \delta y\, \delta z.$$

Adding the contributions from these two surfaces and making use of the definition of the partial derivative (3.1), the combined contribution to the surface integral is

$$\begin{aligned}\iint_{S_1+S_2} \boldsymbol{u} \cdot \boldsymbol{n}\, dS &\approx \left(u_1(x + \frac{\delta x}{2}, y, z) - u_1(x - \frac{\delta x}{2}, y, z)\right) \delta y\, \delta z \\ &\approx \frac{\partial u_1}{\partial x}\, \delta x\, \delta y\, \delta z \\ &\approx \frac{\partial u_1}{\partial x}\, \delta V.\end{aligned}$$

Hence the contribution to $\operatorname{div} \boldsymbol{u}$ defined in (3.14) from surfaces S_1 and S_2 is $\partial u_1/\partial x$. Note that this is now exact since the divergence is defined by taking the limit $\delta V \to 0$. Similarly, the contribution to $\operatorname{div} \boldsymbol{u}$ from the two surfaces perpendicular to the y-axis is $\partial u_2/\partial y$ and that from the surfaces perpendicular to the z-axis is $\partial u_3/\partial z$. These are found simply by permuting the x-, y- and z-axes. Finally, adding all six contributions together gives the definition of $\operatorname{div} \boldsymbol{u}$ in terms of the Cartesian components of $\boldsymbol{u}$:

$$\operatorname{div} \boldsymbol{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}. \tag{3.15}$$

The divergence of $\boldsymbol{u}$ can also be written in terms of the differential operator $\boldsymbol{\nabla}$ defined in (3.12), since

$$\operatorname{div} \boldsymbol{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \cdot (u_1, u_2, u_3) = \boldsymbol{\nabla} \cdot \boldsymbol{u}. \tag{3.16}$$

The form $\boldsymbol{\nabla} \cdot \boldsymbol{u}$ will be used to indicate the divergence of $\boldsymbol{u}$ in the remainder of this book.

Example 3.7

Find the divergence of the vector field $\boldsymbol{u} = \boldsymbol{r}$.

The components of $\boldsymbol{u} = \boldsymbol{r}$ are $\boldsymbol{u} = (x, y, z)$. The divergence of $\boldsymbol{u}$ is therefore

$$\operatorname{div} \boldsymbol{u} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3. \tag{3.17}$$

3.3.1 Physical interpretation of divergence

The physical definition of the divergence (3.14) gives an intuitive meaning in terms of the flux of the vector field out of a small closed surface. This can also be interpreted as the rate of 'expansion' or 'stretching' of the vector field. Consider for example the simple vector field $\boldsymbol{u} = (x, 0, 0)$. This vector field only has a component in the x direction and it is sketched in Figure 3.5(a). It is useful to think of vector fields as representing the motion of a gas. Figure 3.5(a) then represents a gas which is expanding, and the divergence of u, from (3.15), is 1 everywhere. The vector field $\boldsymbol{v} = -\boldsymbol{u} = (-x, 0, 0)$ is shown in Figure 3.5(b). This vector field is contracting, and its divergence is $\boldsymbol{\nabla} \cdot \boldsymbol{v} = -1$. Finally consider the vector field $\boldsymbol{w} = (0, x, 0)$, sketched in Figure 3.5(c). This vector field is neither expanding nor contracting, and its divergence is zero. A vector field $\boldsymbol{w}$ for which $\boldsymbol{\nabla} \cdot \boldsymbol{w} = 0$ everywhere is said to be *solenoidal*.

3.3.2 Laplacian of a scalar field

Suppose that a scalar field ϕ is twice differentiable. Then the gradient of ϕ is a differentiable vector field $\boldsymbol{\nabla}\phi$, so we can take the divergence of $\boldsymbol{\nabla}\phi$ and obtain another scalar field. This scalar field, $\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\phi$ is called the *Laplacian* of ϕ and has its own symbol, $\nabla^2\phi$, so

$$\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\phi = \nabla^2\phi.$$

The Laplacian of ϕ is often referred to as 'del squared ϕ'. The formula for $\nabla^2\phi$ can be found by combining the formulae for div (3.15) and grad (3.7),

$$\begin{aligned} \nabla^2\phi &= \frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial z}\right) \\ &= \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}. \end{aligned} \qquad (3.18)$$

Thus the Laplacian of ϕ is just the sum of the second partial derivatives of ϕ.

The Laplacian can also act on a vector quantity, in which case the result is a vector whose components are the Laplacians of the components of the original vector:

$$\nabla^2\boldsymbol{u} = \left(\nabla^2 u_1, \nabla^2 u_2, \nabla^2 u_3\right). \qquad (3.19)$$

The Laplacian is a very important quantity, occurring in many physical applications including heat transfer and wave motion. These applications will be considered in Chapter 8. The equation $\nabla^2\phi = 0$ is known as *Laplace's equation*.

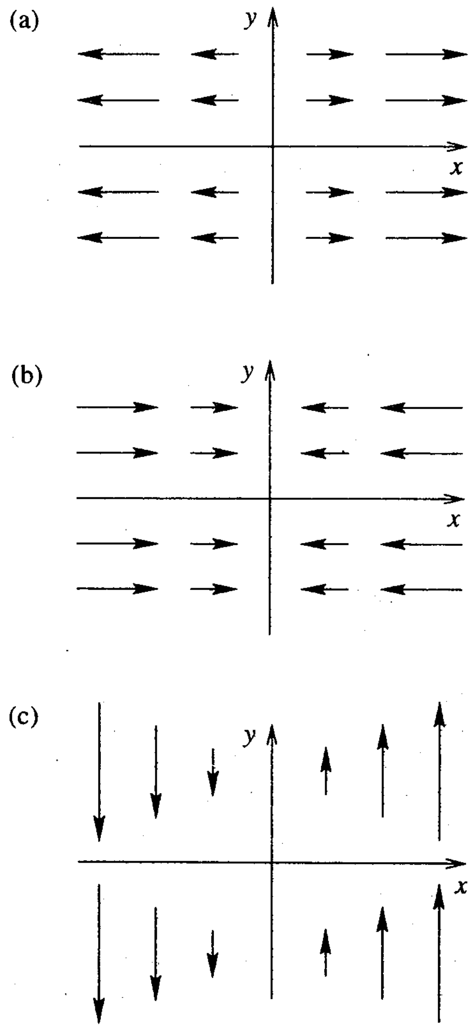


Fig. 3.5. The three vector fields (a) $\boldsymbol{u} = (x, 0, 0)$, (b) $\boldsymbol{v} = (-x, 0, 0)$, (c) $\boldsymbol{w} = (0, x, 0)$.

3.4 Curl of a vector field

The *curl* of a vector field $\boldsymbol{u}$ is a vector field. Its component in the direction of the unit vector $\boldsymbol{n}$ is

$$\boldsymbol{n} \cdot \operatorname{curl} \boldsymbol{u} = \lim_{\delta S \to 0} \frac{1}{\delta S} \oint_{\delta C} \boldsymbol{u} \cdot d\boldsymbol{r}, \tag{3.20}$$

where δS is a small surface element perpendicular to $\boldsymbol{n}$, δC is the closed curve forming the boundary of δS and δC and $\boldsymbol{n}$ are oriented in a right-handed sense, as shown in Figure 3.6. Note that this has a similar form to the definition of divergence (3.14), but with a line integral instead of a surface integral.

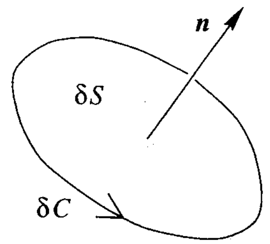


Fig. 3.6. Definition of the curl of a vector field. The small surface δS is enclosed by the curve δC and has unit normal vector $\boldsymbol{n}$.

To obtain an expression for $\operatorname{curl} \boldsymbol{u}$ in terms of the components (u_1, u_2, u_3) of $\boldsymbol{u}$, choose $\boldsymbol{n} = \boldsymbol{e}_3$, the unit vector in the z direction, to determine the z component of $\operatorname{curl} \boldsymbol{u}$. The surface δS then lies in the x, y plane and can be chosen to be a small rectangle with sides of length δx, δy centred on the point (x, y, z), with area $\delta S = \delta x \, \delta y$. The right-hand rule means that the line integral in (3.20) must be taken in the anticlockwise direction. The line integral then has four sections, as shown in Figure 3.7. It is assumed that the vector field $\boldsymbol{u}$ is differentiable with continuous partial derivatives.

Consider first the section C_1 of the line integral, which has its centre at the point $(x, y - \delta y/2, z)$. Here, the line integral is directed in the positive x direction, so $\boldsymbol{u} \cdot d\boldsymbol{r} = u_1 \, dx$. Since the length δx is small, the contribution to the line integral is approximately

$$\int_{C_1} \boldsymbol{u} \cdot d\boldsymbol{r} \approx u_1(x, y - \delta y/2, z) \, \delta x.$$

Similarly, on C_3, centred at $(x, y + \delta y/2, z)$, the integral is directed in the negative x direction, so

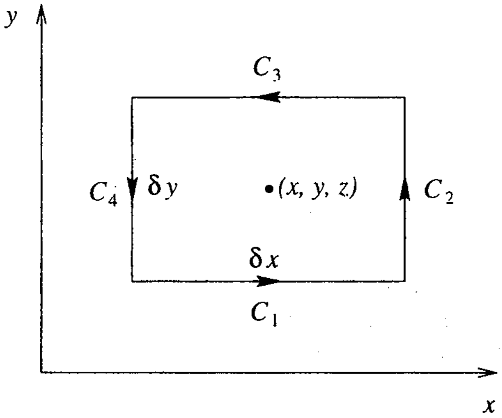


Fig. 3.7. A rectangle of four line segments is used to find an expression for curl $\boldsymbol{u}$ in Cartesian coordinates.

$$\int_{C_3} \boldsymbol{u} \cdot d\boldsymbol{r} \approx -u_1(x, y + \delta y/2, z)\,\delta x.$$

Adding these two contributions together gives

$$\begin{aligned}\int_{C_1+C_3} \boldsymbol{u} \cdot d\boldsymbol{r} &\approx (u_1(x, y - \delta y/2, z) - u_1(x, y + \delta y/2, z))\,\delta x \\ &\approx -\frac{\partial u_1}{\partial y}\,\delta y\,\delta x.\end{aligned}$$

Proceeding in a similar way with the line integrals along C_2, located at $(x + \delta x/2, y, z)$ with $d\boldsymbol{r}$ directed in the positive y direction, and C_4, located at $(x - \delta x/2, y, z)$ with $d\boldsymbol{r}$ directed in the negative y direction, we obtain

$$\begin{aligned}\int_{C_2+C_4} \boldsymbol{u} \cdot d\boldsymbol{r} &\approx (u_2(x + \delta x/2, y, z) - u_2(x - \delta x/2, y, z))\,\delta y \\ &\approx \frac{\partial u_2}{\partial x}\,\delta x\,\delta y.\end{aligned}$$

Adding together all four contributions, dividing by δS and taking the limit $\delta S \to 0$ gives the z component of curl $\boldsymbol{u}$:

$$\boldsymbol{e}_3 \cdot \operatorname{curl} \boldsymbol{u} = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}. \tag{3.21}$$

The other components can be found by permuting x, y and z cyclically ($x \to y \to z \to x$, $u_1 \to u_2 \to u_3 \to u_1$), giving

$$\operatorname{curl} \boldsymbol{u} = \left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right). \tag{3.22}$$

Notice that there is a similarity between this formula and that for the cross product of two vectors (1.4). Curl $\boldsymbol{u}$ can also be written in terms of a determinant,

$$\operatorname{curl} \boldsymbol{u} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u_1 & u_2 & u_3 \end{vmatrix}$$

provided that the determinant is expanded so that the partial derivatives act on the components of $\boldsymbol{u}$. This can also be written as the cross product of the differential operator $\boldsymbol{\nabla}$ and the vector $\boldsymbol{u}$, so

$$\operatorname{curl} \boldsymbol{u} = \boldsymbol{\nabla} \times \boldsymbol{u}.$$

The notation $\boldsymbol{\nabla} \times \boldsymbol{u}$ will be used henceforth.

Example 3.8

The vector field $\boldsymbol{u}$ is defined by $\boldsymbol{u} = (xy, z + x, y)$. Calculate $\boldsymbol{\nabla} \times \boldsymbol{u}$ and find the points where $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$.

The components of $\boldsymbol{\nabla} \times \boldsymbol{u}$ are found using (3.22):

$$\begin{aligned} \boldsymbol{\nabla} \times \boldsymbol{u} &= \left(\frac{\partial y}{\partial y} - \frac{\partial(z+x)}{\partial z}, \frac{\partial(xy)}{\partial z} - \frac{\partial y}{\partial x}, \frac{\partial(z+x)}{\partial x} - \frac{\partial(xy)}{\partial y} \right) \\ &= (1-1, 0-0, 1-x) = (0, 0, 1-x). \end{aligned}$$

Hence $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$ on the plane $x = 1$.

3.4.1 Physical interpretation of curl

From the physical definition of $\boldsymbol{\nabla} \times \boldsymbol{u}$ given in (3.20) and Figure 3.6 it is clear that $\boldsymbol{\nabla} \times \boldsymbol{u}$ is related to the rotation or twisting of the vector field $\boldsymbol{u}$.

Consider the three simple vector fields shown in Figure 3.5. For the first of these, $\boldsymbol{u} = (x, 0, 0)$, the vector field is expanding but there is no sense of rotation, and computing the curl gives $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$. A vector field $\boldsymbol{u}$ for which $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$ everywhere is said to be *irrotational*. Similarly, the second example, $\boldsymbol{v} = (-x, 0, 0)$, is also irrotational. For the third example, $\boldsymbol{w} = (0, x, 0)$, Figure 3.5(c), $\boldsymbol{\nabla} \times \boldsymbol{w} = (0, 0, 1)$, so there is a component of $\boldsymbol{\nabla} \times \boldsymbol{w}$ in the z direction, out of the page. The vector field $\boldsymbol{w}$ has a rotation associated with it in the following sense. Think of $\boldsymbol{w}$ as the velocity of a fluid. Then a small particle placed in this fluid will rotate in an anticlockwise sense as it moves with the fluid, since at any point the velocity component in the y direction to the right of the particle is greater than that on the left. This rotation is about an axis in the z direction, which is in the direction of $\nabla \times \boldsymbol{w}$. The vector $\nabla \times \boldsymbol{w}$ can therefore be related to the rotation of a small particle placed in the velocity field $\boldsymbol{w}$: the rate of rotation depends on the magnitude of $\nabla \times \boldsymbol{w}$, and the axis of rotation is in the direction of $\nabla \times \boldsymbol{w}$. In the context of the motion of a fluid, the curl of the velocity field is often referred to as the *vorticity* of the fluid. This relationship between rotation and curl is made more precise in the following section.

3.4.2 Relation between curl and rotation

Consider a rigid body rotating with angular velocity Ω. Then, as discussed in Section 1.3.1, the velocity $\boldsymbol{v}$ at any point can be written as $\boldsymbol{v} = \boldsymbol{\Omega} \times \boldsymbol{r}$, where the vector $\boldsymbol{\Omega}$ is directed along the axis of rotation (Figure 1.12).

Choosing the z-axis in the direction of $\boldsymbol{\Omega}$ gives $\boldsymbol{\Omega} = (0, 0, \Omega)$ and hence $\boldsymbol{v} = (0, 0, \Omega) \times (x, y, z) = (-\Omega y, \Omega x, 0)$. Computing the curl of this velocity field gives

$$\nabla \times \boldsymbol{v} = \left(-\frac{\partial(\Omega x)}{\partial z}, -\frac{\partial(\Omega y)}{\partial z}, \frac{\partial(\Omega x)}{\partial x} + \frac{\partial(\Omega y)}{\partial y}\right) = (0, 0, 2\Omega).$$

Hence $\nabla \times \boldsymbol{v} = 2\boldsymbol{\Omega}$, i.e. the curl of the velocity field is equal to twice the rotation rate.

3.4.3 Curl and conservative vector fields

Suppose that a vector field $\boldsymbol{u}$ is related to a scalar field ϕ by $\boldsymbol{u} = \nabla\phi$. Recall from Theorem 3.1 that this means that $\boldsymbol{u}$ is conservative. Now consider the curl of $\boldsymbol{u}$,

$$\begin{aligned} \nabla \times \boldsymbol{u} &= \boldsymbol{e}_1\left(\frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial z}\right) - \frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial y}\right)\right) + \boldsymbol{e}_2\left(\frac{\partial}{\partial z}\left(\frac{\partial \phi}{\partial x}\right) - \frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial z}\right)\right) \\ &\quad + \boldsymbol{e}_3\left(\frac{\partial}{\partial x}\left(\frac{\partial \phi}{\partial y}\right) - \frac{\partial}{\partial y}\left(\frac{\partial \phi}{\partial x}\right)\right). \end{aligned}$$

Since the ordering of the cross derivatives of ϕ does not matter (see Section 3.1.1), all of these terms cancel out, giving the result that

$$\nabla \times \nabla\phi = \mathbf{0}. \tag{3.23}$$

Thus any vector field that can be written as the gradient of a scalar field is irrotational.

The converse of this result is also true, so that any irrotational vector field is conservative. This result will be proved in Section 5.2. In combination with Theorem 3.1, this means that the following three statements are equivalent:

- $\boldsymbol{u}$ can be written as the gradient of a potential: $\boldsymbol{u} = \boldsymbol{\nabla}\phi$.
- $\boldsymbol{u}$ is irrotational: $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$.
- $\boldsymbol{u}$ is conservative: the line integral of $\boldsymbol{u}$ around any closed curve is zero.

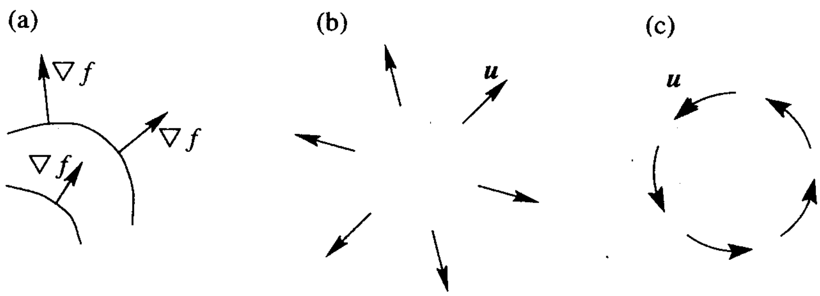


Fig. 3.8. Physical picture associated with (a) gradient, (b) divergence, (c) curl.

Summary of Chapter 3

- The *gradient* of a scalar field f, written grad f or ∇f, is a vector field perpendicular to the surfaces $f =$ constant, pointing in the direction of increasing f, with magnitude equal to the rate of change of f in this direction.
- The components of ∇f are the partial derivatives of f:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right).$$

- The *directional derivative* of f in the direction of the unit vector $\boldsymbol{u}$ is $\nabla f \cdot \boldsymbol{u}$.
- A vector field $\boldsymbol{F}$ is conservative if and only if it can be written as the gradient of a scalar field, $\boldsymbol{F} = \nabla \phi$. The function ϕ is called the *potential* for $\boldsymbol{F}$.
- The *divergence* of a vector field $\boldsymbol{u}$, written div $\boldsymbol{u}$ or $\nabla \cdot \boldsymbol{u}$, is a scalar field,

$$\nabla \cdot \boldsymbol{u} = \lim_{\delta V \to 0} \frac{1}{\delta V} \oiint_{\delta S} \boldsymbol{u} \cdot \boldsymbol{n}\, dS = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}.$$

- The divergence of $\boldsymbol{u}$ corresponds to the amount of stretching or expansion associated with $\boldsymbol{u}$. If $\nabla \cdot \boldsymbol{u} = 0$, $\boldsymbol{u}$ is said to be *solenoidal.*
- The *Laplacian* of a scalar field ϕ, written $\nabla^2 \phi$, is defined as $\nabla^2 \phi = \nabla \cdot \nabla \phi$.
- The *curl* of a vector field $\boldsymbol{u}$ is a vector field, curl $\boldsymbol{u}$ or $\nabla \times \boldsymbol{u}$. Its component in the direction of a unit vector $\boldsymbol{n}$ perpendicular to the surface element δS is

$$\boldsymbol{n} \cdot \nabla \times \boldsymbol{u} = \lim_{\delta S \to 0} \frac{1}{\delta S} \oint_{\delta C} \boldsymbol{u} \cdot d\boldsymbol{r}.$$

- In component form,

$$\nabla \times \boldsymbol{u} = \left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y}\right).$$

- Physically, $\nabla \times \boldsymbol{u}$ corresponds to the rotation or twist of $\boldsymbol{u}$. If $\nabla \times \boldsymbol{u} = \boldsymbol{0}$, $\boldsymbol{u}$ is said to be *irrotational.*
- If $\boldsymbol{u} = \nabla \phi$, then $\boldsymbol{u}$ is irrotational.
- The physical pictures corresponding to grad, div and curl are sketched in Figure 3.8.

EXERCISES

3.9 Find the gradient $\nabla\phi$ and the Laplacian $\nabla^2\phi$ for the scalar field $\phi = x^2 + xy + yz^2$.

3.10 Find the gradient and Laplacian of

$$\phi = \sin(kx)\sin(ly)\exp(\sqrt{k^2 + l^2}z).$$

3.11 Find the unit normal to the surface $xy^2 + 2yz = 4$ at the point $(-2, 2, 3)$.

3.12 For $\phi(x, y, z) = x^2+y^2+z^2+xy-3x$, find $\nabla\phi$ and find the minimum value of ϕ.

3.13 Find the equation of the plane which is tangent to the surface $x^2 + y^2 - 2z^3 = 0$ at the point $(1, 1, 1)$.

3.14 Find both the divergence and the curl of the vector fields
(a) $\boldsymbol{u} = (y, z, x)$;
(b) $\boldsymbol{v} = (xyz, z^2, x - y)$.

3.15 Show that both the divergence and the curl are linear operators, i.e. $\nabla\cdot(c\boldsymbol{u}+d\boldsymbol{v}) = c\nabla\cdot\boldsymbol{u}+d\nabla\cdot\boldsymbol{v}$ and $\nabla\times(c\boldsymbol{u}+d\boldsymbol{v}) = c\nabla\times\boldsymbol{u}+d\nabla\times\boldsymbol{v}$, where $\boldsymbol{u}$ and $\boldsymbol{v}$ are vector fields and c and d are constants.

3.16 For what values, if any, of the constants a and b is the vector field $\boldsymbol{u} = (y\cos x + axz, b\sin x + z, x^2 + y)$ irrotational?

3.17 (a) Show that $\boldsymbol{u} = (y^2z, -z^2\sin y + 2xyz, 2z\cos y + y^2x)$ is irrotational.
(b) Find the corresponding potential function.
(c) Hence find the value of the line integral of $\boldsymbol{u}$ along the curve $x = \sin\pi t/2$, $y = t^2 - t$, $z = t^4$, $0 \le t \le 1$.

4
Suffix Notation and its Applications

4.1 Introduction to suffix notation

This chapter introduces a powerful, compact notation for manipulating vector quantities. In the previous chapters, many of the vector expressions are awkward and cumbersome. This applies particularly to those expressions involving the cross product and the curl, such as the scalar triple product (1.8), the derivation of the alternative expression for the vector triple product (1.9) and the demonstration that $\nabla \times \nabla\phi = \mathbf{0}$ (3.23). Through the use of a new notation, *suffix notation*, such complicated expressions can be written much more concisely and many results can be proved more easily.

In this section, some simple vector equations are written using suffix notation. Consider first the equation $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b}$. This vector equation is equivalent to the three equations for the components of $\boldsymbol{c}$, $c_i = a_i + b_i$ for $i = 1, 2, 3$. In suffix notation, the equation is simply written

$$c_i = a_i + b_i$$

and it is understood that this equation holds for $i = 1$, 2 and 3. The suffix i is called a 'free suffix'. The choice of this free suffix is arbitrary, so the equation could equally well be written $c_j = a_j + b_j$ or $c_k = a_k + b_k$. However, for simplicity and clarity the suffix i will be used for the free suffix in a vector equation in this book. Note that the same free suffix must be used for each term in the equation. This is an important rule which must be followed for an expression in suffix notation to be meaningful: the free suffix must match in each term in the expression.

Now consider the dot product of two vectors, $\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$. This can be written more compactly as

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{j=1}^{3} a_j b_j.$$

In suffix notation, this is written simply as

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_j b_j, \tag{4.1}$$

where the repeated suffix j implies that the term is to be summed from $j = 1$ to $j = 3$. This is known as the *summation convention*: whenever a suffix is repeated in a single term in an equation, summation from 1 to 3 is understood. The repeated suffix is referred to as a 'dummy suffix', and must appear no more than twice in any term in an equation. The choice of the dummy suffix does not matter, so we can write $\boldsymbol{a} \cdot \boldsymbol{b} = a_j b_j = a_k b_k$, since each of these expressions is equivalent to $a_1 b_1 + a_2 b_2 + a_3 b_3$.

Next, suppose that an expression involves two dot products multiplied together, $(\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d})$. In order to indicate which vector is dotted with which, a different dummy suffix must be used for each of the dot products:

$$(\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d}) = a_j b_j c_k d_k.$$

Here, both j and k are repeated, so the summation convention implies summation over both j and k. Again the choice of dummy suffix is arbitrary, so for example we could have written $(\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d}) = a_l b_l c_m d_m$. However, it is essential that no suffix appears more than twice in any term, since this would lead to ambiguity.

Note that it is the suffices that indicate which vector is dotted with which, not the ordering of the components of the vectors. In fact, since the components are just multiplied together, the ordering of terms is arbitrary, so the expression $c_k a_j d_k b_j$ also means $(\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d})$.

Example 4.1

Write the suffix notation expression $a_j b_i c_j$ in ordinary vector notation.

The suffix j is repeated and is therefore a dummy suffix to be summed over. So $a_j b_i c_j$ means

$$\sum_{j=1}^{3} a_j b_i c_j = (\boldsymbol{a} \cdot \boldsymbol{c}) b_i,$$

which is the i component of the vector $(\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b}$.

Example 4.2

Write the vector equation

$$\boldsymbol{u} + (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{v} = |\boldsymbol{a}|^2(\boldsymbol{b} \cdot \boldsymbol{v})\boldsymbol{a}$$

in suffix notation.

First introduce a free suffix i, which is understood to run from 1 to 3, to write the equation in component form; also write $|\boldsymbol{a}|^2$ as a dot product:

$$u_i + (\boldsymbol{a} \cdot \boldsymbol{b})v_i = (\boldsymbol{a} \cdot \boldsymbol{a})(\boldsymbol{b} \cdot \boldsymbol{v})a_i.$$

Now introduce a dummy suffix, which is repeated and therefore summed from 1 to 3, for each of the dot products:

$$u_i + a_j b_j v_i = a_j a_j b_k v_k a_i.$$

Note that two different dummy suffices are used on the right-hand side to avoid ambiguity.

Example 4.3

Show that the product of two $N \times N$ matrices A and B, $C = AB$ can be written in suffix notation as $C_{ij} = A_{ik}B_{kj}$. Hence show that the trace of the matrix AB (defined as the sum of the elements on the diagonal) is the same as the trace of BA.

C_{ij} is the element in the ith row and jth column of the matrix C. The rule for matrix multiplication is that if $C = AB$ then the element C_{ij} is obtained by taking the ith row of A and the jth column of B and multiplying these together term by term, so

$$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} \ldots + A_{iN}B_{Nj} = \sum_{k=1}^{N} A_{ik}B_{kj} = A_{ik}B_{kj},$$

where the repeated index k implies the sum from 1 to N.

The trace of the matrix C is the sum of the elements on the diagonal,

$$Tr(C) = C_{11} + C_{22} \ldots + C_{NN} = C_{jj}.$$

The trace of AB is

$$Tr(AB) = Tr(A_{ik}B_{kj}) = A_{jk}B_{kj}.$$

Similarly the trace of BA is

$$\begin{aligned} Tr(BA) &= B_{jk}A_{kj} \\ &= A_{kj}B_{jk} \quad \text{(since order of terms does not matter)} \\ &= A_{jk}B_{kj} \quad \text{(relabelling } j \leftrightarrow k\text{)} \\ &= Tr(AB). \end{aligned}$$

Note that this proof makes use of the fact that the choice of label used for a dummy suffix is arbitrary, so the labels j and k can be interchanged.

4.2 The Kronecker delta δ_{ij}

The *Kronecker delta* is written δ_{ij} and is defined by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{4.2}$$

The suffices i and j can each take the values 1, 2 or 3, so δ_{ij} has nine elements. From the above definition it follows that three of these are equal to 1 ($\delta_{11} = \delta_{22} = \delta_{33} = 1$) while the remaining six elements are equal to 0 ($\delta_{12} = \delta_{13} = \delta_{21} = \delta_{23} = \delta_{31} = \delta_{32} = 0$). δ_{ij} is an example of an object called a *tensor.* Tensors are described in detail in Chapter 7, but for the time being it is simplest to think of δ_{ij} as the 3×3 identity matrix,

$$\delta_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

From the definition it is clear that δ_{ij} is symmetric, i.e. $\delta_{ij} = \delta_{ji}$.

Consider now the expression $\delta_{ij}a_j$. Notice that the suffix j is repeated, so by the summation convention, summation from $j = 1$ to 3 is understood. Hence

$$\delta_{ij}a_j = \sum_{j=1}^{3} \delta_{ij}a_j = \delta_{i1}a_1 + \delta_{i2}a_2 + \delta_{i3}a_3,$$

and it is clear that the result depends on the value of i. If $i = 1$, then $\delta_{i1} = 1$ while $\delta_{i2} = \delta_{i2} = 0$, so the right-hand side simplifies to a_1. Similarly, if $i = 2$ the result is a_2 and if $i = 3$ the result is a_3. In other words, the right-hand side simplifies to a_i, giving the important equation

$$\delta_{ij}a_j = a_i. \tag{4.3}$$

From the symmetry of δ_{ij} it follows that $\delta_{ij}a_i = a_j$. Because of this property of δ_{ij}, it is sometimes referred to as the 'substitution tensor', since its effect when multiplied by a_j is to replace the j with i.

There is a relationship between δ_{ij} and the dot product. The dot product of $\boldsymbol{a}$ and $\boldsymbol{b}$ can be written $\boldsymbol{a} \cdot \boldsymbol{b} = \delta_{ij}a_ib_j$. In this expression, both the i and the j suffices are repeated, so by the summation convention, both are to be summed from 1 to 3, giving a total of nine terms. However, because of the definition of δ_{ij}, only three of these terms (the ones with $i = j$) are non-zero, so

$$\delta_{ij}a_ib_j = \sum_{i=1}^{3}\sum_{j=1}^{3}\delta_{ij}a_ib_j = a_1b_1 + a_2b_2 + a_3b_3 = \boldsymbol{a} \cdot \boldsymbol{b}.$$

This result can also be demonstrated using (4.3), since $\delta_{ij}b_j = b_i$, so $\boldsymbol{a} \cdot \boldsymbol{b} = a_ib_i = a_i\delta_{ij}b_j = \delta_{ij}a_ib_j$.

Example 4.4

Evaluate δ_{jj}.

Since the suffix j is repeated, the summation convention implies that this expression must be summed from $j = 1$ to 3, so

$$\delta_{jj} = \sum_{j=1}^{3}\delta_{jj} = \delta_{11} + \delta_{22} + \delta_{33} = 3. \tag{4.4}$$

Example 4.5

Simplify $\delta_{ij}\delta_{jk}$.

Here the suffix j is repeated, and must therefore be summed over:

$$\delta_{ij}\delta_{jk} = \sum_{j=1}^{3}\delta_{ij}\delta_{jk} = \delta_{i1}\delta_{1k} + \delta_{i2}\delta_{2k} + \delta_{i3}\delta_{3k}.$$

The result depends on the values of i and k. If, for example, $i = 1$ and $k = 2$, we have $1\times0+0\times1+0\times0 = 0$; but if $i = 1$ and $k = 1$, we have $1\times1+0\times0+0\times0 = 1$. It is apparent that if i and k are different the result is 0, but if i and k are equal the result is 1. This result is therefore simply δ_{ik}, so the solution is $\delta_{ij}\delta_{jk} = \delta_{ik}$. Notice that this result is consistent with the substitution rule described above: the effect of δ_{ij} on δ_{jk} is to replace the j with i.

4.3 The alternating tensor ϵ_{ijk}

This section introduces the quantity which is used for writing cross products in suffix notation. This will prove extremely useful for manipulating expressions involving the cross product and the curl.

The *alternating tensor* is written ϵ_{ijk} and is defined by

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any of } i,\ j,\ k \text{ are equal,} \\ +1 & \text{if } (i,j,k) = (1,2,3),\ (2,3,1) \text{ or } (3,1,2), \\ -1 & \text{if } (i,j,k) = (1,3,2),\ (2,1,3) \text{ or } (3,2,1). \end{cases} \tag{4.5}$$

Since ϵ_{ijk} has three suffices, each of which can take any of the three values 1, 2 or 3, ϵ_{ijk} has 27 elements. However, from the above definition, all but six of these are zero. The six non-zero elements are $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$ and $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$.

There are two important symmetry properties of ϵ_{ijk} which follow directly from its definition:

- ϵ_{ijk} is unchanged if the suffices are reordered by moving them to the left and putting the first suffix third (a cyclic permutation of the suffices), i.e.

$$\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}. \tag{4.6}$$

- The sign of ϵ_{ijk} changes if any two of the suffices are interchanged, e.g.

$$\epsilon_{ijk} = -\epsilon_{jik}. \tag{4.7}$$

The relationship between ϵ_{ijk} and the cross product is as follows:

$$(\boldsymbol{a} \times \boldsymbol{b})_i = \epsilon_{ijk} a_j b_k. \tag{4.8}$$

In this equation, both j and k are repeated, so they are dummy suffices and must be summed over. To check that this agrees with the previous definition of $\boldsymbol{a} \times \boldsymbol{b}$, consider first the case $i = 1$. The right-hand side is then

$$\epsilon_{1jk} a_j b_k = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{1jk} a_j b_k.$$

Since ϵ_{ijk} is only non-zero when all three of its suffices are different, only the two terms $j = 2$, $k = 3$ and $j = 3$, $k = 2$ are non-zero in the double sum. Hence the right-hand side reduces to $\epsilon_{123} a_2 b_3 + \epsilon_{132} a_3 b_2 = a_2 b_3 - a_3 b_2$. This agrees with the previous definition (1.4) for the first component of the cross product. It can be seen by cyclic permutation of indices that the second and third components also agree.

There is also a relation between ϵ_{ijk} and the determinant of a 3×3 matrix. This can be written

$$|M| = \epsilon_{ijk} M_{1i} M_{2j} M_{3k}. \tag{4.9}$$

A related formula is

$$\epsilon_{pqr}|M| = \epsilon_{ijk} M_{pi} M_{qj} M_{rk}. \tag{4.10}$$

An expression for the scalar triple product $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c}$ can be deduced in suffix notation as follows:

$$\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = a_i(\boldsymbol{b} \times \boldsymbol{c})_i = a_i \epsilon_{ijk} b_j c_k = \epsilon_{ijk} a_i b_j c_k. \tag{4.11}$$

A comparison of this neat and elegant expression with the cumbersome formula in terms of the components (1.8) shows the power of suffix notation. The properties of the scalar triple product can also be deduced using suffix notation, as in the following examples.

Example 4.6

Use suffix notation to show that $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c}$.

$$\begin{aligned}
\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} &= \epsilon_{ijk} a_i b_j c_k \\
&= \epsilon_{kij} a_i b_j c_k \quad (\text{using } \epsilon_{ijk} = \epsilon_{kij}) \\
&= (\boldsymbol{a} \times \boldsymbol{b})_k c_k \\
&= \boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c}.
\end{aligned}$$

Example 4.7

Show that $\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} = \boldsymbol{b} \cdot \boldsymbol{c} \times \boldsymbol{a}$.

The demonstration of this result is very similar:

$$\begin{aligned}
\boldsymbol{a} \cdot \boldsymbol{b} \times \boldsymbol{c} &= \epsilon_{ijk} a_i b_j c_k \\
&= \epsilon_{jki} a_i b_j c_k \quad (\text{using } \epsilon_{ijk} = \epsilon_{jki}) \\
&= b_j \epsilon_{jki} c_k a_i \quad (\text{just rearranging terms}) \\
&= b_j (\boldsymbol{c} \times \boldsymbol{a})_j \\
&= \boldsymbol{b} \cdot \boldsymbol{c} \times \boldsymbol{a}.
\end{aligned}$$

Example 4.8

Evaluate ϵ_{iik}.

Since $\epsilon_{ijk} = 0$ if any of i, j, k are equal, it follows that $\epsilon_{iik} = 0$.

Example 4.9

Evaluate $\epsilon_{ijk}\epsilon_{ijk}$.

In this expression all three suffices i, j and k are repeated, and must therefore be summed over, giving a total of 27 terms. Only six of these terms are non-zero, so $\epsilon_{ijk}\epsilon_{ijk} = \epsilon_{123}^2 + \epsilon_{132}^2 + \epsilon_{213}^2 + \epsilon_{231}^2 + \epsilon_{312}^2 + \epsilon_{321}^2 = 6$.

4.4 Relation between ϵ_{ijk} and δ_{ij}

An important relationship between ϵ_{ijk} and δ_{ij} is the following equation:

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{4.12}$$

This equation has four free suffices (i, j, l and m) and therefore represents 81 different equations! The left-hand side is summed over k, because the suffix k appears twice.

The result (4.12) can be demonstrated by the following argument: since the three coordinate axes are equivalent, we need only consider the case $i = 1$. Consider now the possible values for j:

1. If $j = 1$, $\epsilon_{ijk} = \epsilon_{11k} = 0$ and so the l.h.s. is zero; the r.h.s. is $\delta_{1l}\delta_{1m} - \delta_{1m}\delta_{1l}$ which is also zero since the two δ terms cancel.
2. If $j = 2$, $\epsilon_{ijk} = \epsilon_{12k} = 0$ unless $k = 3$, so only the $k = 3$ term contributes to the sum. When $k = 3$, the term ϵ_{klm} is zero unless l and m are 1 and 2. Therefore the l.h.s. takes the value $+1$ if $l = 1$ and $m = 2$, -1 if $l = 2$ and $m = 1$, and zero otherwise. Now the r.h.s. is $\delta_{1l}\delta_{2m} - \delta_{1m}\delta_{2l}$. This is also equal to $+1$ when $l = 1$ and $m = 2$ (from the first term), -1 when $l = 2$ and $m = 1$ (from the second term) and zero otherwise.
3. If $j = 3$, an equivalent argument to the case $j = 2$ applies; the details are left to the reader.

Equation (4.12) is very useful for simplifying expressions involving two cross products.

Example 4.10

Derive the formula (1.9) for the expansion of the vector triple product using suffix notation.

$$(\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}))_i = \epsilon_{ijk} a_j (\boldsymbol{b} \times \boldsymbol{c})_k$$

(writing the first cross product in suffix notation)

$$\begin{aligned}
&= \epsilon_{ijk} a_j \epsilon_{klm} b_l c_m \\
&\quad \text{(writing the second cross product in suffix notation)} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) a_j b_l c_m \quad \text{(using (4.12))} \\
&= a_m b_i c_m - a_j b_j c_i \quad \text{(using (4.3))} \\
&= (\boldsymbol{a} \cdot \boldsymbol{c}) b_i - (\boldsymbol{a} \cdot \boldsymbol{b}) c_i,
\end{aligned}$$

so we have shown that

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}. \tag{4.13}$$

Notice one very important point in this analysis: in the second line, the k component of the vector $\boldsymbol{b} \times \boldsymbol{c}$ is required. When this is written down in suffix notation it is essential that 'new' suffices are used (l and m above) to avoid repeating the existing suffices i and j – recall the essential rule of suffix notation that no suffix may appear more than twice.

EXERCISES

4.1 Write the vector equation $\boldsymbol{a} \times \boldsymbol{b} + (\boldsymbol{a} \cdot \boldsymbol{d})\boldsymbol{c} = \boldsymbol{e}$ in suffix notation.

4.2 Translate the suffix notation equation $\delta_{ij} c_j + \epsilon_{kji} a_k b_j = d_l e_m c_i b_l c_m$ into ordinary vector notation.

4.3 Use suffix notation to show that $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$.

4.4 Simplify the suffix notation expressions
(a) $\delta_{ij}\epsilon_{ijk}$;
(b) $\epsilon_{ijk}\epsilon_{ilm}$;
(c) $\epsilon_{ijk}\epsilon_{ijm}$;
(d) $\epsilon_{ijk}\epsilon_{ijk}$.

4.5 Using suffix notation, find an alternative expression (involving no cross products) for $\boldsymbol{a} \times \boldsymbol{b} \cdot \boldsymbol{c} \times \boldsymbol{d}$.

4.6 If A and B are two $N \times N$ matrices, show that $(AB)^T = B^T A^T$, where A^T is the *transpose* of A defined by interchanging the rows and columns of A.

4.7 Verify the formulae (4.9) and (4.10) for the determinant of a 3×3 matrix.

4.8 Use the formula (4.10) for the determinant of a 3×3 matrix M to show that
(a) $6|M| = \epsilon_{pqr}\epsilon_{ijk} M_{pi} M_{qj} M_{rk}$;
(b) $|M^T| = |M|$;
(c) $|MN| = |M||N|$.

4.5 Grad, div and curl in suffix notation

The differential operators grad, div and curl can be written using suffix notation. To do this, the Cartesian coordinates (x, y, z) will be relabelled (x_1, x_2, x_3). As in the previous section, the use of suffix notation results in a much more compact formulation and simplifies many of the computations.

Consider first the gradient of a scalar field, ∇f. This is defined by

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right).$$

The i component of ∇f is equal to the partial derivative of f with respect to x_i, so in suffix notation this can be written

$$[\nabla f]_i = \frac{\partial f}{\partial x_i}. \tag{4.14}$$

Thus the vector differential operator ∇ defined in (3.12) can be written in suffix notation as

$$[\nabla]_i = \frac{\partial}{\partial x_i}. \tag{4.15}$$

The divergence of a vector field $\boldsymbol{u}$ is

$$\nabla \cdot \boldsymbol{u} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} = \frac{\partial u_j}{\partial x_j}, \tag{4.16}$$

where the summation convention implies the sum over j from 1 to 3. Note that the same expression results from taking the dot product of ∇ defined by (4.15) with the vector $\boldsymbol{u}$ using the suffix notation formula (4.1) for the dot product of two vectors.

The first component of $\nabla \times \boldsymbol{u}$ is, from (3.22),

$$[\nabla \times \boldsymbol{u}]_1 = \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3} = \epsilon_{1jk}\frac{\partial u_k}{\partial x_j},$$

where the repeated j and k imply a double sum, and so the suffix notation expression for $\nabla \times \boldsymbol{u}$ is

$$[\nabla \times \boldsymbol{u}]_i = \epsilon_{ijk}\frac{\partial u_k}{\partial x_j}. \tag{4.17}$$

This can also be obtained simply by taking the cross product of ∇ with $\boldsymbol{u}$ using (4.8).

Example 4.11

Let $\boldsymbol{r}$ be the position vector $\boldsymbol{r} = (x_1, x_2, x_3)$ and $r = |\boldsymbol{r}|$. Use suffix notation to evaluate $\partial x_i/\partial x_j$. Hence find $\boldsymbol{\nabla} r$, $\boldsymbol{\nabla} \cdot \boldsymbol{r}$ and $\boldsymbol{\nabla} \times \boldsymbol{r}$.

$\boldsymbol{r} = (x_1, x_2, x_3)$, so in suffix notation, $r_i = x_i$. The three coordinate axes x_1, x_2, x_3 are independent. Thus the derivative of each of the x_i with respect to one of the others is zero, while the derivative with respect to itself is 1. Thus

$$\begin{aligned} \frac{\partial x_i}{\partial x_j} &= \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j \end{cases} \\ &= \delta_{ij}. \end{aligned} \tag{4.18}$$

To find $\boldsymbol{\nabla} r$, first write $r = |\boldsymbol{r}| = (\boldsymbol{r} \cdot \boldsymbol{r})^{1/2} = (x_j x_j)^{1/2}$, so

$$\begin{aligned} [\boldsymbol{\nabla} r]_i &= \frac{\partial}{\partial x_i}(x_j x_j)^{1/2} \\ &= \frac{1}{2}(x_j x_j)^{-1/2}\frac{\partial}{\partial x_i}(x_j x_j) \\ &= \frac{1}{2r}\, 2x_j \frac{\partial x_j}{\partial x_i} \\ &= \frac{1}{r} x_j \delta_{ij} = \frac{x_i}{r}. \end{aligned}$$

So

$$\boldsymbol{\nabla} r = \boldsymbol{r}/r \tag{4.19}$$

as was shown in Exercise 3.3 without the use of suffix notation. Similarly, the divergence of $\boldsymbol{r}$ is

$$\boldsymbol{\nabla} \cdot \boldsymbol{r} = \frac{\partial x_j}{\partial x_j} = \delta_{jj} = 3$$

and the curl of $\boldsymbol{r}$ is

$$[\boldsymbol{\nabla} \times \boldsymbol{r}]_i = \epsilon_{ijk}\frac{\partial x_k}{\partial x_j} = \epsilon_{ijk}\delta_{jk} = 0,$$

using the result of Exercise 4.4(a).

4.6 Combinations of grad, div and curl

The operators grad, div and curl can be combined together in several different ways. Some of these combinations can be simplified or expanded into alternative expressions. These combinations are considered below, making use of suffix notation. All of these results can also be obtained without suffix notation, by writing out all the components, but in most cases the suffix notation method is much quicker.

- Div grad:

$$\nabla \cdot (\nabla f) = \frac{\partial}{\partial x_j}\left([\nabla f]_j\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial^2 f}{\partial x_j \partial x_j} = \nabla^2 f. \qquad (4.20)$$

  This is the Laplacian of f introduced in Section 3.3.2, the sum of the second partial derivatives of f.

- Curl grad: this combination was shown to be zero in Section 3.4.3. This result can be shown using suffix notation as follows:

$$\begin{aligned} [\nabla \times (\nabla f)]_i &= \epsilon_{ijk}\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_k} \\ &= \epsilon_{ikj}\frac{\partial}{\partial x_k}\frac{\partial f}{\partial x_j} \quad \text{(relabelling } j \leftrightarrow k\text{)} \\ &= -\epsilon_{ijk}\frac{\partial}{\partial x_k}\frac{\partial f}{\partial x_j} \quad \text{(using } \epsilon_{ikj} = -\epsilon_{ijk}\text{)} \\ &= -\epsilon_{ijk}\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_k} \quad \text{(as order of derivatives does not matter)} \\ &= 0, \end{aligned}$$

  since the expression has been manipulated to give minus itself.

- Grad div:

$$[\nabla(\nabla \cdot \boldsymbol{u})]_i = \frac{\partial}{\partial x_i}\left(\frac{\partial u_j}{\partial x_j}\right) = \frac{\partial^2 u_j}{\partial x_i \partial x_j}. \qquad (4.21)$$

  This quantity cannot be simplified further.

- Div curl:

$$\begin{aligned} \nabla \cdot (\nabla \times \boldsymbol{u}) &= \frac{\partial}{\partial x_i}\epsilon_{ijk}\frac{\partial u_k}{\partial x_j} \\ &= \epsilon_{jik}\frac{\partial}{\partial x_j}\frac{\partial u_k}{\partial x_i} \quad \text{(relabelling } i \leftrightarrow j\text{)} \\ &= -\epsilon_{ijk}\frac{\partial}{\partial x_j}\frac{\partial u_k}{\partial x_i} \end{aligned}$$

$$\begin{aligned} &= -\epsilon_{ijk}\frac{\partial}{\partial x_i}\frac{\partial u_k}{\partial x_j} \\ &= 0, \end{aligned} \tag{4.22}$$

using exactly the same argument as for curl grad.

- Curl curl:

$$\begin{aligned} [\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{u})]_i &= \epsilon_{ijk}\frac{\partial}{\partial x_j}\epsilon_{klm}\frac{\partial u_m}{\partial x_l} \\ &= \epsilon_{ijk}\epsilon_{klm}\frac{\partial^2 u_m}{\partial x_j \partial x_l} \\ &= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\frac{\partial^2 u_m}{\partial x_j \partial x_l} \\ &= \frac{\partial^2 u_j}{\partial x_j \partial x_i} - \frac{\partial^2 u_i}{\partial x_j \partial x_j} \\ &= \left[\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u}) - \nabla^2\boldsymbol{u}\right]_i . \end{aligned} \tag{4.23}$$

This result can be used to provide a physical definition for ∇^2 applied to a vector. The previous definition (3.19) was only defined in terms of the components of the vector in Cartesian coordinates. From the above result, $\nabla^2\boldsymbol{u}$ can be defined by

$$\nabla^2\boldsymbol{u} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{u}). \tag{4.24}$$

These five combinations of grad, div and curl are the only ones that make sense. For example, the combination grad curl has no meaning since curl is a vector but grad can only act on a scalar. Combinations of three or more of the operators grad, div and curl can be evaluated using the above results, as in the following example.

Example 4.12

Show that

$$\boldsymbol{\nabla}\times(\nabla^2\boldsymbol{u}) = \nabla^2(\boldsymbol{\nabla}\times\boldsymbol{u}). \tag{4.25}$$

Using the result (4.24),

$$\begin{aligned} \boldsymbol{\nabla}\times(\nabla^2\boldsymbol{u}) &= \boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{u}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{u})\right) \\ &= -\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{u})\right) \quad (\text{since } \boldsymbol{\nabla}\times\boldsymbol{\nabla} = 0) \\ &= -\boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{u})\right) + \nabla^2(\boldsymbol{\nabla}\times\boldsymbol{u}) \quad (\text{using (4.23)}) \\ &= \nabla^2(\boldsymbol{\nabla}\times\boldsymbol{u}) \quad (\text{since } \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\times\boldsymbol{u} = 0). \end{aligned}$$

So the operators ∇^2 and $\boldsymbol{\nabla}\times$ commute.

4.7 Grad, div and curl applied to products of functions

Another useful application of suffix notation is in computing the action of grad, div and curl on products of vector and scalar fields. As in the previous section, these results can also be obtained by writing out all the components, but the suffix notation method is much more compact and elegant. Some of these results are straightforward applications of the usual rule for the differentiation of a product and can simply be written down without any calculation, but many of them are not so obvious.

In the following, f and g are differentiable scalar fields and $\boldsymbol{u}$ and $\boldsymbol{v}$ are differentiable vector fields.

$$[\boldsymbol{\nabla}(fg)]_i = \frac{\partial}{\partial x_i}(fg) = f\frac{\partial g}{\partial x_i} + g\frac{\partial f}{\partial x_i} = [f\boldsymbol{\nabla} g + g\boldsymbol{\nabla} f]_i, \quad \text{so}$$

$$\boldsymbol{\nabla}(fg) = f\boldsymbol{\nabla} g + g\boldsymbol{\nabla} f. \tag{4.26}$$

$$\begin{aligned} \boldsymbol{\nabla}\cdot(f\boldsymbol{u}) &= \frac{\partial}{\partial x_i}(fu_i) \\ &= \frac{\partial f}{\partial x_i}u_i + f\frac{\partial u_i}{\partial x_i} \\ &= \boldsymbol{\nabla} f\cdot\boldsymbol{u} + f\boldsymbol{\nabla}\cdot\boldsymbol{u}. \end{aligned} \tag{4.27}$$

$$\begin{aligned} [\boldsymbol{\nabla}\times(f\boldsymbol{u})]_i &= \epsilon_{ijk}\frac{\partial}{\partial x_j}(fu_k) \\ &= \epsilon_{ijk}\frac{\partial f}{\partial x_j}u_k + f\epsilon_{ijk}\frac{\partial u_k}{\partial x_j} \\ &= [\boldsymbol{\nabla} f\times\boldsymbol{u} + f\boldsymbol{\nabla}\times\boldsymbol{u}]_i. \end{aligned} \tag{4.28}$$

$$\begin{aligned} \boldsymbol{\nabla}\cdot(\boldsymbol{u}\times\boldsymbol{v}) &= \frac{\partial}{\partial x_i}(\epsilon_{ijk}u_j v_k) \\ &= \epsilon_{ijk}\frac{\partial u_j}{\partial x_i}v_k + \epsilon_{ijk}u_j\frac{\partial v_k}{\partial x_i} \\ &= \left(\epsilon_{kij}\frac{\partial u_j}{\partial x_i}\right)v_k - \left(\epsilon_{jik}\frac{\partial v_k}{\partial x_i}\right)u_j \\ &= (\boldsymbol{\nabla}\times\boldsymbol{u})\cdot\boldsymbol{v} - (\boldsymbol{\nabla}\times\boldsymbol{v})\cdot\boldsymbol{u}. \end{aligned} \tag{4.29}$$

$$\begin{aligned}
[\boldsymbol{\nabla} \times (\boldsymbol{u} \times \boldsymbol{v})]_i &= \epsilon_{ijk} \frac{\partial}{\partial x_j}(\epsilon_{klm} u_l v_m) \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) \frac{\partial}{\partial x_j}(u_l v_m) \\
&= \frac{\partial}{\partial x_j}(u_i v_j) - \frac{\partial}{\partial x_j}(u_j v_i) \\
&= u_i \frac{\partial v_j}{\partial x_j} + v_j \frac{\partial u_i}{\partial x_j} - u_j \frac{\partial v_i}{\partial x_j} - v_i \frac{\partial u_j}{\partial x_j} \\
&= [\boldsymbol{u}(\boldsymbol{\nabla} \cdot \boldsymbol{v}) + \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} - \boldsymbol{v}(\boldsymbol{\nabla} \cdot \boldsymbol{u})]_i, \qquad (4.30)
\end{aligned}$$

where the operator $\boldsymbol{u} \cdot \boldsymbol{\nabla}$ is defined by

$$\boldsymbol{u} \cdot \boldsymbol{\nabla} = u_j \frac{\partial}{\partial x_j} \qquad (4.31)$$

and can act on either a scalar or a vector.

To find an expansion for the expression $\boldsymbol{\nabla}(\boldsymbol{u} \cdot \boldsymbol{v})$, consider first the quantity

$$\begin{aligned}
[\boldsymbol{u} \times (\boldsymbol{\nabla} \times \boldsymbol{v})]_i &= \epsilon_{ijk} u_j \epsilon_{klm} \frac{\partial v_m}{\partial x_l} \\
&= (\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}) u_j \frac{\partial v_m}{\partial x_l} \\
&= u_j \frac{\partial v_j}{\partial x_i} - u_j \frac{\partial v_i}{\partial x_j}.
\end{aligned}$$

Similarly, interchanging $\boldsymbol{u}$ and $\boldsymbol{v}$,

$$[\boldsymbol{v} \times (\boldsymbol{\nabla} \times \boldsymbol{u})]_i = v_j \frac{\partial u_j}{\partial x_i} - v_j \frac{\partial u_i}{\partial x_j}.$$

Adding these two equations gives

$$\begin{aligned}
[\boldsymbol{u} \times (\boldsymbol{\nabla} \times \boldsymbol{v}) + \boldsymbol{v} \times (\boldsymbol{\nabla} \times \boldsymbol{u})]_i &= u_j \frac{\partial v_j}{\partial x_i} - u_j \frac{\partial v_i}{\partial x_j} + v_j \frac{\partial u_j}{\partial x_i} - v_j \frac{\partial u_i}{\partial x_j} \\
&= [\boldsymbol{\nabla}(\boldsymbol{u} \cdot \boldsymbol{v}) - \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u}]_i. \qquad (4.32)
\end{aligned}$$

This can be rearranged to give

$$\boldsymbol{\nabla}(\boldsymbol{u} \cdot \boldsymbol{v}) = \boldsymbol{u} \times (\boldsymbol{\nabla} \times \boldsymbol{v}) + \boldsymbol{v} \times (\boldsymbol{\nabla} \times \boldsymbol{u}) + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \boldsymbol{u}. \qquad (4.33)$$

The effect of applying grad, div or curl to products of more than two scalar or vector functions can be obtained either by the repeated use of the above results, or directly by suffix notation, as in the following example.

Example 4.13

Find an expansion for $\nabla \cdot (fg\boldsymbol{u})$.

$$\begin{aligned}\nabla \cdot (fg\boldsymbol{u}) &= \nabla(fg) \cdot \boldsymbol{u} + (fg)\nabla \cdot \boldsymbol{u} \quad \text{(using (4.27))}\\ &= (f\nabla g + g\nabla f) \cdot \boldsymbol{u} + (fg)\nabla \cdot \boldsymbol{u}\\ &= f\nabla g \cdot \boldsymbol{u} + g\nabla f \cdot \boldsymbol{u} + fg\nabla \cdot \boldsymbol{u}.\end{aligned}$$

Alternatively, using suffix notation,

$$\begin{aligned}\nabla \cdot (fg\boldsymbol{u}) &= \frac{\partial}{\partial x_i}(fgu_i)\\ &= fg\frac{\partial u_i}{\partial x_i} + f\frac{\partial g}{\partial x_i}u_i + \frac{\partial f}{\partial x_i}gu_i\\ &= fg\nabla \cdot \boldsymbol{u} + f\nabla g \cdot \boldsymbol{u} + g\nabla f \cdot \boldsymbol{u}.\end{aligned}$$

Example 4.14

Show that $\boldsymbol{u} \cdot \nabla \boldsymbol{u} = \nabla(|\boldsymbol{u}|^2/2) - \boldsymbol{u} \times (\nabla \times \boldsymbol{u})$.

Apply (4.33) with $\boldsymbol{v} = \boldsymbol{u}$:

$$\nabla(\boldsymbol{u} \cdot \boldsymbol{u}) = 2\boldsymbol{u} \times (\nabla \times \boldsymbol{u}) + 2\boldsymbol{u} \cdot \nabla \boldsymbol{u}.$$

Rearranging this and dividing by 2 gives

$$\boldsymbol{u} \cdot \nabla \boldsymbol{u} = \nabla(|\boldsymbol{u}|^2/2) - \boldsymbol{u} \times (\nabla \times \boldsymbol{u}). \tag{4.34}$$

Example 4.15

Use the results (4.30) and (4.33) to provide a definition of $\boldsymbol{u} \cdot \nabla \boldsymbol{v}$ that is not given in terms of Cartesian components.

By subtracting (4.30) from (4.33), $\boldsymbol{v} \cdot \nabla \boldsymbol{u}$ is eliminated and we obtain

$$\nabla(\boldsymbol{u}\cdot\boldsymbol{v}) - \nabla \times (\boldsymbol{u}\times\boldsymbol{v}) = \boldsymbol{u}\times(\nabla\times\boldsymbol{v}) + \boldsymbol{v}\times(\nabla\times\boldsymbol{u}) + 2\boldsymbol{u}\cdot\nabla\boldsymbol{v} - \boldsymbol{u}(\nabla\cdot\boldsymbol{v}) + \boldsymbol{v}(\nabla\cdot\boldsymbol{u})$$

which can be rearranged to give

$$\begin{aligned}\boldsymbol{u} \cdot \nabla \boldsymbol{v} = \frac{1}{2}\big(&\nabla(\boldsymbol{u} \cdot \boldsymbol{v}) - \nabla \times (\boldsymbol{u} \times \boldsymbol{v}) - \boldsymbol{u} \times (\nabla \times \boldsymbol{v}) - \boldsymbol{v} \times (\nabla \times \boldsymbol{u})\\ &+\boldsymbol{u}(\nabla \cdot \boldsymbol{v}) - \boldsymbol{v}(\nabla \cdot \boldsymbol{u})\big).\end{aligned} \tag{4.35}$$

Summary of Chapter 4

Suffix notation

Suffix notation is a powerful tool for manipulating expressions involving vectors. The rules of suffix notation are as follows:

- Within any term in an equation, any suffix must appear either once or twice. No suffix may appear more that twice.
- A suffix that appears once in any term is called a 'free' suffix. A free suffix takes the values 1, 2 and 3 and represents the components of a vector. For example $\boldsymbol{a} + \boldsymbol{b} = \boldsymbol{c} - \boldsymbol{d}$ is written in suffix notation as $a_i + b_i = c_i - d_i$.
- In a vector equation, the free suffix must be the same in each term. The above equation may also be written $a_j + b_j = c_j - d_j$, or any other suffix may be used, provided the same suffix appears in each term.
- A suffix that appears twice in a term is called a 'dummy' suffix and is summed from 1 to 3. This is known as the *summation convention.* So $a_j b_j$ means $a_1 b_1 + a_2 b_2 + a_3 b_3 = \boldsymbol{a} \cdot \boldsymbol{b}$.
- A pair of dummy suffices can be changed. For example, $a_j b_j$, $a_k b_k$ and $a_m b_m$ are all equal to $\boldsymbol{a} \cdot \boldsymbol{b}$.
- The order of terms in a suffix notation expression does not matter.
- The *Kronecker delta* is defined by $\delta_{ij} = 1$ if $i = j$, 0 if $i \neq j$. Properties include $\delta_{ij} = \delta_{ji}$, $\delta_{ij} a_j = a_i$ and $\delta_{ij} a_i b_j = \boldsymbol{a} \cdot \boldsymbol{b}$.
- The *alternating tensor* ϵ_{ijk} is defined by $\epsilon_{ijk} = 0$ if any of i, j, k are equal, $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$, $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$. Properties include $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$, $\epsilon_{ijk} = -\epsilon_{jik}$.
- The cross product of $\boldsymbol{a}$ and $\boldsymbol{b}$ can be written $[\boldsymbol{a} \times \boldsymbol{b}]_i = \epsilon_{ijk} a_j b_k$.
- ϵ_{ijk} and δ_{ij} are related by $\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$.
- Grad, div and curl can be written in suffix notation as follows:

$$[\boldsymbol{\nabla} f]_i = \frac{\partial f}{\partial x_i}, \qquad \boldsymbol{\nabla} \cdot \boldsymbol{u} = \frac{\partial u_j}{\partial x_j}, \qquad [\boldsymbol{\nabla} \times \boldsymbol{u}]_i = \epsilon_{ijk} \frac{\partial u_k}{\partial x_j}.$$

Combinations of operators and derivatives of products

- $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} f) = \nabla^2 f$.
- $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} f) = \boldsymbol{0}$.
- $\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{u}) = 0$.
- $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{u}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{u}) - \nabla^2 \boldsymbol{u}$.

- $\nabla(fg) = f\nabla g + g\nabla f$.
- $\nabla \cdot (f\boldsymbol{u}) = \nabla f \cdot \boldsymbol{u} + f\nabla \cdot \boldsymbol{u}$.
- $\nabla \times (f\boldsymbol{u}) = \nabla f \times \boldsymbol{u} + f\nabla \times \boldsymbol{u}$.
- $\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}) = (\nabla \times \boldsymbol{u}) \cdot \boldsymbol{v} - (\nabla \times \boldsymbol{v}) \cdot \boldsymbol{u}$.
- $\nabla \times (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{u}(\nabla \cdot \boldsymbol{v}) + \boldsymbol{v} \cdot \nabla\boldsymbol{u} - \boldsymbol{u} \cdot \nabla\boldsymbol{v} - \boldsymbol{v}(\nabla \cdot \boldsymbol{u})$.
- $\nabla(\boldsymbol{u} \cdot \boldsymbol{v}) = \boldsymbol{u} \times (\nabla \times \boldsymbol{v}) + \boldsymbol{v} \times (\nabla \times \boldsymbol{u}) + \boldsymbol{u} \cdot \nabla\boldsymbol{v} + \boldsymbol{v} \cdot \nabla\boldsymbol{u}$.

EXERCISES

4.9 Write in suffix notation the vector equation $\boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{c} = (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{b} - \boldsymbol{d}$.

4.10 Simplify the suffix notation expressions

(a) $\delta_{ij}\delta_{jk}\delta_{ki}$;

(b) $\epsilon_{ijk}\epsilon_{klm}\epsilon_{mni}$.

4.11 Simplify the suffix notation expression $\delta_{ij}a_j b_l c_k \delta_{li}$ and write the result in vector form.

4.12 (a) Show that $\nabla \times (f\nabla f) = \boldsymbol{0}$.

(b) Evaluate $\nabla \cdot (f\nabla f)$.

4.13 Show that the vector $\boldsymbol{u} = \nabla f \times \nabla g$ is solenoidal.

4.14 Verify the formula (4.34) for $\boldsymbol{u} \cdot \nabla\boldsymbol{u}$ by using (4.35).

4.15 Show that $\nabla \cdot \nabla^2\boldsymbol{u} = \nabla^2\nabla \cdot \boldsymbol{u}$,

(a) using suffix notation;

(b) using (4.24).

4.16 The vector fields $\boldsymbol{u}$ and $\boldsymbol{w}$ and the scalar field ϕ are related by the equation

$$\boldsymbol{u} + \nabla \times \boldsymbol{w} = \nabla\phi + \nabla^2\boldsymbol{u},$$

and $\boldsymbol{u}$ is solenoidal. Show that ϕ obeys Laplace's equation.

4.17 Show that $\nabla f(r) = f'(r)\boldsymbol{r}/r$, where $\boldsymbol{r}$ is the position vector $\boldsymbol{r} = (x_1, x_2, x_3)$ and $r = |\boldsymbol{r}|$.

4.18 The vector field $\boldsymbol{u}$ is defined by $\boldsymbol{u} = h(r)\boldsymbol{r}$, where $h(r)$ is an arbitrary differentiable function.

(a) Show that $\nabla \times \boldsymbol{u} = \boldsymbol{0}$.

(b) If $\nabla \cdot \boldsymbol{u} = 0$, find the differential equation satisfied by h.

(c) Solve this differential equation.

4.19 A vector field $\boldsymbol{u}$ with the property that $\boldsymbol{u} = c\nabla \times \boldsymbol{u}$, where c is a constant, is called a *Beltrami field.*

(a) Show that a Beltrami field is solenoidal.

(b) Show that the curl of a Beltrami field is a Beltrami field.

(c) A Beltrami field has the form $\boldsymbol{u} = (\sin y, f, g)$. Find the functions f and g and the possible values of c if it is given that g does not depend on x.

5
Integral Theorems

This chapter describes two important theorems that link the material in Chapter 2 on line, surface and volume integrals with the definitions of the divergence and curl from Chapter 3. These theorems have great physical significance and are widely used in deriving mathematical equations representing physical laws.

5.1 Divergence theorem

Let $\boldsymbol{u}$ be a continuously differentiable vector field, defined in a volume V. Let S be the closed surface forming the boundary of V and let $\boldsymbol{n}$ be the unit outward normal to S. Then the divergence theorem states that

$$\iiint_V \nabla \cdot \boldsymbol{u} \, dV = \oiint_S \boldsymbol{u} \cdot \boldsymbol{n} \, dS. \tag{5.1}$$

Proof

The volume V is divided into a large number of small subvolumes δV_i with surfaces δS_i, as shown in Figure 5.1. The proof of the divergence theorem then follows naturally from the physical definition of the divergence in terms of a surface integral (3.14). Within each of the subvolumes, $\nabla \cdot \boldsymbol{u}$ is defined by

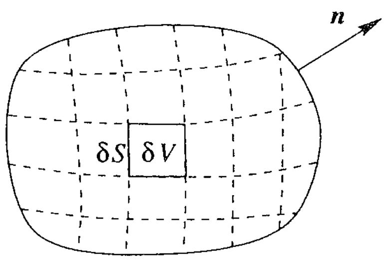


Fig. 5.1. Division of a volume V into small subvolumes δV for the proof of the divergence theorem.

$$\nabla \cdot \boldsymbol{u} \approx \frac{1}{\delta V_i} \oiint_{\delta S_i} \boldsymbol{u} \cdot \boldsymbol{n} \, dS, \tag{5.2}$$

where the approximation becomes exact in the limit $\delta V_i \to 0$. Now multiply both sides of (5.2) by δV_i and add the contributions from all the subvolumes:

$$\sum_i \nabla \cdot \boldsymbol{u} \, \delta V_i \approx \sum_i \oiint_{\delta S_i} \boldsymbol{u} \cdot \boldsymbol{n} \, dS. \tag{5.3}$$

Now take the limit $\delta V_i \to 0$. The l.h.s. becomes the volume integral of $\nabla \cdot \boldsymbol{u}$ over the volume V; this is just the definition of the volume integral. To simplify the r.h.s. consider two adjacent volume elements δV_1 and δV_2 (Figure 5.2). Since

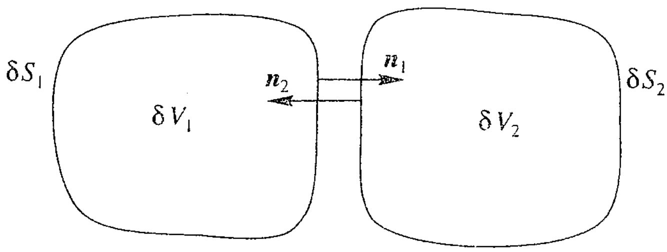


Fig. 5.2. Enlargement of two adjacent volume elements.

the normal vector to each surface points outward, the normal vectors to the two surfaces along their common surface point in opposite directions: $\boldsymbol{n}_1 = -\boldsymbol{n}_2$. Therefore the values of $\boldsymbol{u} \cdot \boldsymbol{n}$ cancel along the common surface: $\boldsymbol{u} \cdot \boldsymbol{n}_1 + \boldsymbol{u} \cdot \boldsymbol{n}_2 = 0$. This means that all the contributions to the sum on the r.h.s. of (5.3) from the interior of the region V cancel out, leaving only the surface integral over the exterior surface S. So in the limit $\delta V_i \to 0$, (5.3) becomes

$$\iiint_V \nabla \cdot \boldsymbol{u}\, dV = \oiint_S \boldsymbol{u} \cdot \boldsymbol{n}\, dS.$$

□

The divergence theorem is sometimes referred to as Gauss's theorem. It has many important applications in physics, and it is important to develop a physical intuition for the meaning of the theorem. Roughly speaking, the divergence theorem states that the total amount of expansion of $\boldsymbol{u}$ within the volume V is equal to the flux of $\boldsymbol{u}$ out of the surface S. This is essentially a conservation law, and the mathematical form of many physical conservation laws is derived from the divergence theorem. An example is given in the following section.

5.1.1 Conservation of mass for a fluid

As an example of the application of the divergence theorem, this section presents the derivation of the law of conservation of mass for a fluid of variable density.

Consider a fluid with density $\rho(\boldsymbol{r}, t)$ flowing with velocity $\boldsymbol{u}(\boldsymbol{r}, t)$. Let V be an arbitrary volume fixed in space, with surface S and outward normal $\boldsymbol{n}$ (Figure 5.3). Then the total mass of the fluid contained in V is the volume integral of ρ:

$$\text{Mass of fluid in } V = \iiint_V \rho\, dv. \tag{5.4}$$

Now the rate at which mass enters V is equal to the surface integral of the flux $\rho\boldsymbol{u}$:

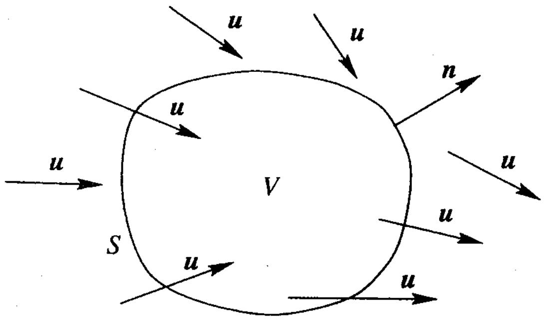


Fig. 5.3. Fluid flows with velocity $\boldsymbol{u}$ through a region V.

$$\text{Rate of mass flow into } V = -\oiint_S \rho \boldsymbol{u} \cdot \boldsymbol{n}\, dS, \tag{5.5}$$

where the minus sign appears because $\boldsymbol{n}$ points outward, so mass enters V if $\boldsymbol{u} \cdot \boldsymbol{n} < 0$.

We can now apply the physical law that mass is conserved: the rate of change of the mass in V must equal the rate at which mass enters V. Mathematically this is

$$\frac{d}{dt} \iiint_V \rho \, dV = -\oiint_S \rho \boldsymbol{u} \cdot \boldsymbol{n}\, dS. \tag{5.6}$$

The surface integral on the r.h.s. can now be written as a volume integral using the divergence theorem. Also, the order of the derivative and the integral on the l.h.s. can be interchanged:

$$\iiint_V \frac{\partial \rho}{\partial t}\, dV = -\iiint_V \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u})\, dV, \tag{5.7}$$

where the time derivative has become a partial derivative since ρ is a function of space and time. These two integrals can now be combined into one:

$$\iiint_V \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u})\; dV = 0. \tag{5.8}$$

Now this result has been obtained without any restrictions on the volume V. Thus it is true for any arbitrary volume V. The only way that this can be true is if the integrand (the quantity inside the integral) is zero everywhere. If there were some point where the integrand were non-zero, a small volume could be drawn around that point, which would contradict (5.8).

Therefore the law for conservation of mass of a fluid is

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0. \tag{5.9}$$

This conservation law takes the following form: the rate of change of the density plus the divergence of the flux is zero. Many other conservation laws can also be written in this form, for example conservation of energy or conservation of electric charge.

By expanding the divergence of $\rho \boldsymbol{u}$, (5.9) can be written in the form

$$\frac{\partial \rho}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho + \rho \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0. \tag{5.10}$$

If the density of the fluid is constant and uniform, i.e. independent of time and space, then this equation simplifies to

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0. \tag{5.11}$$

A fluid obeying (5.11) is said to be *incompressible*.

5.1.2 Applications of the divergence theorem

The divergence theorem has many important applications, in addition to the derivation of the mathematical form of conservation laws shown in the previous section. It can be used to simplify the evaluation of integrals, by converting a complicated volume integral into a simpler surface integral or vice versa. It can also be used to prove some important results, such as the uniqueness of the solution to Laplace's equation

$$\nabla^2\phi = 0. \tag{5.12}$$

Some of these applications are illustrated in the following examples.

Example 5.1

Show that for any closed surface S,

$$\oint\!\!\oint_S (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{n}\, dS = 0.$$

Using the divergence theorem, the surface integral can be converted into a volume integral:

$$\oint\!\!\oint_S (\boldsymbol{\nabla} \times \boldsymbol{u}) \cdot \boldsymbol{n}\, dS = \iiint_V \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{u})\, dV.$$

Since the combination div curl is always zero, this integral is zero.

Example 5.2

Find the relationship between the surface integral

$$\oint\!\!\oint_S \boldsymbol{r} \cdot \boldsymbol{n}\, dS$$

and the volume V contained within the closed surface S.

Applying the divergence theorem,

$$\oint\!\!\oint_S \boldsymbol{r} \cdot \boldsymbol{n}\, dS = \iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{r}\, dV = \iiint_V 3\, dV = 3V,$$

using the result (3.17) that $\boldsymbol{\nabla} \cdot \boldsymbol{r} = 3$. Thus the surface integral is three times the volume V.

Example 5.3

The scalar field ϕ obeys Laplace's equation (5.12) in the region V and obeys $\phi = 0$ on the surface S that encloses V. Show that the only possible solution for ϕ is $\phi = 0$ everywhere within V.

$$\nabla^2\phi = 0 \quad \Rightarrow \quad \phi\nabla^2\phi = 0 \quad \Rightarrow \quad \nabla\cdot(\phi\nabla\phi) - \nabla\phi\cdot\nabla\phi = 0,$$

using (4.27). Now integrate over the volume V and use the divergence theorem to convert the first term to a surface integral:

$$\oiint_S \phi\nabla\phi\cdot\boldsymbol{n}\,dS - \iiint_V |\nabla\phi|^2\,dV = 0.$$

Since $\phi = 0$ on S, the surface integral vanishes. The quantity $|\nabla\phi|^2$ is always greater than or equal to zero, so its integral can only be zero if $\nabla\phi = \mathbf{0}$ everywhere. This means that ϕ must be a constant, and since $\phi = 0$ on S, this constant must be zero, so $\phi = 0$ everywhere within V.

Example 5.4

The scalar field ϕ obeys Laplace's equation (5.12) in the region V and the value of ϕ is given on the surface S that encloses V. Show that the solution to Laplace's equation is unique.

To prove uniqueness, suppose that there are two different solutions, ϕ_1 and ϕ_2, obeying $\nabla^2\phi_1 = 0$ and $\nabla^2\phi_2 = 0$ in V. Since the value of ϕ is specified on S, $\phi_1 = \phi_2$ on S. Now consider the function $\psi = \phi_1 - \phi_2$. This function also obeys Laplace's equation, since $\nabla^2(\phi_1 - \phi_2) = \nabla^2\phi_1 - \nabla^2\phi_2 = 0$. Moreover, $\psi = 0$ on S since $\phi_1 = \phi_2$ on S. Now we can apply the result of Example 5.3 to ψ: the only solution to $\nabla^2\psi = 0$ in V, $\psi = 0$ on S is $\psi = 0$ everywhere. Therefore $\phi_1 = \phi_2$ everywhere, so the solution is unique.

5.1.3 Related theorems linking surface and volume integrals

There are several other relationships between surface and volume integrals that can be derived from the divergence theorem by making different choices for the vector $\boldsymbol{u}$.

- Choose $\boldsymbol{u} = \boldsymbol{a}f$, where $\boldsymbol{a}$ is a constant vector and f is a scalar field. Then $\nabla\cdot\boldsymbol{u} = f\,\nabla\cdot\boldsymbol{a} + \boldsymbol{a}\cdot\nabla f = \boldsymbol{a}\cdot\nabla f$ since $\boldsymbol{a}$ is constant. Applying the divergence theorem gives

  $$\iiint_V \boldsymbol{a}\cdot\nabla f\,dV = \oiint_S \boldsymbol{a}f\cdot\boldsymbol{n}\,dS.$$

  Since $\boldsymbol{a}$ is constant, it can be taken out of the integrals:

  $$\boldsymbol{a}\cdot\left(\iiint_V \nabla f\,dV - \oiint_S f\boldsymbol{n}\,dS\right) = 0.$$

  Now since $\boldsymbol{a}$ is an arbitrary constant vector, this holds for any $\boldsymbol{a}$. This can only be true if the vector quantity within the large brackets is zero (for example, choosing $\boldsymbol{a} = \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ in turn shows that each of the components of the vector in large brackets is zero). The resulting integral theorem is then

$$\iiint_V \boldsymbol{\nabla} f \, dV = \oiint_S f\boldsymbol{n} \, dS. \tag{5.13}$$

- Choose $\boldsymbol{u} = \boldsymbol{a} \times \boldsymbol{v}$, where $\boldsymbol{a}$ is a constant vector and $\boldsymbol{v}$ is a vector field. Then $\boldsymbol{\nabla} \cdot \boldsymbol{u} = (\boldsymbol{\nabla} \times \boldsymbol{a}) \cdot \boldsymbol{v} - (\boldsymbol{\nabla} \times \boldsymbol{v}) \cdot \boldsymbol{a} = -(\boldsymbol{\nabla} \times \boldsymbol{v}) \cdot \boldsymbol{a}$. The divergence theorem gives

$$\iiint_V -(\boldsymbol{\nabla} \times \boldsymbol{v}) \cdot \boldsymbol{a} \, dV = \oiint_S \boldsymbol{a} \times \boldsymbol{v} \cdot \boldsymbol{n} \, dS = \oiint_S \boldsymbol{a} \cdot \boldsymbol{v} \times \boldsymbol{n} \, dS,$$

using the rule that the dot and cross may be interchanged in a scalar triple product. As in the previous example, the dot product with $\boldsymbol{a}$ can be taken out of the integral sign and then cancelled, giving

$$\iiint_V -\boldsymbol{\nabla} \times \boldsymbol{v} \, dV = \oiint_S \boldsymbol{v} \times \boldsymbol{n} \, dS. \tag{5.14}$$

- Choose $\boldsymbol{u} = f\boldsymbol{\nabla} g$, where f and g are two scalar fields. Then $\boldsymbol{\nabla} \cdot \boldsymbol{u} = \boldsymbol{\nabla} f \cdot \boldsymbol{\nabla} g + f\nabla^2 g$, and the divergence theorem gives

$$\iiint_V \boldsymbol{\nabla} f \cdot \boldsymbol{\nabla} g + f\nabla^2 g \, dV = \oiint_S f\boldsymbol{\nabla} g \cdot \boldsymbol{n} \, dS. \tag{5.15}$$

This result is known as Green's First Identity.

- Choose $\boldsymbol{u} = f\boldsymbol{\nabla} g - g\boldsymbol{\nabla} f$. By interchanging f and g in (5.15) and subtracting, we obtain

$$\iiint_V f\nabla^2 g - g\nabla^2 f \, dV = \oiint_S (f\boldsymbol{\nabla} g - g\boldsymbol{\nabla} f) \cdot \boldsymbol{n} \, dS, \tag{5.16}$$

which is known as Green's Second Identity.

Historical note

George Green (1793–1841) was a Nottingham miller who spent less than two years at school and learnt his mathematics by studying library books. In 1828 he published privately his first and greatest work, 'An essay on the application of mathematical analysis to the theories of electricity and magnetism', which includes the two theorems above. As with many geniuses his work was not appreciated until several years after his death. Green's mill in Nottingham has now been restored and is open to the public along with a Science Centre illustrating some of the applications of his work.

EXERCISES

5.1 Use the divergence theorem to evaluate the surface integral

$$\oiint_S \boldsymbol{u} \cdot \boldsymbol{n} \, dS$$

where $\boldsymbol{u} = (x \sin y, \cos^2 x, y^2 - z \sin y)$ and S is the surface of the sphere $x^2 + y^2 + (z-2)^2 = 1$.

5.2 Verify the divergence theorem, by calculating both the volume integral and the surface integral, for the vector field $\boldsymbol{u} = (y, x, z - x)$ and the volume V given by the unit cube $0 \leq x, y, z, \leq 1$.

5.3 An incompressible fluid is contained within a volume V with surface S and $\boldsymbol{u} \cdot \boldsymbol{n} = 0$ on S. Using the divergence theorem, show that

$$\iiint_V \boldsymbol{u} \cdot \boldsymbol{\nabla} \phi \, dV = 0$$

for any differentiable scalar field ϕ.

5.4 Two scalar fields f and g are related by Poisson's equation, $\nabla^2 f = g$. Show that

$$\iiint g \, dV = \oiint \boldsymbol{\nabla} f \cdot \boldsymbol{n} \, dS.$$

5.5 Use the divergence theorem to evaluate the surface integral

$$\iint_S \boldsymbol{v} \cdot \boldsymbol{n} \, dS$$

where $\boldsymbol{v} = (x + y, z^2, x^2)$ and S is the surface of the hemisphere $x^2 + y^2 + z^2 = 1$ with $z > 0$ and $\boldsymbol{n}$ is the upward-pointing normal. Note that the surface S is not closed.

5.6 Following the argument of Section 5.1.1, obtain the equation for conservation of electric charge relating the charge density q and the electric current density $\boldsymbol{j}$.

5.7 Use (5.13) to obtain a definition for $\boldsymbol{\nabla} f$ as the limit of an integral, similar to the definitions of div and curl.

5.2 Stokes's theorem

Stokes's theorem gives an alternative expression for the surface integral of the curl of a vector field. This is analogous to the divergence theorem, so Stokes's theorem could be referred to as the 'curl theorem'. The proof of the theorem is very similar to that for the divergence theorem, being based on the definition of curl in terms of a line integral.

Let C be a closed curve which forms the boundary of a surface S. Then for a continuously differentiable vector field $\boldsymbol{u}$, Stokes's theorem states that

$$\iint_S \nabla \times \boldsymbol{u} \cdot \boldsymbol{n}\, dS = \oint_C \boldsymbol{u} \cdot \boldsymbol{dr}, \tag{5.17}$$

where the direction of the line integral around C and the normal $\boldsymbol{n}$ are oriented in a right-handed sense (Figure 5.4).

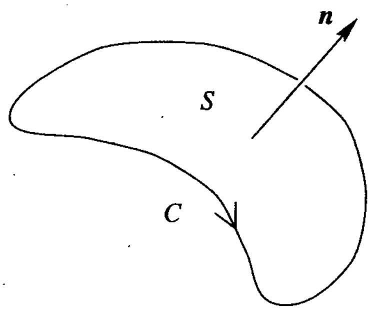


Fig. 5.4. Orientation of the curve C and the surface S for Stokes's theorem.

Proof

To demonstrate the theorem we first divide the surface S into small pieces each with area δS_i and bounding curves δC_i (Figure 5.5). Within each piece of the surface, the definition (3.20) of $\nabla \times \boldsymbol{u}$ is

$$\nabla \times \boldsymbol{u} \cdot \boldsymbol{n} \approx \frac{1}{\delta S_i} \oint_{\delta C_i} \boldsymbol{u} \cdot \boldsymbol{dr},$$

where the approximation is exact in the limit $\delta S_i \to 0$. Multiplying by δS_i and adding the contributions from all the surface elements,

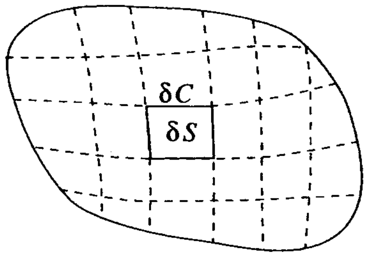


Fig. 5.5. Division of the surface S into small elements δS for the proof of Stokes's theorem.

$$\sum_i \nabla \times \boldsymbol{u} \cdot \boldsymbol{n}\, \delta S_i \approx \sum_i \oint_{\delta C_i} \boldsymbol{u} \cdot d\boldsymbol{r}.$$

Now consider the limit $\delta S_i \to 0$. The l.h.s. gives the surface integral of $\nabla \times \boldsymbol{u} \cdot \boldsymbol{n}$ over the surface S. On the r.h.s. the contributions to the line integrals from neighbouring elements cancel out, because the line elements $d\boldsymbol{r}$ point in opposite directions (Figure 5.6). Therefore only the curves that form part of C

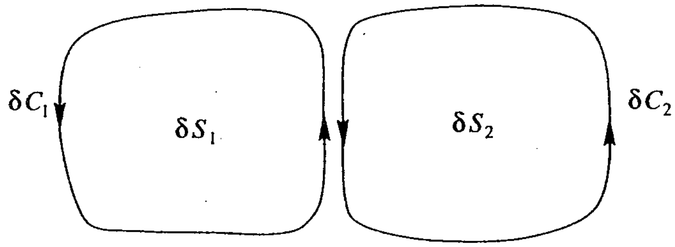


Fig. 5.6. Line integrals along adjoining elements cancel out.

contribute to the sum, so the sum simplifies to the line integral around C:

$$\iint_S \nabla \times \boldsymbol{u} \cdot \boldsymbol{n}\; dS = \oint_C \boldsymbol{u} \cdot d\boldsymbol{r}.$$

□

5.2.1 Applications of Stokes's theorem

Stokes's theorem can be useful for evaluating integrals, by converting line integrals to surface integrals or vice versa. It can also be used to prove other theorems, as in Example 5.5 below, or to formulate physical laws (Example 5.6).

Example 5.5

Show that any irrotational vector field is conservative.

Suppose that $\boldsymbol{u}$ is irrotational, so $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$. Then for any closed curve C,

$$\oint_C \boldsymbol{u} \cdot \boldsymbol{dr} = \iint_S \boldsymbol{\nabla} \times \boldsymbol{u} \cdot \boldsymbol{n}\, dS = 0,$$

where S is any surface spanning C. Thus $\boldsymbol{u}$ is a conservative vector field. Note that this result completes the demonstration of the statement in Section 3.4.3 of the equivalence of the three properties (i) $\boldsymbol{u} = \boldsymbol{\nabla}\phi$, (ii) $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{0}$, (iii) $\boldsymbol{u}$ is conservative.

Example 5.6

Ampère's law states that the total flux of electric current flowing through a loop is proportional to the line integral of the magnetic field around the loop. Use Stokes's theorem to obtain an alternative form of this law that does not involve any integrals.

Let $\boldsymbol{B}$ be the magnetic field strength and $\boldsymbol{j}$ be the current density. The constant of proportionality is μ_0 in SI units. Then Ampère's law states that

$$\oint_C \boldsymbol{B} \cdot \boldsymbol{dr} = \mu_0 \iint_S \boldsymbol{j} \cdot \boldsymbol{n}\, dS$$

for any surface S that spans the loop C. Using Stokes's theorem to transform the l.h.s. gives

$$\iint_S \boldsymbol{\nabla} \times \boldsymbol{B} \cdot \boldsymbol{n}\, dS = \mu_0 \iint_S \boldsymbol{j} \cdot \boldsymbol{n}\, dS.$$

Now if this is true for any loop C, and so any surface S, it follows that

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_0\, \boldsymbol{j}.$$

Note the similarity between this argument and that used when applying the divergence theorem to the conservation of mass of a fluid in Section 5.1.1.

Example 5.7

Use Stokes's theorem to show that for any closed surface S,

$$\oiint_S (\nabla \times \boldsymbol{u}) \cdot \boldsymbol{n}\, dS = 0.$$

Consider the case where a small hole is made in the closed surface. Then by Stokes's theorem, the surface integral of $(\nabla \times \boldsymbol{u}) \cdot \boldsymbol{n}$ over the surface S is equal to the line integral of $\boldsymbol{u} \cdot \boldsymbol{dr}$ around the perimeter of the small hole. As the size of the hole shrinks to zero, so does the value of the line integral, giving the required result. Note that this result was obtained using the divergence theorem in Example 5.1.

Example 5.8

The surface S is defined by $x^2 + 4y^2 = 1$, $-1 \leq z \leq 1$. Use Stokes's theorem to evaluate the surface integral

$$\iint_S (xz^2, -yz^2, 0) \cdot \boldsymbol{n}\, dS.$$

Note that this surface is not simply connected, but Stokes's theorem can still be applied. By imagining a cut in the surface (Figure 5.7), the surface integral is equal to the sum of two line integrals around the two elliptical curves C_1 and C_2 that form the ends of the cylindrical surface. In order to apply Stokes's theorem, the vector field $(xz^2, -yz^2, 0)$ must be written as the curl of another vector field $\boldsymbol{u}$. Seeking a solution of the form $\boldsymbol{u} = (0, 0, h(x, y, z))$, this can be

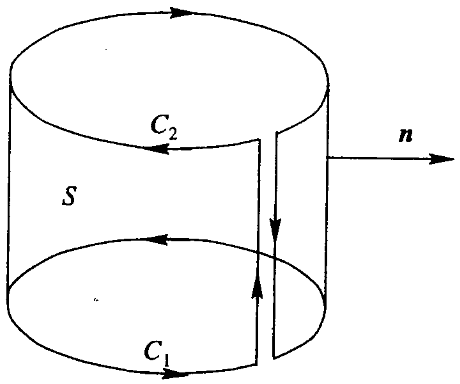


Fig. 5.7. Stokes's theorem can be used to transform the surface integral over the curved surface of the cylinder into two line integrals around the ends of the cylinder, by introducing a cut in the surface. The two line integrals along the cut cancel out.

achieved if

$$xz^2 = \frac{\partial h}{\partial y}, \qquad -yz^2 = -\frac{\partial h}{\partial x}$$

and these two equations are satisfied if $h = xyz^2$. Now since the two curves C_1 and C_2 lie in planes $z =$ constant, so $\boldsymbol{dr} = (dx, dy, 0)$, $\boldsymbol{u} \cdot \boldsymbol{dr} = 0$ on C_1 and C_2 so the value of the integral is zero.

5.2.2 Related theorems linking line and surface integrals

As in the case of the divergence theorem, Stokes's theorem can be used to derive other theorems that relate line integrals to surface integrals by appropriate choices of the vector field $\boldsymbol{u}$.

- Choose $\boldsymbol{u} = \boldsymbol{a}f$, where $\boldsymbol{a}$ is a constant vector and f is a scalar field. Then $\boldsymbol{\nabla} \times \boldsymbol{u} = \boldsymbol{\nabla} f \times \boldsymbol{a} + f\boldsymbol{\nabla} \times \boldsymbol{a} = \boldsymbol{\nabla} f \times \boldsymbol{a}$, so Stokes's theorem gives

$$\iint_S \boldsymbol{\nabla} f \times \boldsymbol{a} \cdot \boldsymbol{n}\, dS = \oint_C \boldsymbol{a} f \cdot \boldsymbol{dr}.$$

  Using the rules for manipulating the scalar triple product and taking out the constant vector $\boldsymbol{a}$ from the integrals gives

$$\boldsymbol{a} \cdot \left(\iint_S -\boldsymbol{\nabla} f \times \boldsymbol{n}\, dS \right) = \boldsymbol{a} \cdot \left(\oint_C f\, \boldsymbol{dr} \right).$$

  As in Section 5.1.3, the constant $\boldsymbol{a}$ can be cancelled, giving

$$\iint_S -\boldsymbol{\nabla} f \times \boldsymbol{n}\, dS = \oint_C f\, \boldsymbol{dr}.$$

- Choose $\boldsymbol{u} = \boldsymbol{a} \times \boldsymbol{v}$, where $\boldsymbol{a}$ is a constant vector and $\boldsymbol{v}$ is a vector field. This case is more complicated, but provides a good example for the use of suffix notation. The line integral in Stokes's theorem is

$$\oint_C \boldsymbol{a} \times \boldsymbol{v} \cdot \boldsymbol{dr} = \boldsymbol{a} \cdot \oint_C \boldsymbol{v} \times \boldsymbol{dr},$$

  interchanging the dot and the cross and taking the constant $\boldsymbol{a}$ outside the integral. Since $\boldsymbol{\nabla} \times (\boldsymbol{a} \times \boldsymbol{v}) = \boldsymbol{a}(\boldsymbol{\nabla} \cdot \boldsymbol{v}) - \boldsymbol{a} \cdot \boldsymbol{\nabla} \boldsymbol{v}$, from formula (4.30), the surface integral is

$$\begin{aligned}
\iint_S (\boldsymbol{a}(\boldsymbol{\nabla} \cdot \boldsymbol{v}) - \boldsymbol{a} \cdot \boldsymbol{\nabla} \boldsymbol{v}) \cdot \boldsymbol{n}\, dS &= \iint_S \left(a_j \frac{\partial v_k}{\partial x_k} - a_k \frac{\partial v_j}{\partial x_k} \right) n_j\, dS \\
&= \iint_S a_j \left(\frac{\partial v_k}{\partial x_k} n_j - \frac{\partial v_k}{\partial x_j} n_k \right) dS
\end{aligned}$$

where j and k have been interchanged in the second term. Applying Stokes's theorem and cancelling the a_j gives a relationship between a surface integral and a line integral in suffix notation:

$$\iint_S \frac{\partial v_k}{\partial x_k} n_j - \frac{\partial v_k}{\partial x_j} n_k \, dS = \left[\oint_C \boldsymbol{v} \times d\boldsymbol{r} \right]_j . \tag{5.18}$$

To obtain the form of this equation in vector notation, consider the quantity $[(\boldsymbol{n} \times \boldsymbol{\nabla}) \times \boldsymbol{v}]_j$. In suffix notation this is

$$\begin{aligned} \epsilon_{jkl}\epsilon_{kmn} n_m \frac{\partial}{\partial x_n} v_l &= (\delta_{lm}\delta_{jn} - \delta_{ln}\delta_{jm}) n_m \frac{\partial v_l}{\partial x_n} \\ &= n_l \frac{\partial v_l}{\partial x_j} - n_j \frac{\partial v_l}{\partial x_l}. \end{aligned}$$

This is now minus the quantity appearing in the surface integral (5.18), so the vector form of (5.18) is

$$\iint_S -(\boldsymbol{n} \times \boldsymbol{\nabla}) \times \boldsymbol{v} \; dS = \oint_C \boldsymbol{v} \times d\boldsymbol{r}. \tag{5.19}$$

Note that suffix notation is the only secure method for obtaining such results, short of writing out all the components of the vector quantities longhand. Attempts to expand using the rules for a vector triple product generally give incorrect results.

- Choose the surface S to be a flat surface lying in the x, y plane, so $\boldsymbol{n} = (0, 0, 1)$ and choose $\boldsymbol{u} = (F(x, y), G(x, y), 0)$. Then

$$\boldsymbol{\nabla} \times \boldsymbol{u} = \left(0, 0, \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right)$$

and $\boldsymbol{u} \cdot d\boldsymbol{r} = F\,dx + G\,dy$. Then Stokes's theorem gives

$$\iint_S \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y} \, dx\, dy = \oint_C F\,dx + G\,dy, \tag{5.20}$$

a result known as Green's theorem.

Summary of Chapter 5

- The divergence theorem states that
$$\iiint_V \boldsymbol{\nabla} \cdot \boldsymbol{u} \, dV = \oiint_S \boldsymbol{u} \cdot \boldsymbol{n} \, dS,$$
where S is the surface enclosing the volume V and $\boldsymbol{n}$ is the outward-pointing unit normal vector.
- Geometrically, the divergence theorem follows naturally from the physical definition of the divergence.
- The divergence theorem has many applications, including simplifying the evaluation of surface or volume integrals, deriving physical conservation laws and showing that Laplace's equation has a unique solution.
- Stokes's theorem states that
$$\iint_S \boldsymbol{\nabla} \times \boldsymbol{u} \cdot \boldsymbol{n} \, dS = \oint_C \boldsymbol{u} \cdot \boldsymbol{dr},$$
where the curve C encloses the surface S and C and $\boldsymbol{n}$ are oriented in a right-handed sense.
- Stokes's theorem follows from the definition of the curl in terms of a line integral.
- A number of other related theorems linking volume, surface and line integrals can be derived from the divergence theorem and Stokes's theorem.

EXERCISES

5.8 Show that

$$\oint_C \boldsymbol{r} \cdot \boldsymbol{dr} = 0$$

for any closed curve C.

5.9 Verify Stokes's theorem by evaluating both the line and surface integrals for the vector field $\boldsymbol{u} = (2x - y, -y^2, -y^2 z)$ and the surface S given by the disk $z = 0$, $x^2 + y^2 \leq 1$.

5.10 Use Stokes's theorem to show that

$$\oint_C f \nabla g \cdot \boldsymbol{dr} = - \oint_C g \nabla f \cdot \boldsymbol{dr}$$

for any closed curve C and differentiable scalar fields f and g.

5.11 If $\boldsymbol{u}$ is irrotational, express the surface integral

$$\iint_S \boldsymbol{u} \times \nabla f \cdot \boldsymbol{n} \, dS$$

as a line integral.

5.12 The magnetic field $\boldsymbol{B}$ in an electrically conducting fluid moving with velocity $\boldsymbol{u}$ obeys the magnetic induction equation

$$\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{B}).$$

Show that the total flux of magnetic field through a surface enclosed by a streamline of the flow (a closed curve which is everywhere parallel to $\boldsymbol{u}$) is independent of time.

5.13 Use (5.18) to show that the area A of a flat surface S enclosed by a curve C is

$$A = 1/2 \left| \oint_C \boldsymbol{r} \times \boldsymbol{dr} \right| .$$