

Chapter 3

Vectors and Differential Calculus

The mathematical description of physical phenomena is often formulated in terms of vectors that depend on one or more variables. For example, the trajectory of a particle under the influence of a force \mathbf{F} is described by a position vector $\mathbf{r}(t)$ that depends on the time. Once we assign a meaning to the differentiation of this vector, we can express Newton's second law for this particle as

$$m \frac{d^2 \mathbf{r}}{dt^2} = \mathbf{F},$$

in which m is the mass of the particle.

Vectors and differential calculus is the amalgamation of vector analysis and differential calculus and allows us to define the differentiation of vectors and various quantities associated with vectors, as discussed in Chapter 2. In this chapter we consider only vector *functions* of a single variable. This will provide the setting for an introductory treatment of derivatives of vectors that can be used to solve problems in classical mechanics. Extensions of these ideas to vector and scalar functions in three dimensions, known as scalar and vector *fields*, respectively, is known as *vector calculus*, and will be covered in a later course. A brief introduction will be provided in this chapter.

We begin this chapter with a summary of the definition of the derivative and the standard results involving derivatives of sums, products and quotients. This enables us to define the derivative of a vector which, in turn, can be extended to compound expressions of scalar and vector functions. We then work through several examples in physics to illustrate the ease with which many standard results can be recovered with a few simple vector manipulations. This chapter concludes with a brief introduction to the calculus of vector fields, which assign a vector to every point in space. The new types of derivatives will be defined and the corresponding derivatives, along with their physical meanings will be explained.

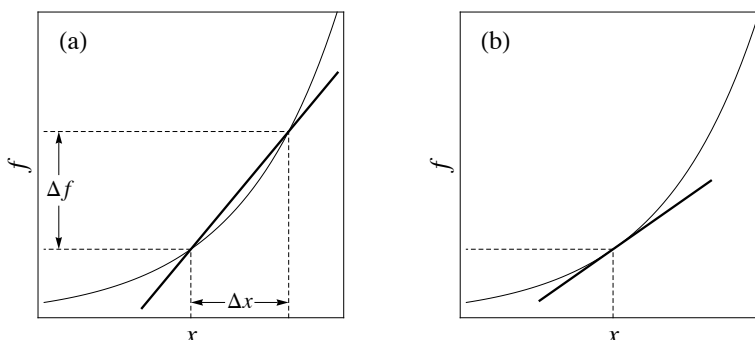


Figure 3.1: The construction of the derivative in (3.1). (a) The line through the point (x, f) with slope $\Delta f / \Delta x$. (b) The effect of taking the limit $\Delta x \rightarrow 0$, which results in a line through (x, f) that is tangent to f at x .

3.1 Summary of Differential Calculus

The **derivative** of a function f of a single independent variable x is defined by the following limit:

$$\frac{df}{dx} \equiv \lim_{\Delta x \rightarrow 0} \left[\frac{f(x + \Delta x) - f(x)}{\Delta x} \right]. \quad (3.1)$$

As the construction in Fig. 3.1 demonstrates, the derivative is the slope of the tangent to f at the point x . All of the properties of derivatives of sums, products, and quotients of functions follow from this definition, as do the derivatives of specific functions, as the following example shows.

EXAMPLE 3.1. Suppose $f(x) = x^2$. The derivative of this function with respect to x can be calculated from first principles by using the definition in (3.1):

$$\begin{aligned} \frac{d(x^2)}{dx} &= \lim_{\Delta x \rightarrow 0} \left[\frac{(x + \Delta x)^2 - x^2}{\Delta x} \right] = \lim_{\Delta x \rightarrow 0} \left[\frac{2x\Delta x + (\Delta x)^2}{\Delta x} \right] \\ &= \lim_{\Delta x \rightarrow 0} (2x + \Delta x) = 2x. \end{aligned} \quad (3.2)$$

Indeed, by using the binomial theorem and essentially the same steps as above, we can obtain the general result

$$\frac{d(x^n)}{dx} = nx^{n-1}, \quad (3.3)$$

for any positive integer n . ■

The fundamental ('first principles') definition in (3.1) can be used to show the following well-known formulae of sums, products, and quotients of functions, and the 'chain rule' for composite functions (i.e. functions of functions):

$$\frac{d}{dx}(af + bg) = a\frac{df}{dx} + b\frac{dg}{dx}, \quad \frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx}, \quad (3.4)$$

$$\frac{d}{dx}\left(\frac{f}{g}\right) = \frac{1}{g^2}\left(\frac{df}{dx}g - f\frac{dg}{dx}\right), \quad \frac{d f(g(x))}{dx} = \frac{df}{dg}\frac{dg}{dx}, \quad (3.5)$$

in which a and b are any constants and f and g are any differentiable functions.

3.2 Derivative of a Vector

Suppose we have a vector $\mathbf{a}(s)$ that is a function of a parameter s . Such vectors occur naturally in physical applications, where \mathbf{a} is a radius vector and s is the time, so the curve traced out by the heads of the vectors is the trajectory of a particle [Fig. 3.2(a)]. The following analysis is not limited to three dimensions, so we will work with an n -dimensional vector. To calculate the derivative of $\mathbf{a}(s)$, we apply the definition (3.1) to this vector:

$$\frac{d\mathbf{a}}{ds} = \lim_{\Delta s \rightarrow 0} \left[\frac{\mathbf{a}(s + \Delta s) - \mathbf{a}(s)}{\Delta s} \right]. \quad (3.6)$$

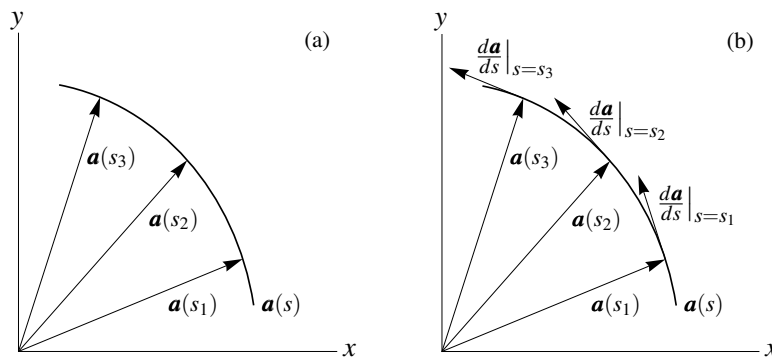


Figure 3.2: (a) A two-dimensional vector $\mathbf{a}(s)$ shown for $s = s_1, s_2, s_3$. The head of this vector traces out a curve as s is varied. (b) The derivatives of this vector at $s = s_1, s_2, s_3$, which are tangent to the curve at these points. This results from the analogous construction shown in Fig. 3.1.

The numerator on the right-hand side is the difference of two vectors, and the denominator is a scalar that multiplies each vector. Thus, using (2.1) and (2.9), we can write this equation in terms of the components of the vector as

$$\begin{aligned}\frac{d\mathbf{a}}{ds} &= \lim_{\Delta s \rightarrow 0} \left[\left(\frac{a_1(s + \Delta s) - a_1(s), a_2(s + \Delta s) - a_2(s), \dots, a_n(s + \Delta s) - a_n(s)}{\Delta s} \right) \right] \\ &= \left(\lim_{\Delta s \rightarrow 0} \left[\frac{a_1(s + \Delta s) - a_1(s)}{\Delta s} \right], \dots, \lim_{\Delta s \rightarrow 0} \left[\frac{a_n(s + \Delta s) - a_n(s)}{\Delta s} \right] \right) \\ &= \left(\frac{da_1}{ds}, \frac{da_2}{ds}, \dots, \frac{da_n}{ds} \right).\end{aligned}\quad (3.7)$$

Thus, the derivative of a vector is obtained as the derivative of each component of that vector:

$$\boxed{\frac{d\mathbf{a}}{ds} = \left(\frac{da_1}{ds}, \frac{da_2}{ds}, \dots, \frac{da_n}{ds} \right).} \quad (3.8)$$

EXAMPLE 3.2. Consider the particle trajectory

$$\mathbf{r}(t) = (x(t), y(t), z(t)) = (\cos t, \sin t, t). \quad (3.9)$$

The physical origin of this trajectory will be discussed later in this chapter. The derivative of this vector, which yields the *velocity* of the particle at each point along the trajectory, is calculated by invoking (3.8):

$$\frac{d\mathbf{r}}{dt} = (-\sin t, \cos t, 1). \quad (3.10)$$

By following the construction in Fig. 3.1, the derivative vector is seen to be tangent to the trajectory at each point [Fig. 3.2(b)]. ■

3.3 Derivatives of Compound Expressions

The derivative (3.8) provides the foundation for calculating all types of expressions involving products of scalar functions and products of vectors. Such products are called *compound expressions*. In this section, we calculate the derivatives of the product of a scalar function and a vector, the dot product of two vectors, and the cross product of two vectors. In each case, the basic method involves reducing the calculation of derivatives of vectors to the derivatives of their components, to which the basic results of differential calculus can be applied.

3.3.1 Product of a Scalar Function and a Vector Function

The multiplication of an n -dimensional vector $\mathbf{a}(s)$ by a scalar function $f(s)$ is written as

$$f(s)\mathbf{a}(s) = (f(s)a_1(s), f(s)a_2(s), \dots, f(s)a_n(s)). \quad (3.11)$$

The derivative of this vector is obtained by using the rule for the differentiation of the product of scalar functions, which is the second equation in (3.4):

$$\begin{aligned} \frac{d(f\mathbf{a})}{ds} &= \left(\frac{d(fa_1)}{ds}, \frac{d(fa_2)}{ds}, \dots, \frac{d(fa_n)}{ds} \right) \\ &= \left(\frac{df}{ds}a_1 + f\frac{da_1}{ds}, \frac{df}{ds}a_2 + f\frac{da_2}{ds}, \dots, \frac{df}{ds}a_n + f\frac{da_n}{ds} \right) \\ &= \left(\frac{df}{ds}a_1, \frac{df}{ds}a_2, \dots, \frac{df}{ds}a_n \right) + \left(f\frac{da_1}{ds}, f\frac{da_2}{ds}, \dots, f\frac{da_n}{ds} \right) \\ &= \frac{df}{ds}(a_1, a_2, \dots, a_n) + f\left(\frac{da_1}{ds}, \frac{da_2}{ds}, \dots, \frac{da_n}{ds} \right) \\ &= \frac{df}{ds}\mathbf{a} + f\frac{d\mathbf{a}}{ds}. \end{aligned} \quad (3.12)$$

Thus, the derivative of (3.11) is analogous to that of a product in ordinary calculus:

$$\boxed{\frac{d(f\mathbf{a})}{ds} = \frac{df}{ds}\mathbf{a} + f\frac{d\mathbf{a}}{ds}.} \quad (3.13)$$

3.3.2 The Dot Product

The derivative of a dot product of n -dimensional vectors $\mathbf{a}(s)$ and $\mathbf{b}(s)$ proceeds with similar steps as in the preceding section:

$$\begin{aligned} \frac{d(\mathbf{a} \cdot \mathbf{b})}{ds} &= \frac{d}{ds}(a_1b_1 + a_2b_2 + \dots + a_nb_n) \\ &= \left(\frac{d(a_1b_1)}{ds} + \frac{d(a_2b_2)}{ds} + \dots + \frac{d(a_nb_n)}{ds} \right) \\ &= \left(\frac{da_1}{ds}b_1 + \frac{da_2}{ds}b_2 + \dots + \frac{da_n}{ds}b_n \right) + \left(a_1\frac{db_1}{ds} + a_2\frac{db_2}{ds} + \dots + a_n\frac{db_n}{ds} \right) \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{da_1}{ds}, \frac{da_2}{ds}, \dots, \frac{da_n}{ds} \right) \cdot (b_1, b_2, \dots, b_n) \\
&\quad + (a_1, a_2, \dots, a_n) \cdot \left(\frac{db_1}{ds}, \frac{db_2}{ds}, \dots, \frac{db_n}{ds} \right) \\
&= \frac{d\mathbf{a}}{ds} \cdot \mathbf{b} + \mathbf{a} \cdot \frac{d\mathbf{b}}{ds}.
\end{aligned} \tag{3.14}$$

The derivative of a dot product is, thus, analogous to that of a product in ordinary calculus. We have preserved the order of the two factors in our calculation, but, in accordance with the properties of this product, this is not necessary. To summarize:

$$\boxed{\frac{d(\mathbf{a} \cdot \mathbf{b})}{ds} = \frac{d\mathbf{a}}{ds} \cdot \mathbf{b} + \mathbf{a} \cdot \frac{d\mathbf{b}}{ds}} \tag{3.15}$$

3.3.3 The Cross Product

The derivative of a cross product also proceeds as in the preceding two sections, but we will now restrict ourselves to three-dimensional vectors and we *must* maintain the order of the factors, as required by property 2 in Sec. 2.2.2. Then, from (2.55), we have

$$\begin{aligned}
\frac{d(\mathbf{a} \times \mathbf{b})}{ds} &= \frac{d}{ds} \left[(a_2b_3 - a_3b_2)\mathbf{i} + (a_3b_1 - a_1b_3)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \right] \\
&= \left(\frac{d(a_2b_3 - a_3b_2)}{ds}, \frac{d(a_3b_1 - a_1b_3)}{ds}, \frac{d(a_1b_2 - a_2b_1)}{ds} \right) \\
&= \left(\frac{da_2}{ds}b_3 + a_2\frac{db_3}{ds} - \frac{da_3}{ds}b_2 - a_3\frac{db_2}{ds}, \right. \\
&\quad \left. \frac{da_3}{ds}b_1 + a_3\frac{db_1}{ds} - \frac{da_1}{ds}b_3 - a_1\frac{db_3}{ds}, \right. \\
&\quad \left. \frac{da_1}{ds}b_2 + a_1\frac{db_2}{ds} - \frac{da_2}{ds}b_1 - a_2\frac{db_1}{ds} \right) \\
&= \left(\frac{da_2}{ds}b_3 - \frac{da_3}{ds}b_2 + a_2\frac{db_3}{ds} - a_3\frac{db_2}{ds}, \right. \\
&\quad \left. \frac{da_3}{ds}b_1 - \frac{da_1}{ds}b_3 + a_3\frac{db_1}{ds} - a_1\frac{db_3}{ds}, \right.
\end{aligned}$$

$$\begin{aligned}
& \left(\frac{da_1}{ds}b_2 - \frac{da_2}{ds}b_1 + a_1 \frac{db_2}{ds} - a_2 \frac{db_1}{ds} \right) \\
&= \left(\frac{da_2}{ds}b_3 - \frac{da_3}{ds}b_2, \frac{da_3}{ds}b_1 - \frac{da_1}{ds}b_3, \frac{da_1}{ds}b_2 - \frac{da_2}{ds}b_1 \right) \\
&+ \left(a_2 \frac{db_3}{ds} - a_3 \frac{db_2}{ds}, a_3 \frac{db_1}{ds} - a_1 \frac{db_3}{ds}, a_1 \frac{db_2}{ds} - a_2 \frac{db_1}{ds} \right) \\
&= \frac{d\mathbf{a}}{ds} \times \mathbf{b} + \mathbf{a} \times \frac{d\mathbf{b}}{ds}, \tag{3.16}
\end{aligned}$$

which again follows the product rule of ordinary differential calculus. In summary,

$$\boxed{\frac{d(\mathbf{a} \times \mathbf{b})}{ds} = \frac{d\mathbf{a}}{ds} \times \mathbf{b} + \mathbf{a} \times \frac{d\mathbf{b}}{ds}}. \tag{3.17}$$

3.4 Examples from Physics

Derivatives of vectors are used throughout classical physics. In this section, we consider three applications that illustrate how some standard results are obtained through vector analysis.

EXAMPLE 3.3. Figure 3.2 shows the relationship between the trajectory defined by a position vector $\mathbf{r}(t)$ and the derivative of that vector, which is the tangent vector to the trajectory. Suppose we have a trajectory such that the direction changes, but the length does not, that is, the trajectory is a circle. As the length of the vector does not change, $|\mathbf{r}(t)|$ is a constant, which means that $|\mathbf{r}(t)|^2$ is also a constant. Taking the derivative of the latter expression using (3.15),

$$\frac{d(|\mathbf{r}|^2)}{dt} = \frac{d(\mathbf{r} \cdot \mathbf{r})}{dt} = \frac{d\mathbf{r}}{dt} \cdot \mathbf{r} + \mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = 2\mathbf{r} \cdot \frac{d\mathbf{r}}{dt}. \tag{3.18}$$

Since the magnitude of $\mathbf{r}(t)$ is assumed to be constant, the left-hand side of this equation vanishes, whereupon we conclude that

$$\mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = 0, \tag{3.19}$$

which means that $\mathbf{r}(t)$ and the velocity vector are *perpendicular* to one another. ■

EXAMPLE 3.4. A somewhat more substantial example is based on Newton's law of gravitation, which is a special case of Newton's second law. We consider a body

of mass m orbiting around a body with a much larger mass M . The trajectory $\mathbf{r}(t)$ of the orbiting body is determined by

$$m \frac{d^2 \mathbf{r}}{dt^2} = -\frac{GMm}{r^2} \hat{\mathbf{r}}, \quad (3.20)$$

where G is the gravitational constant and

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|} \equiv \frac{\mathbf{r}}{r}, \quad (3.21)$$

is the unit vector along the radial direction.

To explore the consequences of this equation of motion, we take the cross product of \mathbf{r} and each side of the equation. After cancelling the factor of m from this equation, we find:

$$\mathbf{r} \times \frac{d^2 \mathbf{r}}{dt^2} = -\frac{GM}{r^2} (\mathbf{r} \times \hat{\mathbf{r}}) = 0. \quad (3.22)$$

This derivative vanishes because \mathbf{r} and $\hat{\mathbf{r}}$ are collinear. Now consider the time-dependence of the quantity $\mathbf{r} \times d\mathbf{r}/dt$:

$$\frac{d}{dt} \left(\mathbf{r} \times \frac{d\mathbf{r}}{dt} \right) = \frac{d\mathbf{r}}{dt} \times \frac{d\mathbf{r}}{dt} + \mathbf{r} \times \frac{d^2 \mathbf{r}}{dt^2} = 0, \quad (3.23)$$

because the first term is a cross product of a vector with itself, and therefore vanishes, and the second term vanishes on account of (3.22). Hence, we conclude that

$$\mathbf{r} \times \frac{d\mathbf{r}}{dt} = \mathbf{c}, \quad (3.24)$$

in which \mathbf{c} is a constant vector (that is, independent of time). Referring to Fig. 3.3, we see that the differential area dA swept out by the trajectory is $dA = \frac{1}{2} |\mathbf{r} \times d\mathbf{r}|$, which, in terms of time derivatives is

$$\frac{dA}{dt} = \frac{1}{2} \left| \mathbf{r} \times \frac{d\mathbf{r}}{dt} \right| = \frac{c}{2}, \quad (3.25)$$

which is a constant. Integrating this equation over the time interval (t_1, t_2) ,

$$\int_{t_1}^{t_2} \frac{dA}{dt} dt = A(t_2) - A(t_1) = \int_{t_1}^{t_2} \frac{c}{2} dt = \frac{c}{2} (t_2 - t_1). \quad (3.26)$$

We have therefore shown that the trajectory sweeps equal areas in equal times. This is Kepler's second law. ■

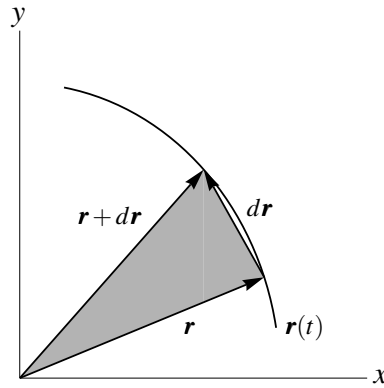


Figure 3.3: The construction used in Example 3.4 to prove Kepler's second law. The shaded area corresponds to $\frac{1}{2}|\mathbf{r} \times d\mathbf{r}|$.

EXAMPLE 3.5. The Lorentz force on a particle of charge q in a magnetic field \mathbf{B} , but no electric field is $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, where $\mathbf{v}(t)$ is the velocity of the particle. Newton's second law for this particle can be written as

$$m \frac{d\mathbf{v}}{dt} = q\mathbf{v} \times \mathbf{B}, \quad (3.27)$$

where m is the mass of the particle. Suppose that $\mathbf{B} = (0, 0, B_z)$, so only the z -component is non-zero. The cross product in the Lorentz force is calculated by using (2.55):

$$q\mathbf{v} \times \mathbf{B} = q \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_x & v_y & v_z \\ 0 & 0 & B_z \end{vmatrix} = qv_y B_z \mathbf{i} - qv_x B_z \mathbf{j}. \quad (3.28)$$

Notice in particular, that there is no z -component of the force, so v_z remains constant. If $v_z = 0$, then the motion of the particle is entirely in the x - y plane.

From the definition of the cross product, the Lorentz force is perpendicular to both \mathbf{v} and \mathbf{B} . Consider the differential work done by a force moving an object by a distance $d\mathbf{r}$:

$$dW = \mathbf{F} \cdot d\mathbf{r} = \mathbf{F} \cdot \mathbf{v} dt, \quad (3.29)$$

which vanishes for the Lorentz force. Hence, the Lorentz force does no work on the particle, so the kinetic energy of the particle does not change. The solution of Newton's second law using (3.28) with $v_z = 0$ shows that the trajectory is a circle.

The Lorentz force must point inward, so for $q < 0$, the right-hand rule requires that the particle motion is in the counterclockwise direction when viewed from the positive z -axis. If $q > 0$, then the particles moves in the clockwise direction when viewed from the positive z -axis. ■

3.5 Calculus of Three-Dimensional Vector Fields¹

The mathematical description of continuous media is often expressed in terms of quantities called ‘fields’, which, to every point (x, y, z) in a region of space assign a scalar $f(x, y, z)$, or a vector $\mathbf{v}(x, y, z)$. These are called scalar and vector fields, respectively. In addition to their spatial dependence, they may also depend on the time. The fundamental equations in many branches of science and engineering are formulated in terms of scalar and vector fields

Vector calculus is the extension of differential and integral calculus to higher spatial dimensions. Differentiation and integration have some straightforward extensions in higher dimensions, but there also some altogether new constructions that underlie the conceptual and computational richness of multivariable calculus. The connection between derivatives and integrals through higher-dimensional analogs of the Fundamental Theorem of Calculus is embodied by the main integral theorems of vector calculus.

Vector calculus is based on the vector differential operator

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right) = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k}, \quad (3.30)$$

which is read as ‘del’ or ‘nabla’. This operator can be applied to scalar and vector functions in a way that formally is similar to scalar multiplication of a vector and vector dot and cross products. Each operation gives a different type of information about the function being differentiated.

The analog of scalar multiplication is the application of (3.30) to a *scalar* function $f(x, y, z)$ and is known as the **gradient** of f :

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k}. \quad (3.31)$$

The result of this operation is a *vector*. The direction of ∇f is the direction in which the change in f has the largest value and the magnitude $|\nabla f|$ is the value of the corresponding derivative.

¹For information only; not examinable.

The construction analogous to the dot product between (3.30) and a vector function

$$\mathbf{v}(x, y, z) = v_x(x, y, z)\mathbf{i} + v_y(x, y, z)\mathbf{j} + v_z(x, y, z)\mathbf{k}, \quad (3.32)$$

is called the **divergence** of \mathbf{v} and is given by

$$\nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}. \quad (3.33)$$

The physical significance of the divergence is the rate at which ‘the density’ described by the vector field (e.g. charge, fluid) exits, or ‘diverges’, from a region of space. The definition of the divergence follows naturally by noting that, in the absence of sources or sinks of matter, the mass density within a region of space can change only by the flow of matter into or out of that region. This property is fundamental in physics and is known as the **principle of continuity**.

Finally, the analog of the cross product between (3.30) and the vector function (3.32) is called the **curl** of \mathbf{v} :

$$\begin{aligned} \nabla \times \mathbf{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ v_x & v_y & v_z \end{vmatrix} \\ &= \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \mathbf{k}. \end{aligned} \quad (3.34)$$

The curl of a vector field is the amount of ‘rotation’ of the contents of given region of space. For example, the vorticity of a fluid field is described by the curl of the velocity field of that fluid.

3.6 Summary

This chapter provided a discussion of how to take derivatives of vectors and associated quantities. The fundamental result is the definition of the derivative of a vector in (3.8). All the other results in this chapter were obtained by applying this formula and reducing vector expressions to enable the rules for derivatives in ordinary calculus to be applied.

