Chapter 4

Differentiation of vectors

4.1 Vector-valued functions

In the previous chapters we have considered real functions of several (usually two) variables $f\colon D \to \mathbb{R}$, where D is a subset of $\mathbb{R}^n$, where n is the number of variables. These are *scalar-valued* functions in the sense that the result of applying such a function is a real number, which is a scalar quantity. We now wish to consider *vector-valued* functions $\mathbf{f}\colon D \to \mathbb{R}^m$. In principal, m can be any positive integer, but we will only consider the cases where $m = 2$ or 3, and the results of applying the function is either a 2D or 3D vector.

4.2 Parametric equations of curves

The simplest type of vector-valued function has the form $\mathbf{f}\colon I \to \mathbb{R}^2$, where $I \subset \mathbb{R}$. Such a function returns a 2D vector $\mathbf{f}(t)$ for each $t \in I$, which may be regarded as the position vector of some point on the plane.

For example, recall the Section Formula from Level 1. This states that the position vector of any point P on the line through points A and B is

$$\mathbf{p} = \frac{\alpha\mathbf{a} + \beta\mathbf{b}}{\alpha + \beta},$$

for any scalars α, β. If we define $t = \beta/(\alpha + \beta)$, then this may be rewritten as

$$\mathbf{p}(t) = (1 - t)\mathbf{a} + t\mathbf{b}.$$

As t changes, we get different points on the line through A and B and in particular, $\mathbf{p}(0) = \mathbf{a}$ and $\mathbf{p}(1) = \mathbf{b}$. We may think of $\mathbf{p}$ as a vector-valued function

$$\mathbf{p}\colon \mathbb{R} \to \mathbb{R}^3,$$

the image of which is the whole of the line, or

$$\mathbf{p}\colon [0, 1] \to \mathbb{R}^3,$$

the image of which is the line segment from A to B.

In general, a curve, in 2D or 3D space, can be represented as the image of a vector-valued function on an interval I; the position vector of a point on the curve is

$$\mathbf{r} = \mathbf{f}(t), \quad t \in I.$$

This is called a *parametric* description of the curve and t is called a *parameter*. This may also be written in component form; if $\mathbf{r} = (x, y, z)$ and $\mathbf{f} = (f_1, f_2, f_3)$ then

$$x = f_1(t), \ y = f_2(t), \ z = f_3(t), \quad t \in I.$$

Standard types of parametric curve

Circle and ellipse   The circle $(x-a)^2+(y-b)^2 = r^2$, having centre (a, b) and radius r, can be parametrised using polar coordinates $x - a = r\cos\theta$ and $y - b = r\sin\theta$. Recall that θ is the angle between the radius and the positive x-axis, measured in an anit-clockwise direction. Hence the circle has parametric form

$$x = a + r\cos\theta, \ y = b + r\sin\theta, \quad \theta \in [0, 2\pi).$$

If this circle were to be thought of as a curve on the xy-plane in 3D space then it would be

$$x = a + r\cos\theta, \ y = b + r\sin\theta, \ z = 0, \quad \theta \in [0, 2\pi).$$

In a similar way, the ellipse

$$\frac{(x-a)^2}{A^2} + \frac{(y-b)^2}{B^2} = 1,$$

has parametric form

$$x = a + A\cos\theta, \ y = b + B\sin\theta, \quad \theta \in [0, 2\pi).$$

Parabola   The parabola $y^2 = 4ax$ can be parametrised as

$$x = at^2, \ y = 2at, \quad t \in (-\infty, \infty).$$

To show that the parametric curve is identical to the parabola we must prove that every point on the parametric curve lies on the parabola and vice versa. For any t, let $x = at^2$ and $y = 2at$ then $y^2 = 4a^2t^2 = 4a(at^2) = 4ax$ so that every point on the parametric curve lies on the parabola. Also, given any point (x, y) on the parabola, define $t = y/2a$ so that $y = 2at$ and then $x = y^2/4a = at^2$, so that (x, y) also lies on the parametric curve.

Line   We have already seen that

$$\mathbf{r} = (1 - t)\mathbf{a} + t\mathbf{b}, \quad t \in [0, 1],$$

is the parametric form of the line segment joining $A(a_1, a_2, a_3)$ and $B(b_1, b_2, b_3)$. This may also be written in component form as

$$x = (1 - t)a_1 + b_1, \ y = (1 - t)a_2 + b_2, \ z = (1 - t)a_3 + b_3, \quad t \in [0, 1].$$

Also, if one is given a point $\mathbf{a}$ on the line and a direction vector $\mathbf{d}$ for the line then the parametric form is

$$\mathbf{r} = \mathbf{a} + t\mathbf{d}, \quad t \in \mathbb{R}.$$

4.3   Differentiation of vector-valued functions

A curve $\mathcal{C}$ is defined by $\mathbf{r} = \mathbf{r}(t)$, a vector-valued function of one (scalar) variable. Let us imagine that $\mathcal{C}$ is the path taken by a particle and t is time. The vector $\mathbf{r}(t)$ is the position vector of the particle at time t and $\mathbf{r}(t + h)$ is the position vector at a later time $t + h$. The *average velocity* of the particle in the time interval $[t, t + h]$ is then

$$\frac{\text{displacement vector}}{\text{length of time interval}} = \frac{\mathbf{r}(t + h) - \mathbf{r}(t)}{h}.$$

See Figure 4.1. In terms of the components of $\mathbf{r}$ this is

$$\left(\frac{x(t + h) - x(t)}{h}, \frac{y(t + h) - y(t)}{h}, \frac{z(t + h) - z(t)}{h} \right).$$
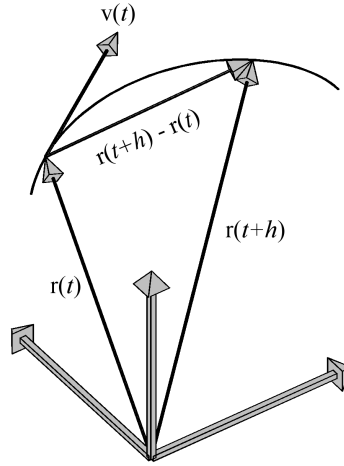
Figure 4.1: Velocity

If each of the scalar function x, y and z are differentiable, then this vector has a limit

$$\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = (\dot{x}, \dot{y}, \dot{z}),$$

which is the *instantaneous velocity* of the particle $\mathbf{v}(t)$. This means that (if the motion is smooth) then

$$\mathbf{v}(t) = \lim_{h \to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h} = \frac{d}{dt}\mathbf{r}(t) = \dot{\mathbf{r}}(t).$$

This vector lies along the tangent to the curve at $\mathbf{r}$ and has length $v(t) = |\mathbf{v}(t)|$ which is the *instantaneous speed* of the particle.

In a similar way, we define the *acceleration*,

$$\mathbf{a}(t) = \frac{d}{dt}\mathbf{v}(t) = \frac{d^2}{dt^2}\mathbf{r}(t) = \ddot{\mathbf{r}}(t).$$

More generally, for any curve $\mathbf{r} = \mathbf{r}(\alpha)$ parametrised by α (say), the vector

$$\mathbf{T} = \frac{d\mathbf{r}}{d\alpha}$$

is called the *tangent vector* to the curve and the unit vector

$$\hat{\mathbf{T}} = \frac{\mathbf{T}}{|\mathbf{T}|},$$

is called the *unit tangent vector*.

Example 4.1 A particle moves with constant angular speed (i.e. rate of change of angle) ω around a circle of radius a and centre $(0, 0)$ and the particle is initially at $(a, 0)$. Show that the position of the particle is $\mathbf{r}(t) = a(\cos \omega t, \sin \omega t)$.

Determine the velocity and speed of the particle at time t and prove that the acceleration of the particle is always directed towards the centre of the circle.

3

Solution :

Answer: Velocity, $\mathbf{v} = a\omega(-\sin\omega t, \cos\omega t)$ and speed $v = a\omega$.

□

Example 4.2 Find the velocity of a particle with position vector $\mathbf{r}(t) = (\cos^2 t, \sin^2 t, \cos 2t)$. Describe the motion of the particle.

Solution :

Answer: The velocity is: $\mathbf{v} = \sin 2t(-1, 1, -2)$. □

Example 4.3 Find the tangent vector and the equation of the tangent to the helix

$$x = \cos\theta, \ y = \sin\theta, \ z = \theta, \quad \theta \in [0, 2\pi),$$

at the point where it crosses the xy-plane.

Solution :

Answer: The tangent vector is $\mathbf{T} = (-\sin\theta, \cos\theta, 1)$ and the equation of the tangent is given by $\mathbf{r} = (1, 0, 0) + t(0, 1, 1), \quad t \in \mathbb{R}$. □

4.4 Vector and scalar fields

A function of two or three variables mapping to a vector is called a *vector field*. In contrast, a function of two or three variables mapping to a scalar is called a *scalar field*. As we saw in Chapter 1 (using different terminology), one can represent the graph of a scalar field as a curve or surface. A vector field $\mathbf{F}(x, y)$ (or $\mathbf{F}(x, y, z)$) is often represented by drawing the vector $\mathbf{F}(\mathbf{r})$ at point $\mathbf{r}$ for representative points in the domain. A good example of a vector field is the velocity at a point in a fluid; at each point we draw an arrow (vector) representing the velocity (the speed and direction) of fluid flow (see Figure 4.2). The length of the arrow represents the fluid speed at each point.
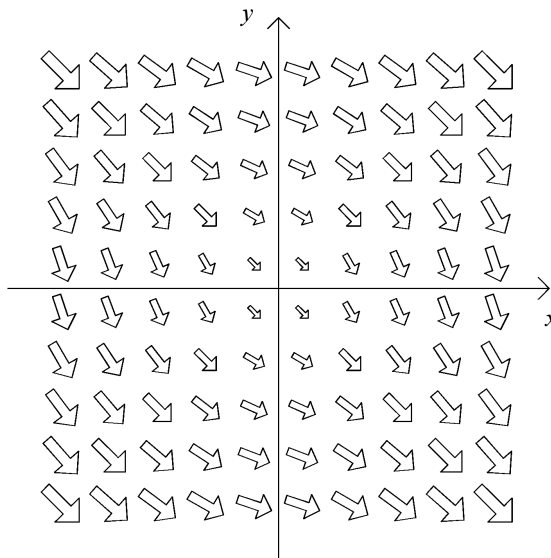

Figure 4.2: Vector field representing fluid velocity

4.5 Different types of derivative

We have already discussed the derivatives and partial derivatives of scalar functions. Next we will consider discuss other different types of "derivatives" of scalar and vector functions; in some cases the result is a scalar and sometimes a vector.

Recall that if $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are vectors and α is a scalar, there are a number of different products that can be made;

Name of product	Formula	Type of result
Scalar multiplication	$\alpha\mathbf{u}$	Vector
Scalar or dot product	$\mathbf{u} \cdot \mathbf{v}$	Scalar
Vector or cross product	$\mathbf{u} \times \mathbf{v}$	Vector

.

Now consider the vector differential operator

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

This is read as *del* or *nabla* and is not to be confused with Δ, the capital Greek letter delta. One can form "products" of this vector with other vectors and scalars, but because it is an operator, it always has to be

the first term if the product is to make sense. For example, if f is a scalar field, we can form the scalar "multiple" with ∇ as the first term

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

the result being a vector.

Below we will introduce the "derivatives" corresponding to the product of vectors given in the above table.

4.5.1 Gradient ("multiplication by a scalar")

This is just the example given above. We define the *gradient* of a scalar field f to be

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

We will use both of the notation grad f and ∇f interchangably.

Remark Note that f must be a *scalar* field for grad f to be defined and grad f itself is a *vector* field.

Example 4.4 Find the gradient of the scalar field $f(x, y, z) = x^2 y + x \cosh yz$.

Solution :

Answer: grad $f = (2xy + \cosh yz, x^2 + xz \sinh yz, xy \sinh yz)$. □

Example 4.5 Let $\mathbf{r} = (x, y, z)$ so that $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$. Show that

$$\nabla(r^n) = nr^{n-2}\mathbf{r},$$

for any integer n and deduce the values of $\operatorname{grad}(r)$, $\operatorname{grad}(r^2)$ and $\operatorname{grad}(1/r)$.

7

Solution :

$\square$

Example 4.6 Determine $\mathrm{grad}(\mathbf{c} \cdot \mathbf{r})$, when c is a constant (vector).

Solution :

Answer: $\mathrm{grad}(\mathbf{c} \cdot \mathbf{r}) = \mathbf{c}$ $\square$

Direction derivative This is the rate of change of a scalar field f in the direction of a *unit* vector $\mathbf{u} = (u_1, u_2, u_3)$. As with normal derivatives it is defined by the limit of a difference quotient, in this case the direction derivative of f at $\mathbf{p}$ in the direction $\mathbf{u}$ is defined to be

$$\lim_{h \to 0^+} \frac{f(\mathbf{p} + h\mathbf{u}) - f(\mathbf{p})}{h}, \tag{$*$}$$

(if the limit exists) and is denoted

$$\frac{\partial f}{\partial \mathbf{u}}(\mathbf{p}).$$

This definition is rarely used directly. The key formula for the directional derivatives of f in the direction $\mathbf{u}$ is

$$\boxed{\frac{\partial f}{\partial \mathbf{u}} = \mathbf{u} \cdot \nabla f = u_1 \frac{\partial f}{\partial x} + u_2 \frac{\partial f}{\partial y} + u_3 \frac{\partial f}{\partial z}}.$$

To prove this, first notice that

$$\frac{d}{dt} f(\mathbf{p} + t\mathbf{u}) = \lim_{h \to 0^+} \frac{f(\mathbf{p} + (t+h)\mathbf{u}) - f(\mathbf{p} + t\mathbf{u})}{h}$$

so that $(*)$ can be obtained as

$$\left. \frac{d}{dt} f(\mathbf{p} + t\mathbf{u}) \right|_{t=0}.$$

Also, using the chain rule, we have

$$\frac{d}{dt} f(\mathbf{p} + t\mathbf{u}) = u_1 \frac{\partial f}{\partial x}(\mathbf{p} + t\mathbf{u}) + u_2 \frac{\partial f}{\partial y}(\mathbf{p} + t\mathbf{u}) + u_3 \frac{\partial f}{\partial z}(\mathbf{p} + t\mathbf{u}) = \mathbf{u} \cdot \nabla f(\mathbf{p} + t\mathbf{u}).$$

Combining these results gives the required formula.

Remarks

1. The formula for a directional derivatives can only be used for *unit* vectors. To calculate the directional derivative along a non-unit vector $\mathbf{v}$, one must use the unit vector having the same direction as $\mathbf{v}$, that is

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|}.$$

2. Partial derivatives are special cases of directional derivatives. For example, the partial x-derivative is the directional derivative in the direction $(1, 0, 0)$.

Example 4.7 Find the directional derivative of $f = x^2 y z^3$ at the point $P(3, -2, -1)$ in the direction of the vector $(1, 2, 2)$.

9

Solution :

Answer: The direction derivative is given by, $\frac{\partial f}{\partial \mathbf{u}}(3, -2, -1) = -38$. $\square$

4.5.2 Divergence of a vector field ("scalar product")

The *divergence* of a vector field $\mathbf{F} = (F_1, F_2, F_3)$ is the *scalar* obtained as the "scalar product" of ∇ and $\mathbf{F}$,

$$\boxed{\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}} \ .$$

It is so called, because it measures the tendency of a vector field to diverge (positive divergence) or converge (negative divergence). In particular, a vector field is said to be *incompressible* (or *solenoidal*) if its divergence is zero.

Figure 4.3 shows the vector fields $\mathbf{F} = (x, y, 0)$, $\mathbf{G} = (x, -y, 0)$ and $\mathbf{H} = (-x, -y, 0)$ in the xy-plane. We have

$$\operatorname{div} \mathbf{F} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} = 2 > 0$$

and similarly, $\operatorname{div} \mathbf{G} = 0$ and $\operatorname{div} \mathbf{H} = -2 < 0$. Notice how the arrows on the plot of $\mathbf{F}$ diverge and on the plot of $\mathbf{H}$ converge.

Example 4.8 Show that the divergence of $\mathbf{F} = (x - y^2, z, z^3)$ is positive at all points in $\mathbb{R}^3$.
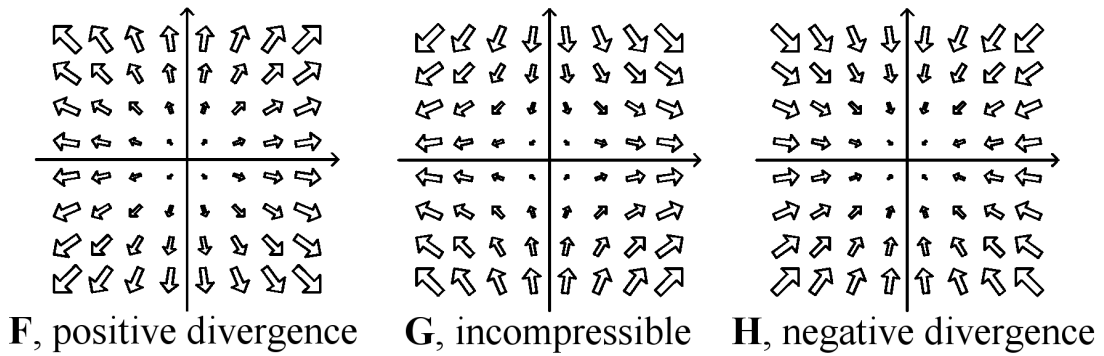
Solution :

$\square$

Figure 4.3: Positive and negative divergence

A particular example of divergence is the *Laplacian* of a scalar field. Given a scalar field f, grad $f = \nabla f$ is a vector field and the divergence of ∇f is the Laplacian of f, written $\nabla^2 f$. This means that

$$\nabla^2 f = \nabla \cdot (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \; .$$

This definition may be extended in a natural way to the Laplacian of a vector field $\mathbf{F} = (F_1, F_2, F_3)$,

$$\nabla^2 \mathbf{F} = (\nabla^2 F_1, \nabla^2 F_2, \nabla^2 F_3) \; .$$

Example 4.9 Find the values of n for which $\nabla^2 (r^n) = 0$.

11

Solution :

Answer: $n = 0$ or $n = -1$.   □

4.5.3   Curl of a vector field ("vector product")

The *curl* of a vector field $\mathbf{F} = (F_1, F_2, F_3)$ is the *vector* obtained as the "vector product" of ∇ and $\mathbf{F}$

$$\boxed{\operatorname{curl}\mathbf{F} = \nabla \times \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}} \; .$$

Like any other vector product, $\operatorname{curl}\mathbf{F}$ can be calculated using a 3×3 determinant,

$$\operatorname{curl}\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k} \; .$$

The curl of a vector field measures its tendency to rotate. In particular, a vector field is said to be *irrotational* if its curl is the zero vector. Figure 4.4 shows the vector fields $\mathbf{F} = (-y, x, 0)$, $\mathbf{G} = (y, x, 0)$ and

12

$\mathbf{H} = (y, -x, 0)$. We have

$$\operatorname{curl} \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -y & x & 0 \end{vmatrix} = 2\mathbf{k}$$

and similarly, $\operatorname{curl} \mathbf{G} = \mathbf{0}$ and $\operatorname{curl} \mathbf{H} = -2\mathbf{k} < 0$. The coefficient of $\mathbf{k}$ in $\operatorname{curl} \mathbf{F}$ being positive indicates anticlockwise rotation.
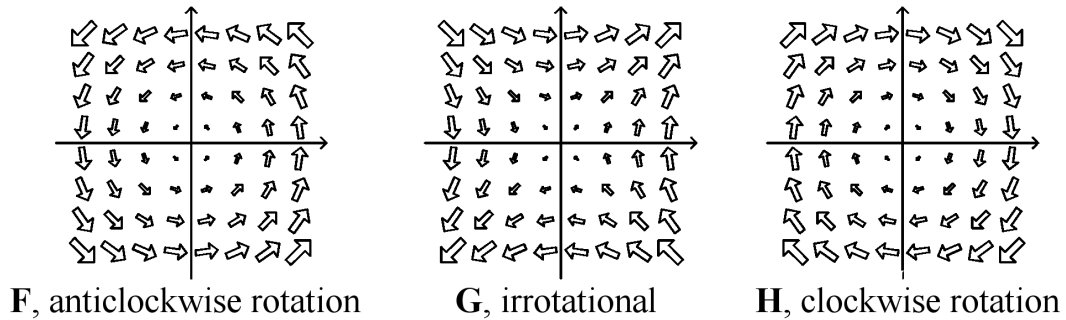

$\mathbf{F}$, anticlockwise rotation $\mathbf{G}$, irrotational $\mathbf{H}$, clockwise rotation

Figure 4.4: Clockwise and anticlockwise rotation

Example 4.10 Determine $\operatorname{curl} \mathbf{F}$ when $\mathbf{F} = (x^2 y, xy^2 + z, xy)$.

Solution :

Answer: $\operatorname{curl} \mathbf{F} = (x - 1, -y, 0)$. □

Example 4.11 If $\mathbf{c}$ is a constant vector, find $\operatorname{curl}(\mathbf{c} \times \mathbf{r})$.

Solution :

Answer: $\operatorname{curl}(\mathbf{c} \times \mathbf{r}) = 2\mathbf{c}$. □

4.6   Nabla identities

There are analogues involving div, grad and curl of the elementary rules of differentiation such as linearity $(f + g)'(x) = f'(x) + g'(x)$ the product rule $(fg)'(x) = f(x)g'(x) + f'(x)g(x)$.

Let f and g be smooth scalar fields and $\mathbf{F}$ and $\mathbf{G}$ smooth vector fields. Then all of the following are straightforward to prove (as illustrated in Example 4.12) just using definitions

$$
\begin{array}{ll}
\operatorname{grad}(f + g) = \operatorname{grad} f + \operatorname{grad} g & \operatorname{grad}(fg) = f(\operatorname{grad} g) + (\operatorname{grad} f)g, \\
\operatorname{div}(\mathbf{F} + \mathbf{G}) = \operatorname{div}\mathbf{F} + \operatorname{div}\mathbf{G} & \operatorname{div}(f\mathbf{F}) = f\operatorname{div}\mathbf{F} + (\operatorname{grad} f) \cdot \mathbf{F}, \\
\operatorname{curl}(\mathbf{F} + \mathbf{G}) = \operatorname{curl}\mathbf{F} + \operatorname{curl}\mathbf{G} & \operatorname{curl}(f\mathbf{F}) = f\operatorname{curl}\mathbf{F} + \operatorname{grad} f \times \mathbf{F}, \\
\operatorname{curl}\operatorname{grad} f = \mathbf{0}, & \operatorname{div}\operatorname{curl}\mathbf{F} = 0.
\end{array}
$$

In particular, note the special cases

$$
\operatorname{grad}(cf) = c\operatorname{grad} f, \quad \operatorname{div}(c\mathbf{F}) = c\operatorname{div}\mathbf{F}, \quad \operatorname{curl}(c\mathbf{F}) = c\operatorname{curl}\mathbf{F},
$$

when c is a (scalar) constant.

All of the identities are easier to remember if written using ∇. For example,

$$
\begin{aligned}
\operatorname{curl}(f\mathbf{F}) &= \nabla \times (f\mathbf{F}) \\
&= f(\nabla \times \mathbf{F}) + (\nabla f) \times \mathbf{F} \\
&= f\operatorname{curl}\mathbf{F} + \operatorname{grad} f \times \mathbf{F}.
\end{aligned}
$$

Example 4.12 Prove the identities

    (i) $\operatorname{curl}\operatorname{grad} f = 0$,   (ii) $\operatorname{curl}(f\mathbf{F}) = f\operatorname{curl}\mathbf{F} + \operatorname{grad} f \times \mathbf{F}$,   (iii) $\operatorname{div}(f\mathbf{F}) = f\operatorname{div}\mathbf{F} + (\operatorname{grad} f)\cdot\mathbf{F}$

Solution :

$\square$

Example 4.13 Let $\mathbf{F} = (x^2 y, yz, x + z)$. Find

$$\text{(i)}\ \operatorname{curl}\operatorname{curl}\mathbf{F}, \quad \text{(ii)}\ \operatorname{grad}\operatorname{div}\mathbf{F}.$$

Solution :

Answer: (i) $\operatorname{curl} \operatorname{curl} \mathbf{F} = (0, 2x, 1)$ and (ii) $\operatorname{grad} \operatorname{div} \mathbf{F} = (2y, 2x, 1)$.   □

Example 4.14 Let $\mathbf{r} = (x, y, z)$ denote a position vector with length $r = \sqrt{x^2 + y^2 + z^2}$ and $\mathbf{c}$ is a constant (vector). Determine

$$\text{(i) } \operatorname{div}(r^n(\mathbf{c} \times \mathbf{r})), \quad \text{(ii) } \operatorname{curl}(r^n(\mathbf{c} \times \mathbf{r})).$$

Solution :

Answer: (i) $\mathrm{div}(r^n(\mathbf{c} \times \mathbf{r})) = 0$ and (ii) $\mathrm{curl}(r^n(\mathbf{c} \times \mathbf{r})) = (n+2)r^n\mathbf{c} - n(\mathbf{r} \cdot \mathbf{c})r^{n-2}\mathbf{r}$. □