Savitribai Phule Pune University, Pune

(Formerly University of Pune)

Board of Studies in Mathematics

(Faculty of Science & Technology)

A Text Book

for

S.Y.B.A. / S.Y.B.Sc. as per NEP-2020
For colleges affilated to Savitribai Phule Pune University, Pune
To be implemented from academic year 2024-25

MTS 241 MN(A): Mathematics for Physical Sciences

Dr. Swapnil V. Kale
(PhD (DRDO-DIAT), PostDoc (IISER-TVM), SET, NET^3(JRF), GATE)
Dept. of Mathematics & Statistics
C. T. Bora College, Shirur, 412210, Pune.
email: swapnilkale@iisertvm.ac.in

Prof. Pratiksha N. Kale
(MSc, BEd)
Dept. of Mathematics & Statistics
C. T. Bora College, Shirur, 412210, Pune.
email:- kalepratiksha021@gmail.com

Course Content

Unit 1: Partial Differentiation **(10 Hours)**

1.1 Definition of function of several variables

1.2 Definition of the partial derivative

1.3 The total differential and total derivative

1.4 Exact and inexact differentials

1.5 The chain rule and Change of variables

Unit 2: Vector Algebra **(10 Hours)**

2.1 Vector Algebra

2.2 Scalars and vectors

2.3 Addition and subtraction of vectors

2.4 Multiplication by a scalar

2.5 Magnitude of a vector

2.6 Angle between the vectors

2.7 Multiplication of vectors

2.8 Scalar Product

2.9 Vector Product

2.10 Scalar Triple Product

2.11 Vector Triple Product

Unit 3: Vector Calculus **(10 Hours)**

3.1 Differentiation of vectors

3.2 Vector operators:
- 3.2.1. Gradient of a scalar field;
- 3.2.2. Divergence of a vector field;
- 3.2.3. Curl of a vector field

3.3 Vector operator formulae
- 3.3.1. Vector operators acting on sums and products
- 3.3.2. Combinations of grad, div and curl

Recommended Book:

1. Mathematical Methods for Physics and Engineering, K. F Riley, Michael Paul Hobson, and Stephen John Bence, Cambridge University Press.

   Section 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 7.1, 7.2, 7.3, 7.5, 7.6, 10.1, 10.7, 10.8

Reference Books:

1. Advanced Engineering Mathematics, Erwin Kreyszig, John Wiley & Sons.
2. Mathematics for Chemistry, Kailas S. Ahire, Rajashri Sawant, SahityaSagar Publication 2023.

Declaration: The authors acknowledge that the content of this book is inspired by the book Mathematical Methods for Physics and Engineering by K. F. Riley, Michael P. Hobson, and Stephen J. Bence, published by Cambridge University Press. They do not claim original authorship of the material presented herein. For the convenience of the reader, the equation numbering and certain other elements have been intentionally aligned with those in the referenced book. While every effort has been made to maintain accuracy, some typographical errors may still be present. In such cases, readers are encouraged to consult the original source for clarification.

Chapter 1. Partial Differentiation

Partial differentiation is a fundamental concept in calculus used when dealing with functions that depend on two or more variables. It involves finding the rate of change of a function with respect to one variable while keeping the other variables constant. For example, if a function depends on both x and y, the partial derivative with respect to x measures how the function changes as x changes, assuming y stays the same. This is especially useful in real-world problems where many factors affect the outcome. Partial derivatives are used to study curves and surfaces, to find maxima and minima of multivariable functions, and to develop mathematical models in science and engineering.

Partial differentiation has many practical applications. In physics, it is used to describe heat flow, wave propagation, and electric or magnetic fields. In engineering, it helps analyze stress and strain in materials, optimize designs, and solve problems in fluid mechanics. In economics, it is used to study how changing one factor (like price or demand) affects outcomes like profit or cost, while keeping other variables constant. In chemistry, partial derivatives are very useful in thermodynamics, where properties such as pressure, volume, temperature, and entropy depend on multiple variables. For example, the rate at which pressure changes with volume at constant temperature is a partial derivative. They are also used in studying chemical kinetics to understand how the concentration of one reactant affects the reaction rate, and in equilibrium analysis where multiple factors influence the balance of a chemical reaction. It is also essential in biology and environmental science when modeling systems with many interacting variables. Overall, partial differentiation is a powerful mathematical tool that helps us understand and solve complex problems involving several changing quantities.

1.1. Definition of function of several variables

Definition (Function of Two Variables):
A function of two variables assigns a real number z to each ordered pair (x, y) in a region of the xy-plane.
We write this as:

$$z = f(x, y)$$

where x and y are real numbers, and $f(x, y)$ is a real-valued function.

Generalization (Function of n Variables):
A function of n variables assigns a real number to each ordered n-tuple $(x_1, x_2, \dots, x_n)$.
We write this as:

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x_1, x_2, \dots, x_n)$$

This means the function takes n real inputs and gives one real output.

1.2. Definition of the partial derivative:

For a function of two variables $f(x, y)$, we may define the derivative with respect to x, for example, by saying that it is that for a one-variable function when y is held fixed and treated as a constant. To signify that a derivative is with respect to x, but at the same time to recognize that a derivative with respect to y also exists, the former is denoted by $\frac{\partial f}{\partial x}$ and is the partial derivative of $f(x, y)$ with respect to x. Similarly, the partial derivative of f with respect to y is denoted by $\frac{\partial f}{\partial y}$.
To define formally the partial derivative of $f(x, y)$ with respect to x, we have

$$\frac{\partial f}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \tag{5.1}$$

provided that the limit exists. This is much the same as for the derivative of a one-variable function. The other partial derivative of $f(x, y)$ is similarly defined as a limit (provided it exists):

$$\frac{\partial f}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}. \tag{5.2}$$

It is common practice in connection with partial derivatives of functions involving more than one variable to indicate those variables that are held constant by writing them as subscripts to the derivative symbol. Thus, the partial derivatives defined in (5.1) and (5.2) would be written respectively as

$$\left(\frac{\partial f}{\partial x}\right)_y \quad \text{and} \quad \left(\frac{\partial f}{\partial y}\right)_x .$$

In this form the subscript shows that which variable is kept constant. Further, the extension of the definitions (5.1), (5.2) to the general n-variable case is straightforward and can be written formally as

$$\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, \dots, x_i + \Delta x_i, \dots, x_n) - f(x_1, x_2, \dots, x_i, \dots, x_n)}{\Delta x_i},$$

provided that the limit exists Just as for one-variable functions, second (and higher) partial derivatives may be defined in a similar way. For a two-variable function $f(x, y)$, they are

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy},$$
$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{xy}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}.$$

Only three of the second derivatives are independent since the relation

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

is always obeyed, provided that the second partial derivatives are continuous at the point in question. This relation often proves useful as a labour-saving device when evaluating second partial derivatives. It can also be shown that for a function of n variables, $f(x_1, x_2, \dots, x_n)$, under the same conditions,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Example: Find the first and second partial derivatives of the function

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

First Partial Derivatives:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(2x^3y^2 + y^3) = 6x^2y^2,$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(2x^3y^2 + y^3) = 4x^3y + 3y^2.$$

Second Partial Derivatives:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(6x^2y^2\right) = 12xy^2,$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(4x^3y + 3y^2\right) = 4x^3 + 6y,$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(4x^3y + 3y^2\right) = 12x^2y,$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(6x^2y^2\right) = 12x^2y.$$

1.3. The total differential and total derivative

Having defined the (first) partial derivatives of a function $f(x, y)$, which give the rate of change of f along the positive x- and y-axes, we next consider the rate of change of $f(x, y)$ in an arbitrary direction. Suppose that we make small simultaneous changes Δx in x and Δy in y, and as a result, f changes to $f + \Delta f$. Then we must have:

$$\begin{aligned}\Delta f &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= \big[f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)\big] + \big[f(x, y + \Delta y) - f(x, y)\big] \\ &= \frac{f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)}{\Delta x} \cdot \Delta x \\ &\quad + \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} \cdot \Delta y \end{aligned} \tag{5.3}$$

In the last step, we observe that the expressions inside the brackets are very similar to the definitions of the partial derivatives given in equations (5.1) and (5.2). To make them exactly equal to the partial derivatives, the values of Δx and Δy would need to be extremely small almost zero. However, even when Δx and Δy are small but not zero, we can still use the following approximate formula:

$$\Delta f \approx \frac{\partial f(x, y)}{\partial x} \cdot \Delta x + \frac{\partial f(x, y)}{\partial y} \cdot \Delta y \tag{5.4}$$

Note that the first term in equation (5.3) is actually closer to the partial derivative $\frac{\partial f(x,y+\Delta y)}{\partial x}$, but in formula (5.4), we have replaced it with $\frac{\partial f(x,y)}{\partial x}$. This is a valid approximation, and both versions give nearly the same result when Δx and Δy are small.

The accuracy of the approximation in equation (5.4) compared to equation (5.3) depends not only on how small Δx and Δy are, but also on the sizes of the higher-order partial derivatives. However, if we allow the small changes Δx and Δy in equation (5.4) to become infinitesimally small, we can define the total differential df of the function $f(x, y)$, without any approximation, as:

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy \tag{5.5}$$

Equation (5.5) can be extended to the case of a function of n variables, $f(x_1, x_2, \ldots, x_n)$, as:

$$df = \frac{\partial f}{\partial x_1}\, dx_1 + \frac{\partial f}{\partial x_2}\, dx_2 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n \tag{5.6}$$

Example: Find the total differential of the function f(x, y) = y exp(x + y).
Solution: Evaluating the first partial derivatives, we find

$$\frac{\partial f}{\partial x} = y\, \exp(x+y), \quad \frac{\partial f}{\partial y} = \exp(x+y) + y\, \exp(x+y).$$

Applying equation (5.5), we then find that the total differential is given by

$$df = [y\, \exp(x+y)]\, dx + [(1+y)\, \exp(x+y)]\, dy.$$

In some situations, despite the fact that several variables x_i, where $i = 1, 2, \ldots, n$, appear to be involved, effectively only one of them is independent. This occurs if there are subsidiary relationships constraining all the x_i to values that depend on one of them, say x_1. These relationships may be represented by equations of the form:

$$x_i = x_i(x_1), \quad i = 2, 3, \ldots, n. \tag{5.7}$$

In principle, the function f can then be expressed as a function of x_1 alone by substituting from (5.7) for $x_2, x_3, \ldots, x_n$, and then the total derivative (or simply the derivative) of f with respect to x_1 is obtained by ordinary differentiation. Alternatively, equation (5.6) can be used to write:

$$\frac{df}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \cdot \frac{dx_2}{dx_1} + \cdots + \frac{\partial f}{\partial x_n} \cdot \frac{dx_n}{dx_1}. \tag{5.8}$$

It should be noted that the left-hand side (LHS) of this equation is the total derivative $\frac{df}{dx_1}$, whereas the partial derivative $\frac{\partial f}{\partial x_1}$ forms only one part of the right-hand side (RHS). In evaluating this partial derivative, one must take into account only the explicit appearances of x_1 in the function f, and no allowance should be made for the fact that changing x_1 also changes $x_2, x_3, \ldots, x_n$. The contribution from these changes is represented by the remaining terms on the RHS of (5.8). Naturally, what has been shown using x_1 in the above argument applies equally well to any other variable x_i, with appropriate modifications.

Example: Find the total derivative of $f(x, y) = x^2 + 3xy$ with respect to x, given that $y = \sin^{-1} x$.

Solution We have:

$$\frac{\partial f}{\partial x} = 2x + 3y, \quad \frac{\partial f}{\partial y} = 3x, \quad \frac{dy}{dx} = \frac{1}{\sqrt{1-x^2}}.$$

Therefore, the total derivative is:

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} = 2x + 3y + 3x \cdot \frac{1}{\sqrt{1-x^2}}.$$

Substituting $y = \sin^{-1} x$, we get:

$$\frac{df}{dx} = 2x + 3\sin^{-1} x + \frac{3x}{\sqrt{1-x^2}}.$$

1.4. Exact and inexact differentials

In the last section we discussed how to find the total differential of a function, i.e., its infinitesimal change in an arbitrary direction, in terms of its gradients $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ in the x- and y-directions (see (5.5)). Sometimes, however, we wish to reverse the process and find the function f that differentiates to give a known differential. Usually, finding such functions relies on inspection and experience.

As an example, it is easy to see that the function whose differential is

$$df = x\,dy + y\,dx$$

is simply $f(x, y) = xy + c$, where c is a constant. Differentials such as this, which integrate directly, are called **exact differentials**, whereas those that do not are **inexact differentials**.

For example, $x\,dy + 3y\,dx$ is not the straightforward differential of any function (see below). Inexact differentials can be made exact, however, by multiplying through by a suitable function called an **integrating factor**.

Example: Show that the differential $x\,dy + 3y\,dx$ is inexact.
If we integrate the differential

$$df = x\,dy + 3y\,dx$$

with respect to x, treating y as a constant, we obtain:

$$f(x, y) = \int 3y\,dx = 3xy + g(y),$$

where $g(y)$ is an arbitrary function of y.
Further, if we integrate with respect to y, treating x as a constant, we obtain:

$$f(x, y) = \int x\,dy = xy + h(x),$$

where $h(x)$ is an arbitrary function of x. These two expressions for $f(x, y)$ are inconsistent for any and every choice of $g(y)$ and $h(x)$, since

$$3xy + g(y) \neq xy + h(x)$$

unless $y = 0$, which is not generally the case. Therefore, the differential $x\,dy + 3y\,dx$ is inexact.

It is naturally of interest to investigate which properties of a differential make it exact. Consider the general differential containing two variables,

$$df = A(x, y)\,dx + B(x, y)\,dy.$$

We see that

$$\frac{\partial f}{\partial x} = A(x, y), \quad \frac{\partial f}{\partial y} = B(x, y)$$

and, using the property that mixed partial derivatives are equal ($f_{xy} = f_{yx}$), we therefore require

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}. \tag{5.9}$$

This is in fact both a necessary and a sufficient condition for the differential to be exact.

Example: Using (5.9) show that x dy + 3y dx is inexact.
To show that the differential $x\,dy + 3y\,dx$ is inexact, we compare it with the general form of a differential:

$$df = A(x, y)\,dx + B(x, y)\,dy,$$

where $A(x, y) = 3y$ and $B(x, y) = x$. c The condition for a differential to be exact is:

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}.$$

Now, compute the partial derivatives:

$$\frac{\partial A}{\partial y} = \frac{\partial (3y)}{\partial y} = 3, \quad \text{and} \quad \frac{\partial B}{\partial x} = \frac{\partial (x)}{\partial x} = 1.$$

Since

$$\frac{\partial A}{\partial y} \neq \frac{\partial B}{\partial x},$$

the condition for exactness is not satisfied. Therefore, the differential $x\,dy + 3y\,dx$ is inexact.

Example: Show that $(y + z)\,dx + x\,dy + x\,dz$ is an exact differential.
Let the differential form be:

$$\omega = (y + z)\,dx + x\,dy + x\,dz$$

We assume that there exists a scalar function $f(x, y, z)$ such that:

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy + \frac{\partial f}{\partial z}\,dz$$

Comparing with ω, we set:

$$\frac{\partial f}{\partial x} = y + z, \quad \frac{\partial f}{\partial y} = x, \quad \frac{\partial f}{\partial z} = x$$

Integrate $\frac{\partial f}{\partial x} = y + z$ with respect to x:

$$f(x, y, z) = x(y + z) + C(y, z)$$

Differentiate f with respect to y:

$$\frac{\partial f}{\partial y} = x + \frac{\partial C}{\partial y}$$

Set equal to given value $\frac{\partial f}{\partial y} = x \Rightarrow \frac{\partial C}{\partial y} = 0$
Differentiate f with respect to z:

$$\frac{\partial f}{\partial z} = x + \frac{\partial C}{\partial z}$$

Set equal to given value $\frac{\partial f}{\partial z} = x \Rightarrow \frac{\partial C}{\partial z} = 0$. So, C is a constant. Therefore,

$$f(x, y, z) = x(y + z) + \text{const}$$

Hence,

$$(y + z)\, dx + x\, dy + x\, dz = d[x(y + z)]$$

is an exact differential.

1.5. The chain rule and change of variables

Some useful theorems of partial differentiation: So far our discussion has centred on a function $f(x, y)$ dependent on two variables, x and y. Equally, however, we could have expressed x as a function of f and y, or y as a function of f and x. To emphasise the point that all the variables are of equal standing, we now replace f by z. This does not imply that x, y and z are coordinate positions (though they might be). Since x is a function of y and z, it follows that

$$dx = \left(\frac{\partial x}{\partial y}\right)_z dy + \left(\frac{\partial x}{\partial z}\right)_y dz \tag{5.11}$$

and similarly, since $y = y(x, z)$,

$$dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz. \tag{5.12}$$

We may now substitute (5.12) into (5.11) to obtain

$$dx = \left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial x}\right)_z dx + \left[\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x + \left(\frac{\partial x}{\partial z}\right)_y\right] dz. \tag{5.13}$$

Now if we hold z constant, so that $dz = 0$, we obtain the reciprocity relation

$$\left(\frac{\partial x}{\partial y}\right)_z = \left(\frac{\partial y}{\partial x}\right)_z^{-1},$$

which holds provided both partial derivatives exist and neither is equal to zero. Note, further, that this relationship only holds when the variable being kept constant, in this case z, is the same on both sides of the equation. Alternatively we can put $dx = 0$ in (5.13). Then the contents of the square brackets also equal zero, and we obtain the cyclic relation

$$\left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial y}\right)_z = -1,$$

which holds unless any of the derivatives vanish. In deriving this result we have used the reciprocity relation to replace $\left(\frac{\partial x}{\partial z}\right)_y^{-1}$ by $\left(\frac{\partial z}{\partial x}\right)_y$.

The Chain rule: So far we have discussed the differentiation of a function $f(x, y)$ with respect to its variables x and y. We now consider the case where x and y are themselves functions of another variable, say u. If we wish to find the derivative $\frac{df}{du}$, we could simply substitute in $f(x, y)$ the expressions for $x(u)$ and $y(u)$, and then differentiate the resulting function of u. Such substitution will quickly give the desired answer in simple cases, but in more complicated examples it is easier to make use of the total differentials described in the previous section.

From equation (5.5), the total differential of $f(x, y)$ is given by

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy,$$

but we now note that by using the formal device of dividing through by du, this immediately implies

$$\frac{df}{du} = \frac{\partial f}{\partial x}\frac{dx}{du} + \frac{\partial f}{\partial y}\frac{dy}{du}, \tag{5.14}$$

which is called the chain rule for partial differentiation.

This expression provides a direct method for calculating the total derivative of f with respect to u, and is particularly useful when an equation is expressed in a parametric form.

Example: Given that $x(u) = 1 + au$ and $y(u) = bu^3$, find the rate of change of $f(x, y) = xe^{-y}$ with respect to u.

We are asked to compute $\dfrac{df}{du}$, where $f(x, y) = xe^{-y}$, and $x = x(u)$, $y = y(u)$. By the chain rule for partial differentiation:

$$\frac{df}{du} = \frac{\partial f}{\partial x}\cdot\frac{dx}{du} + \frac{\partial f}{\partial y}\cdot\frac{dy}{du}$$

First, compute the partial derivatives:

$$\frac{\partial f}{\partial x} = e^{-y}, \quad \frac{\partial f}{\partial y} = -xe^{-y}$$

Now compute the derivatives of $x(u)$ and $y(u)$:

$$\frac{dx}{du} = a, \quad \frac{dy}{du} = 3bu^2$$

Substitute everything into the chain rule formula:

$$\frac{df}{du} = e^{-y} \cdot a + (-xe^{-y}) \cdot 3bu^2$$

Factor out e^{-y}:

$$\frac{df}{du} = e^{-y}\left(a - 3bu^2 x\right)$$

Substitute $x = 1 + au$ and $y = bu^3$:

$$\frac{df}{du} = e^{-bu^3}\left[a - 3bu^2(1 + au)\right]$$

Thus, the rate of change of $f(x, y) = xe^{-y}$ with respect to u is:

$$\boxed{\frac{df}{du} = e^{-bu^3}\left[a - 3bu^2(1 + au)\right]}$$

Equation (5.14) is an example of the chain rule for a function of two variables, each of which depends on a single variable. The chain rule may be extended to functions of many variables, each of which is itself a function of a variable u, i.e.,

$$f(x_1, x_2, x_3, \ldots, x_n), \quad \text{with } x_i = x_i(u).$$

In this case, the chain rule gives

$$\frac{df}{du} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{dx_i}{du} = \frac{\partial f}{\partial x_1}\frac{dx_1}{du} + \frac{\partial f}{\partial x_2}\frac{dx_2}{du} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{du}. \tag{5.15}$$

Change of variables: It is sometimes necessary or desirable to make a change of variables during the course of an analysis, and consequently to have to change an equation expressed in one set of variables into an equation using another set. The same situation arises if a function f depends on one set of variables x_i, so that

$$f = f(x_1, x_2, \ldots, x_n)$$

but the x_i are themselves functions of a further set of variables u_j and are given by the equations

$$x_i = x_i(u_1, u_2, \ldots, u_m).$$

For each different value of i, x_i will be a different function of the u_j. In this case, the chain rule (5.15) becomes

$$\frac{\partial f}{\partial u_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial u_j}, \quad j = 1, 2, \ldots, m, \tag{5.17}$$

and is said to express a change of variables. In general, the number of variables in each set need not be equal, i.e. m need not equal n, but if both the x_i and the u_j are sets of independent variables, then $m = n$.

Example: Plane polar coordinates, ρ and ϕ, and Cartesian coordinates, x and y, are related by the expressions

$$x = \rho \cos \phi, \quad y = \rho \sin \phi,$$

as can be seen from Figure. An arbitrary function $f(x, y)$ can be re-expressed as a function $g(\rho, \phi)$. Transform the expression

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

into one in ρ and ϕ.

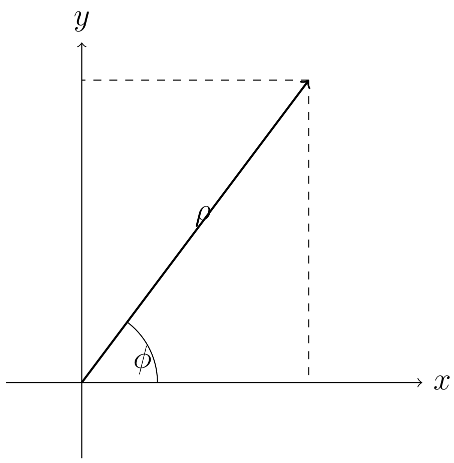

Figure 1: The relationship between Cartesian and plane polar coordinates

Transformation of the Laplacian to Plane Polar Coordinates: Plane polar coordinates ρ and ϕ are related to Cartesian coordinates x and y by:

$$x = \rho \cos \phi, \quad y = \rho \sin \phi.$$

Let $f(x, y) = f(\rho, \phi)$, where the variables ρ and ϕ are defined as:

$$\rho = \sqrt{x^2 + y^2}, \quad \phi = \tan^{-1}\left(\frac{y}{x}\right).$$

We want to compute:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

and express it in terms of ρ and ϕ.

Step 1: First-Order Derivatives Using the Chain Rule

By the multivariable chain rule:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial \rho} \cdot \frac{\partial \rho}{\partial x} + \frac{\partial f}{\partial \phi} \cdot \frac{\partial \phi}{\partial x}$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial \rho} \cdot \frac{\partial \rho}{\partial y} + \frac{\partial f}{\partial \phi} \cdot \frac{\partial \phi}{\partial y}$$

We compute the derivatives:

$$\frac{\partial \rho}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} = \cos\phi, \quad \frac{\partial \rho}{\partial y} = \sin\phi$$

$$\frac{\partial \phi}{\partial x} = -\frac{y}{x^2 + y^2} = -\frac{\sin\phi}{\rho}, \quad \frac{\partial \phi}{\partial y} = \frac{x}{x^2 + y^2} = \frac{\cos\phi}{\rho}$$

Thus,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial \rho}\cos\phi - \frac{1}{\rho}\frac{\partial f}{\partial \phi}\sin\phi$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial \rho}\sin\phi + \frac{1}{\rho}\frac{\partial f}{\partial \phi}\cos\phi$$

Step 2: Second-Order Derivatives

We now compute:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \quad \text{and} \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

The computation is lengthy but follows by applying the product and chain rule to the first derivatives. The result, after simplification, is:

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial f}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2}$$

Thus, the Laplacian in plane polar coordinates is:

$$\nabla^2 f = \frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial f}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \phi^2}$$

Exercise:

Example 1: Using the appropriate properties of ordinary derivatives, perform the following.

(a) Find all the first partial derivatives of the following functions $f(x, y)$:

(i) $f(x, y) = x^2 y$

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2$$

(ii) $f(x, y) = x^2 + y^2 + 4$

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

(iii) $f(x, y) = \sin\left(\frac{x}{y}\right)$

$$\frac{\partial f}{\partial x} = \cos\left(\frac{x}{y}\right) \cdot \frac{1}{y}, \quad \frac{\partial f}{\partial y} = \cos\left(\frac{x}{y}\right) \cdot \left(-\frac{x}{y^2}\right)$$

(iv) $f(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$

$$\frac{\partial f}{\partial x} = \frac{-y}{x^2+y^2}, \quad \frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}$$

(v) $r(x,y,z) = \sqrt{x^2+y^2+z^2}$

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2+z^2}}, \quad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2+z^2}}, \quad \frac{\partial r}{\partial z} = \frac{z}{\sqrt{x^2+y^2+z^2}}$$

(b) For (i), (ii), and (v), find $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y^2}$, and $\frac{\partial^2 f}{\partial x \partial y}$

(i) $f(x,y) = x^2 y$

$$\frac{\partial^2 f}{\partial x^2} = 2y, \quad \frac{\partial^2 f}{\partial y^2} = 0, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2xy) = 2x$$

(ii) $f(x,y) = x^2 + y^2 + 4$

$$\frac{\partial^2 f}{\partial x^2} = 2, \quad \frac{\partial^2 f}{\partial y^2} = 2, \quad \frac{\partial^2 f}{\partial x \partial y} = 0$$

(v) $r(x,y,z) = \sqrt{x^2+y^2+z^2}$ Let $r = (x^2+y^2+z^2)^{1/2}$
Then,

$$\frac{\partial^2 r}{\partial x^2} = \frac{y^2+z^2}{(x^2+y^2+z^2)^{3/2}}, \quad \frac{\partial^2 r}{\partial y^2} = \frac{x^2+z^2}{(x^2+y^2+z^2)^{3/2}}, \quad \frac{\partial^2 r}{\partial x \partial y} = \frac{-xy}{(x^2+y^2+z^2)^{3/2}}$$

(c) For (iv) verify that $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$

Given $f(x,y) = \tan^{-1}\left(\frac{y}{x}\right)$
First,

$$\frac{\partial f}{\partial x} = \frac{-y}{x^2+y^2}, \quad \frac{\partial f}{\partial y} = \frac{x}{x^2+y^2}$$

Now compute mixed partials:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{-(x^2+y^2)(1) + 2y(-y)}{(x^2+y^2)^2} = \frac{-x^2-y^2+2y^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{(x^2+y^2)(1) - 2x \cdot x}{(x^2+y^2)^2} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}$$

Hence,

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

Example 2: Determine which of the following are exact differentials
To determine whether a differential expression

$$M(x,y)\,dx + N(x,y)\,dy$$

is exact, compute the partial derivatives:

$$\frac{\partial M}{\partial y} \quad \text{and} \quad \frac{\partial N}{\partial x}$$

If they are equal, the differential is exact.

(a) $(3x+2)y\,dx + x(x+1)\,dy$

$$M = (3x+2)y, \quad N = x(x+1)$$
$$\frac{\partial M}{\partial y} = 3x+2, \quad \frac{\partial N}{\partial x} = 2x+1$$

Since $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$, not exact.

(b) $y\tan x\,dx + x\tan y\,dy$

$$M = y\tan x, \quad N = x\tan y$$
$$\frac{\partial M}{\partial y} = \tan x, \quad \frac{\partial N}{\partial x} = \tan y$$

Since $\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$, not exact.

(c) $y^2(\ln x + 1)\,dx + 2xy\ln x\,dy$

$$M = y^2(\ln x + 1), \quad N = 2xy\ln x$$
$$\frac{\partial M}{\partial y} = 2y(\ln x + 1), \quad \frac{\partial N}{\partial x} = 2y\ln x + 2y$$

Simplify:

$$\frac{\partial M}{\partial y} = 2y(\ln x + 1), \quad \frac{\partial N}{\partial x} = 2y\,(\ln x + 1)$$

The given differential is exact.

(d) $y^2(\ln x + 1)\,dy + 2xy\ln x\,dx$

Rewriting:

$$M = 2xy\ln x, \quad N = y^2(\ln x + 1)$$
$$\frac{\partial M}{\partial y} = 2x\ln x, \quad \frac{\partial N}{\partial x} = y^2 \cdot \frac{1}{x}$$

So:

$$\frac{\partial M}{\partial y} = 2x\ln x, \quad \frac{\partial N}{\partial x} = \frac{y^2}{x}$$

Not equal in general. Therefore, not exact.

(e) $\dfrac{x}{x^2+y^2}\,dy - \dfrac{y}{x^2+y^2}\,dx$

Rewriting:

$$M = -\frac{y}{x^2+y^2}, \quad N = \frac{x}{x^2+y^2}$$

$$\frac{\partial M}{\partial y} = \frac{(x^2+y^2)(-1)+2y\cdot y}{(x^2+y^2)^2} = \frac{-x^2-y^2+2y^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}$$

$$\frac{\partial N}{\partial x} = \frac{(x^2+y^2)(1)-2x\cdot x}{(x^2+y^2)^2} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{-x^2+y^2}{(x^2+y^2)^2}$$

Since $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, exact.

Example 3: Show that the differential $df = x^2\,dy - (y^2+xy)\,dx$ is not exact, but that $dg = \dfrac{1}{xy^2}\,df$ is exact.

Solution:

Step 1: Check if df is exact

We write:

$$df = M(x,y)\,dx + N(x,y)\,dy$$

where:

$$M(x,y) = -(y^2+xy), \quad N(x,y) = x^2$$

Compute the partial derivatives:

$$\frac{\partial M}{\partial y} = -2y - x, \quad \frac{\partial N}{\partial x} = 2x$$

Since:

$$\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

the differential df is not exact.

Step 2: Define $dg = \dfrac{1}{xy^2}\,df$

Substitute:

$$dg = \frac{1}{xy^2}\left[x^2\,dy - (y^2+xy)\,dx\right] = \left(\frac{-y^2-xy}{xy^2}\right)dx + \left(\frac{x^2}{xy^2}\right)dy$$

Simplify:

$$M(x,y) = -\left(\frac{1}{x}+\frac{1}{y}\right), \quad N(x,y) = \frac{x}{y^2}$$

Compute:

$$\frac{\partial M}{\partial y} = \frac{1}{y^2}, \quad \frac{\partial N}{\partial x} = \frac{1}{y^2}$$

Since:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

the differential dg is exact.

Example 4: Show that

$$df = y(1 + x - x^2)\, dx + x(x + 1)\, dy$$

is not an exact differential. Find the differential equation that a function $g(x)$ must satisfy if

$$d\phi = g(x)\, df$$

is to be an exact differential. Verify that $g(x) = e^{-x}$ is a solution of this equation and deduce the form of $\phi(x, y)$. (Determining nature of $\phi(x, y)$ is optional).

Solution:

Step 1. Show that df is not exact

Write:

$$M(x, y) = y(1 + x - x^2), \quad N(x, y) = x(x + 1)$$

Compute:

$$\frac{\partial M}{\partial y} = 1 + x - x^2, \quad \frac{\partial N}{\partial x} = 2x + 1$$

Since:

$$\frac{\partial M}{\partial y} \neq \frac{\partial N}{\partial x}$$

the differential df is **not exact**.

Step 2. Multiply by integrating factor $g(x)$:

Let:

$$d\phi = g(x)\, df = g(x)\left[y(1 + x - x^2)\, dx + x(x + 1)\, dy\right]$$

Then,

$$M = g(x)\, y(1 + x - x^2), \quad N = g(x)\, x(x + 1)$$

Now compute:

$$\frac{\partial M}{\partial y} = g(x)(1 + x - x^2), \quad \frac{\partial N}{\partial x} = g'(x)\, x(x + 1) + g(x)(2x + 1)$$

For exactness:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \Rightarrow g(x)(1 + x - x^2) = g'(x)\, x(x + 1) + g(x)(2x + 1)$$

Rewriting:

$$g'(x)\, x(x + 1) + g(x)(2x + 1 - 1 - x + x^2) = 0 \Rightarrow g'(x)\, x(x + 1) + g(x)(x + x^2) = 0$$

Simplify:

$$g'(x)\, x(x + 1) + g(x)\, x(1 + x) = 0 \Rightarrow x(x + 1)\,[g'(x) + g(x)] = 0$$

Assuming $x(x + 1) \neq 0$, we get the differential equation:

$$g'(x) + g(x) = 0$$

Step 3. Verify that $g(x) = e^{-x}$ is a solution

Compute:

$$g'(x) = -e^{-x}, \quad g(x) = e^{-x} \Rightarrow g'(x) + g(x) = -e^{-x} + e^{-x} = 0$$

Thus, $g(x) = e^{-x}$ satisfies the differential equation.
Step 4. Find $\phi(x, y)$
We now compute:

$$d\phi = e^{-x} \left[y(1 + x - x^2)\, dx + x(x + 1)\, dy \right]$$

Let:

$$M = e^{-x} y(1 + x - x^2), \quad N = e^{-x} x(x + 1)$$

Integrate M w.r.t. x, treating y as constant:

$$\phi(x, y) = \int e^{-x} y(1 + x - x^2)\, dx + h(y)$$

Factor y out:

$$= y \int (1 + x - x^2) e^{-x}\, dx + h(y)$$

Use integration by parts or symbolic integration:

$$\int (1 + x - x^2) e^{-x} dx = -e^{-x}(x^2 + 1)$$

Therefore,

$$\phi(x, y) = -y(x^2 + 1)e^{-x} + h(y)$$

Differentiate with respect to y to find $h'(y)$:

$$\frac{\partial \phi}{\partial y} = -(x^2 + 1)e^{-x} + h'(y)$$

Compare with $N = e^{-x} x(x + 1)$, so:

$$-(x^2 + 1)e^{-x} + h'(y) = e^{-x} x(x + 1) \Rightarrow h'(y) = e^{-x}(x(x + 1) + x^2 + 1)$$

Not consistent unless function of y only. Hence, we need to adjust integration to match both sides by integrating N w.r.t. y:
Alternatively, integrate N w.r.t. y:

$$\phi(x, y) = \int e^{-x} x(x + 1)\, dy = e^{-x} x(x + 1)y + k(x)$$

Differentiate with respect to x:

$$\frac{\partial \phi}{\partial x} = -e^{-x} x(x + 1)y + e^{-x}(2x + 1)y + k'(x)$$

Simplify:

$$= e^{-x} y(-x(x + 1) + 2x + 1) + k'(x) = e^{-x} y(1 + x - x^2) + k'(x)$$

Compare with $M = e^{-x} y(1 + x - x^2) \Rightarrow k'(x) = 0$ So,

$$\phi(x, y) = e^{-x} x(x + 1)y + C$$

Hence, the required function is:

$$\phi(x, y) = e^{-x} x(x + 1)y$$

Example 5: Implicit Differentiation Problem: For the equation $3y = z^3 + 3xz$, which defines z implicitly as a function of x and y.

(a) Evaluate the first and second partial derivatives of z,

(b) Also, verify that

$$x\frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x^2} = 0$$

Step 1: First-order Partial Derivatives
Given:

$$3y = z^3 + 3xz$$

Differentiate both sides with respect to x:

$$0 = 3z^2\frac{\partial z}{\partial x} + 3z + 3x\frac{\partial z}{\partial x}$$

$$0 = \left(3z^2 + 3x\right)\frac{\partial z}{\partial x} + 3z \Rightarrow \frac{\partial z}{\partial x} = \frac{-3z}{3z^2 + 3x} = \frac{-z}{z^2 + x}$$

Now differentiate both sides with respect to y:

$$3 = 3z^2\frac{\partial z}{\partial y} + 3x\frac{\partial z}{\partial y} = (3z^2 + 3x)\frac{\partial z}{\partial y} \Rightarrow \frac{\partial z}{\partial y} = \frac{1}{z^2 + x}$$

Step 2: Second-order Partial Derivatives
(i) Second partial w.r.t. x:
Differentiate $\frac{\partial z}{\partial x} = \frac{-z}{z^2+x}$ w.r.t. x:
Use quotient rule:

$$\frac{\partial^2 z}{\partial x^2} = \frac{-(z^2 + x)\cdot\frac{\partial z}{\partial x} + z(2z\frac{\partial z}{\partial x} + 1)}{(z^2 + x)^2}$$

Simplify numerator:

$$-(z^2 + x)\frac{\partial z}{\partial x} + z(2z\frac{\partial z}{\partial x} + 1) = \left[-(z^2 + x) + 2z^2\right]\frac{\partial z}{\partial x} + z = (z^2 - x)\frac{\partial z}{\partial x} + z$$

So:

$$\frac{\partial^2 z}{\partial x^2} = \frac{(z^2 - x)\frac{\partial z}{\partial x} + z}{(z^2 + x)^2}$$

(ii) Second partial w.r.t. y:
Differentiate $\frac{\partial z}{\partial y} = \frac{1}{z^2+x}$ w.r.t. y:

$$\frac{\partial^2 z}{\partial y^2} = \frac{-2z\frac{\partial z}{\partial y}}{(z^2 + x)^2} = \frac{-2z}{(z^2 + x)^3}$$

Step 3: Verify the PDE
We check:

$$x\frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x^2} = 0$$

Substitute:

$$x\cdot\left(\frac{-2z}{(z^2 + x)^3}\right) + \frac{(z^2 - x)\left(\frac{-z}{z^2+x}\right) + z}{(z^2 + x)^2}$$

First term:

$$\frac{-2xz}{(z^2 + x)^3}$$

Second term:

$$\frac{-z(z^2-x)+z(z^2+x)}{(z^2+x)^3} = \frac{-z^3+xz+z^3+xz}{(z^2+x)^3} = \frac{2xz}{(z^2+x)^3}$$

So sum is:

$$\frac{-2xz+2xz}{(z^2+x)^3} = 0$$

Hence, verified:

$$x\frac{\partial^2 z}{\partial y^2} + \frac{\partial^2 z}{\partial x^2} = 0.$$

⋆ ⋆ ⋆

Chapter 2. Vector algebra

2.1. Vector algebra

Vector algebra is a branch of mathematics that deals with quantities called vectors, which have both magnitude and direction. Unlike scalar quantities, which only have size, vectors can represent movement or influence in a particular direction. Common operations in vector algebra include vector addition, scalar multiplication, dot product, and cross product. These operations help us analyze and solve problems involving direction-based quantities.

Vector algebra has many real-life applications. In physics, vectors are used to describe forces, velocities, and accelerations. In engineering, vectors help in the design and analysis of structures, machines, and circuits by studying how different forces interact. In computer graphics, vector algebra is used to create 2D and 3D images, simulate motion, and perform lighting calculations. In navigation and GPS systems, vectors help calculate directions and shortest paths. In medical imaging and robotics, vectors are used to determine positions and orientations. In chemistry, vector algebra plays an important role in studying molecular geometry, where bond angles and the spatial arrangement of atoms can be described using vectors. It is also used to model dipole moments, forces between charged particles, and the movement of electrons in electric or magnetic fields. In quantum chemistry, vector spaces are essential for describing the behavior of particles. Thus, vector algebra is a powerful tool in understanding systems that involve both size and direction across many scientific fields.

2.2. Scalars and vectors

Definition (Scalars and vectors): The simplest kind of physical quantity is one that can be completely specified by its magnitude, a single number, together with the units in which it is measured. Such a quantity is called a scalar and examples include temperature, time and density.

A vector is a quantity that requires both a magnitude (≥ 0) and a direction in space to specify it completely; we may think of it as an arrow in space. A familiar example is force, which has a magnitude (strength) measured in newtons and a direction of application. The large number of vectors that are used to describe the physical world include velocity, displacement, momentum and electric field. Vectors are also used to describe quantities such as angular momentum and surface elements (a surface element has an area and a direction defined by the normal to its tangent plane); in such cases their definitions may seem somewhat arbitrary (though in fact they are standard) and not as physically intuitive as for vectors such as force. A vector is denoted by bold type, the convention of this book, or by underlining, the latter being much used in handwritten work.

2.3. Addition and subtraction of vectors

The resultant or vector sum of two displacement vectors is the displacement vector that results from performing first one and then the other displacement, as shown in Figure; this process is known as *vector addition*. However, the principle of addition has physical meaning for vector quantities other than displacements; for example, if two forces act on the same body, then the resultant force acting on the body is the vector sum of the two. The addition of vectors only makes physical sense if they are of a like kind, for example, if they are both forces acting in three dimensions.

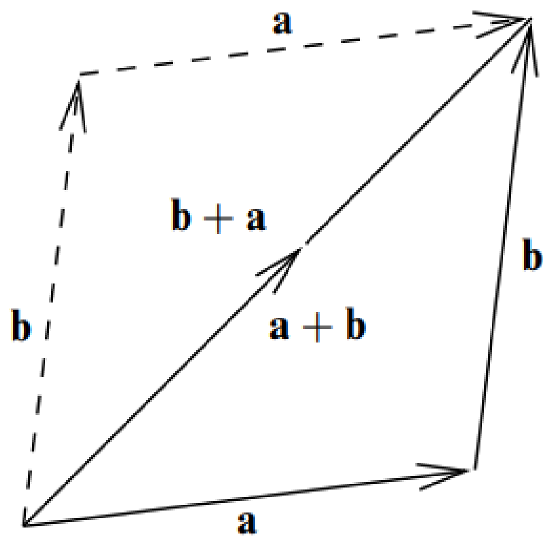

Figure 2: Addition of two vectors showing the commutation relation.

It may be seen from Figure that vector addition is *commutative*, i.e.,

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}. \tag{7.1}$$

The generalisation of this procedure to the addition of three (or more) vectors is clear and leads to the *associativity* property of addition (see Figure), e.g.,

$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}. \tag{7.2}$$

Thus, it is immaterial in what order any number of vectors are added. The subtraction of two vectors is very similar to their addition (see Figure), that is,

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$$

where $-\mathbf{b}$ is a vector of equal magnitude but exactly opposite direction to vector $\mathbf{b}$.

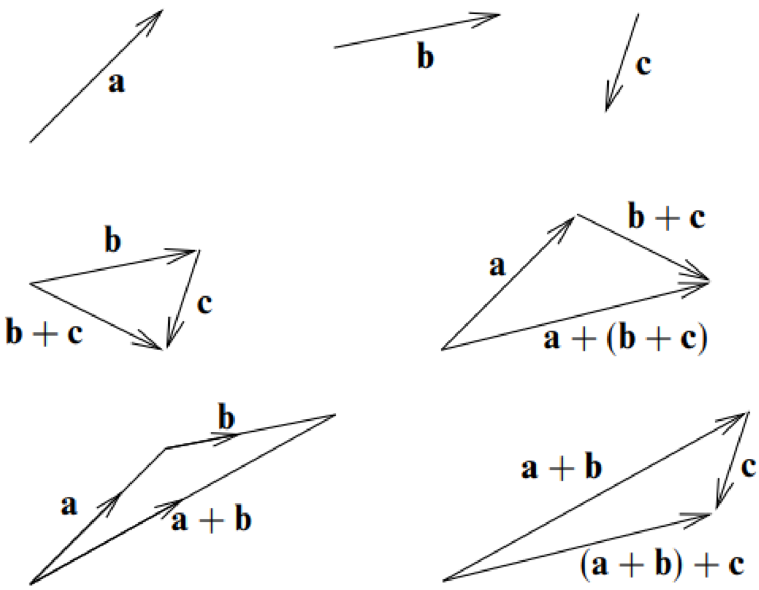

Figure 3: Addition of three vectors showing the associativity relation.

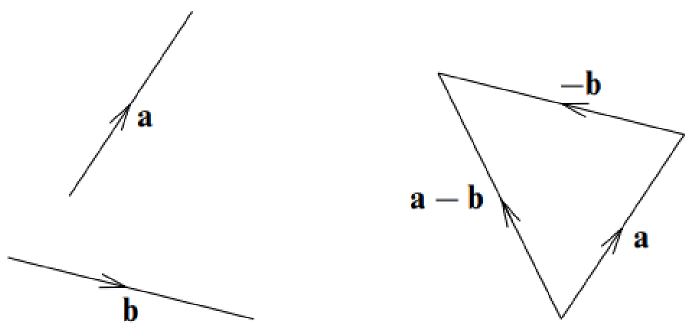

Figure 4: Subtraction of two vectors.

The subtraction of two equal vectors yields the zero vector, $\mathbf{0}$, which has zero magnitude and no associated direction.

2.4. Multiplication by a scalar

Multiplication of a vector by a scalar (not to be confused with the *scalar product*, to be discussed in Subsection 7.6.1) gives a vector in the same direction as the original but of a proportional magnitude. This can be seen in Figure. The scalar may be positive, negative or zero. It can also be complex in some applications. Clearly, when the scalar is negative we obtain a vector pointing in the opposite direction to the original vector.

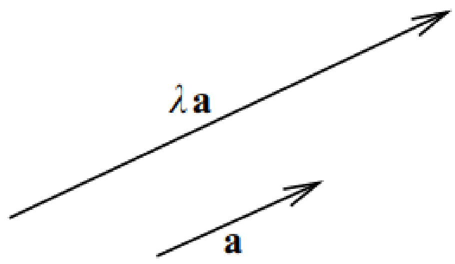

Figure 5: Scalar multiplication for $\lambda > 1$ stretches the vector

Multiplication by a scalar is associative, commutative and distributive over addition. These properties may be summarised for arbitrary vectors **a** and **b** and arbitrary scalars λ and μ by:

$$(\lambda\mu)\mathbf{a} = \lambda(\mu\mathbf{a}) = \mu(\lambda\mathbf{a}), \tag{7.3}$$

$$\lambda(\mathbf{a}+\mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}, \tag{7.4}$$

$$(\lambda+\mu)\mathbf{a} = \lambda\mathbf{a} + \mu\mathbf{a}. \tag{7.5}$$

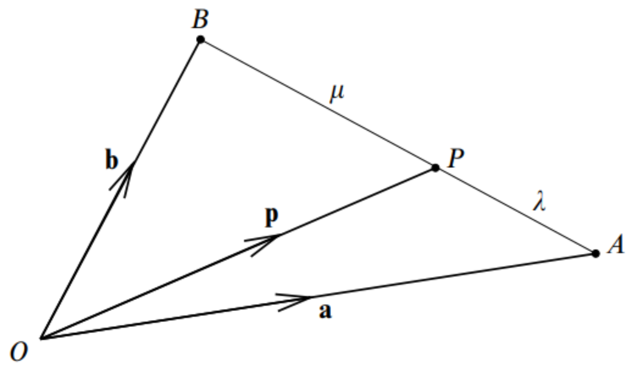

Figure 6: An illustration of the ratio theorem. Point P divides segment AB in the ratio $\lambda : \mu$

Example: A point P divides a line segment AB in the ratio $\lambda : \mu$ (see Figure 7). If the position vectors of the points A and B are $\vec{a}$ and $\vec{b}$, respectively, find the position vector of the point P.
Solution: The position vector $\vec{p}$ of point P dividing the segment AB in the ratio $\lambda : \mu$ is given by the section formula in vector form:

$$\vec{p} = \frac{\mu\vec{a} + \lambda\vec{b}}{\lambda + \mu}$$

This formula is valid for internal division.

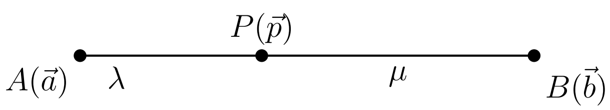

Figure 7: Point P dividing AB in the ratio $\lambda : \mu$

As is conventional for vector geometry problems, we denote the vector from the point A to the point B by $\vec{AB}$. If the position vectors of the points A and B, relative to some origin O, are $\vec{a}$ and $\vec{b}$, it should be clear that

$$\vec{AB} = \vec{b} - \vec{a}.$$

Now, from Figure we see that one possible way of reaching the point P from O is first to go from O to A, and then go along the line AB for a distance equal to the fraction $\frac{\lambda}{\lambda+\mu}$ of its total length. We may express this in terms of vectors as:

$$\begin{aligned} \vec{OP} = \vec{p} &= \vec{a} + \frac{\lambda}{\lambda+\mu}\vec{AB} \\ &= \vec{a} + \frac{\lambda}{\lambda+\mu}(\vec{b} - \vec{a}) \\ &= \left(1 - \frac{\lambda}{\lambda+\mu}\right)\vec{a} + \frac{\lambda}{\lambda+\mu}\vec{b} \\ &= \frac{\mu}{\lambda+\mu}\vec{a} + \frac{\lambda}{\lambda+\mu}\vec{b} \end{aligned} \tag{7.6}$$

which expresses the position vector of the point P in terms of those of A and B. We would, of course, obtain the same result by considering the path from O to B and then to P.

Example: The vertices of triangle ABC have position vectors $\vec{a}, \vec{b}$, and $\vec{c}$ relative to some origin O (see Figure). Find the position vector of the centroid G of the triangle.

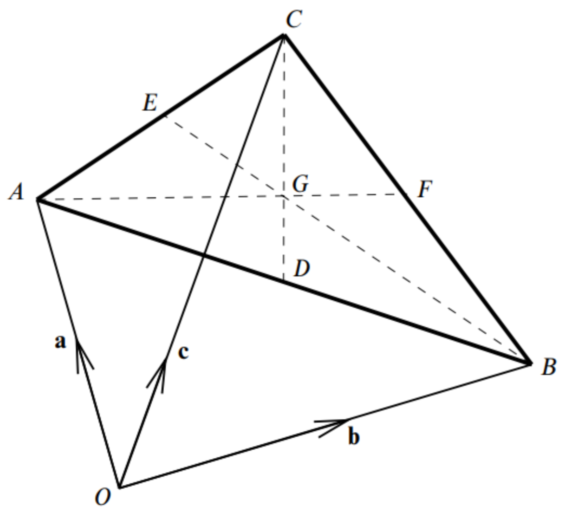

Figure 8: The centroid of a triangle. The triangle is defined by the points A, B and C that have position vectors a, b and c. The broken lines CD, BE, AF connect the vertices of the triangle to the mid-points of the opposite sides; these lines intersect at the centroid G of the triangle.

Solution: The centroid G of triangle ABC is the point of intersection of the medians, and it lies at the average of the position vectors of the three vertices. Therefore, the position vector $\vec{g}$ of the centroid G is given by:

$$\vec{g} = \frac{1}{3}(\vec{a} + \vec{b} + \vec{c})$$

Centroid of a Triangle: From Figure, the points D and E bisect the lines AB and AC, respectively. Thus, from the ratio theorem (7.6), with $\lambda = \mu = \frac{1}{2}$, the position vectors of D and E relative to the origin are:

$$\mathbf{d} = \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b},$$
$$\mathbf{e} = \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{c}.$$

Using the ratio theorem again, we may write the position vector of a general point on the line CD that divides the line in the ratio $\lambda : (1 - \lambda)$ as:

$$\begin{aligned} \mathbf{r} &= (1-\lambda)\mathbf{c} + \lambda\mathbf{d}, \\ &= (1-\lambda)\mathbf{c} + \frac{1}{2}\lambda(\mathbf{a} + \mathbf{b}). \end{aligned} \tag{7.7}$$

Similarly, the position vector of a general point on the line BE can be expressed as:

$$\begin{aligned} \mathbf{r} &= (1-\mu)\mathbf{b} + \mu\mathbf{e}, \\ &= (1-\mu)\mathbf{b} + \frac{1}{2}\mu(\mathbf{a} + \mathbf{c}). \end{aligned} \tag{7.8}$$

Thus, at the intersection of the lines CD and BE, we require from (7.7) and (7.8):

$$(1-\lambda)\mathbf{c} + \frac{1}{2}\lambda(\mathbf{a} + \mathbf{b}) = (1-\mu)\mathbf{b} + \frac{1}{2}\mu(\mathbf{a} + \mathbf{c}).$$

By equating the coefficients of the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, we find:

$$\begin{aligned} \text{Coefficient of } \mathbf{a}: \quad & \frac{1}{2}\lambda = \frac{1}{2}\mu, \\ \text{Coefficient of } \mathbf{b}: \quad & \frac{1}{2}\lambda = 1 - \mu, \\ \text{Coefficient of } \mathbf{c}: \quad & 1 - \lambda = \frac{1}{2}\mu. \end{aligned}$$

These equations are consistent and have the solution:

$$\lambda = \mu = \frac{2}{3}.$$

Substituting these values into either (7.7) or (7.8), we find that the position vector of the centroid G is given by:

$$\mathbf{g} = \frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c}).$$

Example: Two particles have velocities $\mathbf{v}_1 = \mathbf{i} + 3\mathbf{j} + 6\mathbf{k}$ and $\mathbf{v}_2 = \mathbf{i} - 2\mathbf{k}$, respectively. Find the velocity $\mathbf{u}$ of the second particle relative to the first.
Solution:

$$\begin{aligned} \mathbf{u} &= \mathbf{v}_2 - \mathbf{v}_1 \\ &= (\mathbf{i} - 2\mathbf{k}) - (\mathbf{i} + 3\mathbf{j} + 6\mathbf{k}) \\ &= \mathbf{i} - 2\mathbf{k} - \mathbf{i} - 3\mathbf{j} - 6\mathbf{k} \\ &= -3\mathbf{j} - 8\mathbf{k} \end{aligned}$$

2.5. Magnitude of a vector

The magnitude of the vector **a** is denoted by $|\mathbf{a}|$ or simply a. In terms of its components in three-dimensional Cartesian coordinates, the magnitude of **a** is given by

$$a \equiv |\mathbf{a}| = \sqrt{a_x^2 + a_y^2 + a_z^2}. \tag{7.13}$$

Hence, the magnitude of a vector is a measure of its length. Such an analogy is useful for displacement vectors, but magnitude is better described, for example, by 'strength' for vectors such as force or by 'speed' for velocity vectors. The scalar product of **a** and **b** is given by

$$\mathbf{a} \cdot \mathbf{b} = ab\cos\theta. \tag{7.14}$$

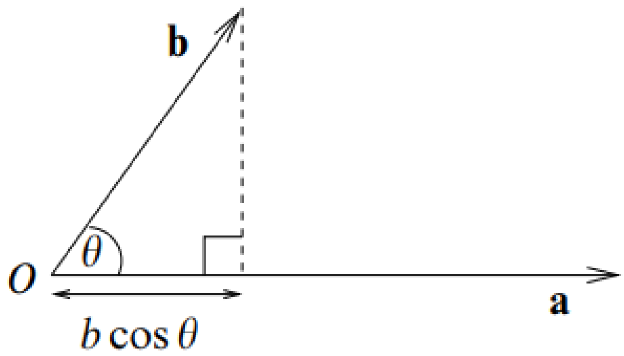

Figure 9: Projection of **b** onto **a** is $\mathbf{b}\cos\theta$
Scalar product: $\mathbf{a} \cdot \mathbf{b} = ab\cos\theta$

For instance, in the previous example, the speed of the second particle relative to the first is given by

$$u = |\mathbf{u}| = \sqrt{(-3)^2 + (-8)^2} = \sqrt{73}.$$

A vector whose magnitude equals unity is called a *unit vector*. The unit vector in the direction of **a** is usually denoted by $\hat{\mathbf{a}}$ and may be evaluated as

$$\hat{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|}. \tag{7.14}$$

The unit vector is a useful concept because a vector written as $\lambda\hat{\mathbf{a}}$ then has magnitude λ and direction $\hat{\mathbf{a}}$. Thus, magnitude and direction are explicitly separated.

2.6. Angle between the vectors

The angle between two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ can be found using the dot product formula:

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}|\,\mathbf{b}|}$$

or

$$\theta = \cos^{-1}\left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}\right)$$

Where:

- $\boldsymbol{a} \cdot \boldsymbol{b}$ is the dot product of the vectors,
- $|\boldsymbol{a}|$ and $|\boldsymbol{b}|$ are the magnitudes (lengths) of the vectors,
- θ is the angle between them.

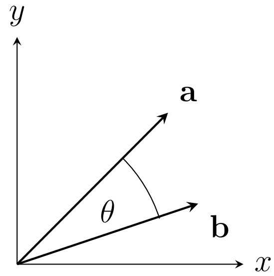

Figure 10: Angle θ between vectors $\mathbf{a}$ and $\mathbf{b}$.

This angle θ tells us how far apart the directions of the vectors are. It is widely used in geometry, physics, and computer science.

2.7. Multiplication of vectors

We have already considered multiplying a vector by a scalar. Now we consider the concept of multiplying one vector by another vector. It is not immediately obvious what the product of two vectors represents, and in fact, two products are commonly defined: the *scalar product* and the *vector product.* As their names imply, the scalar product of two vectors is just a number, whereas the vector product is itself a vector. Although neither the scalar nor the vector product is what we might normally think of as a product, their use is widespread, and numerous examples will be described elsewhere in this book.

2.8. Scalar Product

The scalar product (or dot product) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a} \cdot \mathbf{b}$ and is given by

$$\mathbf{a} \cdot \mathbf{b} \equiv |\mathbf{a}|\,|\mathbf{b}| \cos\theta, \quad 0 \leq \theta \leq \pi. \tag{7.15}$$

Here, θ is the angle between the two vectors, placed tail to tail or head to head. Thus, the value of the scalar product $\mathbf{a} \cdot \mathbf{b}$ equals the magnitude of $\mathbf{a}$ multiplied by the projection of $\mathbf{b}$ onto $\mathbf{a}$ (see Figure 7.8). From (7.15) we see that the scalar product has the particularly useful property that

$$\mathbf{a} \cdot \mathbf{b} = 0 \tag{7.16}$$

is a necessary and sufficient condition for $\mathbf{a}$ to be perpendicular to $\mathbf{b}$ (unless either of them is zero). It should be noted in particular that the Cartesian basis vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, being mutually orthogonal unit vectors, satisfy the equations

$$\mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} = 1, \tag{7.17}$$

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0. \tag{7.18}$$

Examples of scalar products arise naturally throughout physics and in particular in connection with energy. Perhaps the simplest is the work done, $\mathbf{F}\cdot\mathbf{r}$, in moving the point of application of a constant force $\mathbf{F}$ through a displacement $\mathbf{r}$; notice that, as expected, if the displacement is perpendicular to the direction of the force, then $\mathbf{F}\cdot\mathbf{r} = 0$ and no work is done.
A second simple example is afforded by the potential energy $-\mathbf{m}\cdot\mathbf{B}$ of a magnetic dipole, represented in strength and orientation by a vector $\mathbf{m}$, placed in an external magnetic field $\mathbf{B}$.
As the name implies, the scalar product has a magnitude but no direction. The scalar product is commutative and distributive over addition:

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}, \tag{7.19}$$

$$\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}. \tag{7.20}$$

Example: Four non-coplanar points A, B, C, D are positioned such that the line AD is perpendicular to BC and BD is perpendicular to AC. Show that CD is perpendicular to AB.
Solution: Four non-coplanar points A, B, C, and D are positioned such that

$$\overrightarrow{AD} \perp \overrightarrow{BC} \quad \text{and} \quad \overrightarrow{BD} \perp \overrightarrow{AC}.$$

And, we have show that

$$\overrightarrow{CD} \perp \overrightarrow{AB}.$$

Since $\overrightarrow{AD} \perp \overrightarrow{BC}$, we have

$$\overrightarrow{AD}\cdot\overrightarrow{BC} = 0,$$

and since $\overrightarrow{BD} \perp \overrightarrow{AC}$,

$$\overrightarrow{BD}\cdot\overrightarrow{AC} = 0.$$

Rewrite the vectors as differences of position vectors:

$$\overrightarrow{AD} = \mathbf{d}-\mathbf{a}, \quad \overrightarrow{BC} = \mathbf{c}-\mathbf{b},$$

$$\overrightarrow{BD} = \mathbf{d}-\mathbf{b}, \quad \overrightarrow{AC} = \mathbf{c}-\mathbf{a},$$

where $\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{d}$ are the position vectors of points A, B, C, and D respectively.

The given conditions become

$$(\mathbf{d}-\mathbf{a})\cdot(\mathbf{c}-\mathbf{b}) = 0, \quad (\mathbf{d}-\mathbf{b})\cdot(\mathbf{c}-\mathbf{a}) = 0.$$

Consider the dot product

$$(\mathbf{d}-\mathbf{c})\cdot(\mathbf{b}-\mathbf{a}) =?$$

Expand and rearrange:

$$(\mathbf{d}-\mathbf{c})\cdot(\mathbf{b}-\mathbf{a}) = \mathbf{d}\cdot\mathbf{b}-\mathbf{d}\cdot\mathbf{a}-\mathbf{c}\cdot\mathbf{b}+\mathbf{c}\cdot\mathbf{a}.$$

Similarly, expand the given orthogonality relations:

$$(\mathbf{d}-\mathbf{a})\cdot(\mathbf{c}-\mathbf{b}) = \mathbf{d}\cdot\mathbf{c}-\mathbf{d}\cdot\mathbf{b}-\mathbf{a}\cdot\mathbf{c}+\mathbf{a}\cdot\mathbf{b} = 0,$$

$$(\mathbf{d}-\mathbf{b})\cdot(\mathbf{c}-\mathbf{a}) = \mathbf{d}\cdot\mathbf{c}-\mathbf{d}\cdot\mathbf{a}-\mathbf{b}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{a} = 0.$$

Add these two equations:

$$2\mathbf{d}\cdot\mathbf{c}-(\mathbf{d}\cdot\mathbf{a}+\mathbf{d}\cdot\mathbf{b})-(\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c})+(\mathbf{a}\cdot\mathbf{b}+\mathbf{b}\cdot\mathbf{a}) = 0.$$

Noting that dot product is commutative, simplify:

$$2\mathbf{d}\cdot\mathbf{c}-\mathbf{d}\cdot(\mathbf{a}+\mathbf{b})-(\mathbf{a}+\mathbf{b})\cdot\mathbf{c}+2\mathbf{a}\cdot\mathbf{b} = 0.$$

Rearranging this gives

$$(\mathbf{d}-\mathbf{c})\cdot(\mathbf{b}-\mathbf{a}) = 0,$$

which means

$$\overrightarrow{CD} \perp \overrightarrow{AB}.$$

Example: Find the angle between the vectors

$$\mathbf{a} = \mathbf{i}+2\mathbf{j}+3\mathbf{k} \quad \text{and} \quad \mathbf{b} = 2\mathbf{i}+3\mathbf{j}+4\mathbf{k}.$$

Solution: The angle θ between two vectors $\mathbf{a}$ and $\mathbf{b}$ is given by the formula

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|\,|\mathbf{b}|}.$$

First, calculate the scalar (dot) product $\mathbf{a}\cdot\mathbf{b}$:

$$\mathbf{a}\cdot\mathbf{b} = (1)(2)+(2)(3)+(3)(4) = 2+6+12 = 20.$$

Next, find the magnitudes of $\mathbf{a}$ and $\mathbf{b}$:

$$|\mathbf{a}| = \sqrt{1^2+2^2+3^2} = \sqrt{1+4+9} = \sqrt{14},$$

$$|\mathbf{b}| = \sqrt{2^2+3^2+4^2} = \sqrt{4+9+16} = \sqrt{29}.$$

Therefore,

$$\cos\theta = \frac{20}{\sqrt{14}\sqrt{29}} = \frac{20}{\sqrt{406}}.$$

Finally, the angle θ is

$$\theta = \cos^{-1}\left(\frac{20}{\sqrt{406}}\right).$$

2.9. Vector Product

The *vector product* (or *cross product*) of two vectors $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a}\times\mathbf{b}$ and is defined to be a vector of magnitude $|\mathbf{a}||\mathbf{b}|\sin\theta$ in a direction perpendicular to both $\mathbf{a}$ and $\mathbf{b}$:

$$|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta.$$

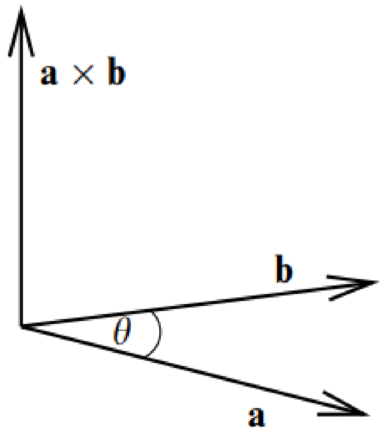

Figure 11: The vector product. The vectors a, b and ab form a right-handed set.

The direction is found by rotating $\mathbf{a}$ into $\mathbf{b}$ through the smallest possible angle. The sense of rotation is that of a right-handed screw that moves forward in the direction $\mathbf{a} \times \mathbf{b}$ (see Figure). Again, θ is the angle between the two vectors placed tail to tail or head to head. With this definition, $\mathbf{a}, \mathbf{b}$ and $\mathbf{a} \times \mathbf{b}$ form a right-handed set.

A more directly usable description of the relative directions in a vector product is provided by a right hand whose first two fingers and thumb are held to be as nearly mutually perpendicular as possible. If the first finger is pointed in the direction of the first vector and the second finger in the direction of the second vector, then the thumb gives the direction of the vector product. The vector product is distributive over addition, but anticommutative and non-associative:

$$(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \times \mathbf{c}) + (\mathbf{b} \times \mathbf{c}), \tag{7.23}$$

$$\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b}), \tag{7.24}$$

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c}). \tag{7.25}$$

From its definition, we see that the vector product has the very useful property that if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, then $\mathbf{a}$ is parallel or antiparallel to $\mathbf{b}$ (unless either of them is zero). We also note that

$$\mathbf{a} \times \mathbf{a} = \mathbf{0}. \tag{7.26}$$

Example: Show that if $\mathbf{a} = \mathbf{b} + \lambda\mathbf{c}$, for some scalar λ, then

$$\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c}.$$

Solution: Using the distributive property of the cross product,

$$\mathbf{a} \times \mathbf{c} = (\mathbf{b} + \lambda\mathbf{c}) \times \mathbf{c} = \mathbf{b} \times \mathbf{c} + (\lambda\mathbf{c}) \times \mathbf{c}.$$

Now, since $\mathbf{c} \times \mathbf{c} = \mathbf{0}$, and scalar multiplication factors out:

$$\lambda\mathbf{c} \times \mathbf{c} = \lambda(\mathbf{c} \times \mathbf{c}) = \lambda \cdot \mathbf{0} = \mathbf{0}.$$

So we have:

$$\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c} + \mathbf{0} = \mathbf{b} \times \mathbf{c}.$$

Finding area of a parallelogram: An example of the use of the vector product is that of finding the area, A, of a parallelogram with sides $\mathbf{a}$ and $\mathbf{b}$, using the formula

$$A = |\mathbf{a} \times \mathbf{b}|. \tag{7.28}$$

Since the basis vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are mutually perpendicular unit vectors, forming a right-handed set, their vector products are easily seen to be

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \mathbf{0}, \tag{7.29}$$

$$\mathbf{i} \times \mathbf{j} = -\mathbf{j} \times \mathbf{i} = \mathbf{k}, \tag{7.30}$$

$$\mathbf{j} \times \mathbf{k} = -\mathbf{k} \times \mathbf{j} = \mathbf{i}, \tag{7.31}$$

$$\mathbf{k} \times \mathbf{i} = -\mathbf{i} \times \mathbf{k} = \mathbf{j}. \tag{7.32}$$

Using these relations, it is straightforward to show that the vector product of two general vectors $\mathbf{a}$ and $\mathbf{b}$ is given in terms of their components with respect to the basis set $\mathbf{i}, \mathbf{j}, \mathbf{k}$, by

$$\mathbf{a} \times \mathbf{b} = (a_y b_z - a_z b_y)\,\mathbf{i} + (a_z b_x - a_x b_z)\,\mathbf{j} + (a_x b_y - a_y b_x)\,\mathbf{k}. \tag{7.33}$$

For the reader who is familiar with determinants, we record that this can also be written as

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}.$$

That the cross product $\mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$ can be verified in component form by forming its dot products with each of the two vectors and showing that it is zero in both cases.

Example: Find the area A of the parallelogram with sides

$$\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}, \quad \mathbf{b} = 4\mathbf{i} + 5\mathbf{j} + 6\mathbf{k}.$$

Solution: The area is given by the magnitude of the cross product:

$$A = |\mathbf{a} \times \mathbf{b}|.$$

Compute the cross product using the determinant:

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{vmatrix} = \mathbf{i}(2 \cdot 6 - 3 \cdot 5) - \mathbf{j}(1 \cdot 6 - 3 \cdot 4) + \mathbf{k}(1 \cdot 5 - 2 \cdot 4)$$

$$= \mathbf{i}(12 - 15) - \mathbf{j}(6 - 12) + \mathbf{k}(5 - 8) = -3\mathbf{i} + 6\mathbf{j} - 3\mathbf{k}.$$

Now compute the magnitude:

$$|\mathbf{a} \times \mathbf{b}| = \sqrt{(-3)^2 + 6^2 + (-3)^2} = \sqrt{9 + 36 + 9} = \sqrt{54} = 3\sqrt{6}.$$

Hence, the area of the parallelogram is $A = 3\sqrt{6}$.

2.10. Scalar Triple Product

Now that we have defined the scalar and vector products, we can extend our discussion to define products of three vectors. Again, there are two possibilities: the *scalar triple product* and the *vector triple product*.

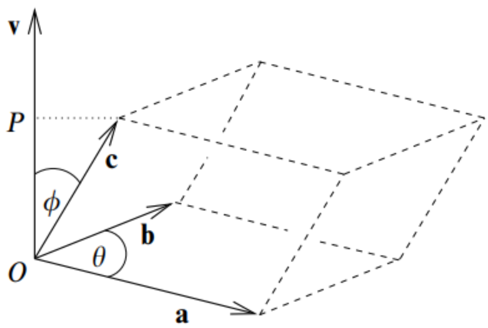

Figure 12: The scalar triple product gives the volume of a parallelepiped.

The scalar triple product is denoted by

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] \equiv \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}),$$

and, as its name suggests, it is just a number. It is most simply interpreted as the volume of a parallelepiped whose edges are given by $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ (see Figure 7.11).

The vector $\mathbf{v} = \mathbf{a} \times \mathbf{b}$ is perpendicular to the base of the solid and has magnitude $v = ab \sin\theta$, i.e., the area of the base. Further, $\mathbf{v} \cdot \mathbf{c} = vc \cos\varphi$. Thus, since $c\cos\varphi = OP$ is the vertical height of the parallelepiped, it is clear that

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \text{area of base} \times \text{height} = \text{volume}.$$

It follows that, if the vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are coplanar, then

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0.$$

Expressed in terms of the components of each vector with respect to the Cartesian basis set $\mathbf{i}, \mathbf{j}, \mathbf{k}$, the scalar triple product is

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = a_x(b_y c_z - b_z c_y) + a_y(b_z c_x - b_x c_z) + a_z(b_x c_y - b_y c_x). \tag{7.34}$$

This can also be written as a determinant:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.$$

By writing the vectors in component form, it can be shown that

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c},$$

so that the dot and cross symbols can be interchanged without changing the result. More generally, the scalar triple product is unchanged under cyclic permutation of the vectors

$\mathbf{a}, \mathbf{b}, \mathbf{c}$. Other permutations give the negative of the original scalar triple product. These results can be summarised by:

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] = -[\mathbf{a}, \mathbf{c}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}, \mathbf{c}] = -[\mathbf{c}, \mathbf{b}, \mathbf{a}]. \tag{7.35}$$

Example: Find the volume V of the parallelepiped with sides

$$\mathbf{a} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}, \quad \mathbf{b} = 4\mathbf{i} + 5\mathbf{j} + 6\mathbf{k}, \quad \mathbf{c} = 7\mathbf{i} + 8\mathbf{j} + 10\mathbf{k}.$$

Solution: The volume is given by the absolute value of the scalar triple product:

$$V = |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})| .$$

First compute the cross product $\mathbf{b} \times \mathbf{c}$:

$$\mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = \mathbf{i}(5 \cdot 10 - 6 \cdot 8) - \mathbf{j}(4 \cdot 10 - 6 \cdot 7) + \mathbf{k}(4 \cdot 8 - 5 \cdot 7)$$

$$= \mathbf{i}(50 - 48) - \mathbf{j}(40 - 42) + \mathbf{k}(32 - 35) = 2\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}.$$

Next compute the dot product with $\mathbf{a}$:

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (1)(2) + (2)(2) + (3)(-3) = 2 + 4 - 9 = -3.$$

Taking the absolute value, the volume is

$$V = | - 3| = 3.$$

hence, the volume of the parallelepiped is $V = 3$.

Another useful formula involving both the scalar and vector products is Lagranges identity

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) \equiv (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}). \tag{7.36}$$

2.11. Vector Triple Product

By the vector triple product of three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ we mean the vector $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$. Clearly, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ is perpendicular to $\mathbf{a}$ and lies in the plane of $\mathbf{b}$ and $\mathbf{c}$ and so can be expressed in terms of them (see (7.37) below). We also note, from (7.25), that the vector triple product is not associative, i.e.

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}.$$

Two useful formulae involving the vector triple product are

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}, \tag{7.37}$$

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}, \tag{7.38}$$

which may be derived by writing each vector in component form.
It can also be shown that for any three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}.$$

Exercise: Which of the following statements about general vectors $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are true?

1. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = (\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c}$.
2. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.
3. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.
4. $\mathbf{d} = \lambda\mathbf{a} + \mu\mathbf{b}$ implies $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d} = 0$.
5. $\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c}$ implies $\mathbf{c} \cdot \mathbf{a} - \mathbf{c} \cdot \mathbf{b} = |\mathbf{c}|\,|\mathbf{a} - \mathbf{b}|$.
6. $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{b}) = \mathbf{b}\big[\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})\big]$.

Solutions:

1. $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = (\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c}$.

   Solution: True. Since $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$,

   $$(\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c} = -(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = -\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}),$$

   but dot product is commutative, and $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ is a scalar, so

   $$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = -\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}),$$

   which implies this is only true if $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = 0$. Actually, this shows

   $$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = -(\mathbf{b} \times \mathbf{a}) \cdot \mathbf{c},$$

   so the statement as given is false unless zero. So the statement is **False**.

2. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

   Solution: False. The vector triple product is not associative:

   $$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}.$$

3. $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$.

   Solution: True. This is the standard vector triple product identity.

4. $\mathbf{d} = \lambda\mathbf{a} + \mu\mathbf{b}$ implies $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d} = 0$.

   Solution: True. Since $\mathbf{a} \times \mathbf{b}$ is perpendicular to any linear combination of $\mathbf{a}$ and $\mathbf{b}$, the dot product is zero.

5. $\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c}$ implies $\mathbf{c} \cdot \mathbf{a} - \mathbf{c} \cdot \mathbf{b} = |\mathbf{c}|\,|\mathbf{a} - \mathbf{b}|$.

   Solution: False. From $\mathbf{a} \times \mathbf{c} = \mathbf{b} \times \mathbf{c}$ it follows $(\mathbf{a} - \mathbf{b}) \times \mathbf{c} = \mathbf{0}$, so $\mathbf{a} - \mathbf{b}$ is parallel to $\mathbf{c}$. This implies $\mathbf{a} - \mathbf{b} = \lambda \mathbf{c}$ for some scalar λ, so

$$\mathbf{c} \cdot (\mathbf{a} - \mathbf{b}) = \lambda |\mathbf{c}|^2,$$

   but this is not equal to $|\mathbf{c}|\,|\mathbf{a} - \mathbf{b}|$ in general, unless $\mathbf{a} - \mathbf{b}$ is in the same direction as $\mathbf{c}$ with positive scalar λ. So the equality as stated is generally false.

6. $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{b}) = \mathbf{b}\big[\mathbf{b} \cdot (\mathbf{c} \times \mathbf{a})\big]$.

   Solution: False. This expression is more complicated, and does not simplify to the right side in general. It can be verified by expanding using vector identities, and it differs by additional terms.

Example: A unit cell of diamond is a cube of side A, with carbon atoms at: each corner,the center of each face, and in addition, at positions displaced by $\frac{1}{4}A(\hat{\imath} + \hat{\jmath} + \hat{k})$ from each of those already mentioned; where $\hat{\imath}, \hat{\jmath}, \hat{k}$ are unit vectors along the cube axes. One corner of the cube is taken as the origin of coordinates. What are the vectors joining the atom at $\frac{1}{4}A(\hat{\imath} + \hat{\jmath} + \hat{k})$ to its four nearest neighbours? Determine the angle between the carbon bonds in diamond.
Solution: Let the atom at position

$$\vec{r}_0 = \frac{A}{4}(\hat{\imath} + \hat{\jmath} + \hat{k})$$

We want to find the relative position vectors $\vec{r}_i - \vec{r}_0$ for the four nearest neighbors. The diamond structure consists of two interpenetrating face-centered cubic (FCC) lattices, displaced by $\frac{1}{4}A(\hat{\imath} + \hat{\jmath} + \hat{k})$. Each atom has 4 nearest neighbors arranged tetrahedrally. The four nearest neighbor atoms to $\vec{r}_0$ are located at:

$$\begin{aligned}
\vec{r}_1 &= \frac{A}{4}(\hat{\imath} + \hat{\jmath} - \hat{k}) \\
\vec{r}_2 &= \frac{A}{4}(\hat{\imath} - \hat{\jmath} + \hat{k}) \\
\vec{r}_3 &= \frac{A}{4}(-\hat{\imath} + \hat{\jmath} + \hat{k}) \\
\vec{r}_4 &= \frac{A}{4}(3\hat{\imath} + 3\hat{\jmath} + 3\hat{k}) = \vec{r}_0 + \frac{A}{2}(\hat{\imath} + \hat{\jmath} + \hat{k})
\end{aligned}$$

The vectors from $\vec{r}_0$ to the nearest neighbors are:

$$\begin{aligned}
\vec{v}_1 &= \vec{r}_1 - \vec{r}_0 = \frac{A}{4}(0, 0, -2) = -\frac{A}{2}\hat{k} \\
\vec{v}_2 &= \vec{r}_2 - \vec{r}_0 = \frac{A}{4}(0, -2, 0) = -\frac{A}{2}\hat{\jmath} \\
\vec{v}_3 &= \vec{r}_3 - \vec{r}_0 = \frac{A}{4}(-2, 0, 0) = -\frac{A}{2}\hat{\imath} \\
\vec{v}_4 &= \vec{r}_4 - \vec{r}_0 = \frac{A}{2}(\hat{\imath} + \hat{\jmath} + \hat{k})
\end{aligned}$$

To determine the bond angle between any two of these vectors (say, between $\vec{v}_1$ and $\vec{v}_2$), use the dot product:

$$\cos\theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|\|\vec{v}_2\|}$$

Since $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ point toward the vertices of a tetrahedron, the angle between any two bond directions is:

$$\cos\theta = -\frac{1}{3} \quad \Rightarrow \quad \theta = \cos^{-1}(-1/3) \approx 109.47^\circ$$

Answer

- The four vectors joining the atom at $\frac{1}{4}A(\hat{\imath} + \hat{\jmath} + \hat{k})$ to its nearest neighbors are:

$$\pm\frac{A}{4}(\pm1, \pm1, \pm1)$$

  (with combinations of signs such that each vector differs from the original position by flipping the sign of one coordinate).

- The angle between the carbon bonds in diamond is:

$$\boxed{109.47^\circ}$$

Example: Find the angle between the position vectors to the points $(3, -4, 0)$ and $(-2, 1, 0)$, and find the direction cosines of a vector perpendicular to both.

Solution: Let

$$\vec{a} = \langle 3, -4, 0\rangle, \quad \vec{b} = \langle -2, 1, 0\rangle$$

1. Angle Between the Vectors:
The angle θ between two vectors is given by:

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{|\vec{a}||\vec{b}|}$$

Compute the dot product:

$$\vec{a} \cdot \vec{b} = (3)(-2) + (-4)(1) + (0)(0) = -6 - 4 = -10$$

Magnitudes:

$$|\vec{a}| = \sqrt{3^2 + (-4)^2 + 0^2} = \sqrt{9 + 16} = \sqrt{25} = 5$$
$$|\vec{b}| = \sqrt{(-2)^2 + 1^2 + 0^2} = \sqrt{4 + 1} = \sqrt{5}$$

Now,

$$\cos\theta = \frac{-10}{5\sqrt{5}} = \frac{-2}{\sqrt{5}} \quad \Rightarrow \quad \theta = \cos^{-1}\left(\frac{-2}{\sqrt{5}}\right)$$

2. A Vector Perpendicular to Both:
A vector perpendicular to both $\vec{a}$ and $\vec{b}$ is given by the cross product:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ 3 & -4 & 0 \\ -2 & 1 & 0 \end{vmatrix} = \hat{k}(3 \cdot 1 - (-4)(-2)) = \hat{k}(3 - 8) = -5\hat{k}$$

So, the perpendicular vector is $\vec{n} = \langle 0, 0, -5 \rangle$

3. Direction Cosines:

The direction cosines of a vector $\vec{n} = \langle a, b, c \rangle$ are:

$$\cos\alpha = \frac{a}{|\vec{n}|}, \quad \cos\beta = \frac{b}{|\vec{n}|}, \quad \cos\gamma = \frac{c}{|\vec{n}|}$$

For $\vec{n} = \langle 0, 0, -5 \rangle$, we get:

$$|\vec{n}| = \sqrt{0^2 + 0^2 + (-5)^2} = 5$$

So the direction cosines are:

$$\cos\alpha = 0, \quad \cos\beta = 0, \quad \cos\gamma = \frac{-5}{5} = -1$$

Hence,

- The angle θ between the vectors is:

$$\theta = \cos^{-1}\left(\frac{-2}{\sqrt{5}}\right)$$

- A vector perpendicular to both is $\langle 0, 0, -5 \rangle$
- The direction cosines are:

$$\boxed{(0,\ 0,\ -1)}$$

⋆ ⋆ ⋆

Chapter 3. Vector Calculus

Vector calculus is a branch of mathematics that combines vector algebra and calculus to study vector fieldsquantities that have both magnitude and direction, and can vary from one point in space to another. It involves operations such as the gradient, which shows the direction and rate of the steepest increase of a scalar field; the divergence, which measures how much a field spreads out from a point; and the curl, which shows the rotation or circulation of a field. These concepts are essential for understanding how vector quantities change in space.

Vector calculus is widely used in real-world applications. In physics, it helps describe electric and magnetic fields, fluid flow, and heat transfer. In engineering, it is used in structural analysis, fluid dynamics, and electromagnetics. In computer graphics and robotics, vector calculus supports motion planning, lighting, and 3D rendering. In weather forecasting, it models wind velocity fields and temperature distribution. In medical imaging, such as MRI and CT scans, vector fields help reconstruct internal body structures. In chemistry, vector calculus plays a crucial role in studying molecular forces and potential energy fields. It is used to analyze the behavior of molecules under various physical conditions, understand electrostatic potential distributions, and describe reaction pathways. In quantum chemistry, it supports the formulation of wave functions and operators in three-dimensional space. Overall, vector calculus is a vital mathematical tool used across science and technology to describe systems involving continuous change in multiple directions.

3.1. Differentiation of vectors

Let us consider a vector $\vec{a}$ that is a function of a scalar variable u. By this we mean that with each value of u, we associate a vector $\vec{a}(u)$.
For example, in Cartesian coordinates,

$$\vec{a}(u) = a_x(u)\,\hat{\imath} + a_y(u)\,\hat{\jmath} + a_z(u)\,\hat{k},$$

where $a_x(u), a_y(u)$, and $a_z(u)$ are scalar functions of u, and are the components of the vector $\vec{a}(u)$ in the x-, y- and z-directions respectively. We note that if $\vec{a}(u)$ is continuous at some point $u = u_0$, then this implies that each of the Cartesian components $a_x(u), a_y(u)$, and $a_z(u)$ is also continuous there.

Let us consider the derivative of the vector function $\vec{a}(u)$ with respect to u. The derivative of a vector function is defined in a similar manner to the ordinary derivative of a scalar function $f(x)$ (as given in Chapter 2).
The small change in the vector $\vec{a}(u)$ resulting from a small change Δu in the value of u is given by:

$$\Delta\vec{a} = \vec{a}(u + \Delta u) - \vec{a}(u) \quad \text{(see Figure)}.$$

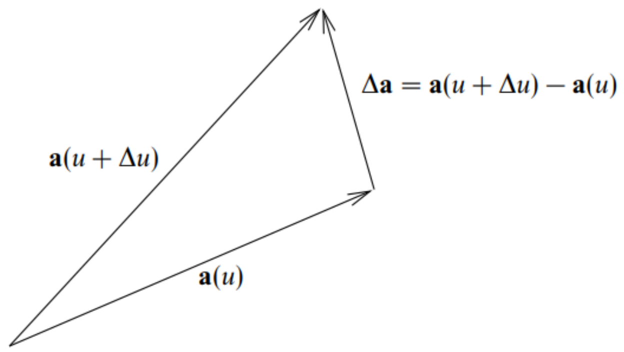

Figure 13: A small change in a vector a(u) resulting from a small change in u

The derivative of $\vec{a}(u)$ with respect to u is defined to be

$$\frac{d\vec{a}}{du} = \lim_{\Delta u \to 0} \frac{\vec{a}(u + \Delta u) - \vec{a}(u)}{\Delta u} \tag{10.1}$$

Assuming that the limit exists, the vector $\vec{a}(u)$ is said to be *differentiable* at that point. Note that $\frac{d\vec{a}}{du}$ is also a vector, which is not, in general, parallel to $\vec{a}(u)$.
In Cartesian coordinates, if

$$\vec{a}(u) = a_x(u)\,\hat{\imath} + a_y(u)\,\hat{\jmath} + a_z(u)\,\hat{k},$$

then the derivative of the vector is given by:

$$\frac{d\vec{a}}{du} = \frac{da_x}{du}\,\hat{\imath} + \frac{da_y}{du}\,\hat{\jmath} + \frac{da_z}{du}\,\hat{k}.$$

Perhaps the simplest application of the above is in finding the velocity and acceleration of a particle in classical mechanics.
If the time-dependent position vector of the particle with respect to the origin in Cartesian coordinates is given by:

$$\vec{r}(t) = x(t)\,\hat{\imath} + y(t)\,\hat{\jmath} + z(t)\,\hat{k},$$

then the velocity of the particle is the time derivative:

$$\vec{v}(t) = \frac{d\vec{r}}{dt} = \frac{dx}{dt}\,\hat{\imath} + \frac{dy}{dt}\,\hat{\jmath} + \frac{dz}{dt}\,\hat{k}.$$

The direction of the velocity vector is along the tangent to the path $\vec{r}(t)$ at the instantaneous position of the particle, and its magnitude $|\vec{v}(t)|$ is equal to the speed of the particle.

The acceleration of the particle is the derivative of velocity:

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d^2x}{dt^2}\,\hat{\imath} + \frac{d^2y}{dt^2}\,\hat{\jmath} + \frac{d^2z}{dt^2}\,\hat{k}.$$

Example: The position vector of a particle at time t, in Cartesian coordinates, is given by:

$$\vec{r}(t) = 2t^2\,\hat{\imath} + (3t - 2)\,\hat{\jmath} + (3t^2 - 1)\,\hat{k}$$

.Find the speed of the particle at t = 1 and the component of its acceleration in the direction s = i + 2j + k.

Solution:

1. Speed at $t = 1$:

First, compute the velocity vector $\vec{v}(t) = \frac{d\vec{r}}{dt}$:

$$\vec{v}(t) = \frac{d}{dt}(2t^2)\,\hat{\imath} + \frac{d}{dt}(3t-2)\,\hat{\jmath} + \frac{d}{dt}(3t^2-1)\,\hat{k} = 4t\,\hat{\imath} + 3\,\hat{\jmath} + 6t\,\hat{k}$$

At $t = 1$, the velocity is:

$$\vec{v}(1) = 4\,\hat{\imath} + 3\,\hat{\jmath} + 6\,\hat{k}$$

The speed is the magnitude of $\vec{v}(1)$:

$$|\vec{v}(1)| = \sqrt{4^2 + 3^2 + 6^2} = \sqrt{16 + 9 + 36} = \sqrt{61}$$

2. Component of Acceleration in the Direction of $\vec{s} = \hat{\imath} + 2\hat{\jmath} + \hat{k}$:

First, compute the acceleration vector:

$$\vec{a}(t) = \frac{d\vec{v}}{dt} = \frac{d}{dt}(4t)\,\hat{\imath} + \frac{d}{dt}(3)\,\hat{\jmath} + \frac{d}{dt}(6t)\,\hat{k} = 4\,\hat{\imath} + 0\,\hat{\jmath} + 6\,\hat{k}$$

To find the component of $\vec{a}(t)$ in the direction of $\vec{s}$, we use the projection formula:

$$\text{Component of } \vec{a} \text{ along } \vec{s} = \frac{\vec{a}\cdot\vec{s}}{|\vec{s}|}$$

Compute the dot product:

$$\vec{a}\cdot\vec{s} = (4)(1) + (0)(2) + (6)(1) = 4 + 0 + 6 = 10$$

Magnitude of $\vec{s}$:

$$|\vec{s}| = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{1 + 4 + 1} = \sqrt{6}$$

So, the component is:

$$\frac{10}{\sqrt{6}}$$

Hence,

- Speed at $t = 1$: $\boxed{\sqrt{61}}$
- Component of acceleration along $\vec{s} = \hat{\imath} + 2\hat{\jmath} + \hat{k}$: $\boxed{\frac{10}{\sqrt{6}}}$

Differentiation of composite vector expressions: In composite vector expressions, each of the vectors or scalars involved may be a function of some scalar variable u, as we have seen. The derivatives of such expressions are easily found using the definition (10.1) and the rules of ordinary differential calculus. They may be summarised by the following, in which we assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are differentiable vector functions of a scalar u and that ϕ is a differentiable scalar function of u:

$$\frac{d}{du}(\phi \boldsymbol{a}) = \phi \frac{d\boldsymbol{a}}{du} + \frac{d\phi}{du}\boldsymbol{a}, \tag{10.4}$$

$$\frac{d}{du}(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot \frac{d\boldsymbol{b}}{du} + \frac{d\boldsymbol{a}}{du} \cdot \boldsymbol{b}, \tag{10.5}$$

$$\frac{d}{du}(\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a} \times \frac{d\boldsymbol{b}}{du} + \frac{d\boldsymbol{a}}{du} \times \boldsymbol{b}. \tag{10.6}$$

The order of the factors in the terms on the RHS of (10.6) is, of course, just as important as it is in the original vector product

Example: A particle of mass m with position vector $\boldsymbol{r}$ relative to some origin O experiences a force $\boldsymbol{F}$, which produces a torque (moment) $\boldsymbol{T} = \boldsymbol{r} \times \boldsymbol{F}$ about O. The angular momentum of the particle about O is given by $\boldsymbol{L} = \boldsymbol{r} \times m\boldsymbol{v}$, where $\boldsymbol{v}$ is the particles velocity. Show that the rate of change of angular momentum is equal to the applied torque.

Solution: We are given the angular momentum of the particle as:

$$\boldsymbol{L} = \boldsymbol{r} \times m\boldsymbol{v}.$$

To find the rate of change of angular momentum, we differentiate $\boldsymbol{L}$ with respect to time:

$$\frac{d\boldsymbol{L}}{dt} = \frac{d}{dt}(\boldsymbol{r} \times m\boldsymbol{v}).$$

Using the product rule for derivatives of cross products:

$$\frac{d\boldsymbol{L}}{dt} = \frac{d\boldsymbol{r}}{dt} \times m\boldsymbol{v} + \boldsymbol{r} \times \frac{d}{dt}(m\boldsymbol{v}).$$

Since $\frac{d\boldsymbol{r}}{dt} = \boldsymbol{v}$ and m is constant:

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{v} \times m\boldsymbol{v} + \boldsymbol{r} \times m\frac{d\boldsymbol{v}}{dt}.$$

Note that $\boldsymbol{v} \times \boldsymbol{v} = \boldsymbol{0}$, so the first term vanishes:

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{r} \times m\frac{d\boldsymbol{v}}{dt}.$$

Using Newton's second law, $m\frac{d\boldsymbol{v}}{dt} = \boldsymbol{F}$:

$$\frac{d\boldsymbol{L}}{dt} = \boldsymbol{r} \times \boldsymbol{F} = \boldsymbol{T}.$$

Conclusion:

$$\boxed{\frac{d\boldsymbol{L}}{dt} = \boldsymbol{T}}$$

Thus, the rate of change of angular momentum is equal to the applied torque.

If a vector $\boldsymbol{a}$ is a function of a scalar variable s that is itself a function of u, so that $s = s(u)$, then the chain rule (see subsection 2.1.3) gives:

$$\frac{d\boldsymbol{a}(s)}{du} = \frac{ds}{du} \cdot \frac{d\boldsymbol{a}}{ds}. \tag{10.7}$$

The derivatives of more complicated vector expressions may be found by repeated application of the above equations.

One further useful result can be derived by considering the derivative:

$$\frac{d}{du}(\boldsymbol{a} \cdot \boldsymbol{a}) = 2\boldsymbol{a} \cdot \frac{d\boldsymbol{a}}{du};$$

since $\boldsymbol{a} \cdot \boldsymbol{a} = a^2$, where $a = |\boldsymbol{a}|$, we see that

$$\boldsymbol{a} \cdot \frac{d\boldsymbol{a}}{du} = 0 \quad \text{if } \boldsymbol{a} \text{ is constant.} \tag{10.8}$$

In other words, if a vector $\boldsymbol{a}(u)$ has a constant magnitude as u varies, then it is perpendicular to the vector $\frac{d\boldsymbol{a}}{du}$.

Differential of a vector: As a final note on the differentiation of vectors, we can also define the *differential* of a vector, in a similar way to that of a scalar in ordinary differential calculus. In the definition of the vector derivative (10.1), we used the notion of a small change $\Delta\boldsymbol{a}$ in a vector $\boldsymbol{a}(u)$ resulting from a small change Δu in its argument. In the limit $\Delta u \to 0$, the change in $\boldsymbol{a}$ becomes infinitesimally small, and we denote it by the differential $d\boldsymbol{a}$. From (10.1) we see that the differential is given by

$$d\boldsymbol{a} = \frac{d\boldsymbol{a}}{du}\, du. \tag{10.9}$$

Note that the differential of a vector is also a vector. As an example, the infinitesimal change in the position vector of a particle in an infinitesimal time dt is

$$d\boldsymbol{r} = \frac{d\boldsymbol{r}}{dt}\, dt = \boldsymbol{v}\, dt,$$

where $\boldsymbol{v}$ is the particles velocity.

3.2. Vector operators:

Certain differential operations may be performed on scalar and vector fields and have wide-ranging applications in the physical sciences. The most important operations are those of finding the *gradient* of a scalar field and the *divergence* and *curl* of a vector field. It is usual to define these operators from a strictly mathematical point of view, as we do below. In the following chapter, however, we will discuss their geometrical definitions, which rely on the concept of integrating vector quantities along lines and over surfaces.

Central to all these differential operations is the vector operator ∇, which is called *del* (or sometimes *nabla*) and in Cartesian coordinates is defined by

$$\nabla \equiv \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z}. \tag{10.25}$$

3.2.1. Gradient of a scalar field

The gradient of a scalar field $\phi(x, y, z)$ is defined by

$$\operatorname{grad}\phi = \nabla\phi = \boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}. \tag{10.26}$$

Clearly, $\nabla\phi$ is a vector field whose x-, y- and z-components are the first partial derivatives of $\phi(x, y, z)$ with respect to x, y, and z respectively. Also note that the vector field $\nabla\phi$ should not be confused with the vector operator $\phi\nabla$, which has components:

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

Example: Find the gradient of the scalar field

$$\phi = xy^2z^3.$$

Solution: The gradient of a scalar field $\phi(x, y, z)$ is given by

$$\nabla\phi = \boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}.$$

Calculate each partial derivative:

$$\begin{aligned}\frac{\partial\phi}{\partial x} &= y^2z^3,\\ \frac{\partial\phi}{\partial y} &= 2xyz^3,\\ \frac{\partial\phi}{\partial z} &= 3xy^2z^2.\end{aligned}$$

Therefore, the gradient is

$$\nabla\phi = \boldsymbol{i}y^2z^3 + \boldsymbol{j}2xyz^3 + \boldsymbol{k}3xy^2z^2.$$

The gradient of a scalar field ϕ has some interesting geometrical properties. Let us first consider the problem of calculating the rate of change of ϕ in some particular direction. For an infinitesimal vector displacement $d\boldsymbol{r}$, forming its scalar product with $\nabla\phi$ we obtain

$$\begin{aligned}\nabla\phi\cdot d\boldsymbol{r} &= \left(\boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}\right)\cdot(\boldsymbol{i}dx + \boldsymbol{j}dy + \boldsymbol{k}dz)\\ &= \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz\\ &= d\phi, \end{aligned} \tag{10.27}$$

which is the infinitesimal change in ϕ in going from position $\boldsymbol{r}$ to $\boldsymbol{r} + d\boldsymbol{r}$.

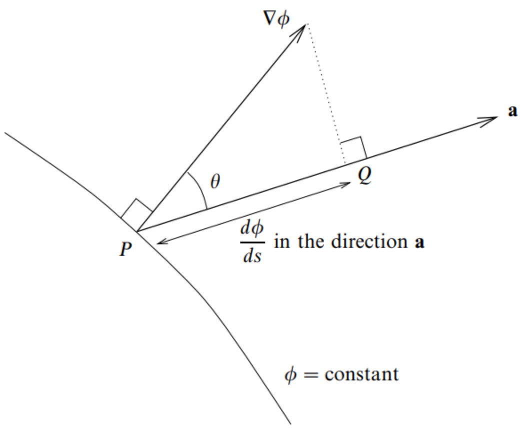

Figure 14: Geometrical interpretation of $\nabla\phi$: PQ is the projection of $\nabla\phi$ in the direction of $\mathbf{a}$, so $\dfrac{d\phi}{ds} = |\nabla\phi|\cos\theta$

In particular, if $\boldsymbol{r}$ depends on some parameter u such that $\boldsymbol{r}(u)$ defines a space curve then the total derivative of ϕ with respect to u along the curve is simply

$$\frac{d\phi}{du} = \nabla\phi \cdot \frac{d\boldsymbol{r}}{du}. \tag{10.28}$$

In the particular case where the parameter u is the arc length s along the curve, the total derivative of ϕ with respect to s along the curve is given by

$$\frac{d\phi}{ds} = \nabla\phi \cdot \hat{\boldsymbol{t}}, \tag{10.29}$$

where $\hat{\boldsymbol{t}}$ is the unit tangent to the curve at the given point. In general, the rate of change of ϕ with respect to the distance s in a particular direction $\boldsymbol{a}$ is given by

$$\frac{d\phi}{ds} = \nabla\phi \cdot \hat{\boldsymbol{a}}, \tag{10.30}$$

and is called the directional derivative. Since $\hat{\boldsymbol{a}}$ is a unit vector we have

$$\frac{d\phi}{ds} = |\nabla\phi|\cos\theta,$$

where θ is the angle between $\hat{\boldsymbol{a}}$ and $\nabla\phi$ as shown in figure 10.5. Clearly, $\nabla\phi$ lies in the direction of the fastest increase in ϕ, and $|\nabla\phi|$ is the largest possible value of $d\phi/ds$. Similarly, the largest rate of decrease of ϕ is $d\phi/ds = -|\nabla\phi|$ in the direction of $-\nabla\phi$.

Example: For the function

$$\phi = x^2 y + yz$$

at the point $(1, 2, -1)$, find its rate of change with distance in the direction

$$\boldsymbol{a} = \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}.$$

At this same point, what is the greatest possible rate of change with distance and in which direction does it occur?

Solution: The rate of change of ϕ in the direction of a unit vector $\hat{\boldsymbol{a}}$ is given by the directional derivative:

$$\frac{d\phi}{ds} = \nabla\phi \cdot \hat{\boldsymbol{a}},$$

where

$$\nabla\phi = \boldsymbol{i}\frac{\partial\phi}{\partial x} + \boldsymbol{j}\frac{\partial\phi}{\partial y} + \boldsymbol{k}\frac{\partial\phi}{\partial z}.$$

Calculate the partial derivatives:

$$\frac{\partial\phi}{\partial x} = 2xy, \quad \frac{\partial\phi}{\partial y} = x^2 + z, \quad \frac{\partial\phi}{\partial z} = y.$$

At the point $(1, 2, -1)$, these evaluate to:

$$\left.\frac{\partial\phi}{\partial x}\right|_{(1,2,-1)} = 2 \times 1 \times 2 = 4,$$

$$\left.\frac{\partial\phi}{\partial y}\right|_{(1,2,-1)} = 1^2 + (-1) = 0,$$

$$\left.\frac{\partial\phi}{\partial z}\right|_{(1,2,-1)} = 2.$$

Therefore,

$$\nabla\phi(1, 2, -1) = 4\boldsymbol{i} + 0\boldsymbol{j} + 2\boldsymbol{k} = 4\boldsymbol{i} + 2\boldsymbol{k}.$$

Next, find the unit vector in the direction of $\boldsymbol{a} = \boldsymbol{i} + 2\boldsymbol{j} + 3\boldsymbol{k}$:

$$|\boldsymbol{a}| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14},$$

$$\hat{\boldsymbol{a}} = \frac{1}{\sqrt{14}}\boldsymbol{i} + \frac{2}{\sqrt{14}}\boldsymbol{j} + \frac{3}{\sqrt{14}}\boldsymbol{k}.$$

The rate of change of ϕ in the direction $\boldsymbol{a}$ is

$$\frac{d\phi}{ds} = \nabla\phi \cdot \hat{\boldsymbol{a}} = (4\boldsymbol{i} + 2\boldsymbol{k}) \cdot \left(\frac{1}{\sqrt{14}}\boldsymbol{i} + \frac{2}{\sqrt{14}}\boldsymbol{j} + \frac{3}{\sqrt{14}}\boldsymbol{k}\right) = \frac{4}{\sqrt{14}} + \frac{6}{\sqrt{14}} = \frac{10}{\sqrt{14}}.$$

The greatest possible rate of change of ϕ at this point is the magnitude of the gradient vector:

$$\max\frac{d\phi}{ds} = |\nabla\phi| = \sqrt{4^2 + 0^2 + 2^2} = \sqrt{20} = 2\sqrt{5}.$$

This maximum rate of change occurs in the direction of the gradient vector $\nabla\phi$, i.e.,

$$\hat{\boldsymbol{n}} = \frac{\nabla\phi}{|\nabla\phi|} = \frac{4\boldsymbol{i} + 2\boldsymbol{k}}{2\sqrt{5}} = \frac{2}{\sqrt{5}}\boldsymbol{i} + \frac{1}{\sqrt{5}}\boldsymbol{k}.$$

We can extend the above analysis to find the rate of change of a vector field (rather than a scalar field as above) in a particular direction. The scalar differential operator $\hat{\boldsymbol{a}} \cdot \nabla$

can be shown to give the rate of change with distance in the direction $\hat{\boldsymbol{a}}$ of the quantity (vector or scalar) on which it acts. In Cartesian coordinates it may be written as

$$\hat{\boldsymbol{a}} \cdot \nabla = a_x \frac{\partial}{\partial x} + a_y \frac{\partial}{\partial y} + a_z \frac{\partial}{\partial z}. \tag{10.31}$$

Thus we can write the infinitesimal change in an electric field in moving from $\boldsymbol{r}$ to $\boldsymbol{r} + d\boldsymbol{r}$ given in (10.20) as

$$d\boldsymbol{E} = (d\boldsymbol{r} \cdot \nabla)\boldsymbol{E}.$$

A second interesting geometrical property of $\nabla\phi$ may be found by considering the surface defined by $\phi(x, y, z) = c$, where c is some constant. If $\hat{\boldsymbol{t}}$ is a unit tangent to this surface at some point then clearly $d\phi/ds = 0$ in this direction and from (10.29) we have

$$\nabla\phi \cdot \hat{\boldsymbol{t}} = 0.$$

In other words, $\nabla\phi$ is a vector normal to the surface $\phi(x, y, z) = c$ at every point, as shown in figure 10.5. If $\hat{\boldsymbol{n}}$ is a unit normal to the surface in the direction of increasing $\phi(x, y, z)$, then the gradient is sometimes written

$$\nabla\phi \equiv \frac{\partial \phi}{\partial n}\hat{\boldsymbol{n}}, \tag{10.32}$$

where $\partial\phi/\partial n \equiv |\nabla\phi|$ is the rate of change of ϕ in the direction $\hat{\boldsymbol{n}}$ and is called the normal derivative.

Example: Find expressions for the equations of the tangent plane and the line normal to the surface

$$\phi(x, y, z) = c$$

at the point P with coordinates (x_0, y_0, z_0). Use the results to find the equations of the tangent plane and the line normal to the surface of the sphere

$$\phi = x^2 + y^2 + z^2 = a^2$$

at the point $(0, 0, a)$.

Solution: Since $\phi(x, y, z) = c$ defines a surface, the gradient vector

$$\nabla\phi = \frac{\partial \phi}{\partial x}\boldsymbol{i} + \frac{\partial \phi}{\partial y}\boldsymbol{j} + \frac{\partial \phi}{\partial z}\boldsymbol{k}$$

is normal to the surface at every point.

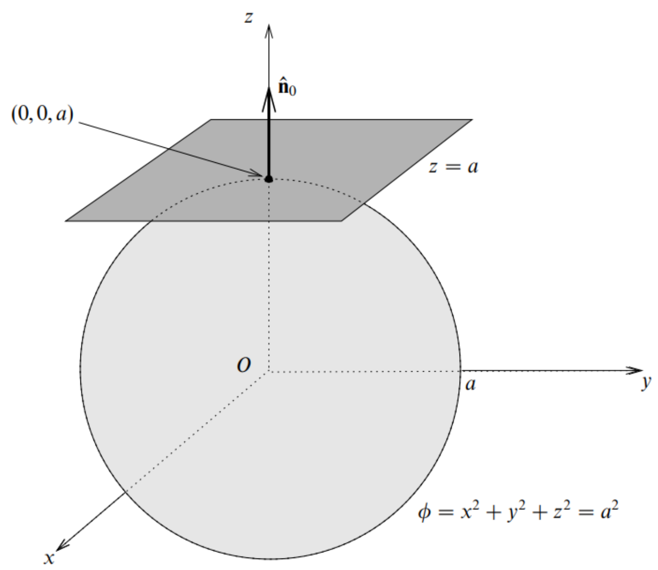

Figure 15: The tangent plane and the normal to the surface of the sphere $\phi = x^2+y^2+z^2 = a^2$ at the point $\mathbf{r}_0 = (0, 0, a)$

Equation of the tangent plane:
At point $P(x_0, y_0, z_0)$, the tangent plane satisfies

$$\nabla\phi(x_0, y_0, z_0) \cdot (\boldsymbol{r} - \boldsymbol{r}_0) = 0,$$

where $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k}$ and $\boldsymbol{r}_0 = x_0\boldsymbol{i} + y_0\boldsymbol{j} + z_0\boldsymbol{k}$.
Writing explicitly,

$$\left.\frac{\partial\phi}{\partial x}\right|_P (x - x_0) + \left.\frac{\partial\phi}{\partial y}\right|_P (y - y_0) + \left.\frac{\partial\phi}{\partial z}\right|_P (z - z_0) = 0.$$

Equation of the normal line:
The line normal to the surface at P passes through P in the direction of $\nabla\phi(x_0, y_0, z_0)$, so its parametric equations are

$$x = x_0 + t\left.\frac{\partial\phi}{\partial x}\right|_P, \quad y = y_0 + t\left.\frac{\partial\phi}{\partial y}\right|_P, \quad z = z_0 + t\left.\frac{\partial\phi}{\partial z}\right|_P,$$

where t is a parameter.

Application to the sphere:
For the sphere,

$$\phi = x^2 + y^2 + z^2 = a^2.$$

The gradient is

$$\nabla\phi = 2x\boldsymbol{i} + 2y\boldsymbol{j} + 2z\boldsymbol{k}.$$

At the point $P = (0, 0, a)$,

$$\nabla\phi(0, 0, a) = 2 \times 0\boldsymbol{i} + 2 \times 0\boldsymbol{j} + 2a\boldsymbol{k} = 2a\boldsymbol{k}.$$

Tangent plane:

$$2 \times 0(x - 0) + 2 \times 0(y - 0) + 2a(z - a) = 0,$$

which simplifies to

$$z = a.$$

Normal line:

$$x = 0 + t \times 0 = 0, \quad y = 0 + t \times 0 = 0, \quad z = a + t \times 2a = a(1 + 2t).$$

Further properties of the gradient operation, which are analogous to those of the ordinary derivative, are listed in subsection 10.8.1 and may be easily proved. In addition to these, we note that the gradient operation also obeys the chain rule as in ordinary differential calculus, i.e., if ϕ and ψ are scalar fields in some region R then

$$\nabla[\phi(\psi)] = \frac{\partial \phi}{\partial \psi} \nabla \psi.$$

3.2.2. Divergence of a vector field

The divergence of a vector field $\mathbf{a}(x, y, z)$ is defined by

$$\operatorname{div} \mathbf{a} = \nabla \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}, \tag{10.33}$$

where a_x, a_y and a_z are the x-, y- and z-components of $\mathbf{a}$. Clearly, $\nabla \cdot \mathbf{a}$ is a scalar field. Any vector field $\mathbf{a}$ for which $\nabla \cdot \mathbf{a} = 0$ is said to be *solenoidal.*

Example: Find the divergence of the vector field

$$\mathbf{a} = x^2 y^2 \, \mathbf{i} + y^2 z^2 \, \mathbf{j} + x^2 z^2 \, \mathbf{k}.$$

Solution: The divergence of a vector field $\mathbf{a} = a_x \, \mathbf{i} + a_y \, \mathbf{j} + a_z \, \mathbf{k}$ is given by

$$\nabla \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}.$$

For the given field:

$$a_x = x^2 y^2, \qquad \frac{\partial a_x}{\partial x} = 2xy^2,$$

$$a_y = y^2 z^2, \qquad \frac{\partial a_y}{\partial y} = 2yz^2,$$

$$a_z = x^2 z^2, \qquad \frac{\partial a_z}{\partial z} = 2zx^2.$$

Therefore, the divergence is:

$$\nabla \cdot \mathbf{a} = 2xy^2 + 2yz^2 + 2zx^2.$$

Now if some vector field $\mathbf{a}$ is itself derived from a scalar field via $\mathbf{a} = \nabla\phi$, then $\nabla \cdot \mathbf{a}$ has the form $\nabla \cdot \nabla\phi$ or, as it is usually written, $\nabla^2\phi$, where ∇^2 (del squared) is the scalar differential operator

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{10.34}$$

$\nabla^2\phi$ is called the *Laplacian* of ϕ and appears in several important partial differential equations of mathematical physics.

Example: Find the Laplacian of the scalar field

$$\phi = xy^2z^3$$

.

Solution: The Laplacian of a scalar field $\phi(x, y, z)$ is given by

$$\nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}.$$

Given:

$$\phi = xy^2z^3$$

Compute second partial derivatives:

$$\frac{\partial^2\phi}{\partial x^2} = \frac{\partial}{\partial x}(y^2z^3) = 0$$

$$\frac{\partial^2\phi}{\partial y^2} = \frac{\partial}{\partial y}(2xyz^3) = 2xz^3$$

$$\frac{\partial^2\phi}{\partial z^2} = \frac{\partial}{\partial z}(3xy^2z^2) = 6xy^2z$$

Therefore,

$$\nabla^2\phi = 0 + 2xz^3 + 6xy^2z = 2xz^3 + 6xy^2z$$

3.2.3. Curl of a vector field

The curl of a vector field $\mathbf{a}(x, y, z)$ is defined by

$$\text{curl } \mathbf{a} = \nabla \times \mathbf{a} = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right)\mathbf{k},$$

where a_x, a_y and a_z are the x-, y- and z-components of $\mathbf{a}$. The right-hand side can be written in a more memorable form as a determinant:

$$\nabla \times \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ a_x & a_y & a_z \end{vmatrix}, \tag{10.35}$$

where it is understood that, on expanding the determinant, the partial derivatives in the second row act on the components of $\mathbf{a}$ in the third row. Clearly, $\nabla \times \mathbf{a}$ is itself a vector field. Any vector field $\mathbf{a}$ for which $\nabla \times \mathbf{a} = 0$ is said to be *irrotational.*

Example: Find the curl of the vector field

$$\mathbf{a} = x^2y^2z^2\,\mathbf{i} + y^2z^2\,\mathbf{j} + x^2z^2\,\mathbf{k}.$$

Solution: We compute the curl using:

$$\nabla \times \mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x^2y^2z^2 & y^2z^2 & x^2z^2 \end{vmatrix}$$

Now compute each component:
i-component:

$$\frac{\partial}{\partial y}(x^2z^2) - \frac{\partial}{\partial z}(y^2z^2) = 0 - \left(y^2 \cdot 2z\right) = -2y^2z$$

j-component:

$$\frac{\partial}{\partial z}(x^2y^2z^2) - \frac{\partial}{\partial x}(x^2z^2) = x^2y^2 \cdot 2z - 2xz^2 = 2x^2y^2z - 2xz^2$$

k-component:

$$\frac{\partial}{\partial x}(y^2z^2) - \frac{\partial}{\partial y}(x^2y^2z^2) = 0 - x^2z^2 \cdot 2y = -2x^2yz^2$$

Therefore, the curl is:

$$\nabla \times \mathbf{a} = -2y^2z\,\mathbf{i} + \left(2x^2y^2z - 2xz^2\right)\mathbf{j} - 2x^2yz^2\,\mathbf{k}$$

For a vector field $\mathbf{v}(x, y, z)$ describing the local velocity at any point in a fluid, $\nabla \times \mathbf{v}$ is a measure of the angular velocity of the fluid in the neighbourhood of that point. If a small paddle wheel were placed at various points in the fluid, then it would tend to rotate in regions where $\nabla \times \mathbf{v} \neq 0$, while it would not rotate in regions where $\nabla \times \mathbf{v} = 0$.

Another insight into the physical interpretation of the curl operator is gained by considering the vector field $\mathbf{v}$ describing the velocity at any point in a rigid body rotating about some axis with angular velocity $\boldsymbol{\omega}$. If $\mathbf{r}$ is the position vector of the point with respect to some origin on the axis of rotation, then the velocity of the point is given by $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$. Without any loss of generality, we may take $\boldsymbol{\omega}$ to lie along the z-axis of our coordinate system, so that $\boldsymbol{\omega} = \omega\mathbf{k}$. The velocity field is then $\mathbf{v} = -\omega y\,\mathbf{i} + \omega x\,\mathbf{j}$. The curl of this vector field is easily found to be

$$\nabla \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix} = 2\omega\,\mathbf{k} = 2\boldsymbol{\omega}. \qquad (10.36)$$

$$\nabla(\phi + \psi) = \nabla\phi + \nabla\psi$$

$$\nabla \cdot (\mathbf{a} + \mathbf{b}) = \nabla \cdot \mathbf{a} + \nabla \cdot \mathbf{b}$$

$$\nabla \times (\mathbf{a} + \mathbf{b}) = \nabla \times \mathbf{a} + \nabla \times \mathbf{b}$$

$$\nabla(\phi\psi) = \phi\nabla\psi + \psi\nabla\phi$$

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}) + (\mathbf{a} \cdot \nabla)\mathbf{b} + (\mathbf{b} \cdot \nabla)\mathbf{a}$$

$$\nabla \cdot (\phi\mathbf{a}) = \phi(\nabla \cdot \mathbf{a}) + \mathbf{a} \cdot \nabla\phi$$

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})$$

$$\nabla \times (\phi\mathbf{a}) = \nabla\phi \times \mathbf{a} + \phi(\nabla \times \mathbf{a})$$

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b}$$

Table 10.1 Vector operators acting on sums and products. The operator ∇ is defined in (10.25); ϕ and ψ are scalar fields, $\mathbf{a}$ and $\mathbf{b}$ are vector fields.

Therefore the curl of the velocity field is a vector equal to twice the angular velocity vector of the rigid body about its axis of rotation.

3.3. Vector operator formulae

3.3.1. Vector operators acting on sums and products

Let ϕ and ψ be scalar fields and $\mathbf{a}$ and $\mathbf{b}$ be vector fields. Assuming these fields are differentiable, the action of ∇ (grad, div, and curl) on various sums and products of them is presented in Table 10.1. These relations can be proved by direct calculation.

Example: Show that

$$\nabla \times (\phi\mathbf{a}) = \nabla\phi \times \mathbf{a} + \phi\nabla \times \mathbf{a}.$$

Solution: Recall that the curl operator can be written as a determinant:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}.$$

Let $\mathbf{F} = \phi\mathbf{a} = (\phi a_x, \phi a_y, \phi a_z)$. Then

$$\nabla \times (\phi\mathbf{a}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \phi a_x & \phi a_y & \phi a_z \end{vmatrix}.$$

Expanding the determinant gives the components:

$$\begin{aligned}
(\nabla \times (\phi\mathbf{a}))_x &= \frac{\partial}{\partial y}(\phi a_z) - \frac{\partial}{\partial z}(\phi a_y), \\
(\nabla \times (\phi\mathbf{a}))_y &= \frac{\partial}{\partial z}(\phi a_x) - \frac{\partial}{\partial x}(\phi a_z), \\
(\nabla \times (\phi\mathbf{a}))_z &= \frac{\partial}{\partial x}(\phi a_y) - \frac{\partial}{\partial y}(\phi a_x).
\end{aligned}$$

Using the product rule for differentiation,

$$\frac{\partial}{\partial y}(\phi a_z) = \frac{\partial \phi}{\partial y} a_z + \phi \frac{\partial a_z}{\partial y},$$

and similarly for the other terms. Substitute to get

$$\begin{aligned}
(\nabla \times (\phi\mathbf{a}))_x &= \left(\frac{\partial \phi}{\partial y} a_z + \phi \frac{\partial a_z}{\partial y} \right) - \left(\frac{\partial \phi}{\partial z} a_y + \phi \frac{\partial a_y}{\partial z} \right) \\
&= \left(\frac{\partial \phi}{\partial y} a_z - \frac{\partial \phi}{\partial z} a_y \right) + \phi \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right),
\end{aligned}$$

and similarly for the y and z components.
Recognizing the terms, we write

$$\nabla \times (\phi\mathbf{a}) = (\nabla\phi) \times \mathbf{a} + \phi(\nabla \times \mathbf{a}),$$

which completes the proof.

Some useful special cases of the relations in Table 10.1 are worth noting. If $\mathbf{r}$ is the position vector relative to some origin and $r = |\mathbf{r}|$, then

$$\begin{aligned}
\nabla\phi(r) &= \frac{d\phi}{dr}\,\hat{\mathbf{r}}, \\
\nabla \cdot [\phi(r)\mathbf{r}] &= 3\phi(r) + r\frac{d\phi(r)}{dr}, \\
\nabla^2\phi(r) &= \frac{d^2\phi(r)}{dr^2} + \frac{2}{r}\frac{d\phi(r)}{dr}, \\
\nabla \times [\phi(r)\mathbf{r}] &= \mathbf{0}.
\end{aligned}$$

These results may be proved straightforwardly using Cartesian coordinates but far more simply using spherical polar coordinates, which are discussed in subsection 10.9.2. Particular cases of these results are

$$\begin{aligned}
\nabla r &= \hat{\mathbf{r}}, \\
\nabla \cdot \mathbf{r} &= 3, \\
\nabla \times \mathbf{r} &= \mathbf{0},
\end{aligned}$$

together with

$$\nabla\left(\frac{1}{r}\right) = -\frac{\hat{\mathbf{r}}}{r^2},$$
$$\nabla\cdot\left(\frac{\hat{\mathbf{r}}}{r^2}\right) = -\nabla^2\left(\frac{1}{r}\right) = 4\pi\delta(\mathbf{r}),$$

where $\delta(\mathbf{r})$ is the Dirac delta function.

3.3.2. Combinations of grad, div and curl

We now consider the action of two vector operators in succession on a scalar or vector field. We can immediately discard four of the nine obvious combinations of grad, div and curl, since they clearly do not make sense. If ϕ is a scalar field and $\mathbf{a}$ is a vector field, these four combinations are $\boldsymbol{\nabla}(\boldsymbol{\nabla}\phi)$, $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\cdot\mathbf{a})$, $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\cdot\mathbf{a})$ and $\boldsymbol{\nabla}(\boldsymbol{\nabla}\times\mathbf{a})$. In each case the second (outer) vector operator is acting on the wrong type of field, i.e., scalar instead of vector or vice versa. In $\boldsymbol{\nabla}(\boldsymbol{\nabla}\phi)$, for example, grad acts on $\boldsymbol{\nabla}\phi$, which is a vector field, but we know that grad only acts on scalar fields (although in fact we will see in Chapter 26 that we can form the outer product of the del operator with a vector to give a tensor, but that need not concern us here). Of the five valid combinations of grad, div and curl, two are identically zero, namely

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\phi = \nabla\times\nabla\phi = \mathbf{0}, \tag{10.37}$$
$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\times\mathbf{a} = \nabla\cdot(\nabla\times\mathbf{a}) = 0. \tag{10.38}$$

From (10.37), we see that if $\mathbf{a}$ is derived from the gradient of some scalar function such that $\mathbf{a} = \nabla\phi$, then it is necessarily irrotational, i.e., $\nabla\times\mathbf{a} = \mathbf{0}$. We also note that if $\mathbf{a}$ is an irrotational vector field then another irrotational vector field is $\mathbf{a} + \nabla\phi + \mathbf{c}$, where ϕ is any scalar field and $\mathbf{c}$ is a constant vector. This follows since

$$\nabla\times(\mathbf{a} + \nabla\phi + \mathbf{c}) = \nabla\times\mathbf{a} + \nabla\times\nabla\phi + \nabla\times\mathbf{c} = \mathbf{0},$$

where we used that $\nabla\times\nabla\phi = \mathbf{0}$ and $\nabla\times\mathbf{c} = \mathbf{0}$ because $\mathbf{c}$ is constant. Similarly, from (10.38) we may infer that if $\mathbf{b}$ is the curl of some vector field $\mathbf{a}$ such that $\mathbf{b} = \nabla\times\mathbf{a}$ then $\mathbf{b}$ is solenoidal ($\nabla\cdot\mathbf{b} = 0$). Obviously, if $\mathbf{b}$ is solenoidal and $\mathbf{c}$ is any constant vector then $\mathbf{b} + \mathbf{c}$ is also solenoidal. The three remaining combinations of grad, div and curl are

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\phi = \nabla\cdot\nabla\phi = \nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}, \tag{10.39}$$

$$\begin{aligned}\boldsymbol{\nabla}\boldsymbol{\nabla}\cdot\mathbf{a} = \nabla(\nabla\cdot\mathbf{a}) &= \left(\frac{\partial^2 a_x}{\partial x^2} + \frac{\partial^2 a_y}{\partial x\partial y} + \frac{\partial^2 a_z}{\partial x\partial z}\right)\mathbf{i} + \left(\frac{\partial^2 a_x}{\partial y\partial x} + \frac{\partial^2 a_y}{\partial y^2} + \frac{\partial^2 a_z}{\partial y\partial z}\right)\mathbf{j} \\ &\quad + \left(\frac{\partial^2 a_x}{\partial z\partial x} + \frac{\partial^2 a_y}{\partial z\partial y} + \frac{\partial^2 a_z}{\partial z^2}\right)\mathbf{k},\end{aligned} \tag{10.40}$$

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{a} = \nabla\times(\nabla\times\mathbf{a}) = \nabla(\nabla\cdot\mathbf{a}) - \nabla^2\mathbf{a}, \tag{10.41}$$

where (10.39) and (10.40) are expressed in Cartesian coordinates. In (10.41), the term $\nabla^2\mathbf{a}$ has the linear differential operator ∇^2 acting on a vector (as opposed to a scalar as in (10.39)), which of course consists of a sum of unit vectors multiplied by components. Two cases arise.

1. If the unit vectors are constants (i.e., they are independent of the values of the coordinates) then the differential operator gives a non-zero contribution only when acting upon the components, the unit vectors being merely multipliers.

2. If the unit vectors vary as the values of the coordinates change (i.e., are not constant in direction throughout the whole space) then the derivatives of these vectors appear as contributions to $\nabla^2 \mathbf{a}$.

Cartesian coordinates are an example of the first case in which each component satisfies

$$(\nabla^2 \mathbf{a})_i = \nabla^2 a_i.$$

In this case, (10.41) can be applied to each component separately:

$$[\nabla \times (\nabla \times \mathbf{a})]_i = [\nabla(\nabla \cdot \mathbf{a})]_i - \nabla^2 a_i. \tag{10.42}$$

However, cylindrical and spherical polar coordinates come in the second class. For them, (10.41) is still true, but the further step to (10.42) cannot be made. More complicated vector operator relations may be proved using the relations given above.

Example: Show that

$$\nabla \cdot (\nabla\phi \times \nabla\psi) = 0,$$

where ϕ and ψ are scalar fields.
Solution: Recall the vector calculus identity:

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

where $\mathbf{A}$ and $\mathbf{B}$ are vector fields.
Let

$$\mathbf{A} = \nabla\phi, \quad \mathbf{B} = \nabla\psi.$$

Since the curl of a gradient is always zero,

$$\nabla \times \nabla\phi = \mathbf{0} \quad \text{and} \quad \nabla \times \nabla\psi = \mathbf{0}.$$

Applying the identity:

$$\nabla \cdot (\nabla\phi \times \nabla\psi) = \nabla\psi \cdot (\nabla \times \nabla\phi) - \nabla\phi \cdot (\nabla \times \nabla\psi) = \nabla\psi \cdot \mathbf{0} - \nabla\phi \cdot \mathbf{0} = 0.$$

Hence,

$$\boxed{\nabla \cdot (\nabla\phi \times \nabla\psi) = 0.}$$

⋆ ⋆ ⋆