Natural Sciences Tripos Part IB
Mathematical Methods I
Michaelmas 2019

Professor Nigel Peake

1 Vector calculus

1.1 Motivation

Scientific quantities are of different kinds:

- *scalars* have only magnitude (and sign), e.g. mass, electric charge, energy,
  temperature

- *vectors* have magnitude and direction, e.g. velocity, magnetic field, temperature gradient

A *field* is a quantity that depends continuously on position (and possibly on time). Examples:

- air pressure in this room (scalar field)

- electric field in this room (vector field)

Vector calculus is concerned with scalar and vector fields. The spatial variation of fields is described by *vector differential operators*, which appear in the *partial differential equations* governing the fields.

Vector calculus is most easily done in Cartesian coordinates, but other systems (*curvilinear coordinates*) are better suited for many problems because of symmetries or boundary conditions.

1.2 Suffix notation and Cartesian coordinates

1.2.1 Three-dimensional Euclidean space

This is a close approximation to our physical space:

- *points* are the elements of the space

- *vectors* are translatable, directed line segments

- *Euclidean* means that lengths and angles obey the classical results of geometry

Points and vectors have a geometrical existence without reference to any coordinate system. For definite calculations, however, we must introduce coordinates.

Cartesian coordinates:

- measured with respect to an origin O and a system of orthogonal axes $Oxyz$
- points have coordinates $(x, y, z) = (x_1, x_2, x_3)$
- unit vectors $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ in the three coordinate directions, also called $(\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ or $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$

The *position vector* of a point P is

$$\overrightarrow{OP} = \boldsymbol{r} = \boldsymbol{e}_x x + \boldsymbol{e}_y y + \boldsymbol{e}_z z = \sum_{i=1}^{3} \boldsymbol{e}_i x_i$$

1.2.2 Properties of the Cartesian unit vectors

The unit vectors form a *basis* for the space. Any vector $\boldsymbol{a}$ belonging to the space can be written uniquely in the form

$$\boldsymbol{a} = \boldsymbol{e}_x a_x + \boldsymbol{e}_y a_y + \boldsymbol{e}_z a_z = \sum_{i=1}^{3} \boldsymbol{e}_i a_i$$

where a_i are the *Cartesian components* of the vector $\boldsymbol{a}$.

The basis is *orthonormal* (orthogonal and normalized):

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

and *right-handed*:

$$[\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] \equiv \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 \times \boldsymbol{e}_3 = +1$$

This means that

$$\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3 \qquad \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1 \qquad \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2$$

The choice of basis is not unique. Two different bases, if both orthonormal and right-handed, are simply related by a rotation. The Cartesian components of a vector are different with respect to two different bases.

1.2.3 Suffix notation

- x_i for a coordinate, a_i (e.g.) for a vector component, $\boldsymbol{e}_i$ for a unit vector
- in three-dimensional space the symbolic index i can have the values 1, 2 or 3
- quantities (*tensors*) with more than one index, such as a_{ij} or b_{ijk}, can also be defined

Scalar and vector products can be evaluated in the following way:

$$\begin{aligned}\boldsymbol{a} \cdot \boldsymbol{b} &= (\boldsymbol{e}_1 a_1 + \boldsymbol{e}_2 a_2 + \boldsymbol{e}_3 a_3) \cdot (\boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2 + \boldsymbol{e}_3 b_3) \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3 \;=\; \sum_{i=1}^{3} a_i b_i\end{aligned}$$

$$\begin{aligned}\boldsymbol{a} \times \boldsymbol{b} &= (\boldsymbol{e}_1 a_1 + \boldsymbol{e}_2 a_2 + \boldsymbol{e}_3 a_3) \times (\boldsymbol{e}_1 b_1 + \boldsymbol{e}_2 b_2 + \boldsymbol{e}_3 b_3) \\ &= \boldsymbol{e}_1(a_2 b_3 - a_3 b_2) + \boldsymbol{e}_2(a_3 b_1 - a_1 b_3) + \boldsymbol{e}_3(a_1 b_2 - a_2 b_1) \\ &= \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}\end{aligned}$$

1.2.4 Delta and epsilon

Suffix notation is made easier by defining two symbols. The *Kronecker delta symbol* is

$$\boxed{\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}}$$

In detail:

$$\delta_{11} = \delta_{22} = \delta_{33} = 1$$

all others, e.g. $\delta_{12} = 0$

δ_{ij} gives the components of the unit matrix. It is *symmetric*:

$$\delta_{ji} = \delta_{ij}$$

and has the *substitution property*:

$$\sum_{j=1}^{3} \delta_{ij} a_j = a_i \qquad \text{(in matrix notation, } \mathsf{1a} = \mathsf{a}\text{)}$$

The *Levi-Civita permutation symbol* is

$$\boxed{\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \\ 0 & \text{otherwise} \end{cases}}$$

In detail:

$$\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = 1$$

$$\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$$

all others, e.g. $\epsilon_{112} = 0$

An *even (odd) permutation* is one consisting of an even (odd) number of transpositions (interchanges of two neighbouring objects). Therefore ϵ_{ijk} is *totally antisymmetric*: it changes sign when any two indices are interchanged, e.g. $\epsilon_{jik} = -\epsilon_{ijk}$. It arises in vector products and determinants.

ϵ_{ijk} has three indices because the space has three dimensions. The even permutations of $(1, 2, 3)$ are $(1, 2, 3)$, $(2, 3, 1)$ and $(3, 1, 2)$, which also happen to be the cyclic permutations. So $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij}$.

Then we can write

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^{3} a_i b_i = \sum_{i=1}^{3} \sum_{j=1}^{3} \delta_{ij} a_i b_j$$

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk} \boldsymbol{e}_i a_j b_k$$

1.2.5 Einstein summation convention

The summation sign is conventionally omitted in expressions of this type. It is implicit that *a repeated index is to be summed over*. Thus

$$\boxed{\boldsymbol{a} \cdot \boldsymbol{b} = a_i b_i}$$

and

$$\boxed{\boldsymbol{a} \times \boldsymbol{b} = \epsilon_{ijk} \boldsymbol{e}_i a_j b_k} \qquad \text{or} \qquad (\boldsymbol{a} \times \boldsymbol{b})_i = \epsilon_{ijk} a_j b_k$$

The repeated index should occur exactly twice in any term. Examples of *invalid* notation are:

$$\boldsymbol{a} \cdot \boldsymbol{a} = a_i^2 \qquad \text{(should be } a_i a_i\text{)}$$

$$(\boldsymbol{a} \cdot \boldsymbol{b})(\boldsymbol{c} \cdot \boldsymbol{d}) = a_i b_i c_i d_i \qquad \text{(should be } a_i b_i c_j d_j\text{)}$$

The repeated index is a *dummy index* and can be relabelled at will. Other indices in an expression are *free indices*. The free indices in each term in an equation must agree.

Examples:

$$\boldsymbol{a} = \boldsymbol{b} \qquad a_i = b_i$$

$$\boldsymbol{a} = \boldsymbol{b} \times \boldsymbol{c} \qquad a_i = \epsilon_{ijk} b_j c_k$$

$$|\boldsymbol{a}|^2 = \boldsymbol{b} \cdot \boldsymbol{c} \qquad a_i a_i = b_i c_i = b_j c_j$$

$$\boldsymbol{a} = (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{d} + \boldsymbol{e} \times \boldsymbol{f} \qquad a_i = b_j c_j d_i + \epsilon_{ijk} e_j f_k$$

Contraction is an operation by which we set one free index equal to another, so that it is summed over. For example, the contraction of a_{ij} is a_{ii}. Contraction is equivalent to multiplication by a Kronecker delta:

$$a_{ij}\delta_{ij} = a_{11} + a_{22} + a_{33} = a_{ii}$$

The contraction of δ_{ij} is $\delta_{ii} = 1 + 1 + 1 = 3$ (in three dimensions).

If the summation convention is *not* being used, this should be noted explicitly.

1.2.6 Matrices and suffix notation

Matrix (A) times vector (x):

$$\mathsf{y} = \mathsf{A}\mathsf{x} \qquad y_i = A_{ij} x_j$$

Matrix times matrix:

$$\mathsf{A} = \mathsf{B}\mathsf{C} \qquad A_{ij} = B_{ik} C_{kj}$$

Transpose of a matrix:

$$(\mathsf{A}^{\mathrm{T}})_{ij} = A_{ji}$$

Trace of a matrix:

$$\mathrm{tr}\,\mathsf{A} = A_{ii}$$

Determinant of a (3×3) matrix:

$$\det \mathsf{A} = \epsilon_{ijk} A_{1i} A_{2j} A_{3k}$$

(or many equivalent expressions).

1.2.7 Product of two epsilons

The general identity

$$\epsilon_{ijk}\epsilon_{lmn} = \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix}$$

can be established by the following argument. The value of both the LHS and the RHS:

- is 0 when any of (i, j, k) are equal or when any of (l, m, n) are equal
- is 1 when $(i, j, k) = (l, m, n) = (1, 2, 3)$
- changes sign when any of (i, j, k) are interchanged or when any of (l, m, n) are interchanged

These properties ensure that the LHS and RHS are equal for any choices of the indices. Note that the first property is in fact implied by the third.

We contract the identity once by setting $l = i$:

$$\begin{aligned}
\epsilon_{ijk}\epsilon_{imn} &= \begin{vmatrix} \delta_{ii} & \delta_{im} & \delta_{in} \\ \delta_{ji} & \delta_{jm} & \delta_{jn} \\ \delta_{ki} & \delta_{km} & \delta_{kn} \end{vmatrix} \\
&= \delta_{ii}(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + \delta_{im}(\delta_{jn}\delta_{ki} - \delta_{ji}\delta_{kn}) \\
&\qquad + \delta_{in}(\delta_{ji}\delta_{km} - \delta_{jm}\delta_{ki}) \\
&= 3(\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}) + (\delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn}) \\
&\qquad + (\delta_{jn}\delta_{km} - \delta_{jm}\delta_{kn}) \\
&= \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}
\end{aligned}$$

This is the most useful form to remember:

$$\boxed{\epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}}$$

Given any product of two epsilons with one common index, the indices can be permuted cyclically into this form, e.g.:

$$\epsilon_{\alpha\beta\gamma}\epsilon_{\mu\nu\beta} = \epsilon_{\beta\gamma\alpha}\epsilon_{\beta\mu\nu} = \delta_{\gamma\mu}\delta_{\alpha\nu} - \delta_{\gamma\nu}\delta_{\alpha\mu}$$

Further contractions of the identity:

$$\begin{aligned}\epsilon_{ijk}\epsilon_{ijn} &= \delta_{jj}\delta_{kn} - \delta_{jn}\delta_{kj} \\ &= 3\delta_{kn} - \delta_{kn} \\ &= 2\delta_{kn}\end{aligned}$$

$$\epsilon_{ijk}\epsilon_{ijk} = 6$$

Example ....

▷ *Show that* $(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d}) = (\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d}) - (\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c})$.

$$\begin{aligned}(\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{c}\times\boldsymbol{d}) &= (\boldsymbol{a}\times\boldsymbol{b})_i(\boldsymbol{c}\times\boldsymbol{d})_i \\ &= (\epsilon_{ijk}a_jb_k)(\epsilon_{ilm}c_ld_m) \\ &= \epsilon_{ijk}\epsilon_{ilm}a_jb_kc_ld_m \\ &= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})a_jb_kc_ld_m \\ &= a_jb_kc_jd_k - a_jb_kc_kd_j \\ &= (a_jc_j)(b_kd_k) - (a_jd_j)(b_kc_k) \\ &= (\boldsymbol{a}\cdot\boldsymbol{c})(\boldsymbol{b}\cdot\boldsymbol{d}) - (\boldsymbol{a}\cdot\boldsymbol{d})(\boldsymbol{b}\cdot\boldsymbol{c})\end{aligned}$$

....

1.3 Vector differential operators

1.3.1 The gradient operator

We consider a scalar field (function of position, e.g. temperature)

$$\Phi(x, y, z) = \Phi(\boldsymbol{r})$$

Taylor's theorem for a function of three variables:

$$\begin{aligned}\Phi(x + \delta x, y + \delta y, z + \delta z) &= \Phi(x, y, z)\\ &+ \frac{\partial \Phi}{\partial x}\delta x + \frac{\partial \Phi}{\partial y}\delta y + \frac{\partial \Phi}{\partial z}\delta z + O(\delta x^2, \delta x \delta y, \dots)\end{aligned}$$

Equivalently

$$\Phi(\boldsymbol{r} + \delta\boldsymbol{r}) = \Phi(\boldsymbol{r}) + (\boldsymbol{\nabla}\Phi) \cdot \delta\boldsymbol{r} + O(|\delta\boldsymbol{r}|^2)$$

where the *gradient* of the scalar field is

$$\boldsymbol{\nabla}\Phi = \boldsymbol{e}_x \frac{\partial \Phi}{\partial x} + \boldsymbol{e}_y \frac{\partial \Phi}{\partial y} + \boldsymbol{e}_z \frac{\partial \Phi}{\partial z}$$

also written $\operatorname{grad}\Phi$. In suffix notation

$$\boxed{\boldsymbol{\nabla}\Phi = \boldsymbol{e}_i \frac{\partial \Phi}{\partial x_i}}$$

For an *infinitesimal* increment we can write

$$\mathrm{d}\Phi = (\boldsymbol{\nabla}\Phi) \cdot \mathrm{d}\boldsymbol{r}$$

We can consider $\boldsymbol{\nabla}$ (*del*, *grad* or *nabla*) as a *vector differential operator*

$$\boldsymbol{\nabla} = \boldsymbol{e}_x \frac{\partial}{\partial x} + \boldsymbol{e}_y \frac{\partial}{\partial y} + \boldsymbol{e}_z \frac{\partial}{\partial z} = \boldsymbol{e}_i \frac{\partial}{\partial x_i}$$

which acts on Φ to produce $\boldsymbol{\nabla}\Phi$.

Example . . . .

▷ *Find* $\boldsymbol{\nabla} f$, *where* $f(r)$ *is a function of* $r = |\boldsymbol{r}|$. By definition

$$\boldsymbol{\nabla} f = \boldsymbol{e}_x \frac{\partial f}{\partial x} + \boldsymbol{e}_y \frac{\partial f}{\partial y} + \boldsymbol{e}_z \frac{\partial f}{\partial z}$$

By the chain rule

$$\frac{\partial f}{\partial x} = \frac{\mathrm{d} f}{\mathrm{d} r} \frac{\partial r}{\partial x}$$

To find the latter term on the RHS we differentiate:

$$r^2 = x^2 + y^2 + z^2$$

$$2r \frac{\partial r}{\partial x} = 2x$$

Thus

$$\frac{\partial f}{\partial x} = \frac{\mathrm{d} f}{\mathrm{d} r} \frac{x}{r}, \text{ etc.}$$

and

$$\boldsymbol{\nabla} f = \frac{\mathrm{d} f}{\mathrm{d} r} \left(\frac{x}{r}, \frac{y}{r}, \frac{z}{r} \right) = \frac{\mathrm{d} f}{\mathrm{d} r} \frac{\boldsymbol{r}}{r}$$

. . . .

1.3.2 Geometrical meaning of the gradient

The rate of change of Φ with distance s in the direction of the unit vector $\boldsymbol{t}$, evaluated at the point $\boldsymbol{r}$, is

$$\begin{aligned} \lim_{s \to 0} \frac{\Phi(\boldsymbol{r} + \boldsymbol{t}s) - \Phi(\boldsymbol{r})}{s} &= \lim_{s \to 0} \frac{(\nabla \Phi) \cdot (\boldsymbol{t}s) + \mathcal{O}(s^2)}{s} \\ &= \boldsymbol{t} \cdot \boldsymbol{\nabla} \Phi \end{aligned}$$

This is known as a *directional derivative*.

Notes:

- the directional derivative is maximal when $\boldsymbol{t} \parallel \boldsymbol{\nabla}\Phi$
- the directional derivative is zero when $\boldsymbol{t} \perp \boldsymbol{\nabla}\Phi$
- the directions $\boldsymbol{t} \perp \boldsymbol{\nabla}\Phi$ therefore lie in the plane tangent to the surface given by $\Phi = \text{constant}$

We conclude that:

- $\boldsymbol{\nabla}\Phi$ is a vector field, pointing in the direction of increasing Φ
- the unit vector normal to the surface $\Phi = \text{constant}$ is $\boldsymbol{n} = \boldsymbol{\nabla}\Phi/|\boldsymbol{\nabla}\Phi|$
- the rate of change of Φ with arclength s along a curve is $\mathrm{d}\Phi/\mathrm{d}s = \boldsymbol{t}\cdot\boldsymbol{\nabla}\Phi$, where $\boldsymbol{t} = \mathrm{d}\boldsymbol{r}/\mathrm{d}s$ is the unit tangent vector to the curve

Example . . . .

▷ *Find the unit normal at the point $\boldsymbol{r} = (x, y, z)$ to the surface*

$$\Phi(\boldsymbol{r}) \equiv xy + yz + zx = -c^2$$

where c is a constant. Then find the points where the plane tangent to the surface is parallel to the (x, y) plane.

First part:

$$\boldsymbol{\nabla}\Phi = (y+z, x+z, y+x)$$

$$\boldsymbol{n} = \frac{\boldsymbol{\nabla}\Phi}{|\boldsymbol{\nabla}\Phi|} = \frac{(y+z, x+z, y+x)}{\sqrt{2(x^2+y^2+z^2+xy+xz+yz)}}$$

Second part:

$$\boldsymbol{n} \parallel \boldsymbol{e}_z \quad \text{when} \quad y+z = x+z = 0$$

$$\Rightarrow -z^2 = -c^2 \qquad \Rightarrow z = \pm c$$

solutions: $(-c, -c, c), \quad (c, c, -c)$

. . . .

1.3.3 Related vector differential operators

We now consider a general vector field (e.g. electric field)

$$\boldsymbol{F}(\boldsymbol{r}) = \boldsymbol{e}_x F_x(\boldsymbol{r}) + \boldsymbol{e}_y F_y(\boldsymbol{r}) + \boldsymbol{e}_z F_z(\boldsymbol{r}) = \boldsymbol{e}_i F_i(\boldsymbol{r})$$

The *divergence* of a vector field is the scalar field

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{F} &= \left(\boldsymbol{e}_x\frac{\partial}{\partial x} + \boldsymbol{e}_y\frac{\partial}{\partial y} + \boldsymbol{e}_z\frac{\partial}{\partial z}\right)\cdot(\boldsymbol{e}_x F_x + \boldsymbol{e}_y F_y + \boldsymbol{e}_z F_z)\\ &= \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}\end{aligned}$$

also written $\operatorname{div}\boldsymbol{F}$. Note that the Cartesian unit vectors are independent of position and do not need to be differentiated. In suffix notation

$$\boxed{\boldsymbol{\nabla}\cdot\boldsymbol{F} = \frac{\partial F_i}{\partial x_i}}$$

The *curl* of a vector field is the vector field

$$\begin{aligned}\boldsymbol{\nabla}\times\boldsymbol{F} &= \left(\boldsymbol{e}_x\frac{\partial}{\partial x}+\boldsymbol{e}_y\frac{\partial}{\partial y}+\boldsymbol{e}_z\frac{\partial}{\partial z}\right)\times(\boldsymbol{e}_xF_x+\boldsymbol{e}_yF_y+\boldsymbol{e}_zF_z)\\ &= \boldsymbol{e}_x\left(\frac{\partial F_z}{\partial y}-\frac{\partial F_y}{\partial z}\right)+\boldsymbol{e}_y\left(\frac{\partial F_x}{\partial z}-\frac{\partial F_z}{\partial x}\right)\\ &\quad+\boldsymbol{e}_z\left(\frac{\partial F_y}{\partial x}-\frac{\partial F_x}{\partial y}\right)\\ &= \begin{vmatrix}\boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z\\ \partial/\partial x & \partial/\partial y & \partial/\partial z\\ F_x & F_y & F_z\end{vmatrix}\end{aligned}$$

also written $\operatorname{curl}\boldsymbol{F}$. In suffix notation

$$\boxed{\boldsymbol{\nabla}\times\boldsymbol{F}=\boldsymbol{e}_i\epsilon_{ijk}\frac{\partial F_k}{\partial x_j}} \qquad \text{or} \qquad (\boldsymbol{\nabla}\times\boldsymbol{F})_i=\epsilon_{ijk}\frac{\partial F_k}{\partial x_j}$$

The *Laplacian* of a scalar field is the scalar field

$$\nabla^2\Phi=\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\Phi)=\frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial y^2}+\frac{\partial^2\Phi}{\partial z^2}=\frac{\partial^2\Phi}{\partial x_i\partial x_i}$$

The Laplacian differential operator (*del squared*) is

$$\boxed{\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}=\frac{\partial^2}{\partial x_i\partial x_i}}$$

It appears very commonly in partial differential equations. The Laplacian of a vector field can also be defined. In suffix notation

$$\nabla^2\boldsymbol{F}=\boldsymbol{e}_i\frac{\partial^2F_i}{\partial x_j\partial x_j}$$

because the Cartesian unit vectors are independent of position.

The *directional derivative* operator is

$$\boxed{\boldsymbol{t}\cdot\boldsymbol{\nabla}=t_x\frac{\partial}{\partial x}+t_y\frac{\partial}{\partial y}+t_z\frac{\partial}{\partial z}=t_i\frac{\partial}{\partial x_i}}$$

$\boldsymbol{t}\cdot\boldsymbol{\nabla}\Phi$ is the rate of change of Φ with distance in the direction of the unit vector $\boldsymbol{t}$. This can be thought of either as $\boldsymbol{t}\cdot(\boldsymbol{\nabla}\Phi)$ or as $(\boldsymbol{t}\cdot\boldsymbol{\nabla})\Phi$.

Example . . . .

▷ *Find the divergence and curl of the vector field* $\boldsymbol{F} = (x^2y, y^2z, z^2x)$.

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \frac{\partial}{\partial x}(x^2y) + \frac{\partial}{\partial y}(y^2z) + \frac{\partial}{\partial z}(z^2x) = 2(xy + yz + zx)$$

$$\boldsymbol{\nabla} \times \boldsymbol{F} = \begin{vmatrix} \boldsymbol{e}_x & \boldsymbol{e}_y & \boldsymbol{e}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ x^2y & y^2z & z^2x \end{vmatrix} = (-y^2, -z^2, -x^2)$$

. . . .

Example . . . .

▷ *Find* $\nabla^2 r^n$.

$$\begin{aligned} \boldsymbol{\nabla} r^n &= \left(\frac{\partial r^n}{\partial x}, \frac{\partial r^n}{\partial y}, \frac{\partial r^n}{\partial z}\right) \\ &= nr^{n-2}(x, y, z) \qquad \left(\text{recall } \frac{\partial r}{\partial x} = \frac{x}{r}, \text{ etc.}\right) \\ &= nr^{n-2}\boldsymbol{r} \end{aligned}$$

$$\begin{aligned} \nabla^2 r^n = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} r^n &= \frac{\partial}{\partial x}(nr^{n-2}x) + \frac{\partial}{\partial y}(nr^{n-2}y) + \frac{\partial}{\partial z}(nr^{n-2}z) \\ &= nr^{n-2} + n(n-2)r^{n-4}x^2 + \cdots \\ &= 3nr^{n-2} + n(n-2)r^{n-2} \\ &= n(n+1)r^{n-2} \end{aligned}$$

. . . .

1.3.4 Vector invariance

A scalar is *invariant* under a change of basis, while vector components transform in a particular way under a rotation.

Fields constructed using $\boldsymbol{\nabla}$ share these properties, e.g.:

- $\boldsymbol{\nabla}\Phi$ and $\boldsymbol{\nabla} \times \boldsymbol{F}$ are vector fields and their components depend on the basis
- $\boldsymbol{\nabla} \cdot \boldsymbol{F}$ and $\nabla^2\Phi$ are scalar fields and are invariant under a rotation

grad, div and ∇^2 can be defined in spaces of any dimension, but curl (like the vector product) is a three-dimensional concept.

1.3.5 Vector differential identities

Here Φ and Ψ are arbitrary scalar fields, and $\boldsymbol{F}$ and $\boldsymbol{G}$ are arbitrary vector fields.

Two operators, one field:

$$\begin{aligned}
\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\Phi) &= \nabla^2\Phi \\
\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\Phi) &= \boldsymbol{0} \\
\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{F}) &= 0 \\
\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{F}) &= \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{F}) - \nabla^2\boldsymbol{F}
\end{aligned}$$

One operator, two fields:

$$\begin{aligned}
\boldsymbol{\nabla}(\Phi\Psi) &= \Psi\boldsymbol{\nabla}\Phi + \Phi\boldsymbol{\nabla}\Psi \\
\boldsymbol{\nabla}(\boldsymbol{F}\cdot\boldsymbol{G}) &= (\boldsymbol{G}\cdot\boldsymbol{\nabla})\boldsymbol{F} + \boldsymbol{G}\times(\boldsymbol{\nabla}\times\boldsymbol{F}) + (\boldsymbol{F}\cdot\boldsymbol{\nabla})\boldsymbol{G} + \boldsymbol{F}\times(\boldsymbol{\nabla}\times\boldsymbol{G}) \\
\boldsymbol{\nabla}\cdot(\Phi\boldsymbol{F}) &= (\boldsymbol{\nabla}\Phi)\cdot\boldsymbol{F} + \Phi\boldsymbol{\nabla}\cdot\boldsymbol{F} \\
\boldsymbol{\nabla}\cdot(\boldsymbol{F}\times\boldsymbol{G}) &= \boldsymbol{G}\cdot(\boldsymbol{\nabla}\times\boldsymbol{F}) - \boldsymbol{F}\cdot(\boldsymbol{\nabla}\times\boldsymbol{G}) \\
\boldsymbol{\nabla}\times(\Phi\boldsymbol{F}) &= (\boldsymbol{\nabla}\Phi)\times\boldsymbol{F} + \Phi\boldsymbol{\nabla}\times\boldsymbol{F} \\
\boldsymbol{\nabla}\times(\boldsymbol{F}\times\boldsymbol{G}) &= (\boldsymbol{G}\cdot\boldsymbol{\nabla})\boldsymbol{F} - \boldsymbol{G}(\boldsymbol{\nabla}\cdot\boldsymbol{F}) - (\boldsymbol{F}\cdot\boldsymbol{\nabla})\boldsymbol{G} + \boldsymbol{F}(\boldsymbol{\nabla}\cdot\boldsymbol{G})
\end{aligned}$$

Example . . . .

▷ *Show that* $\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{F}) = 0$ *for any (twice-differentiable) vector field* $\boldsymbol{F}$.

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{F}) = \frac{\partial}{\partial x_i}\left(\epsilon_{ijk}\frac{\partial F_k}{\partial x_j}\right) = \epsilon_{ijk}\frac{\partial^2 F_k}{\partial x_i\partial x_j} = 0$$

(since ϵ_{ijk} is antisymmetric on (i,j) while $\partial^2 F_k/\partial x_i\partial x_j$ is symmetric) .

Example . . . .

▷ *Show that* $\boldsymbol{\nabla}\times(\boldsymbol{F}\times\boldsymbol{G}) = (\boldsymbol{G}\cdot\boldsymbol{\nabla})\boldsymbol{F}-\boldsymbol{G}(\boldsymbol{\nabla}\cdot\boldsymbol{F})-(\boldsymbol{F}\cdot\boldsymbol{\nabla})\boldsymbol{G}+\boldsymbol{F}(\boldsymbol{\nabla}\cdot\boldsymbol{G})$.

$$\begin{aligned}
\boldsymbol{\nabla}\times(\boldsymbol{F}\times\boldsymbol{G}) &= \boldsymbol{e}_i\epsilon_{ijk}\frac{\partial}{\partial x_j}(\epsilon_{klm}F_lG_m) \\
&= \boldsymbol{e}_i\epsilon_{kij}\epsilon_{klm}\frac{\partial}{\partial x_j}(F_lG_m) \\
&= \boldsymbol{e}_i(\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl})\left(G_m\frac{\partial F_l}{\partial x_j}+F_l\frac{\partial G_m}{\partial x_j}\right) \\
&= \boldsymbol{e}_iG_j\frac{\partial F_i}{\partial x_j}-\boldsymbol{e}_iG_i\frac{\partial F_j}{\partial x_j}+\boldsymbol{e}_iF_i\frac{\partial G_j}{\partial x_j}-\boldsymbol{e}_iF_j\frac{\partial G_i}{\partial x_j} \\
&= (\boldsymbol{G}\cdot\boldsymbol{\nabla})\boldsymbol{F}-\boldsymbol{G}(\boldsymbol{\nabla}\cdot\boldsymbol{F})+\boldsymbol{F}(\boldsymbol{\nabla}\cdot\boldsymbol{G})-(\boldsymbol{F}\cdot\boldsymbol{\nabla})\boldsymbol{G}
\end{aligned}$$

. . . .

Notes:

- be clear about which terms to the right an operator is acting on (use brackets if necessary)

- you cannot simply apply standard vector identities to expressions involving $\boldsymbol{\nabla}$, e.g. $\boldsymbol{\nabla}\cdot(\boldsymbol{F}\times\boldsymbol{G}) \neq (\boldsymbol{\nabla}\times\boldsymbol{F})\cdot\boldsymbol{G}$

- $(\boldsymbol{G}\cdot\boldsymbol{\nabla})\boldsymbol{F} = G_j\dfrac{\partial}{\partial x_j}\boldsymbol{e}_kF_k$

Related results:

- if a vector field $\boldsymbol{F}$ is *irrotational* ($\boldsymbol{\nabla}\times\boldsymbol{F}=\boldsymbol{0}$) in a region of space, it can be written as the gradient of a *scalar potential*: $\boldsymbol{F}=\boldsymbol{\nabla}\Phi$. e.g. a ‘conservative’ force field such as gravity

- if a vector field $\boldsymbol{F}$ is *solenoidal* ($\boldsymbol{\nabla}\cdot\boldsymbol{F}=0$) in a region of space, it can be written as the curl of a *vector potential*: $\boldsymbol{F}=\boldsymbol{\nabla}\times\boldsymbol{G}$. e.g. the magnetic field

1.4 Integral theorems

These very important results derive from the *fundamental theorem of calculus* (integration is the inverse of differentiation):

$$\int_a^b \frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = f(b) - f(a)$$

1.4.1 The gradient theorem

$$\boxed{\int_C (\boldsymbol{\nabla}\Phi)\cdot\mathrm{d}\boldsymbol{r} = \Phi(\boldsymbol{r}_2) - \Phi(\boldsymbol{r}_1)}$$

where C is any curve from $\boldsymbol{r}_1$ to $\boldsymbol{r}_2$.

Outline proof:

$$\begin{aligned}\int_C (\boldsymbol{\nabla}\Phi)\cdot\mathrm{d}\boldsymbol{r} &= \int_C \mathrm{d}\Phi\\ &= \Phi(\boldsymbol{r}_2) - \Phi(\boldsymbol{r}_1)\end{aligned}$$

1.4.2 The divergence theorem (Gauss's theorem)

$$\boxed{\int_V (\boldsymbol{\nabla}\cdot\boldsymbol{F})\,\mathrm{d}V = \oint_S \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}}$$

where V is a volume bounded by the closed surface S (also called ∂V). The right-hand side is the *flux* of $\boldsymbol{F}$ through the surface S. The vector surface element is $\mathrm{d}\boldsymbol{S} = \boldsymbol{n}\,\mathrm{d}S$, where $\boldsymbol{n}$ is the outward unit normal vector.

Outline proof: first prove for a cuboid:

$$\begin{aligned}
\int_V (\boldsymbol{\nabla} \cdot \boldsymbol{F})\,\mathrm{d}V &= \int_{z_1}^{z_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} \left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\
&= \int_{z_1}^{z_2} \int_{y_1}^{y_2} \left[F_x(x_2, y, z) - F_x(x_1, y, z) \right] \mathrm{d}y\,\mathrm{d}z \\
&\qquad + \text{two similar terms} \\
&= \int_S \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{S}
\end{aligned}$$

An arbitrary volume V can be subdivided into small cuboids to any desired accuracy. When the integrals are added together, the fluxes through internal surfaces cancel out, leaving only the flux through S.

A *simply connected* volume (e.g. a ball) is one with no holes. It has only an outer surface. A *multiply connected* volume (e.g. a spherical shell) may have more than one surface; all the surfaces must be considered.

Related results:

$$\int_V (\boldsymbol{\nabla}\Phi)\,\mathrm{d}V = \oint_S \Phi\,\mathrm{d}\boldsymbol{S}$$

$$\int_V (\boldsymbol{\nabla}\times\boldsymbol{F})\,\mathrm{d}V = \oint_S \mathrm{d}\boldsymbol{S}\times\boldsymbol{F}$$

The rule is to replace $\boldsymbol{\nabla}$ in the volume integral with $\boldsymbol{n}$ in the surface integral, and $\mathrm{d}V$ with $\mathrm{d}S$ (note that $\mathrm{d}\boldsymbol{S} = \boldsymbol{n}\,\mathrm{d}S$).

Example . . . .

▷ *A submerged body is acted on by a hydrostatic pressure $p = -\rho g z$, where ρ is the density of the fluid, g is the gravitational acceleration and z is the vertical coordinate. Find a simplified expression for the pressure force acting on the body.*

$$\boldsymbol{F} = -\oint_S p \, \mathrm{d}\boldsymbol{S}$$

$$\begin{aligned} F_z = \boldsymbol{e}_z \cdot \boldsymbol{F} &= \oint_S (-\boldsymbol{e}_z p) \cdot \mathrm{d}\boldsymbol{S} \\ &= \int_V \boldsymbol{\nabla} \cdot (-\boldsymbol{e}_z p) \, \mathrm{d}V \\ &= \int_V \frac{\partial}{\partial z}(\rho g z) \, \mathrm{d}V \\ &= \rho g \int_V \mathrm{d}V \\ &= Mg \qquad (M \text{ is the mass of fluid displaced by the body}) \end{aligned}$$

Similarly

$$F_x = \int_V \frac{\partial}{\partial x}(\rho g z) \, \mathrm{d}V = 0$$

$$F_y = 0$$

$$\boldsymbol{F} = \boldsymbol{e}_z \, Mg$$

Archimedes' principle: buoyancy force equals weight of displaced fluid

. . . .

1.4.3 The curl theorem (Stokes's theorem)

$$\boxed{\int_S (\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \mathrm{d}\boldsymbol{S} = \oint_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}}$$

where S is an open surface bounded by the closed curve C (also called ∂S). The right-hand side is the *circulation* of $\boldsymbol{F}$ around the curve C. Whichever way the unit normal $\boldsymbol{n}$ is defined on S, the line integral follows the direction of a right-handed screw around $\boldsymbol{n}$.

Special case: for a planar surface in the (x, y) plane, we have *Green's theorem in the plane*:

$$\iint_A \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathrm{d}x\, \mathrm{d}y = \int_C (F_x\, \mathrm{d}x + F_y\, \mathrm{d}y)$$

where A is a region of the plane bounded by the curve C, and the line integral follows a positive sense.

Outline proof: first prove Green's theorem for a rectangle:

$$
\begin{aligned}
\int_A (\boldsymbol{\nabla} \times \boldsymbol{F}) \cdot \mathrm{d}\boldsymbol{S} &= \int_{y_1}^{y_2} \int_{x_1}^{x_2} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \mathrm{d}x \, \mathrm{d}y \\
&= \int_{y_1}^{y_2} \left[F_y(x_2, y) - F_y(x_1, y) \right] \mathrm{d}y \\
&\quad - \int_{x_1}^{x_2} \left[F_x(x, y_2) - F_x(x, y_1) \right] \mathrm{d}x \\
&= \int_{x_1}^{x_2} F_x(x, y_1) \, \mathrm{d}x + \int_{y_1}^{y_2} F_y(x_2, y) \, \mathrm{d}y \\
&\quad + \int_{x_2}^{x_1} F_x(x, y_2) \, \mathrm{d}x + \int_{y_2}^{y_1} F_y(x_1, y) \, \mathrm{d}y \\
&= \int_C (F_x \, \mathrm{d}x + F_y \, \mathrm{d}y) \\
&= \oint_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}
\end{aligned}
$$

An arbitrary surface S can be subdivided into small planar rectangles to any desired accuracy. When the integrals are added together, the circulations along internal curve segments cancel out, leaving only the circulation around C.

Notes:

- in Stokes's theorem S is an open surface, while in Gauss's theorem it is closed
- many different surfaces are bounded by the same closed curve, while only one volume is bounded by a closed surface
- a multiply connected surface (e.g. an annulus) may have more than one bounding curve

1.4.4 Geometrical definitions of grad, div and curl

The integral theorems can be used to assign coordinate-independent meanings to grad, div and curl.

Apply the gradient theorem to an arbitrarily small line segment $\delta\boldsymbol{r} = \boldsymbol{t}\,\delta s$ in the direction of any unit vector $\boldsymbol{t}$. Since the variation of $\boldsymbol{\nabla}\Phi$ and $\boldsymbol{t}$ along the line segment is negligible,

$$(\boldsymbol{\nabla}\Phi)\cdot\boldsymbol{t}\,\delta s \approx \delta\Phi$$

and so

$$\boldsymbol{t}\cdot(\boldsymbol{\nabla}\Phi) = \lim_{\delta s\to 0}\frac{\delta\Phi}{\delta s}$$

This definition can be used to determine the component of $\boldsymbol{\nabla}\Phi$ in any desired direction.

Similarly, by applying the divergence theorem to an arbitrarily small volume δV bounded by a surface δS, we find that

$$\boldsymbol{\nabla} \cdot \boldsymbol{F} = \lim_{\delta V \to 0} \frac{1}{\delta V} \int_{\delta S} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{S}$$

Finally, by applying the curl theorem to an arbitrarily small open surface δS with unit normal vector $\boldsymbol{n}$ and bounded by a curve δC, we find that

$$\boldsymbol{n} \cdot (\boldsymbol{\nabla} \times \boldsymbol{F}) = \lim_{\delta S \to 0} \frac{1}{\delta S} \int_{\delta C} \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{r}$$

The gradient therefore describes the rate of change of a scalar field with distance. The divergence describes the net source or efflux of a vector field per unit volume. The curl describes the circulation or rotation of a vector field per unit area.

1.5 Orthogonal curvilinear coordinates

Cartesian coordinates can be replaced with any independent set of coordinates $q_1(x_1, x_2, x_3)$, $q_2(x_1, x_2, x_3)$, $q_3(x_1, x_2, x_3)$, e.g. cylindrical or spherical polar coordinates.

Curvilinear (as opposed to rectilinear) means that the coordinate 'axes' are curves. Curvilinear coordinates are useful for solving problems in curved geometry (e.g. geophysics).

1.5.1 Line element

The infinitesimal *line element* in Cartesian coordinates is

$$\mathrm{d}\boldsymbol{r} = \boldsymbol{e}_x \,\mathrm{d}x + \boldsymbol{e}_y \,\mathrm{d}y + \boldsymbol{e}_z \,\mathrm{d}z$$

In general curvilinear coordinates we have

$$\boxed{\mathrm{d}\boldsymbol{r} = \boldsymbol{h}_1 \,\mathrm{d}q_1 + \boldsymbol{h}_2 \,\mathrm{d}q_2 + \boldsymbol{h}_3 \,\mathrm{d}q_3}$$

where

$$\boxed{\boldsymbol{h}_i = \boldsymbol{e}_i h_i = \frac{\partial \boldsymbol{r}}{\partial q_i}} \qquad \text{(no sum)}$$

determines the displacement associated with an increment of the coordinate q_i.

$h_i = |\boldsymbol{h}_i|$ is the *scale factor* (or *metric coefficient*) associated with the coordinate q_i. It converts a coordinate increment into a length. Any point at which $h_i = 0$ is a *coordinate singularity* at which the coordinate system breaks down.

$\boldsymbol{e}_i$ is the corresponding unit vector. This notation generalizes the use of $\boldsymbol{e}_i$ for a Cartesian unit vector. For Cartesian coordinates, $h_i = 1$ and $\boldsymbol{e}_i$ are constant, but in general both h_i and $\boldsymbol{e}_i$ *depend on position*.

The Einstein summation convention does not work well with orthogonal curvilinear coordinates.

1.5.2 The Jacobian

The *Jacobian* of (x, y, z) with respect to (q_1, q_2, q_3) is defined as

$$J = \frac{\partial(x, y, z)}{\partial(q_1, q_2, q_3)} = \begin{vmatrix} \partial x/\partial q_1 & \partial x/\partial q_2 & \partial x/\partial q_3 \\ \partial y/\partial q_1 & \partial y/\partial q_2 & \partial y/\partial q_3 \\ \partial z/\partial q_1 & \partial z/\partial q_2 & \partial z/\partial q_3 \end{vmatrix}$$

This is the determinant of the *Jacobian matrix* of the transformation from coordinates (x_1, x_2, x_3) to (q_1, q_2, q_3). The columns of the above matrix are the vectors $\boldsymbol{h}_i$ defined above. Therefore the Jacobian is equal to the scalar triple product

$$J = [\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3] = \boldsymbol{h}_1 \cdot \boldsymbol{h}_2 \times \boldsymbol{h}_3$$

Given a point with curvilinear coordinates (q_1, q_2, q_3), consider three small displacements $\delta \boldsymbol{r}_1 = \boldsymbol{h}_1 \, \delta q_1$, $\delta \boldsymbol{r}_2 = \boldsymbol{h}_2 \, \delta q_2$ and $\delta \boldsymbol{r}_3 = \boldsymbol{h}_3 \, \delta q_3$ along the three coordinate directions. They span a parallelepiped of volume

$$\delta V = |[\delta \boldsymbol{r}_1, \delta \boldsymbol{r}_2, \delta \boldsymbol{r}_3]| = |J|\, \delta q_1\, \delta q_2\, \delta q_3$$

Hence the *volume element* in a general curvilinear coordinate system is

$$\boxed{\mathrm{d}V = \left|\frac{\partial(x,y,z)}{\partial(q_1,q_2,q_3)}\right| \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3}$$

The Jacobian therefore appears whenever changing variables in a multiple integral:

$$\int \Phi(\boldsymbol{r})\, \mathrm{d}V = \iiint \Phi\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z = \iiint \Phi \left|\frac{\partial(x,y,z)}{\partial(q_1,q_2,q_3)}\right| \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3$$

The limits on the integrals also need to be considered. Note that if $|J| = 0$ anywhere in the range of variables, the coordinate transformation is invalid.

Jacobians are defined similarly for transformations in any number of dimensions. If curvilinear coordinates (q_1, q_2) are introduced in the (x, y)-plane, the area element is

$$\mathrm{d}A = |J|\, \mathrm{d}q_1\, \mathrm{d}q_2$$

with

$$J = \frac{\partial(x,y)}{\partial(q_1,q_2)} = \begin{vmatrix} \partial x/\partial q_1 & \partial x/\partial q_2 \\ \partial y/\partial q_1 & \partial y/\partial q_2 \end{vmatrix}$$

The equivalent rule for a one-dimensional integral is

$$\int f(x)\, \mathrm{d}x = \int f(x(q)) \left|\frac{\mathrm{d}x}{\mathrm{d}q}\right| \mathrm{d}q$$

The modulus sign appears here if the integrals are carried out over a positive range (the upper limits are greater than the lower limits).

1.5.3 Properties of Jacobians

Consider now three sets of variables α_i, β_i and γ_i, with $1 \leqslant i \leqslant n$, none of which need be Cartesian coordinates. According to the chain rule of partial differentiation,

$$\frac{\partial \alpha_i}{\partial \gamma_j} = \sum_{k=1}^{n} \frac{\partial \alpha_i}{\partial \beta_k} \frac{\partial \beta_k}{\partial \gamma_j}$$

(Under the summation convention we may omit the Σ sign.) The left-hand side is the ij-component of the Jacobian matrix of the transformation from α_i to γ_i. The equation states that this matrix is the product of the Jacobian matrices of the transformations from α_i to β_i and from β_i to γ_i. Taking the determinant of this matrix equation, we find

$$\frac{\partial(\alpha_1, \cdots, \alpha_n)}{\partial(\gamma_1, \cdots, \gamma_n)} = \frac{\partial(\alpha_1, \cdots, \alpha_n)}{\partial(\beta_1, \cdots, \beta_n)} \frac{\partial(\beta_1, \cdots, \beta_n)}{\partial(\gamma_1, \cdots, \gamma_n)}$$

This is the *chain rule for Jacobians*: the Jacobian matrix of a composite transformation is the product of the Jacobian matrices of the transformations of which it is composed.

In the special case in which $\gamma_i = \alpha_i$ for all i, the left-hand side is 1 (the determinant of the unit matrix) and we obtain

$$\frac{\partial(\alpha_1, \cdots, \alpha_n)}{\partial(\beta_1, \cdots, \beta_n)} = \left[\frac{\partial(\beta_1, \cdots, \beta_n)}{\partial(\alpha_1, \cdots, \alpha_n)}\right]^{-1}$$

The Jacobian of an inverse transformation is therefore the reciprocal of that of the forward transformation. This is a multidimensional generalization of the result $\mathrm{d}x/\mathrm{d}y = (\mathrm{d}y/\mathrm{d}x)^{-1}$.

1.5.4 Orthogonality of coordinates

Calculus in general curvilinear coordinates is difficult. We can make things easier by choosing the coordinates to be orthogonal:

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}$$

and right-handed:

$$\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3$$

The squared line element is then

$$\begin{aligned} |\mathrm{d}\boldsymbol{r}|^2 &= |\boldsymbol{e}_1\, h_1\, \mathrm{d}q_1 + \boldsymbol{e}_2\, h_2\, \mathrm{d}q_2 + \boldsymbol{e}_3\, h_3\, \mathrm{d}q_3|^2 \\ &= h_1^2\, \mathrm{d}q_1^2 + h_2^2\, \mathrm{d}q_2^2 + h_3^2\, \mathrm{d}q_3^2 \end{aligned}$$

There are no cross terms such as $\mathrm{d}q_1\, \mathrm{d}q_2$.

When oriented along the coordinate directions:

- line element $\mathrm{d}\boldsymbol{r} = \boldsymbol{e}_1\, h_1\, \mathrm{d}q_1$
- surface element $\mathrm{d}\boldsymbol{S} = \boldsymbol{e}_3\, h_1 h_2\, \mathrm{d}q_1\, \mathrm{d}q_2$,
- volume element $\mathrm{d}V = h_1 h_2 h_3\, \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3$

Note that, for orthogonal coordinates, $J = \boldsymbol{h}_1 \cdot \boldsymbol{h}_2 \times \boldsymbol{h}_3 = h_1 h_2 h_3$

1.5.5 Commonly used orthogonal coordinate systems

Cartesian coordinates (x, y, z):

$$-\infty < x < \infty, \qquad -\infty < y < \infty, \qquad -\infty < z < \infty$$

$$\boldsymbol{r} = (x, y, z)$$

$$\boldsymbol{h}_x = \frac{\partial \boldsymbol{r}}{\partial x} = (1, 0, 0)$$

$$\boldsymbol{h}_y = \frac{\partial \boldsymbol{r}}{\partial y} = (0, 1, 0)$$

$$\boldsymbol{h}_z = \frac{\partial \boldsymbol{r}}{\partial z} = (0, 0, 1)$$

$$h_x = 1, \qquad \boldsymbol{e}_x = (1, 0, 0)$$

$$h_y = 1, \qquad \boldsymbol{e}_y = (0, 1, 0)$$

$$h_z = 1, \qquad \boldsymbol{e}_z = (0, 0, 1)$$

$$\boldsymbol{r} = x\,\boldsymbol{e}_x + y\,\boldsymbol{e}_y + z\,\boldsymbol{e}_z$$

$$\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

Orthogonal. No singularities.

Cylindrical polar coordinates (ρ, ϕ, z):

$$0 < \rho < \infty, \qquad 0 \leqslant \phi < 2\pi, \qquad -\infty < z < \infty$$

$$\boldsymbol{r} = (x, y, z) = (\rho\cos\phi, \rho\sin\phi, z)$$

$$\boldsymbol{h}_\rho = \frac{\partial \boldsymbol{r}}{\partial \rho} = (\cos\phi, \sin\phi, 0)$$

$$\boldsymbol{h}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = (-\rho\sin\phi, \rho\cos\phi, 0)$$

$$\boldsymbol{h}_z = \frac{\partial \boldsymbol{r}}{\partial z} = (0, 0, 1)$$

$$h_\rho = 1, \qquad \boldsymbol{e}_\rho = (\cos\phi, \sin\phi, 0)$$

$$h_\phi = \rho, \qquad \boldsymbol{e}_\phi = (-\sin\phi, \cos\phi, 0)$$

$$h_z = 1, \qquad \boldsymbol{e}_z = (0, 0, 1)$$

$$\boldsymbol{r} = \rho\,\boldsymbol{e}_\rho + z\,\boldsymbol{e}_z$$

$$\mathrm{d}V = \rho\,\mathrm{d}\rho\,\mathrm{d}\phi\,\mathrm{d}z$$

Orthogonal. Singular on the axis $\rho = 0$.

Warning 1. *Many authors use r for ρ and θ for ϕ. This is confusing because r and θ then have different meanings in cylindrical and spherical polar coordinates. Instead of ρ, which is useful for other things, some authors use R, s or ϖ.*

Spherical polar coordinates (r, θ, ϕ):

$$0 < r < \infty, \qquad 0 < \theta < \pi, \qquad 0 \leqslant \phi < 2\pi$$

$$\boldsymbol{r} = (x, y, z) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$$

$$\boldsymbol{h}_r = \frac{\partial \boldsymbol{r}}{\partial r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

$$\boldsymbol{h}_\theta = \frac{\partial \boldsymbol{r}}{\partial \theta} = (r\cos\theta\cos\phi, r\cos\theta\sin\phi, -r\sin\theta)$$

$$\boldsymbol{h}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} = (-r\sin\theta\sin\phi, r\sin\theta\cos\phi, 0)$$

$$h_r = 1, \qquad \boldsymbol{e}_r = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$$

$$h_\theta = r, \qquad \boldsymbol{e}_\theta = (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta)$$

$$h_\phi = r\sin\theta, \qquad \boldsymbol{e}_\phi = (-\sin\phi, \cos\phi, 0)$$

$$\boldsymbol{r} = r\,\boldsymbol{e}_r$$

$$\mathrm{d}V = r^2 \sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi$$

Orthogonal. Singular on the axis $r = 0$, $\theta = 0$ and $\theta = \pi$.

Notes:

- cylindrical and spherical are related by $\rho = r\sin\theta$, $z = r\cos\theta$
- *plane polar coordinates* are the restriction of cylindrical coordinates to a plane $z =$ constant

1.5.6 Vector calculus in orthogonal coordinates

A scalar field $\Phi(\boldsymbol{r})$ can be regarded as function of (q_1, q_2, q_3):

$$
\begin{aligned}
\mathrm{d}\Phi &= \frac{\partial\Phi}{\partial q_1}\,\mathrm{d}q_1 + \frac{\partial\Phi}{\partial q_2}\,\mathrm{d}q_2 + \frac{\partial\Phi}{\partial q_3}\,\mathrm{d}q_3 \\
&= \left(\frac{\boldsymbol{e}_1}{h_1}\frac{\partial\Phi}{\partial q_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial\Phi}{\partial q_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial\Phi}{\partial q_3}\right) \\
&\qquad \cdot (\boldsymbol{e}_1\,h_1\,\mathrm{d}q_1 + \boldsymbol{e}_2\,h_2\,\mathrm{d}q_2 + \boldsymbol{e}_3\,h_3\,\mathrm{d}q_3) \\
&= (\boldsymbol{\nabla}\Phi)\cdot\mathrm{d}\boldsymbol{r}
\end{aligned}
$$

We identify

$$
\boxed{\boldsymbol{\nabla}\Phi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial\Phi}{\partial q_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial\Phi}{\partial q_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial\Phi}{\partial q_3}}
$$

Thus

$$
\boldsymbol{\nabla} q_i = \frac{\boldsymbol{e}_i}{h_i} \qquad \text{(no sum)}
$$

We now consider a vector field in orthogonal coordinates:

$$
\boldsymbol{F} = \boldsymbol{e}_1 F_1 + \boldsymbol{e}_2 F_2 + \boldsymbol{e}_3 F_3
$$

Finding the divergence and curl are non-trivial because both F_i and $\boldsymbol{e}_i$ depend on position. First, consider

$$
\boldsymbol{\nabla}\times(q_2\boldsymbol{\nabla} q_3) = (\boldsymbol{\nabla} q_2)\times(\boldsymbol{\nabla} q_3) = \frac{\boldsymbol{e}_2}{h_2}\times\frac{\boldsymbol{e}_3}{h_3} = \frac{\boldsymbol{e}_1}{h_2 h_3}
$$

which implies

$$
\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{e}_1}{h_2 h_3}\right) = 0
$$

as well as cyclic permutations of this result. Second, we have

$$
\boldsymbol{\nabla}\times\left(\frac{\boldsymbol{e}_1}{h_1}\right) = \boldsymbol{\nabla}\times(\boldsymbol{\nabla} q_1) = \boldsymbol{0}, \qquad \text{etc.}
$$

Now, to work out $\boldsymbol{\nabla}\cdot\boldsymbol{F}$, write

$$
\begin{aligned}
\boldsymbol{F} &= \left(\frac{\boldsymbol{e}_1}{h_2h_3}\right)(h_2h_3F_1) + \cdots \text{cyclic permutations} \\
\boldsymbol{\nabla}\cdot\boldsymbol{F} &= \left(\frac{\boldsymbol{e}_1}{h_2h_3}\right)\cdot\boldsymbol{\nabla}(h_2h_3F_1) + \cdots \\
&= \left(\frac{\boldsymbol{e}_1}{h_2h_3}\right)\cdot\left[\frac{\boldsymbol{e}_1}{h_1}\frac{\partial}{\partial q_1}(h_2h_3F_1) + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial}{\partial q_2}(h_2h_3F_1)\right. \\
&\qquad \left. + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial}{\partial q_3}(h_2h_3F_1)\right] + \cdots \\
&= \frac{1}{h_1h_2h_3}\frac{\partial}{\partial q_1}(h_2h_3F_1) + \cdots \text{cyclic permutations}
\end{aligned}
$$

Similarly, to work out $\boldsymbol{\nabla}\times\boldsymbol{F}$, write

$$
\begin{aligned}
\boldsymbol{F} &= \left(\frac{\boldsymbol{e}_1}{h_1}\right)(h_1F_1) + \cdots \text{cyclic permutations} \\
\boldsymbol{\nabla}\times\boldsymbol{F} &= \boldsymbol{\nabla}(h_1F_1)\times\left(\frac{\boldsymbol{e}_1}{h_1}\right) + \cdots \\
&= \left[\frac{\boldsymbol{e}_1}{h_1}\frac{\partial}{\partial q_1}(h_1F_1) + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial}{\partial q_2}(h_1F_1) + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial}{\partial q_3}(h_1F_1)\right] \\
&\qquad \times\left(\frac{\boldsymbol{e}_1}{h_1}\right) + \cdots \\
&= \left[\frac{\boldsymbol{e}_2}{h_1h_3}\frac{\partial}{\partial q_3}(h_1F_1) - \frac{\boldsymbol{e}_3}{h_1h_2}\frac{\partial}{\partial q_2}(h_1F_1)\right] + \cdots \text{cyclic permutations}
\end{aligned}
$$

The appearance of the scale factors inside the derivatives can be understood with reference to the geometrical definitions of grad, div and curl:

$$
\begin{aligned}
\boldsymbol{t}\cdot(\boldsymbol{\nabla}\Phi) &= \lim_{\delta s\to 0}\frac{\delta\Phi}{\delta s} \\
\boldsymbol{\nabla}\cdot\boldsymbol{F} &= \lim_{\delta V\to 0}\frac{1}{\delta V}\int_{\delta S}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S} \\
\boldsymbol{n}\cdot(\boldsymbol{\nabla}\times\boldsymbol{F}) &= \lim_{\delta S\to 0}\frac{1}{\delta S}\int_{\delta C}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{r}
\end{aligned}
$$

To summarize:

$$\boldsymbol{\nabla}\Phi = \frac{\boldsymbol{e}_1}{h_1}\frac{\partial\Phi}{\partial q_1} + \frac{\boldsymbol{e}_2}{h_2}\frac{\partial\Phi}{\partial q_2} + \frac{\boldsymbol{e}_3}{h_3}\frac{\partial\Phi}{\partial q_3}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{F} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}(h_2h_3F_1) + \frac{\partial}{\partial q_2}(h_3h_1F_2) + \frac{\partial}{\partial q_3}(h_1h_2F_3)\right]$$

$$\begin{aligned}
\boldsymbol{\nabla}\times\boldsymbol{F} &= \frac{\boldsymbol{e}_1}{h_2h_3}\left[\frac{\partial}{\partial q_2}(h_3F_3) - \frac{\partial}{\partial q_3}(h_2F_2)\right] \\
&\quad + \frac{\boldsymbol{e}_2}{h_3h_1}\left[\frac{\partial}{\partial q_3}(h_1F_1) - \frac{\partial}{\partial q_1}(h_3F_3)\right] \\
&\quad + \frac{\boldsymbol{e}_3}{h_1h_2}\left[\frac{\partial}{\partial q_1}(h_2F_2) - \frac{\partial}{\partial q_2}(h_1F_1)\right] \\
&= \frac{1}{h_1h_2h_3}\begin{vmatrix} h_1\,\boldsymbol{e}_1 & h_2\,\boldsymbol{e}_2 & h_3\,\boldsymbol{e}_3 \\ \partial/\partial q_1 & \partial/\partial q_2 & \partial/\partial q_3 \\ h_1F_1 & h_2F_2 & h_3F_3 \end{vmatrix}
\end{aligned}$$

$$\begin{aligned}
\nabla^2\Phi &= \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial\Phi}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3h_1}{h_2}\frac{\partial\Phi}{\partial q_2}\right)\right. \\
&\quad \left. + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial\Phi}{\partial q_3}\right)\right]
\end{aligned}$$

For the vector Laplacian, we can use the following definition along with the expressions above

$$\nabla^2\boldsymbol{F} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{F}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{F})$$

Example . . . .

▷ *Determine the Laplacian operator in spherical polar coordinates.*

$$h_r = 1, \qquad h_\theta = r, \qquad h_\phi = r\sin\theta$$

$$\begin{aligned}
\nabla^2\Phi &= \frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}\left(r^2\sin\theta\frac{\partial\Phi}{\partial r}\right) + \frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{\partial}{\partial\phi}\left(\frac{1}{\sin\theta}\frac{\partial\Phi}{\partial\phi}\right)\right] \\
&= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2}
\end{aligned}$$

. . . .

1.5.7 Grad, div, curl and ∇^2 in cylindrical and spherical polar coordinates

Cylindrical polar coordinates:

$$\boldsymbol{\nabla}\Phi = \boldsymbol{e}_\rho \frac{\partial\Phi}{\partial\rho} + \frac{\boldsymbol{e}_\phi}{\rho}\frac{\partial\Phi}{\partial\phi} + \boldsymbol{e}_z \frac{\partial\Phi}{\partial z}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{F} = \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho F_\rho) + \frac{1}{\rho}\frac{\partial F_\phi}{\partial\phi} + \frac{\partial F_z}{\partial z}$$

$$\boldsymbol{\nabla}\times\boldsymbol{F} = \frac{1}{\rho}\begin{vmatrix} \boldsymbol{e}_\rho & \rho\,\boldsymbol{e}_\phi & \boldsymbol{e}_z \\ \partial/\partial\rho & \partial/\partial\phi & \partial/\partial z \\ F_\rho & \rho F_\phi & F_z \end{vmatrix}$$

$$\nabla^2\Phi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\Phi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\Phi}{\partial\phi^2} + \frac{\partial^2\Phi}{\partial z^2}$$

Spherical polar coordinates:

$$\boldsymbol{\nabla}\Phi = \boldsymbol{e}_r \frac{\partial\Phi}{\partial r} + \frac{\boldsymbol{e}_\theta}{r}\frac{\partial\Phi}{\partial\theta} + \frac{\boldsymbol{e}_\phi}{r\sin\theta}\frac{\partial\Phi}{\partial\phi}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 F_r) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\,F_\theta) + \frac{1}{r\sin\theta}\frac{\partial F_\phi}{\partial\phi}$$

$$\boldsymbol{\nabla}\times\boldsymbol{F} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \boldsymbol{e}_r & r\,\boldsymbol{e}_\theta & r\sin\theta\,\boldsymbol{e}_\phi \\ \partial/\partial r & \partial/\partial\theta & \partial/\partial\phi \\ F_r & rF_\theta & r\sin\theta\,F_\phi \end{vmatrix}$$

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

Example . . . .

▷ *Evaluate* $\boldsymbol{\nabla}\cdot\boldsymbol{r}$, $\boldsymbol{\nabla}\times\boldsymbol{r}$ *and* $\nabla^2 r^n$ *using spherical polar coordinates.*

$$\boldsymbol{r} = \boldsymbol{e}_r r$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{r} = \frac{1}{r^2\sin\theta}\frac{\partial}{\partial r}(r^2\sin\theta\cdot r) = 3$$

$$\boldsymbol{\nabla}\times\boldsymbol{r} = \frac{1}{r^2\sin\theta}\begin{vmatrix} \boldsymbol{e}_r & r\,\boldsymbol{e}_\theta & r\sin\theta\,\boldsymbol{e}_\phi \\ \partial/\partial r & \partial/\partial\theta & \partial/\partial\phi \\ r & 0 & 0 \end{vmatrix} = \boldsymbol{0}$$

$$\nabla^2 r^n = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial r^n}{\partial r}\right) = n(n+1)r^{n-2}$$

. . . .