Natural Sciences Tripos Part IB
Mathematical Methods I
Michaelmas 2017

Dr Henrik Latter

Contents

1 Vector calculus **1**

1.1 Motivation 1

1.2 Suffix notation and Cartesian coordinates 1

1.2.1 Three-dimensional Euclidean space 1

1.2.2 Properties of the Cartesian unit vectors 2

1.2.3 Suffix notation 3

1.2.4 Delta and epsilon 4

1.2.5 Einstein summation convention 5

1.2.6 Matrices and suffix notation 6

1.2.7 Product of two epsilons 7

1.3 Vector differential operators 9

1.3.1 The gradient operator 9

1.3.2 Geometrical meaning of the gradient 10

1.3.3 Related vector differential operators 12

1.3.4 Vector invariance 14

1.3.5 Vector differential identities 15
1.4 Integral theorems 17
1.4.1 The gradient theorem 17
1.4.2 The divergence theorem (Gauss's theorem) 17
1.4.3 The curl theorem (Stokes's theorem) 22
1.4.4 Geometrical definitions of grad, div and curl . . . 24
1.5 Orthogonal curvilinear coordinates 25
1.5.1 Line element 25
1.5.2 The Jacobian 26
1.5.3 Properties of Jacobians 28
1.5.4 Orthogonality of coordinates 28
1.5.5 Commonly used orthogonal coordinate systems . . 30
1.5.6 Vector calculus in orthogonal coordinates 33
1.5.7 Grad, div, curl and ∇^2 in cylindrical and spherical polar coordinates 36

2 Partial differential equations 38
2.1 Motivation 38
2.2 Linear PDEs of second order 38
2.2.1 Definition 38
2.2.2 Principle of superposition 39
2.2.3 Classic examples 40
2.3 Physical examples of occurrence 40
2.3.1 Examples of Laplace's (or Poisson's) equation . . 40
2.3.2 Examples of the diffusion equation 41
2.3.3 Examples of the wave equation 44
2.3.4 Examples of other second-order linear PDEs . . . 45

2.3.5 Examples of nonlinear PDEs 45
2.4 Separation of variables (Cartesian coordinates) 46
2.4.1 Diffusion equation 46
2.4.2 Wave equation (example) 48
2.4.3 Helmholtz equation (example) 50

3 Green's functions 52
3.1 Impulses and the delta function 52
3.1.1 Physical motivation 52
3.1.2 Step function and delta function 53
3.2 Other definitions 54
3.3 More on generalized functions 57
3.4 Differential equations containing delta functions 57
3.5 Inhomogeneous linear second-order ODEs 60
3.5.1 Complementary functions and particular integral . 60
3.5.2 Initial-value and boundary-value problems 60
3.5.3 Green's function for an initial-value problem . . . 61
3.5.4 Green's function for a boundary-value problem . . 66
3.6 Unlectured remarks 69

4 The Fourier transform 70
4.1 Motivation 70
4.2 Fourier series 71
4.3 Approaching the Fourier transform 72
4.4 Examples 77
4.5 Basic properties of the Fourier transform 80
4.6 The delta function and the Fourier transform 83
4.7 The convolution theorem 84

4.7.1 Definition of convolution 84
4.7.2 Interpretation and examples 84
4.7.3 The convolution theorem 86
4.7.4 Correlation 87
4.8 Parseval's theorem 88
4.9 Power spectra 90

5 Matrices and linear algebra 92
5.1 Motivation 92
5.2 Vector spaces 92
5.2.1 Abstract definition of scalars and vectors 92
5.2.2 Span and linear dependence 94
5.2.3 Basis and dimension 95
5.2.4 Examples 96
5.2.5 Change of basis 97
5.3 Matrices 98
5.3.1 Array viewpoint 98
5.3.2 Linear operators 99
5.3.3 Change of basis again 101
5.4 Inner product (scalar product) 102
5.4.1 Definition 102
5.4.2 The Cauchy–Schwarz inequality 103
5.4.3 Inner product and bases 105
5.5 Hermitian conjugate 106
5.5.1 Definition and simple properties 106
5.5.2 Relationship with inner product 107
5.5.3 Adjoint operator 108

5.5.4 Special square matrices 108
5.6 Eigenvalues and eigenvectors 109
5.6.1 Basic properties 109
5.6.2 Eigenvalues and eigenvectors of Hermitian matrices 111
5.6.3 Related results 112
5.6.4 The degenerate case 114
5.7 Diagonalization of a matrix 116
5.7.1 Similarity 116
5.7.2 Diagonalization 117
5.7.3 Diagonalization of a normal matrix 117
5.7.4 Orthogonal and unitary transformations 118
5.7.5 Uses of diagonalization 119
5.8 Quadratic and Hermitian forms 120
5.8.1 Quadratic form 120
5.8.2 Quadratic surfaces 122
5.8.3 Hermitian form 125
5.9 Stationary property of the eigenvalues 125

6 Elementary analysis **127**
6.1 Motivation 127
6.2 Sequences 127
6.2.1 Limits of sequences 127
6.2.2 Cauchy's principle of convergence 128
6.3 Convergence of series 129
6.3.1 Meaning of convergence 129
6.3.2 Classic examples 129
6.3.3 Absolute and conditional convergence 130

6.3.4 Necessary condition for convergence 131
6.3.5 The comparison test 131
6.3.6 D'Alembert's ratio test 132
6.3.7 Cauchy's root test 133
6.4 Functions of a continuous variable 133
6.4.1 Limits and continuity 133
6.4.2 The O notation 135
6.5 Taylor's theorem for functions of a real variable 136
6.6 Analytic functions of a complex variable 137
6.6.1 Complex differentiability 137
6.6.2 The Cauchy–Riemann equations 137
6.6.3 Analytic functions 138
6.6.4 Consequences of the Cauchy–Riemann equations . 140
6.7 Taylor series for analytic functions 141
6.8 Zeros, poles and essential singularities 142
6.8.1 Zeros of complex functions 142
6.8.2 Poles 143
6.8.3 Laurent series 145
6.8.4 Behaviour at infinity 146
6.9 Convergence of power series 146
6.9.1 Circle of convergence 146
6.9.2 Determination of the radius of convergence 147
6.9.3 Examples 148

7 Ordinary differential equations 149

7.1 Motivation 149
7.2 Classification 149

7.3 Homogeneous linear second-order ODEs 150
7.3.1 Linearly independent solutions 150
7.3.2 The Wronskian 151
7.3.3 Calculation of the Wronskian 151
7.3.4 Finding a second solution 152
7.4 Series solutions 154
7.4.1 Ordinary and singular points 154
7.4.2 Series solutions about an ordinary point 155
7.4.3 Series solutions about a regular singular point . . 160
7.4.4 Irregular singular points 164

Schedule

Vector calculus. Suffix notation. Contractions using δ_{ij} and ϵ_{ijk}. Reminder of vector products, grad, div, curl, ∇^2 and their representations using suffix notation. Divergence theorem and Stokes's theorem. Vector differential operators in orthogonal curvilinear coordinates, e.g. cylindrical and spherical polar coordinates. Jacobians. [6 lectures]

Partial differential equations. Linear second-order partial differential equations; physical examples of occurrence, the method of separation of variables (Cartesian coordinates only). [2]

Green's functions. Response to impulses, delta function (treated heuristically), Green's functions for initial and boundary value problems. [3]

Fourier transform. Fourier transforms; relation to Fourier series, simple properties and examples, convolution theorem, correlation functions, Parseval's theorem and power spectra. [2]

Matrices. N-dimensional vector spaces, matrices, scalar product, transformation of basis vectors. Eigenvalues and eigenvectors of a matrix; degenerate case, stationary property of eigenvalues. Orthogonal and unitary transformations. Quadratic & Hermitian forms, quadric surfaces. [5]

Elementary analysis. Idea of convergence and limits. O notation. Statement of Taylor's theorem with discussion of remainder. Convergence of series; comparison and ratio tests. Power series of a complex variable; circle of convergence. Analytic functions: Cauchy–Riemann equations, rational functions and $\exp(z)$. Zeros, poles and essential singularities. [3]

Series solutions of ordinary differential equations. Homogeneous equations; solution by series (without full discussion of logarithmic singularities), exemplified by Legendre's equation. Classification of singular points. Indicial equations and local behaviour of solutions near singular points. [3]

Course websites

http://www.student-systems.admin.cam.ac.uk/moodle
NST Part IB: Mathematics

www.damtp.cam.ac.uk/user/hl278/NST

Assumed knowledge

I will assume familiarity with the following topics at the level of Course A of Part IA Mathematics for Natural Sciences.

- Algebra of complex numbers
- Algebra of vectors (including scalar and vector products)
- Algebra of matrices
- Eigenvalues and eigenvectors of matrices
- Taylor series and the geometric series
- Calculus of functions of several variables
- Line, surface and volume integrals
- The Gaussian integral
- First-order ordinary differential equations
- Second-order linear ODEs with constant coefficients
- Fourier series

Textbooks

The following textbooks are recommended. The first grew out of the NST Maths course, so it will be particularly close to the lectures in places.

- K. F. Riley, M. P. Hobson and S. J. Bence (2006). *Mathematical Methods for Physics and Engineering*, 3rd edition. Cambridge University Press
- G. Arfken and H. J. Weber (2005). *Mathematical Methods for Physicists*, 6th edition. Academic Press

Questions

- Please ask questions in lecture, especially brief ones (typos, jargon)
- Longer questions can be e-mailed to me: `hl278@cam.ac.uk`. Turnaround time: approximately 1-3 days

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

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

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

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:

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

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

1.2.4 Delta and epsilon

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

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*:

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

The *Levi-Civita permutation symbol* is

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

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

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

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

and

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

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

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

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:

$\boldsymbol{a} = \boldsymbol{b}$	$a_i = b_i$
$\boldsymbol{a} = \boldsymbol{b} \times \boldsymbol{c}$	$a_i = \epsilon_{ijk} b_j c_k$
$\lvert\boldsymbol{a}\rvert^2 = \boldsymbol{b} \cdot \boldsymbol{c}$	$a_i a_i = b_i c_i$
$\boldsymbol{a} = (\boldsymbol{b} \cdot \boldsymbol{c})\boldsymbol{d} + \boldsymbol{e} \times \boldsymbol{f}$	$a_i = b_j c_j d_i + \epsilon_{ijk} e_j f_k$

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:

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):

$\mathsf{y} = \mathsf{A}\mathsf{x}$

Matrix times matrix:

$\mathsf{A} = \mathsf{B}\mathsf{C}$

Transpose of a matrix:

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

Trace of a matrix:

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

Determinant of a (3×3) matrix:

$\det \mathsf{A} =$

(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 $l = i$:

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

This is the most useful form to remember:

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:

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

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

Example .....

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

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

.....

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:

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

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

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

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

. . . . . .

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

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

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 $\Phi =$ 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 $\Phi =$ 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:

$$\boldsymbol{\nabla}\Phi =$$

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

Second part:

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

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

solutions:

. . . . . .

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

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

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

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

$\boldsymbol{t}\cdot\boldsymbol{\nabla}\Phi$ 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* $\boldsymbol{F} = (x^2y, y^2z, z^2x)$.

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

$$\boldsymbol{\nabla} \times \boldsymbol{F} =$$

.....

Example .....

▷ *Find* $\nabla^2 r^n$.

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

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

.....

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:

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

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

.....

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

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)

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:

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.*

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

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

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)

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:

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

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

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

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

where

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

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

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

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,

(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:

and right-handed:

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):

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

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

$$\boldsymbol{h}_x = \frac{\partial \boldsymbol{r}}{\partial x} =$$

$$\boldsymbol{h}_y = \frac{\partial \boldsymbol{r}}{\partial y} =$$

$$\boldsymbol{h}_z = \frac{\partial \boldsymbol{r}}{\partial z} =$$

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

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

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

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

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

Orthogonal. No singularities.

Cylindrical polar coordinates (ρ, ϕ, z):

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

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

$$\boldsymbol{h}_\rho = \frac{\partial \boldsymbol{r}}{\partial \rho} =$$

$$\boldsymbol{h}_\phi = \frac{\partial \boldsymbol{r}}{\partial \phi} =$$

$$\boldsymbol{h}_z = \frac{\partial \boldsymbol{r}}{\partial z} =$$

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

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

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

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

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

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):

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

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

$\boldsymbol{h}_r = \dfrac{\partial \boldsymbol{r}}{\partial r} =$

$\boldsymbol{h}_\theta = \dfrac{\partial \boldsymbol{r}}{\partial \theta} =$

$\boldsymbol{h}_\phi = \dfrac{\partial \boldsymbol{r}}{\partial \phi} =$

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

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

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

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

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

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

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:

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

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

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

To summarize:

$$\boldsymbol{\nabla}\Phi =$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{F} =$$

$$\begin{aligned}\boldsymbol{\nabla}\times\boldsymbol{F} = \frac{\boldsymbol{e}_1}{h_2h_3}\left[\frac{\partial}{\partial q_2}(h_3F_3)-\frac{\partial}{\partial q_3}(h_2F_2)\right] \\ + \frac{\boldsymbol{e}_2}{h_3h_1}\left[\frac{\partial}{\partial q_3}(h_1F_1)-\frac{\partial}{\partial q_1}(h_3F_3)\right] \\ + \frac{\boldsymbol{e}_3}{h_1h_2}\left[\frac{\partial}{\partial q_1}(h_2F_2)-\frac{\partial}{\partial q_2}(h_1F_1)\right]\end{aligned}$$

$$=$$

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

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.*

$\boldsymbol{r} =$

$\boldsymbol{\nabla} \cdot \boldsymbol{r} =$

$\boldsymbol{\nabla} \times \boldsymbol{r} =$

$\nabla^2 r^n =$

. . . . . .

2 Partial differential equations

2.1 Motivation

The variation in space and time of scientific quantities is usually described by differential equations. If the quantities depend on space and time, or on more than one spatial coordinate, then the governing equations are *partial differential equations* (PDEs).

Many of the most important PDEs are *linear* and classical methods of analysis can be applied. The techniques developed for linear equations are also useful in the study of nonlinear PDEs.

2.2 Linear PDEs of second order

2.2.1 Definition

We consider an unknown function $u(x, y)$ (the *dependent variable*) of two *independent variables* x and y. A *partial differential equation* is any equation of the form

involving u and any of its derivatives evaluated at the same point.

- the *order* of the equation is the highest order of derivative that appears
- the equation is *linear* if F depends linearly on u and its derivatives

A linear PDE of second order has the general form

$$Lu = f$$

where L is a differential operator such that

$$Lu = a\frac{\partial^2 u}{\partial x^2} + b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} + d\frac{\partial u}{\partial x} + e\frac{\partial u}{\partial y} + gu$$

where a, b, c, d, e, f, g are functions of x and y.

- if $f = 0$ the equation is said to be *homogeneous*
- if a, b, c, d, e, g are independent of x and y the equation is said to have *constant coefficients*

These ideas can be generalized to more than two independent variables, or to systems of PDEs with more than one dependent variable.

2.2.2 Principle of superposition

L defined above is an example of a *linear operator*:

and

where u and v any functions of x and y, and α is any constant.

The principle of superposition:

- if u and v satisfy the homogeneous equation $Lu = 0$, then αu and $u + v$ also satisfy the homogeneous equation
- similarly, any linear combination of solutions of the homogeneous equation is also a solution
- if the *particular integral* u_p satisfies the inhomogeneous equation $Lu = f$ and the *complementary function* u_c satisfies the homogeneous equation $Lu = 0$, then $u_\mathrm{p} + u_\mathrm{c}$ also satisfies the inhomogeneous equation:

The same principle applies to any other type of linear equation (e.g. algebraic, ordinary differential, integral).

2.2.3 Classic examples

Laplace's equation:

Diffusion equation (heat equation):

$\lambda =$ diffusion coefficient (or *diffusivity*)

Wave equation:

$c =$ wave speed

All involve the vector-invariant operator ∇^2, the form of which depends on the number of active dimensions.

Inhomogeneous versions of these equations are also found. The inhomogeneous term usually represents a 'source' or 'force' that generates or drives the quantity u.

2.3 Physical examples of occurrence

2.3.1 Examples of Laplace's (or Poisson's) equation

Gravitational acceleration is related to gravitational potential by

The source (divergence) of the gravitational field is mass:

where G is Newton's constant and ρ is the mass density. Combine:

External to the mass distribution, we have Laplace's equation $\nabla^2\Phi = 0$. With the source term the inhomogeneous equation is called *Poisson's equation*.

Analogously, in electrostatics, the electric field is related to the electrostatic potential by

The source of the electric field is electric charge:

where ρ is the charge density and ϵ_0 is the permittivity of free space. Combine to obtain Poisson's equation:

The vector field ($\boldsymbol{g}$ or $\boldsymbol{E}$) is said to be *generated* by the potential Φ. A scalar potential is easier to work with because it does not have multiple components and its value is independent of the coordinate system. The potential is also directly related to the energy of the system.

2.3.2 Examples of the diffusion equation

A conserved quantity with density (amount per unit volume) Q and flux density (flux per unit area) $\boldsymbol{F}$ satisfies the *conservation equation*:

This is more easily understood when integrated over the volume V bounded by any (fixed) surface S:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V Q\,\mathrm{d}V = \int_V \frac{\partial Q}{\partial t}\,\mathrm{d}V = -\int_V \boldsymbol{\nabla}\cdot\boldsymbol{F}\,\mathrm{d}V = -\int_S \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}$$

The amount of Q in the volume V changes only as the result of a net flux through S.

Often the flux of Q is directed down the gradient of Q through the linear relation (Fick's law)

$$\boldsymbol{F} = -\lambda \boldsymbol{\nabla} Q$$

Combine to obtain the diffusion equation (if λ is independent of $\boldsymbol{r}$):

In a steady state Q satisfies Laplace's equation.

Example . . . . . .

▷ *Heat conduction in a solid.* Conserved quantity: energy. Heat per unit volume: CT (C is heat capacity per unit volume, T is temperature). Heat flux density: $-K\boldsymbol{\nabla}T$ (K is thermal conductivity). Thus

$$\frac{\partial}{\partial t}(CT) + \boldsymbol{\nabla} \cdot (-K\boldsymbol{\nabla}T) = 0$$

or

e.g. copper: $\lambda \approx 1\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ . . . . . .

Further example: concentration of a contaminant in a gas. $\lambda \approx 0.2\,\mathrm{cm}^2\,\mathrm{s}^{-1}$

2.3.3 Examples of the wave equation

Waves on a string. The string has uniform mass per unit length ρ and uniform tension T. The transverse displacement $y(x,t)$ is small ($|\partial y/\partial x| \ll 1$) and satisfies Newton's second law for an element δs of the string:

$$(\rho\,\delta s)\frac{\partial^2 y}{\partial t^2} = \delta F_y \approx \frac{\partial F_y}{\partial x}\delta x$$

Now,

$$F_y = T\sin\theta \approx T\tan\theta \approx T\frac{\partial y}{\partial x}, \quad \text{and also } \delta s \approx \delta x$$

where θ is the small angle between the x-axis and the tangent to the string. Combine to obtain the one-dimensional wave equation

e.g. violin (D-)string: $T \approx 40\,\mathrm{N}$, $\rho \approx 1\,\mathrm{g\,m^{-1}}$: $c \approx 200\,\mathrm{m\,s^{-1}}$

Electromagnetic waves. Maxwell's equations for the electromagnetic field in a vacuum:

where $\boldsymbol{E}$ is electric field, $\boldsymbol{B}$ is magnetic field, μ_0 is the permeability of free space and ϵ_0 is the permittivity of free space. Eliminate $\boldsymbol{E}$:

$$\frac{\partial^2 \boldsymbol{B}}{\partial t^2} = -\boldsymbol{\nabla} \times \frac{\partial \boldsymbol{E}}{\partial t} = -\frac{1}{\mu_0 \epsilon_0} \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{B})$$

Now use the identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{B}) = \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{B}) - \nabla^2 \boldsymbol{B}$ and Maxwell's equation $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ to obtain the (vector) wave equation

$\boldsymbol{E}$ obeys the same equation. c is the speed of light. $c \approx 3 \times 10^8 \,\mathrm{m\,s^{-1}}$

Further example: sound waves in a gas. $c \approx 300 \,\mathrm{m\,s^{-1}}$

2.3.4 Examples of other second-order linear PDEs

Schrödinger's equation (quantum-mechanical wavefunction of a particle of mass m in a potential V):

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V(\boldsymbol{r}) \psi$$

Helmholtz equation (arises in wave problems):

$$\nabla^2 u + k^2 u = 0$$

Klein–Gordon equation (arises in relativistic quantum mechanics):

$$\frac{\partial^2 u}{\partial t^2} = c^2 (\nabla^2 u - m^2 u)$$

2.3.5 Examples of nonlinear PDEs

Burgers' equation (describes shock waves):

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \lambda \nabla^2 u$$

Nonlinear Schrödinger equation (describes solitons, e.g. in optical fibre communication):

$$\mathrm{i}\frac{\partial\psi}{\partial t} = -\nabla^2\psi - |\psi|^2\psi$$

These equations require different methods of analysis.

2.4 Separation of variables (Cartesian coordinates)

2.4.1 Diffusion equation

We consider the one-dimensional diffusion equation (e.g. conduction of heat along a metal bar):

The bar could be considered *finite* (having two ends), *semi-infinite* (having one end) or *infinite* (having no end). Typical *boundary conditions* at an end are:

- u is specified (*Dirichlet* boundary condition)
- $\partial u/\partial x$ is specified (*Neumann* boundary condition)

If a boundary is removed to infinity, we usually require instead that u be *bounded* (i.e. remain finite) as $x \to \pm\infty$. This condition is needed to eliminate unphysical solutions.

To determine the solution fully we also require *initial conditions*. For the diffusion equation this means specifying u as a function of x at some initial time (usually $t = 0$).

Try a solution in which the variables appear in separate factors:

Substitute this into the PDE:

(recall that a prime denotes differentiation of a function with respect to its argument). Divide through by λXT to *separate the variables*:

$$\frac{T'(t)}{\lambda T(t)} = \frac{X''(x)}{X(x)}$$

The LHS depends only on t, while the RHS depends only on x. Both must therefore equal a constant, which we call $-k^2$ (for later convenience here – at this stage the constant could still be positive or negative). The PDE is separated into two *ordinary differential equations* (ODEs)

$$T' + \lambda k^2 T = 0$$

$$X'' + k^2 X = 0$$

with general solutions

Finite bar: suppose the Neumann boundary conditions $\partial u/\partial x = 0$ (i.e. zero heat flux) apply at the two ends $x = 0, L$. Then the admissible solutions of the X equation are

$$X = C\cos(kx), \qquad k = \frac{n\pi}{L}, \qquad n = 0, 1, 2, \ldots$$

These boundary conditions rule out solutions with negative k^2. Combine the factors to obtain the *elementary solution* (let $C = 1$ WLOG)

$$u = A\cos\left(\frac{n\pi x}{L}\right)\exp\left(-\frac{n^2\pi^2\lambda t}{L^2}\right)$$

Each elementary solution represents a 'decay mode' of the bar. The decay rate is proportional to n^2. The $n = 0$ mode represents a uniform temperature distribution and does not decay.

The principle of superposition allows us to construct a *general solution* as a linear combination of elementary solutions:

$$u(x,t) = \sum_{n=0}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right) \exp\left(-\frac{n^2\pi^2\lambda t}{L^2}\right)$$

The coefficients A_n can be determined from the *initial conditions*. If the temperature at time $t = 0$ is specified, then

$$u(x,0) = \sum_{n=0}^{\infty} A_n \cos\left(\frac{n\pi x}{L}\right)$$

is known. The coefficients A_n are just the Fourier coefficients of the initial temperature distribution.

Semi-infinite bar: The solution $X \propto \cos(kx)$ is valid for any real value of k so that u is bounded as $x \to \infty$. We may assume $k \geqslant 0$ WLOG. The solution decays in time unless $k = 0$. The general solution is a linear combination in the form of an integral:

$$u = \int_0^{\infty} A(k) \cos(kx) \exp(-\lambda k^2 t)\, \mathrm{d}k$$

This is an example of a Fourier integral, studied in section 4.

Notes:

- separation of variables doesn't work for all linear equations, by any means, but it does for some of the most important examples
- it is not supposed to be obvious that the most general solution can be written as a linear combination of separated solutions

2.4.2 Wave equation (example)

Boundary conditions:

$$u = 0 \qquad \text{at } x = 0, L$$

Initial conditions:

$$u, \frac{\partial u}{\partial t} \qquad \text{specified at } t = 0$$

Trial solution:

Elementary solution:

$$u = \left[A \sin\left(\frac{n\pi ct}{L}\right) + B \cos\left(\frac{n\pi ct}{L}\right) \right] \sin\left(\frac{n\pi x}{L}\right)$$

General solution:

$$u = \sum_{n=1}^{\infty} \left[A_n \sin\left(\frac{n\pi ct}{L}\right) + B_n \cos\left(\frac{n\pi ct}{L}\right) \right] \sin\left(\frac{n\pi x}{L}\right)$$

At $t = 0$:

$$u = \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi x}{L}\right)$$

$$\frac{\partial u}{\partial t} = \sum_{n=1}^{\infty} \left(\frac{n\pi c}{L}\right) A_n \sin\left(\frac{n\pi x}{L}\right)$$

The Fourier series for the initial conditions determine all the coefficients A_n, B_n.

2.4.3 Helmholtz equation (example)

A three-dimensional example in a cube:

Boundary conditions:

$u = 0$ on the boundaries

Trial solution:

The variables are separated: each term must equal a constant. Call the constants $-k_x^2, -k_y^2, -k_z^2$, since exponentials, the solutions with positive constants, cannot satisfy the boundary conditions:

$$X'' + k_x^2 X = 0$$

$$Y'' + k_y^2 Y = 0$$

$$Z'' + k_z^2 Z = 0$$

$$k^2 = k_x^2 + k_y^2 + k_z^2$$

$$X \propto \sin(k_x x), \qquad k_x = \frac{n_x \pi}{L}, \qquad n_x = 1, 2, 3, \ldots$$

$$Y \propto \sin(k_y y), \qquad k_y = \frac{n_y \pi}{L}, \qquad n_y = 1, 2, 3, \ldots$$

$$Z \propto \sin(k_z z), \qquad k_z = \frac{n_z \pi}{L}, \qquad n_z = 1, 2, 3, \ldots$$

Elementary solution:

The solution is possible only if k^2 is one of the discrete values

These are the eigenvalues of the problem.