Copier’s Message

These notes may contain errors. In fact, they almost certainly do since they were just copied down by me during lectures and everyone makes mistakes when they do that. The fact that I had to type pretty fast to keep up with the lecturer didn’t help. So obviously don’t rely on these notes.

If you do spot mistakes, I’m only too happy to fix them if you email me at mdj27@cam.ac.uk with a message about them. Messages of gratitude, chocolates and job offers will also be gratefully received.

Whatever you do, don’t start using these notes instead of going to the lectures, because the lecturers don’t just write (and these notes are, or should be, a copy of what went on the blackboard) – they talk as well, and they will explain the concepts and processes much, much better than these notes will. Also beware of using these notes at the expense of copying the stuff down yourself during lectures – it really makes you concentrate and stops your mind wandering if you’re having to write the material down all the time. However, hopefully these notes should help in the following ways;

- you can catch up on material from the odd lecture you’re too ill/drunk/lazy to go to;
- you can find out in advance what’s coming up next time (if you’re that sort of person) and the general structure of the course;
- you can compare them with your current notes if you’re worried you’ve copied something down wrong or if you write so badly you can’t read your own handwriting. Although if there is a difference, it might not be your notes that are wrong!

These notes were taken from the course lectured by Dr Evans in Lent 2010. If you get a different lecturer (increasingly likely as time goes on) the stuff may be rearranged or the concepts may be introduced in a different order, but hopefully the material should be pretty much the same. If they start to mess around with what goes in what course, you may have to start consulting the notes from other courses. And I won’t be updating these notes (beyond fixing mistakes) – I’ll be far too busy trying not to fail my second/third/ith year courses.

Good luck – Mark Jackson

Schedules

These are the schedules for the year 2009/10, i.e. everything in these notes that was examinable in that year. The numbers in brackets after each topic give the subsection of these notes where that topic may be found, to help you look stuff up quickly.

Curves in $\mathbb{R}^3$
- Parameterised curves and arc length (1.1), tangents and normals to curves in $\mathbb{R}^3$ (1.1.2, 1.8), the radius of curvature (1.8).

Integration in $\mathbb{R}^2$ and $\mathbb{R}^3$
- Line integrals (1.3). Surface (1.5) and volume (1.6) integrals: definitions, examples using Cartesian, cylindrical (1.7.1) and spherical (1.7.2) coordinates; change of variables (1.5.3, 1.6.2).

Vector operators
- Directional derivatives (1.2). The gradient of a real-valued function: definition (1.2.1); interpretation as normal to level surfaces (1.4.1); examples including the use of cylindrical, spherical *and general orthogonal curvilinear * coordinates.
Divergence, curl (2.1.1) and $\nabla^2$ (2.1.4) in Cartesian coordinates, examples (2.1.2); formulae for these operators (statement only) in cylindrical (2.5.1), spherical (2.5.2) and general orthogonal curvilinear (2.5)* coordinates. Solenoidal fields, irrotational fields and conservative fields (2.1.4); scalar potentials (2.3.1). Vector derivative identities (2.1.2).

Integration theorems

Divergence theorem (2.2.3), Green’s theorem (2.2.1), Stokes’s theorem (2.2.2), Green’s second theorem (2.2.5): statements; informal proofs (2.4); examples; application to fluid dynamics (2.3.2), and to electromagnetism including statement of Maxwell's equations (3.1).

Laplace’s equation

Laplace’s equation in $\mathbb{R}^2$ and $\mathbb{R}^3$ (3.3): uniqueness theorem (3.3.1) and maximum principle (3.3.3). Solution of Poisson’s equation by Gauss’s method (for spherical and cylindrical symmetry) (3.2, 3.4.1) and as an integral (3.3.4).

Cartesian tensors in $\mathbb{R}^3$

Tensor transformation laws (4.1.1), addition, multiplication (4.2.1), contraction (4.2.3), with emphasis on tensors of second rank (4.3). Isotropic second and third rank tensors (4.4). Symmetric and antisymmetric tensors (4.2.4). Revision of principal axes and diagonalization (4.3.2). Quotient theorem (4.5). Examples including inertia (4.3.2) and conductivity.

Contents

0. Introduction .......................................................................................................................... 3
0.1 Course outline .................................................................................................................. 3
0.2 Recall of some key ideas from calculus (Differential Equations) ........................................ 3
1. Curves and surfaces in $\mathbb{R}^3$ ......................................................................................... 5
  1.1 Curves and tangent vectors ............................................................................................ 5
  1.2 Gradients and directional derivatives ............................................................................ 6
  1.3 Integration along curves ............................................................................................... 7
  1.4 Surfaces and normals .................................................................................................... 9
  1.5 Integrals over surfaces .................................................................................................. 10
  1.6 Integration over volumes .............................................................................................. 12
  1.7 Orthogonal curvilinear coordinates on $\mathbb{R}^3$ .......................................................... 13
  1.8 Geometry of curves and surfaces ................................................................................ 14
2. Vector differential operators and integral theorems .......................................................... 15
  2.1 Grad, div, curl and $\nabla$ ............................................................................................... 15
  2.2 Integral theorems ......................................................................................................... 18
  2.3 Some constructions and comments............................................................................... 21
  2.4 Informal proofs of integral theorems ........................................................................... 24
  2.5 Vector differential operators in orthogonal curvilinear coordinates ............................... 26
3. Applications: Electromagnetism, gravitation and potential theory .................................... 27
  3.1 The laws of electromagnetism ....................................................................................... 27
  3.2 Electrostatics and gravitation ....................................................................................... 28
  3.3 Laplace’s equation and Poisson’s equation .................................................................... 30
  3.4 Summary of methods for solving Laplace’s equations and Poisson’s equations .......... 34
4. Tensors ............................................................................................................................ 35
  4.1 Introduction .................................................................................................................. 35
  4.2 Operations involving tensors ....................................................................................... 37
  4.3 Second rank tensors .................................................................................................... 38
0. Introduction

Vector calculus is the study of scalar fields ϕ(x) and vector fields v(x) and their behaviour on smooth curves, surfaces etc. This includes notions of differentiation and integration (to be developed).

This is the basic language of theoretical physics, e.g.

- gravitational field g(x)
- electromagnetic fields E, B
- fluid velocity field v(x)

The first two explain all of non-relativistic physics above the atomic level. The third is challenging physically and mathematically.

Vector calculus also underlies the differential geometry of curves and surfaces. Ultimately these strands meet up again in Einstein’s geometrical description of gravity.

0.1 Course outline

The course is divided into 4 broad sections;

1. Curves, surfaces, gradients of functions, and integration
2. \( \nabla, \text{grad, div, curl and integral theorems} \)
3. Applications: Laplace’s equations and potential theory
4. Tensors

0.2 Recall of some key ideas from calculus (Differential Equations)

0.2.1 Single variable

- A function \( f : \mathbb{R} \to \mathbb{R}, y = f(x) \) is differentiable at \( x = a \) iff \( f(a + h) = f(a) + mh + o(h) \).
- The derivative at \( a, m \leq dy/dx = f'(a) = m \).

(The notation \( R(h) = O(h^N) \) means that \( |R(h)| \leq Ch^N \) for C constant and h sufficiently small. The notation \( R(h) = o(h) \) means that \( R(h)/h \to 0 \) as \( h \to 0 \).)

- If all derivatives \( f''(x), f'''(x), \ldots, f^{(n)}(x), \ldots \) exist, we call f smooth. In this course all functions will be smooth ‘almost everywhere’ (i.e. it can fail at isolated points).
- The notation for derivative

\[
\delta y = \frac{dy}{dx} \delta x + o(\delta x)
\]

will also be used.

- Taylor’s theorem; for any smooth f, we have

\[
f(a + h) = f(a) + f'(a)h + \frac{1}{2} f''(a)h^2 + \cdots + \frac{1}{n!} f^{(n)}(a)h^n + O(h^{n+1}).
\]

- The derivative gives linearisation of the function near \( a \);

\[
f(x) = f(a) + m(x - a) + o(x - a).
\]

- Chain rule; if f and g are smooth, then \( f \circ g \) is smooth. In other words, if both functions are from \( \mathbb{R} \) to \( \mathbb{R} \) with \( y = f(x) \) and \( x = g(u) \), then

\[
\frac{dy}{du} = \frac{dy}{dx} \frac{dx}{du}
\]
Integration; $F : \mathbb{R} \to \mathbb{R}$ smooth implies the (Riemann) integral $\int_a^b F(x) \, dx$ exists. Also the Fundamental Theorem of Calculus;

$$F(x) = \frac{df}{dx} \Rightarrow \int_a^b \frac{df}{dx} \, dx = f(b) - f(a).$$

0.2.2 Multiple variables

- Now consider $f : \mathbb{R}^n \to \mathbb{R}$, $y = f(x_i)$ with $i = 1, 2, ..., n$.
- This is differentiable at $x_i = a_i$ if $f(a_i + h_i) = f(a_i) + m_i h_i + o(h)$, using the summation convention, where $h = \sqrt{\sum_i h_i^2}$.
- This gives $m_i = \frac{\partial f}{\partial x_i} a_i$ (the partial derivative – keep all other $x_j$ fixed)
- We also write

$$\delta y = \frac{\partial y}{\partial x_i} \delta x_i + o(\delta x)$$

- $f$ is smooth if all partial derivatives of all orders exist, i.e.

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \, \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}, ...$$

- Taylor’s theorem;

$$f(a_i + h_i) = f(a_i) + \frac{\partial f}{\partial x_i}(a_i) h_i + \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}(a_i) h_i h_j + O(h^3).$$

(Refer back to DE’s and classification of stationary points).

0.2.3 General case

- $f : \mathbb{R}^n \to \mathbb{R}^m$ (could restrict to $n, m = 1, 2$ or $3$)
- $y_r = f_r(x_i)$ is differentiable at $x_i = a_i$ iff $f_r(a_i + h_i) = f_r(a_i) + m_{ri} h_i + o(h)$
- $m_{ri} = \frac{\partial f_r}{\partial x_i} a_i$
- We also write

$$\delta y_r = \frac{\partial y_r}{\partial x_i} \delta x_i + o(\delta x)$$

- Chain rule; if $f$ and $g$ are smooth, then $f \circ g$ is smooth. In other words, if $y = f(x)$ and $x = g(u)$, with $y : \mathbb{R}^n \to \mathbb{R}^m$ and $x : \mathbb{R}^p \to \mathbb{R}^n$, then

$$\frac{\partial y_r}{\partial u_a} = \frac{\partial y_r}{\partial x_i} \frac{\partial x_i}{\partial u_a}$$

which follows from the relation above together with

$$\delta x_i = \frac{\partial x_i}{\partial u_a} \delta u_a + o(\delta u)$$

0.2.4 Inverse functions

Now consider $p = n = m$ and suppose $u = y$, with the functions $g$ and $f$ being inverse to one another. On the left of the chain rule, we have $\partial u_b / \partial u_a = \delta_{ab}$. Hence

$$\frac{\partial u_b}{\partial x_i} \frac{\partial x_i}{\partial u_a} = \delta_{ab}; \quad \frac{\partial x_i}{\partial u_a} \frac{\partial u_a}{\partial x_j} = \delta_{ij}$$

i.e. inverse matrices.

In the simplest case of one variable, the relation above collapses to

$$\left(\frac{du}{dx}\right)^{-1} = \frac{dx}{du}$$

Next case; $\mathbb{R}^2$ and take $x_1, x_2$ to be usual Cartesian coordinates, while $u_1 = \rho, u_2 = \varphi$, polar coordinates. Then $x_1 = \rho \cos \varphi$ and $x_2 = \rho \sin \varphi$ whereas $\rho = \sqrt{x_1^2 + x_2^2}$ and $\varphi = \tan^{-1}(x_2/x_1)$. 

The matrix of partial derivatives
\[
\begin{pmatrix}
\frac{\partial x_1}{\partial \rho} & \frac{\partial x_1}{\partial \varphi} \\
\frac{\partial x_2}{\partial \rho} & \frac{\partial x_2}{\partial \varphi}
\end{pmatrix}
= \begin{pmatrix}
\cos \varphi & -\rho \sin \varphi \\
\sin \varphi & \rho \cos \varphi
\end{pmatrix}
\]
whereas
\[
\begin{pmatrix}
\frac{\partial \rho}{\partial x_1} & \frac{\partial \rho}{\partial x_2} \\
\frac{\partial \varphi}{\partial x_1} & \frac{\partial \varphi}{\partial x_2}
\end{pmatrix}
= \frac{1}{\rho} \begin{pmatrix}
\rho \cos \varphi & \rho \sin \varphi \\
-\sin \varphi & \cos \varphi
\end{pmatrix}
\]
by the inverse matrix rule.

Note that in \( \mathbb{R}^n \) there is a special case of Cartesian coordinates characterised by
(i) vector addition \( \equiv \) addition of coordinates
(ii) length given by Pythagoras; \( |x|^2 = \sum x_i^2 \) for \( x = x_1, x_2, ..., x_n \) (contrast this with polars in \( \mathbb{R}^2 \)).

Note that if \( \{x_i\} \) and \( \{x_i'\} \) are two sets of Cartesian coordinates, they are related by
\[ x'_i = R_{ij} x_j \]
with \( R_{ij} \) orthogonal.

1. Curves and surfaces in \( \mathbb{R}^3 \)

We use vector notation, with position; \( x = x_1 e_1 = x_2 e_2 + x_3 e_3 \) or \( r = xi + yj + zk \)
with orthonormal basis vectors corresponding to Cartesian coordinates \( (x_1, x_2, x_3) \) or \( (x, y, z) \).

1.1 Curves and tangent vectors

1.1.1 Parametrised curves

We can describe curves by relations amongst the coordinates, e.g.
\[ z = x^3, y = x^2 \text{ (a twisted cubic)} \text{ and } x^2/4 + y^2 = 1, z = 3, y \geq 0 \text{ (top half of an ellipse)}. \]

But it is usually better to deal with parametrised curves, e.g.
\[ r(u) = ui + u^2j + u^3k \text{ for } -\infty < u < \infty \]
\[ r(u) = 2 \cos u i + \sin u j + 3k \text{ for } 0 \leq u \leq \pi \text{ (these correspond to the curves above)}. \]

In general, a parametrised curve \( C \) is defined to be a smooth function \( i \rightarrow \mathbb{R}^3, r(u) \) with \( i \) some interval in (may be all of) \( \mathbb{R} \). For finite \( i \) with \( \alpha \leq u \leq \beta \), the curve has end points \( a = r(\alpha) \) and \( b = r(\beta) \).

1.1.2 The tangent vector

The tangent vector to \( C \) at any point is
\[ \frac{dr}{du} = v(u), \]
using definition of derivative as local linearisation,
\[ r(\alpha + \delta u) = r(\alpha) + v(\alpha) \delta u + o(\delta u), \]
and comparing this with \( a + v(\alpha) \lambda \), the tangent line to \( C \) at \( a = r(\alpha) \).
Interpretation relies on $\nu(\alpha) \neq 0$ – then we say $C$ is regular at $\alpha$; always assume this unless we state otherwise.

If parameter is $t$, time, then tangent vector $\nu(t) = \dot{x}(t)$ is the velocity.

### 1.1.3 Arc length

Length / inner product on $\mathbb{R}^3$ \implies we can determine the arc length $s$ measured along $C$ (from some chosen point).

For points differing by parameter change $\delta u$,
\[
\delta r = r(u + \delta u) - r(u) = \nu \delta u + o(\delta u)
\]

The increase in arc length $\delta s$ corresponding to increase $\delta u$ is then
\[
\delta s = \sqrt{\delta r \cdot \delta r} + o(|\delta r|) = |\nu| \delta u + o(\delta u)
\]

\[
\implies \frac{ds}{du} = |\nu|
\]

### 1.1.4 Arc length as a parameter

There is a large freedom in choosing parametrisation of curve $C$, using any smooth invertible function $u \mapsto \tilde{u}(u)$ (range changes from $l$ to $\tilde{l}$). We can now regard $r(\tilde{u}(u))$ and the chain rule gives a new tangent vector
\[
\tilde{v} = \frac{dr}{d\tilde{u}} = \frac{dr}{du} \cdot \frac{du}{d\tilde{u}} = (\frac{d\tilde{u}}{du})^{-1} \nu.
\]

So direction of tangent vector doesn’t change under $u \mapsto \tilde{u}$. Choosing $\tilde{u} = s$, arc length, then the tangent vector $t = (ds/du)^{-1} \nu$ which is a unit vector.

**Example.** Circular helix
\[
\begin{align*}
\mathbf{r} &= \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \cos u \\ a \sin u \\ bu \end{pmatrix}, -\infty < u < \infty \\
\mathbf{v} &= \frac{d\mathbf{r}}{du} = \begin{pmatrix} -a \sin u \\ a \cos u \\ b \end{pmatrix} \\
\implies |\mathbf{v}|^2 &= a^2 + b^2 \implies \frac{ds}{du} = \sqrt{a^2 + b^2} \implies s = u\sqrt{a^2 + b^2}
\end{align*}
\]

where $s$ is the arc length measured from the point where $u = 0$.

### 1.2 Gradients and directional derivatives

#### 1.2.1 Definitions

Differentiability of function $f$, using vector notation, can be written
\[
\delta f = f(x + \delta x) - f(x) = (\nabla f) \cdot \delta x + o(|\delta x|).
\]

This defines the **gradient** $\nabla f$ (pronounced “grad $f$”), a vector field. With Cartesian coordinates and basis vectors, we have
\[
\nabla f = \frac{\partial f}{\partial x_i} e_i
\]
since with $\delta x = \delta x_i e_i$ we have $\nabla f \cdot \delta x = (\partial f / \partial x_i) \delta x_i$ as expected.

Grad also arises through the chain rule when we consider how $f(x)$ changes along a curve $x(u)$
\[
\frac{d}{du} (f(x(u))) = (\nabla f) \cdot v = \frac{\partial f}{\partial x_i} \frac{dx_i}{du}
\]
in Cartesian coordinates, where $v = dx/du$. 
1.2.2 Examples and relation to change of $f$

*Example.* Let $T(x)$ be the temperature experienced by some particle moving on a trajectory $x(t)$. Then the rate of change $dT/dt = \nabla T \cdot \mathbf{v}$ where $\mathbf{v}(t)$ is the velocity.

*Example.* Special case a curve a straight line through point $a$ in direction given by unit vector $\mathbf{n}$.

$$\frac{df}{du}(f(a + un)) = \mathbf{n} \cdot \nabla f$$

is the directional derivative of $f$ along $\mathbf{n}$ (at the point $a$).

*Note.* $|\mathbf{n} \cdot \nabla f| \leq |\nabla f|$. So $f$ changes most rapidly when $\mathbf{n} \parallel \nabla f$ and it's increasing or decreasing along $\pm n$. $f$ is unchanged to 1st order for $\mathbf{n} \perp \nabla f$.

*Example.* Let $f(x, y, z) = x + e^{xy} \sin z$. Then

$$\nabla f = \begin{pmatrix} e^{xy} \sin z \\ xe^{xy} \sin z \\ e^{xy} \cos z \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix}$$

At $a = (0,1,0)$ we have $\nabla f = (1,0,1)$.

$f$ increases/decreases *fastest* along $\mathbf{n} = \pm \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, and $f$ is unchanged (to 1st order) for $\mathbf{n} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ or $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, perpendicular to $\nabla f$.

1.2.3 Some elementary properties of $\nabla$

- $\nabla (af + \beta g) = a \nabla f + \beta \nabla g$ ($\alpha, \beta$ numbers)
- $\nabla f g) = (\nabla f) g + f(\nabla g)$

1.3 Integration along curves

To get an informal definition of such integrals, we think of dividing up a curve $x(u)$ into segments labelled by $l$, each of length $\delta s_l$, with $\delta s_l \leq l$. Then for any quantity $\phi$ (to be specified more precisely below) we define

$$\int_C \phi \ ds = \lim_{\delta s \to 0} \sum_l \phi(x_l) \delta s_l$$

where $x_l$ is some representative point on each line segment $l$.

1.3.1 Scalar integral of a scalar function

A function $f(x)$ has an integral

$$\int_C f(x) ds = \lim_{\delta s \to 0} \sum_l f(x_l) \delta s_l.$$ 

Relating arc length $s$ to parameter of curve $u$, we have

$$\int_C f(x) ds = \int_\alpha^\beta f(x(u)) \frac{dx}{du} du$$

for $C$ a curve with end-points $x(\alpha)$ and $x(\beta)$.

1.3.2 Line integral of a vector field $\mathbf{F}(x)$

Unit tangent vector $\mathbf{t}$ to curve which gives natural way of integrating a vector field $\mathbf{F}(x)$ along $C$ and getting a scalar;

$$\int_C \mathbf{F} \cdot dx = \int_C \mathbf{F} \cdot \mathbf{t} \ ds = \int_\alpha^\beta \mathbf{F}(x(u)) \cdot \frac{dx}{du} du$$
Choice of orientation (arrow on $C$) ⇔ choice of sign of tangent direction ⇔ ‘starting at $a$’ and ‘ending at $b$’. Changing orientation changes sign of integral.

Underlying definition as limit of sum involves choosing points $x_i$ on $C$ then

$$\int_C F \cdot dx = \lim_{n \to \infty} \sum F(x_i) \delta x_i.$$  

For a force $F$, each term is work done along $\delta x_i$, and $\int_C F \cdot dx$ is total work along $C$.

Examples. Integrate $F(r) = \begin{pmatrix} xe^y \\ z^2 \\ xy \end{pmatrix}$ along two paths;

$C_1 : r = \begin{pmatrix} t \\ 0 \\ t \end{pmatrix}$ for $0 \leq t \leq 1$ (a straight line). Then $\dot{r} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ and

$$\int_{C_1} F \cdot dr = \int_0^1 \begin{pmatrix} t e^t \\ t^2 \\ t^3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} dt = \int_0^1 (te^t + 2t^2) dt = \frac{5}{3}$$

$C_2 : r = \begin{pmatrix} t^2 \\ 0 \\ t \end{pmatrix}$ for $0 \leq t \leq 1$ (a curve). Then $\dot{r} = \begin{pmatrix} 2t \\ 0 \\ 3t^2 \end{pmatrix}$ and

$$\int_{C_2} F \cdot dr = \int_0^1 \begin{pmatrix} t e^{t^2} \\ t^6 \\ t^3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2t \\ 3t^2 \end{pmatrix} dt = \int_0^1 (te^{t^2} + 2t^7 + 3t^6) dt = e + \frac{1}{4}$$

1.3.3 Line integral of a gradient

From the examples, we see that $\int_C F \cdot dx$ depends on $C$ in general, rather than just its end points. But if $F = \nabla f$ for $f(x)$ some function, then

$$\int_C (\nabla f) \cdot dx = \int_a^b (\nabla f) \cdot \frac{dx}{du} du = \int_a^b \frac{d}{du} f(x(u)) = f(x(b)) - f(x(a)) = f(b) - f(a),$$

and the same result holds for any curve with these end points.

In mechanics, the force $F(x)$ is conservative iff $\int_C F \cdot dx$ depends only on the end points of $C$. Therefore $F = -\nabla \varphi$ where $\varphi(x)$ is the potential energy. Energy is conserved along a trajectory $x(t)$;

$$\frac{d}{dt} \left( \frac{1}{2} m \dot{x}^2 + \varphi \right) = 2m \ddot{x} + \nabla \varphi \cdot \ddot{x} = (F + \nabla \varphi) \cdot \ddot{x} = 0.$$

1.3.4 Differential notation

We know how to evaluate $\int_C$ of expressions like $F \cdot dx = F_1 dx_1 + F_2 dx_2 + F_3 dx_3$ called differentials. For a function $f(x)$ we define

$$df = \frac{\partial f}{\partial x_i} dx_i = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \frac{\partial f}{\partial x_3} dx_3$$

Observe that $d(af + \beta g) = \alpha df + \beta dg$ for $\alpha, \beta$ constant and $d(fg) = (df)g + f(dg)$. A differential $F \cdot dx$ is called exact iff $F \cdot dx = df$ for some $f \Rightarrow F_i = \partial f / \partial x_i \Rightarrow F = \nabla f$. Then

$$\int_C F \cdot dx = \int_C df = f(b) - f(a).$$
1.3.5 Comments

There are other kinds of line integrals, e.g. $\int_c F \, ds$ or $\int_c f \, dx$ which are vector valued, but they can always be reduced to cases above by taking components.

May need to evaluate integrals along piecewise smooth curves e.g.

$$\int_{C_1+C_2+C_3} F \cdot dx = \int_{C_1} F \cdot dx + \int_{C_2} F \cdot dx + \int_{C_3} F \cdot dx$$

provided orientations are compatible.

1.4 Surfaces and normals

A smooth surface in $\mathbb{R}^3$ can be defined by a smooth function $f(r) = f(x, y, z) = 0$, or get a family of functions $f(r) = c$ for some constant.

1.4.1 Tangent vectors and grads

Fix $c$ and consider a point $a$ on this surface $f(a) = c$. Let $r(u)$ be any curve in the surface through $a$, so $dr/du$ is the tangent vector to the curve, and so to the surface.

$$\frac{df(r(u))}{du} = 0 = (\nabla f) \cdot \frac{dr}{du}$$

Since this is true for any curve and hence any tangent direction, $\nabla f$ at $a$ is normal to the surface at point $a$.

Example. $f(r) = |r|^2 = x^2 + y^2 + z^2$. $f = c$ gives a sphere for $c > 0$, and $\nabla f = 2(x, y, z) = 2r$ which is indeed normal to the sphere.

In general we also allow a surface to have a boundary, e.g. impose $z \geq 0$ in the example, which gives a hemisphere with boundary a circle.

1.4.2 Parametrisation of a surface

An alternative definition of a surface uses a parametrisation; position vector $r(u, v)$ sweeps out a surface as the parameters $u, v$ change.

Example. $r(\theta, \phi) = (\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta)$ parametrises the unit sphere in terms of the polar angles $\theta, \phi$.

In general for a smooth function $r(u, v)$ we have

$$\delta r = \frac{\partial r}{\partial u} \delta u + \frac{\partial r}{\partial v} \delta v + o\left(\sqrt{\delta u^2 + \delta v^2}\right).$$

For a smooth surface we require $\partial r/\partial u$ and $\partial r/\partial v$ to be linearly independent, i.e.

$$\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \neq 0.$$

This defines a surface. $\partial r/\partial u$ tangent to curve of fixed $v$. $\partial r/\partial v$ tangent to curve of fixed $u$.

For a small rectangle

$$\delta u \delta v$$

is what we get, and the vector area of the parallelogram is

$$\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \delta u \delta v.$$
1.5 Integrals over surfaces

1.5.1 Integral of a function

We can define an integral over a surface $S$ as a limit of a sum by approximating $S$ as a collection of small simple sets e.g. triangles, parallelograms etc., labelled by $I$ and with area $\delta A_i$.

Choose points $x_i$ in each set and define

$$\int_S f(x) \, dA = \lim_{l \to 0} \sum_I f(x_i) \delta A_i$$

where $l$ is the maximum distance between points in some set $I$. This is useful conceptually but we must now refine it to give a method of calculation.

1.5.2 Integrals over subsets of $\mathbb{R}^2$

Consider the special case when the surface is a subset $D$ of $\mathbb{R}^2$. We can relate the concept in 1.5.1 to successive integration over $x$ and $y$ in $\mathbb{R}^2$. Approximate $D$ by rectangles of size $\delta x \times \delta y$, i.e. area $\delta A = \delta x \delta y$.

Taking $\delta y \to 0$ with $\delta x$ fixed gives a narrow strip at $x$ fixed which contributes $\delta x \int_{y_x} f(x,y) \, dy$. Then summing over strips with $\delta x \to 0$ gives

$$\int_D f(x,y) \, dA = \int_{y} \left( \int_{x_y} f(x,y) \, dx \right) \, dy.$$  

But we can also interchange the roles of $x$ and $y$, so

$$\int_D f(x,y) \, dA = \int_{x} \left( \int_{y_x} f(x,y) \, dy \right) \, dx.$$  

Example. Integrate $f(x,y) = xy$ over the domain $D$ shown.

$$\int_D f(x,y) \, dA = \int_{0}^{2} \left( \int_{0}^{1-x/2} xy \, dy \right) \, dx = \int_{0}^{2} \frac{1}{6}x \left( \frac{1}{2} - \frac{1}{2}x \right)^2 \, dx = \frac{1}{6}.$$  

Or,

$$\int_D f(x,y) \, dA = \int_{0}^{1} \left( \int_{0}^{2-2y} xy \, dx \right) \, dy = \int_{0}^{1} \frac{1}{2}y(2-2y)^2 \, dy = \frac{1}{6}.$$  

1.5.3 Changes of variables

Suppose we have a smooth, invertible transformation between $(x,y)$ and $(u,v)$ with $D$ and $D'$ subsets in one-to-one correspondence as shown. Refer to section 1.4, but now with

$$r(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ 0 \end{pmatrix}.$$  

Corresponding to any small rectangle $\delta u \delta v$ in the $uv$-plane, we have a small parallelogram in the $xy$-plane, with area

$$\delta A = \left| \frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v} \right| \delta u \delta v.$$  

Now
where
\[ \frac{\partial x}{\partial u} \times \frac{\partial x}{\partial v} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \]

is called the Jacobian.

Summing over parallelograms in \( D \) is equivalent to summing over rectangles in \( D' \), with the additional Jacobian factor included, and in the limit, \( \delta u, \delta v \to 0 \). Hence
\[ \int_{D} f(x, y) \, dx \, dy = \int_{D'} f(x(u, v), y(u, v), \|J\|) \, du \, dv \]
is the change of variable formula.

Note that the modulus of the Jacobian appears here. Also,
\[ f\left(\frac{x, y}{u, v}\right) = \frac{1}{f\left(\frac{u, v}{x, y}\right)} . \]

**Example.** Plane polar coordinates \((\rho, \varphi)\) (as in the introduction) so \((x, y) = (\rho \cos \varphi, \rho \sin \varphi)\) and \( f = e^{-\frac{1}{2}(x^2+y^2)} = e^{-\frac{1}{2}\rho^2} \) integrated over \( D : x, y \geq 0 \iff D' : \rho \geq 0, 0 \leq \varphi \leq \frac{\pi}{2} \). The Jacobian
\[ \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = \rho. \]

Then
\[ \int_{D} f(x, y) \, dx \, dy = \int_{0}^{\infty} dx \int_{0}^{\infty} dy \, e^{-\frac{1}{2}(x^2+y^2)} = \int_{0}^{\infty} d\rho \int_{0}^{\pi/2} d\varphi \rho e^{-\frac{1}{2}\rho^2} = \frac{\pi}{2} \left[ -e^{-\frac{1}{2}\rho^2} \right]_{0}^{\infty} = \frac{\pi}{2}. \]

But in this case
\[ \text{LHS} = \left( \int_{0}^{\infty} dx \, e^{-\frac{1}{2}x^2} \right)^2 \implies \int_{0}^{\infty} e^{-\frac{1}{2}x^2} \, dx = \sqrt{\frac{\pi}{2}}. \]

**1.5.4 Scalar surface integral of a vector field**

On a surface \( S \) the unit normal \( \mathbf{n} \) (with \(|\mathbf{n}| = 1\)) is unique up to a sign at each point. If a smoothly-varying normal can be chosen at all points of \( S \), then \( S \) is called orientable and there are two possible orientations corresponding to \( + \) ambiguity in choice of \( \mathbf{n} \). (It is possible to have non-orientable surfaces but we will not deal with these. E.g. Möbius band.)

E.g. for a sphere,
outward normal \( \rightarrow \) inward normal \( \leftarrow \)

Once we have chosen a normal field, or orientation, for \( S \), we can use it to integrate a vector field \( \mathbf{F}(\mathbf{x}) \) over \( S \) and produce a scalar;
\[ \int_{S} \mathbf{F} \cdot d\mathbf{A} = \int_{S} \mathbf{F} \cdot \mathbf{n} \, d\mathbf{A} \]

Definition as limit of a sum can be given by triangulating \( S \) as in 1.5.1 but using vector area; \( \delta A_i = \delta A_i \mathbf{n} \) for each set in the triangulation. Then (all notation as before)
\[ \int_{S} \mathbf{F} \cdot d\mathbf{A} = \lim_{l \to \infty} \sum_{i} \mathbf{F}(\mathbf{x}_i) \cdot \delta A_i \]

Physical interpretation; consider \( \mathbf{v}(\mathbf{x}) \) velocity field of a fluid. In a small time \( \delta t \), volume of fluid crossing area \( \delta A \) is volume of cylinder shown, i.e. \( \delta t \mathbf{v} \cdot \delta A \implies \text{volume crosses surface } S \text{ at a rate } \int_{S} \mathbf{v} \cdot d\mathbf{A} \).
To evaluate such integrals, consider the surface parametrised as \( \mathbf{r}(u, v) \) so that for small variations \( \delta u \) and \( \delta v \) the corresponding area is given approximately by parallelogram

\[
\delta A = \left( \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) \delta u \delta v
\]

(just as in 1.4 but now using vector area). By usual limiting argument we have

\[
\int_S \mathbf{F} \cdot d\mathbf{A} = \int_{S'} \mathbf{F} \cdot \left( \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) du \, dv
\]

where \( S' \) is the region in the \( uv \)-plane corresponding to \( S \).

**Example.** \( S \) hemisphere of radius \( a \), so

\[
\mathbf{r}(\theta, \varphi) = \begin{pmatrix} a \cos \varphi \sin \theta \\ a \sin \varphi \sin \theta \\ a \cos \theta \end{pmatrix} = a \mathbf{e}_r
\]

for \( 0 \leq \theta \leq \pi/2 \) and \( 0 \leq \varphi \leq 2\pi \). Then

\[
\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} = a^2 (\sin \theta) \mathbf{e}_r
\]

Vector field \( \mathbf{F}(\mathbf{r}) = (y, -x, z) \) then

\[
\mathbf{F}(\mathbf{r}) = \frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \varphi} = a^3 \sin \theta \cos^2 \theta.
\]

\[
\Rightarrow \int_S \mathbf{F} \cdot d\mathbf{A} = \int_0^{2\pi} d\varphi \int_0^{\pi/2} d\theta \ a^3 \sin \theta \cos^2 \theta = \frac{2\pi}{3} a^3.
\]

### 1.6 Integration over volumes

**1.6.1 Basic concepts**

In now familiar fashion we can define the integral of a function \( f(x) \) over a region \( V \) in \( \mathbb{R}^3 \) as the limit of a sum;

\[
\int_V f(x) \, dV = \lim_{I \rightarrow 0} \sum_i f(x_i) \, \delta V_i
\]

after dividing \( V \) up into small regions with volume \( \delta V_i \) etc.

We can relate this definition to successive integration over \( x, y, z \) by taking small regions to be cuboids \( \delta x, \delta y, \delta z \). Letting \( \delta x, \delta y \rightarrow 0 \) with \( \delta z \) fixed gives a contribution to the sum

\[
\delta z \int_{D_z} f(x, y, z) \, dx \, dy
\]

= \( \delta z \) (area integral over region \( D_z \) in \( xy \)-plane at fixed \( z \))

Finally \( \delta z \rightarrow 0 \) gives

\[
\int_V f(x, y, z) \, dV = \int_Z \left( \int_{D_z} f(x, y, z) \, dx \, dy \right) \, dz
\]

with range \( Z \) as shown in sketch.

E.g. volume integral in physics \( \int_V \rho(\mathbf{r}) \, dV \) gives total mass in \( V \) if \( \rho(\mathbf{r}) \) is mass density or total charge in \( V \) if \( \rho(\mathbf{r}) \) is charge density.

**1.6.2 Change of variables in \( \mathbb{R}^3 \)**

If a volume \( V \) is parametrised by \( \mathbf{r}(u, v, w) \) then
and the vectors on the RHS must be linearly independent. This means that the Jacobian

\[
\frac{\partial r}{\partial u} \cdot \frac{\partial r}{\partial v} \times \frac{\partial r}{\partial w} = \int \begin{vmatrix} x/\partial_u & y/\partial_u & z/\partial_u \\ x/\partial_v & y/\partial_v & z/\partial_v \\ x/\partial_w & y/\partial_w & z/\partial_w \end{vmatrix} \neq 0.
\]

A small cuboid of volume \(\delta u \delta v \delta w\) in parameter space corresponds to a region approximated by a parallelipiped, as shown, with volume

\[
\delta V = \int \begin{vmatrix} x, y, z \end{vmatrix}_{u, v, w} \delta u \delta v \delta w.
\]

Approximation is exact in the limit \(\delta u, \delta v, \delta w \to 0\), and so summing over cuboids and taking limit gives

\[
\int_V f(r) \, dV = \int_{V'} f(r(u, v, w)) \left| \frac{\partial}{\partial u} \begin{vmatrix} x, y, z \end{vmatrix}_{u, v, w} \right| \, du \, dv \, dw
\]

where \(V\) and \(V'\) are in one-to-one correspondence.

### 1.6.3 Generalisation to \(\mathbb{R}^n\)

Integration over region \(D\) in \(\mathbb{R}^n\) amounts to successive integration over \(x_1, x_2, \ldots, x_n\). To change variables to \(u_1, u_2, \ldots, u_n\) in some region \(D'\) use

\[
\int_D f(x) \, dx_1 \ldots dx_n = \int_{D'} f(x(u)) |J| \, du_1 \ldots du_n
\]

where

\[
J = \det \left( \frac{\partial x_i}{\partial u_j} \right)
\]

is an \(n \times n\) determinant. This follows from the relation

\[
\delta x_i = \frac{\partial x_i}{\partial u_j} \delta u_j + o(\delta u)
\]

using standard interpretation of \(\det\).

### 1.7 Orthogonal curvilinear coordinates on \(\mathbb{R}^3\)

As above, consider region in \(\mathbb{R}^3\) parametrised by \(r(u, v, w)\). Use differential notation; replace \(\delta\) by \(d\) and drop all \(o(\delta)\) terms. Then

\[
dr = \frac{\partial r}{\partial u} du + \frac{\partial r}{\partial v} dv + \frac{\partial r}{\partial w} dw = h_u e_u \, du + h_v e_v \, dv + h_w e_w \, dw
\]

where \(e_u, e_v, e_w\) are unit vectors and \(h_u, h_v, h_w\) are positive scalars.

If the unit vectors \(e_u, e_v, e_w\) are orthogonal (rather than just linearly independent) then \(u, v, w\) are called orthogonal curvilinear coordinates.

Arc length is then given by line element;

\[
ds^2 = \frac{\partial r}{\partial u} \cdot dr = h_u^2 du^2 + h_v^2 dv^2 + h_w^2 dw^2
\]

If we choose the order appropriately, we can assume that we have a 'right-handed' orthonormal basis \((e_w = e_u \times e_v)\) and then

\[
\frac{\partial r}{\partial u} \cdot \frac{\partial r}{\partial v} \times \frac{\partial r}{\partial w} = h_u h_v h_w.
\]

Then change of variable formula involves the volume element
For our purposes we focus on two examples.

1.7.1 Cylindrical polars

\[ \mathbf{r}(\rho, \varphi, z) = \left( \begin{array}{c} \rho \cos \varphi \\ \rho \sin \varphi \\ z \end{array} \right), \quad \rho \geq 0, \quad 0 \leq \varphi \leq 2\pi \]

\[ \mathbf{d}r = \mathbf{e}_\rho \, d\rho + \rho \mathbf{e}_\varphi \, d\varphi + \mathbf{e}_z \, dz \]

\[ h_\rho = h_z = 1, \quad h_\varphi = \rho \]

\[ dV = \rho \, d\rho \, d\varphi \, dz \]

using various formulae. Also common to write \( \hat{\rho}, \hat{\varphi}, \hat{z} \) for basis vectors.

This is plane polars with area element \( dA = \rho \, d\rho \, d\varphi \) and a coordinate \( z \) added.

1.7.2 Spherical polars

\[ \mathbf{r}(r, \theta, \varphi) = \left( \begin{array}{c} r \sin \theta \cos \varphi \\ r \sin \theta \sin \varphi \\ r \cos \theta \end{array} \right), \quad r \geq 0, \quad 0 \leq \varphi \leq 2\pi, \quad 0 \leq \theta \leq \pi \]

\[ \mathbf{d}r = \mathbf{e}_r \, dr + r \mathbf{e}_\theta \, d\theta + (r \sin \theta) \mathbf{e}_\varphi \, d\varphi \]

\[ h_r = 1, \quad h_\theta = r, \quad h_\varphi = r \sin \theta \]

\[ dV = r^2 \sin \theta \, dr \, d\theta \, d\varphi \]

Also common to write \( \hat{r}, \hat{\theta}, \hat{\varphi} \) for unit vectors.

1.8 Geometry of curves and surfaces

Consider a curve \( \mathbf{x}(s) \) parametrised by arc length \( s \) and use dash for \( d/ds \). Then \( \mathbf{t} = \mathbf{x}'(s) \) is unit tangent vector at each point on the curve.

Let \( \mathbf{t}'(s) = \kappa(s) \mathbf{n}(s) \), and this defines a unit vector \( \mathbf{n}(s) \) called the principal normal and a function \( \kappa(s) \) called the curvature, unique up to signs, provided \( \mathbf{t}'(s) \neq 0 \).

Now \( \mathbf{t}^2 = 1 \) so \( \mathbf{t} \cdot \mathbf{t}' = 0 \), so \( \mathbf{n} \) is perpendicular to \( \mathbf{t} \). Define \( \mathbf{b}(s) = \mathbf{t}(s) \times \mathbf{n}(s) \) called the binormal, so giving a right handed set of orthonormal vectors \( \mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s) \) at each point on the curve.

Similar argument to above; \( \mathbf{n}^2 = 1 \Rightarrow \mathbf{n} \cdot \mathbf{n}' = 0 \Rightarrow \mathbf{n}' \) perpendicular to \( \mathbf{n} \), and \( \mathbf{b}^2 = 1 \Rightarrow \mathbf{b}' \) perpendicular to \( \mathbf{b} \). But also \( \mathbf{t} \cdot \mathbf{n} = 0 \Rightarrow \mathbf{t}' \cdot \mathbf{n} + \mathbf{t} \cdot \mathbf{n}' = 0 \Rightarrow \mathbf{t} \cdot \mathbf{n}' = -\kappa \). Hence \( \mathbf{n}' = -\kappa \mathbf{t} + \tau \mathbf{b} \) which defines the torsion \( \tau(s) \).

Continuing; \( \mathbf{b} \cdot \mathbf{n} = 0 \Rightarrow \mathbf{b}' \cdot \mathbf{n} + \mathbf{b} \cdot \mathbf{n}' = 0 \Rightarrow \mathbf{n} \cdot \mathbf{b}' = \tau \) (from above). But also \( \mathbf{b} \cdot \mathbf{t} = 0 \Rightarrow \mathbf{b}' \cdot \mathbf{t} + \mathbf{b} \cdot \mathbf{t}' = 0 \Rightarrow \mathbf{b}' \) perpendicular to \( \mathbf{t} \), since \( \mathbf{b} \cdot \mathbf{t}' = 0 \) follows from the definition of \( \mathbf{n} \). Hence \( \mathbf{b}' = -\tau \mathbf{n} \).

In summary; \( \mathbf{t}' = \kappa \mathbf{n}, \quad \mathbf{n}' = -\kappa \mathbf{t} + \tau \mathbf{b}, \quad \mathbf{b}' = -\tau \mathbf{n} \) (the Frenet-Serret formulae). These are 1st-order differential equations for \( \mathbf{t}, \mathbf{n}, \mathbf{b} \) that allow the curve to be described by \( \kappa(s) \) and \( \tau(s) \) up to translation and rotation.

1.8.1 Interpretation of curvature and torsion

To interpret \( \kappa \) and \( \tau \), consider Taylor expansion of \( \mathbf{x}(s) \) about \( s = 0 \). Let \( \mathbf{t}, \mathbf{n}, \mathbf{b}, \kappa, \tau \) and their derivatives be understood to be evaluated at \( s = 0 \). Then Taylor’s theorem to 2nd order gives

\[ \mathbf{x}(s) = \mathbf{x}(0) + st + \frac{1}{2} s^2 \kappa \mathbf{n} + O(s^3) \]
Compare with a circle of radius $\rho$ in the $t, n$ plane; $a + \rho(1 - \cos \theta)n + (\rho \sin \theta)t$ (choice of centre to give convenient comparison with the curve). For $\theta$ small, this becomes

$$a + \frac{1}{2} \rho \theta^2 n + \rho \theta t + O(\theta^3) = a + \frac{1}{2} s^2 n + s t + O(s^3)$$

where $s = \rho \theta$ is the arc length. $a$ position vector on circumference.

Comparison shows curve and circle match to this order, so $\rho = 1/\kappa$, the radius of curvature.

To next order, we get a Taylor expansion

$$x(s) = x(0) + \left( s - \frac{1}{6} \kappa^2 s^3 \right) t + \left( \frac{1}{2} \kappa s^2 + \frac{1}{6} \kappa' s^3 \right) n + \frac{1}{6} \kappa \tau s^3 b + O(s^4)$$

Thus torsion controls how curve leaves $t, n$ plane.

### 1.8.2 Surfaces (non-examinable remarks)

Given a surface, at any point we choose a plane containing the normal and in this plane the surface becomes a curve. For each plane through this point we get a different curve, and we can compute the curvature $\kappa$ of each one. It can be shown that $\kappa_1 \leq \kappa \leq \kappa_2$ where $\kappa_1, \kappa_2$ are called principal curvatures.

The Gaussian curvature is defined by $K = \kappa_1 \kappa_2$. Gauss’s remarkable theorem (Theorema Egregium) is that $K$ is intrinsic to the surface – depends only on measuring lengths and angles in the surface.

## 2. Vector differential operators and integral theorems

### 2.1 Grad, div, curl and $\nabla$

#### 2.1.1 Definitions

The gradient or grad of a smooth function $f(x)$ is defined by

$$df = (\nabla f) \cdot dx \Leftrightarrow \nabla f = e_i \frac{\partial f}{\partial x_i} = i \frac{\partial f}{\partial x} + j \frac{\partial f}{\partial y} + k \frac{\partial f}{\partial z} = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

using Cartesian coordinates with basis vectors obeying $e_i \cdot e_j = \delta_{ij}$ and $e_i \times e_j = \varepsilon_{ijk} e_k$.

We can regard grad of a function as arising from applying

$$\nabla = e_i \frac{\partial}{\partial x_i} = i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z}$$

which is both a differential operator and a vector.

Now $\nabla$ can be applied to smooth vector field $F(x)$ and combined with scalar or vector product. We define

- divergence (or div) of $F$, a scalar field:

$$\nabla \cdot F = \left(e_i \frac{\partial F_i}{\partial x_i} \right) = \frac{\partial F_i}{\partial x_i}$$

- curl of $F$, a vector field:

$$\nabla \times F = \left[ e_1 \frac{\partial}{\partial x_1} F_1 \frac{\partial}{\partial x_2} F_2 \frac{\partial}{\partial x_3} F_3 \right]$$

Shorthand form for curl:

$$\varepsilon_{ijk} \frac{\partial}{\partial x_i} F_j$$
Example. \( \mathbf{F} = xe^z \mathbf{i} + y^2 \sin x \mathbf{j} + xyz \mathbf{k} \)

\[
\nabla \cdot \mathbf{F} = \frac{\partial}{\partial x} (xe^z) + \frac{\partial}{\partial y} (y^2 \sin x) + \frac{\partial}{\partial z} (xyz) = e^z + 2y \sin x + xy
\]

\[
\nabla \times \mathbf{F} = \left( \frac{\partial}{\partial y} (xyz) - \frac{\partial}{\partial z} (y^2 \sin x) \right) \mathbf{i} + \left( \frac{\partial}{\partial z} (xe^z) - \frac{\partial}{\partial x} (xyz) \right) \mathbf{j} + \left( \frac{\partial}{\partial x} (y^2 \sin x) - \frac{\partial}{\partial y} (xe^z) \right) \mathbf{k}
\]

\[
= xz \mathbf{i} + (xe^z - yz) \mathbf{j} + y^2 \cos x \mathbf{k}
\]

Note that, by contrast,

\[
\mathbf{F} \cdot \nabla = F_i \frac{\partial}{\partial x_i}, \quad \mathbf{F} \times \nabla = \epsilon_{kij} F_i \frac{\partial}{\partial x_j}.
\]

These are now scalar and vector operators. E.g. with \( \mathbf{F} \) as above,

\[
\mathbf{F} \cdot \nabla = xe^z \frac{\partial}{\partial x} + y^2 \sin x \frac{\partial}{\partial y} + xyz \frac{\partial}{\partial z}
\]

If \( \mathbf{n} \) is some unit vector (constant) then \( \mathbf{n} \cdot \nabla \) gives the directional derivative on functions;

\[
(\mathbf{n} \cdot \nabla) f = n_i \frac{\partial}{\partial x_i} (f) = \mathbf{n} \cdot (\nabla f)
\]

Using this notation we can express Taylor’s Theorem very compactly;

\[
f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + (\mathbf{h} \cdot \nabla) f \bigg|_a + \frac{1}{2} (\mathbf{h} \cdot \nabla)^2 f \bigg|_a + \cdots + \frac{1}{r!} (\mathbf{h} \cdot \nabla)^r f \bigg|_a + O(h^{r+1})
\]

To compare with previous version, rewrite term-by-term. E.g.

\[
(h \cdot \nabla)^2 f = \left( h_i \frac{\partial}{\partial x_i} \right) \left( h_j \frac{\partial}{\partial x_j} \right) f = h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}
\]

2.1.2 Properties and simple examples

Grad, div and curl are linear operators;

\[
\nabla (\alpha \varphi + \beta \psi) = \alpha \nabla \varphi + \beta \nabla \psi
\]

\[
\nabla \cdot (\alpha \mathbf{F} + \beta \mathbf{G}) = \alpha \nabla \cdot \mathbf{F} + \beta \nabla \cdot \mathbf{G}
\]

\[
\nabla \times (\alpha \mathbf{F} + \beta \mathbf{G}) = \alpha \nabla \times \mathbf{F} + \beta \nabla \times \mathbf{G}
\]

where \( \alpha, \beta \) are any real numbers, \( \varphi, \psi \) smooth scalar fields and \( \mathbf{F}, \mathbf{G} \) smooth vector fields.

These operators are also first order and so satisfy Leibniz identities. Some have simple structure;

\[
\nabla (\varphi \psi) = (\nabla \varphi) \psi + \varphi (\nabla \psi)
\]

\[
\nabla \cdot (\varphi \mathbf{F}) = (\nabla \varphi) \cdot \mathbf{F} + \varphi \nabla \cdot \mathbf{F}
\]

\[
\nabla \times (\varphi \mathbf{F}) = (\nabla \varphi) \times \mathbf{F} + \varphi \nabla \times \mathbf{F}
\]

These (and others below) can be justified using components. E.g.

\[
(\nabla \times (\varphi \mathbf{F}))_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} (\varphi F_k) = \epsilon_{ijk} \left( \frac{\partial \varphi}{\partial x_j} F_k + \varphi \frac{\partial F_k}{\partial x_j} \right) = \left( (\nabla \varphi) \times \mathbf{F} \right)_i + (\varphi \nabla \times \mathbf{F})_i
\]

Other identities are more involved, e.g.

\[
\nabla \cdot (\mathbf{F} \times \mathbf{G}) = (\nabla \times \mathbf{F}) \cdot \mathbf{G} - \mathbf{F} \cdot (\nabla \times \mathbf{G})
\]

(compare with \( \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} = \mathbf{b} \cdot \mathbf{c} \times \mathbf{a} - \mathbf{a} \cdot \mathbf{b} \times \mathbf{c} \)). Again with components,

\[
\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \frac{\partial}{\partial x_i} (F \times G)_i = \frac{\partial}{\partial x_i} \left( \epsilon_{ijk} F_j G_k \right) = \epsilon_{ijk} \left( \frac{\partial F_j}{\partial x_i} G_k + F_j \frac{\partial G_k}{\partial x_i} \right) = (\nabla \times \mathbf{F})_k G_k - F_j (\nabla \times \mathbf{F})_j
\]

as required.

Finally, the most elaborate are;

\[
\nabla \times (\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \nabla) \mathbf{F} - (\mathbf{F} \cdot \nabla) \mathbf{G} + \mathbf{F} (\nabla \cdot \mathbf{G}) - \mathbf{G} (\mathbf{F} \cdot \nabla)
\]
1. \( \nabla f(r) \) where \( r^2 = x_ix_i \). Apply definition;
\[
e_i \frac{\partial f}{\partial x_i} = e_i \frac{\partial r}{\partial x_i} f'(r)
\]
But \( \frac{\partial r}{\partial x_i} = x_i/r \). So
\[
\nabla f(r) = e_i \frac{x_i}{r} f'(r) = \frac{f'(r)}{r} \mathbf{x}
\]
2. \( \nabla \cdot \mathbf{x} = \frac{\partial x_i}{\partial x_i} = \delta_{ii} = 3 \), and
\[
(\nabla \times \mathbf{x})_i = \epsilon_{ijk} \frac{\partial x_k}{\partial x_j} = \epsilon_{ijk} \delta_{jk} = 0.
\]
3. \( \nabla \times ((\mathbf{a} \cdot \mathbf{x}) \mathbf{x}) = \nabla (\mathbf{a} \cdot \mathbf{x}) \times \mathbf{x} + (\mathbf{a} \cdot \nabla) \times \mathbf{x} = \mathbf{a} \times \mathbf{x}
\]
4. \( \nabla \times (f(r) \mathbf{x}) = (\nabla f) \times \mathbf{x} + f(\nabla \times \mathbf{x}) = (f'(r)/r) \mathbf{x} \times \mathbf{x} = \mathbf{0}
\]

### 2.1.3 \( \nabla \) as a vector and cartesian coordinates

Cartesian coordinates are not unique, but any two sets \( x_i \) and \( x'_i \) with basis vectors \( e_i \) and \( e'_i \) are related by \( x'_i = R_{ij} x_j \) and \( e'_i = R_{ij} e_j \) where \( R \) is orthogonal. These imply that \( \mathbf{x} = x_i e_i = x'_i e'_i \) because \( R_{ik} R_{jk} = \delta_{ij} \).

For any vector \( \mathbf{v} = v_i e_i \) we have \( v'_i = R_{ij} v_j \Rightarrow v = v'_i e'_i \) and this can be taken as a definition of what we mean by a vector. Does \( \nabla \) (del) satisfy this?

Note that \( x'_i = R_{ij} x_j \Rightarrow x_i = (R^{-1})_{ij} x'_j = R_{ji} x'_i \). Now by chain rule, the components of \( \nabla \), namely \( \partial / \partial x_i \) and \( \partial / \partial x'_i \), are related by
\[
\frac{\partial}{\partial x'_i} = \frac{\partial x_j}{\partial x'_i} \frac{\partial}{\partial x_j} R_{ij} \frac{\partial}{\partial x'_j} = \nabla = \frac{\partial}{\partial x'_i} e_i = \frac{\partial}{\partial x'_i} e'_i.
\]

### 2.1.4 Double derivatives involving \( \nabla \)

There are a number of ways of applying \( \nabla \), \( \nabla \cdot \) and \( \nabla \times \) successively and two of these always give 0; \( \nabla \times \nabla \cdot \mathbf{f} = 0 \) on any scalar field \( f \), and \( \nabla \cdot \nabla \times \mathbf{f} = 0 \) on any vector field \( \mathbf{f} \). To justify these statements,
\[
\nabla \times (\nabla f) = e_i e_{ijk} \frac{\partial}{\partial x_j} (\nabla f)_k = e_i e_{ijk} \frac{\partial^2 f}{\partial x_j \partial x_k} = 0
\]
by (anti)symmetry on \( j \) and \( k \). Similarly,
\[
\nabla \cdot (\nabla \times \mathbf{A}) = \frac{\partial}{\partial x_i} (\nabla \times \mathbf{A})_i = e_{ijk} \frac{\partial A_k}{\partial x_i} \frac{\partial}{\partial x_j} = 0.
\]

In fact, \( \nabla \times \mathbf{F} = 0 \iff \mathbf{F} = \nabla f \) for some \( f \) although the region in which \( f \) is single-valued may be smaller than the region where \( \mathbf{F} \) is defined. Similarly, \( \nabla \cdot \mathbf{B} = 0 \iff \mathbf{B} = \nabla \times \mathbf{A} \iff \mathbf{B} \) is solenoidal.

\( \mathbf{F} \) is called conservative or irrotational if \( \nabla \times \mathbf{F} = 0 \) and \( f \) is called the scalar potential. Sometimes a sign is included; e.g. if \( \mathbf{F} \) is a force with \( \nabla \times \mathbf{F} = 0 \), then we write \( \mathbf{F} = -\nabla \varphi \) with \( \varphi \) the potential energy. The definition of ‘conservative’ above agrees with the earlier definition of line integrals.

We will justify \( \nabla \times \mathbf{F} = 0 \Rightarrow \mathbf{F} = \nabla f \) below using Stokes’ theorem. In fluid mechanics, if \( \mathbf{u} \) is an irrotational velocity field, then \( \mathbf{u} = \nabla \varphi \) defines the velocity potential \( \varphi \). Similarly if \( \mathbf{B} = \nabla \times \mathbf{A} \) we call \( \mathbf{A} \) the vector potential.

Other combinations of grad, div and curl are important;
\[
\nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}
\]
is called the Laplacian. On scalars, $\nabla^2 \varphi = (\nabla \cdot \nabla \varphi)$ and $\nabla^2 A = \nabla(\nabla \cdot A) - \nabla \times (\nabla \times A)$, and these can be proved using components.

2.2 Integral theorems

These closely related results all generalise the Fundamental Theorem of Calculus. In this section we give statements, and return to proofs later.

2.2.1 Green's theorem (in the plane)

For smooth functions $P(x, y)$ and $Q(x, y)$,

$$\int_A \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \int_C (P \, dx + Q \, dy)$$

where $A$ is a region in the $xy$-plane within the boundary $\partial A = C$, a piecewise smooth curve which is traversed anticlockwise.

E.g. $P = x^2 y$, $Q = xy^2$ implies

$$\int_A (y^2 - x^2) dA = \int_C x^2 y \, dx + xy^2 \, dy$$

and for $C$ as shown, each side is $104a^4/105$.

Theorem also holds if boundary $C$ is not connected, provided interior boundaries are traversed clockwise. This can be related to the piecewise case by the following construction;

The contributions along the two coincident line segments cancel out.

2.2.2 Stokes' theorem

For a smooth vector field $F(x)$,

$$\int_S \nabla \times F \cdot dS = \int_C F \cdot dx$$

where $S$ is a smooth surface with boundary $\partial S = C$, a piecewise smooth closed curve and with compatible orientations as shown;

In detail, $n$ (which is the normal to $S$ which defines the area element $dS = n \, dS$), and $t$ (which is tangent to the curve) are related by $t \times n$ pointing outward from $S$.

Example. $S$ a hemispherical surface with $x^2 + y^2 + z^2 = a^2, z \geq 0$. $F = (0, x, 0)$ so $\nabla \times F = (0, 0, 1)$. On $S$, $dS = n \, dS = \frac{1}{a} (x, y, z) \, dS$. Then

$$\nabla \times F \cdot dS = \frac{1}{a} z \, dS = a^2 \cos \theta \sin \theta \, d\theta \, d\varphi$$

$$\int_S \nabla \times F \cdot dS = 2\pi a^2 \int_0^{\pi/2} \cos \theta \sin \theta \, d\theta = \pi a^2$$

$$\int_C F \cdot dx = \int_0^{2\pi} a^2 \cos^2 \varphi \, d\varphi = \pi a^2$$

Thus Stokes' theorem holds.

Disconnected boundaries. For Stokes', as with Green's theorem, the boundary may be disconnected, and we can use a similar construction to relate this situation to the case of a connected boundary.
With curves as shown,
\[ \int_C F \cdot dx = \int_{C_1} F \cdot dx + \int_{C_2} F \cdot dx + \int_{C_3} F \cdot dx \]
because integrals along parallel segments are equal.

Prescription given above for compatible orientations now applies to all boundaries. Informally, travelling around \( C \) in direction given by \( \mathbf{t} \), and with \( \mathbf{n} \) defining ‘up’, then surface is on the left.

**Example.** \( S \) part of hemisphere with \( \alpha \leq \theta \leq \pi/2 \); radius \( a \) as before, and \( F = \begin{pmatrix} 0 \\ x \\ 0 \end{pmatrix} \) so \( \nabla \times F = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \). Then
\[
\int_S \nabla \times F \cdot dS = 2\pi a^2 \int_0^\pi \cos \theta \sin \theta \, d\theta = \pi a^2 (1 - \sin^2 \alpha)
\]
Now \( dS = C + C' \), and on \( C \), \( x = a \begin{pmatrix} \sin \alpha \cos \varphi \\ \sin \alpha \sin \varphi \\ \cos \alpha \end{pmatrix} \) from \( \varphi = 2\pi \) to 0; \( \theta = \alpha \). Then
\[
\int_C F \cdot dx + \int_{C'} F \cdot dx = \pi a^2 - \pi(a \sin \alpha)^2
\]
as required.

### 2.2.3 Divergence (or Gauss’s) Theorem

For a smooth vector field \( \mathbf{F}(x) \),
\[
\int_V \nabla \cdot \mathbf{F} \, dV = \int_S \mathbf{F} \cdot dS
\]
where \( V \) is a volume with boundary \( \partial V = S \), a piecewise smooth closed surface with \( \mathbf{n} \) the outward normal (i.e. smooth surfaces intersecting in piecewise smooth curves, such as the surface of a cube).

**Example.** \( V \) a solid hemisphere, \( x^2 + y^2 + z^2 \leq a^2 \) and \( z \geq 0 \). Boundary \( S = S_1 + S_2 \), hemisphere and disc.

Consider \( F = \begin{pmatrix} 0 \\ 0 \\ z + a \end{pmatrix} \) and \( \nabla \cdot F = 1 \).

Then
\[
\int_V \nabla \cdot F \, dV = \frac{2}{3} \pi a^3
\]
On \( S_1 \); \( dS = n \, dS = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \frac{1}{a} \, dS \)
and \( F \cdot dS = z(z + a)/a \, dS = a^3 \cos \theta (\cos \theta + 1) \sin \theta \, d\theta \, d\phi \)
\[
\Rightarrow \int_{S_1} F \cdot dS = 2\pi a^3 \int_0^\pi (\cos \theta + \cos^2 \theta) \sin \theta \, d\theta = \frac{5}{3} a^3
\]
On \( S_2 \), \( dS = n \, dS = -\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \, dS \) so \( F \cdot dS = -a \, dS \) and hence
\[
\int_{S_2} F \cdot dS = -\pi a^3
\]
\[
\Rightarrow \int_{S_1} \mathbf{F} \cdot dS + \int_{S_2} \mathbf{F} \cdot dS = \frac{2}{3} a^3
\]

as required.

### 2.2.4 Links between Green’s, Stokes’ and the divergence theorems

Let \( S \) be a surface in the \( xy \)-plane, with boundary \( C = \partial S = (x(s), y(s)) \) parametrised by arc length \( s \). Then

\[
\mathbf{t} = \left( \frac{dx}{ds}, \frac{dy}{ds}, 0 \right)
\]

is the unit tangent to \( C \). Let

\[
\mathbf{n} = \left( \frac{dy}{ds}, -\frac{dx}{ds}, 0 \right)
\]

be a unit normal to \( C \). Given \( P(x, y), Q(x, y) \), define

\[
\mathbf{F} = \begin{pmatrix} P \\ Q \\ 0 \end{pmatrix} \Rightarrow \nabla \times \mathbf{F} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \end{pmatrix}
\]

Then

\[
\mathbf{F} \cdot \mathbf{t} \, ds = P \, dx + Q \, dy
\]

\[
(\nabla \times \mathbf{F}) \cdot \mathbf{k} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}
\]

So Stokes’ theorem implies Green’s theorem.

But alternatively we can define \( \mathbf{G} = \begin{pmatrix} Q \\ -P \\ 0 \end{pmatrix} \) so

\[
\nabla \cdot \mathbf{G} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}
\]

\[
\mathbf{G} \cdot \mathbf{n} \, dS = P \, dx + Q \, dy.
\]

Hence Green’s theorem is equivalent to the statement

\[
\int_{\partial S} \nabla \cdot \mathbf{G} \, dS = \int_{S} \mathbf{G} \cdot \mathbf{n} \, dS
\]

This is exactly the two-dimensional version of the divergence theorem.

### Aside on notation

Sometimes use \( \int \) to denote integration over a closed curve or surface. Also surface integrals sometimes written \( \iint \) and volume integrals \( \iiint \).

### 2.2.5 Extensions of standard theorems

Let \( \mathbf{a} \) be a constant vector, and consider \( \nabla \cdot (\mathbf{a} f) = \mathbf{a} \cdot \nabla f \) for any smooth scalar function \( f(x) \). The divergence theorem gives

\[
\int_{S} f \mathbf{a} \cdot dS = \int_{V} \mathbf{a} \cdot \nabla f
\]

with \( S = \partial V \) as usual. Since \( \mathbf{a} \) is constant,

\[
\mathbf{a} \cdot \int_{S} f \, dS = \mathbf{a} \cdot \int_{V} \nabla f \, dV.
\]

But since \( \mathbf{a} \) is arbitrary,
\[ \int_S f \, dS = \int_V \nabla f \, dV. \]

There are many similar results, e.g.

\[ \int_C f \, dx = -\int_S \nabla f \times dS \]

with \( C = \partial S \).

Another kind of extension is a generalisation of integration by parts. E.g.

\[ \int_V (F \cdot \nabla \varphi) \, dV = \int_V (\nabla \cdot (F \varphi) - \varphi \nabla \cdot F) \, dV = \int_S F \varphi \, \cdot \, dS - \int_V \varphi \nabla \cdot F \, dV. \]

In particular for \( F = \nabla \psi \), the identity above becomes

\[ \int_V (\nabla \psi \cdot \nabla \varphi) \, dV = \int_S \varphi \nabla \psi \cdot dS - \int_V \varphi \nabla^2 \psi \, dV. \]

all with \( S = \partial V \). This is sometimes called Green’s First Identity (or Theorem). Now taking this and combining for the same result with \( \varphi \leftrightarrow \psi \), we get

\[ \int_V (\varphi \nabla^2 \psi - \psi \nabla^2 \varphi) \, dV = \int_S (\varphi \nabla \psi - \psi \nabla \varphi) \cdot dS. \]

This is sometimes called Green’s Second Identity (or Theorem). Can alternatively show this by divergence theorem applied to \( \varphi \nabla \psi - \psi \nabla \varphi \).

These identities will be applied later with \( \partial V = S \) disconnected – just need to be careful with orientations.

### 2.3 Some constructions and comments

#### 2.3.1 Scalar potentials and conservative fields

Given a smooth vector field \( F(x) \) with \( \nabla \times F = 0 \), we can define a scalar potential \( \varphi(x) \) with \( F = -\nabla \varphi \) (sign: mechanics conventions) as follows. Choose a reference point \( x_0 \) at which \( \varphi(x_0) = 0 \), by definition. Then let

\[ \varphi(x) = -\int_C F(x') \cdot dx' \]

with \( C \) some curve from \( x_0 \) to \( x \). (\( x' \) to distinguish variable of integration.)

This depends on \( x \) if \( F \) is defined and smooth on a subset \( D \) of \( \mathbb{R}^3 \) with the property that any two curves \( C_1 \) and \( C_2 \) from \( x_0 \) to \( x \) (any point) can be smoothly deformed into one another, sweeping out some surface \( S \) within \( D \). A set \( D \) with this property is called simply connected.

With these assumptions, \( \partial S = C_1 - C_2 \) (with orientations), and Stokes’ theorem implies

\[ \int_{C_1} F \cdot dx' + \int_{C_2} F \cdot dx' = \int_{\partial S} F \cdot dS = \int_S \nabla \times F \cdot dS = 0. \]

Hence Stokes’ theorem \( \Rightarrow \varphi(x) \) well defined for \( D \) simply connected. Now

\[ \varphi(x + \delta x) - \varphi(x) = \int_{\partial C} F \cdot dx' = -F(x) \cdot \delta x + o(|\delta x|) = \nabla \varphi(x) \cdot \delta x + o(|\delta x|) \]

by the definition of gradient. Since this is true for any \( \delta x \), we have \( -\nabla \varphi = F \) as claimed.

Physically, if \( F \) is a conservative force, the potential is obtained by computing the work done.
Examples. First of\( F\) and\( \varphi\) for\( D\) simply connected;

- \(D = \mathbb{R}^3\) and \(F\) constant \(\Rightarrow \varphi = - F \cdot (x - x_0)\)
- \(D = \mathbb{R}^3 - \{0\}\) and \(F = \alpha x/r^3 \Rightarrow \varphi = -\alpha/r\)

**Example** with \(D\) not simply connected;

\[
D = \mathbb{R}^3 - \{\text{z axis}\}, \quad F = \beta \left( -\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2}, 0 \right)
\]

Then \(\nabla \times F = 0\) and \(F = -\nabla \varphi\) with \(\varphi = \beta \tan^{-1}(y/x) = -\beta x\) where \(x\) is the azimuthal angle in spherical polars (usually called \(\varphi\!\)).

In this case \(D\) is not simply connected, by considering \(C_1, C_2\) as shown. Also \(x\) is not single valued (it’s an angle) unless we restrict to a smaller region \(D_0\). In this case we can choose \(D_0 = \mathbb{R}^3 - \{\text{the plane with } y = 0 \text{ and } x \geq 0\}\).

Now angle is well defined; \(0 < \chi < 2\pi\).

In general, if \(D\) is not simply connected we can always restrict to a subset \(D_0\) which is, e.g. any solid sphere is simply connected.

In summary: in general the conditions (i) \(F = -\nabla \varphi\), (ii) \(\nabla \times F = 0\), and (iii) \(\int_C F \cdot dx \neq 0\)

are independent of path, are exactly equivalent whenever \(D\) is simply connected. If it is not, then they are equivalent on some simply connected subset \(D_0\), but on \(D\) can find that

\[
\int_C F \cdot dx = 0
\]

for closed curve \(C\), even though \(\nabla \times F = 0\), because \(C\) need not be the boundary of a surface \(S\) lying entirely in \(D\). Hence no contradiction with Stokes’ theorem.

### 2.3.2 Interpretation of curl and divergence for fluid flow

A fluid is described by a velocity field \(\mathbf{v} = \mathbf{v}(x)\). Interpretations of \(\nabla \times \mathbf{v}\) and \(\nabla \cdot \mathbf{v}\) arise by comparing with two simple kinds of motion.

Consider a rigid body rotating with angular velocity \(\omega\) about an axis through the point \(x_0\). Then

\[
\dot{x} = \mathbf{v}(x) = \omega \times (x - x_0)
\]

If \(\mathbf{n} \perp x - x_0\) then \(|\mathbf{v}| = |x - x_0||\omega \cdot \mathbf{n}|\). With this \(\mathbf{v}(x)\) we have \(\nabla \times \mathbf{v} = 2\omega\). This suggests that for general \(\mathbf{v}(x)\) we should interpret \(\nabla \times \mathbf{v}\) as twice the local angular velocity of the fluid.

To confirm this, consider a small circle \(S\) with \(|x - x_0| \leq a\) and normal \(\mathbf{n}\).

Average of tangential component of \(\mathbf{v}\) around \(C = \partial S\) is

\[
\frac{1}{2\pi a} \int_C \mathbf{v} \cdot dx = \frac{1}{2\pi a} \int_S \nabla \times \mathbf{v} \cdot dS \approx \frac{1}{2\pi a} (4\pi a^2) (\nabla \times \mathbf{v}) \cdot \mathbf{n}|_{x_0} = \frac{a}{2} (\nabla \times \mathbf{v}) \cdot \mathbf{n}|_{x_0}
\]

by Stokes’ theorem, for \(a\) small. By comparison with rigid body case,

\(|\mathbf{v}| = |x - x_0||\omega \cdot \mathbf{n}|\) we see \(\frac{a}{2} (\nabla \times \mathbf{v}) \cdot \mathbf{n}\) is the local rate of rotation at \(x_0\). About \(n\).

Now consider instead a dilatation about \(x_0\);

\[
\dot{x} = \mathbf{v}(x) = \alpha (x - x_0)
\]

with constant \(\alpha\) being the proportional rate of increase of length. Integrated form;

\[
x(t) - x_0 = e^{\alpha t} (x(0) - x_0)
\]

and so for volumes, \(\mathbf{v}(t) = e^{2\alpha t} \mathbf{v}(0)\) or \(\dot{\mathbf{v}} = 3\alpha \mathbf{v}'\). Now for this \(\mathbf{v}(x)\) we have \(\nabla \cdot \mathbf{v} = 3\alpha\).
For a general velocity field $\mathbf{v}(\mathbf{x})$ we interpret, by comparison, $\nabla \cdot \mathbf{v}$ is three times the local rate of proportional increase in length. We establish this by considering a small sphere $V$ with $|\mathbf{x} - \mathbf{x}_0| = a$.

Average of normal component of $\mathbf{v}$ over $S = \partial V$ is

$$\frac{1}{4\pi a^2} \int_S \mathbf{v} \cdot dS = \frac{1}{4\pi a^2} \int_V \nabla \cdot \mathbf{v} \, dV = \frac{1}{4\pi a^2} (\frac{4}{3} \pi a^3) \nabla \cdot \mathbf{v} |_{\mathbf{x}_0} = a \left(\frac{1}{3} \nabla \cdot \mathbf{v}\right) |_{\mathbf{x}_0}$$

for a small. Compare with dilation above with $|\mathbf{v}| = |\mathbf{x} - \mathbf{x}_0|/a$ and we get $\frac{1}{3} \nabla \cdot \mathbf{v} |_{\mathbf{x}_0}$ the local rate of proportional increase of length.

2.3.3 Conservation laws

General form of conservation equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0$$

where $\rho(\mathbf{x}, t)$ is a scalar and $\mathbf{j}(\mathbf{x}, t)$ is a vector. Let

$$Q(t) = \int_V \rho(\mathbf{x}, t) \, dV$$

with $V$ some volume at fixed time $t$. Then

$$\frac{dQ}{dt} = \int_V \frac{\partial \rho}{\partial t} \, dV = \int_V - (\nabla \cdot \mathbf{j}) \, dV$$

So by the divergence theorem (at fixed $t$),

$$\frac{dQ}{dt} = - \int_S \mathbf{j} \cdot dS$$

where $S = \partial V$. Have a conserved quantity $Q$ (i.e. $dQ/dt = 0$) if $\mathbf{j} = 0$ on $S$, or if $|\mathbf{j}| \to 0$ rapidly enough that $\int_S \mathbf{j} \cdot dS \to 0$ as size of $V$ increases. E.g. $S$ a sphere of radius $R \to \infty$ so $V \to \mathbb{R}^3$.

Examples.

(i) Electromagnetism – conservation of charge. $\rho(\mathbf{x}, t)$ charge density and $\mathbf{j}(\mathbf{x}, t)$ current density. $Q(\mathbf{t})$ total charge in $V$ and $\int_S \mathbf{j} \cdot dS$ is the total charge crossing $S$ per unit time.

(ii) Fluid flow – conservation of mass. $\rho(\mathbf{x}, t)$ mass density and $\mathbf{j} = \rho \mathbf{v}$. Note that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

and if $\rho$ independent of $\mathbf{x}$ and $t$ this becomes $\nabla \cdot \mathbf{v} = 0$ – compare with interpretation above.

2.3.4 Quantum mechanics (non-examinable; strictly for IB)

$\rho(\mathbf{x}, t) = |\Phi(\mathbf{x}, t)|^2$ is probability density for measuring a particle with position $\mathbf{x}$ at time $t$, where $\Phi(\mathbf{x}, t)$ is complex-valued, called the wave function. The probability of finding the particle in volume $V$ at time $t$ is $\int_V \rho(\mathbf{x}, t) \, dV$ and hence

$$\int_{\mathbb{R}^3} \rho(\mathbf{x}, t) \, dV = 1$$

But $\Psi(\mathbf{x}, t)$ obeys Schrodinger’s equation: for a particle of mass $m$ in a potential $U(\mathbf{x})$

$$\frac{\hbar^2}{2m} \nabla^2 \Psi + U(\mathbf{x}) \Psi = i \hbar \frac{\partial \Psi}{\partial t}$$

We can normalise $\Psi$ to get

$$\int_{\mathbb{R}^3} \rho(\mathbf{x}, 0) \, dV = 1$$

(Schrodinger’s equation linear) and then
because of conservation equation with
\[ j = -\frac{i\hbar}{2m}(\bar{\Psi}\nabla\Psi - \Psi\nabla\bar{\Psi}) \]
and result follows for suitable boundary conditions.

2.4 Informal proofs of integral theorems

2.4.1 Green's theorem

We will show (using 2nd version in 2.2.1 with \( G = G_x i + G_y j = Q i - Pf \))
\[ \int_A \left( \frac{\partial G_x}{\partial x} + \frac{\partial G_y}{\partial y} \right) dA = \int_C G \cdot n \, ds \]
where \( n \) is outward normal to \( C = \partial A \). Start with simple shape for \( C \) and \( A \) where for each \( x \) we have a single interval \( Y_x \) of values of \( y \) with \( y \) in \( A \)

More precisely, \( y_+(x) \geq y \geq y_-(x) \). Now consider
\[ \int_A \frac{\partial G_y}{\partial y} \, dA = \int_A dx \int_{Y_y} dy \frac{\partial G_y}{\partial y} = \int_X (G_y(x,y_+(x)) - G_y(x,y_-(x))) \, dx = \int_C n \cdot (G_y j) \]

because \( n \cdot j \, ds = dx \) at \( (x,y_+(x)) \), but \( -dx \) at \( (x,y_-(x)) \).

A similar argument with \( x \) and \( y \) interchanged, shows
\[ \int_A \frac{\partial G_x}{\partial x} \, dA = \int_C n \cdot (G_x f) \]

This proves the result for the simple shape described. To extend it to more general \( A \) and \( C \), we can proceed in a number of ways.

(i) Allow \( Y_x \) to consist of several disjoint intervals, \( y_+^{(a)}(x) \geq y \geq y_-^{(a)}(x) \).
We then integrate over \( x \) expressions like \( G_y \left( x, y_+^{(1)}(x) \right) - G_y \left( x, y_-^{(1)}(x) \right) + G_y \left( x, y_+^{(2)}(x) \right) - G_y \left( x, y_-^{(2)}(x) \right) \). Argument goes through as before because \( n \cdot j \, ds = dx \) at \( (x,y_+^{(a)}(x)) \), but \( -dx \) at \( (x,y_-^{(a)}(x)) \).

(ii) Alternatively – subdivide \( A \) using lines with \( x \) or \( y \) constant, so that in each subregion, any remaining boundary is a monotonic function \( y(x) \) or \( x(y) \). Then previous result applies to each subregion and contributions from all internal lines cancel on summing.

2.4.2 Stokes' theorem

We can reduce the result for a surface \( S \),
\[ \int_S \nabla \times F \, dS = \int_{\partial S} F \cdot dx \]
to Green's theorem for a corresponding region \( A \) in the \( uv \)-plane, given some parametrisation \( x(u,v) \) for \( S \). Let
\[ F_u = \mathbf{F} \cdot \frac{\partial \mathbf{x}}{\partial u}, \quad F_v = \mathbf{F} \cdot \frac{\partial \mathbf{x}}{\partial v} \]

(subscripts not partial derivatives). Then

\[ \frac{\partial F_v}{\partial u} - \frac{\partial F_u}{\partial v} = (\nabla \times \mathbf{F}) \cdot \left( \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right) \]

since each side can be written

\[ \left( \frac{\partial F_v}{\partial x_i} - \frac{\partial F_u}{\partial x_j} \right) \frac{\partial x_i}{\partial u} \frac{\partial x_j}{\partial v} \]

On LHS this comes from the chain rule

\[ \frac{\partial F_v}{\partial u} = \frac{\partial}{\partial u} \left( F_j \frac{\partial x_j}{\partial v} \right) = \frac{\partial F_j}{\partial x_i} \frac{\partial x_i}{\partial u} \frac{\partial x_j}{\partial v} + F_j \frac{\partial^2 x_j}{\partial u \partial v} \]

and similarly with \( u \leftrightarrow v \). On RHS use standard vector identities.

On boundary we have \( \mathbf{F}(\mathbf{x}) \) with \( \mathbf{x}(u, v) \) and \( u(t), v(t) \) for some parameter \( t \). Hence

\[ \mathbf{F} \cdot \frac{d\mathbf{x}}{dt} = F_u \frac{du}{dt} + F_v \frac{dv}{dt} \]

\[ \int_{\partial S} \mathbf{F} \cdot d\mathbf{s} = \int_{\partial A} F_u \, du + F_v \, dv \]

But from above,

\[ \int_S \nabla \times \mathbf{F} \cdot dS = \int_A \left( \frac{\partial F_v}{\partial u} \frac{\partial F_u}{\partial v} \right) dA \]

2.4.3 Divergence theorem

For \( \mathbf{F} = F_x \mathbf{i} + F_y \mathbf{j} + F_z \mathbf{k} \) in three dimensions, we can show

\[ \int_V \nabla \cdot \mathbf{F} \, dV = \int_S \mathbf{F} \cdot dS \quad (S = \partial V) \]

by generalising arguments in two dimensions used in 2.4.1 above. As before, start with simplest case: assume that points \((x, y, z)\) in \( V \) lie in an interval \( Z_{xy} : Z_{xy}(x, y) \geq z \geq Z_{xy}(x, y) \) for each \( x, y \).

Then

\[ \int_V \frac{\partial F_x}{\partial z} \, dV = \int_A \left( \int_{z_{xy}} \frac{\partial F_x}{\partial z} \, dz \right) dA \]

where \( A \) is the projection of \( V \) onto the \( xy \)-plane.

\[ = \int_A \left( F_x(x, y, z_+(x, y)) - F_x(x, y, z_-(x, y)) \right) dA = \int_S (\mathbf{F} \cdot \mathbf{k}) \cdot n \, dS \]

This last step follows because

\[ \mathbf{k} \cdot n \, dS = \begin{cases} \frac{dA}{dS}, & \text{at } (x, y, z_+(x, y)) \\ -\frac{dA}{dS}, & \text{at } (x, y, z_-(x, y)) \end{cases} \]

as shown.

Similarly for the \( F_x \) and \( F_y \) terms – gives proof for simple kind of region \( V \). Extensions to more general \( V \) can be obtained by generalising methods (i) and (ii) in 2.4.1 above.
2.5 Vector differential operators in orthogonal curvilinear coordinates

Recall (from 1.7) that with \( r(u, v, w) \) we have

\[
\frac{\partial r}{\partial u} = h_u e_u, \quad \frac{\partial r}{\partial v} = h_v e_v, \quad \frac{\partial r}{\partial w} = h_w e_w
\]

and this defines a right-handed orthonormal basis. Since \( df = \nabla f \cdot dr \) and

\[
dr = \frac{\partial r}{\partial u} du + \frac{\partial r}{\partial v} dv + \frac{\partial r}{\partial w} dw
\]

we read off a formula for the gradient

\[
\nabla f = \frac{1}{h_u} \frac{\partial f}{\partial u} e_u + \frac{1}{h_v} \frac{\partial f}{\partial v} e_v + \frac{1}{h_w} \frac{\partial f}{\partial w} e_w
\]

To find curl and div we can apply

\[
\nabla \times A = \frac{1}{h_u h_v h_w} \begin{vmatrix} h_u e_u & h_v e_v & h_w e_w \\ \frac{\partial}{\partial u} (h_u A_u) & \frac{\partial}{\partial v} (h_v A_v) & \frac{\partial}{\partial w} (h_w A_w) \\ \frac{\partial}{\partial u} A_u & \frac{\partial}{\partial v} A_v & \frac{\partial}{\partial w} A_w \end{vmatrix}
\]

\[
\nabla \cdot A = \frac{1}{h_u h_v h_w} \left( \frac{\partial}{\partial u} (h_u h_v A_u) + \cdots \right)
\]

\[
\nabla^2 f = \frac{1}{h_u h_v h_w} \left( \frac{\partial}{\partial u} \left( \frac{h_v h_w}{h_u} \frac{\partial}{\partial u} A_u \right) + \cdots \right)
\]

All ... terms obtained by cycling in \( u, v, w \).

Check: familiar expressions obtained for \( x, y, z \) with \( h_x = h_y = h_z = 1 \).

We need to be able to apply results above in two other cases.

2.5.1 Cylindrical polars

\( \rho, \varphi, z \) with \( h_\rho = h_z = 1 \) and \( h_\varphi = \rho \). Basis vectors

\[
e_\rho = \cos \varphi \mathbf{i} + \sin \varphi \mathbf{j}, \quad e_\varphi = -\sin \varphi \mathbf{i} + \cos \varphi \mathbf{j}, \quad e_z = \mathbf{k}
\]

Note \( \partial e_\rho / \partial \varphi = e_\varphi \) and \( \partial e_\varphi / \partial \varphi = -e_\rho \), and all other derivatives vanish. This can be used to check

\[
\nabla f = \frac{\partial f}{\partial \rho} e_\rho + \frac{1}{\rho} \frac{\partial f}{\partial \varphi} e_\varphi + \frac{\partial f}{\partial z} e_z
\]

\[
\nabla \times A = \begin{vmatrix} e_\rho & \rho e_\varphi & e_z \\ \partial / \partial \rho & \partial / \partial \varphi & \partial / \partial z \\ A_\rho & \rho A_\varphi & A_z \end{vmatrix}
\]

\[
\nabla \cdot A = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho A_\rho) + \frac{1}{\rho} \frac{\partial A_\varphi}{\partial \varphi} + \frac{\partial A_z}{\partial z}
\]

in agreement with general results.

2.5.2 Spherical polars

\( r, \theta, \varphi \) with \( h_r = 1, h_\theta = r, h_\varphi = r \sin \theta \). The basis vectors

\[
e_r = \sin \theta (\cos \varphi \mathbf{i} + \sin \varphi \mathbf{j}) + \cos \theta \mathbf{k}
\]

\[
e_\theta = \cos \theta (\cos \varphi \mathbf{i} + \sin \varphi \mathbf{j}) - \sin \theta \mathbf{k}
\]

\[
e_\varphi = -\sin \varphi \mathbf{i} + \cos \varphi \mathbf{j}
\]
2.5.3 Sketch proof of formulas for div and curl in general case (non-examinable)

Rather than calculating $\partial e_u / \partial v$ etc..., note

$$\nabla f = \frac{\partial f}{\partial r} e_r + \frac{1}{r} \frac{\partial f}{\partial \theta} e_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} e_\phi$$

$$\nabla \times A = \left| \begin{array}{ccc}
 e_r & r e_\theta & r \sin \theta e_\phi \\
 \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\
 A_r & r A_\theta & r \sin \theta A_\phi \\
 \end{array} \right|$$

$$\nabla \cdot A = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (r \sin \theta A_\phi) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} A_\phi.$$
This is because
\[
\frac{d\rho}{dt} = \frac{d}{dt}(\nabla \cdot E) = \epsilon_0 \frac{d}{d\tau} (c^2 \nabla \times B - c^2 \mu_0 j) = -\nabla \cdot j
\]
as required.

(ii) When \( \rho = 0, j = 0 \), for Maxwell implies
\[
\left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) E = \left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) B = 0
\]

Using the divergence or Stokes’ theorem we can re-cast Maxwell’s equations in integral form, e.g. if \( V \) is a volume bounded by a closed curve surface \( S \) then:
\[
\int_S E \cdot dS = \frac{1}{\epsilon_0} \int_V \rho dV = \frac{Q_V}{\epsilon_0}
\]
is the flux of \( E \) out of \( S \), where \( Q_V \) is the total charge in \( V \). Also
\[
\int_S B \cdot dS = 0
\]
so there is zero net magnetic flux.

3.1.2 Static charges and steady current

If \( \rho \) and \( j \) are independent of \( t \), then \( E \) and \( B \) decouple in equations. We study;

**Electrostatics**

\[\nabla \times E = 0, \quad \nabla \cdot E = \frac{\rho}{\epsilon_0} \]

\[\Rightarrow E = \nabla \phi, \quad \nabla^2 \phi = -\frac{\rho}{\epsilon_0}\]

(Poisson’s equations, for scalar \( \phi(x) \).)

**Magnetostatics**

\[\nabla \cdot B = 0, \quad \nabla \times B = \mu_0 j \]

\[\Rightarrow B = \nabla \times A\]

where \( A \) is a vector potential. Note that we can change \( A \) to \( A + \nabla X \) without affecting \( B \). With a suitable choice of \( X \) we can arrange to have \( \nabla^2 A = -\mu_0 j \) (the vector Poisson equations).

3.2 Electrostatics and gravitation

3.2.1 Basic relations

There is an exact parallel between electrostatic force on a charge \( q \) and the gravitational force on a mass \( m \).

<table>
<thead>
<tr>
<th><strong>Electric</strong></th>
<th><strong>Gravitational</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{F} = q \mathbf{E} ) (force per unit charge)</td>
<td>( \mathbf{F} = m \mathbf{g} ) (force per unit mass)</td>
</tr>
<tr>
<td>( \mathbf{E} ) is the electric field</td>
<td>( \mathbf{g} ) is the gravitational field</td>
</tr>
<tr>
<td>( \nabla \times \mathbf{E} = 0 ) (( \mathbf{E} ) is conservative)</td>
<td>( \nabla \times \mathbf{g} = 0 ) (( \mathbf{g} ) is conservative)</td>
</tr>
<tr>
<td>( \mathbf{E} = -\nabla \phi ) (( \phi ) = electrostatic potential)</td>
<td>( \mathbf{g} = -\nabla \varphi ) (( \varphi ) = gravitational potential)</td>
</tr>
<tr>
<td>( \mathbf{V} = q \phi ) (( \phi ) = energy per unit charge)</td>
<td>( \mathbf{V} = m \varphi ) (( \varphi ) = energy per unit mass)</td>
</tr>
<tr>
<td>( \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \Rightarrow \nabla^2 \phi = -\frac{\rho}{\epsilon_0} )</td>
<td>( \nabla \cdot \mathbf{g} = -4\pi G \rho \Rightarrow \nabla^2 \varphi = 4\pi G \rho )</td>
</tr>
</tbody>
</table>

like masses repel so sign negative

like masses attract so sign positive

\( \rho \) is the electric charge density or mass density.
3.2.2 Spherically symmetric fields – electrostatic case

Consider \( \rho(r) \) a function only of \( r \) (in spherical polars). The electrostatic case and assume spherical symmetry for fields. \( \varphi(r) \) and \( E = E(r) \hat{r} \) with \( E(r) = -\varphi'(r) \).

Apply Gauss’s Theorem to a sphere of radius \( R \), with normal \( \mathbf{n} = \hat{r} \), we get

\[
\int_{S_R} E \cdot dS = \frac{1}{\varepsilon_0} \int_{V_R} \rho \, dV
\]

\[
E \cdot \mathbf{n} = E(R), \quad 4\pi R^2 E(R) = \frac{Q(R)}{\varepsilon_0}
\]

with \( Q(R) \) the total charge in \( V_R \)

\[
\Rightarrow E(r) = \frac{1}{4\pi \varepsilon_0} \frac{Q(r)}{r^2}
\]

Now determine \( \varphi(r) \) by integrating.

Consider charge contained within radius \( a \). Then for \( r > a \) we have

\[
E(r) = \frac{Q}{4\pi \varepsilon_0 r^2}
\]

Field is just what we obtain if all charge \( Q \) were at \( r = 0 \). We can define a point charge as a distribution with a very small footprint. Then formulae above hold for all \( r \) large compared to a force on \( \alpha \), and the force on a charge \( q \) is given by Coulomb’s Law:

\[
qE = \frac{qQ}{4\pi \varepsilon_0 r^2} \hat{r}
\]

Rather than using Gauss’s flux method (above) we can also solve \( \nabla^2 \varphi = -\rho/\varepsilon_0 \) directly, given \( \varphi(r) \), but Gauss’s method is normally faster.

Note: that with \( \rho \) contained within radius \( a \), for \( r > a \) we need to solve \( \nabla^2 \varphi = 0 \), Laplace’s equation.

3.2.3 Spherically symmetric fields – gravitational case

Consider a spherically symmetric distribution of matter with density \( \rho(r) \). Spherical symmetry

\( \Rightarrow \)

\( \varphi(r) \) and gravitational field \( \mathbf{g}(r) = g(r) \hat{r} \) with \( g(r) = -\varphi'(r) \). Note \( g(r) \) is not \( |\mathbf{g}(r)| \)

We need to solve

\[
\nabla^2 \varphi = \frac{4\pi G}{\rho}
\]

but this can be done by applying divergence theorem to first determinant \( g(r) \). (Consider 3.1.1.) For a sphere of radius \( R \), \( \mathbf{g} = g(r) \hat{r} \) and \( \mathbf{g} \cdot \mathbf{n} = g(R) \). Then

\[
\int_{S_R} \mathbf{g} \cdot d\mathbf{S} = -4\pi G \int_{V_R} \rho \, dV
\]

and hence \( 4\pi R^2 g(R) = -4\pi GM(R) \) where \( M(R) \) is the total mass contained in \( V_R \) \( (r \leq R) \). Thus

\[
g(r) = -\frac{GM(r)}{r^2}
\]

Note. If all mass contained within radius \( a \) then

\[
g(r) = -\frac{GM}{r^2} \hat{r} \text{ for } r > a
\]

with \( M \) the total mass (this is the inverse square law).

A point mass \( M \) is obtained in limit \( a \to 0 \), and then the inverse square law holds for all \( r > 0 \).

The potential is
with integration constant fixed by $\varphi(r) \to 0$ as $r \to \infty$. The resulting force on a mass $m$ at position $r$ is given by Newton's law of gravitation:

$$ F = -\frac{GMm}{r^2} \hat{r} $$

**Example.** Sphere of constant density $\rho$ and radius $a$ (“planet”). Then

$$ \rho(r) = \begin{cases} \rho_0, & r \leq a \\ 0, & r > a \end{cases} $$

Mass contained within radius $r$ is

$$ M(r) = \int_0^r 4\pi r'^2 \rho(r') dr' = \begin{cases} \frac{4}{3} \pi r^3 \rho_0, & r \leq a \\ \frac{4}{3} \pi a^3 \rho_0, & r > a \end{cases} = \begin{cases} \frac{M}{a^3}, & r \leq a \\ M, & r > a \end{cases} $$

where $M$ is the total mass. Therefore

$$ g(r) = \begin{cases} -\frac{GM}{3\pi \rho_0} r, & r \leq a \\ -\frac{GM}{a^3}, & r > a \end{cases} $$

Potential;

$$ \varphi(r) = \begin{cases} \frac{GM}{2a^3}, & r \leq a \\ -\frac{GM}{r}, & r > a \end{cases} $$

with the constants chosen for continuity at $r = a$, and so that $\varphi \to 0$ as $r \to \infty$.

### 3.3 Laplace’s equation and Poisson’s equation

Laplace’s equation is $\nabla^2 \varphi = 0$ and Poisson’s is $\nabla^2 \varphi = -\rho$. To recover electrostatic conventions, let $\rho = \rho/\varepsilon_0$, and for gravitational conventions, $\rho = -4\pi \rho$.

These equations arise throughout mathematics and physics. This is expected from the fact that, up to an overall factor, the Laplacian $\nabla^2$ is the only way to construct a rotationally- and translationally-invariant operator from terms involving the spatial derivatives.

#### 3.3.1 Uniqueness theorem

**Theorem 3.1 (Uniqueness theorem).** Consider a volume $V$ with boundary $\partial V = S$, a closed surface with outward normal $n$. Let $\varphi(x)$ be a smooth field satisfying $\nabla^2 \varphi = 0$ on $V$, and either

(i) $\varphi(x) = f(x)$ on $S$

or (ii) $\partial \varphi / \partial n = -n \cdot \nabla \varphi = g(x)$ on $S$

where $f(x)$ and $g(x)$ are prescribed functions. Then for (i) $\varphi(x)$ is unique, and for (ii) $\varphi(x)$ is unique up to the addition of a constant.

**Proof.** Let $\varphi_1(x)$ and $\varphi_2(x)$ be any solutions to the problem posed above, each with boundary conditions (i) or (ii). Then $\psi = \varphi_1 - \varphi_2$ satisfies $\nabla^2 \psi = \nabla^2 \varphi_1 - \nabla^2 \varphi_2 = 0$ and

(i) $\psi = 0$ on $S$

(ii) $\partial \psi / \partial n = 0$ on $S$.

By the divergence theorem,
\[ \int_V \nabla \cdot (\psi \nabla \psi) \, dV = \int_S \psi \nabla \psi \cdot dS \]

But \( \nabla \cdot (\psi \nabla \psi) = \nabla \psi \cdot \nabla \psi + \psi \nabla^2 \psi = |\nabla \psi|^2 \) and so
\[ \int_V |\nabla \psi|^2 \, dV = \int_S \psi \frac{\partial \psi}{\partial n} \, dS = 0 \]

by either (i) or (ii). Since \(|\nabla \psi|^2 \geq 0\) we deduce that \(\nabla \psi = 0\) so \(\psi = C\), a constant, on \(V\). Now
(i) \(\psi = 0\) on \(S \Rightarrow c = 0\) and \(\varphi_1 = \varphi_2\) on \(V\).
(ii) No further information can be deduced from \(\partial \psi / \partial n = 0\) on \(S\), so \(\varphi_1 = \varphi_2 + C\), so the solution is unique up to a constant. QED.

Notes. The boundary conditions of type (i) and (ii) are known as Dirichlet and Neumann respectively.

The uniqueness results extend to Poisson’s equation, \(\nabla^2 \varphi = -\rho\) on \(V\) with boundary conditions as before. This follows from the fact that \(\nabla^2 \varphi_1 = -\rho\) and \(\nabla^2 \varphi_2 = -\rho \Rightarrow \nabla^2 \psi = 0\) where \(\psi = \varphi_1 - \varphi_2\), so the same proof applies.

Note that we cannot impose (i) and (ii) simultaneously – the problem is then over-determined. But we can use various combinations of (i) and (ii) (e.g. on different parts of the boundary).

Note that with Neumann boundary condition \(\partial \psi / \partial n = g\) for Poisson’s equation we must have
\[ \int_V \rho \, dV + \int_S g \, dS = 0 \]

by divergence theorem applied to \(\nabla \varphi\).

Our discussion in three dimensions works equally well in two dimensions. The proof extends to unbounded domains \(V\) if \(\psi = \varphi_1 - \varphi_2\) obeys appropriate conditions. E.g. in three dimensions
\[ \psi(x) = O \left( \frac{1}{|x|} \right), \quad \frac{d\psi}{dn} = O \left( \frac{1}{|x|^2} \right) \]

since then
\[ \int_S \psi \frac{d\psi}{dn} \, dS \to 0 \]

for a sphere with radius \(\to \infty\).

### 3.3.2 Point sources and averaging

In three dimensions, consider
\[ \psi(x) = \frac{1}{|x-x_0|}, \quad \nabla \psi(x) = -\frac{1}{|x-x_0|^3} (x-x_0), \quad \nabla^2 \psi = 0 \]

all for \(x \neq x_0\).

This is a solution of Laplace’s equation for \(x \neq x_0\), or a solution of Poisson’s equation for a point source if \(x = x_0\) (compare with point mass or charge)

If \(S_R\) is the surface of a sphere \(|x-x_0| = R\), then \(\nabla \psi = -n/R^2\) where \(n\) is the unit normal to \(S_R\) at \(x\). Also
\[ \int_{S_R} \nabla \psi \cdot dS = \frac{1}{R^2} \int_{S_R} dS = 4\pi \]

(all familiar). More generally, if \(S\) is any closed surface bounding a volume \(V\), then
\[ -\int_{S} \nabla \psi \cdot dS = \begin{cases} 0, & x \notin V \\ 4\pi, & x \in V \end{cases} \]

First case follows from divergence theorem for \(V\) since \(\nabla^2 \psi = 0\). Second case follows from divergence theorem applied to \(W\) region with \(\partial W = S - S_R\) (with orientations) for some suitable \(R\), as shown.
\[ \nabla^2 \psi = 0 \text{ in } W \] and so result follows by previous case.

Now let \( \varphi(x) \) be any smooth function defined on a region containing \( |x - x_0| \leq R \) with boundary \( S_R \).

\[ - \int_{S_R} \varphi \nabla \psi \cdot dS = \frac{1}{R^2} \int_{S_R} \varphi(x) \, dS = 4\pi \bar{\varphi}(R) \]

where by definition \( \bar{\varphi}(R) \) is the average of \( \varphi \) over \( S_R \).

As \( R \to 0 \) we expect \( \bar{\varphi}(R) \to \varphi(x_0) \) and this is so:

\[ \left| \int_{S_R} (\varphi(x) - \varphi(x_0)) \, dS \right| \leq 4\pi R^2 \max_{x \text{ on } S_R} |\varphi(x) - \varphi(x_0)| \]

\[ \Rightarrow |\bar{\varphi}(R) - \varphi(x_0)| \leq \max_{x \text{ on } S_R} |\varphi(x) - \varphi(x_0)| \to 0 \text{ as } R \to 0 \]

as required.

### 3.3.3 The maximum (or minimum) principle

Let \( \varphi(x) \) be a smooth function satisfying \( \nabla^2 \varphi = 0 \) in a volume \( V \) with boundary \( \partial V = S \). Then: \( \varphi \) has no local maximum or minimum inside \( V \) (not on \( S \)) and \( \varphi \) attains its maximum and minimum values on the boundary \( S \).

To confirm this result (but not quite prove it) we use differential approach to analysing a stationary point \( x_0 \). We have \( \nabla \varphi = 0 \) at \( x_0 \) and the character of the stationary point can be deduced from the eigenvalues of the Hessian

\[ H_{ij} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j} |_{x_0} \]

The sum of the eigenvalues

\[ \text{Tr}(H) = H_{ii} = \frac{\partial^2 \varphi}{\partial x_i \partial x_i} |_{x_0} = 0. \]

If \( H \neq 0 \), then there must be eigenvalues of each sign, and hence \( x_0 \) is not a local maximum or minimum. However, this leaves open the case \( H = 0 \).

To give a general proof, consider spheres \( V_1 : |x - x_0| \leq R_1 \) and \( V_2 : |x - x_0| \leq R_2 \) with \( R_2 > R_1 \) and \( S_1 = \partial V_1, S_2 = \partial V_2 \) and outward normals. Then

\[ \int_{S_1} \nabla \varphi \cdot dS = \int_{V_1} \nabla^2 \varphi \, dV = 0, \quad \int_{S_2} \nabla \varphi \cdot dS = \int_{V_2} \nabla^2 \varphi \, dV = 0 \]

But then

\[ \int_{S_1} \psi \nabla \varphi \cdot dS = \int_{S_2} \psi \nabla \varphi \cdot dS = 0 \]

with

\[ \psi(x) = \frac{1}{|x - x_0|} = \begin{cases} 1/R_1 & \text{on } S_1, \\ 1/R_2 & \text{constant on } S_2. \end{cases} \]

By divergence theorem

\[ \int_{S_2} (\psi \nabla \varphi - \varphi \nabla \psi) \cdot dS - \int_{S_1} (\psi \nabla \varphi - \varphi \nabla \psi) \cdot dS = \int_W (\psi \nabla^2 \varphi - \varphi \nabla^2 \psi) \, dV = 0 \]

with \( W = V_2 - V_1 \) and \( \partial W = S_2 - S_1 \) (with orientations), since \( \nabla^2 \varphi = \nabla^2 \psi = 0 \) in \( W \). Hence
for any $R_2$ and $R_1$. Thus, for any $R$,

$$\varphi(R) = \varphi(x_0).$$

For small $R$,

$$\bar{\varphi}(R) = \varphi(x_0).$$

Now if $\varphi(x_0)$ is a local maximum for the function, we can find $R$ such that $\varphi(x) < \varphi(x_0)$ for $|x - x_0| = R$. But then $\bar{\varphi}(R) < \varphi(x_0)$ which is a contradiction. This proves the maximum principle, and minimum principle follows similarly.

For second version of maximum principle, note that if $\varphi$ is not constant and attains its maximum value at $x_0$, then we can find $R$ with $\varphi(x) \leq \varphi(x_0)$ where $\varphi|\{x - x_0\} = R$ and with strict inequality $\varphi(x) < \varphi(x_0)$ somewhere on the sphere, for some $x_0$. Then $\bar{\varphi}(R) < \varphi(x_0)$, another contradiction.

Note that we attain proof of uniqueness of solutions of Poisson’s equation with Dirichlet boundary conditions. Since $\nabla^2(\varphi_1 - \varphi_2) = 0$ and $\varphi_1 - \varphi_2 = 0$ on a boundary, the maximum/minimum principle says that $0$ is maximum and minimum value of $\varphi_1 - \varphi_2$ everywhere in the interior; i.e. $\varphi_1 = \varphi_2$.

### 3.3.4 Integral solutions of Poisson’s equations

$\nabla^2 \varphi = -\rho$ in a volume $V$ with boundary $S = \partial V$. To find a solution at point $x_0$ in the interior of $V$, use

$$\psi(x) = \frac{1}{|x - x_0|}.$$

Consider a sphere $V_\varepsilon$ with equation $|x - x_0| \leq \varepsilon$ and boundary $\partial V_\varepsilon = S_\varepsilon$, and let $W = V - V_\varepsilon$ so $\partial W = \partial V - S_\varepsilon$ (with orientations).

By divergence theorem (or Green’s second identity),

$$\int_S (\varphi \nabla \psi - \psi \nabla \varphi) \cdot dS = \int_{S_\varepsilon} (\varphi \nabla \psi - \psi \nabla \varphi) \cdot dS = \int_W \varphi \nabla^2 \psi - \psi \nabla^2 \varphi \, dV$$

$$\left| \int_{S_\varepsilon} \psi \nabla \varphi \cdot dS \right| = \frac{1}{\varepsilon} \int_{V_\varepsilon} \nabla^2 \varphi \, dV \leq \frac{4}{\varepsilon} \pi \varepsilon^3 \max_{x \in V_\varepsilon} |\rho(x)| \to 0 \text{ as } \varepsilon \to 0.$$

Also,

$$-\int_{S_\varepsilon} \varphi \nabla \psi \cdot dS = 4\pi \bar{\varphi}(\varepsilon) \to 4\pi \varphi(x_0) \text{ as } \varepsilon \to 0$$

$$\Rightarrow \int_{V_\varepsilon} \psi \rho \, dV \to \int_V \psi \rho \, dV \text{ as } \varepsilon \to 0$$

(the limit exists). We deduce that

$$4\pi \varphi(x_0) = \int_V \frac{1}{|x - x_0|} \rho(x) \, dV - \int_S \left( \varphi(x) \frac{\partial}{\partial n} \frac{1}{|x - x_0|} - \frac{1}{|x - x_0|} \frac{\partial \varphi(x)}{\partial n} \right) \, dS$$

This gives solutions of Poisson’s equations, but here we have both $\varphi$ and $\partial \varphi / \partial n$ on boundary which need to be specified.

(i) Solutions in infinite volume. With boundary conditions

$$|\varphi(x)| = O(|x|^{-1}), \quad |\frac{\partial \varphi}{\partial n}| = O(|x|^{-2})$$

the surface term above $\to 0$ if $S$ is a sphere with radius $\to \infty$. Then
\[ \varphi(x) = \int \frac{1}{4\pi |x - x'|} \rho(x') \, dV' \]

with integral \( V' \) over all space.

Solution can be interpreted as the superposition of potentials;

\[ \frac{1}{4\pi} \frac{1}{|x - x'|} \rho(x') \delta V' \]

due to point sources at each \( x' \).

(ii) Solutions in finite volume. If we have Dirichlet or Neumann boundary conditions \( \varphi = f \) or \( \partial \varphi/\partial n = g \) on \( S \), we have

\[ G(x, x')_D = \frac{1}{4\pi} \frac{1}{|x - x'|} + H(x, x')_D \]

or

\[ G(x, x')_N = \frac{1}{4\pi} \frac{1}{|x - x'|} + H(x, x')_N \]

where

\[ \nabla^2 H(x, x')_D = \nabla^2 H(x, x')_N = 0 \text{ in } V \]

and

\[ G(x, x')_D = 0 \text{ or } \frac{\partial G}{\partial n}, \quad (x, x')_N = 0 \text{ on } S \]

Then we can show previous argument unaffected but

\[ \varphi(x) = \int_V G(x, x')_D \rho(x') \, dV' - \int_S f(x') \frac{\partial G_D}{\partial n} \, dS' \]

or

\[ \varphi(x) = \int_V G(x, x')_N \rho(x') \, dV - \int_S g(x') G(x, x')_N \, dS \]

3.4 Summary of methods for solving Laplace's equations and Poisson's equations

3.4.1 Gauss's flux method

\[ \nabla^2 \varphi = -\rho, \quad \int_D \nabla \varphi \cdot dS = -\int_D \rho \, dV \]

for a region \( D \) in \( \mathbb{R}^3 \) and similarly in \( \mathbb{R}^2 \). If \( \rho(r) \) and \( \varphi(r) \) depend only on radial coordinate \( r \) then \( \nabla \varphi = \varphi'(r) \hat{r} \). Also, by the divergence theorem with \( D \) a solid sphere or disc, we get

\[ -\varphi'(r) = \begin{cases} \frac{1}{4\pi r^2} \int_D \rho \, dV - \frac{1}{r^2} \int_0^r r'^2 \rho(r') \, dr' \text{ in } \mathbb{R}^3 \\
\frac{1}{2\pi r} \int_D \rho \, dA - \frac{1}{r} \int_0^r r' \rho(r') \, dr' \text{ in } \mathbb{R}^2 \end{cases} \]

assuming \( \rho(r) \) is bounded as \( r \to 0 \), and so \( \varphi'(r) \to 0 \) as \( r \to 0 \).

First set of expressions applies to point sources with \( \rho(r) = 0 \text{ for } r > 0 \) and singularity at \( r = 0 \). (Spherical case in \( \mathbb{R}^3 \) discussed previously.)

For circular symmetry in \( \mathbb{R}^2 \) or cylindrical symmetry in \( \mathbb{R}^3 \), the same approach gives field due to a point charge \( \sigma \) at origin, or line of charge-per-unit-length \( \sigma \) along the \( z \)-axis.

Resulting electric field \( \rho = \sigma / \varepsilon_0 \)

\[ E = -\nabla \varphi = \frac{\sigma}{2\pi \varepsilon_0} \frac{1}{r} \hat{r} \]
where $r$ is the radial coordinate in cylindrical polars in $\mathbb{R}^3$. Thus
\[ \varphi = -\frac{\sigma}{2\pi \varepsilon_0} \ln r \]

To derive this in $\mathbb{R}^3$, first replace $\rho$ with $\rho / \varepsilon_0$ in Poisson’s equation and then apply flux argument to a cylinder of radius $R$ and height $l$ as shown. By symmetry, $E = E(r)\hat{r}$ and $2\pi RIE(R) = l\sigma / \varepsilon_0$, which gives result for $E$.

### 3.4.2 Symmetry and separable solutions

For $\varphi(r)$ depending only on radial coordinate $r$, we have
\[
\nabla^2 \varphi = \begin{cases} \varphi'' + \frac{2\varphi'}{r} & \text{in } \mathbb{R}^3 \\ \varphi'' + \frac{\varphi'}{r} & \text{in } \mathbb{R}^2 \end{cases}
\]

Then $\nabla^2 \varphi = -\rho$ can be solved directly and result for $\varphi'(r)$ agrees with result from flux method.

For Laplace’s equation ($\rho = 0$) we get
\[
\nabla^2 \varphi = 0 \Rightarrow \varphi = \begin{cases} A + \frac{B}{r} & \text{in } \mathbb{R}^3 \\ A + B \ln r & \text{in } \mathbb{R}^2 \end{cases}
\]

Solutions without spherical or circular symmetry can be found in separable form;
\[ \varphi = R(r)\Theta(\theta) \]

In $\mathbb{R}^2$, with $r$ and $\theta$ polar coordinates, we find $\nabla^2 \varphi = 0$ for $R = r^n$ or $r^{-n}$ with $\Theta = \cos n\theta$ or $\sin n\theta$ ($n \geq 1$).

In $\mathbb{R}^3$, with $r$, $\theta$, two of the spherical polar coordinates (solution rotationally invariant about $z$-axis), we find $R = r^n$ or $r^{-(n+1)}$ with $\Theta = P_n(\cos \theta)$, where $P_n$ is a Legendre polynomial.

### 3.4.3 Green’s function approach

In 3.3.4 we found solution of Poisson’s equation in $\mathbb{R}^3$ (with boundary conditions on $\varphi$ as $|x| \to \infty$)
\[ \varphi(x) = \int G(x, x')\rho(x') \, dV' \]

where
\[ G(x, x') = \frac{1}{4\pi |x - x'|} \]

In $\mathbb{R}^2$,
\[ G(x, x') = -\frac{1}{2\pi} \ln |x - x'|. \]

These are solutions for point sources at $x = x'$.

Finding such a Green’s function $G$ is a powerful general method for solving differential equations (see IB Maths Methods).

### 4. Tensors

#### 4.1 Introduction

#### 4.1.1 Definition of a tensor

A vector $\mathbf{v}$ is specified by its components $v_i$ with respect to an orthonormal basis $e_i$ in $\mathbb{R}^3$ or a set of Cartesian coordinates $x_i$. Under a change of coordinates or basis
\[ x'_i = R_{ij} x_j, \quad e'_i = R_{ij} e_j \]

with \( R_{ip} R_{jp} = R_{qj} R_{qj} = \delta_{ij} \) and \( \det R = 1 \), two conditions which mean \( R \) is a rotation, the components of \( \mathbf{v} \) change; \( v'_i = R_{ij} v_j \).

Tensors are geometrical objects which obey a generalisation of this transformation rule.

A tensor of rank \( n \) is an object \( T \) with components \( T_{ij...k} \) (where there are \( n \) indices) with respect to Cartesian coordinates \( x_i \), and such that under a change in coordinates \( x'_i = R_{ij} x_j \) the components change according to

\[ T'_{ij...k} = R_{ip} R_{jq} ... R_{kr} T_{pq...r} \]

This is the tensor transformation rule.

A tensor of rank \( n = 0 \) (no indices) has \( T' = T \), thus it is a scalar.

A tensor of rank \( n = 1 \) has \( T'_{ij} = R_{ij} T_j \) and thus is a vector.

A tensor of rank \( n = 2 \) has \( T'_{ij} = R_{ij} R_{jq} T_{pq} \) and components \( (T_{ij}) \), and can be regarded as a \( 3 \times 3 \) matrix.

But for general \( n \), a tensor is some multidimensional array with \( 3^n \) components.

4.1.2 How tensors arise

(i) In working with vectors \( u, v, ..., w \) (\( n \) vectors in total), we may encounter expressions such as

\[ T_{ij...k} = u_i v_j ... w_k \]

This is a tensor, by checking the transformation rule. E.g. for \( n = 2 \), \( T_{ij} = u_i v_j \) and \( T'_{ij} = u'_i v'_j = (R_{ip} u_p)(R_{jq} v_q) = R_{ip} R_{jq} u_p v_q = R_{ip} R_{jq} T_{pq} \).

(ii) Two special examples: \( \delta_{ij} \) and \( \varepsilon_{ijk} \) are invariant tensors, meaning that their components are same in any Cartesian coordinate system.

\[ \delta'_{ij} = R_{ip} R_{jq} \delta_{pq} = R_{ij} \delta_{ij} \]

\[ \varepsilon'_{ijk} = R_{ip} R_{jq} R_{kr} \varepsilon_{pqr} = \varepsilon_{ijk} (\det R) = \varepsilon_{ijk} \]

(check the latter for \( i = 1, j = 2, k = 3 \) etc)

(iii) Second-rank tensors arise as matrices or linear maps between vectors. If \( v_i = M_{ij} u_j \) and \( v'_i = M'_i j u'_j \) in two Cartesian coordinate systems, then

\[ v'_i = R_{ip} v_p = R_{ip} M_{pq} u_q = M'_{ij} u'_j = M'_{ij} R_{jq} u_q \]

This holds for all \( u_q \), and so \( R_{ip} M_{pq} = M_{ij} R_{jq} \) or \( M'_{ij} = R_{ip} R_{jq} M_{pq} \). So \( M_{ij} \) is a 2\(^{nd}\)-rank tensor, and in matrix notation \( M' = R M R^T \) (standard behaviour for a matrix under an orthogonal change of basis).

There are many physical examples;

(a) In a conducting medium the current \( I_i \) (vector) resulting from an electric field \( E_i \) (vector) is given by \( J_i = \sigma_{ij} E_j \) where \( \sigma_{ij} \) (2\(^{nd}\)-rank tensor) is the conductivity tensor. This is a general version of Ohm’s Law. The medium may conduct differently according to the direction of \( E_i \).

(b) For a rigid body the angular velocity \( \omega_i \) and angular momentum \( L_i \) (both vectors) are related by \( L_i = I_{ij} \omega_j \) with \( I_{ij} \) the inertia tensor.

(iv) Given a scalar field \( \varphi(x) \) (a smooth function on \( \mathbb{R}^3 \)) then

\[ T_{ij...k} = \frac{\partial^n \varphi}{\partial x_i \partial x_j ... \partial x_k} \]

is a tensor of rank \( n \) at each point \( x \), i.e. a tensor field.

Tensor transformation property is checked using chain rule;

\[ x'_i = R_{ip} x_p, \quad x_p = R_{ip} x'_i \]
So e.g. $n = 2$

$$T_{ij}' = \frac{\partial^2 \varphi}{\partial x_i' \partial x_j'} = R_{ip} R_{jq} \frac{\partial^2 \varphi}{\partial x_p \partial x_q} = R_{ip} R_{jq} T_{pq}$$

as required.

### 4.2 Operations involving tensors

#### 4.2.1 Addition and scalar multiplication

Addition of tensors $T$ and $S$ of same rank $n$ is defined by $(T + S)_{ij...k} = T_{ij...k} + S_{ij...k}$

Multiplication of a tensor $T$ by a scalar $\alpha$ is defined by $(\alpha T)_{ij...k} = \alpha(T_{ij...k})$.

To show that the results are indeed tensors, we must check the transformation rule.

#### 4.2.2 Tensor products

If $T$ and $S$ are tensors of ranks $n$ and $m$, the tensor product

$$(T \otimes S)_{(ij...k)}^{(pq...r)} = T_{ij...k} S_{pq...r}$$

Again, checking this is a tensor means checking the transformation rule.

The tensor product definition extends immediately to any number of tensors, e.g. $T \otimes S \otimes P$.

This generalises our earlier discussion using vectors; we can write

$$T = u \otimes v \otimes \ldots \otimes w$$

which we wrote before as $T_{ij...k} = u_i v_j \ldots w_k$.

#### 4.2.3 Contraction

Given $T$ a tensor of rank $n$ with components $T_{ijp...q}$, we define a new tensor $S$ of rank $n - 2$ by

$$S_{p...q} = \delta_{ij} T_{ijp...q} = T_{ip...q}$$

This is called contracting on original index pair $i, j$. It can be done on any index pair and produces different results in general.

E.g. for $n = 2$, $S = \delta_{ij} T_{ij} = T_{ii}$ which is a scalar (it has rank $n - 2 = 0$). To check this is a scalar, $T_{ii}' = R_{ik} R_{ii} T_{kk} = \delta_{kk} T_{kk} = T_{kk}$.

The general case works in just the same way, with indices $p, \ldots, q$ playing no role in contraction.

#### 4.2.4 Symmetric and antisymmetric tensors

A tensor $T$ of rank $n$ obeying $T_{ijp...q} = \pm T_{ijp...q}$ is said to be symmetric/antisymmetric on the index pair $i, j$. This property holds in every coordinate system if it holds in any one, since

$$T_{k\ell r...s} = R_{k(i} R_{j)l} R_{rp} \ldots R_{sq} T_{ijp...q} = \pm R_{l(j} R_{i)k} R_{rp} \ldots R_{sq} T_{ijp...q} = \pm T_{k\ell r...s}$$

by the transformation rule.

The definitions apply to any index pair. We say a tensor is totally symmetric or antisymmetric on a subset of indices if swapping any two index values (in the subset) produces a $\pm$ sign.

Hence $\delta_{ij}$ is symmetric and $\epsilon_{ijk}$ is totally antisymmetric (in all indices).

There are symmetric tensors of arbitrary rank $n$ in $\mathbb{R}^n$, but any antisymmetric tensor of rank 3 can be written $T_{ijk} = \lambda \epsilon_{ijk}$ (since only non-vanishing components of $T$ are those with $i, j, k$ distinct).

Similarly, there are no totally antisymmetric tensors of rank $n > 3$ in $\mathbb{R}^3$. 

$$\Rightarrow \frac{\partial x_p}{\partial x_i'} = R_{ip}, \quad \frac{\partial}{\partial x_i'} = R_{ip} \frac{\partial}{\partial x_p}$$

Thus, e.g. $n = 2$.
4.2.5 Tensor fields and derivatives

A tensor field of rank \( n \), \( T_{ij...k}(x) \), is a tensor-valued function which depends smoothly on position \( x \). Taking \( m \) derivatives gives a tensor field

\[
\frac{\partial}{\partial x_p} \frac{\partial}{\partial x_q} ... \frac{\partial}{\partial x_r} T_{ij...k}
\]

of rank \( n + m \). This follows from the chain rule;

\[
\frac{\partial}{\partial x_i'} = R_{ip} \frac{\partial}{\partial x_p}
\]
as in 4.1.2.

4.3 Second rank tensors

4.3.1 Decomposition of a second rank tensor

Any tensor \( T_{ij} \) can be written as a sum of its symmetric and antisymmetric parts;

\[
T_{ij} = S_{ij} + A_{ij}, \quad S_{ij} = \frac{1}{2}(T_{ij} + T_{ji}) = S_{ji}, \quad A_{ij} = \frac{1}{2}(T_{ij} - T_{ji}) = -A_{ji}
\]

Check; \( T_{ij} \) has 9 independent components, but \( S_{ij} \) has 6 and \( A_{ij} \) only 3.

The symmetric part can be further reduced;

\[
S_{ij} = P_{ij} + \frac{1}{3} \delta_{ij} Q
\]

where \( P_{ii} = 0 \) (so \( P_{ij} \) is traceless), \( \delta_{ii} = 3 \), and \( Q \) is \( S_{ii} \) = the trace of \( S_{ij} \)

Check; \( S_{ij} \) has 6 independent components, but \( P_{ij} \) has 5 and \( Q \) only 1.

The antisymmetric part can be re-expressed as a vector by setting

\[
A_{ij} = \varepsilon_{ijk} B_k \iff B_k = \frac{1}{2} \varepsilon_{ijk} A_{ij}
\]

with equivalence following from \( \varepsilon_{ijk} \varepsilon_{pqk} = \delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp} \) and \( \varepsilon_{ijk} \varepsilon_{ijp} = 2 \delta_{kp} \).

In summary,

\[
T_{ij} = P_{ij} + \varepsilon_{ijk} B_k + \frac{1}{3} \delta_{ij} Q
\]

where \( P_{ij} \) is symmetric traceless, \( B_k \) is a vector and \( Q \) (trace) is a scalar. Also,

\[
B_k = \frac{1}{2} \varepsilon_{ijk} T_{ij}, \quad Q = T_{kk}
\]

because only the parts of \( T_{ij} \) with correct symmetry contribute.

Example. Consider a vector field \( u(x) \) and its first derivatives

\[
\frac{\partial u_j}{\partial x_i} = T_{ij} = P_{ij} + \varepsilon_{ijk} B_k + \frac{1}{3} \delta_{ij} Q
\]

where

\[
Q = T_{kk} = \frac{\partial u_k}{\partial x_k} = \nabla \cdot u, \quad B_k = \frac{1}{2} \varepsilon_{kij} T_{ij} = \frac{1}{2} \varepsilon_{kij} \frac{\partial u_j}{\partial x_i} = \frac{1}{2} (\nabla \times u)_k
\]

Geometrical interpretation of these expressions was discussed previously. Now in addition we have

\[
P_{ij} = \frac{1}{2} \left( \frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right) - \frac{1}{3} \delta_{ij} \frac{\partial u_k}{\partial x_k}
\]

This contains all other information about the first derivatives of \( u \).
4.3.2 Diagonalisation of symmetric tensors

Let $M' = (M'_{ij})$ and $M = (M_{pq})$ be components of a 2nd rank tensor, related by a rotation $R = (R_{ip})$. The tensor transformation rule in matrix form is $M' = RM'R^T$. In general;

(i) $M$ need not be diagonalisable
(ii) if it is, the change of basis need not be orthogonal.

But if $M$ is symmetric, then it is always diagonalisable by an orthogonal change of basis.

The new basis consists of eigenvectors $e_i'$ of $M$ with eigenvalues $\lambda_i$ say, and then

$$M' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

The orthogonal directions given by $e_i'$ are called the principal axes of the tensor.

**Example (Inertia tensor).** Recall the velocity of a point $x$ rotating with angular velocity $\omega$ about $0$ is $v = \omega \times x$, and the angular momentum for a mass $m$ at this point is

$$x \times m v = m x \times (\omega \times x) = m(x^2 \omega - (\omega \cdot x)x)$$

For a rigid body occupying volume with density $\rho(x)$, the total angular momentum about $0$ is

$$L = \int_V \rho(x)(x^2 \omega - (\omega \cdot x)x) \, dV$$

or $L_i = I_{ij}\omega_j$, where

$$I_{ij} = \int_V \rho(x)(x^2 \delta_{ij} - x_i x_j) \, dV$$

is the inertia tensor about $0$. (Normally take $0$ at the centre of mass of the body.)

Since $I_{ij}$ is symmetric, it can be diagonalised. The eigenvalues $\lambda_i$ are called the principal moments of inertia.

If $I_{ij}$ is diagonal in coordinate system $x_i$, then $\omega = \omega e_i$ and $L = \lambda_i \omega e_i = \lambda_i \omega$ (no sum!) for each $i = 1,2,3$.

Principal axes can often be identified on basis of symmetry. Choosing coordinates with these axes, we expect $I_{11} = \lambda_1, I_{22} = \lambda_2, I_{33} = \lambda_3$ and $I_{ij} = 0$ for $i \neq j$.

**Example (sphere).** Consider a sphere $V : |x| \leq a$ with density $\rho(x) = \rho_0(1 + \alpha \cos^2 \theta)$ in spherical polars. For $\alpha = 0$ we have spherical symmetry; for $\alpha \neq 0$ we still have symmetry under rotations about $x_3$ axis.

Direct calculation:

$$I_{33} = \int_V \rho(x)(x_1^2 + x_2^2) \, dV = \int_V \rho(x)r^4 \sin^2 \theta \, dV$$

$$= \frac{1}{5} a^5 \times 2\pi \rho_0 \times \int_0^\pi (\sin \theta (1 + (\alpha - 1) \cos^2 \theta - \alpha \cos^4 \theta)$$

since $dV = r^2 \sin \theta \, dr \, d\theta \, d\varphi$. This gives us

$$= \frac{2\pi \rho_0}{5} a^5 \left(2 \left(1 + \frac{1}{3}(\alpha - 1) - \frac{1}{5}\alpha\right)\right) = \frac{8\pi \rho_0}{15} a^5(1 + \frac{2}{5}\alpha)$$

Now also

$$I_{11} = \int_V \rho(x)(x_1^2 + x_2^2) \, dV = \int_V \rho(x)r^2 (\sin^2 \theta \sin^2 \varphi + \cos^2 \theta) \, dV$$

$$= \frac{1}{5} a^5 \times \rho_0 \times \pi \int_0^\pi \sin \theta (1 + \cos^2 \theta)(1 + \alpha \cos^2 \theta) \, \, d\vartheta = \frac{8\pi \rho_0}{15} a^5 \left(1 + \frac{2\alpha}{5}\right)$$

Finally, off-diagonal elements vanish. For example;
\[ I_{13} = -\int_V \rho(x) x_1 x_3 \, dV \]

contains

\[ \int_0^{2\pi} \cos \varphi \, d\varphi = 0. \]

Note that for \( \alpha = 0 \) we get

\[ I_{11} = I_{22} = I_{33} = \frac{8\pi \rho_0}{15} \alpha^5 = \frac{2}{5} Ma^2 \]

where \( M \) is the mass of the sphere.

### 4.4 Invariant and isotropic tensors

#### 4.4.1 Definitions and key results

**Definition.** A tensor \( T \) is invariant under a particular rotation \( R \) if \( T'_{ij...k} = R_{ip} R_{jq} \cdots R_{kr} T_{pq...r} = T_{ij...k} \). If \( T \) is invariant under *any* rotation \( R \), it is called *isotropic*.

**Key results; in \( \mathbb{R}^3 \)**

- The most general isotropic tensor of rank 2 is \( T_{ij} = \alpha \delta_{ij} \).
- The most general isotropic tensor of rank 3 is \( T_{ij} = \beta \varepsilon_{ijk} \).

These properties ensure that the definitions and results are independent of the Cartesian coordinate system.

All isotropic tensors of higher rank are obtained by combining \( \delta \) and \( \varepsilon \) using tensor product and contraction. E.g. the most general isotropic tensor of rank 4 is \( \alpha \delta_{ij} \delta_{kl} + \beta \varepsilon_{ijkl} + \gamma \delta_{il} \delta_{jk} \) for some \( \alpha, \beta, \gamma \).

#### 4.4.2 Proof of key results for rank 2 and 3

For rank 2, isotropic \( \Rightarrow T_{ij} = R_{ip} R_{jq} T_{pq} \) for any orthogonal \( R \). Consider

\[ (R_{ij}) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \]

a rotation through \( \pi/2 \) about the \( x_3 \) axis. Then

\[ T_{11} = R_{ip} R_{1q} T_{pq} = T_{22}, \]
\[ T_{22} = R_{2p} R_{2q} T_{pq} = T_{11}, \]
\[ T_{13} = R_{ip} R_{3q} T_{pq} = T_{23}, \]
\[ T_{23} = R_{2p} R_{3q} T_{pq} = -T_{13}. \]

Hence \( T_{13} = T_{23} = 0 \), and similarly for all other off-diagonal elements, using \( \pi/2 \) rotations about the \( x_1 \) or \( x_2 \) axes. All diagonal entries must also coincide; \( T_{11} = T_{22} = T_{33} \) and this finishes the proof.

For rank 3, isotropic \( \Rightarrow T_{ijk} = R_{ip} R_{jq} R_{kr} T_{pqr} \). Use rotation \( R \) through \( \pi/2 \) about the \( x_3 \) axis as before, and find

\[ T_{133} = R_{ip} R_{3q} R_{3r} T_{pqr} = T_{233}, \]

but \( T_{233} = -T_{133} = 0 \) just as in the rank 2 case.

Now \( T_{111} = T_{222} \) (since \( R_{12} = 1 \)) but \( T_{222} = -T_{111} \) (since \( R_{21} = -1 \)). Hence both vanish.

By rotations about other axes, we deduce similarly that the only non-zero components are \( T_{ijk} \) with \( i, j, k \) distinct. Then \( T_{123} = R_{1p} R_{2q} R_{3r} T_{pqr} = -T_{213} \). Using all possible \( \pi/2 \) rotations, we find \( T_{ijk} \) is totally antisymmetric and hence a multiple of \( \varepsilon_{ijk} \).
Our results for isotropic tensors in \( \mathbb{R}^3 \) can be generalised easily to \( \mathbb{R}^n \), where we find \( \delta_{ij} \) and \( \varepsilon_{ij...k} \) (\( n \) values). In particular, note that for \( n = 2 \) we have two rank 2 isotropic tensors.

### 4.4.3 Application to evaluating integrals

Consider

\[
T_{ij...k} = \int_V f(x)x_i x_j ... x_k \, dV
\]

Under a change of variables \( x_i' = R_{ip} x_p \) (\( R \) a rotation). The Jacobian

\[
\det \left( \frac{\partial x_i'}{\partial x_p} \right) = \det R_{ip} = 1
\]

and hence \( dV' = dV \) implying

\[
T_{ij...k} = \int_{V'} f(x')x_i' x_j' ... x_k' \, dV'
\]

If, for this \( R \), (i) \( V' = V \) (the region of integration is invariant) and (ii) \( f(x') = f(x) \), then

\[
T_{ij...k} = \int_V f(x)R_{ip} x_p R_{jq} x_q ... R_{kr} x_r \, dV = R_{ip} R_{jq} ... R_{kr} T_{pq...r}
\]

i.e. \( T \) is invariant under \( R \). If this holds for any \( R \), then \( T \) is isotropic.

**Example.** \( V \) a sphere, centre \( 0 \), and \( f(x) \) spherically symmetric (invariant under all rotations) then

\[
T_{ij} = \int_V f(x)x_i x_j \, dV = I \delta_{ij}
\]

(since isotropic). So we just need \( 3I = T_{ii} = \int_V f(x)x^2 \, dV \), a scalar integral.

### 4.5 The quotient theorem and some related ideas

An array of components \( T_{ij...k} \) is determined uniquely by specifying

(i) \( T_{ij...k} u_{ij} v_{ij} ... w_k \) for arbitrary vectors \( u, v, w \) or

(ii) \( T_{ij...k} A_{ij...k} \) for an arbitrary tensor \( A \).

### 4.5.1 The quotient theorem

Let \( T_{i...jp...q} \) be an array defined for each Cartesian coordinate system. If for any tensor \( A_{p...q} \) of rank \( m \), \( B_{i...j} = T_{i...jp...q} A_{p...q} \) is a tensor of rank \( n \), then \( T_{i...jp...q} \) is a tensor of rank \( m + n \).

**Proof.**

\[
B'_{k...l} = R_{kl} ... R_{ij} B_{i...j} = R_{kl} ... R_{ij} T_{i...jp...q} A_{p...q} = T'_{k...l...r...s} A'_{r...s} = T'_{k...l...r...s} R_{rp} ... R_{sq} A_{p...q}
\]

Since this holds for arbitrary \( A_{p...q} \), we have

\[
R_{kl} ... R_{ij} T_{i...jp...q} = T'_{k...l...r...s} R_{rp} ... R_{sq}
\]

Multiplying by \( R_{ap} ... R_{bq} \) now gives

\[
T'_{k...l...a...b} = R_{kl} ... R_{ij} R_{ap} ... R_{bq} T_{i...jp...q}
\]

which is the required tensor transformation rule.

The special case \( m = n = 1 \) (\( A \) and \( B \) vectors and \( T \) a matrix) was discussed and proved in 4.1.2. So e.g. in conductivity, \( J_i = \sigma_{ij} E_j \) and \( \sigma_{ij} \) must be a tensor. Similarly, in elasticity

\[
\sigma_{ij} = C_{ijkl} e_{kl}
\]

where \( \sigma \) represents stress and \( e \) strain. Since \( \sigma_{ij}, e_{ij} \) are tensors, \( C_{ijkl} \) also a tensor.

A common special case is \( m = 0 \) (\( B \) is a scalar): if \( B = T_{i...j} A_{i...j} \) is a scalar, then quotient theorem ensures that \( A \) a tensor \( \Rightarrow T \) a tensor.
**4.5.2 Tensor divergence theorem**

Let $V$ be a volume bounded by a smooth surface $S = \partial V$ and let $T_{ij...kl}$ be a smooth tensor field. Then

$$\int_{S} T_{ij...kl} n_i \, dS = \int_{V} \frac{\partial}{\partial x_i} (T_{ij...kl}) \, dV$$

where $n_i$ is the outward normal to $S$.

*Proof.* Apply divergence theorem to the vector field $v_1 = a_i b_j ... c_k T_{ij...kl}$ for arbitrary constant vectors $\mathbf{a}, \mathbf{b}, ..., \mathbf{c}$; then remove vectors at the end (since they are arbitrary).

**4.5.3 A coordinate-free approach to tensors (non-examinable)**

A tensor of rank $n$ can be defined as a map on $n$ vectors to $\mathbb{R}$;

$$T(\mathbf{u}, \mathbf{v}, ... \mathbf{w})$$

which is multi-linear (i.e. linear in each argument). To make contact with component definition, write $\mathbf{u} = u_i \mathbf{e}_i, \mathbf{v} = v_j \mathbf{e}_j, \mathbf{w} = w_k \mathbf{e}_k$ and we find

$$T(\mathbf{u}, \mathbf{v}, ... \mathbf{w}) = u_i v_j ... w_k T(\mathbf{e}_i, \mathbf{e}_j, ..., \mathbf{e}_k) = u_i v_j w_k T_{ij...k}$$

where $T_{ij...k} = T(\mathbf{e}_i, \mathbf{e}_j, ..., \mathbf{e}_k)$. The transformation rule for components in moving to a new basis $\mathbf{e}'_i = R_{ip} \mathbf{e}_p$ is now easily checked (using multi-linearity)