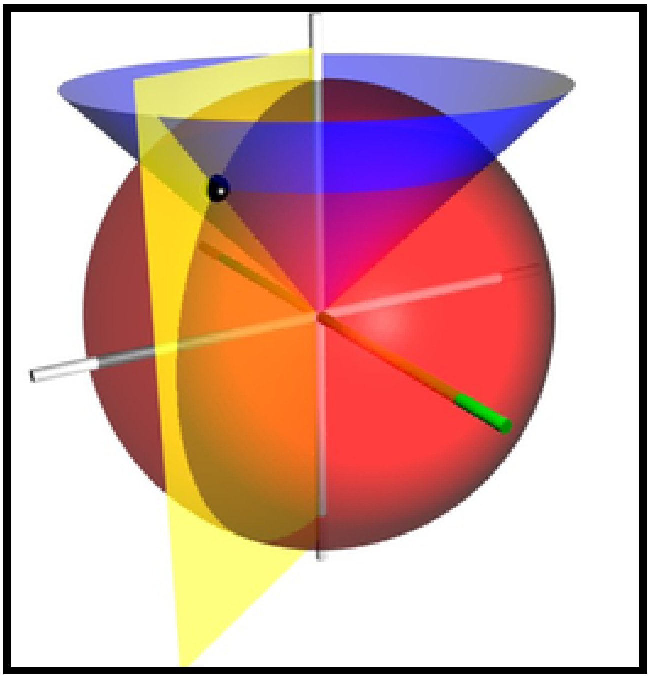
University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

School of Mathematics and Statistics
Scoil na Matamaitice agus na Staitisticí

Vector Integral and Differential Calculus (ACM 20150)



Dr Lennon Ó Náraigh

Lecture notes in Vector Calculus, September 2017

Vector Integral and Differential Calculus (ACM20150)

- Subject: Applied and Computational Mathematics
- School: Mathematical Sciences
- Module coordinator: Dr Lennon Ó Náraigh
- Credits: 5
- Level: 2
- Semester: First

This module introduces the fundamental concepts and methods in the differentiation and integration of vector-valued functions.

- **Fundamentals:** Vectors and scalars, the dot and cross products, the geometry of lines and planes
- **Taylor's theorem in several dimensions**
- **Curves in three-dimensional space:** Differentiation of curves, the tangent vector
- **Partial derivatives and vector fields:**
- Introduction to partial derivatives, scalar and (Cartesian) vector fields, the operators div, grad, and curl in the Cartesian framework, applications of vector differentiation in electromagnetism and fluid mechanics
- **Mutli-variate integration:** Area and volume as integrals, integrals of vector and scalar fields, Stokes's and Gauss's theorems (statement and proof)
- **Consequences of Stokes's and Gauss's theorems:** Green's theorems, the connection between vector fields that are derivable from a potential and irrotational vector fields
- **Curvilinear coordinate systems:** Basic concepts, the metric tensor, scale factors, div, grad, and curl in a general orthogonal curvilinear system, special curvilinear systems including spherical and cylindrical polar coordinates

Further topics may include:

- Introduction to differential forms, exact and inexact differential forms
- **Advanced integration** Integrating the Gaussian function using polar coordinates, the gamma function, the volume of a four-ball by appropriate coordinate parameterization, the volume of a ball in an arbitrary (finite) number of dimensions using the gamma function

What will I learn?

On completion of this module students should be able to

1. Write down parametric equations for lines and planes, and perform standard calculations based on these equations (e.g. points/lines of intersection, condition for lines to be skew);
2. Compute Taylor expansions for functions of one and several variables;
3. Differentiate scalar and vector fields expressed in a Cartesian framework;
4. Perform operations involving div, grad, and curl;
5. Perform line, surface, and volume integrals. The geometric objects involved in the integrals may be lines, arbitrary curves, simple surfaces, and simple volumes, e.g. cubes, spheres, cylinders, and pyramids;
6. State precisely and prove Gauss's and Stokes's theorems;
7. Derive corollaries of these theorems, including Green's theorems and the necessary and sufficient condition for a vector field to be derivable from a potential;
8. Compute the scale factors for arbitrary orthogonal curvilinear coordinate systems;
9. Apply the formulas for div, grad, and curl in arbitrary orthogonal curvilinear coordinate systems

Editions

First edition: September 2010

Second edition: September 2011

Third edition: September 2012

Fourth edition: September 2013

Fifth edition: September 2014

Sixth edition: September 2015 – with substantial modifications compared to previous years

Seventh edition: September 2016

This edition: September 2017

Contents

Chapter 1

Introduction

1.1 Overview

Here is the executive summary of the module:

> This module involves the study of vector and scalar fields in two- and three-dimensional space. A *field* is an object that assigns a vector or a scalar to each point in space. We need to find out how to integrate and differentiate these things, hence vector integral and differential calculus.

In more detail, a field is a map that assigns either a scalar or a vector to each point in the map domain, giving scalar- and vector-fields, respectively. We study this concept in depth:

1. We review concepts of vectors, scalars, and vector algebra.
2. We review Taylor series in a single dimension and introduce Taylor series in several dimensions. It is quite simply impossible to do this module without knowledge of such series and the associated Taylor expansions.
3. We formulate the derivative of a scalar field, based on the *gradient* operator.
4. We learn how to differentiate vector fields, using the *divergence* and *curl* operators.
5. We define line and area integrals – generalization of integration on $\mathbb{R}$.
6. We state and prove two fundamental theorems of vector integration – Gauss' and Stokes' theorems.[1] These can be crudely thought of as generalizations of integration by parts.

[1] Sir George Gabriel Stokes F.R.S. Born in Skreen Co. Sligo, 1829, died in Cambridge, England, 1903.

7. These topics are formulated against the backdrop of Cartesian space (that is, a triple $(x, y, z) \in \mathbb{R}^3$ labels points in space). However, the integration theorems enable us to generalize div, grad, and curl to differentiation on curved surfaces ('manifolds').

1.2 Learning and Assessment

Learning:

- Thirty six classes, three per week.[2]
- Optional tutorial.
- Graded homework assignments (see the assessment section below) and non-graded homework assignments.
- One of the main goals of any degree in the mathematical sciences is to put problems into a quantitative framework and hence to solve them. This will be accomplished through the homework assignments, and *autonomous student learning*, or what used to be known less grandiosely as *studying*. One way of studying is by independently practising the non-graded homework assignments and by subsequently going through the answers in the tutorial classes. You can also try reading a textbook about the subject. The Schaum's textbook (see below) contains lots of further supplementary problems for practice for these purposes.

Assessment:

- **Two** homework assignments, for a total of 10%;
- **One** in-class test, for a total of 20%;
- One end-of-semester exam, 70%;

Percentage scores-to-grades conversion:

In UCD, percentage scores have no meaning, and it is the letter grades that matter. The reason for this is that letter grades, although qualitative in nature, have a very precise meaning. Thus, each letter grade comes with a 'grade descriptor' which defines what that particular letter grade means (e.g. 'A' means 'Excellent', 'B' means 'Very Good' etc.). The precise grade descriptors used for this module (indeed, for all UCD modules) are found in the following document:

`http://www.ucd.ie/registry/assessment/staff_info/modular grades explained staff.pdf`

[2]Really, thirty four classes because of the October Bank Holiday

Now, sometimes it is more convenient to award a numerical score to a piece of assessment and hence to assign percentage scores, and from there to assign letter grades using a percentage-to-grades conversion scheme. But the manner in which the percentage score is translated into a qualitative grade is not set in stone and will vary from discipline to discipline. According to official UCD policy,

> The award of grades is a matter of academic judgement against agreed criteria (learning outcomes and grade descriptors) and should not be simply a mathematical exercise... clear grade descriptors will assist you in this effort....
>
> There may be circumstances where an examiner wishes to convert percentage marks to grades using a different scale [from the one used elsewhere in UCD] for pedagogical reasons related to the form of the assessment... this grading scheme is a matter for the individual examiner. It is good practice to describe the grading scheme and the rationale to students in advance.

This policy can be found in the UCD Registry document alluded to earlier. In any case, any percentage-to-grades conversion scheme should have the following characteristics:

- It should accurately reflect differences in student performance.
- It should be clear to students so that they should be able to chart their progress.
- It should be fair.

For these reasons, the percentage-to-grades conversion scheme of the School of Mathematics and Statistics is hereby adopted for this module:

`http://mathsci.ucd.ie/tl/grading/en06`

Policy on late submission of homework:

The official UCD policy will be absolutely strictly adhered to[3]. As such, coursework that is late by up to one week after the due date will have the grade awarded reduced by 10% (e.g. from 87% to 77%); coursework submitted up to two weeks after the due date will have the grade reduced by 20% (e.g. from 87% to 67%). Coursework received more than two weeks after the due date will not be accepted.

In addition, it will not be possible for me to accept late coursework via email in a scan or otherwise – it will be the student's responsibility to submit such late homework directly to me in paper form. There are many reasons for this – not least because of incidents in the past where I have had to print

[3] `http://www.ucd.ie/registry/academicsecretariat/docs/latesub_po.pdf`

down pages and pages of late submissions consisting of jpeg after jpeg of mobile-phone photographs of homework.

Policy on extenuating circumstances:

This module adheres to the policy on extenuating circumstances as outlined in the UCD Science Handbook:

- Serious issues (serious illness, hospitalization, etc.) are dealt with through a formal process administered by the College of Science;
- Minor issues (e.g. assignment/midterm missed due to minor illness) are dealt with through direct contact with the lecturer – via email, with supporting documentation as necessary. In order to keep track of such extenuating circumstances, please use the phase 'minor extenuating circumstances ACM20150' in the subject line in the email.

Plagiarism:

Plagiarism is a **serious academic offence**. While plagiarism may be easy to commit unintentionally, it is defined by the act not the intention.

- All students are responsible for being familiar with the Universitys policy statement on plagiarism and are encouraged, if in doubt, to seek guidance from an academic member of staff.
- The University encourages students to adopt good academic practice by maintaining academic integrity in the presentation of all academic work.
- For more detailed information see:

  `http://www.ucd.ie/governance/resources/policypage-plagiarismpolicy/`

Office hours:

I do not keep specific office hours. If you have a question, you can visit me whenever you like – from 09:00-17:00. I am usually in my office if not lecturing. The office is E0.88 on the ground floor of Science East. It is a little complicated to get to, although there are directions on the homepage of my website. Thus, emailing me in the first instance might be a better idea:

`onaraigh@maths.ucd.ie`

Textbooks

- Lecture notes will be put on the web. These are self-contained. They will be available *before* class. It is anticipated that you will print them and bring them with you to class. You can then annotate them and follow the proofs and calculations done on the board. Thus, you are still expected to attend class, and I will occasionally deviate from the content of the notes, give hints about solving the homework problems, or give a revision tips for the final exam.

- There are some books for extra reading, if desired:

  – *Vector analysis and an introduction to tensor analysis*, M. R. Spiegel, Schaum's Outline Series, McGraw–Hill (Three copies in library, 515.63 SPI).

  – *Mathematical methods for physicists*, G. B. Arfken, H. J. Weber, and F. Harris, Wiley, Fifth Edition (Two copies of seventh edition in library, 510.2453 ARF).

  – *Vectors, tensors and the basic equations of fluid mechanics*, R. Aris, Dover (One copy in library, 532.00151 ARI; also available for $£8.00$ on Amazon.co.uk).

1.3 A modern perspective on vector calculus

Before beginning the lecture course, let us discuss a contemporary problem that uses the techniques of vector calculus.

The advection-diffusion equation: The concentration C of a chemical in the atmosphere, a pollutant on the sea-surface, or of a blob of dye in a container of fluid is a function of space and time:

$$C = C\left(\boldsymbol{x}, t\right), \qquad \boldsymbol{x} = \begin{cases} (x, y) \in \Omega \subset \mathbb{R}^2, & \text{or} \\ (x, y, z) \in \Omega \subset \mathbb{R}^3. & \end{cases}$$

This concentration is stirred around by the flow field

$$\boldsymbol{u} = \left(u\left(\boldsymbol{x}, t\right), v\left(\boldsymbol{x}, t\right)\right)$$

in two dimensions, or

$$\boldsymbol{u} = \left(u\left(\boldsymbol{x}, t\right), v\left(\boldsymbol{x}, t\right), w\left(\boldsymbol{x}, t\right)\right)$$

in three dimensions. The flow is assumed to be incompressible: this means that density is conserved along streamlines; mathematically,

$$\nabla \cdot \boldsymbol{u} = 0.$$

At the same time, the concentration is 'diffused', so that regions where the concentration possesses high gradients are smoothed out, on a timescale

$$T = [\text{Length scale of variation}]^2 / D$$

where D is the *diffusion coefficient*. The law that expresses these two processes is called the *advection-diffusion equation*

$$\underbrace{\frac{\partial C}{\partial t}}_{\text{Instantaneous changes in concentration}} + \underbrace{\boldsymbol{u} \cdot \nabla C}_{\text{Stirring by the flow}} = \underbrace{D \nabla^2 C}_{\text{Diffusion}}. \tag{1.1}$$

The integral theorems discussed previously can be used to show that

$$\frac{d}{dt} \int_\Omega C(\boldsymbol{x}, t)\, \mathrm{d}^n x = 0 + \text{Boundary terms},$$

hence, the total amount of chemical is *conserved*. If we multiply Eq. (1.1) by $C(\boldsymbol{x}, t)$ and integrate over the flow domain Ω, we obtain, using the same integral theorems as before,

$$\frac{d}{dt} \int_\Omega \tfrac{1}{2} C^2(\boldsymbol{x}, t)\, \mathrm{d}^n x = -D \int_\Omega |\nabla C(\boldsymbol{x}, t)|^2 \mathrm{d}^n x + \text{Boundary terms}.$$

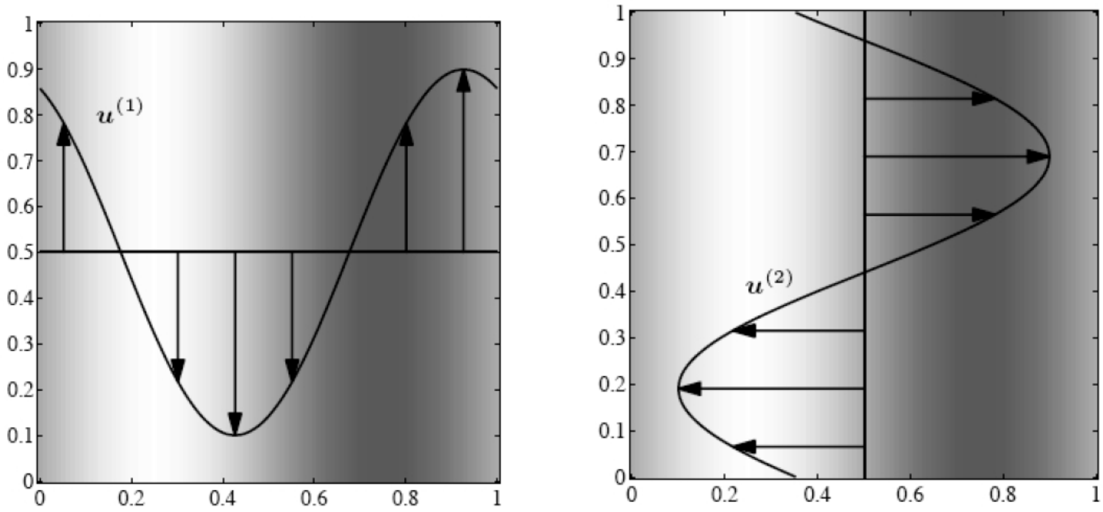
If the flow and the concentration gradients satisfy certain conditions on the boundary, the last term in this equation vanishes, and we are left with

$$\frac{d}{dt} \int_\Omega \tfrac{1}{2} C^2(\boldsymbol{x}, t)\, \mathrm{d}^n x = -D \int_\Omega |\nabla C(\boldsymbol{x}, t)|^2 \mathrm{d}^n x,$$

and the *variance* in the concentration, away from its mean value, decays to zero. Thus, the chemical becomes better and better mixed, over time.

The question of *how fast the mixing is* depends on the character of the flow. You will no doubt be aware of a certain experiment involving coffee and milk: if you add a drop of milk to a cup of black coffee and do not stir, the two components will eventually mix, but over a long interval. If you add the milk and then stir, the homogenization is faster. Mixing times therefore depend on the flow. It turns out that if the flow $\boldsymbol{u}$ is *chaotic* (in a sense described below), then the mixing is as close to optimal as can be imagined. A flow $\boldsymbol{u}$ is chaotic if two initially neighbouring fluid particles separate away from each other exponentially fast in time, under the influence of the flow. The average rate of separation is called the *Lyapunov exponent*, Λ_0.

One popular model of mixing in two dimensions is the *random-phase sine flow*, which is a succession



(a) First half-period

(b) Second half-period

Figure 1.1: Schematic description of the random-phase sine flow in each quasi-period.

of unidirectional quasi-periodic 'whisking' motions:

$$u = A_0 \sin(ky + \phi_j), \qquad v = 0, \tag{1.2}$$

in the first half-period of the flow, and

$$u = 0, \qquad v = A_0 \sin(kx + \psi_j), \tag{1.3}$$

in the second (See Fig. 1.1). Here ϕ_j and ψ_j are random phases that change after each whisking motion and A_0 and k are positive constants ('amplitude' and 'wavenumber' respectively). Particles drawn along by this flow satisfy the trajectory equation

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}(\boldsymbol{x}, t),$$

and can be tracked numerically. The time-averaged rate of separation along trajectories gives rise to the Lyapunov exponent $\Lambda_0(\boldsymbol{x})$ (Fig. 1.2), which varies in space but not in time (the spatial variations label the trajectories). The decay rate of the concentration can also be measured (it is exponential). The energy of the flow is the space-time average

$$E = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t \int_\Omega \mathrm{d}^2x |\boldsymbol{u}(\boldsymbol{x}, t)|^2.$$

By the end of this course, you should be able to see that

$$E = \tfrac{1}{2}A_0^2,$$

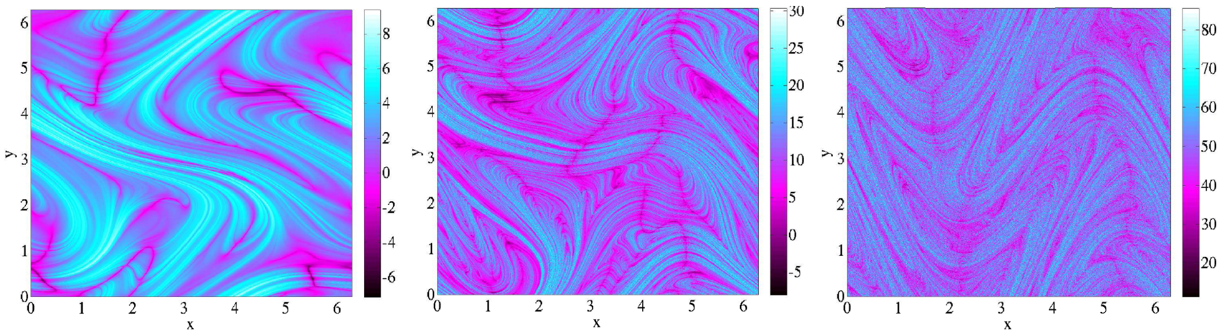
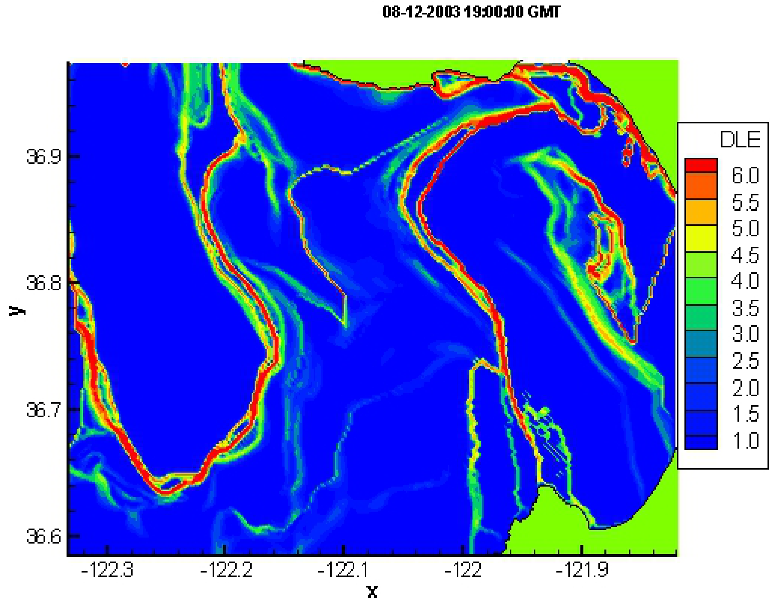


Figure 1.2: The Lyapunov exponent $\Lambda_0(\boldsymbol{x})$ for different trajectories (the constant Λ_0 is the average over all trajectories, and is positive). The larger the value of A_0, the larger the values taken by $\Lambda_0(\boldsymbol{x})$.

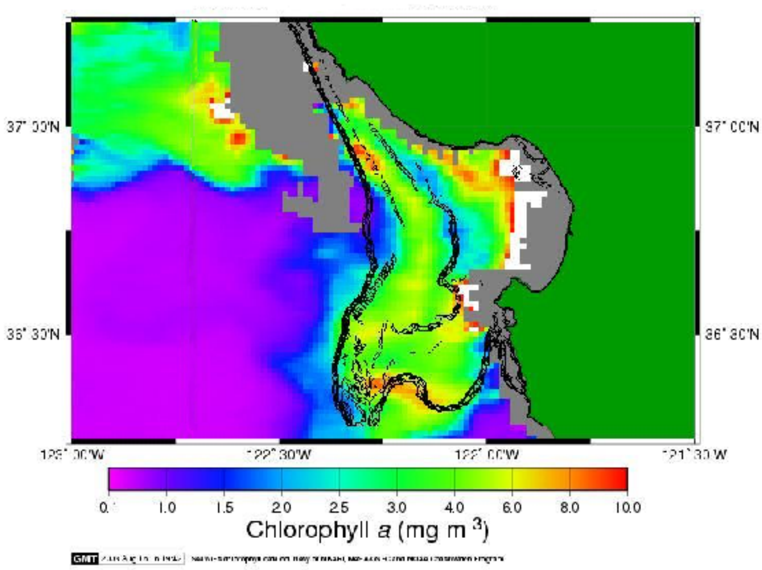
independent of wavenumber. Referring to Fig. 1.2, the more energy you put into the flow, the better mixed it becomes. This in part answers the question about stirring the cup of coffee: stirring, that is inputting mechanical energy (in the correct, chaotic, fashion), increases the Lyapunov exponent, and hence promotes mixing.

Note, from Fig. 1.2 that the Lyapunov exponent $\Lambda_0(\boldsymbol{x})$ can be calculated numerically for a given flow, and in fact is an *averaged separation rate*, averaged over an infinitely long time interval. There is a finite-time analogue, the finite-time Lyapunov exponent, when the temporal averaging is over a finite interval τ, and denoted by $\Lambda_0(\boldsymbol{x};\tau)$. Ridges in this quantity are called *Lagrangian coherent structures* [4]. A *ridge* is a local maximum *in only one direction*. Just as a ridge in a mountain range is a barrier to transport, so too is a ridge in the FTLE: particles cannot flow through them. Ridges can be found in the ocean and act as barriers to pollution dispersal, or to the uniform distribution of micro-organisms (as in Fig. 1.3). Before people discovered Lagrangian coherent structures, they thought tides would wash away pollution. However, these structures persist through tides, and they represent a permanent barrier.

[4] See 'Finding Order in the Apparent Chaos of Currents', New York Times, 28 September 2009.



(a)



(b)

Figure 1.3: (a) A *snapshot* of the FTLE in Monterey Bay, CA, *at a particular point in time* (they can evolve in time); (b) A snapshot of the distribution of sea-surface chlorphyll: this very clearly is contained within the transport barriers represented by the ridges in the FTLE. (Taken from the webpage http://www.cds.caltech.edu/~shawn)/LCS-tutorial/).

Chapter 2

Vectors – revision

Overview

We review some basics of vector algebra that have already been covered in MATH 10340.

2.1 The connection between vectors and Cartesian coordinates

A vector is a quantity with magnitude and direction. A point P in space can be labelled by coordinates (a_1, a_2, a_3) with respect to some Cartesian coordinate frame with origin O. The distance from O to P is thus

$$|OP| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Associated with O and P is a direction – from O to P. Thus, we identify $\overrightarrow{OP}$ as a vector with direction from O to P, with magnitude $\sqrt{a_1^2 + a_2^2 + a_3^2}$. We can also identify the vector by its coordinates, writing

$$\overrightarrow{OP} \equiv (a_1, a_2, a_3) \equiv \boldsymbol{a}.$$

Two vectors, $\overrightarrow{OP} = \boldsymbol{a} = (a_1, a_2, a_3)$ and $\overrightarrow{OQ} = \boldsymbol{b} = (b_1, b_2, b_3)$ can be added together in an obvious way:

$$\boldsymbol{a} + \boldsymbol{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

This is consistent with the parallelogram law of vector addition – see Fig. 2.1. We also have the notion of scalar multiplication: if $\lambda \in \mathbb{R}$, and if $\overrightarrow{OP} = \boldsymbol{a} = (a_1, a_2, a_3)$, then

$$\lambda\left(\overrightarrow{OP}\right) = \lambda\boldsymbol{a} = \lambda(a_1, a_2, a_3) \stackrel{\text{def}}{=} (\lambda a_1, \lambda a_2, \lambda a_3).$$

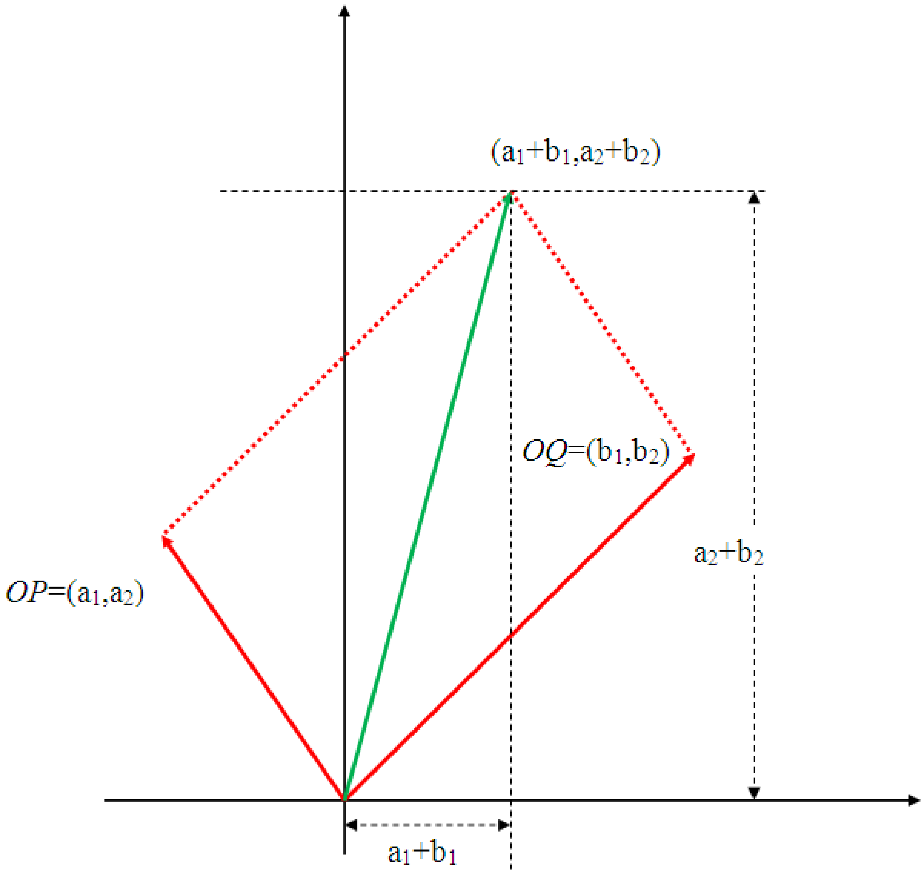


Figure 2.1: Parallelogram law for vector addition

In this way, we identify **unit vectors** (vectors of length one) that point along the three distinguished, mutually perpendicular directions of the Cartesian frame:

$$\hat{\boldsymbol{x}} = (1, 0, 0), \qquad \hat{\boldsymbol{y}} = (0, 1, 0), \qquad \hat{\boldsymbol{z}} = (0, 0, 1).$$

This introduces further consistency to the identification of triples (e.g. (a_1, a_2, a_3)) with vectors, since

$$\overrightarrow{OP} = \boldsymbol{a} = (a_1, a_2, a_3) = a_1(1, 0, 0) + a_2(0, 1, 0) + a_3(0, 0, 1) = a_1\hat{\boldsymbol{x}} + a_2\hat{\boldsymbol{y}} + a_3\hat{\boldsymbol{z}}.$$

2.2 The dot product

Definition 2.1 *Take two vectors* $\boldsymbol{a} = (a_1, a_2, a_3)$ *and* $\boldsymbol{b} = (b_1, b_2, b_3)$. *The dot product is a combination of these two vectors that returns a scalar, and is defined as follows:*

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

The dot product inherits many of the usual properties of ordinary multiplication:

1. Commutative: $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$,

2. Distributive: $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$. Also, $(\boldsymbol{a} + \boldsymbol{b}) \cdot \boldsymbol{c} = \boldsymbol{a} \cdot \boldsymbol{c} + \boldsymbol{b} \cdot \boldsymbol{c}$,

for all $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ in $\mathbb{R}^3$. Here $\mathbb{R}^3$ denotes all triples (x, y, z), where x, y, and z are real numbers; equivalently, it denotes all points in three-dimensional space.

The dot product can also be used to compute the length (magnitude) of a vector, as

$$\mathrm{mag}(\boldsymbol{a}) = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}}.$$

Henceforth, we denote $\mathrm{mag}(\boldsymbol{a})$ as $|\boldsymbol{a}|$.

Using the properties of dot-product multiplication, we can prove the following theorem:

Theorem 2.1 *Let* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *be vectors in* $\mathbb{R}^3$. *Then*

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

where $0 \leq \theta \leq \pi$ *is the angle between* $\boldsymbol{a}$ *and* $\boldsymbol{b}$.

Proof: Introduce $\boldsymbol{c} := \boldsymbol{a} - \boldsymbol{b}$. We apply the laws of dot-product multiplication to obtain

$$\begin{aligned} \boldsymbol{c} \cdot \boldsymbol{c} &= (\boldsymbol{a} - \boldsymbol{b}) \cdot (\boldsymbol{a} - \boldsymbol{b}), && (2.1) \\ &= \boldsymbol{a} \cdot \boldsymbol{a} - 2\boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{b} \cdot \boldsymbol{b}, && (2.2) \\ |\boldsymbol{c}|^2 &= |\boldsymbol{a}|^2 - 2\boldsymbol{a} \cdot \boldsymbol{b} + |\boldsymbol{b}|^2. && (2.3) \end{aligned}$$

However, we refer to the triangle in Fig. 2.2,

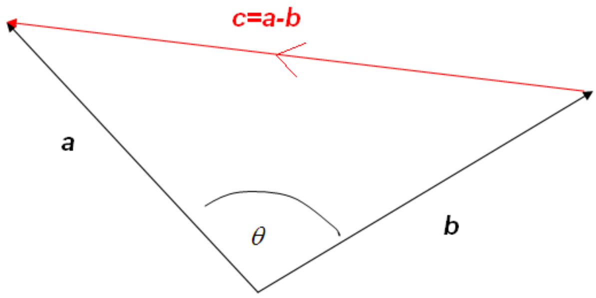


Figure 2.2: Sketch for applying the cosine rule to the dot-product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$

and we apply the cosine rule, to obtain

$$|\boldsymbol{c}|^2 = |\boldsymbol{a}|^2 + |\boldsymbol{b}|^2 - 2|\boldsymbol{a}||\boldsymbol{b}|\cos\theta. \tag{2.4}$$

Equating Equations (2.3) and (2.4), we obtain

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

as required. ∎

Corollary 2.1 *Two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are* **orthogonal** *(perpendicular) if and only if*

$$\boldsymbol{a}\cdot\boldsymbol{b} = 0.$$

Example: consider

$$\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}} = (1,0,0)\cdot(0,1,0) = 1\times 0 + 0\times 0 + 0\times 0 = 0.$$

Not surprisingly, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ have zero dot product (and hence, are orthogonal), as they point along different mutually-perpendicular axes. Also, $\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{x}} = 1$ etc. The vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ are called an **orthonormal triad**.

2.3 The vector or cross product

Given vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we have seen how to form a scalar. We can also form a third vector from these two vectors, using the cross or vector product:

Definition 2.2 (Cross product)

$$\begin{aligned}
\boldsymbol{a} \times \boldsymbol{b} &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}, \qquad (2.5) \\
&= \hat{\boldsymbol{x}}\left(a_2 b_3 - a_3 b_2\right) - \hat{\boldsymbol{y}}\left(a_1 b_3 - a_3 b_1\right) + \hat{\boldsymbol{z}}\left(a_1 b_2 - a_2 b_1\right), \\
&= \hat{\boldsymbol{x}}\left(a_2 b_3 - a_3 b_2\right) + \hat{\boldsymbol{y}}\left(a_3 b_1 - a_1 b_3\right) + \hat{\boldsymbol{z}}\left(a_1 b_2 - a_2 b_1\right),.
\end{aligned}$$

Properties of the vector or cross product:

1. Skew-symmetry: $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$,
2. Linearity: $(\lambda \boldsymbol{a}) \times \boldsymbol{b} = \boldsymbol{a} \times (\lambda \boldsymbol{b}) = \lambda\left(\boldsymbol{a} \times \boldsymbol{b}\right)$, for $\lambda \in \mathbb{R}$.
3. Distributive: $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$.

These results readily follow from the determinant definition. Result (1) is particularly weird. Note:

$$\begin{aligned}
\boldsymbol{a} \times \boldsymbol{a} &= -\boldsymbol{a} \times \boldsymbol{a}, \qquad \text{Result (1)}, \\
2\boldsymbol{a} \times \boldsymbol{a} &= 0, \\
\boldsymbol{a} \times \boldsymbol{a} &= 0.
\end{aligned}$$

Example: Let

$$\boldsymbol{a} = \hat{\boldsymbol{x}} + 3\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}, \qquad \boldsymbol{b} = 2\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + 2\hat{\boldsymbol{z}}.$$

Then

$$\boldsymbol{a} \times \boldsymbol{b} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ 1 & 3 & 1 \\ 2 & -1 & 2 \end{vmatrix} = 7\hat{\boldsymbol{x}} - 7\hat{\boldsymbol{z}}.$$

Example: The orthonormal triad $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ satisfy

$$\begin{aligned}\hat{\boldsymbol{x}} \times \hat{\boldsymbol{y}} &= \hat{\boldsymbol{z}}, \\ \hat{\boldsymbol{y}} \times \hat{\boldsymbol{z}} &= \hat{\boldsymbol{x}}, \\ \hat{\boldsymbol{z}} \times \hat{\boldsymbol{x}} &= \hat{\boldsymbol{y}}. \end{aligned} \tag{2.6}$$

2.4 Geometrical treatment of cross product

So far, our treatment of the cross product has been in terms of a particular choice of Cartesian axes. However, the definition of the cross product is in fact independent of any choice of such axes. To demonstrate this, we re-construct the cross product.

Step 1: Finding the length of $\boldsymbol{a} \times \boldsymbol{b}$

Note that

$$\begin{aligned}|\boldsymbol{a} \times \boldsymbol{b}|^2 + (\boldsymbol{a} \cdot \boldsymbol{b})^2 &= (a_2 b_3 - a_3 b_2)^2 + (a_3 b_1 - a_1 b_3)^2 + (a_1 b_2 - a_2 b_1)^2 \\ &+ (a_1 b_1 + a_2 b_2 + a_3 b_3)^2, \\ &= \left(a_1^2 + a_2^2 + a_3^2\right)\left(b_1^2 + b_2^2 + b_3^2\right), \\ &= |\boldsymbol{a}|^2 |\boldsymbol{b}|^2. \end{aligned}$$

Hence,

$$\begin{aligned}|\boldsymbol{a} \times \boldsymbol{b}|^2 &= |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 - (\boldsymbol{a} \cdot \boldsymbol{b})^2, \\ &= |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 \left(1 - \cos^2\theta\right), \\ &= |\boldsymbol{a}|^2 |\boldsymbol{b}|^2 \sin^2\theta \end{aligned}$$

and

$$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta,$$

where $0 \le \theta \le \pi$, such that the relation $|\boldsymbol{a} \times \boldsymbol{b}| \ge 0$ is satisfied.

Step 2: Finding the direction of $\boldsymbol{a} \times \boldsymbol{b}$

Note that

$$
\begin{aligned}
\boldsymbol{a} \cdot (\boldsymbol{a} \times \boldsymbol{b}) &= a_1 (a_2 b_3 - a_3 b_2) + a_2 (a_3 b_1 - a_1 b_3) + a_3 (a_1 b_2 - a_2 b_1), \\
&= 0.
\end{aligned}
$$

Similarly, $\boldsymbol{b} \cdot (\boldsymbol{a} \times \boldsymbol{b}) = 0$. Hence, $\boldsymbol{a} \times \boldsymbol{b}$ is a vector perpendicular to both $\boldsymbol{a}$ and $\boldsymbol{b}$. It remains to find the sense of $\boldsymbol{a} \times \boldsymbol{b}$. Indeed, this is arbitrary and must be fixed. We fix it such that we have a right-handed system, and such that the following rule-of-thumb is satisfied (Figure 2.3).

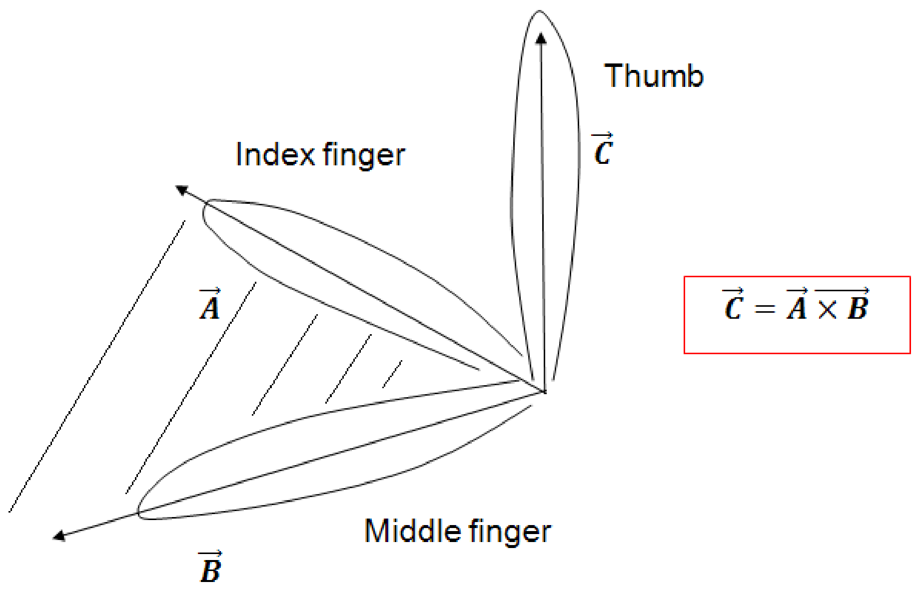


Figure 2.3: The right-hand rule.

Choosing a right-hand rule means that relations (2.6) are satisfied ($\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ form a 'right-handed' system). This also corresponds to putting a plus sign in front of the determinant in the original definition of the cross product.

> In summary, $\boldsymbol{a} \times \boldsymbol{b}$ is a vector of magnitude $|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$, that is normal to both $\boldsymbol{a}$ and $\boldsymbol{b}$, and whose sense is determined by the right-hand rule.

2.4.1 The cross product as an area

Consider a parallelogram, whose two adjacent sides are made up of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ (Fig. 2.4). The

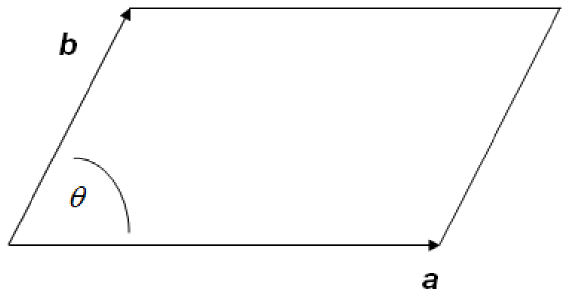


Figure 2.4: The cross product as an area

area of the parallelogram is

$$\begin{aligned} A &= (\text{base length}) \times (\text{perpendicular height})\,, \\ &= (\text{base length})\,|\boldsymbol{b}|\sin\theta, \\ &= |\boldsymbol{a}||\boldsymbol{b}|\sin\theta, \\ &= |\boldsymbol{a}\times\boldsymbol{b}|. \end{aligned}$$

2.5 The scalar triple product and volume

We can form a scalar from the three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ by combining the operations just defined:

$$\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})\,. \tag{2.7}$$

This is the so-called ‘scalar triple product’.

Theorem 2.2 *The scalar triple product $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})$ is identically equal to*

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Proof: By brute force,

$$\begin{aligned} & (a_1\hat{\boldsymbol{x}} + a_2\hat{\boldsymbol{y}} + a_3\hat{\boldsymbol{z}}) \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \\ = & \ (a_1\hat{\boldsymbol{x}} + a_2\hat{\boldsymbol{y}} + a_3\hat{\boldsymbol{z}})\cdot\left[(b_2c_3 - c_2b_3)\,\hat{\boldsymbol{x}} + (b_3c_1 - b_1c_3)\,\hat{\boldsymbol{y}} + (b_1c_2 - b_2c_1)\,\hat{\boldsymbol{z}}\right] \\ = & \ a_1\,(b_2c_3 - c_2b_3) + a_2\,(b_3c_1 - b_1c_3) + a_3\,(b_1c_2 - b_2c_1)\,, \end{aligned}$$

which is the determinant of the theorem.

Now consider a parallelepiped spanned by the vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ (Figure 2.5)

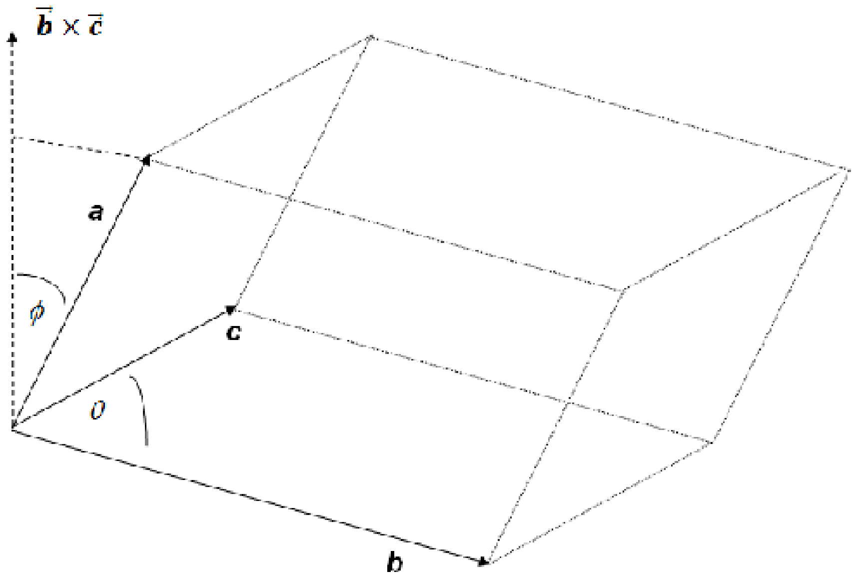


Figure 2.5: The scalar triple product as a volume

$$
\begin{aligned}
\text{Volume of parallelepiped} &= \text{(Perpendicular height)} \times \text{(Base area)} \\
&= (|\boldsymbol{a}|\cos\varphi) \times (|\boldsymbol{b}||\boldsymbol{c}|\sin\theta)\,, \\
&= (|\boldsymbol{a}|\cos\varphi)\,(|\boldsymbol{b}\times\boldsymbol{c}|)\,, \\
&= \boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})\,. \quad \blacksquare
\end{aligned}
$$

Corollary 2.2 *Three nonzero vectors* $\boldsymbol{a}$, $\boldsymbol{b}$, *and* $\boldsymbol{c}$ *are coplanar if and only if* $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = 0$.

Proof: The equation $\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = 0$ holds iff the volume of the parallelepiped spanned by the three vectors is zero iff the perpendicular height is zero, iff the three vectors are coplanar. $\blacksquare$

2.6 The vector triple product

Given three vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, we can form yet another vector,

$$\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c})\,. \tag{2.8}$$

The brackets are important because the cross product is not associative, e.g.

$$\hat{\boldsymbol{x}}\times(\hat{\boldsymbol{x}}\times\hat{\boldsymbol{y}}) = \hat{\boldsymbol{x}}\times\hat{\boldsymbol{z}} = -\hat{\boldsymbol{y}},$$

but

$$(\hat{\boldsymbol{x}} \times \hat{\boldsymbol{x}}) \times \hat{\boldsymbol{y}} = 0 \times \hat{\boldsymbol{y}} = 0.$$

Theorem 2.3 *The vector triple product satisfies*

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}\,(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}\,(\boldsymbol{a} \cdot \boldsymbol{b})\,, \tag{2.9}$$

a result that can be recalled by the mnemonic 'BAC minus CAB'.

Proof: Without loss of generality, we prove the result in a frame wherein the x- and y-axes of our frame lie in the plane generated by $\boldsymbol{b}$ and $\boldsymbol{c}$. In fact, we may take

$$\begin{aligned} \boldsymbol{c} &= \hat{\boldsymbol{x}} c_1, \\ \boldsymbol{b} &= b_1 \hat{\boldsymbol{x}} + b_2 \hat{\boldsymbol{y}}, \end{aligned}$$

and

$$\boldsymbol{a} = a_1 \hat{\boldsymbol{x}} + a_2 \hat{\boldsymbol{y}} + a_3 \hat{\boldsymbol{z}}.$$

The result then follows by a brute-force calculation of the LHS and the RHS of Eq. (2.9). ∎

Chapter 3

The geometry of lines and planes

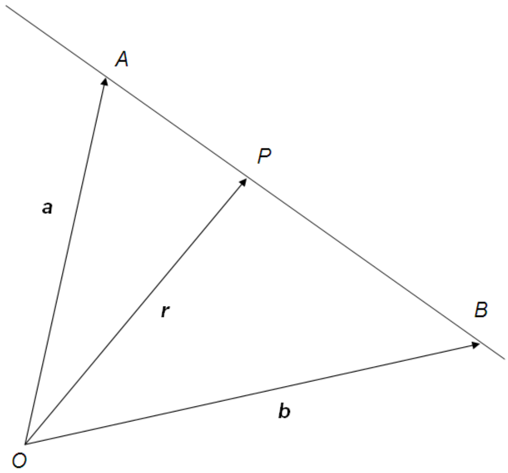
Overview

In this section we show how vector operations can be used to describe lines and planes in three-dimensional space. Some of this material will have been covered already in MATH 10340 but it is of vital importance to this module, so it is repeated here. The reason why these ideas are so important is that they carry over to general (smooth) curves and surfaces, which can be approximated to arbitrary precision by collections of line segments and planar surfaces.

3.1 The equation of a line

Problem: Find the equation of a straight line which passes through two given points A and B having position vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ w.r.t. an origin O.

Solution: Let $\boldsymbol{r}$ be the position vector of any point P on the line through A and B. From the figure,



$$\overrightarrow{OA} + \overrightarrow{AP} = \overrightarrow{OP} \implies \boldsymbol{a} + \overrightarrow{AP} = \boldsymbol{r}$$
$$\implies \overrightarrow{AP} = \boldsymbol{r} - \boldsymbol{a},$$

and

$$\overrightarrow{OA} + \overrightarrow{AB} = \overrightarrow{OB} \implies \boldsymbol{a} + \overrightarrow{AB} = \boldsymbol{b}$$
$$\implies \overrightarrow{AB} = \boldsymbol{b} - \boldsymbol{a}.$$

But $\overrightarrow{AP}$ and $\overrightarrow{AB}$ are colinear, hence

$$\overrightarrow{AP} = t\overrightarrow{AB} = t\left(\boldsymbol{b} - \boldsymbol{a}\right),$$

where t is some real number. Putting these equations together gives

$$\boldsymbol{r} = \boldsymbol{a} + t\left(\boldsymbol{b} - \boldsymbol{a}\right).$$

Thus, two vectors are sufficient to specify a line in space: a vector $\boldsymbol{r}_0 := \boldsymbol{a}$ whose tip lies on the line, and a vector $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{a}$ that lies along the line. We therefore write

$$\boldsymbol{r} = \boldsymbol{r}_0 + t\boldsymbol{e}.$$

A straight line is thus a **one-parameter curve**.

Now let's go over to the Cartesian form of the line:

$$\begin{aligned} x &= x_0 + te_x, \\ y &= y_0 + te_y, \\ z &= z_0 + te_z. \end{aligned}$$

Eliminating t between the equations (if possible) gives

$$\frac{x - x_0}{e_x} = \frac{y - y_0}{e_y} = \frac{z - z_0}{e_z}.$$

If the line lies entirely in the x-y plane, then $z = 0$ and the elimination is carried out on only the x- and y-variables:

$$\frac{x - x_0}{e_x} = \frac{y - y_0}{e_y} \implies y = y_0 + \frac{e_y}{e_x}\left(x - x_0\right),$$

which is the standard equation of the line in a plane with slope $m = e_y/e_x$.

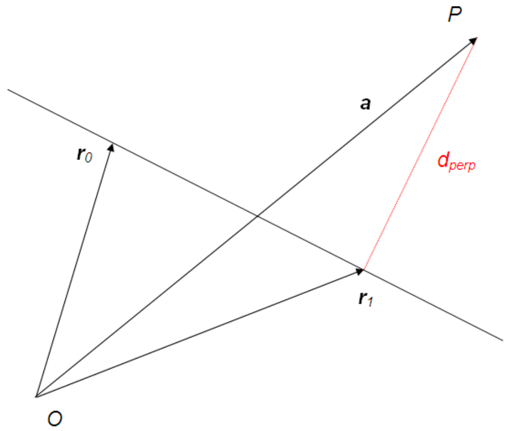
3.2 The perpendicular distance between a point and a line

Problem: Let $\boldsymbol{r}(t) = \boldsymbol{r}_0 + t\boldsymbol{e}$ be the equation of a straight line L, and let P be a point with position vector $\boldsymbol{a}$ (henceforth written as $P(\boldsymbol{a})$). Find the shortest distance between the line and the point.

Solution: The shortest distance between the point P and the line L is in fact the perpendicular distance. Suppose that a perpendicular dropped from P to L intersects L at position vector $\boldsymbol{r}_1$, such that

$$\boldsymbol{r}_1 = \boldsymbol{r}_0 + t_1\boldsymbol{e} \tag{3.1}$$

(Refer to the figure).



By construction of this perpendicular, the line $\boldsymbol{a} - \boldsymbol{r}_1$ is perpendicular to the line, or

$$\boldsymbol{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_1) = 0.$$

Hence,

$$\boldsymbol{e} \cdot \boldsymbol{a} = \boldsymbol{e} \cdot \boldsymbol{r}_1 = \boldsymbol{e} \cdot (\boldsymbol{r}_0 + t_1\boldsymbol{e}) = \boldsymbol{e} \cdot \boldsymbol{r}_0 + t_1\boldsymbol{e}^2.$$

Solving for t_1,

$$t_1 = \frac{\boldsymbol{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_0)}{|\boldsymbol{e}|^2}.$$

Substitute this expression into Equation (3.1):

$$\boldsymbol{r}_1 = \boldsymbol{r}_0 + \boldsymbol{e}\frac{\boldsymbol{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_0)}{|\boldsymbol{e}|^2} := \boldsymbol{r}_0 + \hat{e}\left[\hat{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_0)\right],$$

where $\hat{e} = \boldsymbol{e}/|\boldsymbol{e}|$ is a unit vector along the line L. The perpendicular distance is thus

$$\begin{aligned}
d_\perp &= |\boldsymbol{a} - \boldsymbol{r}_1|, \\
&= |(\boldsymbol{a} - \boldsymbol{r}_0) - [\hat{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_0)]\,\hat{e}|, \\
&= |(\boldsymbol{r}_0 - \boldsymbol{a}) + [\hat{e} \cdot (\boldsymbol{a} - \boldsymbol{r}_0)]\,\hat{e}|, \\
&= |(\boldsymbol{r}_0 - \boldsymbol{a}) - [\hat{e} \cdot (\boldsymbol{r}_0 - \boldsymbol{a})]\,\hat{e}|.
\end{aligned}$$

This is a valid final answer. However, a little more manipulation yields

$$d_\perp^2 = (\boldsymbol{r}_0 - \boldsymbol{a})^2 - |\hat{\boldsymbol{e}} \cdot (\boldsymbol{r}_0 - \boldsymbol{a})|^2 = (\boldsymbol{r}_0 - \boldsymbol{a})^2 \left(1 - \cos^2\theta\right)$$
$$= (\boldsymbol{r}_0 - \boldsymbol{a})^2 \sin^2\theta = |(\boldsymbol{r}_0 - \boldsymbol{a}) \times \hat{\boldsymbol{e}}|^2,$$

hence

$$d_\perp = |(\boldsymbol{r}_0 - \boldsymbol{a}) \times \hat{\boldsymbol{e}}|.$$

Now suppose the line lies in the x-y plane only. The equation of the line is $\alpha x + \beta y + \gamma = 0$, with slope $m = -\alpha/\beta$. But $e_y/e_x = m = m/1$, hence

$$\boldsymbol{e} = (e_x, e_y) = (1, m), \qquad \hat{\boldsymbol{e}} = \frac{(1,m)}{\sqrt{1+m^2}} = \frac{(1, -\alpha/\beta)}{\sqrt{1+\alpha^2/\beta^2}} = \frac{(\beta, -\alpha)}{\sqrt{\alpha^2+\beta^2}}$$

Thus,

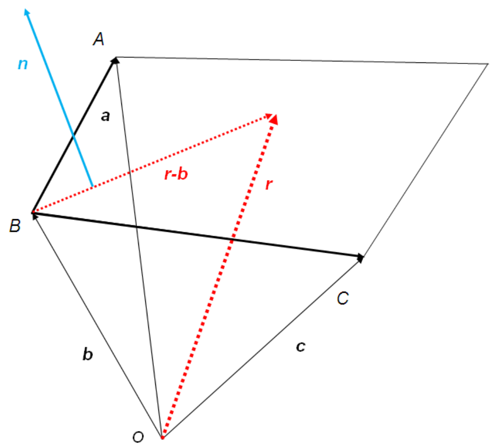
$$\begin{aligned}
d_\perp^2 &= \frac{1}{\alpha^2+\beta^2}\left|\underbrace{(x_0, y_0, 0) \times (\beta, -\alpha, 0)}_{=\boldsymbol{r}_0 \times \hat{\boldsymbol{e}}} - \underbrace{(a_1, a_2, 0) \times (\beta, -\alpha, 0)}_{=\boldsymbol{a} \times \hat{\boldsymbol{e}}}\right|^2, \\
&= \frac{1}{\alpha^2+\beta^2}|-(y_0\beta + x_0\alpha) + (a_2\beta + a_1\alpha)|^2, \\
&= \frac{1}{\alpha^2+\beta^2}|\gamma + (a_2\beta + a_1\alpha)|^2, \\
&= \frac{|a_1\alpha + a_2\beta + \gamma|^2}{\alpha^2+\beta^2}.
\end{aligned}$$

3.3 The equation of a plane

Problem: Find the equation of a plane which passes through three given points A, B, and C having position vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ w.r.t. an origin O.

Solution: By construction, the vectors

$$\boldsymbol{v}_1 := \overrightarrow{BA} = \boldsymbol{a} - \boldsymbol{b},$$
$$\boldsymbol{v}_2 := \overrightarrow{BC} = \boldsymbol{c} - \boldsymbol{b},$$



lie in the plane we call Π (see the accompanying figure). A normal to this plane is

$$\boldsymbol{n} = \boldsymbol{v}_2 \times \boldsymbol{v}_1 = (\boldsymbol{c} - \boldsymbol{b}) \times (\boldsymbol{a} - \boldsymbol{b}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{b} \times \boldsymbol{c} + \boldsymbol{c} \times \boldsymbol{a}.$$

The equation of a plane is subset of all vectors $\boldsymbol{r} = (x, y, z)$ in $\mathbb{R}^3$, such that the vector $\boldsymbol{r} - \boldsymbol{b}$ is perpendicular to $\boldsymbol{n}$:

$$\Pi = \left\{ \boldsymbol{r} \in \mathbb{R}^3 | \, (\boldsymbol{r} - \boldsymbol{b}) \cdot (\boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{b} \times \boldsymbol{c} + \boldsymbol{c} \times \boldsymbol{a}) = 0 \right\}.$$

Simplifying, the general vector $\boldsymbol{r}$ lies in the plane Π if and only if

$$\boldsymbol{r} \cdot \underbrace{(\boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{b} \times \boldsymbol{c} + \boldsymbol{c} \times \boldsymbol{a})}_{\text{normal vector}} = \underbrace{\boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a})}_{\text{a constant}}.$$

Note that the general equation of a plane in three dimensions is

$$\boxed{\Pi(\boldsymbol{n}, \boldsymbol{r}_0) = \left\{ \boldsymbol{r} \in \mathbb{R}^3 | \, (\boldsymbol{r} - \boldsymbol{r}_0) \cdot \boldsymbol{n} = 0 \right\}.} \tag{3.2}$$

The plane is thus parametrized by the normal vector $\boldsymbol{n}$ and a reference vector $\boldsymbol{r}_0$ whose tip lies in the plane (Fig. 3.1). If $n_z \neq 0$, we have the Cartesian expression

$$z = z_0 + (y_0 - y)\frac{n_y}{n_z} + (x_0 - x)\frac{n_x}{n_z}.$$

Thus, a point $z = z(x, y)$ on a surface is labelled by two parameters, x and y. A plane is therefore a **two-parameter** object, just as a line was a one-parameter curve.

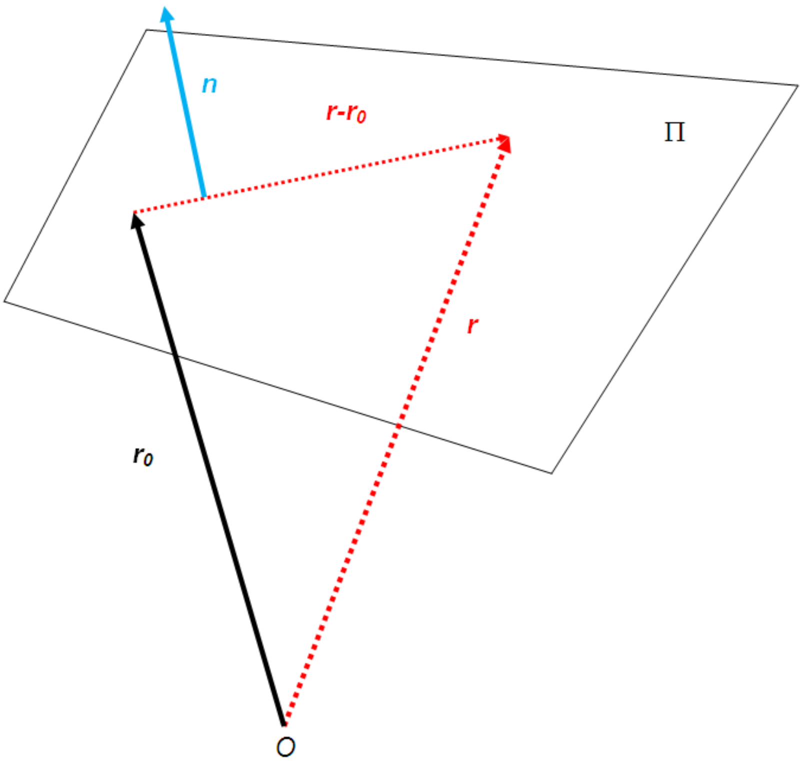


Figure 3.1: Figure for the general equation of a plane

3.4 Skew lines and intersecting lines in three dimensions

Skew lines are a very nice application of three-dimensional geometry. In two dimensions, two non-parallel lines definitely intersect. However, in three dimensions, they need not intersect: they can "go around" one another.

We first of all consider two intersecting lines. In other words, we **assume intersection** and compute explicitly the point of intersection in terms of the parameters of the two lines.

$$
\begin{aligned}
\boldsymbol{r}_L(t) &= \boldsymbol{r}_0 + t\boldsymbol{e}, \\
\boldsymbol{r}_M(u) &= \boldsymbol{s}_0 + u\boldsymbol{f},
\end{aligned}
$$

and we show that their point of intersection $\boldsymbol{r}_L(t_0) = \boldsymbol{r}_M(u_0)$ is given by the solution of the equation

$$
\left(\begin{array}{c} (\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{e} \\ (\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{f} \end{array}\right) = \left(\begin{array}{cc} -|\mathbf{e}|^2 & \boldsymbol{e}\cdot\boldsymbol{f} \\ -\boldsymbol{e}\cdot\boldsymbol{f} & |\boldsymbol{f}|^2 \end{array}\right)\left(\begin{array}{c} t_0 \\ u_0 \end{array}\right).
$$

By assumption, the point of intersection exists. Hence,

$$\boldsymbol{r}_0 + t_0\boldsymbol{e} = \boldsymbol{s}_0 + u_0\boldsymbol{f},$$

for parameter values t_0 and u_0. Re-arrange,

$$\boldsymbol{r}_0 - \boldsymbol{s}_0 = u_0\boldsymbol{f} - t_0\boldsymbol{e}.$$

Take the scalar product of this equation with $\boldsymbol{e}$:

$$(\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{e} = u_0\boldsymbol{f}\cdot\boldsymbol{e} - t_0|\boldsymbol{e}|^2;$$

do the same thing with vector $\boldsymbol{f}$:

$$(\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{f} = u_0|\boldsymbol{f}|^2 - t_0\boldsymbol{e}\cdot\boldsymbol{f}.$$

Gather up:

$$\begin{aligned}(\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{e} &= -t_0|\boldsymbol{e}|^2 + u_0\boldsymbol{f}\cdot\boldsymbol{e},\\ (\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{f} &= -t_0\boldsymbol{e}\cdot\boldsymbol{f} + u_0|\boldsymbol{f}|^2,\end{aligned}$$

which is the required result:

$$\begin{pmatrix}(\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{e}\\ (\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{f}\end{pmatrix} = \begin{pmatrix}-|\boldsymbol{e}|^2 & \boldsymbol{e}\cdot\boldsymbol{f}\\ -\boldsymbol{e}\cdot\boldsymbol{f} & |\boldsymbol{f}|^2\end{pmatrix}\begin{pmatrix}t_0\\ u_0\end{pmatrix}.$$

In the previous example, we were told that the point of intersection exists. It was then fairly straightforward to compute that point. We now look at the converse, and formulate **a general condition for the point of intersection to exist**. We start with the two lines

$$\begin{aligned}\boldsymbol{r}_L(t) &= \boldsymbol{r}_0 + t\boldsymbol{e},\\ \boldsymbol{r}_M(u) &= \boldsymbol{s}_0 + u\boldsymbol{f}.\end{aligned}$$

From the first part, a candidate point (t_0, u_0) for the intersection is the solution of the equation

$$\begin{pmatrix}(\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{e}\\ (\boldsymbol{r}_0 - \boldsymbol{s}_0)\cdot\boldsymbol{f}\end{pmatrix} = \begin{pmatrix}-|\boldsymbol{e}|^2 & \boldsymbol{e}\cdot\boldsymbol{f}\\ -\boldsymbol{e}\cdot\boldsymbol{f} & |\boldsymbol{f}|^2\end{pmatrix}\begin{pmatrix}t_0\\ u_0\end{pmatrix},$$

provided the solution exists. Now the determinant of this matrix is

$$-|\boldsymbol{e}|^2|\boldsymbol{f}|^2 + (\boldsymbol{e}\cdot\boldsymbol{f})^2 = -|\boldsymbol{e}\times\boldsymbol{f}|^2.$$

Thus, if $\boldsymbol{e}\times\boldsymbol{f}\neq 0$, the pair $(t_0, u_0)^T$, with

$$\begin{pmatrix} t_0 \\ u_0 \end{pmatrix} = \begin{pmatrix} -|\boldsymbol{e}|^2 & \boldsymbol{e}\cdot\boldsymbol{f} \\ -\boldsymbol{e}\cdot\boldsymbol{f} & |\boldsymbol{f}|^2 \end{pmatrix}^{-1} \begin{pmatrix} (\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot\boldsymbol{e} \\ (\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot\boldsymbol{f} \end{pmatrix}$$

is certainly a candidate solution. In plugging the solution of the matrix equation back into the intersection condition

$$\boldsymbol{r}_0 - \boldsymbol{s}_0 = u_0\boldsymbol{f} - t_0\boldsymbol{e},$$

we must be very careful: the only way to go from the solution of the matrix equation to the intersection condition is if $\boldsymbol{r}_0-\boldsymbol{s}_0$ lies entirely in the plane generated by $\boldsymbol{e}$ and $\boldsymbol{f}$. However, in general,

$$\boldsymbol{r}_0 - \boldsymbol{s}_0 = \alpha\boldsymbol{e} + \beta\boldsymbol{f} + \gamma\boldsymbol{e}\times\boldsymbol{f}.$$

Thus, we require $\gamma = 0$, or

$$(\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot(\boldsymbol{e}\times\boldsymbol{f}) = 0.$$

Therefore, a set of sufficient conditions for the lines to intersect is the following:

$$\boldsymbol{e}\times\boldsymbol{f} \neq 0 \quad \text{AND} \tag{3.3}$$

$$(\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot(\boldsymbol{e}\times\boldsymbol{f}) = 0. \tag{3.4}$$

Condition (3.4) states that $\boldsymbol{r}_0-\boldsymbol{s}_0$ lies entirely in the plane generated by $\boldsymbol{e}$ and $\boldsymbol{f}$ and thus, $(\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot\boldsymbol{e}$ and $(\boldsymbol{r}_0-\boldsymbol{s}_0)\cdot\boldsymbol{f}$ can not both be zero. In geometrical language, the condition is that $\boldsymbol{e}$ and $\boldsymbol{f}$ must be non-parallel AND the difference $\boldsymbol{r}_0-\boldsymbol{s}_0$ must lie entirely in the plane generated by $\boldsymbol{e}$ and $\boldsymbol{f}$. Lines that satisfy the first condition but not the second are called **skew lines**.

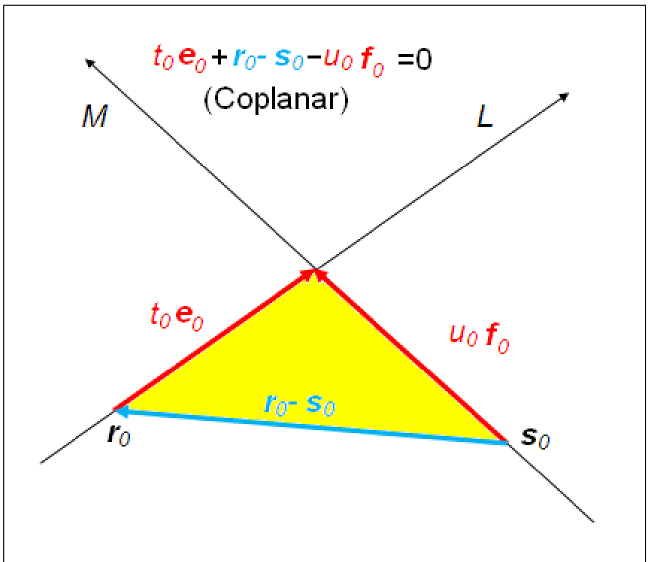


Figure 3.2: Vectors coplanar

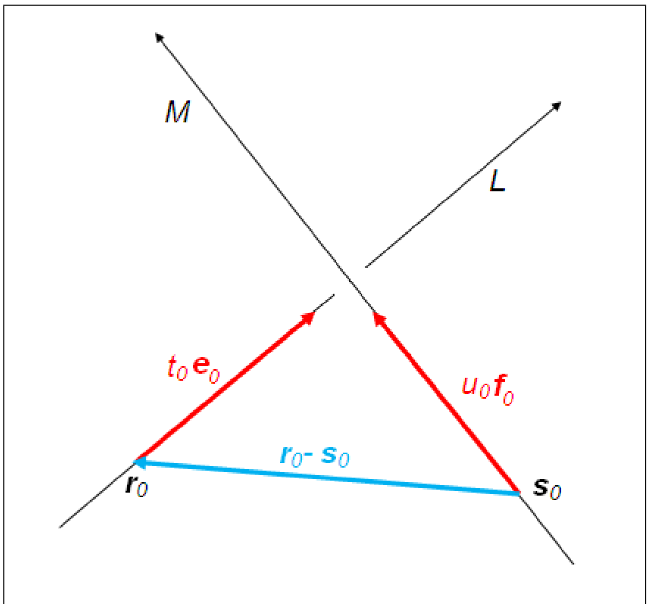


Figure 3.3: Skew lines: Vectors non-coplanar

Chapter 4

Ordinary derivatives of vectors

Overview

In many applications, we must consider a vector in $\mathbb{R}^3$ that varies continuously as a single parameter is varied. In particular, in mechanics, the position $\boldsymbol{x}$ of a particle is a function of time. Such a situation is called a curve: a curve γ is a map

$$\begin{aligned} \gamma : \mathbb{R} &\to \mathbb{R}^3, \\ t &\to \boldsymbol{x}_\gamma(t) = \left(x_\gamma(t), y_\gamma(t), z_\gamma(t)\right). \end{aligned}$$

Here $x_\gamma(\cdot)$, $y_\gamma(\cdot)$, and $z_\gamma(\cdot)$ are functions of time that give the Cartesian coordinates of the particle. Although not technically correct, in this section we drop the curve label γ and write $\boldsymbol{x}_\gamma(t) = \boldsymbol{x}(t) = (x(t), y(t), z(t))$. Such sloppiness even has a formal name: it is called *an abuse of notation*.

4.1 Definitions and properties

Definition 4.1 (Derivative of a curve) *Let $\boldsymbol{x}(t)$ be a curve parametrized by the parameter t. The derivative of the curve with respect to t is defined as*

$$\frac{d\boldsymbol{x}}{dt} = \lim_{\Delta t \to 0} \frac{\boldsymbol{x}(t + \Delta t) - \boldsymbol{x}(t)}{\Delta t},$$

provided the limit exists.

A similar definition holds for the higher derivatives. Since

$$\boldsymbol{x}(t) = (x(t), y(t), z(t)) = \hat{\boldsymbol{x}} x(t) + \hat{\boldsymbol{y}} y(t) + \hat{\boldsymbol{z}} z(t),$$

where $\hat{\boldsymbol{x}}$ etc. are constant vectors, this derivative can also be written as

$$\frac{d\boldsymbol{x}}{dt} = \hat{\boldsymbol{x}}\frac{dx}{dt} + \hat{\boldsymbol{y}}\frac{dy}{dt} + \hat{\boldsymbol{z}}\frac{dz}{dt}.$$

It should be clear that curves inherit all the properties of real-valued functions. In particular,

Theorem 4.1 *The following properties are satisfied, for arbitrary differentiable curves $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, and $\boldsymbol{C}(t)$:*

1.
$$\frac{d}{dt}\left[\boldsymbol{A}(t) + \boldsymbol{B}(t)\right] = \frac{d\boldsymbol{A}}{dt} + \frac{d\boldsymbol{B}}{dt},$$

2.
$$\frac{d}{dt}\left[\boldsymbol{A}(t) \cdot \boldsymbol{B}(t)\right] = \boldsymbol{A}(t) \cdot \frac{d\boldsymbol{B}}{dt} + \boldsymbol{B}(t) \cdot \frac{d\boldsymbol{A}}{dt},$$

3.
$$\frac{d}{dt}\left[\boldsymbol{A}(t) \times \boldsymbol{B}(t)\right] = \boldsymbol{A}(t) \times \frac{d\boldsymbol{B}}{dt} + \frac{d\boldsymbol{A}}{dt} \times \boldsymbol{B}(t),$$

   (note the order!)

4. *For a scalar function $f(t)$,*
$$\frac{d}{dt}\left[f(t)\boldsymbol{A}(t)\right] = f(t)\frac{d\boldsymbol{A}}{dt} + \boldsymbol{A}\frac{df}{dt},$$

5.
$$\frac{d}{dt}\left[\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})\right] = \boldsymbol{A} \cdot \left(\boldsymbol{B} \times \frac{d\boldsymbol{C}}{dt}\right) + \boldsymbol{A} \cdot \left(\frac{d\boldsymbol{B}}{dt} \times \boldsymbol{C}\right) + \frac{d\boldsymbol{A}}{dt} \cdot (\boldsymbol{B} \times \boldsymbol{C}),$$

6.
$$\frac{d}{dt}\left[\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C})\right] = \boldsymbol{A} \times \left(\boldsymbol{B} \times \frac{d\boldsymbol{C}}{dt}\right) + \boldsymbol{A} \times \left(\frac{d\boldsymbol{B}}{dt} \times \boldsymbol{C}\right) + \frac{d\boldsymbol{A}}{dt} \times (\boldsymbol{B} \times \boldsymbol{C}),$$

Here we move the derivative 'operator' sequentially through the product.

The proofs are straightforward because the vectors $\boldsymbol{A} = A_1\hat{\boldsymbol{x}} + A_2\hat{\boldsymbol{y}} + A_3\hat{\boldsymbol{y}} := (A_1, A_2, A_3)$ etc. inherit their differentiability properties from their components. For example,

$$\begin{aligned}
\frac{d}{dt}\left(\boldsymbol{A}(t)\cdot\boldsymbol{B}(t)\right) &= \frac{d}{dt}\sum_{i=1}^{3} A_i(t)B_i(t), \\
&= \sum_{i=1}^{3}\frac{d}{dt}\left[A_i(t)B_i(t)\right], \\
&= \sum_{i=1}^{3}\left[A_i(t)\frac{dB_i}{dt} + \frac{dA_i}{dt}B_i\right], \\
&= \boldsymbol{A}\cdot\frac{d\boldsymbol{B}}{dt} + \frac{d\boldsymbol{A}}{dt}\cdot\boldsymbol{B}. \quad \blacksquare
\end{aligned}$$

Theorem 4.2 *Let $\boldsymbol{x}(t)$ be a curve in $\mathbb{R}^3$. Then $d\boldsymbol{x}(t)/dt$ is everywhere tangent to the curve.*

Proof: Take a point $\boldsymbol{x}(t)$ on the curve and a neighbouring point $\boldsymbol{x}(t+\Delta t)$, also on the curve, where Δt is small. Form the difference

$$\frac{\boldsymbol{x}(t+\Delta t) - \boldsymbol{x}(t)}{\Delta t}.$$

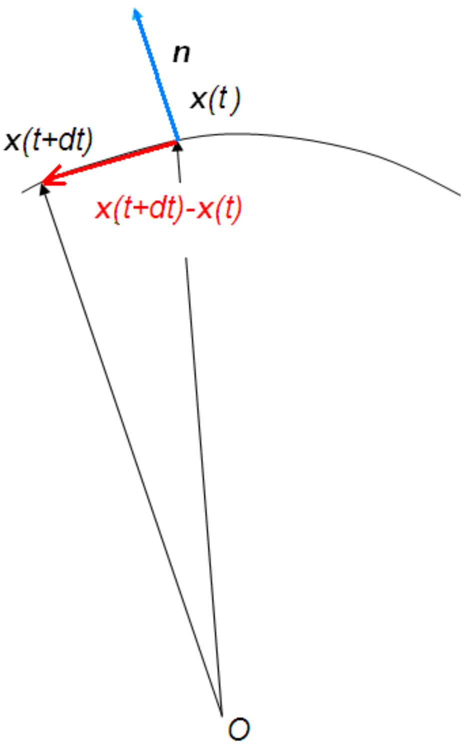
As the interval Δt is made smaller, the difference $\boldsymbol{x}(t+\Delta t) - \boldsymbol{x}(t)$ comes to lie parallel to the curve (as in the figure on the right), hence

$$\left(\frac{\boldsymbol{x}(t+\Delta t) - \boldsymbol{x}(t)}{\Delta t}\right)\cdot\boldsymbol{n} \to 0, \text{ as } \Delta t \to 0,$$

where $\boldsymbol{n}$ is a unit normal vector to the curve at the point $\boldsymbol{x}(t)$. In other words,

$$\frac{d\boldsymbol{x}}{dt}\cdot\boldsymbol{n} = 0,$$

and the vector $d\boldsymbol{x}/dt$ is therefore everywhere tangent to the curve $\boldsymbol{x}$. Thus, $d\boldsymbol{x}/dt$ is often called the **tangent vector** or the **velocity vector**. $\blacksquare$



4.2 Frenet–Serret frame

We introduce the notion of **arc length**. Consider a curve $\boldsymbol{x}(t)$. Along the curve, a small line element has length

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2.$$

Hence, the **arc length** along the curve, measured from a reference value $\boldsymbol{x}_0 = \boldsymbol{x}(t=0)$ is

$$s(t) = \int_0^{s(t)} \mathrm{d}s = \int_0^{s(t)} \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2} =$$
$$\int_0^t \sqrt{\left(\frac{dx}{dt'}\right)^2 + \left(\frac{dy}{dt'}\right)^2 + \left(\frac{dz}{dt'}\right)^2}\mathrm{d}t' = \int_0^t \left|\frac{d\boldsymbol{x}}{dt'}\right| \mathrm{d}t'.$$

This is a straightforward integration because $|d\boldsymbol{x}(t)/dt|$ is a simple function of time. Moreover,

$$\frac{ds}{dt} = \left|\frac{d\boldsymbol{x}}{dt}\right| \geq 0,$$

and the arclength is an **increasing function of time**. There is thus an inverse function $t = t(s)$, enabling a **reparametrization** of the curve according to arclength:

$$\tilde{\boldsymbol{x}}(s) = \boldsymbol{x}(t(s)).$$

Hence,

$$\frac{d\tilde{\boldsymbol{x}}}{ds} = \frac{d\boldsymbol{x}}{dt}\frac{dt}{ds} = \frac{d\boldsymbol{x}}{dt}\frac{1}{\left|\frac{d\boldsymbol{x}}{dt}\right|}, \qquad \text{(Chain Rule)}$$

and $d\tilde{\boldsymbol{x}}/ds$ is a unit vector tangent to the curve:

$$\boldsymbol{T} = \frac{d\tilde{\boldsymbol{x}}}{ds}.$$

Now

$$\boldsymbol{T}\cdot\boldsymbol{T} = 1,$$

hence

$$0 = \boldsymbol{T}\cdot\frac{d\boldsymbol{T}}{ds} + \frac{d\boldsymbol{T}}{ds}\cdot\boldsymbol{T} \implies \boldsymbol{T}\cdot\frac{d\boldsymbol{T}}{ds} = 0,$$

and $d\boldsymbol{T}/ds$ is perpendicular to the tangent vector $\boldsymbol{T}$. We therefore define a new unit vector $\boldsymbol{N} \propto d\boldsymbol{T}/ds$ that is normal to the tangent:

$$\frac{d\boldsymbol{T}}{ds} = \kappa(s)\boldsymbol{N},$$

and $\boldsymbol{N}$ is the **principal normal** to the curve and κ is the **curvature**.

Now our goal should be clear: we are deriving a triple of axes that move with the curve. $\boldsymbol{T}$ defines an axis everywhere parallel to the curve; $\boldsymbol{N}$ defines an axis that is everywhere perpendicular to the curve. In three dimensions, three axes are necessary: we therefore form a third unit vector

$$\boldsymbol{B} := \boldsymbol{T} \times \boldsymbol{N}.$$

The triple $(\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B})$ of axes along the curve $\tilde{\boldsymbol{x}}(s)$ parametrized by the arclength s is called the **Frenet–Serret frame**.

Note:

$$\begin{aligned}\frac{d\boldsymbol{B}}{ds} &= \boldsymbol{T} \times \frac{d\boldsymbol{N}}{ds} + \frac{d\boldsymbol{T}}{ds} \times \boldsymbol{N}, \\ &= \boldsymbol{T} \times \frac{d\boldsymbol{N}}{ds} + \kappa \boldsymbol{N} \times \boldsymbol{N}, \\ &= \boldsymbol{T} \times \frac{d\boldsymbol{N}}{ds}\end{aligned}$$

Hence

$$\boldsymbol{T} \cdot \left(\frac{d\boldsymbol{B}}{ds}\right) = \boldsymbol{T} \cdot \left(\boldsymbol{T} \times \frac{d\boldsymbol{N}}{ds}\right) = 0,$$

and $\boldsymbol{T}$ is perpendicular to $d\boldsymbol{B}/ds$. But $\boldsymbol{B} \cdot \boldsymbol{B} = 1$, hence

$$\boldsymbol{B} \cdot \left(\frac{d\boldsymbol{B}}{ds}\right) = 0.$$

Thus, $d\boldsymbol{B}/ds$ is perpendicular to $\boldsymbol{T}$ and $\boldsymbol{B}$, and must therefore lie along $\boldsymbol{N}$:

$$\frac{d\boldsymbol{B}}{ds} \propto \boldsymbol{N}.$$

We write

$$\frac{d\boldsymbol{B}}{ds} = -\tau(s)\boldsymbol{N},$$

where τ is the **torsion**. Finally, since $(\boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B})$ form a right-handed system (by construction), and since $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$, we may perform a cyclic permutation and obtain

$$\boldsymbol{N} = \boldsymbol{B} \times \boldsymbol{T}.$$

Operating on this with d/ds, we obtain

$$\begin{aligned}\frac{d\boldsymbol{N}}{ds} &= \boldsymbol{B} \times \frac{d\boldsymbol{T}}{ds} + \frac{d\boldsymbol{B}}{ds} \times \boldsymbol{T}, \\ &= \boldsymbol{B} \times (\kappa \boldsymbol{N}) - \tau(\boldsymbol{N} \times \boldsymbol{T}), \\ &= -\kappa \boldsymbol{T} + \tau \boldsymbol{B}.\end{aligned}$$

Let us assemble our results:

Theorem 4.3 *For the curve $\boldsymbol{x}(s)$ parametrized by arclength s, with*

- $\boldsymbol{T}$ *– unit tangent vector to curve ,*
- $\boldsymbol{N}$ *– unit vector normal to $\boldsymbol{T}$,*
- $\boldsymbol{B}$ *– a second unit vector normal to $\boldsymbol{T}$, $\boldsymbol{B} = \boldsymbol{T} \times \boldsymbol{N}$,*

the curvature κ and torsion τ are defined via the following equations:

$$\begin{aligned} \frac{d\boldsymbol{T}}{ds} &= \kappa(s)\boldsymbol{N}, \\ \frac{d\boldsymbol{B}}{ds} &= -\tau(s)\boldsymbol{N}, \\ \frac{d\boldsymbol{N}}{ds} &= \tau\boldsymbol{B} - \kappa\boldsymbol{T}. \end{aligned}$$

This framework is summarized graphically in Fig. 4.1.

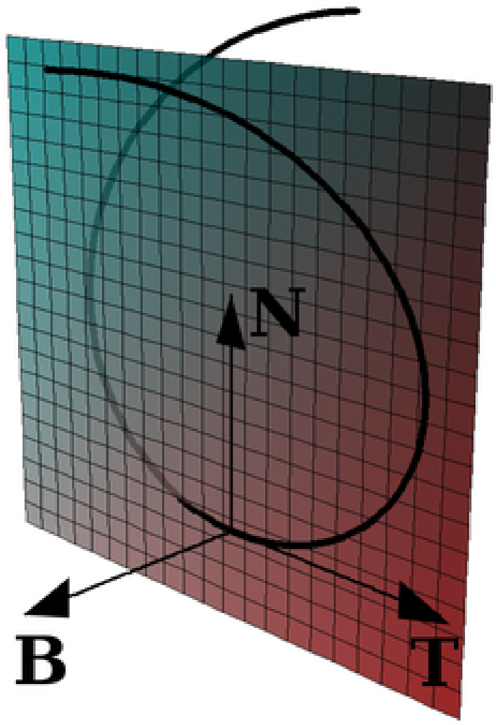


Figure 4.1: The Frenet–Serret frame along a curve. The plane shown the *osculating plane*, and this is the plane normal to the vector $\boldsymbol{B}$. From http://en.wikipedia.org/wiki/Frenet-Serret_formulas (3^{rd} August 2010)

Chapter 5

Frenet–Serret – Worked examples

Overview

We go through some worked examples that apply the Frenet–Serret formulas.

Example: Find the curvature and torsion of the curve $y = f(x)$ in two dimensions.

Solution: As we know from school, a curve in two dimensions can always be written in the form

$$y = f(x).$$

In other words,

$$\boldsymbol{x} = (x, f(x)). \tag{5.1}$$

Now here, x is simply a label, which indicates that the first variable in the bracket pair $(x, f(x))$ ranges over the whole real line (or some interval thereof). Thus, we can re-write the curve (5.1) as

$$\boldsymbol{x} = (t, f(t)).$$

The unit tangent vector is available immediately as

$$\boldsymbol{T} = \dot{\boldsymbol{x}}/|\dot{\boldsymbol{x}}|,$$

where

$$\dot{\boldsymbol{x}} := \frac{d\boldsymbol{x}}{dt} = (1, f'(t)), \qquad |\dot{\boldsymbol{x}}| = \sqrt{1 + f'(t)^2}.$$

Henceforth, to save chalk/ink/typing we write f instead of $f(t)$ &c, the functional dependence of

f on t being understood. Hence,

$$\boldsymbol{T} = \frac{(1, f')}{\sqrt{1 + f'^2}}. \tag{5.2}$$

To find the principal normal vector, we are going to have to differentiate Eq. (5.2):

$$\begin{aligned}
\frac{d\boldsymbol{T}}{dt} &= \left(\frac{d}{dt}(1 + f'^2)^{-1/2}, \frac{d}{dt}\frac{f'}{(1 + f'^2)^{1/2}}\right), \\
&= \left(-\frac{f'f''}{(1 + f'^2)^{3/2}}, \frac{(1 + f'^2)^{1/2}f'' - f'(1 + f'^2)^{-1/2}f'f''}{1 + f'^2}\right), \\
&= \left(-\frac{f'f''}{(1 + f'^2)^{3/2}}, \frac{(1 + f'^2)f'' - f'f'f''}{(1 + f'^2)^{3/2}}\right), \\
&= \left(-\frac{f'f''}{(1 + f'^2)^{3/2}}, \frac{f''}{(1 + f'^2)^{3/2}}\right), \\
&= \frac{f''}{(1 + f'^2)^{3/2}}\left(-f', 1\right).
\end{aligned}$$

Also,

$$\begin{aligned}
\frac{d\boldsymbol{T}}{ds} &= \frac{d\boldsymbol{T}}{dt}\Big/\left|\frac{d\boldsymbol{x}}{dt}\right|, \\
&= \frac{f''}{(1 + f'^2)^{3/2}}\frac{(-f', 1)}{\sqrt{1 + f'^2}}, \\
&= \kappa\boldsymbol{N}.
\end{aligned}$$

Actually, there was some ambiguity in our identification of the curvature in the derivation of the FS formulae – there are separate notions of **signed** and **unsigned** curvature. Here, we identify

$$\kappa_s := \frac{f''}{(1 + f'^2)^{3/2}}$$

as the **signed curvature** of the curve (since it can take either sign). Also, we identify

$$\boldsymbol{N}_s := \frac{(-f', 1)}{\sqrt{1 + f'^2}}$$

as the **signed principal normal vector**. The unsigned curvature is $\kappa_{us} := |\kappa_s|$, such that

$$\kappa_s\boldsymbol{N}_s = |\kappa_s|\mathrm{sign}(\kappa_s)\boldsymbol{N}_s = \kappa_{us}\mathrm{sign}(\kappa_s)\boldsymbol{N}_s.$$

This gives an unsigned normal vector,

$$\boldsymbol{N}_{us} = \mathrm{sign}(\kappa_s)\boldsymbol{N}_s,$$

such that

$$\frac{d\boldsymbol{T}}{ds} = \kappa_{us}\boldsymbol{N}_{us}, \qquad \kappa_{us} \geq 0.$$

To confuse matters more, there is further ambiguity in our choice of $(\boldsymbol{N}_s, \kappa_s)$: we can have either

$$\kappa_s = \pm\frac{f''}{(1+f'^2)^{3/2}}, \qquad \boldsymbol{N}_s = \pm\frac{(-f', 1)}{\sqrt{1+f'^2}}.$$

Choosing the positive sign means that the definition of (signed) curvature agrees with the ordinary notion of curvature, as being a quantity proportional to the second derivative of the curve.

Because $(\boldsymbol{T}, \boldsymbol{N})$ live in the x-y plane for all time, it follows that $\boldsymbol{B}$ is in the z-direction:

$$\boldsymbol{B} = \hat{\boldsymbol{z}}.$$

Now

$$\tau \propto \frac{d\boldsymbol{B}}{dt},$$

hence

$$\tau = 0.$$

This makes sense: the torsion is actually a measure of how much the curve "twists" out of the plane generated by $(\boldsymbol{T}, \boldsymbol{N})$. Since the curve lies in this plane for all time, it is impossible for it to "twist" out of this plane, hence $\tau = 0$:

Theorem 5.1 $\tau = 0$ *for a curve that lives entirely in the* x-y *plane.*

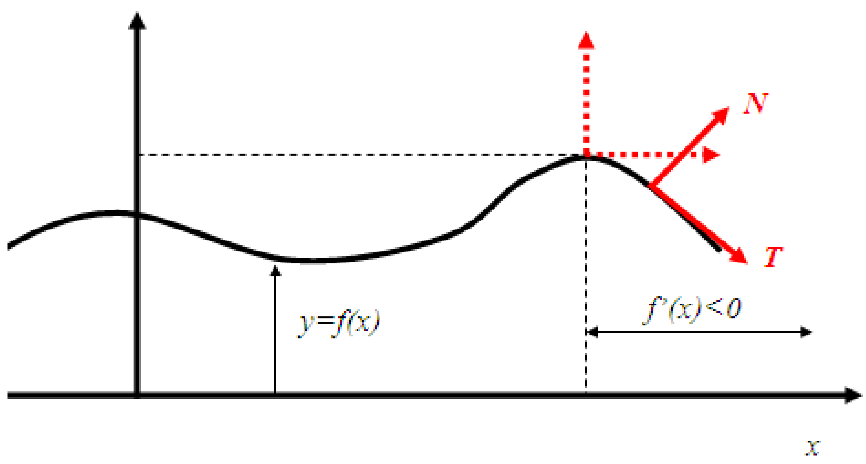


Figure 5.1: Normal and tangent vectors for a two-dimensional curve.

Example: Find the curvature and torsion of the a right-handed helix, given by the following parametric equations

$$
\begin{aligned}
x(t) &= r\cos t, && (5.3a)\\
y(t) &= r\sin t, && (5.3b)\\
z(t) &= vt, \qquad t\in[0,\infty), \qquad r,v>0. && (5.3c)
\end{aligned}
$$

Note that Equations (5.3) correspond to a right-handed helix. For, imagine a particle that follows the path (5.3). The particle does circular motion in the x-y plane and, at the same time, it moves up the z-axis. Moreover, if you coil your four fingers in the sense of the circular motion, your thumb points in the positive z-direction – the same direction of travel as the particle. Thus, the trajectory satisfies the right-hand rule.

Solution: We first compute the tangent vector:

$$
\begin{aligned}
\frac{d\boldsymbol{x}}{dt} &= \frac{d}{dt}\left(r\cos t, r\sin t, vt\right),\\
&= \left(-r\sin t, r\cos t, v\right),\\
\left|\frac{d\boldsymbol{x}}{dt}\right| &= \sqrt{r^2+v^2},\\
\frac{d\boldsymbol{x}}{dt}\Big/\left|\frac{d\boldsymbol{x}}{dt}\right| &= \frac{\left(-r\sin t, r\cos t, v\right)}{\sqrt{r^2+v^2}}.
\end{aligned}
$$

Hence

$$
\boldsymbol{T} = \frac{\left(-r\sin t, r\cos t, v\right)}{\sqrt{r^2+v^2}}.
$$

Also,

$$
\frac{d\boldsymbol{T}}{dt} = \frac{\left(-r\cos t, -r\sin t, v\right)}{\sqrt{r^2+v^2}},
$$

and

$$
\begin{aligned}
\frac{d\boldsymbol{T}}{dt}\Big/\left|\frac{d\boldsymbol{x}}{dt}\right| &= \frac{\left(-r\cos t, -r\sin t, 0\right)}{r^2+v^2},\\
&= \frac{r}{r^2+v^2}\left(-\cos t, -\sin t, 0\right),\\
&= \frac{d\boldsymbol{T}}{ds},\\
&= \kappa_s\boldsymbol{N}_s.
\end{aligned}
$$

Hence,

$$\boldsymbol{N}_s = \pm\left(-\cos t, -\sin t, 0\right), \qquad \kappa_s = \pm\frac{r}{r^2+v^2}.$$

Here, by taking the positive sign, the unsigned and signed curvatures agree:

$$\kappa_{us} = \kappa_s = \frac{r}{r^2+v^2} := \kappa;$$

hence, the signed and unsigned normal vectors also agree:

$$\boldsymbol{N}_{us} = \boldsymbol{N}_s = -\left(\cos t, \sin t, 0\right) := \boldsymbol{N}. \tag{5.4}$$

This means that $\boldsymbol{N}$ is an inward-pointing unit normal (the sign choice here is free and arbitrary choice). See Fig. 5.2 for more details. Here, the binormal points in the direction of motion (increasing z), which is a consequence of our having chosen the principal normal vector to be inward-pointing.

Next, we compute the torsion. We have,

$$\begin{aligned}
\boldsymbol{B} &= \boldsymbol{T}\times\boldsymbol{N},\\
&= \frac{1}{\sqrt{r^2+v^2}}\begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ -r\sin t & r\cos t & v \\ -\cos t & -\sin t & 0 \end{vmatrix},\\
&= \frac{1}{\sqrt{r^2+v^2}}\left(v\sin t, -v\cos t, r\right).
\end{aligned}$$

Also,

$$\begin{aligned}
\frac{d\boldsymbol{B}}{dt} &= \frac{1}{\sqrt{r^2+v^2}}\left(v\cos t, v\sin t, 0\right),\\
\frac{d\boldsymbol{B}}{dt}\Big/\left|\frac{d\boldsymbol{x}}{dt}\right| &= \frac{v}{r^2+v^2}\left(\cos t, \sin t, 0\right),\\
&= -\frac{v}{r^2+v^2}\left(-\cos t, -\sin t, 0\right),\\
\frac{d\boldsymbol{B}}{ds} &= -\tau\boldsymbol{N}.
\end{aligned}$$

Hence,

$$\tau = \frac{v}{r^2+v^2}.$$

Thus, the conventional minus sign in the formula $d\boldsymbol{B}/ds = -\tau N$ conspires to make the torsion of a right-handed helix positive. Note also that the torsion remains positive regardless of whether we take $(+\boldsymbol{N}, +\kappa)$ or $(-\boldsymbol{N}, -\kappa)$ to be the normal-curvature pair.

Note finally that for a helix,

$$\frac{\tau}{\kappa} = \frac{v}{r}.$$

We therefore have the following result:

Theorem 5.2 *The ratio of the torsion to the curvature is constant (t-independent) for a helix.*

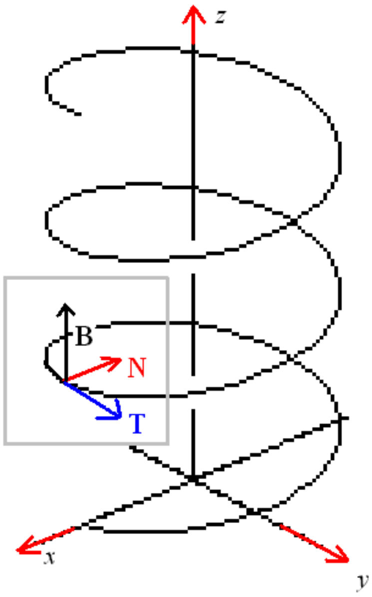


Figure 5.2: Frenet–Serret frame for a right-handed helix.

Example: Find the curvature and torsion of the general curve

$$x = t - t^3/3,\ \ y = t^2,\ \ z = t + t^3/3.$$

Solution: We have

$$\begin{aligned}
\boldsymbol{x} &= (t - t^3/3)\hat{\boldsymbol{x}} + t^2\hat{\boldsymbol{y}} + (t + t^3/3)\hat{\boldsymbol{z}}, \\
\frac{d\boldsymbol{x}}{dt} &= (1 - t^2)\hat{\boldsymbol{x}} + 2t\hat{\boldsymbol{y}} + (1 + t^2)\hat{\boldsymbol{z}}, \\
\left|\frac{d\boldsymbol{x}}{dt}\right|^2 &= (1 - t^2)^2 + 4t^2 + (1 + t^2)^2, \\
&= 2(1 + 2t^2 + t^4) = 2(1 + t^2)^2. \\
\frac{ds}{dt} &= \left|\frac{d\boldsymbol{x}}{dt}\right| = \sqrt{2}(1 + t^2). \\
\boldsymbol{T} &= \frac{d\boldsymbol{x}}{ds} = \frac{d\boldsymbol{x}}{dt}\bigg/\frac{ds}{dt} = \frac{(1 - t^2)\hat{\boldsymbol{x}} + 2t\hat{\boldsymbol{y}} + (1 + t^2)\hat{\boldsymbol{z}}}{\sqrt{2}(1 + t^2)}.
\end{aligned}$$

Next,

$$
\begin{aligned}
\frac{d\boldsymbol{T}}{dt} &= \frac{\{-2t\hat{\boldsymbol{x}}+2\hat{\boldsymbol{y}}+2t\hat{\boldsymbol{z}}\}(1+t^2)-2t\{(1-t^2)\hat{\boldsymbol{x}}+2t\hat{\boldsymbol{y}}+(1+t^2)\hat{\boldsymbol{z}}\}}{\sqrt{2}(1+t^2)^2},\\
&= \frac{-4t\hat{\boldsymbol{x}}+2(1-t^2)\hat{\boldsymbol{y}}}{\sqrt{2}(1+t^2)^2}.\\
\frac{d\boldsymbol{T}}{ds} &= \frac{d\boldsymbol{T}}{dt}\Big/\frac{ds}{dt},\\
&= \frac{-4t\hat{\boldsymbol{x}}+2(1-t^2)\hat{\boldsymbol{y}}}{2(1+t^2)^3} = \frac{-2t\hat{\boldsymbol{x}}+(1-t^2)\hat{\boldsymbol{y}}}{(1+t^2)^3}.
\end{aligned}
$$

Using the second FS equation,

$$
\left(\frac{d\boldsymbol{T}}{ds}\right)^2 = \kappa^2\boldsymbol{N}^2 = \kappa^2,
$$

$$
\kappa^2 = \left|\frac{d\boldsymbol{T}}{ds}\right|^2 = \frac{4t^2+(1-t^2)^2}{(1+t^2)^6} = \frac{(1+t^2)^2}{(1+t^2)^6} = \frac{1}{(1+t^2)^4}.
$$

Again, we take $\kappa_s = \kappa_{us} = 1/(1+t^2)^2$: because both curvatures are positive in this example, there is no need for the labels 's' and 'us'. Thus, we unambiguously use the formula

$$
\boldsymbol{N} = \frac{1}{\kappa}\frac{d\boldsymbol{T}}{ds}
$$

and compute

$$
\boldsymbol{N} = \frac{(1+t^2)^2\{-2t\hat{\boldsymbol{x}}+(1-t^2)\hat{\boldsymbol{y}}\}}{(1+t^2)^3} = \frac{-2t\hat{\boldsymbol{x}}+(1-t^2)\hat{\boldsymbol{y}}}{1+t^2}.
$$

Furthermore,

$$
\boldsymbol{B} = \boldsymbol{T}\times\boldsymbol{N},
$$

hence

$$
\begin{aligned}
\boldsymbol{B} &= \frac{1}{\sqrt{2}(1+t^2)^2}\begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ (1-t^2) & 2t & (1+t^2) \\ -2t & (1-t^2) & 0 \end{vmatrix}\\
&= \frac{-(1-t^2)(1+t^2)\hat{\boldsymbol{x}}-2t(1+t^2)\hat{\boldsymbol{y}}+(1+t^2)^2\hat{\boldsymbol{z}}}{\sqrt{2}(1+t^2)^2}\\
&= \frac{(t^2-1)\hat{\boldsymbol{x}}-2t\hat{\boldsymbol{y}}+(1+t^2)\hat{\boldsymbol{z}}}{\sqrt{2}(1+t^2)}.
\end{aligned}
$$

Next, we compute

$$
\begin{aligned}
\frac{d\boldsymbol{B}}{dt} &= \frac{(1+t^2)\{2t\hat{\boldsymbol{x}} - 2\hat{\boldsymbol{y}} + 2t\hat{\boldsymbol{z}}\} - 2t\{(t^2-1)\hat{\boldsymbol{x}} - 2t\hat{\boldsymbol{y}} + (1+t^2)\hat{\boldsymbol{z}}\}}{\sqrt{2}(1+t^2)^2} \\
&= \frac{\sqrt{2}\{2t\hat{\boldsymbol{x}} - (1-t^2)\hat{\boldsymbol{y}}\}}{(1+t^2)^2}. \\
\frac{d\boldsymbol{B}}{ds} &= \frac{d\boldsymbol{B}}{dt}\bigg/\frac{ds}{dt}, \\
&= \frac{2t\hat{\boldsymbol{x}} - (1-t^2)\hat{\boldsymbol{y}}}{(1+t^2)^3}, \\
&= -\frac{1}{(1+t^2)^2}\frac{-2t\hat{\boldsymbol{x}} + (1-t^2)\hat{\boldsymbol{y}}}{(1+t^2)}, \\
&= -\frac{\boldsymbol{N}}{(1+t^2)^2}.
\end{aligned}
$$

Now

$$\frac{d\boldsymbol{B}}{ds} = -\tau\boldsymbol{N}.$$

Therefore

$$\tau = \frac{1}{(1+t^2)^2}.$$

In conclusion, the space curve $x = t - t^3/3,\ y = t^2,\ z = t + t^3/3$ has curvature

$$\kappa = \frac{1}{(1+t^2)^2}$$

and torsion

$$\tau = \frac{1}{(1+t^2)^2}.$$

Chapter 6

Partial derivatives and fields

Overview

In this section we formulate the theory of scalar functions of several variables and learn how to differentiate such functions. The focus here in this chapter is on the elementary theory of partial derivatives. Although elementary partial differentiation is covered elsewhere (e.g. MATH 20060), it is repeated briefly here: it is important to get it right! Don't worry – we will move on to more involved topics in later chapters.

6.1 Partial derivatives

Definition 6.1 (Scalar field) *A function $\phi\left(x_1,x_2,\cdots,x_n\right)$ of n variables is a map from a subset of $\mathbb{R}^n$ to $\mathbb{R}$:*

$$
\begin{aligned}
\phi : (\Omega \subset \mathbb{R}^n) &\rightarrow \mathbb{R} \\
(x_1,x_2,\cdots,x_n) &\rightarrow \phi\left(x_1,x_2,\cdots,x_n\right).
\end{aligned}
$$

The function ϕ assigns to each point $(x_1,x_2,\cdots x_n)\in\Omega$ a real number (scalar), and is therefore called a **scalar field**.

Examples:

- The elevation above sea level at any point in Ireland is a function of latitude and longitude;
- The pressure of an ideal gas is a function of temperature and density (Boyle's Law);

- The quantity theory of money says that the GDP of an economy is a function of the velocity of money and the quantity of (broad) money in circulation.

In this section, we shall consider functions of two variables (x, y); the generalization to three or more variables is straightforward. When we are given such a function, it is natural to ask how the function varies as x changes, and as y changes. Equivalently, we want to know how the function changes as we move in the 'x-direction', and in the 'y'-direction'. Thus, we make small variations in the x-coordinate, keeping y fixed:

$$\phi\left(x+\delta x, y\right).$$

Then, we form the quotient

$$\frac{\phi\left(x+\delta x, y\right)-\phi\left(x, y\right)}{\delta x}.$$

Taking $\delta x \to 0$, we obtain the *partial derivative of ϕ w.r.t. x (keeping y fixed)*:

$$\frac{\partial \phi}{\partial x}\left(x, y\right)=\lim_{\delta x \to 0} \frac{\phi\left(x+\delta x, y\right)-\phi\left(x, y\right)}{\delta x}.$$

Similarly, we have a partial derivative with w.r.t. y keeping x fixed: First, we form the quotient

$$\frac{\phi\left(x, y+\delta y\right)-\phi\left(x, y\right)}{\delta y},$$

then we take the limit as $\delta y \to 0$:

$$\frac{\partial \phi}{\partial y}\left(x, y\right)=\lim_{\delta y \to 0} \frac{\phi\left(x, y+\delta y\right)-\phi\left(x, y\right)}{\delta y}.$$

Thus

To form a partial derivative in the x-direction, you treat y as a constant and do ordinary differentiation on the x-variable.

6.1.1 Worked examples

Example: Let $\phi\left(x, y\right)=x^2+y^2$. Compute the first partial derivatives of ϕ.

Solution: Let us hold y fixed and differentiate w.r.t. x:

$$\frac{\partial \phi}{\partial x}=\frac{\partial}{\partial x}\left(x^2+y^2\right)=\frac{\partial}{\partial x}\left(x^2+\mathrm{Const.}\right)=\frac{\partial}{\partial x}\left(x^2\right)=2x.$$

Now hold x fixed and differentiate w.r.t. y:

$$\frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y}\left(x^2 + y^2\right) = \frac{\partial}{\partial y}\left(\text{Const.} + y^2\right) = \frac{\partial}{\partial y}\left(y^2\right) = 2y.$$

> Example: Let $\phi(x, y) = x/y$. Compute the first partial derivatives of ϕ.

Let us hold y fixed and differentiate w.r.t. x:

$$\frac{\partial \phi}{\partial x} = \frac{\partial}{\partial x}\frac{x}{y} = \frac{\partial}{\partial x}\frac{x}{\text{Const.}} = \frac{1}{\text{Const.}} = \frac{1}{y}$$

Now hold x fixed and differentiate w.r.t. y:

$$\frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y}\frac{x}{y} = \frac{\partial}{\partial y}\frac{\text{Const.}}{y} = -\frac{\text{Const.}}{y^2} = -\frac{x}{y^2}$$

> Example: The function $\phi(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$ is a function of three variables. Compute all three first partial derivatives of ϕ.

Solution: Let us hold y **and** z fixed and differentiate w.r.t. x:

$$\begin{aligned}\frac{\partial \phi}{\partial x} = \frac{\partial}{\partial x}\frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{\partial}{\partial x}\frac{1}{\sqrt{x^2 + \text{Const.}}} &= \frac{\partial}{\partial x}\left(x^2 + \text{Const.}\right)^{-1/2} \\ = -\tfrac{1}{2}\left(x^2 + \text{Const.}\right)^{-3/2}(2x) &= -\frac{x}{\left(x^2 + \text{Const.}\right)^{3/2}} = -\frac{x}{\left(x^2 + y^2 + z^2\right)^{3/2}}\end{aligned}$$

Now hold x **and** z fixed and differentiate w.r.t. y:

$$\begin{aligned}\frac{\partial \phi}{\partial y} = \frac{\partial}{\partial y}\frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{\partial}{\partial x}\frac{1}{\sqrt{\text{Const.} + y^2}} &= \frac{\partial}{\partial x}\left(\text{Const.} + y^2\right)^{-1/2} \\ = -\tfrac{1}{2}\left(\text{Const.} + y^2\right)^{-3/2}(2y) &= -\frac{y}{\left(\text{Const.} + y^2\right)^{3/2}} = -\frac{y}{\left(x^2 + y^2 + z^2\right)^{3/2}}\end{aligned}$$

Lastly, we hold x and y fixed and differentiate w.r.t. z:

$$\begin{aligned}\frac{\partial \phi}{\partial z} = \frac{\partial}{\partial z}\frac{1}{\sqrt{x^2 + y^2 + z^2}} = \frac{\partial}{\partial z}\frac{1}{\sqrt{\text{Const.} + z^2}} &= \frac{\partial}{\partial z}\left(\text{Const.} + z^2\right)^{-1/2} \\ = -\tfrac{1}{2}\left(\text{Const.} + z^2\right)^{-3/2}(2z) &= -\frac{z}{\left(\text{Const.} + z^2\right)^{3/2}} = -\frac{z}{\left(x^2 + y^2 + z^2\right)^{3/2}}\end{aligned}$$

Pedantic notation

1. When the function ϕ is in fact a function of a single variable only ($\phi = \phi\,(x)$, say) there is no difference between $\partial/\partial x$ and d/dx. In that case, $\partial\phi/\partial x = d\phi/dx = \phi'\,(x)$.

2. To save chalk, we will sometimes write $\partial\phi/\partial x$ as $\partial_x \phi$ or even ϕ_x. A similar notation holds for partial derivatives w.r.t. y and z.

Chapter 7

Introduction to Maclaurin series

Overview

This is an informal presentation of how to write a function $f(x)$ as a power series. A more rigorous, formal treatment, follows later. The idea is that the informal treatment will help you to understand the whole approach. Only with this informal understanding (as well as fluency in the mathematical techniques involved) can the more formal understanding be built up.

7.1 The basic idea of the Taylor and Maclaurin series

Suppose we have a differentiable function $f(x)$. We can try to write it as a **power series centred at zero** as follows:

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + \cdots \tag{7.1}$$

The question is – what are the values of c_n? As it turns out, we can determine these in a **systematic** way by repeated term-by-term differentiation of the power series. We have,

- $f(0) = c_0$ – got by setting $x = 0$ in Equation (7.1)
- $f'(x) = c_1 + 2c_2 x + 3c_3 x^2 + 4c_4 x^3 + \cdots$, hence

$$c_1 = f'(0).$$

- $f''(x) = 2c_2 + 3 \cdot 2c_3 x + 4 \cdot 3c_4 x^2 + \cdots$, hence $f''(0) = 2c_2$, hence

$$c_2 = \tfrac{1}{2} f''(0).$$

- $f'''(x) = 3 \cdot 2c_3 + 4 \cdot 3 \cdot 2c_4 x + \cdots$, hence $f'''(0) = 3 \cdot 2c_3$, hence
$$c_3 = \tfrac{1}{3\cdot 2} f'''(0).$$
- $f^{(4)}(x) = 4 \cdot 3 \cdot 2c_4 + \cdots$, hence $f^{(4)}(0) = 4 \cdot 3 \cdot 2c_4$, hence
$$c_4 = \tfrac{1}{4\cdot 3\cdot 2} f^{(4)}(0).$$

Guess the pattern:

$$\boxed{c_n = \frac{1}{n!} f^{(n)}(0), \qquad n = 0, 1, 2, \cdots} \tag{7.2}$$

Going back to Equation (7.1), this gives

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} c_n x^n, \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \end{aligned} \tag{7.3}$$

This expression is called the **Maclaurin series** of $f(x)$.

A Macalurin series is an example of a power series centred at zero. We can guess what the corresponding series should look like if the centre is at $a \neq 0$:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n. \tag{7.4}$$

This is the **Taylor series** of $f(x)$ centred at a.

Watch movie: https://www.youtube.com/watch?v=IL93Sh8LpxM

7.2 Radius of convergence of a generic power series

Here, we look at the radius of convergence of a generic power series, without reference to any particular Maclaurin series. However, to fix ideas, we will look at a specific example:

$$\sum_{n=0}^{\infty} \frac{(-1)^n n}{4^n} (x+3)^n := \sum_{n=0}^{\infty} a_n,$$

where the **general term** is a_n, with

$$a_n = \frac{(-1)^n n}{4^n} (x+3)^n$$

We apply the **ratio test**, which says to form the ratio

$$L = \lim_{n=\infty} \left| \frac{a_{n+1}}{a_n} \right|,$$

Then,

- The series converges if $L < 1$,
- The series diverges if $L > 1$,
- The test is inconclusive if $L = 1$ (another test should then be tried)

Here, L is related to the **radius of convergence** R of the power series, as we will see in the specific example. A proof of the ratio test is provided in Appendix A.

We now apply the ratio test, identifying

$$a_n = \frac{(-1)^n n}{4^n} (x+3)^n, \qquad a_{n+1} = \frac{(-1)^{n+1}(n+1)}{4^{n+1}} (x+3)^{n+1}$$

Hence,

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{\frac{(-1)^{n+1}(n+1)}{4^{n+1}} (x+3)^{n+1}}{\frac{(-1)^n n}{4^n} (x+3)^n} \right|, \\ &\overset{\text{algebra!}}{=} \lim_{n\to\infty} \tfrac{1}{4} \left| \frac{n+1}{n} \right| |x+3|, \\ &= \tfrac{1}{4}|x+3|. \end{aligned}$$

We now require $L < 1$ for convergence, hence

$$\tfrac{1}{4}|x+3| < 1,$$

hence

$$|x+3| < \mathbf{4}$$

Thus, $\mathbf{R} = \mathbf{4}$ is the **radius of convergence** because the series converges for all values of x in an interval (radius) of width 4 centred at 3, i.e. for $|x+3| < 4$.

Remark: The radius of convergence can also be identified as follows:

$$
\begin{aligned}
L &= \lim_{n\to\infty} \left| \frac{\frac{(-1)^{n+1}(n+1)}{4^{n+1}}(x+3)^{n+1}}{\frac{(-1)^n n}{4^n}(x+3)^n} \right|, \\
&\overset{\text{algebra!}}{=} \lim_{n\to\infty} \tfrac{1}{4} \left| \frac{n+1}{n} \right| |x+3|, \\
&:= |x+3|(1/R),
\end{aligned}
$$

hence

$$
R = \frac{1}{\lim_{n\to\infty} \frac{1}{4}\left(\frac{n+1}{n}\right)} = 4.
$$

Watch movie: https://www.youtube.com/watch?v=r8bK9j4vKco

7.3 The Maclaurin series of e^x

Here, let $f(x) = \mathrm{e}^x$ and recall Equation (7.3), viz. $f(x) = \sum_{n=0}^{\infty}(1/n!)f^{(n)}x^n$. We have

- $f(x) = \mathrm{e}^x$, hence $f(0) = 1$,
- $f'(x) = \mathrm{e}^x$, hence $f'(0) = 1$,
- $f''(x) = \mathrm{e}^x$ (again!), hence $f''(0) = 1$.

Indeed,

$$
f^{(n)}(x) = \mathrm{e}^x
$$

for all $n = 0, 1, 2, \cdots$, hence

$$
f^{(n)}(0) = \mathrm{e}^0 = 1, \qquad n = 0, 1, 2, \cdots .
$$

Plug into Equation (**??**)

$$
f(x) = \mathrm{e}^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.
$$

We can use the ratio test to check the radius of convergence:

$$
\begin{aligned}
L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|, \qquad a_n = \frac{1}{n!}x^n, \\
&= \lim_{n\to\infty} \left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right|, \\
&= |x| \lim_{n\to\infty} \frac{1}{n}.
\end{aligned}
$$

Hence, $L = 0$ and the series converges for all real values of x, in other words, the radius of convergence is infinite. Done otherwise,

$$\begin{aligned} L &= \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|, \qquad a_n = \frac{1}{n!}x^n, \\ &= |x| \lim_{n\to\infty}\frac{1}{n}, \\ &= |x|/R, \end{aligned}$$

hence

$$\begin{aligned} R &= \frac{1}{\lim_{n\to\infty}\frac{1}{n}}, \\ &= \lim_{n\to\infty}(n), \\ &= \infty. \end{aligned}$$

We can write down the Taylor series (i.e. arbitrary centre) of $f(x) = \mathrm{e}^x$, setting the centre at a. As before, we have

$$f^{(n)}(x) = \mathrm{e}^x \implies f^{(n)}(a) = \mathrm{e}^a,$$

for all $n = 0, 1, 2, \cdots$. We use the expression (7.4) for the Taylor series of a generic function $f(x)$:

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n, \\ &= \sum_{n=0}^{\infty}\frac{\mathrm{e}^a}{n!}(x-a)^n, \\ &= \mathrm{e}^a\sum_{n=0}^{\infty}\frac{1}{n!}(x-a)^n, \end{aligned}$$

where e^a comes outside the sum because it is a constant.

Watch movie: https://www.youtube.com/watch?v=Rmt2vXW3s2U

7.4 Maclaurin series of $\sin(x)$

We take

$$f(x) = \sin(x) = \sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n.$$

We start differentiating:

- $f(x) = \sin(x)$, hence $f(0) = 0$,

- $f'(x) = \cos(x)$, hence $f'(0) = 1$,
- $f''(x) = -\sin(x)$, hence $f''(0) = 0$,
- $f'''(x) = -\cos(x)$, hence $f'''(0) = -1$,
- $f^{(4)}(x) = \sin(x)$, hence $f^{(4)}(0) = 0$.

Thus,

$$f(x) = f^{(4)}(x),$$

so the above pattern is cyclic and will repeat after every four derivatives. Thus, start with

$$\begin{aligned} f(x) = f(0) + f'(0)x + \tfrac{1}{2}f''(0)x^2 + \tfrac{1}{3!}f'''(0)x^3 + \tfrac{1}{4!}f^{(4)}(0)x^4 \\ + \tfrac{1}{5!}f^{(5)}(0)x^5 + \tfrac{1}{6!}f^{(6)}(0)x^6 + \tfrac{1}{7!}f^{(7)}(0)x^7 + \cdots, \end{aligned}$$

and fill in:

$$f(x) = 0 + 1 \cdot x + 0 - \tfrac{1}{3!}x^3 + 0 + 1 \cdot \tfrac{1}{5!}x^5 + 0 - 1 \cdot \tfrac{1}{7!}x^7 + \cdots.$$

Guess the pattern:

$$f(x) = \sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}.$$

We can again use the ratio test to check the radius of convergence:

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|, \qquad a_n = \frac{(-1)^n}{(2n+1)!}x^{2n+1}, \\ &= \lim_{n\to\infty} \left| \frac{(-1)^{n+1}x^{2(n+1)+1}}{[2(n+1)+1]!} \frac{(2n+1)!}{(-1)^n x^{2n+1}} \right|, \\ &= |x|^2 \lim_{n\to\infty} \frac{1}{(2n+3)(2n+2)}, \\ &= 0. \end{aligned}$$

Hence, $L = 0$ and the series converges for all real values of x, in other words, the radius of convergence is infinite, $R = \infty$.

As extra practice, try to guess / derive the similar expression for $\cos(x)$:

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} = 1 - \tfrac{1}{2!}x^2 + \tfrac{1}{4!}x^4 - \tfrac{1}{6!}x^6 + \cdots. \tag{7.5}$$

Watch movie: https://www.youtube.com/watch?v=909IKnRCUiM

7.5 Maclaurin series of $1/(1-x)$

We take

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

We start differentiating:

- $f(x) = (1-x)^{-1}$, hence $f(0) = 1$,
- $f'(x) = (-1)(-1)(1-x)^{-2}$, hence $f'(0) = 1$,
- $f''(x) = (-1)(-2)(1-x)^{-3}$, hence $f''(0) = 2$,
- $f'''(x) = (-1)(-3)(-2)(1-x)^{-4}$, hence $f'''(0) = 3 \times 2 \times 1$,

Guessing the pattern, we have $f^{(n)}(0) = n!$, hence

$$f(x) = \sum_{n=0}^{\infty} x^n.$$

We can again use the ratio test to check the radius of convergence:

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|, \qquad a_n = x^n, \\ &= \lim_{n\to\infty} |x|, \\ &= |x|. \end{aligned}$$

Hence, $L = |x|$ and the series converges for $L < 1$, i.e. for $|x| < 1$. In other words, the radius of convergence is $R = 1$, with $|x| < R = 1$ for convergence.

Of course, this result can also be derived (for $|x| < 1$) by noting that $\sum_{n=0}^{\infty} x^n$ is a geometric progression.

7.6 Maclaurin series – one more example

In this example a clever shortcut is used to avoid computing too many complicated derivatives. You should try to use a similar shortcut whenever you suspect that your answer might involve a geometric progression, which has a known series expansion.

Take

$$f(x) = \frac{x^3}{x+8}.$$

We know

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n.$$

Here, **please don't use**

$$f^{(n)}(x) = \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{x^3}{x+8}\right).$$

Instead, notice that

$$\frac{1}{1-r} = \sum_{n=0}^{\infty} r^n, \qquad |r| < 1$$

is the geometric progression. So,

$$\begin{aligned}
\frac{1}{x+8} &= \frac{1}{8\left(1+\frac{x}{8}\right)}, \\
&= \tfrac{1}{8}\frac{1}{1-\left(-\frac{x}{8}\right)}, \\
&\implies r = -\tfrac{1}{8}x.
\end{aligned}$$

Thus,

$$\frac{1}{x+8} = \tfrac{1}{8}\sum_{n=0}^{\infty} \tfrac{1}{8^n}(-1)^n x^n,$$

provided $|r| < 1$, i.e. $|x/8| < 1$, hence $|x| < 8$. The **radius of convergence** is therefore $R = 8$, and the series converges for all $|x| < 8$. Note that correspondingly, there is an interval of convergence,

$$-8 < x < 8.$$

Putting x^3 together now with $1/(x+8)$ we have

$$\begin{aligned}
\frac{x^3}{x+8} &= \tfrac{1}{8}x^3 \sum_{n=0}^{\infty} \tfrac{1}{8^n}(-1)^n x^n, \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n x^{n+3}}{8^{n+1}}.
\end{aligned}$$

Watch movie: https://www.youtube.com/watch?v=rAIoScFPX54

7.7 Maclaurin series – one more example involving the radius of convergence

In this section we look for the radius of convergence for the Maclaurin series of

$$f(x) = \cos(x^2).$$

Here, **please don't use**

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

directly. Instead, let's start with the well-known formula

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

(i.e. Equation (7.5)) and let $x \to x^2$ everywhere, to obtain

$$\begin{aligned} \cos(x^2) &= \sum_{n=0}^{\infty} \frac{(-1)^n (x^2)^{2n}}{(2n)!}, \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{4n}}{(2n)!}. \end{aligned}$$

Thereby, identify

$$a_n = \frac{(-1)^n x^{4n}}{(2n)!}$$

and compute

$$\begin{aligned} L &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|, \\ &= \lim_{n\to\infty} \left| \frac{(-1)^{n+1} x^{4(n+1)}}{[2(n+1)]!} \frac{(2n)!}{(-1)^n x^{4n}} \right|, \\ &= \lim_{n\to\infty} \left| \frac{-x^4}{(2n+2)(2n+1)} \right|, \\ &= 0. \end{aligned}$$

Thus, $L = 0$ (and therefore, $L < 1$) for all real values of x and hence, the series converges for all $|x| < \infty$ and therefore, the radius of convergence is again infinite, i.e. $R = \infty$.

Watch movie: https://www.youtube.com/watch?v=Iv9Kqk7rxJw

Chapter 8

Taylor's theorem in one and several dimensions

Overview

We now introduce Taylor series in a more rigorous and precise way. The discussion starts with Taylor polynomials, which are a way of approximating a smooth function by a polynomial. Under certain conditions, this approximation becomes exact in the limit when the degree of the polynomial tends to infinity. This is Taylor's theorem. We further use this theory to extend Taylor's theorem to several dimensions, introducing along the way the notion of the differential of a function of several variables.

8.1 Taylor's theorem in one dimension

Let

$$\begin{aligned} f : [\alpha, \beta] &\rightarrow \mathbb{R}, \\ x &\mapsto f(x) \end{aligned}$$

Suppose that $f(x)$ is N-times differentiable on (α, β) N continuous derivatives. We start with Taylor's theorem in one dimension for $f(x)$. We work with the so-called **integral form of the remainder**:

Theorem 8.1 *Suppose that a and x are points in the interval (α, β). Then,*

$$f(x) = f(a)+f'(a)(x-a)+\tfrac{1}{2}f''(a)(x-a)^2+\cdots+\frac{1}{(k-1)!}f^{(k-1)}(a)(x-a)^{k-1}+R_k(x), \qquad 1 \le k \le N \tag{8.1}$$

where the remainder $R_k(x)$ can be written in integral form as

$$R_k(x) = \frac{1}{(k-1)!}\int_a^x f^{(k)}(t)\,(x-t)^{k-1}\,\mathrm{d}t. \tag{8.2}$$

Proof: We show this by mathematical induction. We first of all check that the proposition is true for $k = 1$. By the fundamental theorem of calculus we have

$$\int_a^x f'(t)\,\mathrm{d}t = f(x) - f(a),$$

hence

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t,$$

which agrees with the statement of the theorem at $k = 1$.

Assume now that the statement is true for a general integer k, with $1 \le k < N$:

$$f(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{(k-1)!}f^{(k-1)}(a)(x-a)^{k-1} + R_k(x), \tag{8.3}$$

with

$$R_k(x) = \frac{1}{(k-1)!}\int_a^x f^{(k)}(t)\,(x-t)^{k-1}\,\mathrm{d}t.$$

Apply integration by parts to the remainder:

$$\begin{aligned}
R_k(x) &= \frac{1}{(k-1)!}\int_a^x \underbrace{f^{(k)}(t)}_{=u}\,\underbrace{(x-t)^{k-1}\,\mathrm{d}t}_{=\mathrm{d}v}, \\
&= \frac{1}{(k-1)!}\left[-\frac{1}{k}f^{(k)}(t)(x-t)^k\Big|_a^x + \frac{1}{k}\int_a^x f^{(k+1)}(t)(x-t)^k\,\mathrm{d}t\right], \\
&= \frac{1}{k!}f^{(k)}(a)(x-a)^k + \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k\,\mathrm{d}t.
\end{aligned}$$

Thus, Equation (8.3) becomes

$$\begin{aligned}
f(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{(k-1)!}f^{(k-1)}(a)(x-a)^{k-1} \\
+ \frac{1}{k!}f^{(k)}(a)(x-a)^k + \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k\,\mathrm{d}t,
\end{aligned} \tag{8.4}$$

which can be written as

$$f(x) = f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \cdots + \frac{1}{(k-1)!}f^{(k-1)}(a)(x-a)^{k-1}$$
$$+ \frac{1}{k!}f^{(k)}(a)(x-a)^k + R_{k+1}(x), \qquad R_{k+1}(x) = \frac{1}{k!}\int_a^x f^{(k+1)}(t)(x-t)^k \,\mathrm{d}t, \quad (8.5)$$

and thus, the proposition is true for $k+1$. We therefore have:

- The proposition is true for $k = 1$,
- If the proposition is true for k then it is true for $k+1$,

and hence, by the principle of mathematical induction, the proposition is true for all $k = 1, 2, \cdots, N$ and thus, Taylor's remainder theorem is proved. ∎

The only thing stopping us extending the proof to all integers $k = 1, 2, \cdots$ is the assumption that $f(x)$ has only N continuous derivatives. However, from now on we will assume that the function $f(x)$ has as arbitrarily many continuous derivatives.

8.2 Applications of Taylor's Theorem

Why is a Taylor series useful? It turns out that the truncations of a Taylor series for $f(x)$ are a pretty good approximation for $f(x)$. Thus, we can write

$$f(x) \approx \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(x-a)^k := T_{f,a}(x)$$

where the truncation on the right-hand side is called a **Taylor polynomial**:

Definition 8.1 (Taylor polynomials) *The polynomial*

$$\begin{aligned} T_{f,a}(x) &= \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(x-a)^k, \\ &= f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2 + \cdots + \tfrac{1}{n!}f^{(n)}(a)(x-a)^n \end{aligned}$$

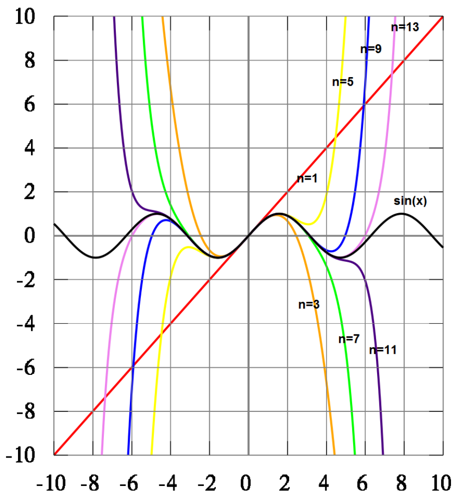
is the n^{th} *Taylor polynomial of* f *at* a.

Even better, we have a handle on the error involved in the approximation $f(x) \approx T_{f,a}(x)$ – it is precisely the remainder, and we have

$$\begin{aligned} f(x) - T_{f,a}(x) &= R_{n+1}(x), \\ R_{n+1}(x) &= \frac{1}{n!}\int_a^x f^{(n+1)}(t)(x-t)^n \,\mathrm{d}t, \\ |R_{n+1}(x)| &= \frac{1}{n!}\left|\int_a^x f^{(n+1)}(t)(x-t)^n \,\mathrm{d}t\right|, \\ &\le \frac{1}{n!}\max_{t\in[\alpha,\beta]}|f^{(n+1)}(t)|\,|x-a|^{n+1}. \end{aligned} \tag{8.6}$$

The example in the figure on the left shows the Taylor polynomials centred at $x = 0$ of degree $n = 1, 3, 5, 7, 9, 11$, and 13 and the approximation they give to $\sin(x)$.[a] It can be seen that the approximation to $\sin(x)$ near $x = 0$ given by the Taylor polynomials gets better and better as the degree of the polynomial is increased.

[a] "Sintay SVG" by IkamusumeFan - Own work. Licensed under CC BY-SA 3.0 via Commons - https://commons.wikimedia.org/wiki/File:Sintay_SVG.svg#/ı



We can give this observation some theoretical backbone by looking at Equation (8.6) with $a = 0$ and

$$f^{(n)}(0) = \pm\sin(0) \text{ or } \pm\cos(0) \implies |f^{(n)}(0)| \le 1$$

to obtain

$$|R_n(x)| \le \frac{1}{n!}x^{n+1}$$

so that in the limit $n \to \infty$, $R_n(x) \to 0$ – in this example, we are **guaranteed** that the approximation given by the Taylor polynomials gets better and better (for arbitrary finite x) as the degree of the polynomials is increased indefinitely.

Example: Estimate e^x correct to three significant figures for $x = 0.01$.

We make various approximations to $f(x) = \mathrm{e}^x$ for $x = 0.01$, based on truncations of the Taylor polynomials of e^x centred at zero:

$$T_n(x) = \sum_{n=0}^{n} \frac{1}{k!}x^k.$$

n	$T_n(x)$	Value
0	1	1
1	$1+x$	1.01
2	$1+x+(x^2/2)$	1.01005
3	$1+x+(x^2/2)+(x^3/6)$	$1.0100501666666\cdots$
4	$1+(x^2/2)+(x^3/6)+(x^4/24)$	$1.010050167083334\cdots$

Table 8.1: Approximations of e^x by truncations of the Maclaurin series at $x = 0.01$ (Maclaurin series $=$ Taylor series centred at $a = 0$)

Based on Table 8.1,

$$\mathrm{e}^x = 1.010, \qquad \text{(correct to four significant figures)},$$

and the approximation

$$\mathrm{e}^x \approx 1 + x$$

is sufficient to capture this function-evaluation to the prescribed level of accuracy.

Note that the remainder is given by the following expressions for the Taylor polynomials $T_n(x)$ centred at $a = 0$:

$$\mathrm{e}^x - T_n(x) = R_{n+1}(x), \qquad R_{n+1}(x) = \frac{1}{n!}\int_0^x f^{(n+1)}(t)(x-t)^n \mathrm{d}t.$$

We have $f(x) = \mathrm{e}^x$, hence

$$\begin{aligned}
|R_{n+1}(x)| &= \frac{1}{n!}\left|\int_0^x \mathrm{e}^t (x-t)^n \mathrm{d}t\right|, \\
&\leq \frac{1}{n!}\left(\max_{[0,x]} \mathrm{e}^x\right)|x|^{n+1}, \\
&\leq \frac{1}{n!}\max\{1, \mathrm{e}^x\}|x|^{n+1}, \\
&\to 0 \text{ as } n \to \infty,
\end{aligned}$$

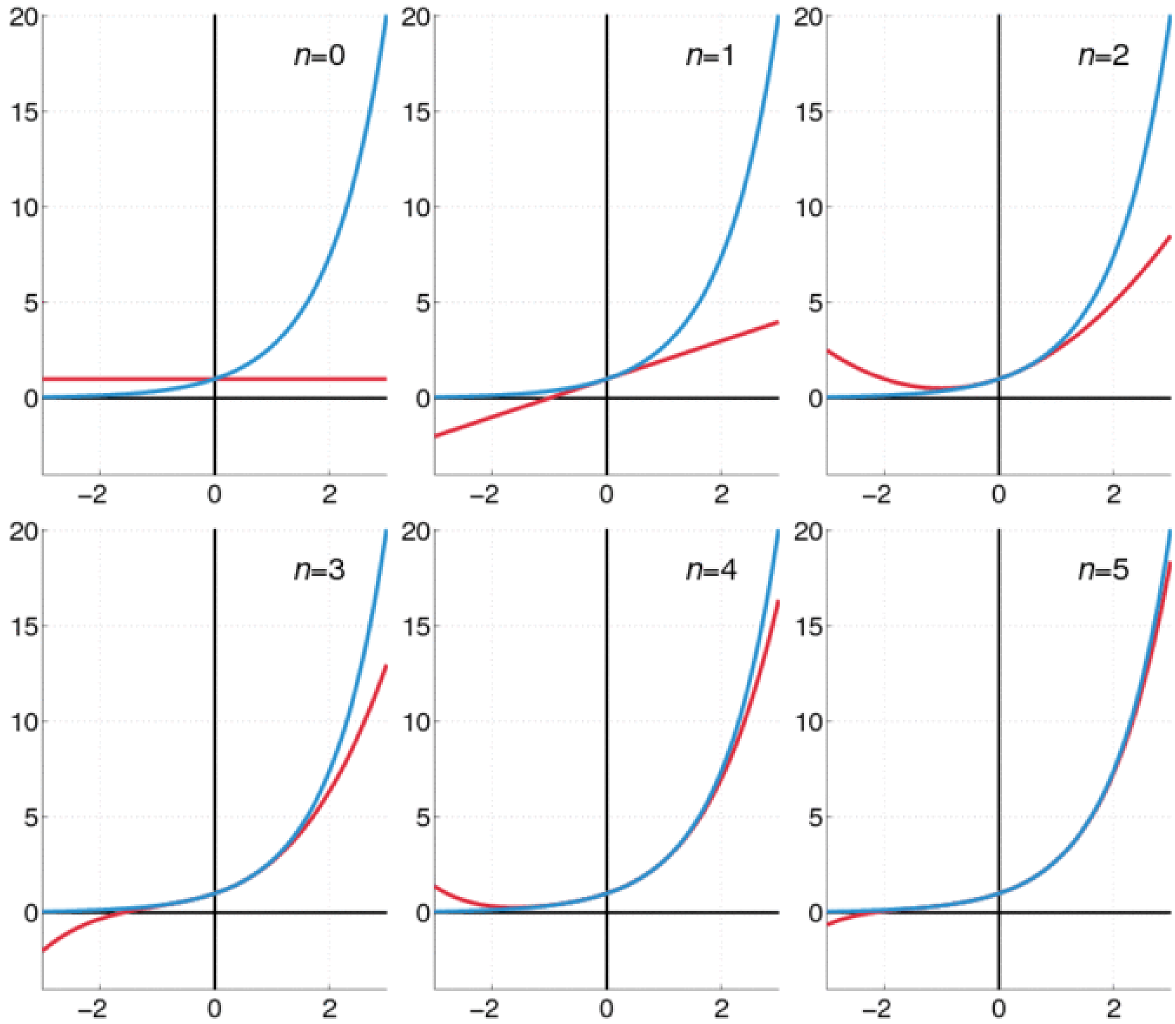
and thus the remainder vanishes as $n \to \infty$.[1] Thus, we can write

$$\mathrm{e}^x = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{1}{k!}x^k = \sum_{k=0}^{\infty}\frac{1}{k!}x^k. \tag{8.7}$$

This is shown schematically in the figure below. The convergence of infinite series such as the one

[1] I determine that $(1/n!)x^{n+1} \to 0$ as $n \to \infty$ for all real values of x by using Stirling's approdximation of the factorial $n!$ for large values of n, i.e. $n! \sim \sqrt{2\pi n}(n/e)^n$ as $n \to \infty$.

in Equation (8.7) is the next topic for discussion.



8.3 Convergence of Taylor series

Suppose that

- $f(x)$ is C^∞ on the interval of interest (α, β),
- The series
$$\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a)^k$$
is a convergent power series,
- The remainder $R_n(x)$ tends to zero as $n \to \infty$ for all $x \in (\alpha, \beta)$

then we can take $n \to \infty$ in the approximation of $f(x)$ by the Taylor polynomials, and the approximation becomes exact:

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(a)(x-a)^k \tag{8.8}$$

Definition 8.2 (Taylor series, Maclaurin series) *Equation* (8.8) *is called the Taylor series of* $f(x)$ *centred at* a. *The special case* $a = 0$ *is called the Maclaurin series of* $f(x)$.

Example: We have already shown that

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \qquad |x| < 1.$$

Use Taylor's Theorem to arrive at this result in a different way.

Solution: Let $f(x) = (1-x)^{-1}$. We have

$$\begin{aligned} f'(x) &= (-1)(-1)(1-x)^{-2}, \\ f''(x) &= (-2)(-1)(1-x)^{-3}, \\ f'''(x) &= (-2)(-1)(-3)(-1)(1-x)^{-4}, \\ f^{(n)}(x) &= n(n-1)\cdots 1(1-x)^{-n-1}, \end{aligned}$$

hence

$$\begin{aligned} f(0) &= 1, \\ f'(x) &= 1, \\ f''(x) &= 2, \\ f'''(x) &= 3 \times 2 \times 1. \end{aligned}$$

Guessing the pattern, we have $f^{(n)}(0) = n!$. We compute the remainder, setting $a = 0$ in Equation (8.2):

$$\begin{aligned} R_n(x) &= \frac{1}{(n-1)!}\int_0^x n!(1-t)^{-n-1}(x-t)^{n-1}\,\mathrm{d}t, \\ &= n\int_0^x \frac{(x-t)^{n-1}}{(1-t)^{n+1}}\,\mathrm{d}t, \\ &= \left.\frac{(1-t)^{-n}(x-t)^n}{x-1}\right|_0^x, \\ &= \frac{x^n}{1-x}. \end{aligned}$$

Therefore, we need $|x| < 1$ such that $R_n \to 0$ as $n \to \infty$, in which case we can write

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0) x^n = \sum_{n=0}^{\infty} x^n.$$

Each coefficient in the power series is equal to unity and hence the radius of convergence is also equal to unity. Thus,

$$|x| < 1 \qquad \text{(for convergence)},$$

which is exactly the same interval on which the remainder vanishes.

8.3.1 The mathematician's nightmare

Sometimes we get sloppy in calculating Taylor series. Indeed, after a while, one can arrive at the complacent conclusion that a function that is infinitely continuously differentiable will definitely have a Taylor series. This is mistaken. For, consider the function

$$f(x) = \begin{cases} \mathrm{e}^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

All the derivatives of $f(x)$ are zero at zero. Thus, a Maclaurin series of $f(x)$ at $x \neq 0$ would give

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)x^n/n! = 0.$$

Yet $f(x) = 0$ only at $x = 0$. This is a contradiction! The only way out of the contradiction is to conclude that the Taylor series of $f(x)$ centred at zero does not converge to $f(x)$. Thus, just because a function is infinitely continuously differentiable does not guarantee that the existence of a 'proper' Taylor series – in principle we need to rigorously analyse the remainder as well.

Of course, it is a bit of a pain to analyse the remainder in great depth all of the time, and the above example is highly exceptional and a bit contrived. Therefore, in the rest of this module – and in many modules in Applied Mathematics, we shall simply use Taylor's theorem in a procedural way, with an expectation of convergence. If our results look dodgy we shall re-examine this approach and analyse things in proper depth.

Things get better when we move to the complex plane, taking $x \to z = x + \mathrm{i}y$. There, a function is complex-differentiable if and only if it has a Taylor-series representation. You will learn about such functions in MATH 30040 (Functions of One Complex Variable).

8.4 Big-O notation

Mindful of the discussion above, we consider a smooth function $f(x)$ and assume it can be written in terms of its Taylor series:

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2}f''(x)h^2 + \cdots .$$

For small values of h, from our discussion about Taylor polynomials, we know that

$$f(x+h) \approx f(x) + f'(x)h$$

is a good approximation to $f(x+h)$. We also know that the error incurred in this approximation is Taylor's remainder R_2, with

$$\begin{aligned} R_2 &= \int_x^{x+h} f''(t)(x-t)\,\mathrm{d}t, \\ |R_2| &\leq \tfrac{1}{2} \sup_{(x,x+h)} |f''(t)| h^2, \end{aligned}$$

and thus the error is proportional to h^2.

Definition 8.3 (Big-O notation) *Based on the discussion above, we write*

$$f(x+h) = f(x) + f'(x)h + O(h^2)$$

to mean that $f(x)$ can be approximated by $f(x) + f'(x)h$ where the error or leftover term is proprtional to h^2. Similarly,

$$f(x+h) = f(x) + f'(x)h + \tfrac{1}{2}f''(x)h^2 + \cdots + \tfrac{1}{n!}f^{(n)}(x)h^n + O(h^{n+1})$$

means that $f(x+h)$ can be approximated by a Taylor polynomial of degree n in h, where the leftover term is proportional to h^{n+1}.

8.5 Taylor's theorem in several dimensions

We consider here an expression for the first-order terms in Taylor's expansion in multivariate calculus. This result is a simple consequence of single-variable version of Taylor's theorem, together with the standard rules of partial derivatives. We prove the following theorem:

Theorem 8.2 (Taylor's theorem in two dimensions) *Let $f(x, y)$ be a function of two variables.*

For $f(x,y)$ sufficiently smooth,

$$f(x+\delta x, y+\delta y) = f(x,y) + f_x(x,y)\delta x + f_y(x,y)\delta y + O(\delta x^2, \delta y^2, \delta x \delta y),$$

where $f_x = \partial f/\partial x$, and $f_y = \partial f/\partial y$.

Proof: Call

$$F(x) := f(x, y+\delta y), \qquad \text{fixed } y.$$

From the single-variable version of Taylor's theorem,

$$F(x+\delta x) = F(x) + F'(x)\delta x + O(\delta x^2);$$

in other words,

$$f(x+\delta x, y+\delta y) = f(x, y+\delta y) + f_x(x, y+\delta y)\delta x + O(\delta x^2). \tag{8.9}$$

Now introduce

$$G_0(y) = f(x,y), \qquad \text{fixed } x,$$

and

$$G_1(y) = f_x(x,y), \qquad \text{fixed } x.$$

Hence,

$$G_0(y+\delta y) = G_0(y) + G_0'(y)\delta y + O(\delta y^2) \implies f(x, y+\delta y) = f(x,y) + f_y(x,y)\delta y + O(\delta y^2);$$

similarly,

$$G_1(y+\delta y) = G_1(y) + G_1'(y)\delta y + O(\delta y^2) \implies f_x(x, y+\delta y) = f_x(x,y) + \partial_y f_x(x,y) + O(\delta y^2).$$

Consider again Eq. (8.9):

$$f(x+\delta x, y+\delta y) = \underbrace{f(x, y+\delta y)}_{=f(x,y)+f_y(x,y)\delta y+O(\delta y^2)} + \delta x \underbrace{[f_x(x, y+\delta y)]}_{f_x(x,y)+\partial_y f_x(x,y)+O(\delta y^2)} + O(\delta x^2).$$

Hence,

$$f(x+\delta x, y+\delta y) = f(x,y) + f_x(x,y)\delta x + f_y(x,y)\delta y + f_{xy}(x,y)\delta x\delta y + O(\delta x^2, \delta y^2).$$

The terms that are proportional to δx^2, δy^2 and $\delta x \delta y$ are bundled up into an error term as follows:

$$f(x+\delta x, y+\delta y) = f(x,y) + f_x(x,y)\delta x + f_y(x,y)\delta y + O(\delta x^2, \delta y^2, \delta x \delta y). \qquad \blacksquare$$

The proof extends readily to higher dimensions. You will encounter the above approximation in MATH 20060 wherein the following definition will be provided:

Definition 8.4 *The quantity $f_x(x,y)\delta x + f_y(x,y)\delta y$ is called the differential of f at the point (x,y).*

Chapter 9

Div, grad, and curl and all that

Overview

We introduce the **gradient and curl operators** as being distinguished ways of differentiating scalar and vector fields that carry special physical significance.

9.1 The gradient operator in three dimensions

Definition 9.1 (The gradient operator) *Let ϕ be a function of three variables,*

$$\begin{aligned} \phi : \left(\Omega \subset \mathbb{R}^3\right) &\rightarrow \mathbb{R} \\ (x, y, z) &\rightarrow \phi\left(x, y, z\right). \end{aligned}$$

Then the gradient operator acting on ϕ is a vector with the following form:

$$\mathrm{grad}\,\phi := \hat{\boldsymbol{x}}\frac{\partial \phi}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial \phi}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial \phi}{\partial z}.$$

In class, we will write this vector as $\nabla\phi$, and call it 'grad ϕ' or 'nabla ϕ'.

Example: Compute the gradient of the function $\phi\left(x, y, z\right) = x^2 + y^2 + z^2$.

Solution: We know that $\partial_x \phi = 2x$, $\partial_y \phi = 2y$, and $\partial_z \phi = 2z$. Hence,

$$\nabla\phi = \hat{\boldsymbol{x}}\partial_x\phi + \hat{\boldsymbol{y}}\partial_y\phi + \hat{\boldsymbol{z}}\partial_z\phi = 2\hat{\boldsymbol{x}}x + 2\hat{\boldsymbol{y}}y + 2z\hat{\boldsymbol{z}} = 2\left(x, y, z\right) = 2\boldsymbol{x},$$

where $\boldsymbol{x}$ is a position vector.

Example: Compute the gradient of the function $\phi(x,y,z) = 1/\sqrt{x^2+y^2+z^2}$.

Solution: We know that

$$\begin{aligned}
\partial_x \phi &= -\frac{x}{\left(x^2+y^2+z^2\right)^{3/2}}, \\
\partial_y \phi &= -\frac{y}{\left(x^2+y^2+z^2\right)^{3/2}}, \\
\partial_z \phi &= -\frac{z}{\left(x^2+y^2+z^2\right)^{3/2}}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\nabla\phi = \hat{\boldsymbol{x}}\left(-\frac{x}{\left(x^2+y^2+z^2\right)^{3/2}}\right) + \hat{\boldsymbol{y}}\left(-\frac{y}{\left(x^2+y^2+z^2\right)^{3/2}}\right)\hat{\boldsymbol{z}}\left(-\frac{z}{\left(x^2+y^2+z^2\right)^{3/2}}\right) \\
= -\frac{(x,y,z)}{\left(x^2+y^2+z^2\right)^{3/2}} = -\frac{\boldsymbol{x}}{|\boldsymbol{x}|^3}.
\end{aligned}$$

9.2 The physical meaning of the gradient

In three dimensions, the surface is specified by an equation of the type

$$\phi(x,y,z) = 0.$$

This is the generic equation for a surface because if ϕ is sufficiently smooth, it can be inverted in the neighbourhood of a given point and an expression of the kind $z = z(x,y)$ can be found, which gives a surface (Fig. 9.1). Suppose that $\boldsymbol{x} = (x,y,z)$ satisfies $\phi = 0$. Then $\nabla\phi$ evaluated at $\boldsymbol{x}$ is normal to the surface. To prove this, we take $\boldsymbol{x} + \delta\boldsymbol{x}$, a neighbouring point of $\boldsymbol{x}$ that still resides on the surface. We form the difference

$$\begin{aligned}
0 &= 0 - 0, \\
&= \phi(\boldsymbol{x} + \delta\boldsymbol{x}) - \phi(\boldsymbol{x}), \\
&= \phi(x+\delta x, y+\delta y, z+\delta z) - \phi(x,y,z), \\
&= \frac{\partial\phi}{\partial x}(x,y,z)\delta x + \frac{\partial\phi}{\partial y}(x,y,z)\delta y + \frac{\partial\phi}{\partial z}(x,y,z)\delta z + \text{H.O.T.}, \\
&= \nabla\phi\cdot\delta\boldsymbol{x}.
\end{aligned}$$

But $\boldsymbol{x}$ and $\boldsymbol{x} + \delta\boldsymbol{x}$ are vectors whose tip lies on the surface (Fig. 9.2). Hence, $\delta\boldsymbol{x}$ is tangent to the surface, and $\nabla\phi\cdot\delta\boldsymbol{x} = 0$, so $\nabla\phi(x,y,z)$ is normal to the surface.

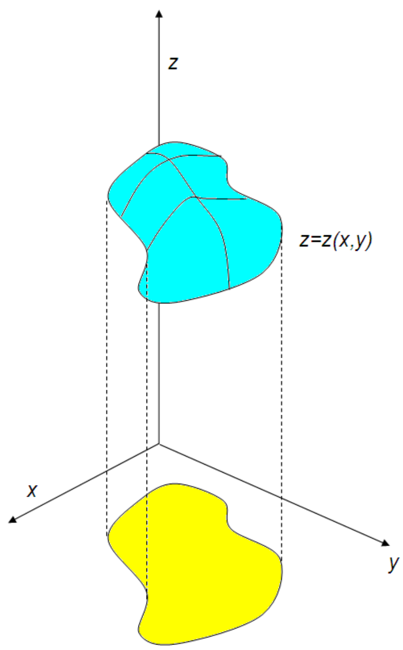


Figure 9.1: A schematic description of a surface described by $z = z(x, y)$.

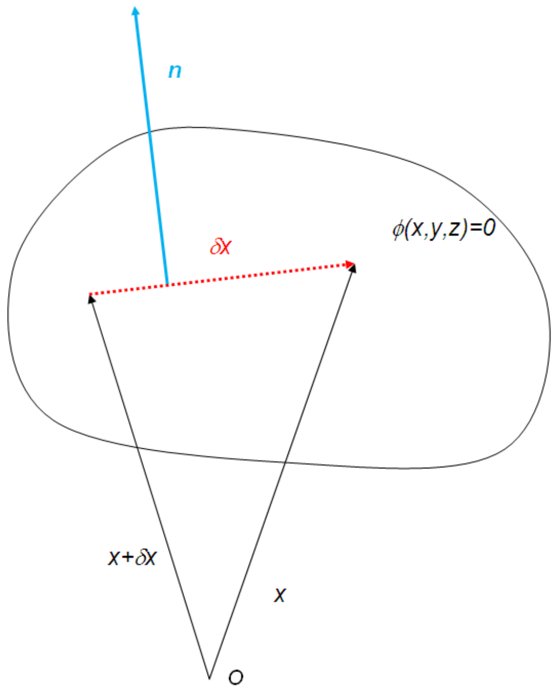


Figure 9.2: A schematic description of the normal to a surface, $\boldsymbol{n} = \nabla\phi$.

The directional derivative

Suppose we have a scalar field $\phi(x, y, z)$ and we want to know how it changes in a given, fixed direction $\boldsymbol{e}$. The way to do this is to form the difference

$$\delta\phi = \phi(\boldsymbol{x} + t\boldsymbol{e}) - \phi(\boldsymbol{x}),$$

where t is a parameter that takes all real values. In particular, let t be small. Then we have

$$\begin{aligned}
\delta\phi &= \phi(x + te_x, y + te_y z, z + te_z) - \phi(x, y, z), \\
&= \frac{\partial\phi}{\partial x}te_x + \frac{\partial\phi}{\partial y}te_y + \frac{\partial\phi}{\partial z}te_z, \\
&= (\nabla\phi)\cdot\boldsymbol{e}t, \\
&= \boldsymbol{e}\cdot\nabla\phi\, t.
\end{aligned}$$

Hence,

$$\frac{\delta\phi}{t} = \boldsymbol{e}\cdot\nabla\phi.$$

The Taylor approximation becomes exact when $t \to 0$:

$$\boldsymbol{e}\cdot\nabla\phi = \lim_{t\to 0}\frac{\delta\phi}{t}.$$

This is **the directional derivative in the direction e**:

$$\frac{d\phi}{de} := \boldsymbol{e}\cdot\nabla\phi.$$

9.3 Vector fields and the divergence operator

Definition 9.2 (Vector Field) *A vector field $\boldsymbol{v}(x,y,z)$ in $\mathbb{R}^3$ is a map that assigns to each element of its domain $\Omega\subset\mathbb{R}^3$ a uniquely determined vector, also in $\mathbb{R}^3$. In map language,*

$$\begin{aligned}\boldsymbol{v}:\ \left(\Omega\subset\mathbb{R}^3\right) &\rightarrow \mathbb{R}^3,\\ (x,y,z) &\rightarrow \boldsymbol{v}(x,y,z).\end{aligned}$$

Since $\boldsymbol{v}(x,y,z)$ is a vector, we can write

$$\begin{aligned}\boldsymbol{v}(x,y,z) &= \hat{\boldsymbol{x}}v_1(x,y,z)+\hat{\boldsymbol{y}}v_2(x,y,z)+\hat{\boldsymbol{z}}v_3(x,y,z)\\ &= = (v_1(x,y,z),v_2(x,y,z),v_3(x,y,z)).\end{aligned}$$

Example: if $\phi(x,y,z)$ is a scalar field, then

$$\nabla\phi=\hat{\boldsymbol{x}}\partial_x\phi+\hat{\boldsymbol{y}}\partial_y\phi+\hat{\boldsymbol{z}}\partial_z\phi$$

is a vector field.

Definition 9.3 (Divergence) *The divergence of a vector field $\boldsymbol{v}(x,y,z)$ is a scalar computed as follows:*

$$\operatorname{div}\boldsymbol{v}=\frac{\partial v_1}{\partial x}+\frac{\partial v_2}{\partial y}+\frac{\partial v_3}{\partial z}.$$

Formally, this is like 'dotting' ∇ with $\boldsymbol{v}$, so we write

$$\operatorname{div}\boldsymbol{v}=\nabla\cdot\boldsymbol{v}.$$

There is one crucial difference between ordinary dot products and the divergence: for ordinary vectors $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{B}\cdot\boldsymbol{A}$; for vector fields, $\nabla\cdot\boldsymbol{v}$ is NOT equal to $\boldsymbol{v}\cdot\nabla$.

9.3.1 Examples

Example: Compute $\nabla \cdot \boldsymbol{v}$, where $\boldsymbol{v}(x,y,z) = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z$.

Solution: If $\boldsymbol{v}(x,y,z) = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z$, then

$$\mathrm{div}\,\boldsymbol{v} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3.$$

Example: Consider the vector field

$$v_1 = \frac{\partial \psi}{\partial y},\ v_2 = -\frac{\partial \psi}{\partial x},\ v_3 = 0.$$

Show that $\nabla \cdot \boldsymbol{v} = 0$.

Solution:

$$\begin{aligned}\mathrm{div}\,\boldsymbol{v} &= \frac{\partial}{\partial x}\frac{\partial \psi}{\partial y} + \frac{\partial}{\partial y}\left(-\frac{\partial \psi}{\partial x}\right) + \frac{\partial}{\partial z}0,\\ &= \frac{\partial}{\partial x}\frac{\partial \psi}{\partial y} - \frac{\partial}{\partial y}\frac{\partial \psi}{\partial x}.\end{aligned}$$

We now use the remarkable fact that **the partial derivatives of smooth functions commute**, $\partial_{xy}\psi = \partial_{yx}\psi$ to obtain

$$\mathrm{div}\,\boldsymbol{v} = 0.$$

Definition 9.4 *A vector field whose divergence is zero is called* **incompressible***.*

Example: Consider the vector field

$$\boldsymbol{v} = \hat{\boldsymbol{x}}\partial_x\phi + \hat{\boldsymbol{y}}\partial_y\phi + \hat{\boldsymbol{z}}\partial_z\phi,$$

where $\phi(x,y,z)$ is some scalar field. Show that $\nabla \cdot \boldsymbol{v} = \nabla^2\phi$.

Solution: The divergence of $\boldsymbol{v}$ is

$$\begin{aligned}\nabla \cdot \boldsymbol{v} &= \partial_x v_1 + \partial_y v_2 + \partial_z v_3,\\ &= \partial_x\partial_x\phi + \partial_y\partial_y\phi + \partial_z\partial_z\phi,\\ &= \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}.\end{aligned}$$

This particular operation on the scalar field ϕ is quite common in physics and is therefore given its own name: it is called the **Laplacian**, and given the notation ∇^2 (or Δ):

Definition 9.5 *The Laplacian of the scalar field ϕ is given by*

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

9.4 The physical meaning of the divergence

Consider a fluid that flows in a three-dimensional container. We take a small cuboid of sides of length Δx, Δy and Δz as a control volume, one of whose vertices lies at (x, y, z). Fluid flows into and out of the cuboid with velocity $\boldsymbol{v}(x, y, z, t)$ (the t is for time). The amount of mass leaving the system through the x-direction per unit time is

$$\begin{aligned}
\frac{\text{Mass Out} - \text{Mass In}}{\text{Time}} \text{ in the } x \text{ direction} &= (\rho v_1)(x + \Delta x, y, z)\,\Delta y \Delta z - (\rho v_1)(x, y, z)\,\Delta y \Delta z, \\
&= \frac{\partial}{\partial x}(\rho v_1)\bigg|_{(x,y,z)} \Delta x \Delta y \Delta z + \text{H.O.T.},
\end{aligned}$$

where $\rho(x, y, z, t)$ is the scalar **fluid density** (Fig. 9.3). Similarly,

$$\begin{aligned}
\frac{\text{Mass Out} - \text{Mass In}}{\text{Time}} \text{ in the } y \text{ direction} &= \frac{\partial}{\partial y}(\rho v_2)\bigg|_{(x,y,z)} \Delta x \Delta y \Delta z + \text{H.O.T.}, \\
\frac{\text{Mass Out} - \text{Mass In}}{\text{Time}} \text{ in the } z \text{ direction} &= \frac{\partial}{\partial z}(\rho v_3)\bigg|_{(x,y,z)} \Delta x \Delta y \Delta z + \text{H.O.T.},
\end{aligned}$$

Adding them,

$$\begin{aligned}
\frac{\text{Total Mass Out} - \text{Total Mass In}}{\text{Time}} &= \left[\frac{\partial}{\partial x}(\rho v_1) + \frac{\partial}{\partial y}(\rho v_3) + \frac{\partial}{\partial z}(\rho v_3)\right]_{(x,y,z)} \Delta x \Delta y \Delta z \\
&= \nabla \cdot (\rho \boldsymbol{v})\bigg|_{(x,y,z)} \Delta x \Delta y \Delta z.
\end{aligned}$$

Now in this control volume, matter is not created or destroyed, so the change in the mass in the control volume over time must be balanced by changes in the density over time:

$$\begin{aligned}
\frac{\text{Total Mass Out} - \text{Total Mass In}}{\text{Time}} &= \frac{\text{Change in mass}}{\text{Time}} \\
&= -\frac{\partial}{\partial t}(\rho \Delta x \Delta y \Delta z).
\end{aligned}$$

Why is there a minus sign here? Well, if Total Mass Out > Total Mass in, then the LHS will be a positive quantity. At the same time, $\partial\rho/\partial t$ will be negative (the box is losing mass). Therefore, in order for the signs to balance, we need

$$\text{sign}\,(\text{LHS}) = +1 = \text{sign}\,(\text{RHS}) = \text{sign}\,[(-1)\partial\rho/\partial t]\,.$$

Finally, we equate these two identical changes and take the constant volume element $\Delta x\Delta y\Delta z$ outside the time derivative:

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

This result is called the **continuity equation** and holds whenever some continuum quantity is conserved (e.g. mass or charge).

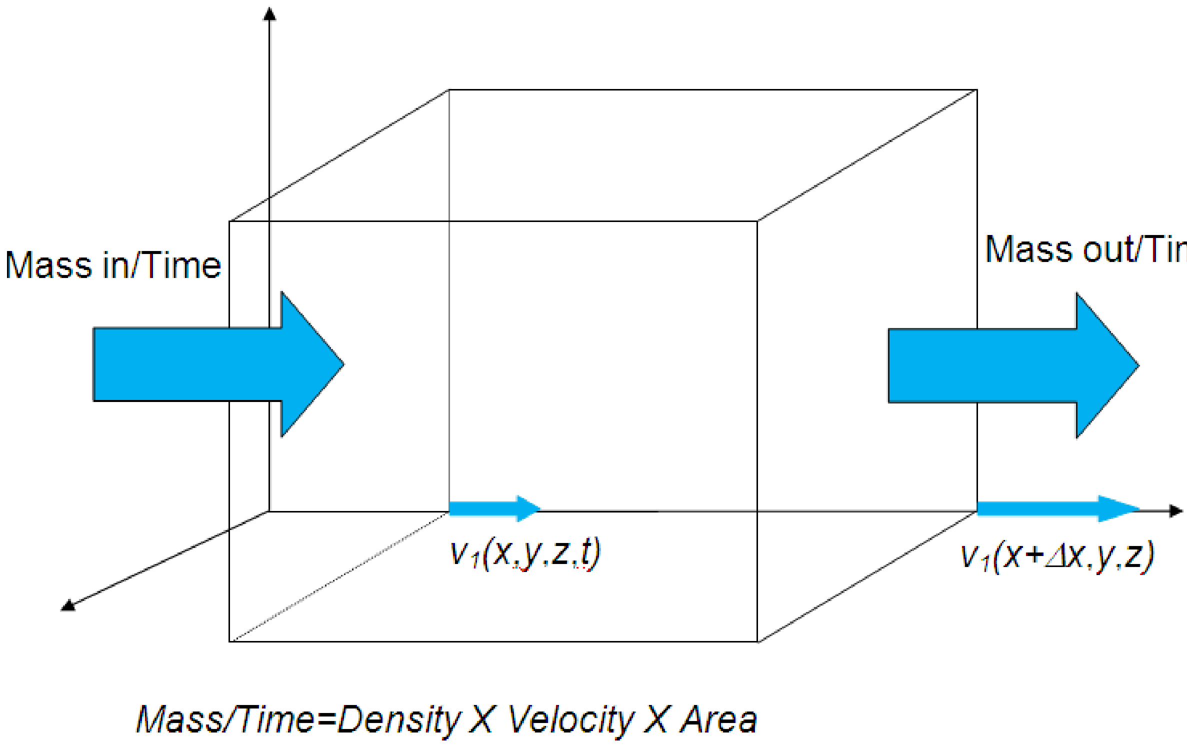


Figure 9.3: The physical meaning of the divergence

9.5 The curl of a vector field

Definition 9.6 *Let $\boldsymbol{v}(x,y,z) = \hat{\boldsymbol{x}}v_1(x,y,z) + \hat{\boldsymbol{y}}v_2(x,y,z) + \hat{\boldsymbol{z}}v_3(x,y,z)$ be a vector field. The curl of $\boldsymbol{v}$ is a new vector field formed as follows*

$$
\begin{aligned}
\mathrm{curl}\,\boldsymbol{v} &:= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \partial_x & \partial_y & \partial_z \\ v_1 & v_2 & v_3 \end{vmatrix}, \\
&= \hat{\boldsymbol{x}}\left(\partial_y v_3 - \partial_z v_2\right) + \hat{\boldsymbol{y}}\left(\partial_z v_1 - \partial_x v_3\right) + \hat{\boldsymbol{z}}\left(\partial_x v_2 - \partial_y v_1\right). \qquad (9.1)
\end{aligned}
$$

Because this expression for the curl is like the ordinary cross product of two vectors, we write

$$\mathrm{curl}\,\boldsymbol{v} = \nabla \times \boldsymbol{v}.$$

There is one crucial difference between ordinary cross products and the curl: for ordinary vectors $\boldsymbol{A} \times \boldsymbol{B} = -\boldsymbol{B} \times \boldsymbol{A}$; for vector fields, $\nabla \times \boldsymbol{v}$ is NOT equal to $-\boldsymbol{v} \times \nabla$.

9.6 The physical meaning of the curl

Consider the following vector field:

$$\boldsymbol{v}(x,y,z) = y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}}.$$

Imagine that this represents the velocity of a fluid in a container. We can plot the vector field by drawing a little arrow at random points (x,y) in two-dimensional space (Fig. 9.4). The arrow should have length $\sqrt{x^2+y^2}$ and point in the direction $y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}}$. Simply by inspection, we see that the field is rotating. If we stick a paddle wheel anywhere in the fluid, it will be carried by the flow and rotate clockwise. Using the right-hand rule, we expect the curl to be into the page. If we are to keep a right-handed coordinate system, the negative z-direction must point into the page. Moreover, we can apply the formula

$$\omega = \frac{v}{r}$$

for ordinary circular motion, to the vector field $\boldsymbol{v}$, giving

$$\omega = \frac{|y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}}|}{\sqrt{x^2+y^2}} = 1$$

Thus, the ‘amount of rotation’ in the vector field is constant (independent of position), and the sense of rotation is into the page.

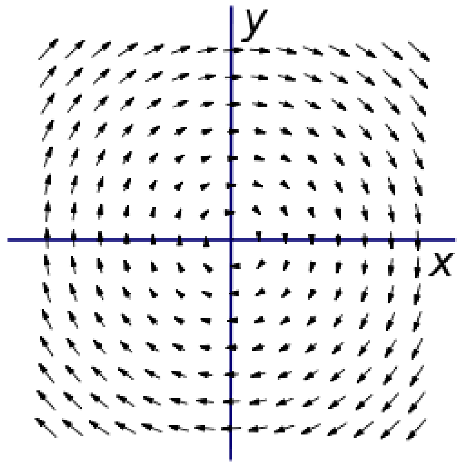


Figure 9.4: The vector field $\boldsymbol{v} = y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}}$.

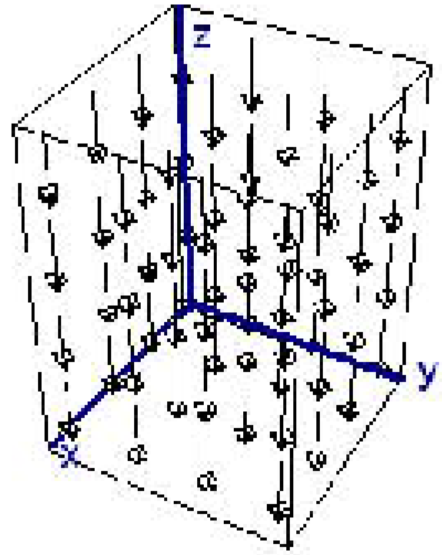


Figure 9.5: The curl of the vector field $\boldsymbol{v} = y\hat{\boldsymbol{x}} - x\hat{\boldsymbol{y}}$.

Now, we calculate the curl:

$$\nabla \times \boldsymbol{v} = 0\hat{\boldsymbol{x}} + 0\hat{\boldsymbol{y}} + \left[\frac{\partial}{\partial x}(-x) - \frac{\partial}{\partial y}y\right]\hat{\boldsymbol{z}} = -2\hat{\boldsymbol{z}};$$

it is indeed in the negative z-direction. It is also a constant! Thus, the curl corresponds to our intuitive idea about the amount of rotation in a vector field.

Plotting the curl of $\boldsymbol{v}$ is not very interesting (Fig. 9.5). Nevertheless, we see the physical meaning of curl: **it tells us by how much a vector field is rotating, and in what sense**.

9.7 Formulas involving div, grad, and curl

Theorem 9.1 *Let $\phi(x, y, z)$ and $\psi(x, y, z)$ be differentiable scalar fields and let $\boldsymbol{u}(x, y, z)$ and $\boldsymbol{v}(x, y, z)$ be differentiable vector fields. Then the following identities hold:*

1. $\nabla(\phi + \psi) = \nabla\phi + \nabla\psi$;
2. $\nabla \cdot (\boldsymbol{u} + \boldsymbol{v}) = \nabla \cdot \boldsymbol{u} + \nabla \cdot \boldsymbol{v}$;
3. $\nabla \times (\boldsymbol{u} + \boldsymbol{v}) = \nabla \times \boldsymbol{u} + \nabla \times \boldsymbol{v}$;
4. $\nabla \cdot (\phi\boldsymbol{u}) = (\nabla\phi) \cdot \boldsymbol{u} + \phi(\nabla \cdot \boldsymbol{u})$;
5. $\nabla \times (\phi\boldsymbol{u}) = (\nabla\phi) \times \boldsymbol{u} + \phi(\nabla \times \boldsymbol{u})$;

6. $\nabla \cdot (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{v} \cdot (\nabla \times \boldsymbol{u}) - \boldsymbol{u} \cdot (\nabla \times \boldsymbol{v})$;

7. $\nabla \times (\boldsymbol{u} \times \boldsymbol{v}) = (\boldsymbol{v} \cdot \nabla)\, \boldsymbol{u} - \boldsymbol{v}\, (\nabla \cdot \boldsymbol{u}) - (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{v} + \boldsymbol{u}\, (\nabla \cdot \boldsymbol{v})$;

8. $\nabla\, (\boldsymbol{u} \cdot \boldsymbol{v}) = (\boldsymbol{v} \cdot \nabla)\, \boldsymbol{u} + (\boldsymbol{u} \cdot \nabla)\, \boldsymbol{v} + \boldsymbol{v} \times (\nabla \times \boldsymbol{u}) + \boldsymbol{u} \times (\nabla \times \boldsymbol{v})$;

9. $\nabla \times (\nabla \times \boldsymbol{u}) = \nabla\, (\nabla \cdot \boldsymbol{u}) - \nabla^2 \boldsymbol{u}$.

Properties 1–3 are obvious; the others are tricky and some of them will appear as exercises. Note that if λ and μ are a scalars (*constant* real numbers), then

$$\nabla\, (\lambda\phi + \mu\psi) = \lambda\nabla\phi + \mu\nabla\psi,$$

and similarly, for vector fields,

$$\nabla \cdot (\lambda\boldsymbol{u} + \mu\boldsymbol{v}) = \lambda\nabla \cdot \boldsymbol{u} + \mu\nabla \cdot \boldsymbol{v},$$

and

$$\nabla \times (\lambda\boldsymbol{u} + \mu\boldsymbol{v}) = \lambda\nabla \times \boldsymbol{u} + \mu\nabla \times \boldsymbol{v}.$$

This is the property of **linearity**. The operations div, grad, and curl thus take vector or scalar fields and map them linearly to other vector or scalar fields. They are thus called **linear operators**. In the next chapter we will gain more proficiency in handling these operators.

Chapter 10

Techniques in vector differentiation

Overview

In this section we gain more familiarity with the vector operators div, grad, and curl by doing a number of examples.

10.1 Calculations

Example: If $\boldsymbol{u} = 2x^2\hat{\boldsymbol{x}} - 3yz\hat{\boldsymbol{y}} + xz^2\hat{\boldsymbol{z}}$ and $\phi = 2z - x^3y$, find $\boldsymbol{u} \cdot \nabla\phi$ at the point $(1, -1, 1)$ and $\boldsymbol{u} \times \nabla\phi$ at the point $(1, -1, 1)$.

Solution: We have

$$\begin{aligned}
\boldsymbol{u} &= 2x^2\hat{\boldsymbol{x}} - 3yz\hat{\boldsymbol{y}} + xz^2\hat{\boldsymbol{z}}, \\
\phi &= 2z - x^3y, \\
\mathrm{grad}\,\phi &= -3x^2y\,\hat{\boldsymbol{x}} - x^3\,\hat{\boldsymbol{y}} + 2\,\hat{\boldsymbol{z}}.
\end{aligned}$$

At $(1, -1, 1)$,

$$\begin{aligned}
\boldsymbol{u} &= 2\hat{\boldsymbol{x}} + 3\hat{\boldsymbol{y}} + \hat{\boldsymbol{z}}, \\
\mathrm{grad}\,\phi &= 3\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + 2\hat{\boldsymbol{z}}, \\
\boldsymbol{u} \cdot \mathrm{grad}\,\phi &= 6 - 3 + 2 = 5.
\end{aligned}$$

Also,

$$
\begin{aligned}
\boldsymbol{u} \times \operatorname{grad} \phi &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ 2 & 3 & 1 \\ 3 & -1 & 2 \end{vmatrix} \\
&= 7\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} - 11\ \hat{\boldsymbol{z}}.
\end{aligned}
$$

Example: If $\nabla\phi = 2xyz^3\hat{\boldsymbol{x}} + x^2z^3\hat{\boldsymbol{y}} + 3x^2yz^2\hat{\boldsymbol{z}}$, find $\phi(x, y, z)$ if $\phi(1, -2, 2) = 4$.

Solutioni: We have

$$
\begin{aligned}
\frac{\partial\phi}{\partial x} &= 2xyz^3, \\
\phi &= x^2yz^3 + f(y, z), \\
\frac{\partial\phi}{\partial y} &= x^2z^3, \\
\phi &= x^2yz^3 + g(z, x), \\
\frac{\partial\phi}{\partial z} &= 3x^2yz^2, \\
\phi &= x^2yz^3 + h(x, y).
\end{aligned}
$$

Therefore

$$f(y, z) = g(z, x) = h(x, y) = c \text{ (constant)}$$

and

$$\phi(x, y, z) = x^2yz^3 + c.$$

$$\phi(1, -2, 2) = -16 + c = 4, \qquad c = 20,$$

$$\phi(x, y, z) = x^2yz^3 + 20.$$

Example: Find the unit outward drawn normal to the surface $(x - 1)^2 + y^2 + (z + 2)^2 = 9$ at the point $(3, 1, -4)$.

Solution: Let $\phi = (x - 1)^2 + y^2 + (z + 2)^2 - 9$. Then

$$\operatorname{grad} \phi = 2(x - 1)\,\hat{\boldsymbol{x}} + 2y\,\hat{\boldsymbol{y}} + 2(z + 2)\,\hat{\boldsymbol{z}}.$$

At $(3, 1, -4)$, $\operatorname{grad} \phi = 4\hat{\boldsymbol{x}} + 2\hat{\boldsymbol{y}} - 4\hat{\boldsymbol{z}}$.

Unit outward drawn normal:

$$\boldsymbol{n} = \frac{\operatorname{grad}\phi}{|\operatorname{grad}\phi|} = \frac{4\hat{\boldsymbol{x}} + 2\hat{\boldsymbol{y}} - 4\hat{\boldsymbol{z}}}{\sqrt{16+4+16}} = \frac{2\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} - 2\hat{\boldsymbol{z}}}{3}.$$

Example: Find the equation for the tangent plane and the equation (not just the direction) of the normal line to the surface $z = x^2 + y^2$ at the point $(2, -1, 5)$

Solution: Let $\phi = x^2 + y^2 - z$. Then

$$\operatorname{grad}\phi = 2x\hat{\boldsymbol{x}} + 2y\hat{\boldsymbol{y}} - \hat{\boldsymbol{z}}.$$

At $(2, -1, 5)$, $\operatorname{grad}\phi = 4\hat{\boldsymbol{x}} - 2\hat{\boldsymbol{y}} - \hat{\boldsymbol{z}}$.

Normal (not necessarily a unit normal):

$$\boldsymbol{n} = \operatorname{grad}\phi = 4\hat{\boldsymbol{x}} - 2\hat{\boldsymbol{y}} - \hat{\boldsymbol{z}}.$$

Let $\boldsymbol{r}_0 = 2\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}} + 5\hat{\boldsymbol{z}}$. The tangent plane at $\boldsymbol{r}_0$ is given by

$$\begin{gathered}
(\boldsymbol{r} - \boldsymbol{r}_0)\cdot\boldsymbol{n} = 0, \\
4(x-2) - 2(y+1) - (z-5) = 0, \\
4x - 2y - z = 5.
\end{gathered}$$

Normal line:

$$\begin{gathered}
\boldsymbol{r} = \boldsymbol{r}_0 + \lambda\boldsymbol{n}, \\
\boldsymbol{r} - \boldsymbol{r}_0 = \lambda\boldsymbol{n}, \\
(x-2)\,\hat{\boldsymbol{x}} + (y+1)\,\hat{\boldsymbol{y}} + (z-5)\,\hat{\boldsymbol{z}} = \lambda(4\hat{\boldsymbol{x}} - 2\hat{\boldsymbol{y}} - \hat{\boldsymbol{z}}), \\
x - 2 = 4\lambda, \qquad y + 1 = -2\lambda, \qquad z - 5 = -\lambda, \\
\frac{x-2}{4} = \frac{y+1}{-2} = \frac{z-5}{-1} \ (= -\lambda).
\end{gathered}$$

Show that $\boldsymbol{u} = (6xy + z^3)\hat{\boldsymbol{x}} + (3x^2 - z)\hat{\boldsymbol{y}} + (3xz^2 - y)\hat{\boldsymbol{z}}$ is irrotational ($\nabla \times \boldsymbol{u} = 0$). Find ϕ such that $\boldsymbol{u} = \nabla\phi$.

Solution: We have

$$\begin{aligned}
\operatorname{curl}\boldsymbol{u} &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ 6xy + z^3 & 3x^2 - z & 3xz^2 - y \end{vmatrix} \\
&= (-1+1)\,\hat{\boldsymbol{x}} - (3z^2 - 3z^2)\,\hat{\boldsymbol{y}} + (6x - 6x)\,\hat{\boldsymbol{z}} = \boldsymbol{0}.
\end{aligned}$$

Suppose $\boldsymbol{u} = \operatorname{grad}\phi$.

$$\begin{aligned}
\frac{\partial \phi}{\partial x} &= 6xy + z^3, \\
\phi &= 3x^2y + xz^3 + f(y,z), \\
\frac{\partial \phi}{\partial y} &= 3x^2 - z, \\
\phi &= 3x^2y - yz + g(z,x), \\
\frac{\partial \phi}{\partial z} &= 3xz^2 - y, \\
\phi &= xz^3 - yz + h(x,y).
\end{aligned}$$

Therefore

$$f(y,z) = -yz + c, \qquad g(z,x) = xz^3 + c, \qquad h(x,y) = 3x^2y + c.$$

and

$$\phi(x,y,z) = 3x^2y + xz^3 - yz + c.$$

Caution: In this example, the final answer is of the form $\phi(x,y,z) = f(y,z) + g(z,x) + h(x,y)$. However, this is not true in general, e.g. $\boldsymbol{u} = \nabla\phi$, where $\phi = z\mathrm{e}^{xy}$ does not decompose into a sum like the one in this example.

10.2 Proofs

Example: Show that $\nabla\times(\nabla\phi)=0$, for any differentiable scalar field $\phi(x,y,z)$.

Solution: We have

$$\begin{aligned}\nabla\times(\nabla\phi) &= \begin{vmatrix}\hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}}\\ \partial_x & \partial_y & \partial_z\\ \partial_x\phi & \partial_y\phi & \partial_z\phi\end{vmatrix}\\ &= \hat{\boldsymbol{x}}\left(\partial_x\partial_y\phi-\partial_y\partial_x\phi\right)-\hat{\boldsymbol{y}}\left(\partial_x\partial_z\phi-\partial_z\partial_x\right)+\hat{\boldsymbol{z}}\left(\partial_x\partial_y\phi-\partial_y\partial_x\phi\right).\end{aligned}$$

Since the scalar field is smooth, the partial derivatives commute, and this sum is zero. This exercise shows the implication

$$\boldsymbol{u}=\nabla\phi \implies \nabla\times\boldsymbol{u}=0.$$

In one of the previous examples we encountered $\nabla\times[(6xy+z^3)\hat{\boldsymbol{x}}+(3x^2-z)\hat{\boldsymbol{y}}+(3xz^2-y)\hat{\boldsymbol{z}}]=0 \implies \exists\phi=3x^2y+xz^3-yz$ such that

$$\nabla\phi=\left[(6xy+z^3)\hat{\boldsymbol{x}}+(3x^2-z)\hat{\boldsymbol{y}}+(3xz^2-y)\hat{\boldsymbol{z}}\right]=\boldsymbol{u}.$$

In fact, the implication always goes both ways:

Theorem 10.1 *A vector field $\boldsymbol{u}(x,y,z)$ is irrotational if and only if it can be written as the gradient of a scalar field,*

$$\boldsymbol{u}(x,y,z)=\nabla\phi(x,y,z).$$

We shall prove this statement later in the module using Stokes's Theorem.

Example: Prove that

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

for any differentiable vector field $\boldsymbol{u}(x,y,z)$.

$$\mathrm{curl}\,(\mathrm{curl}\,\boldsymbol{u})=\begin{vmatrix}\hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}}\\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z}\\ \dfrac{\partial u_3}{\partial y}-\dfrac{\partial u_2}{\partial z} & \dfrac{\partial u_1}{\partial z}-\dfrac{\partial u_3}{\partial x} & \dfrac{\partial u_2}{\partial x}-\dfrac{\partial u_1}{\partial y}\end{vmatrix}.$$

Let $\boldsymbol{B} = \operatorname{curl}(\operatorname{curl}\boldsymbol{u})$. Then

$$\begin{aligned}
B_1 &= \frac{\partial^2 u_2}{\partial x \partial y} - \frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_3}{\partial x \partial z} - \frac{\partial^2 u_1}{\partial z^2} \\
&= \frac{\partial}{\partial x}\left(\underbrace{\frac{\partial u_1}{\partial x}}_{***} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z} \right) - \left(\underbrace{\frac{\partial^2 u_1}{\partial x^2}}_{***} + \frac{\partial^2 u_1}{\partial y^2} + \frac{\partial^2 u_1}{\partial z^2} \right) \\
&= \frac{\partial}{\partial x}(\operatorname{div}\boldsymbol{u}) - \operatorname{grad}^2 u_1.
\end{aligned}$$

Similarly

$$B_2 = \frac{\partial}{\partial y}(\operatorname{div}\boldsymbol{u}) - \operatorname{grad}^2 u_2 \quad \text{and} \quad B_3 = \frac{\partial}{\partial z}(\operatorname{div}\boldsymbol{u}) - \operatorname{grad}^2 u_3.$$

Therefore

$$\begin{aligned}
\operatorname{curl}(\operatorname{curl}\boldsymbol{u}) &= \boldsymbol{B} = B_1\,\hat{\boldsymbol{x}} + B_2\,\hat{\boldsymbol{y}} + B_3\,\hat{\boldsymbol{z}} \\
&= \left(\hat{\boldsymbol{x}}\frac{\partial}{\partial x} + \hat{\boldsymbol{y}}\frac{\partial}{\partial y} + \hat{\boldsymbol{z}}\frac{\partial}{\partial z} \right)(\operatorname{div}\boldsymbol{u}) - \operatorname{grad}^2(u_1\,\hat{\boldsymbol{x}} + u_2\,\hat{\boldsymbol{y}} + u_3\,\hat{\boldsymbol{z}}) \\
&= \operatorname{grad}(\operatorname{div}\boldsymbol{u}) - \operatorname{grad}^2\boldsymbol{u}.
\end{aligned}$$

10.3 Physical application: fluid flow in two dimensions

In a general three-dimensional setting, a vector field $\boldsymbol{u}(x, y, z)$ (and possibly time) describes the velocity of a fluid at location $\boldsymbol{x} = (x, y, z)$. The **vorticity** $\boldsymbol{\omega}(x, y, z)$ measures the amount of rotation in the fluid, and its sense:

$$\omega = \nabla \times \boldsymbol{u} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \partial_x & \partial_y & \partial_z \\ u_1 & u_2 & u_3 \end{vmatrix}$$

Prove that in a two-dimensional fluid, where

$$\boldsymbol{u}(x, y, z) = (u_1(x, y), u_2(x, y), 0)$$

the vorticity is given by

$$\boldsymbol{\omega}(x, y) = \omega(x, y)\hat{\boldsymbol{z}}, \qquad \omega(x, y) = \partial_x u_2 - \partial_y u_1.$$

We have

$$
\begin{aligned}
\boldsymbol{\omega} &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \partial_x & \partial_y & \partial_z \\ u_1(x,y) & u_2(x,y) & 0 \end{vmatrix}, \\
&= \hat{\boldsymbol{x}}\left(\partial_y 0 - \partial_z u_2(x,y)\right) - \hat{\boldsymbol{y}}\left(\partial_x 0 - \partial_z u_1(x,y)\right) + \hat{\boldsymbol{z}}\left(\partial_x u_2 - \partial_y u_1\right) \\
&= \hat{\boldsymbol{z}}\left(\partial_x u_2 - \partial_y u_1\right),
\end{aligned}
$$

as required. The vorticity has magnitude $\omega = \partial_x u_2 - \partial_y u_1$ and points in the z-direction.

Suppose that the two-dimensional fluid is incompressible if $\nabla \cdot \boldsymbol{u} = 0$, i.e.

$$\partial_x u_1 + \partial_y u_2 = 0.$$

Prove that the necessary and sufficient condition for the fluid to be incompressible is the existence of a **streamfunction** $\psi(x,y)$, such that

$$u_1 = \partial_y \psi, \qquad u_2 = -\partial_x \psi.$$

Necessity: Assume that the flow is incompressible, $\partial_x u_1 + \partial_y u_2 = 0$. We show that a streamfunction exists by construction: Let

$$\psi(x,y) = \int_a^y u_1(x,\lambda)\,\mathrm{d}\lambda - \int_b^x u_2(\mu,a)\mathrm{d}\mu,$$

where a is an arbitrary y-value in the domain of the fluid and λ is a dummy variable of integration. Similarly, b is an arbitrary x-value and μ is a dummy variable. By construction, and by the Fundamental Theorem of Calculus,

$$\frac{\partial \psi}{\partial y} = u_1(x,y).$$

Now

$$
\begin{aligned}
\frac{\partial \psi}{\partial x} &= \frac{\partial}{\partial x}\int_a^y u_1(x,\lambda)\,\mathrm{d}\lambda - u_2(x,a), \\
&= \int_a^y \frac{\partial u_1}{\partial x}(x,\lambda)\,\mathrm{d}\lambda - u_2(x,a), \\
&= -\int_a^y \frac{\partial u_2}{\partial \lambda}(x,\lambda)\,\mathrm{d}\lambda - u_2(x,a), \qquad \text{(By incompressibility)} \\
&= -\left[u_2(x,y) - u_2(x,a)\right] - u_2(x,a), \\
&= -u_2(x,y).
\end{aligned}
$$

Hence, ψ is a streamfunction because the flow $(u_1(x,y), u_2(x,y))$ can be derived from it.

Sufficiency: Assume that the streamfunction exists. Then

$$u_1 = \frac{\partial \psi}{\partial y},\ u_2 = -\frac{\partial \psi}{\partial x},$$

and

$$\begin{aligned}
\operatorname{div} \boldsymbol{u} &= \frac{\partial}{\partial x}\frac{\partial \psi}{\partial y} + \frac{\partial}{\partial y}\left(-\frac{\partial \psi}{\partial x}\right), \\
&= \frac{\partial}{\partial x}\frac{\partial \psi}{\partial y} - \frac{\partial}{\partial y}\frac{\partial \psi}{\partial x}.
\end{aligned}$$

Using the fact that the partial derivatives of smooth functions commute, $\partial_{xy}\psi = \partial_{yx}\psi$, we obtain

$$\operatorname{div} \boldsymbol{u} = 0.$$

Thus, a two-dimensional flow is incompressible if and only if it has a streamfunction.

Prove that

$$\nabla^2 \psi = -\omega.$$

Hence, demonstrate that in an irrotational fluid, $\nabla^2 \psi = 0$.

We have

$$\begin{aligned}
\omega &= \partial_x u_2 - \partial_y u_1, \\
&= \partial_x \left(-\partial_x \psi\right) - \partial_y \left(+\partial_y \psi\right), \\
&= -\left(\partial_x^2 + \partial_y^2\right)\psi,
\end{aligned}$$

and $\omega = -\nabla^2 \psi$. If the flow is irrotational, its curl is zero, and $\omega = 0$. Hence, in an irrotational flow, $\nabla^2 \psi = 0$. No streamfunction exists in three-dimensional flows and this simple analysis no longer holds there.

Chapter 11

Line integrals

Overview

Logically, the next step after differentiating vector and scalar fields is to integrate them. We start with **line integrals**. These are the easiest because they are most like ordinary single-variable integration.

11.1 The definition

Formally, we have the following small increment of displacement:

$$\mathrm{d}\boldsymbol{x} = \hat{\boldsymbol{x}}\mathrm{d}x + \hat{\boldsymbol{y}}\mathrm{d}y + \hat{\boldsymbol{z}}\mathrm{d}z,$$

which gives rise to the following possible integrals for a scalar field $\phi(x, y, z)$ and a vector field $\boldsymbol{v}(x, y, z)$:

$$\int_C \phi(x, y, z)\mathrm{d}\boldsymbol{x},$$
$$\int_C \boldsymbol{v}(x, y, z) \cdot \mathrm{d}\boldsymbol{x},$$
$$\int_C \boldsymbol{v}(x, y, z) \times \mathrm{d}\boldsymbol{x},$$

where C denotes a **contour**, that is, a curve $\boldsymbol{x}_C(t) : \mathbb{R} \to \mathbb{R}^3$. Let us work with the first kind of integral and introduce the formal definition (generalizing to the other integrals will be left to exercises and examples).

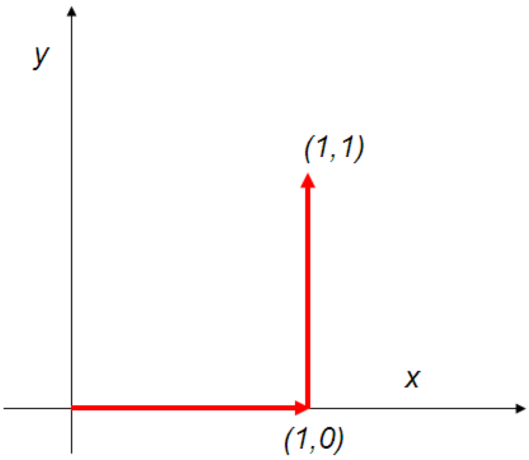
Definition 11.1 (Line integrals) *Let $\boldsymbol{x}_C : [t_1, t_2] \to \mathbb{R}^3$ be some* **piecewise smooth** *curve. Then the line integral $\int_C \phi(x, y, z)\mathrm{d}\boldsymbol{x}$ along the curve C is defined as follows:*

$$\begin{aligned}\int_C \phi(x,y,z)\mathrm{d}\boldsymbol{x} &:= \int_{t_1}^{t_2} \phi\left(\boldsymbol{x}_C(t)\right)\frac{d\boldsymbol{x}_C}{dt}\mathrm{d}t,\\ &= \hat{\boldsymbol{x}}\int_{t_1}^{t_2} \phi\left(\boldsymbol{x}_C(t)\right)\frac{dx_C}{dt}\mathrm{d}t + \hat{\boldsymbol{y}}\int_{t_1}^{t_2} \phi\left(\boldsymbol{x}_C(t)\right)\frac{dy_C}{dt}\mathrm{d}t + \hat{\boldsymbol{z}}\int_{t_1}^{t_2} \phi\left(\boldsymbol{x}_C(t)\right)\frac{dz_C}{dt}\mathrm{d}t\end{aligned}$$

11.2 Worked examples

Example: Let $\phi(x, y) = x^2 + y^2$. Compute the line integral along the curve

$$\begin{aligned} C_1 &: (0,0) \to (1,0)\,, &&\text{in a straight line,}\\ C_2 &: (1,0) \to (1,1)\,, &&\text{in a straight line.}\end{aligned}$$



The path $(0,0) \to (1,0) \to (1,1)$.

Solution: Break up the integration into two parts. In the first part, the curve is

$$\boldsymbol{x}_{C1}(t) = (t, 0)\,, \qquad t \in [0,1]\,, \qquad \frac{d\boldsymbol{x}_{C1}}{dt} = (1,0)\,.$$

Hence,

$$\int_{C1} \phi(x,y,z)\mathrm{d}\boldsymbol{x} = \hat{\boldsymbol{x}}\int_0^1 t^2\mathrm{d}t + \hat{\boldsymbol{y}}\int_0^1 0\,\mathrm{d}t = \tfrac{1}{3}\hat{\boldsymbol{x}}.$$

In the second part, the curve is

$$\boldsymbol{x}_{C2}(t) = (1, t)\,, \qquad t \in [0,1]\,, \qquad \frac{d\boldsymbol{x}_{C1}}{dt} = (0,1)\,.$$

Hence,

$$\int_{C2} \phi(x,y,z)\mathrm{d}\boldsymbol{x} = \hat{\boldsymbol{y}}\int_0^1 (1+t^2)\,\mathrm{d}t = \tfrac{4}{3}\hat{\boldsymbol{y}}.$$

Putting them together,

$$\int_C \phi(x,y,z)\mathrm{d}\boldsymbol{x} = \int_{C1} \phi(x,y,z)\mathrm{d}\boldsymbol{x} + \int_{C2} \phi(x,y,z)\mathrm{d}\boldsymbol{x} = \tfrac{1}{3}\hat{\boldsymbol{x}} + \tfrac{4}{3}\hat{\boldsymbol{y}}.$$

The most common line integrals in physics are of the form $\int_C \boldsymbol{v}\cdot \mathrm{d}\boldsymbol{x}$, as in the following example:

> Consider a vector field
>
> $$\boldsymbol{v} = 3xy\hat{\boldsymbol{x}} - y^2\hat{\boldsymbol{y}}$$
>
> integrated along the curve C defined by $y = 2x^2$, from the origin $(0,0)$ to the point $(1,2)$. Evaluate $\int_C \boldsymbol{v}\cdot \mathrm{d}\boldsymbol{x}$.

Solution: The curve has the parametric form

$$\boldsymbol{x}_C(t) = \left(t, 2t^2\right), \qquad t\in[0,1] \qquad \frac{d\boldsymbol{x}_C}{dt} = (1,4t)\,.$$

We compute

$$\begin{aligned}
\int_C \boldsymbol{v}\cdot \mathrm{d}\boldsymbol{x} &= \int_C \left[v_1(x,y)\mathrm{d}x + v_2(x,y)\mathrm{d}y\right], \\
&= \int_C v_1(\boldsymbol{x}_C(t))\frac{\mathrm{d}x_C}{dt}\mathrm{d}t + \int_C v_2(\boldsymbol{x}_C(t))\frac{\mathrm{d}y_C}{dt}\mathrm{d}t, \\
&= \int_0^1 3(t)(2t^2)(1)\,\mathrm{d}t + \int_0^1 [-(2t^2)^2](4t)\,\mathrm{d}t, \\
&= \int_0^1 6t^3\,\mathrm{d}t - \int_0^1 16t^5\,\mathrm{d}t, \\
&= -\tfrac{7}{6}.
\end{aligned}$$

In mechanics, there is the notion of force. Suffice to say, force is a vector field in two or three dimensions, $\boldsymbol{F}(x,y)$, or $\boldsymbol{F}(x,y,z)$. The **work done**, W, as a particle is moved along a trajectory $\boldsymbol{x}_C(t)$ through the force field $\boldsymbol{F}(\boldsymbol{x})$ is the line integral of the force field along the trajectory:

$$W = \int_C \boldsymbol{F}(\boldsymbol{x}) \cdot \mathrm{d}\boldsymbol{x}.$$

Example: Consider a force

$$\boldsymbol{F} = -k\left(\hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y\right).$$

Compare the work done moving against this force field when going from $(1,1)$ to $(4,4)$ along the following straight-line paths:

$$\begin{aligned}
(1,1) &\rightarrow (4,1) \rightarrow (4,4), \\
(1,1) &\rightarrow (1,4) \rightarrow (4,4), \\
(1,1) &\rightarrow (4,4), \text{ along } x = y.
\end{aligned}$$

Solution: Consider for example the third path. The curve is

$$\boldsymbol{x}_C(t) = (1+t, 1+t), \qquad t \in [0,3], \qquad \frac{d\boldsymbol{x}_C}{dt} = (1,1),$$

and the integral is

$$\begin{aligned}
\int_C \boldsymbol{F} \cdot \mathrm{d}\boldsymbol{x} &= \int_C \left[F_1(x,y)\mathrm{d}x + F_2(x,y)\mathrm{d}y\right], \\
&= \int_0^3 F_1(\boldsymbol{x}_C(t))\frac{dx_C}{dt}\mathrm{d}t + \int_0^3 F_2(\boldsymbol{x}_C(t))\frac{dy_C}{dt}\mathrm{d}t, \\
&= -2k\int_0^3 (1+t)\mathrm{d}t, \\
&= -15k.
\end{aligned}$$

The other two cases are left as an exercise but you should get the same answer in all three cases. Here is why. The force field $\boldsymbol{F} = -k(\hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y)$ can be written as

$$\boldsymbol{F} = -k\left(\hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y\right) = -\nabla\left[\frac{k}{2}\left(x^2+y^2\right)\right] := -\nabla\mathcal{U}(x,y).$$

Thus, along any path C,

$$\begin{aligned}\int_C \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{x} &= -\int_C \nabla\mathcal{U}\cdot \mathrm{d}\boldsymbol{x},\\ &= -\int_{t_1}^{t_2} \nabla\mathcal{U}\left(\boldsymbol{x}_C(t)\right)\cdot \frac{\mathrm{d}\boldsymbol{x}_C}{dt}\mathrm{d}t,\end{aligned}$$

$$\begin{aligned}\int_C \boldsymbol{F}\cdot \mathrm{d}\boldsymbol{x} &= -\int_{t_1}^{t_2} \frac{d}{dt}\mathcal{U}\left(\boldsymbol{x}_C(t)\right)\mathrm{d}t,\\ &= -\left[\mathcal{U}(\boldsymbol{x}_C(t_2)) - \mathcal{U}(\boldsymbol{x}_C(t_1))\right],\end{aligned}$$

and the line integral is independent of the path and depends only on the initial and final points. Recall from previous lectures that a vector field $\boldsymbol{F}$ is irrotational if and only if it can be written in the form $\boldsymbol{F} = -\nabla\mathcal{U}$. Thus, we have the following string of statements:

A vector field $\boldsymbol{F}$ is irrotational if and only if

- $\nabla\times\boldsymbol{F} = 0$ if and only if
- $\boldsymbol{F} = -\nabla\mathcal{U}$ if and only if
- The line integral $\int_C \boldsymbol{F}\cdot\mathrm{d}\boldsymbol{x}$ depends only on the initial and final points of the path C and is independent of the details of the path between these terminal points.

Consider also a closed path C, for which $\boldsymbol{x}_C(t_2) = \boldsymbol{x}_C(t_1)$. For an irrotational vector field $\boldsymbol{u}(\boldsymbol{x})$ integrated over such a path,

$$\begin{aligned}\int_C \boldsymbol{u}(\boldsymbol{x})\cdot\mathrm{d}\boldsymbol{x} &:= \oint_C \boldsymbol{u}(\boldsymbol{x})\cdot\mathrm{d}\boldsymbol{x}\\ &= -\int_C \nabla\mathcal{U}\cdot\mathrm{d}\boldsymbol{x}\\ &= -\left[\mathcal{U}(\boldsymbol{x}_C(t_2)) - \mathcal{U}(\boldsymbol{x}_C(t_1))\right]\\ &= -\left[\mathcal{U}(\boldsymbol{x}_C(t_2)) - \mathcal{U}(\boldsymbol{x}_C(t_2))\right]\\ &= 0.\end{aligned}$$

Example: In contrast, consider the force

$$\boldsymbol{G} = -k\left(\hat{\boldsymbol{x}}y - \hat{\boldsymbol{y}}x\right).$$

Compute the curl of the force. Show that the work done is path-dependent.

Solution: The curl of the force is

$$\nabla \times \boldsymbol{G} = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ \partial_x & \partial_y & \partial_z \\ -ky & kx & 0 \end{vmatrix} = 2k\hat{\boldsymbol{z}}.$$

Let's integrate along the paths

$$\begin{aligned} C &: (1,1) \to (4,1) \to (4,4), \\ D &: (1,1) \to (1,4) \to (4,4). \end{aligned}$$

First path, first component:

$$\boldsymbol{x}_{C1}(t) = (1+t, 1), \qquad t \in [0,3], \qquad \frac{d\boldsymbol{x}_C}{dt} = (1,0),$$

and

$$\boldsymbol{G} \cdot \mathrm{d}\boldsymbol{x} = \boldsymbol{G} \cdot \frac{d\boldsymbol{x}_C}{dt}\mathrm{d}t = G_x \mathrm{d}t = -ky\mathrm{d}t = -k\mathrm{d}t.$$

Integrating gives $-3k$. First path, second component:

$$\boldsymbol{x}_{C2}(t) = (4, 1+t), \qquad t \in [0,3], \qquad \frac{d\boldsymbol{x}_C}{dt} = (0,1),$$

and

$$\boldsymbol{G} \cdot \mathrm{d}\boldsymbol{x} = \boldsymbol{G} \cdot \frac{d\boldsymbol{x}_C}{dt}\mathrm{d}t = G_y \mathrm{d}t = +kx\mathrm{d}t = 4k\mathrm{d}t.$$

Integrating gives $12k$. Adding up gives

$$\int_C \boldsymbol{G} \cdot \mathrm{d}\boldsymbol{x} = 9k.$$

Second path, first component:

$$\boldsymbol{x}_{D1}(t) = (1, 1+t), \qquad t \in [0,3], \qquad \frac{d\boldsymbol{x}_C}{dt} = (0,1),$$

and

$$\boldsymbol{G}\cdot\mathrm{d}\boldsymbol{x}=\boldsymbol{G}\cdot\frac{d\boldsymbol{x}_D}{dt}\mathrm{d}t=G_y\mathrm{d}t=+kx\mathrm{d}t=k\mathrm{d}t.$$

Integrating gives $+3k$. Second path, second component:

$$\boldsymbol{x}_{D2}(t)=(1+t,4)\,,\qquad t\in[0,3],\qquad \frac{d\boldsymbol{x}_C}{dt}=(1,0),$$

and

$$\boldsymbol{G}\cdot\mathrm{d}\boldsymbol{x}=\boldsymbol{G}\cdot\frac{d\boldsymbol{x}_D}{dt}\mathrm{d}t=G_x\mathrm{d}t=-ky\mathrm{d}t=-4k\mathrm{d}t.$$

Integrating gives $-12k$. Adding up gives

$$\int_C\boldsymbol{G}\cdot\mathrm{d}\boldsymbol{x}=-9k.$$

and the two paths differ.

Chapter 12

Theory of integration

Overview

We review the theory of Riemann integration in one dimension. This enables us to extend integration theory to higher dimensions. We will then have the theoretical background to be able to do practical integration calculations in subsequent chapters.

12.1 Riemann integration in one dimension – review – or crash course?

We introduce the notion of the **Riemann integral** for a generic continuous function

$$\begin{aligned} f : [a,b] &\to \mathbb{R}, \\ x &\mapsto f(x), \end{aligned}$$

with the positivity property

$$f(x) > 0 \qquad \text{on } [a,b].$$

We consider the problem of finding the area under the curve $y = f(x)$.

We form a uniform partition of $[a,b]$,

$$[a,b] = [x_0, x_1] \cup [x_1, x_2] \cup \cdots \cup [x_{N-1}, x_N], \qquad a = x_0, \qquad b = x_N,$$
$$x_i = a + i\left(\frac{b-a}{N}\right), \; i = 0, 1, \cdots, N.$$

We form the following **upper and lower sums**:

$$\text{Upper}: \qquad \mathcal{U}_N := \sum_{i=0}^{N-1}\left[\sup_{x\in[x_i,x_{i+1}]} f(x)\right](x_{i+1}-x_i)\,,$$
$$\text{Lower}: \qquad \mathcal{L}_N := \sum_{i=0}^{N-1}\left[\inf_{x\in[x_i,x_{i+1}]} f(x)\right](x_{i+1}-x_i)\,.$$

Definition 12.1 *The function f is called* **Riemann-integrable** *with Riemann integral I if*

$$\lim_{N\to\infty}\mathcal{U}_N = \lim_{N\to\infty}\mathcal{L}_N := I. \tag{12.1}$$

See Figure 12.5 for the accompanying sketch.

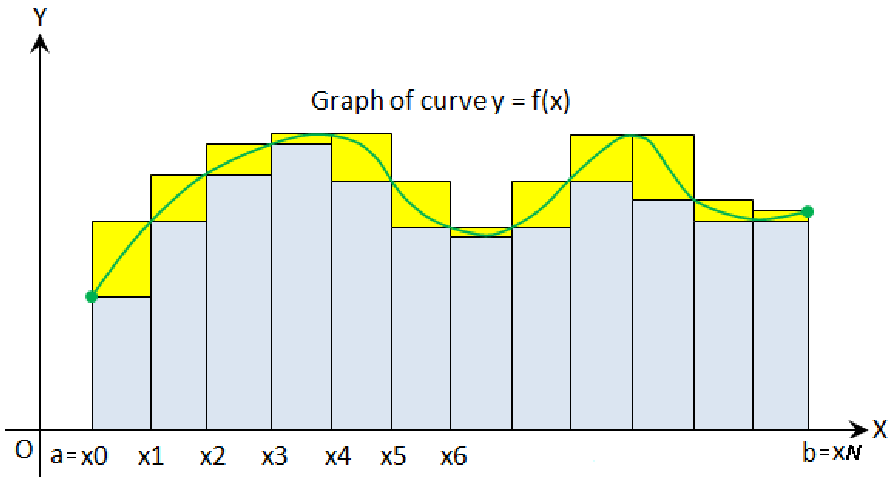


Figure 12.1: Upper and lower sums for Riemann integration

We have the following important result:

Theorem 12.1 (Riemann integrability for continuous functions) *Let $f:[a,b]\to\mathbb{R}$ be continuous on $[a,b]$ with the positivity property $f(x)>0$. Then f is Riemann integrable.*

No proof is provided here – this is not an analysis class, but see Beals (p. 107)

A further result enables an obvious generalization of Riemann integration to non-positive functions:

Theorem 12.2 *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ with the positivity property $f(x) > 0$, and let $c \in (a,b)$. Then*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

Hence, we have the following definition:

Definition 12.2 *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ with $c \in (a,b)$ such that $f(x) \le 0$ on $[a,c]$ and $f(x) \ge 0$ on $[c,b]$. Then*

$$\int_a^b f(x)\,\mathrm{d}x := -\int_a^c [-f(x)]\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x. \tag{12.2}$$

Definition (12.2) is extended in an obvious way to continuous functions with many zeros on a given interval of integration. Moreover, the Riemann-sum definition (12.1) can be applied to non-positive functions immediately without the intermediate step in Equation (12.2).

In MATH10350 you will have seen the statement of the Fundamental Theorem of Calculus. This can be connected to Riemann integration as follows. First, let $f(x)$ be a continuous function on an interval $[a,b]$. Form the Riemann integral

$$F(x) = \int_a^x f(x)\,\mathrm{d}x, \qquad x \in (a,b). \tag{12.3}$$

Then,

Theorem 12.3 *The function $F(x)$ in Equation* (12.3) *is differentiable with continuous derivative and moreover,*

$$\frac{\mathrm{d}F}{\mathrm{d}x} = f(x), \qquad x \in (a,b).$$

This establishes the remarkable – and when you think of it, utterly surprising – connection between Rieann integration and calculus. This theorem is discussed and proved in Beals (p. 111).

12.2 Multivariable integration – heuristic

We explain and develop the theory of multivariable integration by a heuristic approach. This is a great word! Wikipedia defines it to mean

> A heuristic technique, often called simply a heuristic, is any approach to problem solving, learning, or discovery that employs a practical methodology not guaranteed to be optimal or perfect, but sufficient for the immediate goals.

This is perfect for us - we want a shortcut to the general tehory of multivariable integration, mindful that a full rigorous theory should be left to the pure mathematicans – and perhaps a module in Measure Theory such as MATH 40430 (Measure Theory & Integration).

We therefore consider the problem of finding the area of a right-angled triangle (Fig. 12.2). The

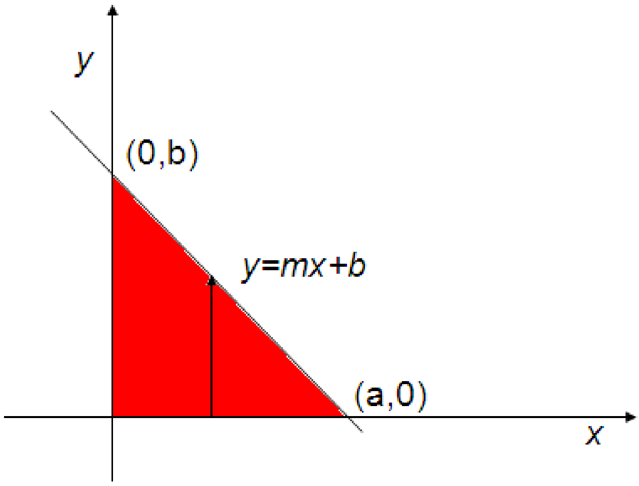


Figure 12.2: Right-angled triangle with vertices at $(0,0)$, $(a,0)$, and $(0,b)$.

first goal here is to compute from first-principles the area of the triangle using Riemann integration, using two separate approaches. We shall show that these approaches give the same answer. This is an illustration of a general principle called **Fubini's theorem**, which we state at towards the end of this chapter.

There are at least two ways of calculating the area of the triangle in Figure 12.2 using Riemann integration. The first involves **a sum over boxes**, whereby we break up the triangle into small boxes and sum over all such boxes. To do this, we fit rows of boxes into the triangle, where each row is parallel to the x-axis. Each box has sides of length Δx (See Figure 12.3).

- First row of boxes: N_1 boxes fit into the first row, with $\Delta y = m(N_1\Delta x) + b$, hence

$$N_1 = \frac{\Delta y - b}{m\Delta x}.$$

- N_2 boxes are placed into the second row, with

$$N_2 = \frac{2\Delta y - b}{m\Delta x}.$$

- One continues thus until the last row is reached, in which precisely one box fits. This is the

N_y^{th} row, and $N_y \Delta y = m\Delta x + b$, hence

$$N_y = \frac{m\Delta x + b}{\Delta y}.$$

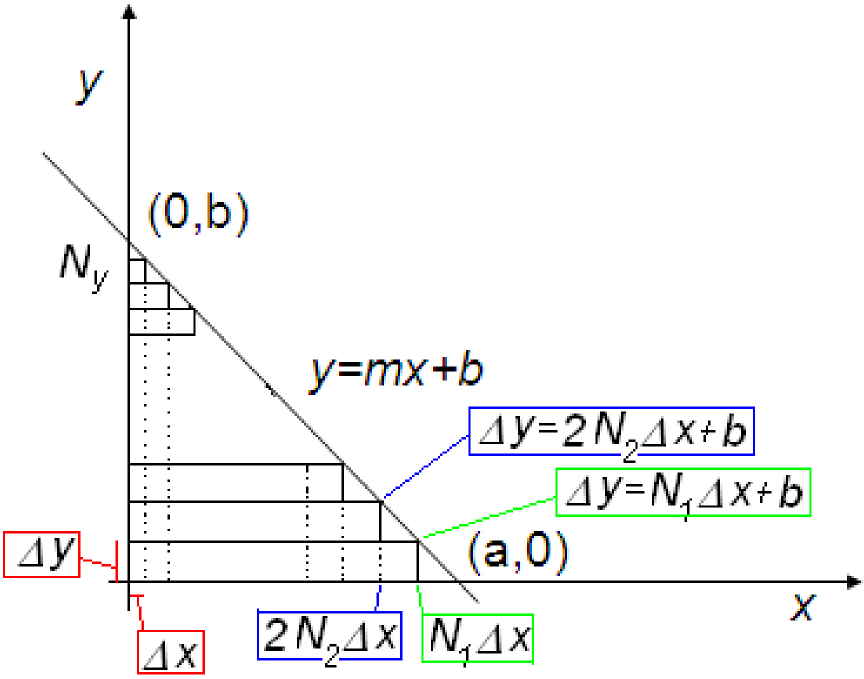


Figure 12.3: The area of a triangle computed as a Riemann sum, where each summand is a small box.

We sum of the total area of the boxes:

$$\begin{aligned}
\text{Area} &= \Delta x \Delta y N_1 + \Delta x \Delta y N_2 + \cdots + \Delta x \Delta y, \\
&= \Delta x \Delta y \left(\frac{\Delta y - b}{m\Delta x}\right) + \Delta x \Delta y \left(\frac{2\Delta y - b}{m\Delta x}\right) + \cdots + \Delta x \Delta y, \\
&= \Delta y \left(\frac{\Delta y - b}{m}\right) + \Delta y \left(\frac{2\Delta y - b}{m}\right) + \cdots + \Delta x \Delta y, \\
&= \frac{\Delta y^2}{m} \sum_{j=1}^{N_y} j - \frac{\Delta y}{m} \sum_{j=1}^{n} b, \\
&= \frac{\Delta y^2}{m} \tfrac{1}{2} N_y (N_y + 1) - \frac{b\Delta y}{m} N_y,
\end{aligned}$$

Now use the formula for N_y:

$$\begin{aligned}
\text{Area} &= \tfrac{1}{2}\frac{\Delta y^2}{m}\left(\frac{m\Delta x+b}{\Delta y}\right)\left(\frac{m\Delta x+b}{\Delta y}+1\right)-\frac{b\Delta y}{m}\left(\frac{m\Delta x+b}{\Delta y}\right),\\
&= \frac{1}{2m}\left(m\Delta x+b\right)\left(m\Delta x+b+\Delta y\right)-\frac{b}{m}(m\Delta x+b),\\
&\overset{\Delta x,\Delta y\to 0}{=} \frac{b^2}{2m}-\frac{b^2}{m},\\
&= -\frac{b^2}{2m}.
\end{aligned}$$

Using $m=-b/a$, this is

$$\text{Area} = \tfrac{1}{2}ab.$$

Another way of calculating the area of the triangle involves **a sum over strips**: we break up the triangle into small vertical strips and sum over all such strips. There are N strips of width Δx, hence $(N+1)\Delta x=1$, and $N=(a/\Delta x)-1$ (Figure 12.4). The height of the j^{th} strip is

$$y_j = mx_j + b = m(j\Delta x)+b.$$

We sum over the area of each strip as follows:

$$\begin{aligned}
\text{Area} &= \sum_{j=1}^{N} y_j\Delta x,\\
&= \sum_{j=1}^{N}(m\Delta x j+b)\Delta x,\\
&= m\Delta x^2\sum_{j=1}^{N} j + b\Delta x\sum_{j=1}^{N}(1),\\
&= \tfrac{1}{2}m\Delta x^2 N(N+1)+b\Delta x N,\\
&= \tfrac{1}{2}m\Delta x^2\left(\frac{a}{\Delta x}\right)\left(\frac{a}{\Delta x}-1\right)+b\Delta x\left(\frac{a}{\Delta x}-1\right),\\
&= \tfrac{1}{2}ma\left(a-\Delta x\right)+(a-\Delta x)b,\\
&\overset{\Delta x\to 0}{=} \tfrac{1}{2}ma^2+ab,\\
&= \tfrac{1}{2}\left(-\frac{b}{a}\right)a^2+ab,\\
&= \tfrac{1}{2}ab.
\end{aligned}$$

Both methods give the same answer. Indeed, we could have computed the area strip-wise by using strips *parallel* to the x-axis, rather than perpendicular, and we would still get the same answer. One can think of this result as being the equivalence of the 'limit over strips' and 'the limit over squares'.

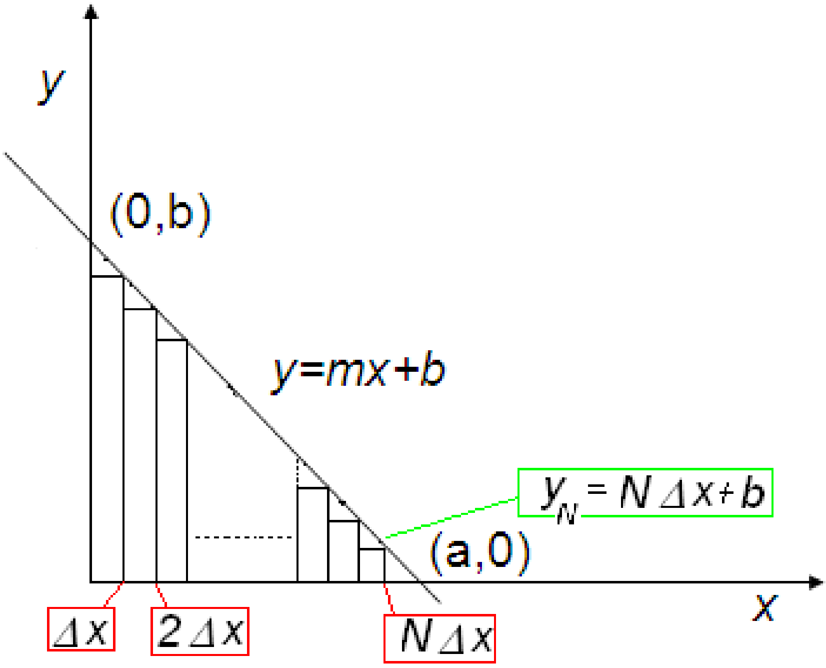


Figure 12.4: The area of a triangle computed as a Riemann sum, where each summand is a small strip.

This fact fits into a much more general result called **Fubini's theorem**

12.3 Fubini's theorem for areas

We now consider the problem of Riemann integration in two dimensions for a more general domain. Let Ω be a region in $\mathbb{R}^2$ whose boundary is a piecewise continuous closed curve (Figure 12.5). The goal here is to compute the area of the region Ω. We start with a **sum over boxes**: we fit a grid of small boxes into the region Ω, of sides of length Δx and Δy. The i^{th} box has its centre at $\boldsymbol{x}_i$. Those sections of Ω not covered by the grid become vanishingly small as Δx and Δy are reduced indefinitely. We therefore have

$$\text{area}(\Omega) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \sum_{\boldsymbol{x}_i \in \Omega} \Delta x \Delta y.$$

We interpret this integral as the double integral over Ω:

$$\int\int_{\Omega} \mathrm{d}x\, \mathrm{d}y \stackrel{\text{def}}{=} \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \sum_{\boldsymbol{x}_i \in \Omega} \Delta x \Delta y;$$

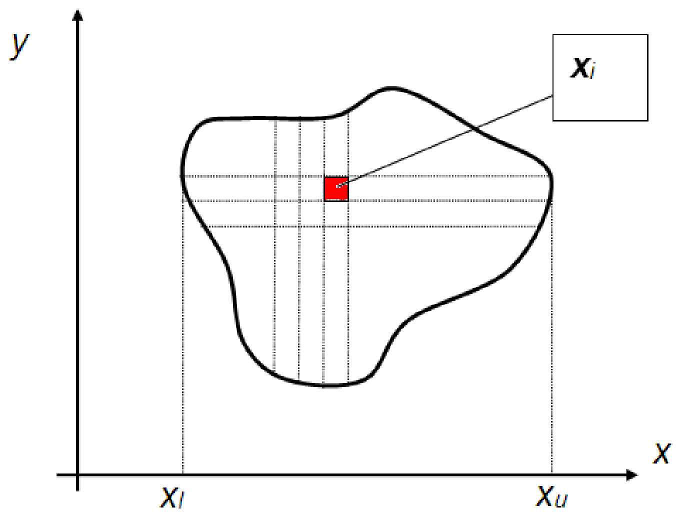


Figure 12.5: Riemann integration in $\mathbb{R}^2$ as a 'limit over boxes'. Each small box has sides of length Δx and Δy.

for sets Ω with piecewise continuous boundaries, the limit can be shown to exist and from Figure 12.5, it follows that the double integral is equal to the area of the region Ω.

However, the area of Ω can also be computed as a 'limit over strips'. Consider Figure 12.6. In this

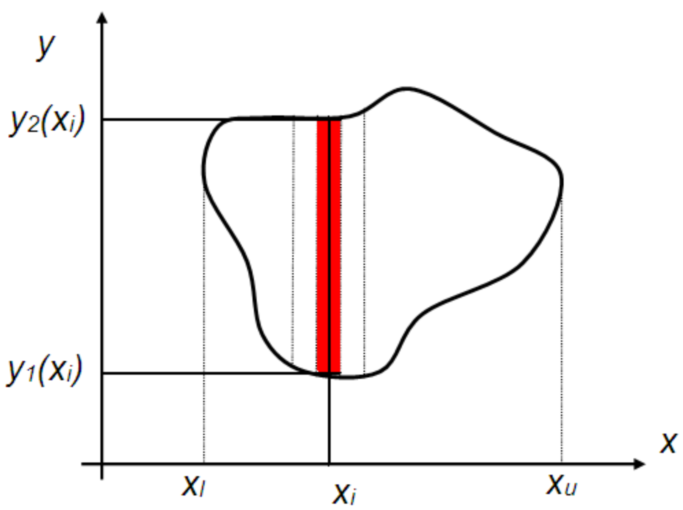


Figure 12.6: Riemann integration in $\mathbb{R}^2$ as a 'limit over strips'. Each small strip has width Δx.

figure, the region Ω is filled in with narrow strips of width Δx. The strips are centred at x_i, with $i = 1, 2, \cdots$. Also, for each x-value, the there are two y-boundary points, denoted by $y_1(x)$ and

$y_2(x)$. Thus,

$$\text{area}(\Omega) = \lim_{\Delta x \to 0} \Delta x \sum_{x_i} \left[y_2(x_i) - y_1(x_i) \right]. \tag{12.4}$$

But

$$y_2(x_i) - y_1(x_i) = \int_{y_1(x_i)}^{y_2(x_i)} \mathrm{d}y.$$

Define

$$F(x) := \int_{y_1(x)}^{y_2(x)} \mathrm{d}y.$$

Thus, Equation (12.4) becomes

$$\text{area}(\Omega) = \lim_{\Delta x \to 0} \sum_{x_i} \Delta x F(x_i).$$

But this is an ordinary (one-dimensional) Riemann integral:

$$\text{area}(\Omega) = \int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \, F(x).$$

One restores the definition of $F(x)$ to obtain

$$\text{area}(\Omega) = \int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \left[\int_{y_1(x)}^{y_2(x)} \mathrm{d}y \right].$$

But

$$\text{area}(\Omega) = \int \int_{\Omega} \mathrm{d}x \, \mathrm{d}y,$$

and we therefore have Fubini's theorem for areas, in the general case:

$$\int \int_{\Omega} \mathrm{d}x \, \mathrm{d}y = \int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \left[\int_{y_1(x)}^{y_2(x)} \mathrm{d}y \right].$$

In other words, the difficult and slightly weird double integral on the left-hand side can be converted into a repeated application of single integrals on the right-hand side, albeit that the limits of integration on the innermost integral are not necessarily constants.

Finally, there is nothing special about strips parallel to the y-axis, so we have, quite generally,

$$\int \int_{\Omega} \mathrm{d}x \, \mathrm{d}y = \int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \left[\int_{y_1(x)}^{y_2(x)} \mathrm{d}y \right] = \int_{y_\mathrm{l}}^{y_\mathrm{u}} \mathrm{d}y \left[\int_{x_1(y)}^{x_2(y)} \mathrm{d}x \right]. \tag{12.5}$$

Here, y_l and y_u refer to the (constant) extent of the domain in the y-direction, and $x_1(y)$ and $x_2(y)$ refer to the boundary curves, where the x-coordinates are viewed as parametric functions of

y. Thus, Fubini's theorem is sometimes restated as the fact that 'the order in which the integration is performed does not matter'.

In Fubini's theorem, the parametric boundary curves $y_1(x)$ and $y_2(x)$ (and their inverses) refer to the upper and lower segments of the boundary in Figure 12.5. Of course, it is possible for certain ranges of x, that three or more parametric paths are required to describe the boundary (draw this!), but this does not affect any of the arguments given in this chapter.

12.4 Integration of functions of several variables

Of course, not only do we wish to compute areas, but also n-fold integrals of functions. Focusing on the two-dimensional case, we wish to define and compute integrals such as

$$\mathcal{I} = \int\int_{\Omega} \mathrm{d}x\,\mathrm{d}y\, f(x,y),$$

where $f(x,y)$ is a continuous function and Ω is a set whose boundary is a closed, piecewise continuous curve. Following on from the previous section, this is naturally defined to be

$$\int\int_{\Omega} \mathrm{d}x\,\mathrm{d}y\, f(x,y) \stackrel{\text{def}}{=} \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \sum_{\boldsymbol{x}_i\in\Omega} f(\boldsymbol{x}_i)\Delta x\Delta y; \tag{12.6}$$

the main purpose of the analytic theory of Riemann integration is to show that this limit exists for continuous functions and for sets whose boundaries are closed and piecewise continuous; the existence of this limit is taken as given in this module (this is not an analysis course).

As in the previous section, we note that the sum in Equation (12.6) can be computed in many ways. We therefore choose to compute the sum in the following way. We identify **columns** of cells as those at a fixed x-value (see Figure 12.7). Then,

- Run over all columns, $I = 1, 2, \cdots$.
- When at the I^{th} column, run over all cells in that column. The cells in this one given column are given a label J. So, for a fixed I, there are many J-values, with $J = 1, 2, \cdots, N_I$. The number N_I depends on the column number: near x_{l} and x_{u} the number N_I is relatively small while closer to the middle of the allowed range of x-values there are relatively many cells in the columns, and N_I is relatively large.

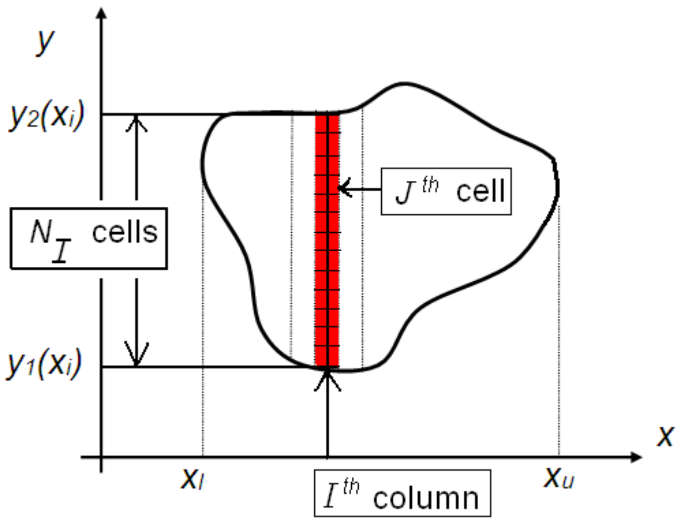


Figure 12.7: Riemann integration for a function $f(x,y)$ $\mathbb{R}^2$ as a 'limit over strips'. As before, each small strip has width Δx.

Thus, we have

$$
\begin{aligned}
\mathcal{I} &= \int\int_\Omega \mathrm{d}x\,\mathrm{d}y\, f(x,y), \\
&= \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \sum_{\boldsymbol{x}_i\in\Omega} f(\boldsymbol{x}_i)\Delta x\Delta y, \\
&= \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \sum_I \Delta x \sum_{J=1}^{N_I} \Delta y\, f(x_I, y_J). \qquad (12.7)
\end{aligned}
$$

If the double limit can be performed sequentially (and it can, but this is beyond the scope of this module, and is the crux of Fubini's theorem), then we can re-write Equation (12.7) as

$$
\mathcal{I} = \lim_{\Delta x\to 0} \sum_I \Delta x \left[\lim_{\Delta y\to 0} \sum_{J=1}^{N_I} \Delta y\, f(x_I, y_J) \right], \qquad (12.8)
$$

and we identify the limit inside the square brackets as

$$
\lim_{\Delta y\to 0} \sum_{J=1}^{N_I} \Delta y\, f(x_I, y_J) = \int_{y_1(x_I)}^{y_2(x_I)} f(x_I, y)\,\mathrm{d}y,
$$

which is an ordinary single integral with respect to y carried out at fixed $x = x_I$. Thus, Equation (12.8) becomes

$$
\mathcal{I} = \lim_{\Delta x\to 0} \sum_I \Delta x \left[\int_{y_1(x_I)}^{y_2(x_I)} f(x_I, y)\,\mathrm{d}y \right],
$$

and the outermost limit (as $\Delta x \to 0$) converts into an ordinary single integral with respect to x:

$$\mathcal{I} = \int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \left[\int_{y_1(x)}^{y_2(x)} f(x,y)\,\mathrm{d}y \right],$$

and Fubini's theorem for the double integral of functions (i.e. not just areas) is recoverd: multiple integration amounts to repeated applications of ordinary (single-variable) integration. For the sake of simplifiying notation, we sometimes omit the parametric dependence of the y-limits on x, and write

$$\int_{x_\mathrm{l}}^{x_\mathrm{u}} \mathrm{d}x \left[\int_{y_1(x)}^{y_2(x)} f(x,y)\,\mathrm{d}y \right] \equiv \int_\Omega \mathrm{d}x \left[\int \mathrm{d}y\, f(x,y) \right]$$

Again, there is nothing special about the column treatment in Figure 12.7: one could just as easily sum over rows of cells, leading to the following result:

$$\mathcal{I} = \int_{y_\mathrm{l}}^{y_\mathrm{u}} \mathrm{d}y \left[\int_{x_1(y)}^{x_2(y)} f(x,y)\,\mathrm{d}x \right] \equiv \int_\Omega \mathrm{d}y \left[\int \mathrm{d}x\, f(x,y) \right],$$

where y_l, y_u, $x_1(y)$ and $x_2(y)$ are the same as defined previously after Equation (12.5). Thus, we have the following general statement of Fubini's theorem:

Theorem 12.4 (Fubini) *Let Ω be a region of $\mathbb{R}^2$ with boundary C, where C is a closed, piecewise differentiable curve. Let $f(x,y)$ be a continuous function on Ω and C. Then*

$$\int\int_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_\Omega \left[\mathrm{d}x \int \mathrm{d}y\, f(x,y) \right] = \int_\Omega \mathrm{d}y \left[\int \mathrm{d}x\, f(x,y) \right]. \tag{12.9}$$

The result extends to n-fold integrals over finite domains in $\mathbb{R}^n$.

Finally, it would be bold (and indeed erroneous) to claim that we have proved Equation (12.9). In particular, the claim after Equation (12.7) about the order in which the limits are taken would need to be shown rigorously with lots epsilons and deltas, or in the much more general context of the more powerful theory of Lebesgue integration and measure theory.

Chapter 13

The limits of integration are not constants any more!

Overview

In this section we focus on computing the area and volume of irregular (i.e. non-cuboid) shapes in two and three dimensions. These involve integrals of the form

$$\int\int_{\text{Area enclosed by some curve}} \mathrm{d}x\,\mathrm{d}y, \qquad \int\int_{\text{Volume enclosed by some surface}} \mathrm{d}x\,\mathrm{d}y\mathrm{d}z.$$

The most novel feature of these problems for the class is the appearance of **non-constant limits of integration**. We have already encountered these limits – in theory – in the previous chapter. We now work out practical strategies to do calculations with such limits.

13.1 Introduction

Imagine an evil genius who is rubbish at elementary maths but a wizard at calculus. She wants to compute the area of a right-angle triangle. She would proceed as follows.

Vertices at $(0,0)$, $(a,0)$, $(0,b)$, where a and b are positive constants. The triangle is thus bounded by the lines $y=0$, $x=0$, and $y=mx+b$, where $m=-b/a$. An element of area in the x-y plane is

$$\mathrm{d}S = \mathrm{d}x\,\mathrm{d}y.$$

Hence,

$$\text{Area of triangle} = \int_{\text{Region bounded by three lines mentioned}} \mathrm{d}x\mathrm{d}y.$$

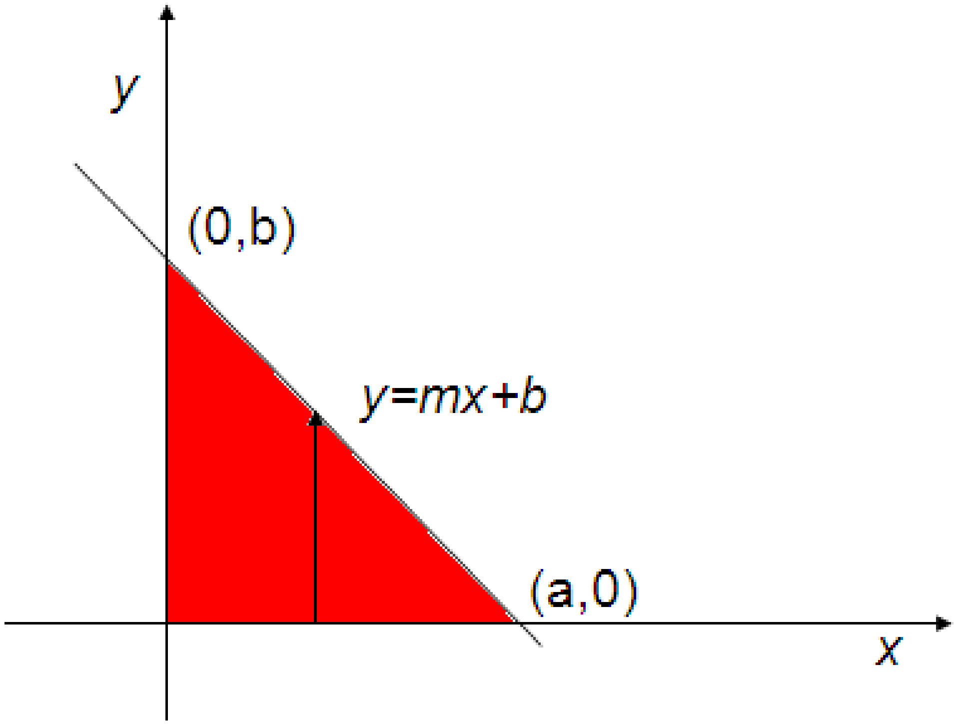


Figure 13.1: Integration domain to compute the area of a right-angled triangle.

Now the variable x is allowed to run between 0 and a, while the variable y is allowed to run between 0 and $mx + b$. Hence,

$$
\begin{aligned}
\text{Area of triangle} &= \int_0^a \mathrm{d}x \int_0^{mx+b} \mathrm{d}y, \\
&= \int_0^a \mathrm{d}x y\Big|_0^{mx+b}, \\
&= \int_0^a \mathrm{d}x \left(mx+b\right), \\
&= \int_0^a \mathrm{d}x \frac{1}{2m}\frac{d}{dx}\left(mx+b\right)^2, \\
&= \tfrac{1}{2m}\left(mx+b\right)^2\Big|_0^a, \\
&= -\frac{1}{2m}b^2, \\
&= \frac{1}{2}ab.
\end{aligned}
$$

13.2 Density integrals

Suppose that the integration domain in Fig. 13.1 instead represents a thin sheet of metal, whose density varies as

$$\rho(x,y) = 1 + \epsilon y \cos(2\pi x/a) = \frac{\text{Mass}}{\text{Unit area}}.$$

Compute the mass of the sheet.

We have,

$$\begin{aligned}
\mathrm{d}m &= \rho(x,y)\mathrm{d}x\,\mathrm{d}y, \\
m &= \int_{\text{Triangle}} \rho(x,y)\mathrm{d}x\,\mathrm{d}y, \\
&= \int_0^a \mathrm{d}x \int_0^{b-(b/a)x} \mathrm{d}y\left[1+\epsilon y\cos(2\pi x/a)\right], \\
&= \int_0^a \mathrm{d}x\left[y + \tfrac{1}{2}y^2\epsilon\cos(2\pi x/a)\right]_{y=0}^{y=b-(b/a)x},
\end{aligned}$$

Hence,

$$\begin{aligned}
m &= \int_0^a \left[b-(b/a)x\right]\mathrm{d}x + \tfrac{1}{2}\epsilon\int_0^a \left[b-(b/a)x\right]^2\cos(2\pi x/a)\mathrm{d}x, \\
&= \tfrac{1}{2}ab + \tfrac{1}{2}\epsilon b^2\int_0^a \cos\left(2\pi x/a\right)\mathrm{d}x - \epsilon(b^2/a)\int_0^a x\cos(2\pi x/a)\mathrm{d}x \\
&\qquad + \tfrac{1}{2}\epsilon(b/a)^2\int_0^a x^2\cos(2\pi x/a)\mathrm{d}x, \\
&= \tfrac{1}{2}ab + \tfrac{1}{2}\epsilon b^2 I_1 + (-\epsilon b^2/a)I_2 + \left(\epsilon b^2/2a^2\right)I_3.
\end{aligned}$$

Now

$$\begin{aligned}
I_1 &= \int_0^a \cos(2\pi x/a)\mathrm{d}x, \\
&= \frac{a}{2\pi}\sin(2\pi x/a)\Big|_0^a, \\
&= \frac{a}{2\pi}\left(\sin(2\pi)-\sin(0)\right) = 0;
\end{aligned}$$

$$\begin{aligned}
I_2 &= \int_0^a x\cos(2\pi x/a)\mathrm{d}x, \\
&= \frac{a^2}{4\pi^2}\int_0^{2\pi} s\cos(s)\mathrm{d}s, \\
&= \frac{a^2}{4\pi^2}\left[s\sin(s)+\cos(s)\right]_0^{2\pi} = 0.
\end{aligned}$$

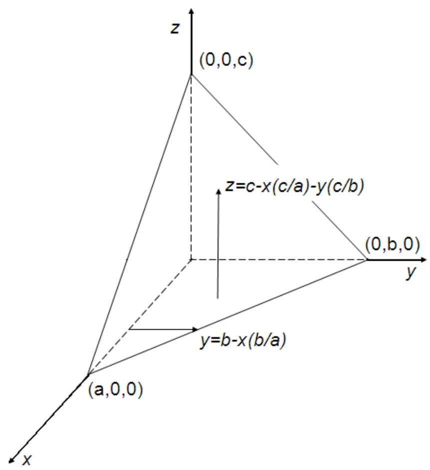
$$
\begin{aligned}
I_3 &= \int_0^a x^2 \cos(2\pi x/a)\mathrm{d}x, \\
&= \frac{a^3}{8\pi^3} \int_0^{2\pi} s^2 \cos(s)\mathrm{d}s, \\
&= \frac{a^3}{8\pi^3} \left[(s^2-2)\sin(s) + 2s\cos(s)\right]_0^{2\pi}, \\
&= \frac{a^3}{8\pi^3} \left[2(2\pi)\cos(2\pi)\right], \\
&= \frac{a^3}{8\pi^3} 4\pi, \\
&= \frac{a^3}{2\pi^2}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
m &= \tfrac{1}{2}ab + \frac{\epsilon b^2}{2a^2}\frac{a^3}{2\pi^2}, \\
&= \tfrac{1}{2}ab + \frac{\epsilon}{4\pi^2} b^2 a, \\
&= \tfrac{1}{2}ab\left(1 + \tfrac{1}{2\pi^2}\epsilon b\right).
\end{aligned}
$$

13.3 Volume integrations

Buoyed by her success, the evil genius of Section 13.1 decides to compute the volume of a a certain pyramid. The pyramid has three edges that come together at right angles; the edges have lengths a, b, and c (tetrahedron). We of course know that the final answer must be $abc/6$. Here below is how our friend would proceed.



The three edges come together to form a right angle at the vertex $(0,0,0)$. The other extremities of the three edges are at $(a,0,0)$, $(0,b,0)$, and $(0,0,c)$. The pyramid is thus bounded by the planes $x=0$, $y=0$, $z=0$, and by a certain other plane, which our friend must work out.

The points $(a,0,0)$, $(0,b,0)$, and $(0,0,c)$ live in this fourth bounding plane. Hence, the vectors

$$
\boldsymbol{v}_1 = (0,0,c) - (a,0,0) = (-a,0,c), \qquad \boldsymbol{v}_2 = (0,b,0) - (a,0,0) = (-a,b,0)
$$

are parallel to the plane and have the same base at the point $(a, 0, 0)$. Thus,

$$\boldsymbol{n} = \boldsymbol{v}_2 \times \boldsymbol{v}_1 = \hat{\boldsymbol{x}} bc + \hat{\boldsymbol{y}} ac + \hat{\boldsymbol{z}} ab = (bc, ac, ab)$$

is normal to the plane, and the plane is therefore defined by

$$[\boldsymbol{x} - (a, 0, 0)] \cdot \boldsymbol{n} = 0,$$

or

$$z = c - \frac{c}{a} x - \frac{c}{b} y.$$

In summary, the four bounding planes are

$$x = 0, \qquad y = 0, \qquad z = 0, \qquad z = c - \frac{c}{a} x - \frac{c}{b} y.$$

The volume element is

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

Hence,

$$\text{Volume of pyramid} = \int_{\text{Region bounded by four planes}} \mathrm{d}x \mathrm{d}y \mathrm{d}z.$$

Now the variable x is allowed to run between 0 and a, while the variable y is allowed to run between 0 and $y = b - x(b/a)$. This is because, in the x-y plane, the integration reduces to the triangle integration in Section 13.1. Finally, the variable z is allowed to run between 0 and $z = c - (c/a)x - (c/b)y$.

Thus,

$$
\begin{aligned}
\text{Volume of pyramid} &= \int_0^a \mathrm{d}x \int_0^{b-x(b/a)} \mathrm{d}y \int_0^{c-x(c/a)-y(c/b)} \mathrm{d}z, \\
&= \int_0^a \mathrm{d}x \int_0^{b-x(b/a)} \mathrm{d}y\, z\Big|_0^{c-x(c/a)-y(c/b)}, \\
&= \int_0^a \mathrm{d}x \int_0^{b-x(b/a)} \mathrm{d}y \left[c - x(c/a) - y(c/b)\right], \\
&= \int_0^a \mathrm{d}x \int_0^{b-x(b/a)} \mathrm{d}y \left(-\frac{b}{2c}\right) \frac{\partial}{\partial y} \left[c - x(c/a) - y(c/b)\right]^2, \\
&= -\frac{b}{2c} \int_0^a \mathrm{d}x \int_0^{b-x(b/a)} \mathrm{d}y \frac{\partial}{\partial y} \left[c - x(c/a) - y(c/b)\right]^2, \\
&= -\frac{b}{2c} \int_0^a \mathrm{d}x \left[c - x(c/a) - y(c/b)\right]^2 \Big|_{y=0}^{y=b-x(b/a)}, \\
&= +\frac{b}{2c} \int_0^a \mathrm{d}x \left[c - x(c/a)\right]^2, \\
&= \frac{b}{2c} \int_0^a \mathrm{d}x \left(-\frac{a}{3c}\right) \frac{\partial}{\partial x} \left[c - x(c/a)\right]^3, \\
&= -\frac{ab}{6c^2} \left[c - x(c/a)\right]^3 \Big|_0^a, \\
&= +\frac{ab}{6c^2} c^3, \\
&= \tfrac{1}{6} abc.
\end{aligned}
$$

Chapter 14

Integrals over surfaces and volumes

Overview

In this section we devise a general method for computing area integrals. It holds for arbitrary shapes. We also introduce volume integrals. Worked examples are provided in each case.

14.1 Parametrization of surface integrals

We focus on surface integrals in three dimensions. That is, we are to integrate a vector field $\boldsymbol{v}(x,y,z)$ over a surface S. The element of surface area actually has an orientation:

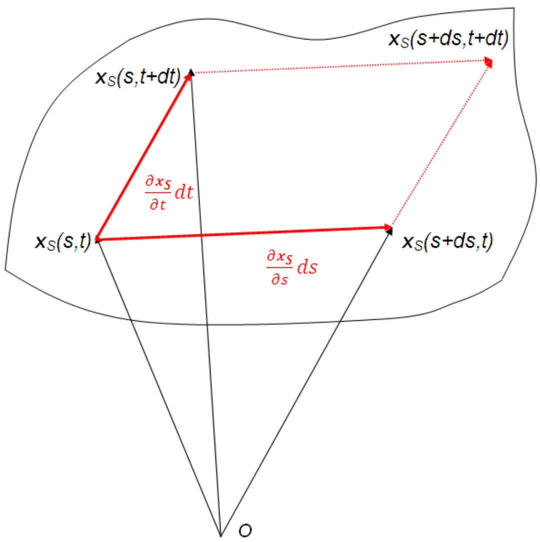
$$\mathrm{d}\boldsymbol{S} = \hat{\boldsymbol{n}}\,\mathrm{d}S,$$

where $\hat{\boldsymbol{n}}$ is normal to the surface at location $\boldsymbol{x}$ on the surface. By convention, we choose $\hat{\boldsymbol{n}}$ to be the outward-pointing normal. We focus on the most commonly-encountered integral:

$$\int_S \boldsymbol{v}(\boldsymbol{x})\cdot \mathrm{d}\boldsymbol{S}.$$

To do line integrals along a curve, we had to introduce a **parametrization** of the curve. We must do a similar thing here: We parametrize the surface S as follows:

$$S = \{\boldsymbol{x}\in\mathbb{R}^3 | \boldsymbol{x} = \boldsymbol{x}_S(s,t), (s,t)\in\Omega_S\},$$
$$\Omega_S = \text{Some subset of } \mathbb{R}^2.$$



Thus, a curve in three dimensions is a one-parameter set, and a surface is a two-parameter set. Refer to the figure on the right and consider the points

$$\boldsymbol{x}_S(s,t), \qquad \boldsymbol{x}_S(s,t+\mathrm{d}t),$$
$$\boldsymbol{x}_S(s+ds,t), \qquad \boldsymbol{x}_S(s+ds,t+dt),$$

vectors whose tips all lie in the surface S. Form the differences

$$\boldsymbol{x}_S(s,t+\mathrm{d}t) - \boldsymbol{x}_S(s,t) = \frac{\partial \boldsymbol{x}_S}{\partial t}\mathrm{d}t,$$

and

$$\boldsymbol{x}_S(s+\mathrm{d}s,t) - \boldsymbol{x}_S(s,t) = \frac{\partial \boldsymbol{x}_S}{\partial s}\mathrm{d}s.$$

These are small vectors that lie in the surface and form the two lengths of a parallelogram. The area described by the four points $\boldsymbol{x}_S(s,t)$,...,$\boldsymbol{x}_S(s+\mathrm{d}s, t+\mathrm{d}t)$ is thus

$$\mathrm{d}\boldsymbol{S} = \frac{\partial \boldsymbol{x}_S}{\partial s} \times \frac{\partial \boldsymbol{x}_S}{\partial t} \mathrm{d}s\,\mathrm{d}t.$$

If the parameters s and t take values in a set Ω_S, then the surface integral $\int_S \boldsymbol{v}(\boldsymbol{x}) \cdot \mathrm{d}\boldsymbol{S}$ is

$$\int_S \boldsymbol{v}(\boldsymbol{x}) \cdot \mathrm{d}\boldsymbol{S} = \int\int_{\Omega_S} \boldsymbol{v}(\boldsymbol{x}_S(s,t)) \cdot \left(\frac{\partial \boldsymbol{x}_S}{\partial t} \times \frac{\partial \boldsymbol{x}_S}{\partial s} \right) \mathrm{d}t\,\mathrm{d}s$$

14.2 Worked examples

If $\boldsymbol{v} = 2y\hat{\boldsymbol{x}} - z\hat{\boldsymbol{y}} + x^2\hat{\boldsymbol{z}}$ and S is the surface of the parabolic cylinder $y^2 = 8x$ in the first (positive) octant bounded by the planes $y = 4$ and $z = 6$, evaluate $\displaystyle\int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}$.

Solution: Let us compute the surface in parametric form. The parametric form of the curve is

$$\begin{aligned} y_S(s,t) &= s, \\ z_S(s,t) &= t, \\ x_S(s,t) &= s^2/8. \end{aligned}$$

where $0 \le s \le 4$ and $0 \le t \le 6$. Hence,

$$\boldsymbol{x}_S(s,t) = (s^2/8, s, t).$$

$$\frac{\partial \boldsymbol{x}_S}{\partial s} = (s/4, 1, 0), \qquad \frac{\partial \boldsymbol{x}_S}{\partial t} = (0, 0, 1)$$

and

$$\mathrm{d}\boldsymbol{S} = \left(\frac{\partial \boldsymbol{x}_S}{\partial s} \times \frac{\partial \boldsymbol{x}_S}{\partial t} \right) \mathrm{d}s\,\mathrm{d}t = \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ s/4 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \mathrm{d}s\mathrm{d}t = [\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}(s/4)]\,\mathrm{d}s\,\mathrm{d}t$$

Hence

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \left(2y\hat{\boldsymbol{x}} - z\hat{\boldsymbol{y}} + x^2\hat{\boldsymbol{z}}\right) \cdot (\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}(s/4))\,\mathrm{d}s\,\mathrm{d}t = (2y + zs/4)\,\mathrm{d}s\,\mathrm{d}t.$$

But $y = s$ and $z = t$, hence

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = (2s + ts/4)\,\mathrm{d}s\,\mathrm{d}t.$$

We let $0 \le s \le 4$ and $0 \le t \le 6$ and integrate. We make use of the following remarkable fact:

$$\int_{s_1}^{s_2} \mathrm{d}s \int_{t_1}^{t_2} \mathrm{d}t\, \phi(s,t) = \int_{t_1}^{t_2} \mathrm{d}t \int_{s_1}^{s_2} \mathrm{d}s\, \phi(s,t),$$

that is, the order of integration can be reversed, for suitable functions ϕ. Such a reversal cannot be done if, in the first integral, the limits t_1 and t_2 depend on s. Here, however, the limits are constants.

$$\begin{aligned}
\int_{s=0}^{s=4} \int_{t=0}^{t=6} (2s + ts/4)\, \mathrm{d}s\, \mathrm{d}t &= \int_{s=0}^{s=4} \int_{t=0}^{t=6} 2s \mathrm{d}s\, \mathrm{d}t + \tfrac{1}{4} \int_{s=0}^{s=4} \int_{t=0}^{t=6} ts \mathrm{d}s\, \mathrm{d}t, \\
&= \int_{t=0}^{t=6} \left(\int_{s=0}^{s=4} 2s \mathrm{d}s \right) \mathrm{d}t + \tfrac{1}{4} \left(\int_{s=0}^{s=4} s \mathrm{d}s \right) \left(\int_{t=0}^{t=6} t \mathrm{d}t \right), \\
&= 16 \times 6 + \left(\tfrac{1}{4} \times \tfrac{1}{4} \times 16 \times 36 \right), \\
&= 132.
\end{aligned}$$

One particularly easy case involves surface integrals over cuboids. Let us consider such an example now:

Example: If $\boldsymbol{v} = x\hat{\boldsymbol{x}} + 2y\hat{\boldsymbol{y}} + 3z\hat{\boldsymbol{z}}$ and S is the unit cube with a vertex at $(0,0,0)$ and situated in the positive octant, compute $\int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}$.

Solution: Refer to Figure 14.1. We divide the area S into its six faces, Fxp, Fxm, Fyp, Fym, Fzp, Fzm. Consider the face Fxp. This is the face contained entirely in a $y - z$ plane, with unit normal $+\hat{\boldsymbol{x}}$, and such that $x = 1$. Consider also Fxm. Again, this face is contained entirely in a $y - z$ plane, with unit normal $-\hat{\boldsymbol{x}}$, and with $x = 0$. Along Fxp,

$$\mathrm{d}\boldsymbol{S} = \mathrm{d}y\, \mathrm{d}z \hat{\boldsymbol{x}},$$

and

$$\boldsymbol{v} \cdot \boldsymbol{S} = x \mathrm{d}S = x \mathrm{d}y\, \mathrm{d}z = \mathrm{d}y\, \mathrm{d}z, \qquad x = 1.$$

Along Fxm, $\mathrm{d}\boldsymbol{S} = -\mathrm{d}y \mathrm{d}z \hat{\boldsymbol{x}}$ and $\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = -x \mathrm{d}S = -x \mathrm{d}y\, \mathrm{d}z = 0$, since $x = 0$ on this face. Hence,

$$\int_{Fxm} + \int_{Fxp} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \int_0^1 \mathrm{d}y \int_0^1 1\, \mathrm{d}z = 1.$$

Similarly,

$$\int_{Fym} + \int_{Fyp} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \int_0^1 \mathrm{d}x \int_0^1 2\, \mathrm{d}z = 2,$$

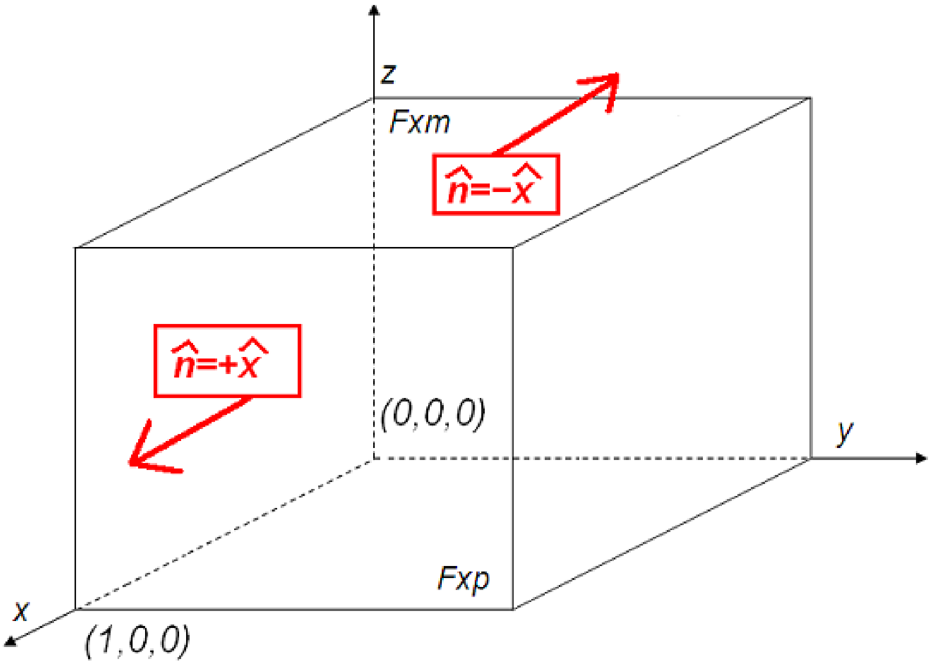


Figure 14.1: Integration over a cuboid.

and

$$\int_{Fzm} + \int_{Fzp} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \int_0^1 \mathrm{d}z \int_0^1 3\,\mathrm{d}x = 3,$$

Putting it all together,

$$\int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \left[\int_{Fxm} + \int_{Fxp} + \int_{Fym} + \int_{Fyp} + \int_{Fzm} + \int_{Fzp}\right] \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = 6.$$

14.3 Volume integrals

Volume integrals are much simpler than the other two, since the volume element $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ is a scalar. For a scalar field $\phi(x, y, z)$, the volume integral

$$\int_\Omega \phi(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

is the ordinary triple integral over the domain $\Omega \subset \mathbb{R}^3$. For a vector field $\boldsymbol{v}(x, y, z)$, the associated volume integral can be broken up into three scalar integrals:

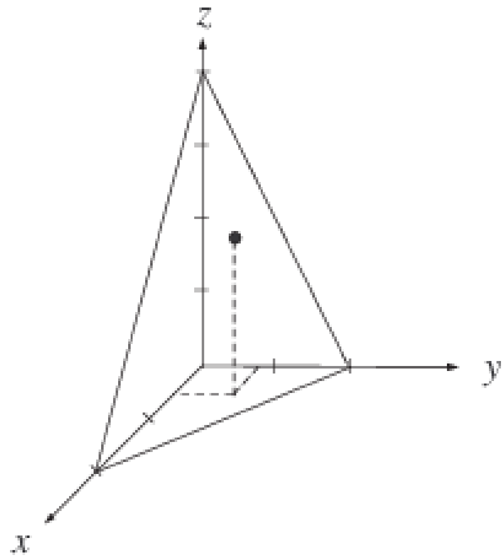
$$\int_\Omega \boldsymbol{v}(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \hat{\boldsymbol{x}} \int_\Omega v_1(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \hat{\boldsymbol{y}} \int_\Omega v_2(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \hat{\boldsymbol{z}} \int_\Omega v_3(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$

since the unit vectors $\hat{\boldsymbol{x}}$ etc. are constants and can be taken outside the integrals.

Example: If $\boldsymbol{v} = (2x^2 - 3z)\hat{\boldsymbol{x}} - 2xy\hat{\boldsymbol{y}} - 4x\hat{\boldsymbol{z}}$, evaluate

$$\int_\Omega \nabla \cdot \boldsymbol{v} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z,$$

where Ω is the closed region bounded by the planes $x = 0$, $y = 0$, $z = 0$ and $2x + 2y + z = 4$.



Solution: Notice that

$$\nabla \cdot \boldsymbol{v} = 4x - 2x = 2x.$$

To find out where the plane $2x+2y+z = 4$ intersects the x and y axes, let $z = 0$. Then $2x+2y = 4$, and the plane intersects the x-axis when y=0, i.e. $x = 2$. Referring to the figure, in order for all values in the domain Ω to be included in the integration,

- x must vary between 0 and 2;
- y must vary between 0 and $y = 2 - x$;
- z must vary between 0 and $z = 4 - 2x - 2y$.

Hence,

$$
\begin{aligned}
\int_V \operatorname{div} \boldsymbol{v} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z &= 2\int_0^2 \mathrm{d}x \int_0^{2-x} \mathrm{d}y \int_0^{4-2x-2y} \mathrm{d}z \, x \\
&= 2\int_0^2 \mathrm{d}x \, x \int_0^{2-x} \mathrm{d}y z\Big|_0^{4-2x-2y}, \\
&= 2\int_0^2 \mathrm{d}x \, x \int_0^{2-x} \mathrm{d}y(4-2x-2y), \\
&= 2\int_0^2 \mathrm{d}x \, x \int_0^{2-x} \mathrm{d}y(4-2x) - 4\int_0^2 \mathrm{d}x \, x \int_0^{2-x} \mathrm{d}y \, y, \\
&= 2\int_0^2 \mathrm{d}x \, x(4-2x)y\Big|_0^{2-x} - 2\int_0^2 \mathrm{d}x \, xy^2\Big|_0^{2-x}, \\
&= 2\int_0^2 \mathrm{d}x \, x \left[2\,(2-x)\,(2-x) - (2-x)^2\right], \\
&= 2\int_0^2 \mathrm{d}x \, x\,(2-x)^2, \\
&= 2\int_0^2 \mathrm{d}x \, \left(4x - 4x^2 + x^3\right), \\
&= 2\left(2x^2 - \tfrac{4}{3}x^3 + \tfrac{1}{4}x^4\right)_0^2 = 8/3.
\end{aligned}
$$

Pedantic note Sometimes, instead of the notation $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ for the volume element, we will write $\mathrm{d}V$, but we mean the same thing. The notation V will sometimes be used to denote a volume or domain in $\mathbb{R}^3$. Thus, it is not unusual to write

$$\int_V \phi(\boldsymbol{x})\,\mathrm{d}V$$

to denote the integration of the scalar field $\phi(\boldsymbol{x})$ over the domain $V \subset \mathbb{R}^3$.

Chapter 15

Integrals over surfaces and volumes – further worked examples

Overview

We provide several worked examples of surface and volume integrals.

15.1 Line integral

Example: If $\boldsymbol{v} = 2y\hat{\boldsymbol{x}} - z\hat{\boldsymbol{y}} + x\hat{\boldsymbol{z}}$, evaluate $\int_C \boldsymbol{v} \times \mathrm{d}\boldsymbol{x}$ along the curve $x = \cos(t)$, $y = \sin(t)$, $z = 2\cos(t)$ from $t = 0$ to $t = \pi/2$.

Solution: We have,

$$\boldsymbol{x}(t) = (\cos(t), \sin(t), 2\cos(t)), \qquad \frac{d\boldsymbol{x}}{dt} = (-\sin(t), \cos(t), -2\sin(t)),$$

and

$$I := \int_C \boldsymbol{v} \times \mathrm{d}\boldsymbol{x} = \int_{t_1}^{t_2} \boldsymbol{v} \times \frac{d\boldsymbol{x}}{dt}\mathrm{d}t.$$

Now

$$\boldsymbol{v} = (2y, -z, x) = (2\sin(t), -2\cos(t), \cos(t)),$$

hence

$$
\begin{aligned}
\boldsymbol{v} \times \frac{d\boldsymbol{x}}{dt} &= \begin{vmatrix} \hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}} \\ 2\sin(t) & -2\cos(t) & \cos(t) \\ -\sin(t) & \cos(t) & -2\sin(t) \end{vmatrix} \\
&= \hat{\boldsymbol{x}} \left[4\sin(t)\cos(t) - \cos^2(t)\right] + \hat{\boldsymbol{y}} \left[4\sin^2(t) - \sin(t)\cos(t)\right],
\end{aligned}
$$

hence

$$
\begin{aligned}
I &= \hat{\boldsymbol{x}} \int_0^{\pi/2} \left[4\sin(t)\cos(t) - \cos^2(t)\right] \mathrm{d}t + \hat{\boldsymbol{y}} \int_0^{\pi/2} \left[4\sin^2(t) - \sin(t)\cos(t)\right] \mathrm{d}t, \\
&= \hat{\boldsymbol{x}} I_1 + \hat{\boldsymbol{y}} I_2.
\end{aligned}
$$

We will need the following trigonometric identities:

$$
\begin{aligned}
\cos^2\theta &= \tfrac{1}{2}\left(1 + \cos 2\theta\right), \\
\sin^2\theta &= \tfrac{1}{2}\left(1 + \sin 2\theta\right), \\
\sin\theta\cos\theta &= \tfrac{1}{2}\sin 2\theta.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
I_1 &= \int_0^{\pi/2} \left[4\sin(t)\cos(t) - \cos^2(t)\right] \mathrm{d}t, \\
&= \int_0^{\pi/2} \left[2\sin(2t) - \tfrac{1}{2}\left(1 + \cos(2t)\right)\right] \mathrm{d}t, \\
&= -\cos(2t)\Big|_0^{\pi/2} - \tfrac{1}{2}t\Big|_0^{\pi/2} - \tfrac{1}{2}\sin(2t)\Big|_0^{\pi/2}, \\
&= 2 - \tfrac{1}{4}\pi.
\end{aligned}
$$

Similarly, $I_2 = \pi - 1/2$, hence

$$
I = \hat{\boldsymbol{x}}\left(2 - \tfrac{1}{4}\pi\right) + \hat{\boldsymbol{y}}\left(\pi - \tfrac{1}{2}\right).
$$

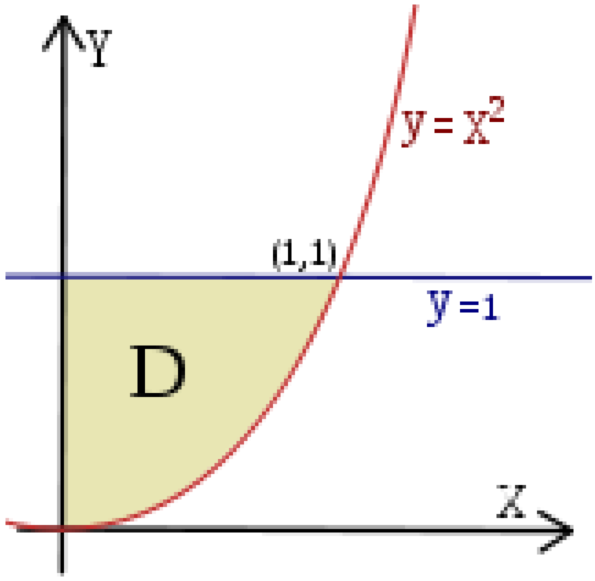
15.2 Straightforward area integral

Example: Consider the region D defined as follows and shown in the accompanying figure:

$$D = \{(x, y) \in \mathbf{R}^2 \ : \ x \geq 0, y \leq 1, y \geq x^2\}$$

Compute

$$\int\int_D (x + 2y)\, \mathrm{d}x\, \mathrm{d}y.$$



Solution: Clearly, x ranges freely between $x = 0$ and $x = 1$, and for a given x, y ranges between $y = x^2$ (lower) and $y = 1$ (upper). Thus, the double integral can be re-expressed as a repeated application of single integrals – with non-constant limits of integration, as follows:

$$\begin{aligned}
\int\int_D (2x + y)\, \mathrm{d}x\, \mathrm{d}y &\overset{\text{Fubini}}{=} \int_0^1 \mathrm{d}x \int_{y=x^2}^{y=1} (x + 2y)\mathrm{d}y, \\
&= \int_0^1 \mathrm{d}x \int_{y=x^2}^{y=1} x\, \mathrm{d}y + 2\int_0^1 \mathrm{d}x \int_{y=x^2}^{y=1} y\, \mathrm{d}y, \\
&= \int_0^1 x\, \mathrm{d}x \int_{y=x^2}^{y=1} \mathrm{d}y + 2\int_0^1 \mathrm{d}x \int_{y=x^2}^{y=1} y\, \mathrm{d}y, \\
&= \int_0^1 \mathrm{d}x\, x(y)_{y=x^2}^{y=1} + 2\int_0^1 \mathrm{d}x \left(\tfrac{1}{2}y^2\right)_{y=x^2}^{y=1}, \\
&= \int_0^1 \mathrm{d}x\, x(1 - x^2) + \int_0^1 \left(1 - x^4\right) \mathrm{d}x, \\
&= \int_0^1 \left(x - x^3 - x^4 + 1\right), \\
&= \tfrac{1}{2}x\Big|_0^1 - \tfrac{1}{4}x^3\Big|_0^1 - \tfrac{1}{5}\Big|_0^1 + x\Big|_0^1, \\
&= \tfrac{1}{2} - \tfrac{1}{4} - \tfrac{1}{5} + 1, \\
&= \tfrac{10}{20} - \tfrac{5}{20} - \tfrac{4}{20} + \tfrac{20}{20}, \\
&= \tfrac{21}{20}.
\end{aligned}$$

15.3 Straightforward volume integral

Example: Compute the average value of the function $f(x, y) = xy$ over the volume

$$V = \{(x, y, z)|x \le 2 \le 4,\ y \le 3 \le 9,\ 0 \le z \le 2 + 3x + 4y\}.$$

Solution: The average value is

$$\mathrm{avg}(f) = \frac{1}{V}\int_V f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$

where V is used here both to denote the region of integration and its volume. We compute the numerator:

$$\begin{aligned}
\text{Num} &= \int_V f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \\
&\overset{\text{Fubini}}{=} \int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y \int_0^{2+3x+4y} xy\,\mathrm{d}z, \\
&= \int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\, xy(2 + 3x + 4y), \\
&= 2\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\, xy + 3\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\, x^2 y + 4\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\, xy^2.
\end{aligned}$$

Carrying on, this is

$$\begin{aligned}
\text{Num} &= 2\int_2^4 x\,\mathrm{d}x \int_3^9 y\,\mathrm{d}y + 3\int_2^4 x^2\,\mathrm{d}x \int_3^9 y\,\mathrm{d}y + 4\int_2^4 x\,\mathrm{d}x \int_3^9 y^2\,\mathrm{d}y, \\
&= 2\left(\tfrac{1}{2}x^2\right)_2^4\left(\tfrac{1}{2}y^2\right)_3^9 + 3\left(\tfrac{1}{3}x^3\right)_2^4\left(\tfrac{1}{2}y^2\right)_3^9 + 4\left(\tfrac{1}{2}x^2\right)_2^4\left(\tfrac{1}{3}y^3\right)_3^9, \\
&= \tfrac{1}{2}(4^2 - 2^2)(9^2 - 3^2) + \tfrac{1}{2}(2)(9^2 - 3^2) + \tfrac{2}{3}(4^2 - 2^2)(9^3 - 3^3), \\
&= \tfrac{1}{2}12 \times 72 + \tfrac{1}{2}56 \times 72 + \tfrac{2}{3}12 \times (729 - 27), \\
&= 6 \times 72 + 28 \times 72 + 8 \times 702, \\
&= 8064.
\end{aligned}$$

We compute the denominator:

$$\begin{aligned}
\text{Den} &= \int_V \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,\\
&\overset{\text{Fubini}}{=} \int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y \int_0^{2+3x+4y} \mathrm{d}z,\\
&= \int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\,(2+3x+4y),\\
&= 2\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y + 3\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\,x + 4\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y\,y\\
&= 2\int_2^4 \mathrm{d}x \int_3^9 \mathrm{d}y + 3\int_2^4 x\,\mathrm{d}x \int_3^9 \mathrm{d}y + 4\int_2^4 \mathrm{d}x \int_3^9 y\,\mathrm{d}y,\\
&= 2(x)_2^4(y)_3^9 + 3(\tfrac{1}{2}x^2)_2^4(y)_3^9 + 4(x)_2^4(\tfrac{1}{2}y^2)_3^9,\\
&= 2(4-2)(9-3) + \tfrac{3}{2}(4^2-2^2)(9-3) + 2(4-2)(9^2-3^2),\\
&= 2\times 2\times 6 + \tfrac{3}{2}12\times 6 + 2\times 2\times 72,\\
&= 2\times 2\times 6 + 12\times 9 + 2\times 2\times 72,\\
&= 420.
\end{aligned}$$

The answer is

$$\mathrm{avg}(f) = \frac{\text{Num}}{\text{Den}} = \tfrac{8064}{420} = \tfrac{96}{5}.$$

15.4 Area integral by parametrization

Example: Find the surface area of the plane $x + 2y + 2z = 12$ cut off by $x = 0$, $y = 0$, and $x^2 + y^2 = 16$.

Solution: See Figure 15.1. Parametrise the surface:

$$\begin{aligned}
x &= s,\\
y &= t,\\
z &= \frac{12 - s - 2t}{2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\boldsymbol{x} &= \left(s, t, 6 - \tfrac{1}{2}s - t\right),\\
\frac{\partial \boldsymbol{x}}{\partial s} &= \left(1, 0, -\tfrac{1}{2}\right),\\
\frac{\partial \boldsymbol{x}}{\partial t} &= (0, 1, -1).
\end{aligned}$$

Area element (vector):

$$\begin{aligned}
\mathrm{d}\boldsymbol{S} &= \frac{\partial \boldsymbol{x}}{\partial s} \times \frac{\partial \boldsymbol{x}}{\partial t}\mathrm{d}s\mathrm{d}t, \\
&= \left(\tfrac{1}{2}, 1, 1\right)\mathrm{d}s\mathrm{d}t.
\end{aligned}$$

Area element (scalar):

$$\begin{aligned}
\mathrm{d}S &= \left|\frac{\partial \boldsymbol{x}}{\partial s} \times \frac{\partial \boldsymbol{x}}{\partial t}\right|\mathrm{d}s\mathrm{d}t, \\
&= \sqrt{2 + \tfrac{1}{4}}\mathrm{d}s\mathrm{d}t.
\end{aligned}$$

The parameters run between

$$\left\{ \begin{array}{c} 0 \le s \\ 0 \le t \end{array} \right\} \text{ and } s^2 + t^2 \le 16.$$

In the $s - t$ plane, this is a quarter-circle of radius 4, centred at 0. We let s run between 0 and 4 and let t run between 0 and $\sqrt{16 - s^2}$. Thus,

$$\begin{aligned}
S &= \sqrt{2 + \tfrac{1}{4}} \int_0^4 \mathrm{d}s \int_0^{\sqrt{16 - s^2}} \mathrm{d}t, \\
&= \sqrt{2 + \tfrac{1}{4}} \int_0^4 \mathrm{d}s \sqrt{16 - s^2}, \\
&= \sqrt{2 + \tfrac{1}{4}} \times \tfrac{1}{4} \times \text{Area of a circle of radius 4}, \\
&= \tfrac{3}{2} \times \tfrac{1}{4} \times 16\pi, \\
&= 6\pi.
\end{aligned}$$

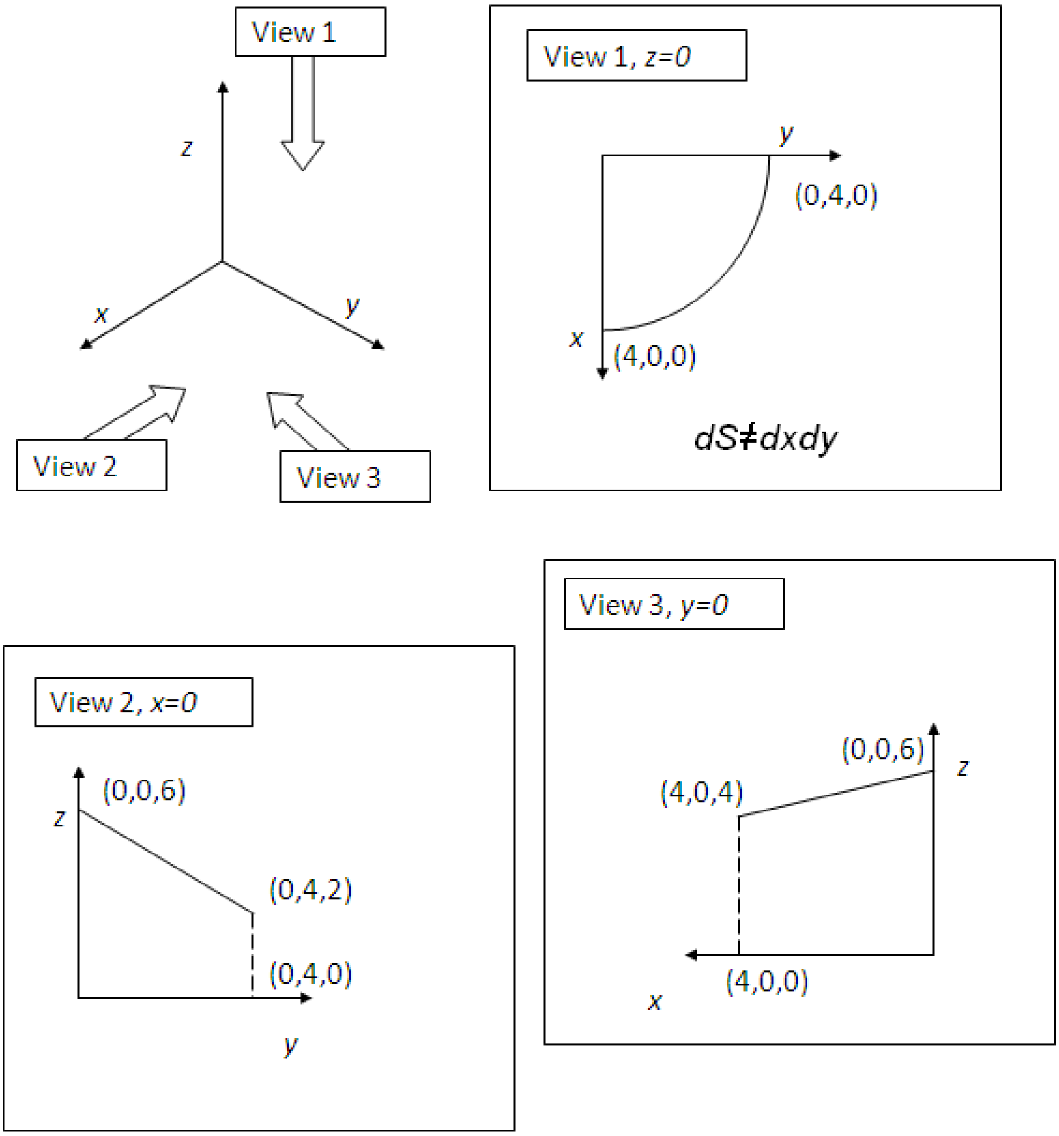

View 1
z
x
y
View 2
View 3
View 1, z=0
y
(0,4,0)
x
(4,0,0)
dS≠dxdy
View 2, x=0
(0,0,6)
z
(0,4,2)
(0,4,0)
y
View 3, y=0
(0,0,6)
z
(4,0,4)
(4,0,0)
x

Figure 15.1:

Chapter 16

Stokes's and Gauss's Theorems

Overview

In ordinary calculus, recall the rule of integration by parts:

$$\int_a^b u\,\mathrm{d}v = (uv)\,|_a^b - \int_a^b v\,\mathrm{d}u.$$

That is, a difficult integral $u\,\mathrm{d}v$ can be split up into an easier integral $v\,\mathrm{d}u$ and a 'boundary term' $u(b)v(b) - u(a)v(a)$. In this section we do something similar for vector integrals.

16.1 Gauss's Theorem (or the Divergence Theorem)

Theorem 16.1 *Let V be a region in space bounded by a closed surface S, and let $\boldsymbol{v}(\boldsymbol{x})$ be a vector field with continuous first derivatives. Then*

$$\boxed{\int_V \nabla\cdot\boldsymbol{v}\,\mathrm{d}V = \int_S \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{S},} \tag{16.1}$$

where $\mathrm{d}\boldsymbol{S}$ is outward-pointing surface-area element associated with the surface S.

Proof: First, consider a box of sides of length Δx, Δy, and Δz, with one vertex positioned at (x, y, z) (Figure 16.1). As in previous exercises, label the faces Fxp, Fxm, Fyp, Fym, Fzp, and Fzm. We compute

$$\sum_{\text{all faces}} \boldsymbol{v}\cdot\Delta\boldsymbol{S},$$

where $\Delta\boldsymbol{S}$ is the area element on each face. For example, in the x-direction, we have a positive

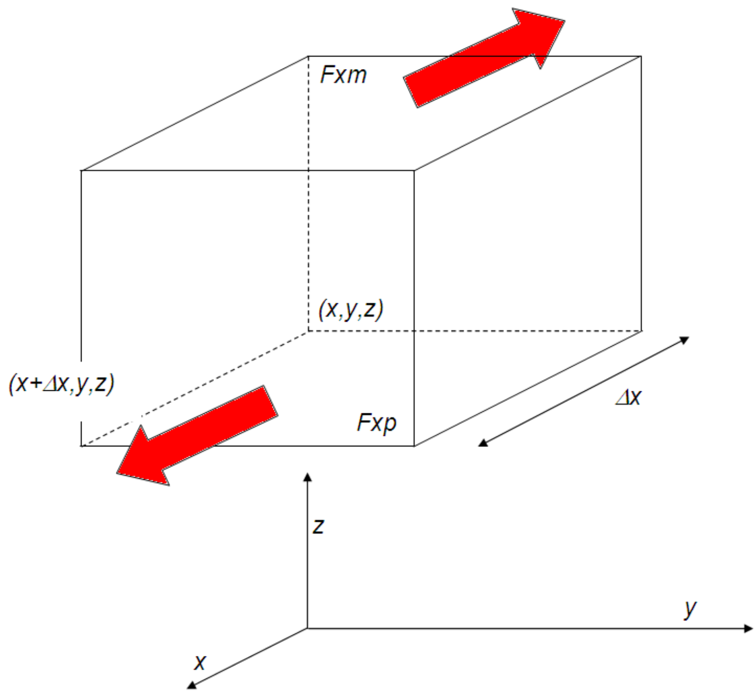


Figure 16.1: Area integration over a box, as applied to Gauss's theorem.

contribution from Fxp and a negative one from Fxm, to give

$$-v_1(x,y,z)\Delta y\Delta z + v_1(x+\Delta x,y,z)\Delta y\Delta z.$$

We immediately write down the other contributions: From Fyp and Fym, we have

$$-v_2(x,y,z)\Delta x\Delta z + v_2(x,y+\Delta y,z)\Delta x\Delta z,$$

and from Fzp and Fzm, we have

$$-v_3(x,y,z)\Delta x\Delta y + v_2(x,y,z+\Delta z)\Delta x\Delta y.$$

Summing over all six contributions (i.e. over all six faces), we have

$$\sum_{\text{all faces}} \boldsymbol{v}\cdot\Delta\boldsymbol{S} =$$
$$v_1(x+\Delta x,y,z)\Delta y\Delta z - v_1(x,y,z)\Delta y\Delta z + v_2(x,y+\Delta y,z)\Delta x\Delta z - v_2(x,y,z)\Delta x\Delta z +$$
$$v_3(x,y,z+\Delta z)\Delta x\Delta y - v_3(x,y,z)\Delta x\Delta y.$$

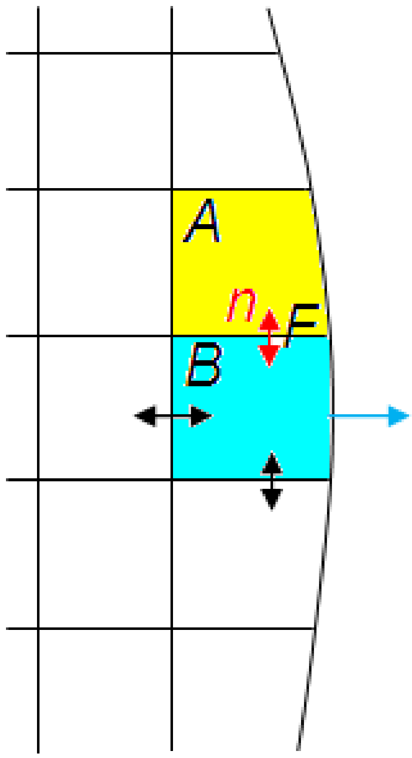
We apply Taylor's theorem to these increments, and omit terms that are $O(\Delta x^2,\Delta y^2,\Delta z^2)$. This becomes rigorous in the limit when the box volume go to zero. In this way, we obtain

$$\sum_{\text{all faces}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \nabla \cdot \boldsymbol{v} \, \mathrm{d}V.$$

Here, we have replaced the Δ's by d's, to indicate the passage to the limiting case as Δx, Δy, and $\Delta z \to 0$.

For the second and final step, consider an arbitrary shape of volume V in three dimensions. We break this volume up into many infinitesimally small boxes. By the previous result, we have

$$\sum_{\text{all boxes}} \nabla \cdot \boldsymbol{v} \, \mathrm{d}V = \sum_{\text{all boxes}} \left(\sum_{\text{all faces}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} \right). \tag{16.2}$$



Consider, however, the two neighbouring boxes shown in the right. Call them A and B These will share a common face, F, with normal vector $\hat{\boldsymbol{n}}$ and area $\mathrm{d}S$. Box A gives a contribution $\hat{\boldsymbol{n}} \cdot \boldsymbol{v}(F) dS$, say, to the sum (16.2), while box B must give a contribution $-\hat{\boldsymbol{n}} \cdot \boldsymbol{v}(F)\mathrm{d}S$. The only place where such a cancellation cannot occur is on exterior faces. Thus,

$$\sum_{\text{all boxes}} \nabla \cdot \boldsymbol{v} \, \mathrm{d}V = \sum_{\text{all exterior faces}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}.$$

But the box volumes are infinitesimally small, so this sum converts into an integral:

$$\int_V \nabla \cdot \boldsymbol{v} \, \mathrm{d}V = \int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}. \quad \blacksquare$$

16.1.1 Green's theorem

A frequently used corollary of Gauss's theorem is a relation called **Green's theorem**. If ϕ and ψ are two scalar fields, then we have the identities

$$\nabla \cdot (\phi \nabla \psi) = \phi \nabla \cdot \nabla \psi + \nabla \phi \cdot \nabla \psi,$$
$$\nabla \cdot (\psi \nabla \phi) = \psi \nabla \cdot \nabla \phi + \nabla \psi \cdot \nabla \phi.$$

Subtracting these equations gives

$$\begin{aligned}\nabla\cdot(\phi\nabla\psi-\psi\nabla\phi) &= \phi\nabla\cdot\nabla\psi-\psi\nabla\cdot\nabla\phi,\\ &= \phi\nabla^2\psi-\psi\nabla^2\phi.\end{aligned}$$

We integrate over a volume V whose boundary is a closed set S. Applying Gauss's theorem gives

$$\begin{aligned}\int_V\left(\phi\nabla^2\psi-\psi\nabla^2\phi\right)\mathrm{d}V &= \int_V\left[\nabla\cdot(\phi\nabla\psi-\psi\nabla\phi)\right]\mathrm{d}V,\\ &= \int_S(\phi\nabla\psi-\psi\nabla\phi)\cdot\mathrm{d}\boldsymbol{S}.\end{aligned}$$

Thus, we arrive at the following result:

Theorem 16.2 (Green's theorem)

$$\int_V\left(\phi\nabla^2\psi-\psi\nabla^2\phi\right)\mathrm{d}V=\int_S(\phi\nabla\psi-\psi\nabla\phi)\cdot\mathrm{d}\boldsymbol{S},$$

where V is a region of $\mathbb{R}^3$ whose boundary is the closed set S.

16.1.2 Other forms of Gauss's theorem

Although the form $\int_V\nabla\cdot\boldsymbol{v}\mathrm{d}V=\int_S\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{S}$ is the most common statement of Gauss's theorem, there are other forms. For example, let

$$\boldsymbol{v}(\boldsymbol{x})=v(\boldsymbol{x})\boldsymbol{a},$$

where $\boldsymbol{a}$ is a **constant vector**. We have

$$\int_V\nabla\cdot\boldsymbol{v}\,\mathrm{d}V=\int_V\nabla\cdot(\boldsymbol{a}\,v)\ \mathrm{d}V=\boldsymbol{a}\cdot\int_V(\nabla v)\,\mathrm{d}V.$$

However, applying Gauss's theorem gives

$$\int_V\nabla\cdot\boldsymbol{v}\,\mathrm{d}V=\int_S v\boldsymbol{a}\cdot\mathrm{d}\boldsymbol{S}=\boldsymbol{a}\cdot\int_S v\,\mathrm{d}\boldsymbol{S}.$$

Equating both sides,

$$\boldsymbol{a}\cdot\int_V\nabla v\,\mathrm{d}V=\boldsymbol{a}\cdot\int_S v\,\mathrm{d}\boldsymbol{S},$$

or

$$\boldsymbol{a}\cdot\left[\int_V\nabla v\,\mathrm{d}V-\int_S v\,\mathrm{d}\boldsymbol{S}\right]=0.$$

Since this holds for arbitrary vector fields of the form $\boldsymbol{v}=v(\boldsymbol{x})\boldsymbol{a}$, it must be true that $[\cdots]=0$, or

Theorem 16.3 (Alternative version of Gauss's theorem, 1)

$$\int_V \nabla v \, \mathrm{d}V = \int_S v \, \mathrm{d}\boldsymbol{S}.$$

Similarly, letting $\boldsymbol{v}(\boldsymbol{x}) = \boldsymbol{a} \times \boldsymbol{u}(\boldsymbol{x})$, where $\boldsymbol{a}$ is a constant vector, gives

Theorem 16.4 (Alternative version of Gauss's theorem, 2)

$$\int_V \nabla \times \boldsymbol{u} \, \mathrm{d}V = \int_S \mathrm{d}\boldsymbol{S} \times \boldsymbol{u}.$$

16.1.3 Worked examples

Example: Evaluate by using Gauss's theorem $\int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}$, where

$$\boldsymbol{v} = 8xz\hat{\boldsymbol{x}} + 2y^2\hat{\boldsymbol{y}} + 3yz\hat{\boldsymbol{z}}$$

and S is the surface of the unit cube in the positive octant, one of whose vertices lies at $(0, 0, 0)$.

We compute:

$$\begin{aligned}
\int_S \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} &= \int_V \mathrm{d}V \, \nabla \cdot \boldsymbol{v}, \\
&= \int_0^1 \mathrm{d}x \int_0^1 \mathrm{d}y \int_0^1 \mathrm{d}z \, (8z + 4y + 3y), \\
&= 1 \cdot 1 \cdot \int_0^1 8z \, \mathrm{d}z + 1 \cdot 1 \cdot \int_0^1 7y \, \mathrm{d}y, \\
&= 4 + \tfrac{7}{2} = \tfrac{15}{2}.
\end{aligned}$$

Example: A fluid is confined in a container of volume V with closed boundary S. The velocity of the fluid is $\boldsymbol{v}(\boldsymbol{x}, t)$. The velocity satisfies the so-called no-throughflow condition

$$\boldsymbol{v} \cdot \hat{\boldsymbol{n}} = 0, \text{ on } S,$$

where $\hat{\boldsymbol{n}}$ is the outward-pointing normal to the surface. Now suppose that a pollutant is introduced to the fluid, of concentration $C(\boldsymbol{x}, t)$. The pollutant must satisfy the equation

$$\frac{\partial C}{\partial t} + \nabla \cdot (\boldsymbol{v} C) = 0.$$

Prove that the total amount of pollutant,

$$P(t) = \int_V C(\boldsymbol{x}, t)\,\mathrm{d}V,$$

stays the same over time (hence P is in fact independent of time).

Proof: We have

$$\begin{aligned}\frac{dP}{dt} &= \frac{d}{dt}\int_V C(\boldsymbol{x}, t)\,\mathrm{d}V,\\ &= \int_V \frac{\partial C(\boldsymbol{x}, t)}{\partial t}\,\mathrm{d}V,\\ &= -\int_V \nabla \cdot (\boldsymbol{v}C)\ \mathrm{d}V,\\ &= -\int_S C(\boldsymbol{x} \in S, t)\boldsymbol{v}(\boldsymbol{x} \in S, t) \cdot \mathrm{d}\boldsymbol{S}.\end{aligned}$$

But

$$\hat{\boldsymbol{n}} \cdot \boldsymbol{v}|_{\boldsymbol{x}\in S} = 0,$$

hence

$$\frac{dP}{dt} = 0,$$

and the amount of pollutant P is constant ('conserved').

16.2 Stokes's Theorem

Theorem 16.5 *Let S be an open, two-sided surface bounded by a closed, non-intersecting curve C, and let $\boldsymbol{v}(\boldsymbol{x})$ be a vector field with continuous derivatives. Then,*

$$\boxed{\oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} = \int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S},} \tag{16.3}$$

where C is treated in the positive direction: an observer walking along the boundary of S, with his head pointing in the direction of the positive normal to S, has the surface on his left.

For the (S, C) pair to which the theorem refers, see Figure 16.2.

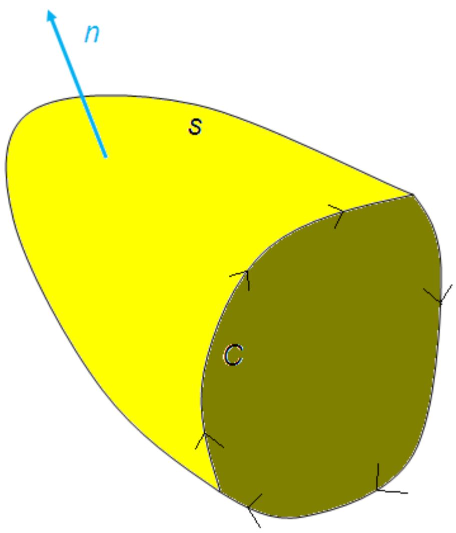
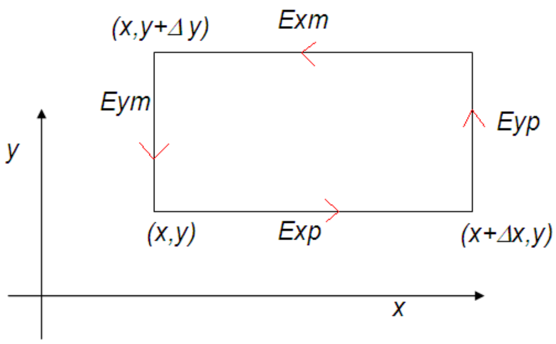


Figure 16.2: Stokes theorem: S is a surface; C is its boundary. The boundary can be given a definite orientation so the curve is called **two-sided**.

Proof: First, consider a rectangle in the x-y plane of sides of length Δx and Δy, with one vertex positioned at (x, y) – see the figure on the right. Label the edges Exp, Exm, Eyp, and Eym. We compute

$$\sum_{\text{all edges}} \boldsymbol{v} \cdot \Delta \boldsymbol{x},$$

where $\Delta \boldsymbol{x}$ is the line element on each edge, and we compute in an anticlockwise sense.



For example, in the x-direction, along Exp we have $\mathrm{d}\boldsymbol{x} = \hat{\boldsymbol{x}}\mathrm{d}x$ and along Exm we have $\mathrm{d}\boldsymbol{x} = -\hat{\boldsymbol{x}}\mathrm{d}x$. Adding up these contributions to $\boldsymbol{v} \cdot \Delta \boldsymbol{x}$ gives

$$[v_1(x, y, z) - v_1(x, y + \Delta y, z)]\, \Delta x.$$

Similarly, the contributions along Eyp and Eym give

$$[v_2(x + \Delta x, y, z) - v_2(x, y, z)]\, \Delta y.$$

Summing over these four contributions (i.e. summing over the four edges), we have

$$\sum_{\text{all edges}} \boldsymbol{v} \cdot \Delta \boldsymbol{x} = [v_1(x, y) - v_1(x, y + \Delta y)]\, \Delta x + [v_2(x + \Delta x, y) - v_2(x, y)]\, \Delta y$$

We apply Taylor's theorem to these increments and omit terms that are $O(\Delta x^2, \Delta y^2)$. This procedure is rigorous in the limit as the rectangle area goes to zero. We obtain

$$\begin{aligned} \sum_{\text{all edges}} \boldsymbol{v} \cdot \Delta \boldsymbol{x} &= [v_1(x, y) - v_1(x, y + \Delta y)]\, \Delta x + [v_2(x + \Delta x, y) - v_2(x, y)]\, \Delta y \\ &= \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right)_{(x,y)} \Delta x \Delta y. \end{aligned}$$

We take the limit as Δx and $\Delta y \to 0$, using $\Delta x \to \mathrm{d}x$ and $\Delta y \to \mathrm{d}y$ to denote this process. The corresponding infinitesimal patch of surface area is $\mathrm{d}\boldsymbol{S} = \hat{\boldsymbol{z}}\, \mathrm{d}x\, \mathrm{d}y$ pointing out of the page, hence

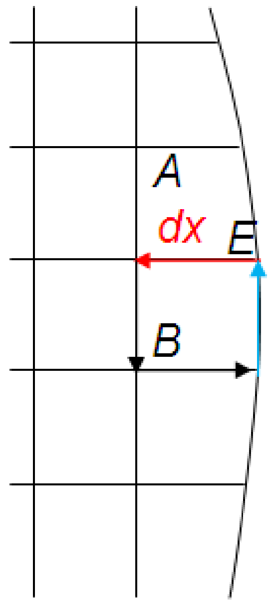
$$\sum_{\text{all edges}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} = (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S}. \tag{16.4}$$

As the final result involves only dot products and cross products, Equation (16.4) is independent of the Cartesian coordinate frame used in the derivation. Thus, Equation (16.4) is valid for an arbitrary patch of surface area in any plane.

For the second and final step, consider a surface S with boundary C. We break this surface up into many infinitesimally small rectangles. Each rectangle is a patch of surface area that lies in an arbitrary plane. The previous result (Equation (16.4)) is applied to all such patches, such that

$$\sum_{\text{all rectangles}} (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} = \sum_{\text{all rectangles}} \left(\sum_{\text{all edges}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} \right). \tag{16.5}$$

Consider, however, two neighbouring rectangles – see the accompanying figure. Call them A and B These will share a common edge, E, with line element $\mathrm{d}\boldsymbol{x}$. Rectangle A gives a contribution a, say, to the sum (16.2), while rectangle B must give a contribution $-a$. The only place where such a cancellation cannot occur is on exterior edges. Thus,



$$\sum_{\text{all rectangles}} (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} = \sum_{\text{all exterior edges}} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}.$$

But the rectangle areas are infinitesimally small, so this sum converts into an integral:

$$\int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} = \oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}. \quad \blacksquare$$

Example: Given a vector $\boldsymbol{v} = -\hat{\boldsymbol{x}}y + \hat{\boldsymbol{y}}x$, using Stokes's theorem, show that the integral around a continuous closed curve in the xy plane

$$\tfrac{1}{2}\oint \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} = \tfrac{1}{2}\oint (x\,\mathrm{d}y - y\,\mathrm{d}x) = S,$$

the area enclosed by the curve.

Proof:

$$\begin{aligned}
\tfrac{1}{2}\oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} &= \tfrac{1}{2}\int_S [\nabla \times (-\hat{\boldsymbol{x}}y + \hat{\boldsymbol{y}}x)] \cdot \mathrm{d}\boldsymbol{S}, \\
&= \tfrac{1}{2}\int_S (2\hat{\boldsymbol{z}}) \cdot \mathrm{d}\boldsymbol{S}, \\
&= \tfrac{1}{2}\int_S (2\hat{\boldsymbol{z}}) \cdot (\mathrm{d}x\,\mathrm{d}y\,\hat{\boldsymbol{z}}), \\
&= \int_S \mathrm{d}x\,\mathrm{d}y = S.
\end{aligned}$$

16.2.1 Green's theorem in the plane

The previous example hints at the following result: let S be a patch of area entirely contained in the xy plane, with boundary C, and let $\boldsymbol{v} = (v_1(x,y), v_2(x,y), 0)$ be a smooth vector field. Then,

$$\begin{aligned}\int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} &= \int_S (\nabla \times \boldsymbol{v}) \cdot (\mathrm{d}x\,\mathrm{d}y\,\hat{\boldsymbol{z}}), \\ &= \int_S \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y.\end{aligned}$$

But by Stokes's theorem,

$$\begin{aligned}\int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} &= \int_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}, \\ &= \int_C (v_1 \mathrm{d}x + v_2 \mathrm{d}y).\end{aligned}$$

Putting these equations together, we have Green's theorem in the plane:

Theorem 16.6

$$\int_S \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y = \int_C (v_1 \mathrm{d}x + v_2 \mathrm{d}y).$$

16.3 Potential theory

A vector field $\boldsymbol{v}$ is irrotational if and only if

- $\nabla \times \boldsymbol{v} = 0$ if and only if
- $\boldsymbol{v} = -\nabla \mathcal{U}$ if and only if
- The line integral $\int_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}$ depends only on the initial and final points of the path C and is independent of the details of the path between these terminal points.

Proving that $\boldsymbol{v} = -\nabla \mathcal{U} \implies \nabla \times \boldsymbol{v} = 0$ was trivial and we have done this already. Until now, we have been unable to prove the converse, namely that $\nabla \times \boldsymbol{v} \implies \boldsymbol{v} = -\nabla \mathcal{U}$. Let us do so now.

Consider an open subset $\Omega \in \mathbb{R}^3$ that is **simply connected**, i.e. contains no 'holes'. Let us take an arbitrary closed, smooth curve C in Ω. Because Ω is simply connected, it is possible to find a surface S that lies entirely in Ω, such that (S, C) have the properties mentioned in Stokes's theorem. Suppose now that $\nabla \times \boldsymbol{v} = 0$ for **all points** $\boldsymbol{x} \in \Omega$. Now, by Stokes's theorem,

$$\begin{aligned}0 &= \int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S}, \\ &= \oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}.\end{aligned}$$

This last result is true for **all closed, piecewise smooth contours** in the domain Ω. The only way for this relationship to be satisfied for **all contours** is if $\boldsymbol{v} = -\nabla\mathcal{U}$, for some function $\mathcal{U}(\boldsymbol{x})$, since then,

$$
\begin{aligned}
\oint_C \boldsymbol{v}\cdot \mathrm{d}\boldsymbol{x} &= -\oint_C (\nabla\mathcal{U})\cdot \mathrm{d}\boldsymbol{x},\\
&= -\left[\mathcal{U}(a) - \mathcal{U}(a)\right],\\
&= 0,
\end{aligned}
$$

for some reference point a on the contour C. Thus, we have proved that a vector field $\boldsymbol{v}$ is irrotational **if and only if** $\boldsymbol{v} = -\nabla\mathcal{U}$.

Simple-connectedness

Simple-connectedness will not be an issue in this module, as we usually work with vector fields defined on the whole of $\mathbb{R}^3$. On the other hand, it is not hard to find a domain Ω that is not simply connected. For example, consider a portion of the xy plane with a hole (Figure 16.3). The closed

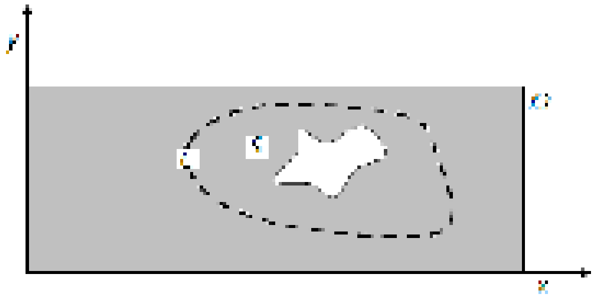


Figure 16.3: The set Ω is not simply connected.

curve C surrounds a region S; however, S is not contained entirely in Ω. We have knowledge of $\nabla\times\boldsymbol{v}$ only in Ω; we are unable to say anything about $\nabla\times\boldsymbol{v}$ in certain parts of the region S, and are therefore unable to apply the arguments of Stokes's theorem to this particular (S, C) pair. Again, it is not hard to find examples of such domains: imagine the domain of the vector field for flow over an aerofoil: such a domain is obviously not simply connected.

A more precise definition of simple-connectedness than the vague condition that 'the set should contain no holes' is the following: for any two closed paths $C_0 : [0,1] \to \Omega$, $C_1 : [0,1] \to \Omega$ based at $\boldsymbol{x}_0$, i.e.

$$\boldsymbol{x}_{C_0}(0) = \boldsymbol{x}_{C_1}(0) = \boldsymbol{x}_0,$$

there exists a continuous map

$$H : [0,1]\times[0,1] \to \Omega,$$

such that

$$\begin{aligned} H(t,0) &= \boldsymbol{x}_{C_0}(t), \qquad 0 \le t \le 1, \\ H(t,1) &= \boldsymbol{x}_{C_1}(t), \qquad 0 \le t \le 1, \\ H(0,s) &= H(1,s) = \boldsymbol{x}_0, \qquad 0 \le s \le 0. \end{aligned}$$

Such a map is called a **homotopy** and C_0 and C_1 are called homotopy equivalent. One can think of this map as a 'continuous deformation of one loop into another'. Because a point is, trivially, a loop, in a simply-connected set, a loop can be continuously deformed into a point. Note in the example Figure 16.3, the loop C cannot be continuously deformed into a point without leaving the set Ω. This is a more relational - or topological way - of describing the 'hole' in the set in Figure 16.3.

16.3.1 Worked examples

Example: In thermodynamics, the energy of a system of gas particles is expressed in differential form:

$$A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y,$$

where

- A is the temperature;
- B is minus the pressure;
- x has the interpretation of entropy;
- y has the interpretation of container volume.

The temperature and the pressure are known to satisfy the following relation:

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

Prove that for any closed path C in xy-space (i.e. in entropy/volume-space),

$$\oint_C [A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y] = 0.$$

Proof: We may regard

$$\boldsymbol{v}(x,y) = (A(x,y), B(x,y))$$

as a vector field, and we may take

$$\mathrm{d}\boldsymbol{S} = \mathrm{d}x\,\mathrm{d}y\hat{\boldsymbol{z}}$$

as an area element, pointing out of the xy-plane. Now let S be the patch of area in xy space enclosed by the curve C. We have

$$\begin{aligned}
\int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} &= \int_S \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y, \\
&= \int_S \left[\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right] \mathrm{d}x\,\mathrm{d}y, \\
&= \int_S \left(\frac{\partial A}{\partial y} - \frac{\partial A}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y, \\
&= 0.
\end{aligned}$$

But by Stokes's theorem,

$$\begin{aligned}
0 &= \int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S}, \\
&= \int_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}, \\
&= \int_C \left[A\mathrm{d}x + B\mathrm{d}y \right],
\end{aligned}$$

as required. Because $A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y$ integrates to zero when the integral is a closed contour, there exists a potential $E(x,y)$, such that

$$\mathrm{d}E = A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y.$$

The function E is called the **thermodynamic energy**. The integral of $\mathrm{d}E$ around a closed path is identically zero, and **the energy is path-independent**.

In general, the **differential form**

$$A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y$$

is **exact** if and only if

- There is a function $\phi(x,y)$, such that

$$A(x,y)\mathrm{d}x + B(x,y)\mathrm{d}y = \frac{\partial \phi}{\partial x}\mathrm{d}x + \frac{\partial \phi}{\partial y}\mathrm{d}y := \mathrm{d}\phi,$$

  if and only if

- The following relation holds:

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

Example: In mechanics, particles experience a **force field** $\boldsymbol{F}(\boldsymbol{x})$. The force is called conservative if a potential function exists:

$$\boldsymbol{F} = -\nabla \mathcal{U}.$$

Thus, a force is conservative if and only if $\nabla \times \boldsymbol{F} = 0$.

Show that the three-dimensional gravitational force

$$\boldsymbol{F} = -\frac{\alpha \boldsymbol{r}}{|\boldsymbol{r}|^3}$$

is a conservative force, where α is a positive constant.

Solution: We compute $\nabla \times \boldsymbol{F}$ by application of the following chain rule:

$$\nabla \times (\phi \boldsymbol{u}) = \phi \nabla \times \boldsymbol{u} + \nabla \phi \times \boldsymbol{u},$$

and we take $\phi = r^{-3}$ and $\boldsymbol{u} = \boldsymbol{r}$:

$$\nabla \times \boldsymbol{F} = -\alpha \left\{ \frac{1}{r^3} (\nabla \times \boldsymbol{r}) + \left[\nabla \left(r^{-3} \right) \right] \times \boldsymbol{r} \right\},$$

Now

$$\nabla \times \boldsymbol{r} = \nabla \times \left(\tfrac{1}{2} \nabla r^2 \right) = 0.$$

Also,

$$\nabla r^{-3} = -\frac{3\boldsymbol{r}}{r^5},$$

Hence,

$$\begin{aligned}\nabla\times\boldsymbol{F} &= -\alpha\left[\frac{1}{r^3}\nabla\times\boldsymbol{r}-\left(\nabla r^{-3}\right)\times\boldsymbol{r}\right],\\ &= -\alpha\left[0-\left(\frac{3\boldsymbol{r}}{r^5}\right)\times\boldsymbol{r}\right],\\ &= 0.\end{aligned}$$

Thus, both contributions to $\nabla\times\boldsymbol{F}$ are zero, so $\nabla\times\boldsymbol{F}=0$, and gravity is conservative.

See if you can show that

$$\mathcal{U}=-\frac{\alpha}{r}$$

is a suitable potential, $\boldsymbol{F}=-\nabla\left(-\alpha r^{-1}\right)$.

Example: Show that the force

$$\boldsymbol{F}=\alpha(x^2\hat{\boldsymbol{x}}+y\hat{\boldsymbol{y}})$$

is a conservative force and construct its potential.

Solution: We have

$$\nabla\times\boldsymbol{F}=\alpha\begin{vmatrix}\hat{\boldsymbol{x}} & \hat{\boldsymbol{y}} & \hat{\boldsymbol{z}}\\ \partial_x & \partial_y & \partial_z\\ x^2 & y & 0\end{vmatrix}=\alpha\hat{\boldsymbol{z}}\left(\partial_x y-\partial_y x^2\right)=0.$$

Next, we take

$$F_x=\alpha x^2=-\partial_x\mathcal{U}.$$

Ordinary integration gives

$$\mathcal{U}(x,y)=-\tfrac{1}{3}\alpha x^3+f(y),$$

where $f(y)$ is a function to be determined. But we also have

$$F_y=\alpha y=-\partial_y\mathcal{U},$$

which gives

$$\mathcal{U}(x,y)=-\tfrac{1}{2}\alpha y^2+g(x).$$

Putting these results together, we have

$$\mathcal{U}(x,y)=-\alpha\left(\tfrac{1}{3}x^3+\tfrac{1}{2}y^2\right)+\text{Const.},$$

and the constant is immaterial because only *gradients* of the potential are important.

Example: Recall that the vorticity $\boldsymbol{\omega}(\boldsymbol{x})$ measures the amount of swirl in a fluid velocity field $\boldsymbol{v}(\boldsymbol{x})$, $\boldsymbol{\omega} = \nabla \times \boldsymbol{v}$. Show that all irrotational flows

$$\boldsymbol{\omega} = 0,$$

are potential flows,

$$\boldsymbol{v} = \nabla\phi.$$

Show that the potential for an incompressible irrotational flow satisfies Laplace's equation:

$$\nabla \cdot \boldsymbol{v} = 0 \text{ and } \boldsymbol{\omega} = 0 \implies \nabla^2\phi = 0.$$

Solution: If the flow is irrotational, then $\nabla \times \boldsymbol{v} = 0$, which implies, by Stokes's theorem,

$$\boldsymbol{v} = \nabla\phi,$$

(note the sign), for some *velocity potential* ϕ. We are to assume that the flow is incompressible:

$$0 = \nabla \cdot \boldsymbol{v} = \nabla \cdot \nabla\phi = \nabla^2\phi.$$

Thus, an incompressible, irrotational flow satisfies

$$\nabla^2\phi = 0.$$

The study of the equation $\nabla^2\phi = 0$ is called **harmonic analysis**.

Chapter 17

Curvilinear coordinate systems

Overview and introduction

So far we have restricted ourselves to Cartesian coordinate systems. A Cartesian coordinate system offers a unique advantage in that the distinguished directions $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ all point in constant directions. However, many physical problems are not well suited to solution in Cartesian coordinates. For instance, in the atmosphere, fluid flow takes place on a sphere, and latitude and longitude are more appropriate labels for position in space. Such a problem naturally leads to the use of spherical polar coordinates. In fact, the coordinate system we use should be chosen to fit the problem in hand, and to exploit any type of symmetry or constraint therein. Then, hopefully, the problem will be more amenable to solution than if we had stubbornly persisted with the Cartesian framework.

Unfortunately, there is a high price to pay for this freedom of choice (for coordinate systems). In an arbitrary coordinate system, the distinguished directions are no longer constant, and the operators div, grad, and curl become very cumbersome. Nevertheless, we must be willing to pay the ultimate price for this freedom, and derive expressions for div, grad, and curl in **orthogonal curvilinear coordinate systems**.

17.1 Coordinate transformations

In three dimensions, three variables are necessary and sufficient to specify the location of a particle. We have used the Cartesian triple (x, y, z), where the equations $x = \mathrm{Const.}$, $y = \mathrm{Const.}$, and $z = \mathrm{Const.}$ describe three mutually perpendicular families of planes. Suppose now we superimpose on these planes a second family of surfaces. These surfaces need not be planes; nor need they be parallel. In the Cartesian framework, a point is specified by the intersection of the three planes; in the new framework, the same point is specified by the intersection of three surfaces. In the new framework, let the new surfaces be described by

$$q_1 = \mathrm{Const.}, \qquad q_2 = \mathrm{Const.}, \qquad q_3 = \mathrm{Const.}.$$

Because the point in question can be described adequately in both frameworks, as the point of intersection of three surfaces, we may write

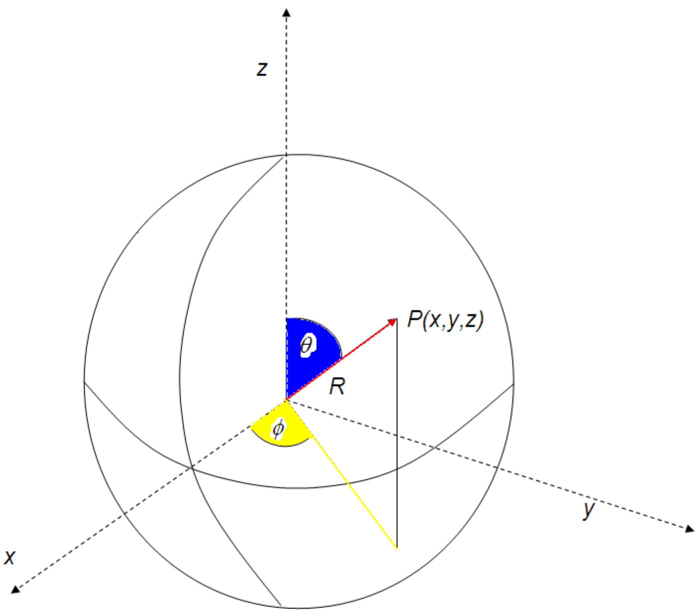
$$x = x(q_1, q_2, q_3), \qquad y = y(q_1, q_2, q_3), \qquad z = z(q_1, q_2, q_3),$$

and

$$q_1 = q_1(x, y, z), \qquad q_2 = q_2(x, y, z), \qquad q_3 = q_3(x, y, z),$$

where each function written here is assumed smooth. That is, there is a smooth, invertible map connecting the two coordinate systems. This map is called a **coordinate transformation**.

Example: Consider **spherical polar coordinates** as shown in the figure. Identify the three intersecting surfaces that make up the coordinate system.



The point P can either be labelled by the Cartesian triple (x, y, z), or by its radial distance R from the origin, together with two angles: the **azimuthal angle** and the **polar angle**. The azimuthal angle φ is the angle between the x-axis and the projection of the radius vector $\boldsymbol{x} \equiv \boldsymbol{r} \equiv \overrightarrow{OP}$ on to the x-y plane. The polar angle θ is the angle between the z-direction and the radius vector. Here are the surfaces generated by these new coordinates:

- The surface $R = \mathrm{Const.}$ is a **sphere** of radius R centred at O (q_1),
- The surface $\theta = \mathrm{Const.}$ is a **cone** whose tip lies at the origin O (q_2),
- The surface $\varphi = \mathrm{Const.}$ is a plane parallel to the z-axis, given by $y = x\tan\varphi$ (q_3).

The point P is given by the intersection of these surfaces, or by the intersection of the planes $x = \mathrm{Const.}$, $y = \mathrm{Const.}$, and $z = \mathrm{Const.}$ (See the figure below.)

These two coordinate systems are related through

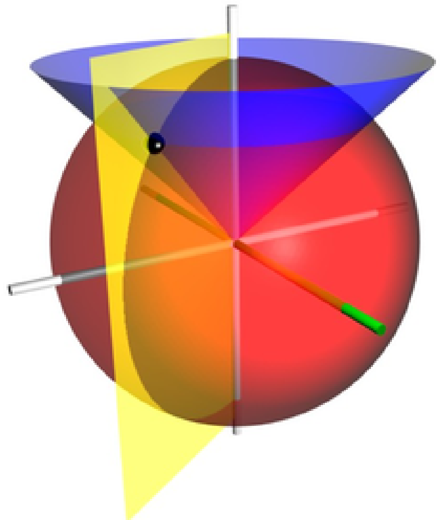
$$
\begin{aligned}
x &= r\sin\theta\cos\varphi,\\
y &= r\sin\theta\sin\varphi,\\
z &= r\cos\theta,
\end{aligned}
$$

with inverse transformation

$$
\begin{aligned}
r &= \sqrt{x^2+y^2+z^2},\\
\theta &= \cos^{-1}\left(z/r\right),\\
\varphi &= \tan^{-1}\left(y/x\right).
\end{aligned}
$$

Note: Particular care must be taken with the inverse $\tan^{-1}(y/x)$. Where necessary, we must add or subtract 2π to the answer to obtain an angle $\varphi \in [0, 2\pi)$.

Surfaces generated by spherical polar coodinates.[a]



[a]From http://en.wikipedia.org/wiki/Spherical_coordinates, 16th August 2010.

17.2 The line element, tangent vectors, scale factors

Recall, in a Cartesian frame, that a small increment of length $\mathrm{d}s$ is given by

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2.$$

The quantity $\mathrm{d}s$ is called the **line element**. Let us take a coordinate transformation

$$x = x(q_1,q_2,q_3), \qquad y = y(q_1,q_2,q_3), \qquad z = z(q_1,q_2,q_3),$$

and

$$q_1 = q_1(x,y,z), \qquad q_2 = q_2(x,y,z), \qquad q_3 = q_3(x,y,z),$$

and compute the line element in terms of the q_i's. This is possible because the line element exists independent of its description in Cartesian coordinates. We have,

$$\mathrm{d}x = \frac{\partial x}{\partial q_1}\mathrm{d}q_1 + \frac{\partial x}{\partial q_2}\mathrm{d}q_2 + \frac{\partial x}{\partial q_3}\mathrm{d}q_3,$$

and similarly for $\mathrm{d}y$ and $\mathrm{d}z$. Thus, in vector notation,

$$\mathrm{d}\boldsymbol{x} = \frac{\partial \boldsymbol{x}}{\partial q_1}\mathrm{d}q_1 + \frac{\partial \boldsymbol{x}}{\partial q_2}\mathrm{d}q_2 + \frac{\partial \boldsymbol{x}}{\partial q_3}\mathrm{d}q_3$$

Substitution of these **differentials** into the definition of the line element gives

$$\begin{aligned}
\mathrm{d}s^2 = \mathrm{d}\boldsymbol{x}\cdot\mathrm{d}\boldsymbol{x} &= \left(\frac{\partial \boldsymbol{x}}{\partial q_1}\mathrm{d}q_1 + \frac{\partial \boldsymbol{x}}{\partial q_2}\mathrm{d}q_2 + \frac{\partial \boldsymbol{x}}{\partial q_3}\mathrm{d}q_3\right)\cdot\left(\frac{\partial \boldsymbol{x}}{\partial q_1}\mathrm{d}q_1 + \frac{\partial \boldsymbol{x}}{\partial q_2}\mathrm{d}q_2 + \frac{\partial \boldsymbol{x}}{\partial q_3}\mathrm{d}q_3\right)\\
&= \left(\frac{\partial \boldsymbol{x}}{\partial q_1}\cdot\frac{\partial \boldsymbol{x}}{\partial q_1}\right)\mathrm{d}q_1^2 + \left(\frac{\partial \boldsymbol{x}}{\partial q_2}\cdot\frac{\partial \boldsymbol{x}}{\partial q_2}\right)\mathrm{d}q_2^2 + \left(\frac{\partial \boldsymbol{x}}{\partial q_3}\cdot\frac{\partial \boldsymbol{x}}{\partial q_3}\right)\mathrm{d}q_3^2\\
&\quad + \left(\frac{\partial \boldsymbol{x}}{\partial q_1}\cdot\frac{\partial \boldsymbol{x}}{\partial q_2}\right)\mathrm{d}q_1\,\mathrm{d}q_2 + \left(\frac{\partial \boldsymbol{x}}{\partial q_1}\cdot\frac{\partial \boldsymbol{x}}{\partial q_3}\right)\mathrm{d}q_1\,\mathrm{d}q_3 + \left(\frac{\partial \boldsymbol{x}}{\partial q_2}\cdot\frac{\partial \boldsymbol{x}}{\partial q_3}\right)\mathrm{d}q_2\,\mathrm{d}q_3\\
&\quad + \left(\frac{\partial \boldsymbol{x}}{\partial q_2}\cdot\frac{\partial \boldsymbol{x}}{\partial q_1}\right)\mathrm{d}q_2\,\mathrm{d}q_1 + \left(\frac{\partial \boldsymbol{x}}{\partial q_3}\cdot\frac{\partial \boldsymbol{x}}{\partial q_1}\right)\mathrm{d}q_3\,\mathrm{d}q_1 + \left(\frac{\partial \boldsymbol{x}}{\partial q_3}\cdot\frac{\partial \boldsymbol{x}}{\partial q_2}\right)\mathrm{d}q_3\,\mathrm{d}q_2.
\end{aligned}$$

Definition 17.1 (Metric tensor) *In more compact form, this is written as*

$$\begin{aligned}
\mathrm{d}s^2 &= g_{11}\mathrm{d}q_1^2 + g_{22}\mathrm{d}q_2^2 + g_{33}\mathrm{d}q_3^2\\
&\quad + g_{12}\mathrm{d}q_1\,\mathrm{d}q_2 + g_{13}\mathrm{d}q_1\,\mathrm{d}q_3 + g_{23}\mathrm{d}q_2\,\mathrm{d}q_3\\
&\quad + g_{21}\mathrm{d}q_2\,\mathrm{d}q_1 + g_{31}\mathrm{d}q_3\,\mathrm{d}q_1 + g_{32}\mathrm{d}q_3\,\mathrm{d}q_2.
\end{aligned}$$

and

$$g_{ij} = \frac{\partial \boldsymbol{x}}{\partial q_i}\cdot\frac{\partial \boldsymbol{x}}{\partial q_j} = \frac{\partial x}{\partial q_i}\frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i}\frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i}\frac{\partial z}{\partial q_j}$$

is called the **metric tensor**.

Hence,

$$\boxed{ds^2 = \sum_{ij} g_{ij}\,\mathrm{d}q_i\,\mathrm{d}q_j.} \tag{17.1}$$

The expression we have derived for the line element is clearly very complicated. Therefore, we restrict ourselves to **orthogonal coordinate systems**:

Definition 17.2 *A coordinate system is orthogonal if g_{ij} is a diagonal matrix.*

The reason for this nomenclature is clear: the vector

$$\frac{\partial \boldsymbol{x}}{\partial q_i} \tag{17.2}$$

is **normal** to the surface $q_i = \text{Constant}$. Thus, the coordinate surfaces are mutually perpendicular if

$$\left(\frac{\partial \boldsymbol{x}}{\partial q_i}\right) \cdot \left(\frac{\partial \boldsymbol{x}}{\partial q_j}\right) = 0, \qquad i \neq j,$$

in which case the metric tensor is diagonal. In this context, we actually call the vectors (17.2) the **tangent vectors** of the coordinate system, because $\partial \boldsymbol{x}/\partial q_1$ is tangent to the surfaces $q_2 =$ Constant and $q_3 =$ Constant etc. Restricting to such coordinate systems, the line element becomes

$$\mathrm{d}s^2 = g_{11}\mathrm{d}q_1^2 + g_{22}\mathrm{d}q_2^2 + g_{33}\mathrm{d}q_3^2,$$

or

$$\mathrm{d}s^2 = h_1^2\mathrm{d}q_1^2 + h_2^2\mathrm{d}q_2^2 + h_3^2\mathrm{d}q_3^2,$$

where

$$h_i = \sqrt{g_{ii}}, \qquad \text{no sum over } i$$

are the **scale factors of the orthogonal coordinate system**. Moreover, we have three mutually orthogonal vectors $\partial \boldsymbol{x}/\partial q_i$, which we may take to form a basis. Indeed, we take unit vectors

$$\hat{\boldsymbol{q}}_i = \frac{\partial \boldsymbol{x}}{\partial q_i} \bigg/ \left|\frac{\partial \boldsymbol{x}}{\partial q_i}\right| = \frac{1}{h_i}\frac{\partial \boldsymbol{x}}{\partial q_i},$$

and thus any vector $\boldsymbol{A}$ can be written as

$$\boldsymbol{A} = \hat{\boldsymbol{q}}_1 A_1 + \hat{\boldsymbol{q}}_2 A_2 + \hat{\boldsymbol{q}}_3 A_3,$$

where

$$A_i = \boldsymbol{A} \cdot \hat{\boldsymbol{q}}_i$$

is the component of the vector $\boldsymbol{A}$ in the $\hat{\boldsymbol{q}}_i$ direction (and NOT in any particular Cartesian direction).

Example: Consider spherical polar coordinates again, where

$$\begin{aligned} x &= r\sin\theta\cos\varphi, \\ y &= r\sin\theta\sin\varphi, \\ z &= r\cos\theta, \end{aligned}$$

with inverse transformation

$$\begin{aligned} r &= \sqrt{x^2+y^2+z^2}, \\ \theta &= \cos^{-1}(z/r), \\ \varphi &= \tan^{-1}(y/x). \end{aligned}$$

Show that this is an orthogonal coordinate system and compute the metric tensor.

Let take the position vector

$$\boldsymbol{x} = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z,$$

and compute the tangent vectors:

$$\begin{aligned} \frac{\partial \boldsymbol{x}}{\partial r} &= \hat{\boldsymbol{x}}\frac{\partial x}{\partial r} + \hat{\boldsymbol{y}}\frac{\partial x}{\partial r} + \hat{\boldsymbol{z}}\frac{\partial x}{\partial r}, \\ &= \hat{\boldsymbol{x}}\frac{\partial}{\partial r}(r\sin\theta\cos\varphi) + \hat{\boldsymbol{x}}\frac{\partial}{\partial r}(r\sin\theta\sin\varphi) + \hat{\boldsymbol{x}}\frac{\partial}{\partial r}(r\cos\theta), \\ &= \hat{\boldsymbol{x}}\sin\theta\cos\varphi + \hat{\boldsymbol{y}}\sin\theta\sin\varphi + \hat{\boldsymbol{z}}\cos\theta. \end{aligned}$$

$$\begin{aligned} \frac{\partial \boldsymbol{x}}{\partial \theta} &= \hat{\boldsymbol{x}}\frac{\partial x}{\partial \theta} + \hat{\boldsymbol{y}}\frac{\partial x}{\partial \theta} + \hat{\boldsymbol{z}}\frac{\partial x}{\partial \theta}, \\ &= \hat{\boldsymbol{x}}\frac{\partial}{\partial \theta}(r\sin\theta\cos\varphi) + \hat{\boldsymbol{y}}\frac{\partial}{\partial \theta}(r\sin\theta\sin\varphi) + \hat{\boldsymbol{z}}\frac{\partial}{\partial \theta}(r\cos\theta), \\ &= r\left[\hat{\boldsymbol{x}}\cos\theta\cos\varphi + \hat{\boldsymbol{y}}\cos\theta\sin\varphi - \hat{\boldsymbol{z}}\sin\theta\right]. \end{aligned}$$

$$\begin{aligned} \frac{\partial \boldsymbol{x}}{\partial \varphi} &= \hat{\boldsymbol{x}}\frac{\partial x}{\partial \varphi} + \hat{\boldsymbol{y}}\frac{\partial y}{\partial \varphi} + \hat{\boldsymbol{z}}\frac{\partial z}{\partial \varphi}, \\ &= \hat{\boldsymbol{x}}\frac{\partial}{\partial \varphi}(r\sin\theta\cos\varphi) + \hat{\boldsymbol{y}}\frac{\partial}{\partial \varphi}(r\sin\theta\sin\varphi) + \hat{\boldsymbol{z}}\frac{\partial}{\partial \varphi}(r\cos\theta), \\ &= r\left[-\hat{\boldsymbol{x}}\sin\theta\sin\varphi + \hat{\boldsymbol{y}}\sin\theta\cos\varphi\right]. \end{aligned}$$

Now compute

$$\begin{aligned} \left(\frac{\partial \boldsymbol{x}}{\partial r}\right)\cdot\left(\frac{\partial \boldsymbol{x}}{\partial \theta}\right) &= r\left(\hat{\boldsymbol{x}}\sin\theta\cos\varphi + \hat{\boldsymbol{y}}\sin\theta\sin\varphi + \hat{\boldsymbol{z}}\cos\theta\right)\cdot\left(\hat{\boldsymbol{x}}\cos\theta\cos\varphi + \hat{\boldsymbol{y}}\cos\theta\sin\varphi - \hat{\boldsymbol{z}}\sin\theta\right), \\ &= r\left[\sin\theta\cos\theta\cos^2\varphi + \sin\theta\cos\theta\sin^2\varphi - \sin\theta\cos\theta\right] = 0. \end{aligned}$$

$$\begin{aligned} \left(\frac{\partial \boldsymbol{x}}{\partial r}\right)\cdot\left(\frac{\partial \boldsymbol{x}}{\partial \varphi}\right) &= r\left(\hat{\boldsymbol{x}}\sin\theta\cos\varphi + \hat{\boldsymbol{y}}\sin\theta\sin\varphi + \hat{\boldsymbol{z}}\cos\theta\right)\cdot\left(-\hat{\boldsymbol{x}}\sin\theta\sin\varphi + \hat{\boldsymbol{y}}\sin\theta\cos\varphi\right), \\ &= r\left[-\sin^2\theta\sin\varphi\cos\varphi + \sin^2\theta\sin\varphi\cos\varphi\right] = 0. \end{aligned}$$

$$
\begin{aligned}
\left(\frac{\partial \boldsymbol{x}}{\partial \varphi}\right) \cdot \left(\frac{\partial \boldsymbol{x}}{\partial \theta}\right) &= r^2 \left(-\hat{\boldsymbol{x}} \sin\theta \sin\varphi + \hat{\boldsymbol{y}} \sin\theta \cos\varphi\right) \cdot \left(\hat{\boldsymbol{x}} \cos\theta \cos\varphi + \hat{\boldsymbol{y}} \cos\theta \sin\varphi - \hat{\boldsymbol{z}} \sin\theta\right), \\
&= r^2 \left[-\sin\theta\cos\theta\sin\varphi\cos\varphi + \sin\theta\cos\theta\sin\varphi\cos\varphi\right] = 0,
\end{aligned}
$$

and the coordinate system is orthogonal. Now we compute the scale factors:

$$
\begin{aligned}
h_r^2 &= \left(\frac{\partial \boldsymbol{x}}{\partial r}\right) \cdot \left(\frac{\partial \boldsymbol{x}}{\partial r}\right), \\
&= \left(\hat{\boldsymbol{x}} \sin\theta \cos\varphi + \hat{\boldsymbol{y}} \sin\theta \sin\varphi + \hat{\boldsymbol{z}} \cos\theta\right) \cdot \left(\hat{\boldsymbol{x}} \sin\theta \cos\varphi + \hat{\boldsymbol{y}} \sin\theta \sin\varphi + \hat{\boldsymbol{z}} \cos\theta\right), \\
&= 1.
\end{aligned}
$$

$$
\begin{aligned}
h_\theta^2 &= \left(\frac{\partial \boldsymbol{x}}{\partial \theta}\right) \cdot \left(\frac{\partial \boldsymbol{x}}{\partial \theta}\right), \\
&= r^2 \left(\hat{\boldsymbol{x}} \cos\theta \cos\varphi + \hat{\boldsymbol{y}} \cos\theta \sin\varphi - \hat{\boldsymbol{z}} \sin\theta\right) \cdot \left(\hat{\boldsymbol{x}} \cos\theta \cos\varphi + \hat{\boldsymbol{y}} \cos\theta \sin\varphi - \hat{\boldsymbol{z}} \sin\theta\right), \\
&= r^2
\end{aligned}
$$

$$
\begin{aligned}
h_\varphi^2 &= \left(\frac{\partial \boldsymbol{x}}{\partial \varphi}\right) \cdot \left(\frac{\partial \boldsymbol{x}}{\partial \varphi}\right), \\
&= r^2 \left(-\hat{\boldsymbol{x}} \sin\theta \sin\varphi + \hat{\boldsymbol{y}} \sin\theta \cos\varphi\right) \cdot \left(-\hat{\boldsymbol{x}} \sin\theta \sin\varphi + \hat{\boldsymbol{y}} \sin\theta \cos\varphi\right), \\
&= r^2 \sin^2\theta.
\end{aligned}
$$

Thus, spherical polar coordinates are orthogonal, the line element is

$$
\mathrm{d}s^2 = \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2 + r^2 \sin^2\theta \mathrm{d}\varphi^2,
$$

and the unit vectors are

$$
\begin{aligned}
\hat{\boldsymbol{r}} &= \hat{\boldsymbol{x}} \sin\theta \cos\varphi + \hat{\boldsymbol{y}} \sin\theta \sin\varphi + \hat{\boldsymbol{z}} \cos\theta, \\
\hat{\boldsymbol{\theta}} &= \hat{\boldsymbol{x}} \cos\theta \cos\varphi + \hat{\boldsymbol{y}} \cos\theta \sin\varphi - \hat{\boldsymbol{z}} \sin\theta. \\
\hat{\boldsymbol{\varphi}} &= -\hat{\boldsymbol{x}} \sin\varphi + \hat{\boldsymbol{y}} \cos\varphi.
\end{aligned}
$$

These unit vectors point in the directions of increasing r, φ, and θ, respectively (Figure 17.1).

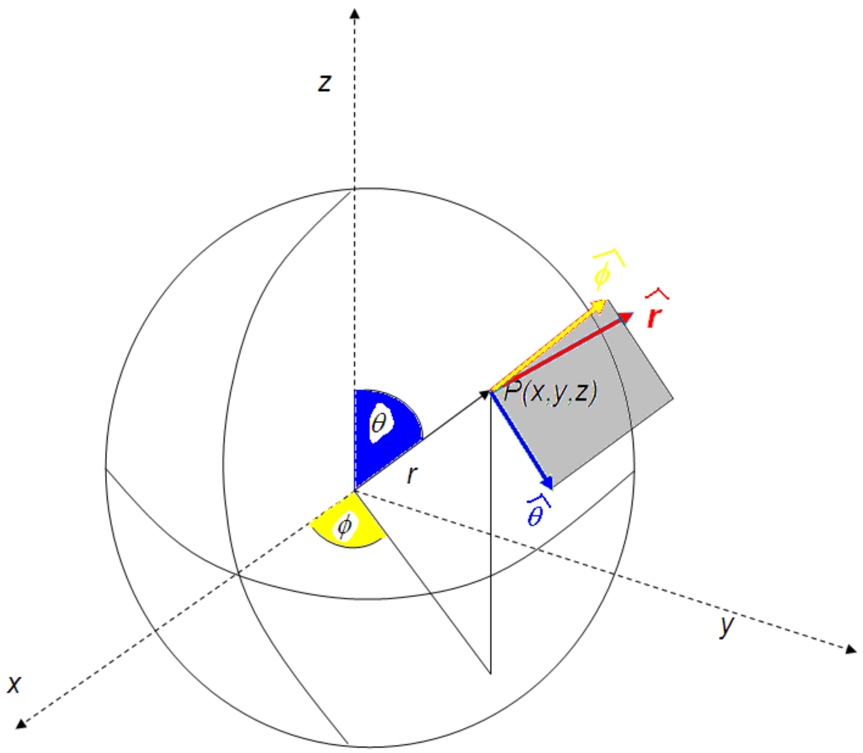


Figure 17.1: The unit vectors for spherical polar coordinates

Note that the unit vectors, although of constant magnitude, vary in direction as the point P is varied. They are not constant vectors, and do not go to zero when differentiated. It is for this reason that developing expressions for div, grad, and curl in curvilinear coordinates is complicated. It is to this issue that we now turn.

17.3 Grad, div, and curl in curvilinear coordinate systems

To avoid confusion, in this section we use the notation ψ for scalar fields. The use of φ to label a function is avoided because it is conventional to use this symbol for the azimuthal coordinate in the spherical polar system.

17.3.1 The gradient

Because $\hat{\boldsymbol{q}}_i$ form an orthogonal basis, any vector (such as $\nabla\psi$) can be written as

$$\nabla\psi = \sum_{i=1}^{3} \hat{\boldsymbol{q}}_i \left[(\nabla\psi) \cdot \hat{\boldsymbol{q}}_i \right].$$

Now consider $(\nabla\psi)\cdot\hat{\boldsymbol{q}}_i$. This is nothing other than the directional derivative of ψ in the q_i-direction:

$$(\nabla\psi)\cdot\hat{\boldsymbol{q}}_i=\lim_{t\to0}\left[\frac{\psi(\boldsymbol{x}+t\hat{\boldsymbol{q}}_i)-\psi(\boldsymbol{x})}{t}\right].$$

For definiteness we focus on $(\nabla\psi)\cdot\hat{\boldsymbol{q}}_1$ and we identify the fixed point $\boldsymbol{x}=\boldsymbol{x}(q_1,q_2,q_3)$. Then, the point $\boldsymbol{x}+t\hat{\boldsymbol{q}}_1$ is identified with the neighbouring point $\boldsymbol{x}(q_1+\delta q_1,q_2+\delta q_2,q_3+\delta q_3)$. Thus,

$$\begin{aligned}
\boldsymbol{x}+t\hat{\boldsymbol{q}}_1 &= \boldsymbol{x}(q_1+\delta q_1,q_2+\delta q_2,q_3+\delta q_3),\\
&= \boldsymbol{x}(q_1,q_2,q_3)+\left(\frac{\partial\boldsymbol{x}}{\partial q_1}\right)_{(q_1,q_2,q_3)}\delta q_1+\left(\frac{\partial\boldsymbol{x}}{\partial q_2}\right)_{(q_1,q_2,q_3)}\delta q_2+\left(\frac{\partial\boldsymbol{x}}{\partial q_3}\right)_{(q_1,q_2,q_3)}\delta q_3,\\
&= \boldsymbol{x}+h_1\hat{\boldsymbol{q}}_1\delta q_1+h_2\hat{\boldsymbol{q}}_2\delta q_2+h_3\hat{\boldsymbol{q}}_3\delta q_3,
\end{aligned}$$

hence

$$t\hat{\boldsymbol{q}}_1=h_1\hat{\boldsymbol{q}}_1\delta q_1+h_2\hat{\boldsymbol{q}}_2\delta q_2+h_3\hat{\boldsymbol{q}}_3\delta q_3.$$

Thus, $\delta q_2=\delta q_3=0$, and $t=h_1\delta q_1$. Hence,

$$\begin{aligned}
(\nabla\psi)\cdot\hat{\boldsymbol{q}}_1 &= \lim_{t\to0}\left[\frac{\psi(\boldsymbol{x}+t\hat{\boldsymbol{q}}_i)-\psi(\boldsymbol{x})}{t}\right],\\
&= \lim_{t\to0}\left[\frac{\psi(q_1+\delta q_1,q_2,q_3)-\psi(q_1,q_2,q_3)}{t}\right],\\
&= \lim_{\delta q_1\to0}\left[\frac{\psi(q_1+\delta q_1,q_2,q_3)-\psi(q_1,q_2,q_3)}{h_1\delta q_1}\right],\\
&= \frac{1}{h_1}\frac{\partial\psi}{\partial q_1},
\end{aligned}$$

and similarly for the other directions $i=2,3$:

$$(\nabla\psi)\cdot\hat{\boldsymbol{q}}_i=\frac{1}{h_i}\frac{\partial\psi}{\partial q_i},\qquad i=1,2,3.$$

Hence,

$$\nabla\psi=\sum_{i=1}^{3}\frac{\hat{\boldsymbol{q}}_i}{h_i}\frac{\partial\psi}{\partial q_i},$$

repeated as follows as a theorem:

Theorem 17.1 (The gradient operator in orthogonal curvilinear coordinates)

$$\boxed{\nabla\psi(q_1,q_2,q_3)=\frac{\hat{\boldsymbol{q}}_1}{h_1}\frac{\partial\psi}{\partial q_1}+\frac{\hat{\boldsymbol{q}}_2}{h_2}\frac{\partial\psi}{\partial q_2}+\frac{\hat{\boldsymbol{q}}_3}{h_3}\frac{\partial\psi}{\partial q_3}.}\qquad(17.3)$$

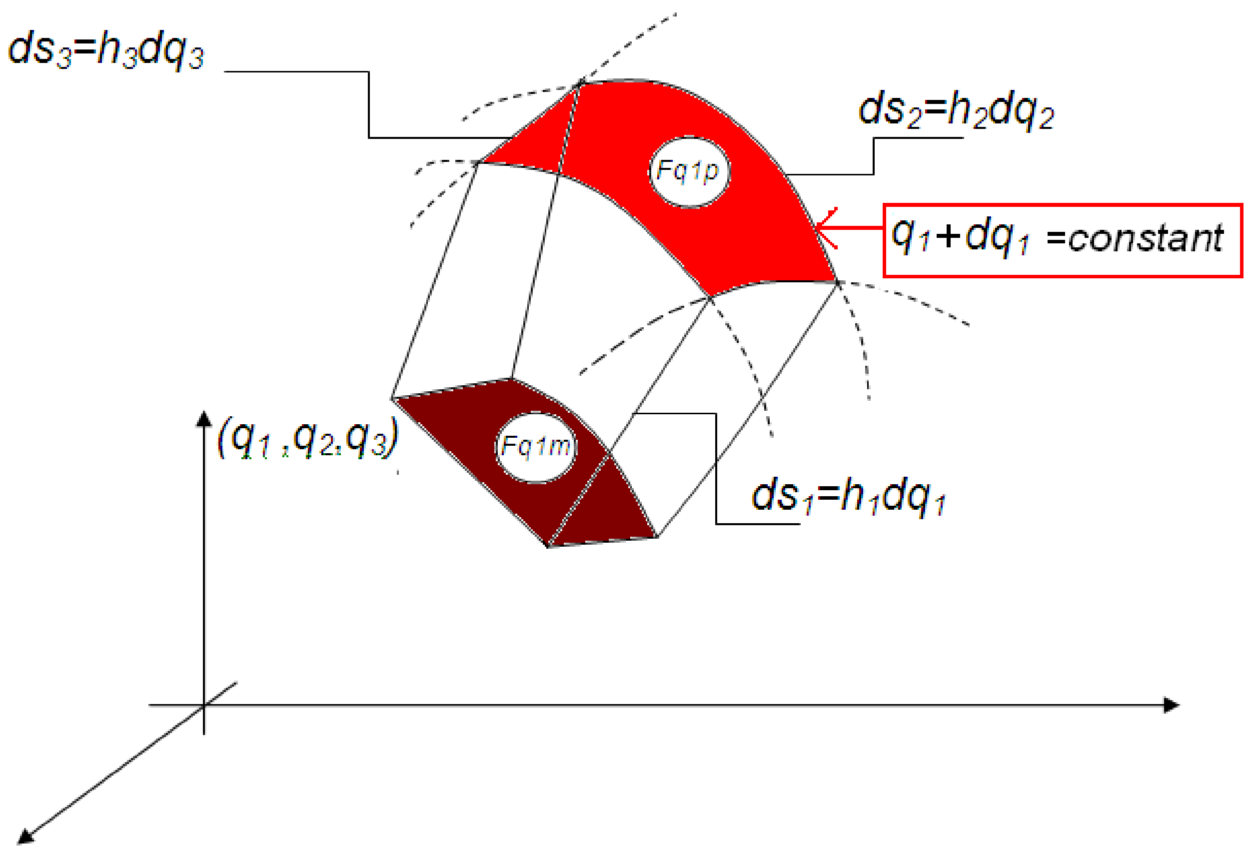


Figure 17.2: The volume element in curvilinear coordinates: this sketch forms a basis for deriving div and grad in curvilinear coordinates.

17.3.2 The divergence

Recall Gauss's theorem: In three dimensions, given a vector field $\boldsymbol{v}(\boldsymbol{x})$ and a volume V with bounding surface S,

$$\int_V \nabla\cdot\boldsymbol{v}\,\mathrm{d}V = \int_S \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{S}.$$

Here, we view Gauss's theorem as a *definition* of divergence:

$$\nabla\cdot\boldsymbol{v} = \lim_{\int_V \mathrm{d}V\to 0}\frac{\int_V \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{S}}{\int_V \mathrm{d}V}. \tag{17.4}$$

Thus,

$$\nabla\cdot\boldsymbol{v}(q_1,q_2,q_3) = \lim_{\int_V \mathrm{d}V\to 0}\frac{\int_V \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{S}}{\int_V \mathrm{d}V},\qquad \mathrm{d}V = h_1h_2h_3\,\mathrm{d}q_1\,\mathrm{d}q_2\,\mathrm{d}q_3.$$

Refer to Figure 17.2: we compute the area integrals associated with a small parallelepiped formed by the intersection of 6 surfaces,

$$q_1 = \text{Constant},\qquad q_1+\mathrm{d}q_1 = \text{Constant},\qquad \text{etc.}$$

On the face labelled F_{q1p} in Figure 17.2, we have

$$\begin{aligned} \mathrm{d}\boldsymbol{S} &= \left[\frac{\partial \boldsymbol{x}}{\partial q_2} \times \frac{\partial \boldsymbol{x}}{\partial q_3}\right]_{(q_1+\mathrm{d}q_1, q_2, q_3)} \mathrm{d}q_2 \mathrm{d}q_3, \\ &= \left[h_2 h_3 \left(\hat{\boldsymbol{q}}_2 \times \hat{\boldsymbol{q}}_3\right) \mathrm{d}q_2 \mathrm{d}q_3\right]_{(q_1+\mathrm{d}q_1, q_2, q_3)}, \\ &= \hat{\boldsymbol{q}}_1 h_2 h_3 \mathrm{d}q_2 \mathrm{d}q_3\big|_{(q_1+\mathrm{d}q_1, q_2, q_3)}. \end{aligned}$$

Hence,

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}\bigg|_{F_{q1p}} = (v_1 h_2 h_3)(q_1 + \mathrm{d}q_1, q_2, q_3) dq_2 dq_3.$$

Similarly, on the face labelled F_{q1m}, we have

$$\mathrm{d}\boldsymbol{S} = -\hat{\boldsymbol{q}}_1 h_2 h_3 \mathrm{d}q_2 \mathrm{d}q_3\big|_{(q_1, q_2, q_3)}.$$

Hence,

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S}\bigg|_{F_{q1m}} = -(v_1 h_2 h_3)(q_1, q_2, q_3) dq_2 dq_3.$$

Adding these contributions gives

$$\left[(v_1 h_2 h_3)(q_1 + \mathrm{d}q_1, q_2, q_3) - (v_1 h_2 h_3)(q_1, q_2, q_3)\right] \mathrm{d}q_2\, \mathrm{d}q_3 = \frac{\partial}{\partial q_1}(v_1 h_2 h_3)\, \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3.$$

Adding up the other contributions gives

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \left[\frac{\partial}{\partial q_1}(v_1 h_2 h_3) + \frac{\partial}{\partial q_2}(v_2 h_3 h_1) + \frac{\partial}{\partial q_3}(v_3 h_1 h_2)\right] \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3$$

Applying the definition of the divergence (17.4) gives

$$\begin{aligned} \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} &= \left[\frac{\partial}{\partial q_1}(v_1 h_2 h_3) + \frac{\partial}{\partial q_2}(v_2 h_3 h_1) + \frac{\partial}{\partial q_3}(v_3 h_1 h_2)\right] \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3, \\ &\overset{\text{divergence}}{=} (\nabla \cdot \boldsymbol{v})\, \mathrm{d}V. \end{aligned}$$

We need to find an expression for the volume $\mathrm{d}V$ of the region shown in Figure 17.2. Geometrically, $\mathrm{d}V$ is the volume of a parallelepiped spanned by the infinitesimal tangent vectors $(\partial \boldsymbol{x}/\partial q_i)\mathrm{d}q_i$, with $i = 1, 2, 3$, hence

$$\begin{aligned} \mathrm{d}V &= \frac{\partial \boldsymbol{x}}{\partial q_1} \cdot \left(\frac{\partial \boldsymbol{x}}{\partial q_2} \times \frac{\partial \boldsymbol{x}}{\partial q_3}\right) \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3, \\ &= h_1 \hat{\boldsymbol{q}}_1 \cdot \left(h_2 \hat{\boldsymbol{q}}_1 \times h_3 \hat{\boldsymbol{q}}_2\right) \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3, \\ &= h_1 h_2 h_3\, \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3, \end{aligned}$$

hence

$$\boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} \stackrel{\text{divergence}}{=} (\nabla \cdot \boldsymbol{v})\ \mathrm{d}V = (\nabla \cdot \boldsymbol{v})\, h_1 h_2 h_3\, \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3.$$

Thus,

Theorem 17.2 (The divergence operator in orthogonal curvilinear coordinates)

$$\boxed{\nabla \cdot \boldsymbol{v} = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}(v_1 h_2 h_3) + \frac{\partial}{\partial q_2}(v_2 h_3 h_1) + \frac{\partial}{\partial q_3}(v_3 h_1 h_2)\right].} \tag{17.5}$$

17.3.3 The curl

This is the last operator to compute. Let us take Stokes's theorem for an areal patch S with boundary C, and integrate a vector field $\boldsymbol{v}$ in the usual manner:

$$\int_S \nabla \times \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} = \oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}.$$

In particular, let S be a patch of area on the surface

$$q_1 = \text{Constant},$$

as shown in Figure 17.3. Thus,

$$\mathrm{d}\boldsymbol{S} = \hat{\boldsymbol{q}}_1 h_2 h_3 \mathrm{d}q_2 \mathrm{d}q_3.$$

We compute the line integral around boundary of this areal patch in a sense given by the right-hand rule:

$$\begin{aligned}\oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} = {} & \underbrace{[\boldsymbol{v} \cdot (\hat{\boldsymbol{q}}_2 h_2 \mathrm{d}q_2)]\,(q_1, q_2, q_3)}_{\text{Contribution from edge } E1} + \underbrace{[\boldsymbol{v} \cdot (\hat{\boldsymbol{q}}_3 h_3 \mathrm{d}q_3)]\,(q_1, q_2 + \mathrm{d}q_2, q_3)}_{E2} \\ & - \underbrace{[\boldsymbol{v} \cdot (\hat{\boldsymbol{q}}_2 h_2 \mathrm{d}q_2)]\,(q_1, q_2 + \mathrm{d}q_2, q_3 + \mathrm{d}q_3)}_{E3} - \underbrace{[\boldsymbol{v} \cdot (\hat{\boldsymbol{q}}_3 h_3 \mathrm{d}q_3)]\,(q_1, q_2, q_3 + \mathrm{d}q_3)}_{E4}\end{aligned}$$

Tidy up a little bit:

$$\begin{aligned}\oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x} = {} & [v_2 h_2]\,(q_1, q_2, q_3)\mathrm{d}q_2 + [v_3 h_3]\,(q_1, q_2 + \mathrm{d}q_2, q_3)\mathrm{d}q_3 \\ & - [v_2 h_2]\,(q_1, q_2 + \mathrm{d}q_2, q_3 + \mathrm{d}q_3)\mathrm{d}q_2 - [v_3 h_3]\,(q_1, q_2, q_3 + \mathrm{d}q_3)\mathrm{d}q_3.\end{aligned}$$

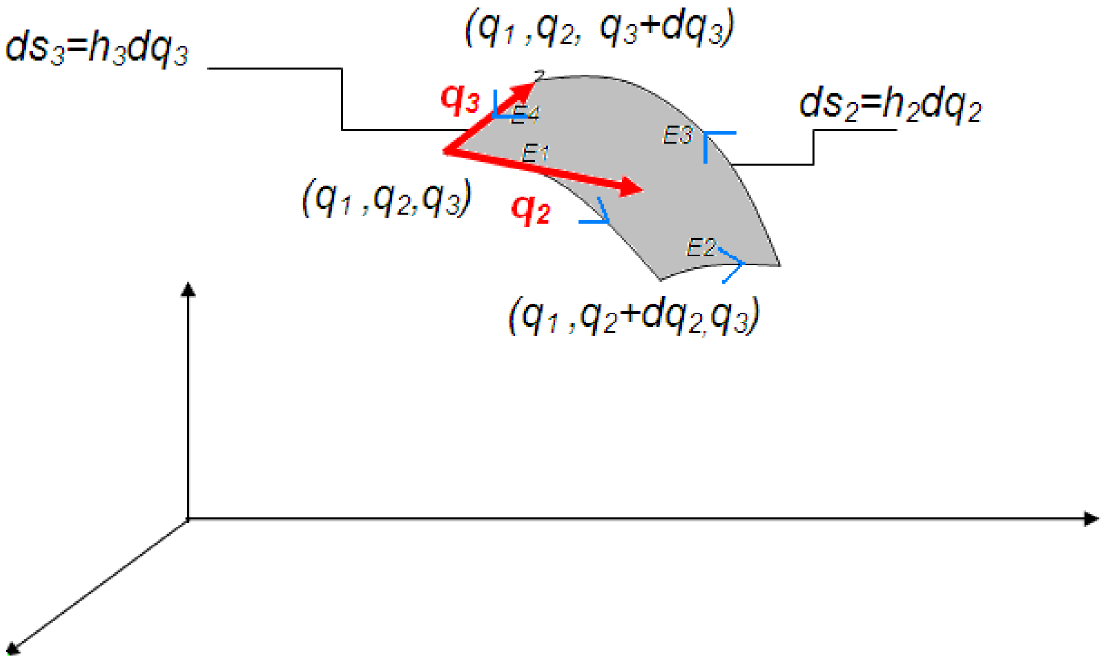


Figure 17.3: The area element in curvilinear coordinates: this sketch forms a basis for deriving the curl operator in curvilinear coordinates.

Pair up terms ready for a Taylor expansion:

$$\oint_C \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{x} = \left[v_2h_2\right](q_1,q_2,q_3)\mathrm{d}q_2 - \left[v_2h_2\right](q_1,q_2+\mathrm{d}q_2,q_3+\mathrm{d}q_3)\mathrm{d}q_2$$
$$+\left[v_3h_3\right](q_1,q_2+\mathrm{d}q_2,q_3)\mathrm{d}q_3 - \left[v_3h_3\right](q_1,q_2,q_3+\mathrm{d}q_3)\mathrm{d}q_3.$$

Expand:

$$\oint_C \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{x} = -\left[\frac{\partial}{\partial q_3}(v_2h_2)\right]\mathrm{d}q_2\mathrm{d}q_3 + \left[\frac{\partial}{\partial q_2}(v_3h_3)\right]\mathrm{d}q_2\mathrm{d}q_3.$$

Note that we have neglected the term

$$-\left[\frac{\partial}{\partial q_2}(v_2h_2)\right]\mathrm{d}q_2\mathrm{d}q_2$$

from the first Taylor expansion because it is second order in the small quantity $\mathrm{d}q_2$. Next, we consider

$$\begin{aligned}\int_S (\nabla\times\boldsymbol{v})\cdot\mathrm{d}\boldsymbol{S} &= \int_S (\nabla\times\boldsymbol{v})\cdot\hat{\boldsymbol{q}}_1 h_2h_3\mathrm{d}q_2\mathrm{d}q_3,\\ &= (\nabla\times\boldsymbol{v})\cdot\hat{\boldsymbol{q}}_1 h_2h_3\mathrm{d}q_2\mathrm{d}q_3,\end{aligned} \tag{17.6}$$

since the areal patch is infinitesimal. But by Stokes's theorem,

$$\begin{aligned}\int_S (\nabla \times \boldsymbol{v}) \cdot \mathrm{d}\boldsymbol{S} &= \oint_C \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{x}, \\ &= -\left[\frac{\partial}{\partial q_3}(v_2 h_2)\right] \mathrm{d}q_2 \mathrm{d}q_3 + \left[\frac{\partial}{\partial q_2}(v_3 h_3)\right] \mathrm{d}q_2 \mathrm{d}q_3. \end{aligned} \tag{17.7}$$

Equate Equations (17.6) and (17.7)

$$(\nabla \times \boldsymbol{v}) \cdot \hat{\boldsymbol{q}}_1 h_2 h_3 \mathrm{d}q_2 \mathrm{d}q_3 = -\left[\frac{\partial}{\partial q_3}(v_2 h_2)\right] \mathrm{d}q_2 \mathrm{d}q_3 + \left[\frac{\partial}{\partial q_2}(v_3 h_3)\right] \mathrm{d}q_2 \mathrm{d}q_3$$

But

$$(\nabla \times \boldsymbol{v}) \cdot \hat{\boldsymbol{q}}_1 = (\nabla \times \boldsymbol{v})_1 ,$$

the component of the curl in the first ($\hat{\boldsymbol{q}}_1$) direction. Hence,

$$(\nabla \times \boldsymbol{v})_1 = \frac{1}{h_2 h_3}\left[\frac{\partial}{\partial q_2}(v_3 h_3) - \frac{\partial}{\partial q_3}(v_2 h_2)\right]$$

By construction, $(\hat{\boldsymbol{q}}_1, \hat{\boldsymbol{q}}_2, \hat{\boldsymbol{q}}_3)$ form a right-handed system. Thus, we may obtain the other components of the curl through cyclic permutations:

$$\begin{aligned}(\nabla \times \boldsymbol{v})_2 &= \frac{1}{h_3 h_1}\left[\frac{\partial}{\partial q_3}(v_1 h_1) - \frac{\partial}{\partial q_1}(v_3 h_3)\right], \\ (\nabla \times \boldsymbol{v})_3 &= \frac{1}{h_1 h_2}\left[\frac{\partial}{\partial q_1}(v_2 h_2) - \frac{\partial}{\partial q_2}(v_1 h_1)\right].\end{aligned}$$

This result may be summarized succinctly in determinant form:

Theorem 17.3 (The curl operator in orthogonal curvilinear coordinates)

$$\boxed{\nabla \times \boldsymbol{v} = \frac{1}{h_1 h_2 h_3}\begin{vmatrix} \hat{\boldsymbol{q}}_1 h_1 & \hat{\boldsymbol{q}}_2 h_2 & \hat{\boldsymbol{q}}_3 h_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 v_1 & h_2 v_2 & h_3 v_3 \end{vmatrix},} \tag{17.8}$$

where $\boldsymbol{v} = v_1 \hat{\boldsymbol{q}}_1 + v_2 \hat{\boldsymbol{q}}_2 + v_3 \hat{\boldsymbol{q}}_3$ is the vector field $\boldsymbol{v}(q_1, q_2, q_3)$ written in curvilinear coordinates.

17.4 Some applications of our results

Example: Compute the Laplacian $\nabla^2\psi$ of a scalar field $\psi(\boldsymbol{x})$ (a) in general, curvilinear coordinates (q_1, q_2, q_3); (b) in spherical polar coordinates.

Let $\boldsymbol{v} = \nabla\psi$. Using the definition of the gradient,

$$v_i = \frac{1}{h_i}\frac{\partial\psi}{\partial q_i},$$

(no sum), and

$$\boldsymbol{v} = \sum_{i=1}^{3} \hat{\boldsymbol{q}}_i v_i.$$

Next, using the definition of divergence,

$$\nabla\cdot\boldsymbol{v} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}(v_1h_2h_3) + \frac{\partial}{\partial q_2}(v_2h_3h_1) + \frac{\partial}{\partial q_3}(v_3h_1h_2)\right],$$

we obtain

$$\nabla\cdot(\nabla\psi) = \frac{1}{h_1h_2h_3}\left\{\frac{\partial}{\partial q_1}\left[\left(\frac{1}{h_1}\frac{\partial\psi}{\partial q_1}\right)h_2h_3\right] + \text{Cyclic permutations}\right\}$$

But $\nabla^2\psi = \nabla\cdot(\nabla\psi)$. We therefore have the following theorem:

Theorem 17.4 (The laplace operator in orthogonal curvilinear coordinates)

$$\boxed{\nabla^2\psi = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial\psi}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3h_1}{h_2}\frac{\partial\psi}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial\psi}{\partial q_3}\right)\right].} \tag{17.9}$$

For the spherical-polar case, let $q_1 \to r$, $q_2 \to \theta$, and $q_3 \to \varphi$. Then,

$$\hat{\boldsymbol{q}}_1 = \hat{\boldsymbol{r}}, \qquad \hat{\boldsymbol{q}}_2 = \hat{\boldsymbol{\theta}}, \qquad \hat{\boldsymbol{q}}_3 = \hat{\boldsymbol{\varphi}},$$

and

$$h_1 = 1, \qquad h_2 = r, \qquad h_3 = r\sin\theta.$$

Substituting these relations into Eq. (17.9), we obtain

$$\nabla^2\psi\,(r,\theta,\varphi) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2}.$$

Example: Using spherical polar coordinates, show that the central force

$$\boldsymbol{F} = \alpha|\boldsymbol{r}|^n\boldsymbol{r}, \qquad n \neq -2.$$

is conservative. Show also that

$$\mathcal{U} = -\frac{\alpha}{n+2}|\boldsymbol{r}|^{n+2}$$

is a potential, $\boldsymbol{F} = -\nabla\mathcal{U}$.

We use the assignments made in the last exercise and compute

$$\begin{aligned}
\nabla\times\boldsymbol{F} &= \frac{1}{h_1h_2h_3}\begin{vmatrix} \hat{\boldsymbol{q}}_1h_1 & \hat{\boldsymbol{q}}_2h_2 & \hat{\boldsymbol{q}}_3h_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1v_1 & h_2v_2 & h_3v_3 \end{vmatrix}, \\
&= \frac{1}{r^2\sin\theta}\begin{vmatrix} \hat{\boldsymbol{r}} & \hat{\boldsymbol{\theta}}r & \hat{\boldsymbol{\varphi}}r\sin\theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\varphi} \\ F_r & rF_\theta & r\sin\theta F_\varphi, \end{vmatrix},
\end{aligned}$$

Note that the force can be written as

$$\boldsymbol{F} = \alpha|\boldsymbol{r}|^n|\boldsymbol{r}|\hat{\boldsymbol{r}},$$

so

$$F_r = \alpha r^{n+1}, \qquad F_\varphi = 0, \qquad F_\vartheta = 0.$$

Thus, a central force only has a radial component, when expressed in spherical polar coordinates. The curl is then

$$\begin{aligned}
\nabla\times\boldsymbol{F} &= \frac{1}{r^2\sin\theta}\begin{vmatrix} \hat{\boldsymbol{r}} & \hat{\boldsymbol{\theta}}r & \hat{\boldsymbol{\varphi}}r\sin\theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\varphi} \\ F_r(r) & 0 & 0 \end{vmatrix}, \\
&= \frac{1}{r^2\sin\theta}\left[\hat{\boldsymbol{r}}\left(\partial_\theta 0 - \partial_\varphi 0\right) - \hat{\boldsymbol{\theta}}r\left(\partial_r 0 - \partial_\varphi F_r(r)\right) + \hat{\boldsymbol{\varphi}}r\sin\theta\left(\partial_r 0 - \partial_\theta F_r(r)\right)\right], \\
&= 0.
\end{aligned}$$

Next, we compute $-\nabla\mathcal{U}$. For spherical polar coordinates,

$$\begin{aligned}
\nabla\psi &= \frac{\hat{\boldsymbol{q}}_1}{h_1}\frac{\partial\psi}{\partial q_1} + \frac{\hat{\boldsymbol{q}}_2}{h_2}\frac{\partial\psi}{\partial q_2} + \frac{\hat{\boldsymbol{q}}_3}{h_3}\frac{\partial\psi}{\partial q_3}, \\
&= \hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r} + \frac{\hat{\boldsymbol{\theta}}}{r}\frac{\partial\psi}{\partial\theta} + \frac{\hat{\boldsymbol{\varphi}}}{r\sin\theta}\frac{\partial\psi}{\partial\varphi}.
\end{aligned}$$

But $\mathcal{U}$ is radially symmetric, $\mathcal{U} = \mathcal{U}(r)$, so

$$\begin{aligned}
-\nabla\mathcal{U} &= -\hat{\boldsymbol{r}}\frac{\partial\mathcal{U}}{\partial r} = -\hat{\boldsymbol{r}}\mathcal{U}'(r), \\
&= -\hat{\boldsymbol{r}}\frac{d}{dr}\left(-\frac{\alpha}{n+2}r^{n+2}\right), \\
&= +\hat{\boldsymbol{r}}\alpha r^{n+1}, \\
&= \alpha r^n\boldsymbol{r}, \qquad \boldsymbol{r} = r\hat{\boldsymbol{r}}, \\
&= \boldsymbol{F}.
\end{aligned}$$

Note in particular, that if

$$\boldsymbol{F} = \frac{\alpha}{|\boldsymbol{r}|^3}\boldsymbol{r},$$

then $n = -3$ and

$$\mathcal{U} = \frac{\alpha}{r}.$$

When $\alpha = -Gm_1m_2$ this is the gravitational force.

Chapter 18

Special Curvilinear coordinate systems

Overview

In this section we study two special coordinate systems that commonly occur in fluid flow, electromagnetism, and quantum mechanics. Because of their importance to your later courses, this is a particularly important chapter.

18.1 Spherical polar coordinates

We have already encountered this system, but let us recall it briefly. The coordinate system is shown in Fig. 18.1. The point P is labelled by its radial distance r from the origin, together with two angles: the azimuthal angle φ, and the angle θ between the z-direction and the radius vector $\boldsymbol{r}$ extending from the origin O to the point P. The Cartesian coordinate system (x, y, z) and the spherical polar coordinate system are related through

$$\begin{aligned} x &= r \sin\theta \cos\varphi, \\ y &= r \sin\theta \sin\varphi, \\ z &= r \cos\theta, \end{aligned}$$

with inverse transformation

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \theta &= \cos^{-1}\left(z/r\right), \\ \varphi &= \tan^{-1}\left(y/x\right). \end{aligned}$$

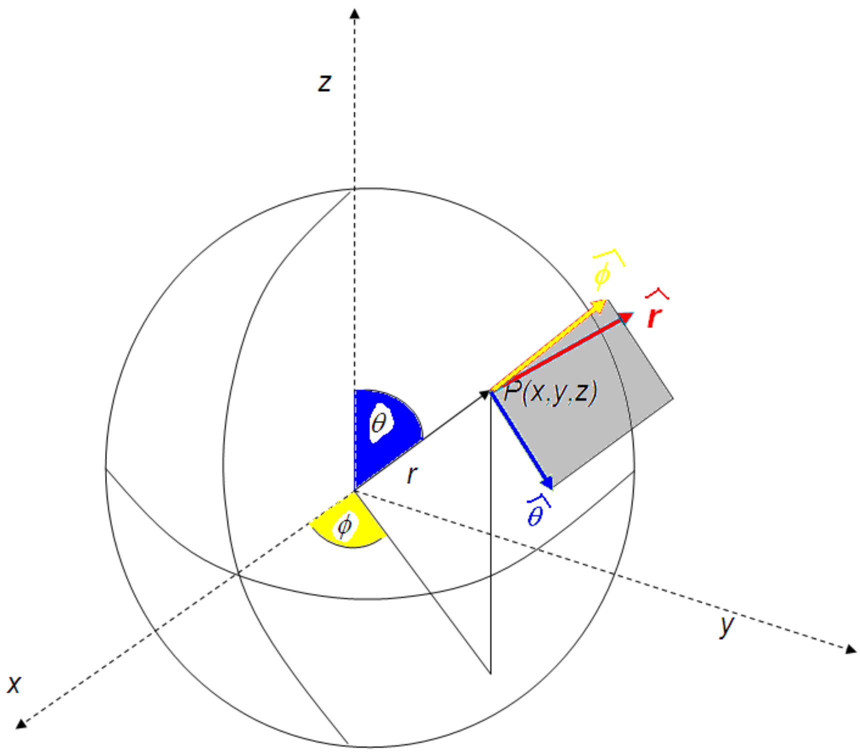


Figure 18.1: Spherical polar coordinates

Recall, we made the identification $q_1 \to r$, $q_2 \to \theta$, and $q_3 \to \varphi$, and we wrote

$$\begin{aligned} \boldsymbol{r} \equiv \boldsymbol{x} &= \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z, \\ &= \hat{\boldsymbol{x}}r\sin\theta\cos\varphi + \hat{\boldsymbol{y}}r\sin\theta\sin\varphi + \hat{\boldsymbol{z}}r\cos\theta. \end{aligned}$$

From this we computed the vectors

$$\hat{\boldsymbol{r}} = \frac{\partial \boldsymbol{x}}{\partial r}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial r}\right|, \qquad \hat{\boldsymbol{\theta}} = \frac{\partial \boldsymbol{x}}{\partial \theta}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial \theta}\right|, \qquad \hat{\boldsymbol{\varphi}} = \frac{\partial \boldsymbol{x}}{\partial \varphi}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial \varphi}\right|.$$

These were found to be

$$\begin{aligned} \hat{\boldsymbol{r}} &= \hat{\boldsymbol{x}}\sin\theta\cos\varphi + \hat{\boldsymbol{y}}\sin\theta\sin\varphi + \hat{\boldsymbol{z}}\cos\theta, \\ \hat{\boldsymbol{\theta}} &= \hat{\boldsymbol{x}}\cos\theta\cos\varphi + \hat{\boldsymbol{y}}\cos\theta\sin\varphi - \hat{\boldsymbol{z}}\sin\theta, \\ \hat{\boldsymbol{\varphi}} &= -\hat{\boldsymbol{x}}\sin\varphi + \hat{\boldsymbol{y}}\cos\varphi. \end{aligned}$$

and are mutually orthogonal:

$$\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{\varphi}} = \hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\varphi}} \cdot \hat{\boldsymbol{\theta}} = 0;$$

Also,

$$\hat{\boldsymbol{r}} \times \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\varphi}} + \text{ Cyclic permutations.}$$

Note that it also follows that

$$\boldsymbol{x} \equiv \boldsymbol{r} = r\hat{\boldsymbol{r}}.$$

We also computed the scale factors

$$h_r = \left|\frac{\partial \boldsymbol{x}}{\partial r}\right|, \qquad h_\theta = \left|\frac{\partial \boldsymbol{x}}{\partial \theta}\right|, \qquad h_\varphi = \left|\frac{\partial \boldsymbol{x}}{\partial \varphi}\right|,$$

which we found to be equal to

$$h_r = 1, \qquad h_\theta = r, \qquad h_\varphi = r\sin\theta.$$

Once we know the scale factors and the unit vectors, we may compute grad, div, curl, and the Laplacian in the spherical polar system. We recall again the identifications $\hat{\boldsymbol{q}}_1 = \hat{\boldsymbol{r}}$, $\hat{\boldsymbol{q}}_2 = \hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{q}}_3 = \hat{\boldsymbol{\varphi}}$, along with

$$q_1 = r, \qquad q_2 = \theta, \qquad q_3 = \varphi.$$

Thus,

- **The gradient**

  General case:

$$\nabla\psi(q_1, q_2, q_3) = \frac{\hat{\boldsymbol{q}}_1}{h_1}\frac{\partial\psi}{\partial q_1} + \frac{\hat{\boldsymbol{q}}_2}{h_2}\frac{\partial\psi}{\partial q_2} + \frac{\hat{\boldsymbol{q}}_3}{h_3}\frac{\partial\psi}{\partial q_3}; \tag{18.1}$$

  Spherical polar coordinates:

$$\boxed{\nabla\psi(r, \theta, \varphi) = \hat{\boldsymbol{r}}\frac{\partial\psi}{\partial r} + \frac{\hat{\boldsymbol{\theta}}}{r}\frac{\partial\psi}{\partial\theta} + \frac{\hat{\boldsymbol{\varphi}}}{r\sin\theta}\frac{\partial\psi}{\partial\varphi}.} \tag{18.2}$$

- **The divergence**

  General case:

$$\nabla\cdot\boldsymbol{v} = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}(v_1h_2h_3) + \frac{\partial}{\partial q_2}(v_2h_3h_1) + \frac{\partial}{\partial q_3}(v_3h_1h_2)\right]; \tag{18.3}$$

  Spherical polar coordinates:

$$\boxed{\nabla\cdot\boldsymbol{v}(r, \theta, \varphi) = \frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2v_r\right) + r\frac{\partial}{\partial\theta}(\sin\theta v_\theta) + r\frac{\partial v_\varphi}{\partial\varphi}\right]} \tag{18.4}$$

- **The curl**

  General case:

$$(\nabla \times \boldsymbol{v})(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} \hat{\boldsymbol{q}}_1 h_1 & \hat{\boldsymbol{q}}_2 h_2 & \hat{\boldsymbol{q}}_3 h_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 v_1 & h_2 v_2 & h_3 v_3 \end{vmatrix}; \tag{18.5}$$

  Spherical polar coordinates:

$$\boxed{(\nabla \times \boldsymbol{v})(r, \theta, \varphi) = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\boldsymbol{r}} & \hat{\boldsymbol{\theta}} r & \hat{\boldsymbol{\varphi}} r \sin\theta \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ v_r & r v_\theta & r \sin\theta v_\varphi, \end{vmatrix}} \tag{18.6}$$

- **The Laplacian**

  General case:

$$\nabla^2 \psi(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial \psi}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial \psi}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial \psi}{\partial q_3} \right) \right]. \tag{18.7}$$

  Spherical polar coordinates:

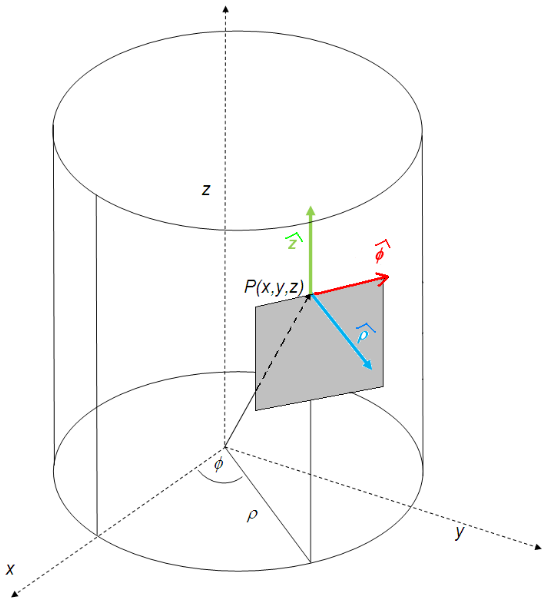
$$\boxed{\nabla^2 \psi\, (r, \theta, \varphi) = \frac{1}{r^2 \sin\theta} \left[\sin\theta \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin\theta} \frac{\partial^2 \psi}{\partial \varphi^2} \right].} \tag{18.8}$$

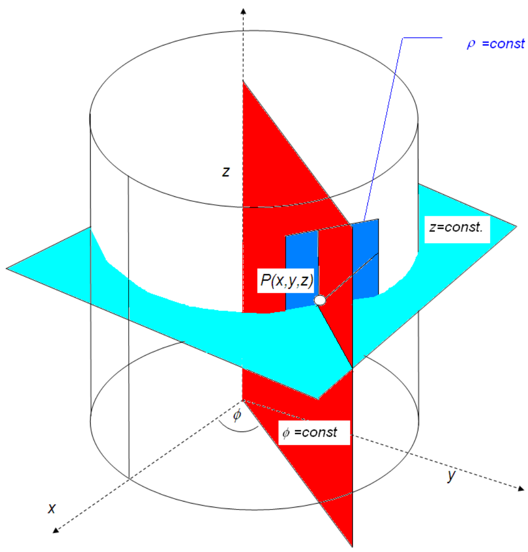
18.2 Cylindrical coordinates

Consider **cylindrical polar coordinates** as shown in the figure. The point P can either be labelled by the Cartesian triple (x, y, z), or by the following quantities:

- The distance z between the point P and its projection on to the xy-plane;
- The distance ρ from the origin O to the projection of P on to the xy-plane;
- The angle φ that the projection makes with the x-axis;

Note that φ is the same as the azimuthal angle in the spherical polar system; otherwise these two systems are different.





The surfaces generated by these new coordinates are two planes and a cylinder:

- The plane $z = \text{Const.}$;
- The plane $y = x \tan\varphi$ (i.e. the plane $\varphi = \text{Const.}$);
- The cylinder $\rho^2 = x^2 + y^2 = \text{Const.}$;

see the accompanying figure.

The Cartesian and the cylindrical coordinate systems are related through

$$\begin{aligned} x &= \rho\cos\varphi, \\ y &= \rho\sin\varphi, \\ z &= z, \end{aligned}$$

with inverse transformation

$$\begin{aligned}\rho &= \sqrt{x^2+y^2},\\ \varphi &= \tan^{-1}(y/x),\\ z &= z.\end{aligned}$$

Thus, we have

$$\begin{aligned}\boldsymbol{r} \equiv \boldsymbol{x} &= \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z,\\ &= \hat{\boldsymbol{x}}\rho\cos\varphi + \hat{\boldsymbol{y}}\rho\sin\varphi + \hat{\boldsymbol{z}}z.\end{aligned}$$

Let's compute the tangent vectors:

$$\hat{\boldsymbol{\rho}} = \frac{\partial \boldsymbol{x}}{\partial \rho}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial \rho}\right|, \qquad \hat{\boldsymbol{\varphi}} = \frac{\partial \boldsymbol{x}}{\partial \varphi}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial \varphi}\right|, \qquad \hat{\boldsymbol{z}} = \frac{\partial \boldsymbol{x}}{\partial z}\bigg/\left|\frac{\partial \boldsymbol{x}}{\partial z}\right|.$$

First,

$$\frac{\partial \boldsymbol{x}}{\partial \rho} = \hat{\boldsymbol{x}}\cos\varphi + \hat{\boldsymbol{y}}\sin\varphi,$$

and this has unit length, hence

$$\hat{\boldsymbol{\rho}} = \hat{\boldsymbol{x}}\cos\varphi + \hat{\boldsymbol{y}}\sin\varphi.$$

Next,

$$\frac{\partial \boldsymbol{x}}{\partial \varphi} = -\hat{\boldsymbol{x}}\rho\sin\varphi + \hat{\boldsymbol{y}}\rho\cos\varphi,$$

which has length ρ, hence

$$\hat{\boldsymbol{\varphi}} = -\hat{\boldsymbol{x}}\sin\varphi + \hat{\boldsymbol{y}}\cos\varphi.$$

Finally, the third tangent vector must simply be $\hat{\boldsymbol{z}}$. We assemble these results:

$$\begin{aligned}\hat{\boldsymbol{\rho}} &= \hat{\boldsymbol{x}}\cos\varphi + \hat{\boldsymbol{y}}\sin\varphi,\\ \hat{\boldsymbol{\varphi}} &= -\hat{\boldsymbol{x}}\sin\varphi + \hat{\boldsymbol{y}}\cos\varphi,\\ \hat{\boldsymbol{z}} &= \hat{\boldsymbol{z}}.\end{aligned}$$

These are quite clearly mutually orthogonal:

$$\hat{\boldsymbol{\rho}}\cdot\hat{\boldsymbol{\varphi}} = \hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\varphi}}\cdot\hat{\boldsymbol{\theta}} = 0.$$

We must also compute the scale factors:

$$h_\rho = \left|\frac{\partial \boldsymbol{x}}{\partial \rho}\right| = |\hat{\boldsymbol{x}}\cos\varphi + \hat{\boldsymbol{y}}\sin\varphi| = 1,$$

$$h_\varphi = \left|\frac{\partial \boldsymbol{x}}{\partial \varphi}\right| = |-\hat{\boldsymbol{x}}\rho\sin\varphi + \hat{\boldsymbol{y}}\rho\cos\varphi| = \rho,$$

and

$$h_z = 1.$$

For convenience, let us assemble these results also:

$$\begin{aligned} h_\rho &= 1, \\ h_\varphi &= \rho, \\ h_z &= 1. \end{aligned}$$

The line element in this coordinate system is thus

$$\mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2\mathrm{d}\varphi^2 + \mathrm{d}z^2.$$

Now, we make the identifications $\hat{\boldsymbol{q}}_1 = \hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{q}}_2 = \hat{\boldsymbol{\varphi}}$, and $\hat{\boldsymbol{q}}_3 = \hat{\boldsymbol{z}}$, along with

$$q_1 = \rho, \qquad q_2 = \varphi, \qquad q_3 = z.$$

Thus,

- **The gradient**

  General case:

  $$\nabla\psi(q_1,q_2,q_3) = \frac{\hat{\boldsymbol{q}}_1}{h_1}\frac{\partial\psi}{\partial q_1} + \frac{\hat{\boldsymbol{q}}_2}{h_2}\frac{\partial\psi}{\partial q_2} + \frac{\hat{\boldsymbol{q}}_3}{h_3}\frac{\partial\psi}{\partial q_3};$$

  Cylindrical polar coordinates:

  $$\boxed{\nabla\psi(\rho,\varphi,z) = \hat{\boldsymbol{\rho}}\frac{\partial\psi}{\partial\rho} + \frac{\hat{\boldsymbol{\varphi}}}{\rho}\frac{\partial\psi}{\partial\varphi} + \hat{\boldsymbol{z}}\frac{\partial\psi}{\partial z}.} \tag{18.9}$$

- **The divergence**

General case:

$$\nabla \cdot \boldsymbol{v} = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} (v_1 h_2 h_3) + \frac{\partial}{\partial q_2} (v_2 h_3 h_1) + \frac{\partial}{\partial q_3} (v_3 h_1 h_2) \right];$$

Cylindrical polar coordinates:

$$\boxed{\nabla \cdot \boldsymbol{v}(\rho, \varphi, z) = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho v_\rho) + \frac{1}{\rho} \frac{\partial v_\varphi}{\partial \varphi} + \frac{\partial v_z}{\partial z}.} \tag{18.10}$$

- **The curl**

General case:

$$(\nabla \times \boldsymbol{v})(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} \hat{\boldsymbol{q}}_1 h_1 & \hat{\boldsymbol{q}}_2 h_2 & \hat{\boldsymbol{q}}_3 h_3 \\ \frac{\partial}{\partial q_1} & \frac{\partial}{\partial q_2} & \frac{\partial}{\partial q_3} \\ h_1 v_1 & h_2 v_2 & h_3 v_3 \end{vmatrix};$$

Cylindrical polar coordinates:

$$\boxed{(\nabla \times \boldsymbol{v})(\rho, \varphi, z) = \frac{1}{\rho} \begin{vmatrix} \hat{\boldsymbol{\rho}} & \hat{\boldsymbol{\varphi}}\rho & \hat{\boldsymbol{z}} \\ \frac{\partial}{\partial \rho} & \frac{\partial}{\partial \varphi} & \frac{\partial}{\partial z} \\ v_\rho & \rho v_\varphi & v_z, \end{vmatrix}} \tag{18.11}$$

- **The Laplacian**

General case:

$$\nabla^2 \psi(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} \left(\frac{h_2 h_3}{h_1} \frac{\partial \psi}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left(\frac{h_3 h_1}{h_2} \frac{\partial \psi}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left(\frac{h_1 h_2}{h_3} \frac{\partial \psi}{\partial q_3} \right) \right].$$

Cylindrical polar coordinates:

$$\boxed{\nabla^2 \psi\,(\rho, \varphi, z) = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial \psi}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 \psi}{\partial \varphi^2} + \frac{\partial^2 \psi}{\partial z^2},} \tag{18.12}$$

and NOTE THE EXPONENT in the ∂_ρ derivative: it is 1 (in the spherical polar coordinate case the corresponding radial exponent is 2).

18.3 Physical applications

18.3.1 Laplace's equation in cylindrical polar coordinates

Example: Solve Laplace's equation $\nabla^2\psi = 0$, in cylindrical coordinates, for $\psi = \psi(\rho)$.

For this particular function,

$$\partial_\varphi \psi = \partial_z \psi = 0,$$

hence,

$$0 = \nabla^2\psi = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) = 0.$$

This means that the $(\cdot) = \mathrm{Const.} := k$, hence,

$$\rho\frac{\partial\psi}{\partial\rho} = k,$$

or

$$\frac{d\psi}{d\rho} = \frac{k}{\rho}.$$

Separating the variables gives

$$\mathrm{d}\psi = k\frac{\mathrm{d}\rho}{\rho};$$

integration gives the final answer:

$$\psi(\rho) = \psi_0 + k\log(\rho), \qquad \rho \neq 0.$$

and there are two constants of integration because the equation is second-order.

18.3.2 Pipe flow

For the flow of an incompressible viscous fluid, the Navier–Stokes equations lead to

$$-\nabla\times(\boldsymbol{v}\times(\nabla\times\boldsymbol{v})) = \nu\nabla^2(\nabla\times\boldsymbol{v}), \tag{18.13}$$

where $\boldsymbol{v}(\boldsymbol{x})$ is the fluid velocity and ν is the constant kinematic viscosity. For axial flow in a cylindrical pipe we take the velocity $\boldsymbol{v}$ to be

$$\boldsymbol{v} = \hat{\boldsymbol{z}}v(\rho). \tag{18.14}$$

Show that the left-hand side of Equation (18.13) is identically zero when the velocity has the form (18.14). Hence, $\boldsymbol{v} = \hat{\boldsymbol{z}} v(\rho)$ must satisfy

$$\nabla^2 \left(\nabla \times \boldsymbol{v}\right) = 0.$$

Let's focus on the left-hand side first. With $\boldsymbol{v} = \hat{\boldsymbol{z}} v(\rho)$, we compute

$$\nabla \times \boldsymbol{v} = \frac{1}{\rho} \begin{vmatrix} \hat{\boldsymbol{\rho}} & \rho\hat{\boldsymbol{\varphi}} & \hat{\boldsymbol{z}} \\ \partial_\rho & \partial_\varphi & \partial_z \\ 0 & 0 & v(\rho) \end{vmatrix} = -\hat{\boldsymbol{\varphi}} \frac{\partial v}{\partial \rho}.$$

Now we take

$$\boldsymbol{v} \times \left(\nabla \times \boldsymbol{v}\right) = \begin{vmatrix} \hat{\boldsymbol{\rho}} & \hat{\boldsymbol{\varphi}} & \hat{\boldsymbol{z}} \\ 0 & 0 & v(\rho) \\ 0 & -\frac{\partial v}{\partial \rho} & 0 \end{vmatrix} = +\hat{\boldsymbol{\rho}} v(\rho) \frac{\partial v}{\partial \rho},$$

and this determinant expansion is legitimate because $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{z}})$ form a right-handed orthonormal triad. Finally, we take the curl of this expression:

$$\nabla \times \left(\boldsymbol{v} \times \left(\nabla \times \boldsymbol{v}\right)\right) = \frac{1}{\rho} \begin{vmatrix} \hat{\boldsymbol{\rho}} & \rho\hat{\boldsymbol{\varphi}} & \hat{\boldsymbol{z}} \\ \partial_\rho & \partial_\varphi & \partial_z \\ v(\rho)\frac{\partial v}{\partial \rho} & 0 & 0 \end{vmatrix} = 0.$$

Thus, for $\boldsymbol{v} = \hat{\boldsymbol{z}} v(\rho)$, the LHS of the fluid equation is identically zero, and we are forced to consider

$$\nabla^2 \left(\nabla \times \boldsymbol{v}\right) = 0, \tag{18.15}$$

Show that Equation (18.15) leads to the ordinary differential equation

$$\frac{1}{\rho} \frac{d}{d\rho} \left(\rho \frac{d^2 v}{d\rho^2}\right) - \frac{1}{\rho^2} \frac{dv}{d\rho} = 0,$$

with solution

$$v = v_0 + a_2 \rho^2,$$

where v_0 and a_2 are constants.

Solution – start with Equation (18.15). Use $\nabla \times \boldsymbol{v} = -\hat{\boldsymbol{\varphi}}(\partial v/\partial \rho)$ to rewrite Equation (18.15) as

$$\nabla^2 \left(-\hat{\boldsymbol{\varphi}} v'(\rho)\right) = 0.$$

Some care is required here because $\hat{\boldsymbol{\varphi}}$ is non-constant and cannot be taken outside the differential operator. However, we can cross both sizes with $\hat{\boldsymbol{z}}$ and take the constant vector $\hat{\boldsymbol{z}}$ inside the operator:

$$0 = \nabla^2 \left(\hat{\boldsymbol{\varphi}} \times \hat{\boldsymbol{z}} v'(\rho)\right).$$

But $(\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{z}})$ are a right-handed orthonormal triad, so $\hat{\boldsymbol{\varphi}} \times \hat{\boldsymbol{z}} = \hat{\boldsymbol{\rho}}$, and we solve

$$0 = \nabla^2 \left(\hat{\boldsymbol{\rho}} v'(\rho)\right) = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial}{\partial \rho}\left(\hat{\boldsymbol{\rho}} v'(\rho)\right)\right) + \frac{1}{\rho^2}\frac{\partial^2}{\partial \varphi^2}\left(\hat{\boldsymbol{\rho}} v'(\rho)\right) + \frac{\partial^2}{\partial z^2}\left(\hat{\boldsymbol{\rho}} v'(\rho)\right). \tag{18.16}$$

Let's consider

$$\hat{\boldsymbol{\rho}} = \hat{\boldsymbol{x}} \cos\varphi + \hat{\boldsymbol{y}} \sin\varphi.$$

Evidently,

$$\frac{\partial \hat{\boldsymbol{\rho}}}{\partial \rho} = \frac{\partial \hat{\boldsymbol{\rho}}}{\partial z} = 0,$$

and

$$\frac{\partial \hat{\boldsymbol{\rho}}}{\partial \varphi} = -\hat{\boldsymbol{x}} \sin\varphi + \hat{\boldsymbol{y}} \cos\varphi, \qquad \frac{\partial^2 \hat{\boldsymbol{\rho}}}{\partial \varphi^2} = -\hat{\boldsymbol{x}} \cos\varphi - \hat{\boldsymbol{y}} \sin\varphi = -\hat{\boldsymbol{\rho}}.$$

Substitute these expressions back into Eq. (18.16):

$$\begin{aligned} 0 &= \frac{\hat{\boldsymbol{\rho}}}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial v'(\rho)}{\partial \rho}\right) + \frac{v'(\rho)}{\rho^2}\frac{\partial^2 \hat{\boldsymbol{\rho}}}{\partial \varphi^2} + 0, \\ &= \frac{\hat{\boldsymbol{\rho}}}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial v'(\rho)}{\partial \rho}\right) - \hat{\boldsymbol{\rho}}\frac{v'(\rho)}{\rho^2}. \end{aligned}$$

Hence,

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial v'(\rho)}{\partial \rho}\right) = \frac{v'(\rho)}{\rho^2},$$

as required.

Substitution of the trial solution $v = v_0 + a_2\rho^2$ into the LHS gives

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho 2a_2\right) = \frac{2a_2}{\rho}.$$

Substitution into the RHS gives

$$\frac{v'(\rho)}{\rho^2} = \frac{2a_2\rho^2}{\rho^2} = \frac{2a_2}{\rho}.$$

Hence, L.H.S. = R.H.S., and

$$v = v_0 + a_2\rho^2 \tag{18.17}$$

is a solution.

Carrying on from the previous parts, show that the **boundary condition**

$$v = 0, \qquad \text{on } \rho = R, \text{ the pipe wall}$$

leads to the final form

$$v(\rho) = v_0 \left(1 - \frac{\rho^2}{R^2}\right).$$

Start with Equation (18.17). Note that if $v(R) = 0$, then

$$v_0 + a_2 R^2 = 0 \implies a_2 = -\frac{v_0}{R^2},$$

hence

$$v = v_0 \left(1 - \frac{\rho^2}{R^2}\right).$$

This is the celebrated **Poiseuille flow**, observed in flows in blood vessels.

Chapter 19

Special integrals involving curvilinear coordinate systems

A mathematician is someone to whom

$$\int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \mathrm{d}x = \sqrt{\pi}$$

is as obvious as $1 + 1 = 2$.

William Thomson, 1st Baron Kelvin of Largs.[1]

Overview

In this section we carry out some special integrations in various spatial dimensions. These require clever substitutions involving curvilinear coordinates.

19.1 The gamma integral

Consider the integral

$$\Gamma(n+1) = \int_0^{\infty} t^n \mathrm{e}^{-t} \mathrm{d}t, \qquad n \in \{0, 1, 2, \cdots\}.$$

If $n = 0$, the integration is easy:

$$\Gamma(1) = \int_0^{\infty} \mathrm{e}^{-t} \mathrm{d}t = -\mathrm{e}^{-t}\Big|_0^{\infty} = 1.$$

[1] William Thomson, b. 1824 Belfast, d. 1907 Largs, Scotland. Kelvin was born in Belfast but moved to Scotland as a child. There is a very impressive statue of Kelvin in the Belfast botanical gardens.

Otherwise, we do integration by parts:

$$\begin{aligned}\Gamma(n+1) &= \int_0^\infty \underbrace{t^n}_{u}\underbrace{\mathrm{e}^{-t}\mathrm{d}t}_{dv},\\ &= -t^n\mathrm{e}^{-t}\Big|_0^\infty - n\int_0^\infty \underbrace{(-\mathrm{e}^{-t})}_{v}\underbrace{t^{n-1}\mathrm{d}t}_{du},\\ &= n\int_0^\infty t^{n-1}\mathrm{e}^{-t}\mathrm{d}t,\\ &= n\Gamma(n-1).\end{aligned}$$

Now, we repeat this integration by parts until we are left with one integral evaluation, $\Gamma(1)$:

$$\Gamma(n+1) = n(n-1)\dots 2.\Gamma(1) = n(n-1)\dots 2.1 = n!$$

Thus, for $n \in \{0, 1, 2, \cdots\}$,

$$n! = \Gamma(n+1) := \int_0^\infty t^n\mathrm{e}^{-t}\mathrm{d}t.$$

Note, however, that the integral

$$\Gamma(x+1) = \int_0^\infty t^x\mathrm{e}^{-t}\mathrm{d}t$$

is valid for all $x \geq 0$, and that

$$\Gamma(x+1) = x\Gamma(x) \qquad x > 0.$$

This gives a generalization of the factorial function to positive real numbers:

$$x! := \Gamma(x+1).$$

Note: Let $t = u^2$ in $\Gamma(x+1) = \int_0^\infty t^x\mathrm{e}^{-t}\mathrm{d}t$. Then,

$$\mathrm{d}t = 2u\mathrm{d}u.$$

The integral, re-expressed in the u-variable, also ranges from 0 to $+\infty$:

$$\Gamma(x+1) = 2\int_0^\infty u^{2x+1}\mathrm{e}^{-u^2}\mathrm{d}u,$$

which is an alternative expression for the Gamma function. Setting $x = 0$ gives

$$1 = \Gamma(1) = 2\int_0^\infty u\mathrm{e}^{-u^2}\mathrm{d}u$$

19.2 The exponential integral

In this section we compute the integral

$$I := \int_{-\infty}^{\infty} \mathrm{e}^{-x^2} \mathrm{d}x.$$

First, let us derive the area element in two-dimensional polar coordinates.

In two dimensions, the spherical polar coordinates are as follows:

$$x = r\cos\varphi, \qquad y = r\sin\varphi,$$

where $r = \sqrt{x^2+y^2}$ is the distance from the origin to the point $P(x,y)$ and $\varphi = \tan^{-1}(y/x)$ is the angle between the x-axis and the radius vector $\boldsymbol{r} = \overrightarrow{OP}$. Based on the identity

$$\boldsymbol{r} \equiv \boldsymbol{x} = r\cos\varphi\hat{\boldsymbol{x}} + r\sin\varphi\hat{\boldsymbol{y}},$$

we compute the tangent vectors:

$$\frac{\partial \boldsymbol{x}}{\partial r} = \cos\varphi\hat{\boldsymbol{x}} + \sin\varphi\hat{\boldsymbol{y}} = \hat{\boldsymbol{r}},$$

since $|\cos\varphi\hat{\boldsymbol{x}} + \sin\varphi\hat{\boldsymbol{y}}|^2 = \cos^2\varphi + \sin^2\varphi = 1$,

$$\frac{\partial \boldsymbol{x}}{\partial \varphi} = -r\sin\varphi\hat{\boldsymbol{x}} + r\cos\varphi\hat{\boldsymbol{y}}, \qquad \hat{\boldsymbol{\varphi}} = -\sin\varphi\hat{\boldsymbol{x}} + \cos\varphi\hat{\boldsymbol{y}},$$

since $|-r\sin\varphi\hat{\boldsymbol{x}} + r\cos\varphi\hat{\boldsymbol{y}}|^2 = r^2$. The scale factors are thus

$$h_r = \left|\frac{\partial \boldsymbol{x}}{\partial r}\right| = 1, \qquad h_\varphi = \left|\frac{\partial \boldsymbol{x}}{\partial \varphi}\right| = r.$$

Hence, the line element is

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\varphi^2.$$

and an infinitesimal patch of area is

$$\mathrm{d}S = h_r h_\varphi \mathrm{d}r\mathrm{d}\varphi = r\mathrm{d}r\mathrm{d}\varphi.$$

But $\mathrm{d}S = \mathrm{d}x\mathrm{d}y$, hence

$$\mathrm{d}x\mathrm{d}y = r\mathrm{d}r\mathrm{d}\varphi.$$

Now we compute I. First, take

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x\right), \\ &= \left(\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x\right)\left(\int_{-\infty}^{\infty} \mathrm{e}^{-y^2}\mathrm{d}y\right), \end{aligned}$$

since x is a 'dummy variable' of integration. Re-write this as

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y\, \mathrm{e}^{-(x^2+y^2)}.$$

Now introduce polar coordinates. To enumerate all points (x, y) in the plane, the angle φ must go between 0 and 2π, and the radius vector r must go from 0 to ∞. Thus,

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y\, \mathrm{e}^{-(x^2+y^2)}, \\ &= \int_0^{2\pi}\int_0^{\infty} r\mathrm{d}r\mathrm{d}\varphi\, \mathrm{e}^{-r^2}, \\ &= \int_0^{2\pi}\mathrm{d}\varphi\int_0^{\infty} r\mathrm{d}r\, \mathrm{e}^{-r^2}, \\ &= 2\pi\int_0^{\infty} r\mathrm{d}r\, \mathrm{e}^{-r^2}, \\ &= 2\pi\int_0^{\infty} \left(-\tfrac{1}{2}\right)\frac{d}{dr}\mathrm{e}^{-r^2}\mathrm{d}r, \\ &= \pi\left[-\mathrm{e}^{-r^2}\right]_0^{\infty}, \\ &= \pi. \end{aligned}$$

Hence $I^2 = \pi$, and

$$I = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \sqrt{\pi}.$$

One final note: Recall

$$\begin{aligned} \Gamma(x+1) &= \int_0^{\infty} t^x\mathrm{e}^{-t}\mathrm{d}t, \\ &= 2\int_0^{\infty} u^{2x+1}\mathrm{e}^{-u^2}\mathrm{d}u. \end{aligned}$$

Take the second form with $x = -1/2$:

$$\Gamma\left(\tfrac{1}{2}\right) = 2\int_0^{\infty} u^0\mathrm{e}^{-u^2}\mathrm{d}u = \int_{-\infty}^{\infty} \mathrm{e}^{-u^2}\mathrm{d}u = \sqrt{\pi}.$$

Thus, for half-integers $n+\frac{1}{2}$, where $n \in \{0,1,2,\cdots\}$,

$$\left(n+\tfrac{1}{2}\right)! = \left(n+\tfrac{1}{2}\right)\left(n-\tfrac{1}{2}\right)\left(n-\tfrac{3}{2}\right)\cdots\tfrac{1}{2}\Gamma\left(\tfrac{1}{2}\right),$$

or

$$\left(n+\tfrac{1}{2}\right)! = \left(n+\tfrac{1}{2}\right)\left(n-\tfrac{1}{2}\right)\left(n-\tfrac{3}{2}\right)\cdots\tfrac{1}{2}\sqrt{\pi}.$$

19.3 The volume of an n-ball

In n dimensions, the ball centred at 0 of radius r is a subset of $\mathbb{R}^n$ such that

$$x_1^2 + x_2^2 + \cdots x_n^2 \leq r^2.$$

We would like to find the volume of this ball:

$$V_n(r) = \int \cdots \int_{x_1^2+\cdots+x_n^2 \leq r^2} \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

In analogy with polar coordinates in two-dimensional space, let us write the volume element as

$$\mathrm{d}x_1 \cdots \mathrm{d}x_n = r^{n-1}\mathrm{d}r\mathrm{d}\Omega_n,$$

where $\mathrm{d}\Omega_n$ is a differential involving angles $\varphi_1, \cdots, \varphi_{n-1}$ that are unspecified polar coordinates on the sphere in $\mathbb{R}^n$. It is not necessary to know what these angles are, suffice to say that

$$\mathrm{d}\Omega_n = f\left(\varphi_1, \cdots, \varphi_{n-1}\right)\mathrm{d}\varphi_1 \cdots \mathrm{d}\varphi_{n-1},$$

where $f(\cdots)$ is some function. The differential $\mathrm{d}\Omega_n$ is the element of **solid angle** in n dimensions, and its integral over all possible values of $\varphi_1, \cdots, \varphi_{n-1}$ gives the **surface area** of the unit sphere in n dimensions, $S_n(1)$. Thus,

$$\begin{aligned} V_n(r) &= \int \cdots \int_{x_1^2+\cdots+x_n^2 \leq r^2} \mathrm{d}x_1 \cdots \mathrm{d}x_n, \\ &= \int \mathrm{d}\Omega_n \int_0^r r^{n-1}\mathrm{d}r, \\ &= \frac{S_n(1) r^n}{n}. \end{aligned}$$

This gives a relationship between surface area and volume in n dimensions.

Now, consider the integral

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{x}^2}\mathrm{d}^n x &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \mathrm{e}^{-x_1^2-\cdots x_n^2}\mathrm{d}x_1\cdots\mathrm{d}x_n,\\
&= \int_{-\infty}^{\infty} \mathrm{e}^{-x_1^2}\mathrm{d}x_1\cdots\int_{-\infty}^{\infty} \mathrm{e}^{-x_n^2}\mathrm{d}x_n,\\
&= I^n,\\
&= \pi^{n/2}. \qquad (*)
\end{aligned}$$

But we can write $\int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{x}^2}\mathrm{d}^n x$ in general spherical polar coordinates as

$$\begin{aligned}
\int_{-\infty}^{\infty} \mathrm{e}^{-\boldsymbol{x}^2}\mathrm{d}^n x &= \int \mathrm{d}\Omega_n \int_0^{\infty} r^{n-1}\mathrm{e}^{-r^2}\,\mathrm{d}r,\\
&= \tfrac{1}{2}S_n(1)\Gamma(n/2). \qquad (**)
\end{aligned}$$

Equating (*) and (**) gives

$$S_n(1) = \frac{2\pi^{n/2}}{\Gamma(n/2)},$$

hence,

$$V_n(r) = \frac{2\pi^{n/2}}{n\Gamma(n/2)}r^n.$$

Check: $n=2$ gives $V_2(r) = \pi r^2$, $n=3$ gives

$$\frac{2\pi\sqrt{\pi}}{3\frac{1}{2}\sqrt{\pi}}r^3 = \tfrac{4}{3}\pi r^3.$$

19.4 The Jacobian

Recall, in two dimensions, in spherical polar coordinates,

$$\mathrm{d}S = \mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\varphi = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix} \mathrm{d}r\,\mathrm{d}\varphi.$$

The determinant

$$J := \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix}$$

is called the **Jacobian** of the coordinate transformation $(x,y) \to (r,\varphi)$. In general, in n dimensions, given a coordinate transformation

$$q_1 = q_1(x_1, \cdots x_n), \cdots, q_n = q_n(x_1, \cdots x_n),$$

the volume element has the form

$$\mathrm{d}V_n = \mathrm{d}x_1 \cdots \mathrm{d}x_n = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix} \mathrm{d}q_1 \cdots \mathrm{d}q_n,$$

and

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix}$$

is the **Jacobian**. For orthogonal coordinate systems, this always reduces to

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix} = h_{q_1} \cdots h_{q_n}.$$

We prove these facts now.

Proof: Form the tangent vectors

$$\boldsymbol{t}_1 = \frac{\partial \boldsymbol{x}}{\partial q_1}, \cdots, \boldsymbol{t}_n = \frac{\partial \boldsymbol{x}}{\partial q_n},$$

where

$$\hat{\boldsymbol{q}}_i = \boldsymbol{t}_i / |\boldsymbol{t}_i|.$$

Recall, in two dimensions,

$$\mathrm{d}V_2 = |\boldsymbol{t}_1 \times \boldsymbol{t}_2| \, \mathrm{d}q_1 \, \mathrm{d}q_2,$$

in three dimensions,

$$\mathrm{d}V_3 = |\boldsymbol{t}_1 \cdot (\boldsymbol{t}_2 \times \boldsymbol{t}_3)| \, \mathrm{d}q_1 \, \mathrm{d}q_2 \, \mathrm{d}q_3.$$

In both cases, we have the formula

$$\mathrm{d}V_n = \begin{vmatrix} | & \cdots & | \\ \boldsymbol{t}_1 & \cdots & \boldsymbol{t}_n \\ | & \cdots & | \end{vmatrix} \mathrm{d}q_1 \cdots \mathrm{d}q_n, \qquad n = 2, 3,$$

where the i^{th} column of this determinant is the column vector $\boldsymbol{t}_i$. Now there is nothing special about dimensions $n = 2$ or $n = 3$, so this formula must hold in an arbitrary spatial dimension:

$$\mathrm{d}V_n = \begin{vmatrix} | & | & | \\ \boldsymbol{t}_1 & | & \boldsymbol{t}_n \\ | & | & | \end{vmatrix} \mathrm{d}q_1 \cdots \mathrm{d}q_n, \qquad n \in \{1, 2, \cdots\}.$$

In other words,

$$\mathrm{d}V_n = \mathrm{d}x_1 \cdots \mathrm{d}x_n = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial x_n}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix} \mathrm{d}q_1 \cdots \mathrm{d}q_n.$$

For orthogonal curvilinear coordinates,

$$\begin{aligned} \mathrm{d}V_n &= \begin{vmatrix} | & | & | \\ \boldsymbol{t}_1 & | & \boldsymbol{t}_n \\ | & | & | \end{vmatrix} \mathrm{d}q_1 \cdots \mathrm{d}q_n, \\ &= \begin{vmatrix} | & | & | \\ \hat{\boldsymbol{q}}_1 & | & \hat{\boldsymbol{q}}_n \\ | & | & | \end{vmatrix} h_1 \cdots h_n \, \mathrm{d}q_1 \cdots \mathrm{d}q_n, \\ &= \left| P \begin{pmatrix} | & | & | \\ \hat{\boldsymbol{q}}_1 & | & \hat{\boldsymbol{q}}_n \\ | & | & | \end{pmatrix} P^T \right| h_1 \cdots h_n \, \mathrm{d}q_1 \cdots \mathrm{d}q_n, \end{aligned}$$

where P is an orthogonal matrix $|PP^T| = 1$ that rotates the matrix

$$\begin{pmatrix} | & | & | \\ \hat{\boldsymbol{q}}_1 & | & \hat{\boldsymbol{q}}_n \\ | & | & | \end{pmatrix}$$

into the identity matrix,

$$P \begin{pmatrix} | & | & | \\ \hat{\boldsymbol{q}}_1 & | & \hat{\boldsymbol{q}}_n \\ | & | & | \end{pmatrix} P^T = \mathbb{I}_n.$$

Thus,

$$\mathrm{d}V_n = h_1 \cdots h_n \, \mathrm{d}q_1 \cdots \mathrm{d}q_n,$$

as required.

In conclusion, for orthogonal coordinates,

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial q_1} & \cdots & \frac{\partial x_1}{\partial q_n} \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial q_1} & \cdots & \frac{\partial x_n}{\partial q_n} \end{vmatrix} = h_1 \cdots h_n.$$

19.5 The ball in $\mathbb{R}^4$

In this section we construct coordinates for the ball in $\mathbb{R}^4$ and compute its volume from this construction. Our approach for developing coordinates is based on an analogy with three dimensional space.

Recall the construction of a ball in $\mathbb{R}^3$. Topologically, we take two identical discs (balls in $\mathbb{R}^2$) and let them sit one on top of the other. We glue the boundary edges of these two balls together. We then 'inflate' the glued-together object so that the two discs are pushed in opposite directions into the third dimension. In terms of coordinates, this construction is summarized by the augmentation of the two-dimensional coordinate system

$$\begin{aligned} x &= r\cos\varphi, \\ y &= r\sin\varphi, \qquad 0 \le \varphi < 2\pi, \qquad r = \sqrt{x^2+y^2}, \end{aligned}$$

to the following form:

$$\begin{aligned} z &= r\cos\theta, \\ x &= \sin\theta\, r\cos\varphi, \\ y &= \sin\theta\, r\sin\varphi, \end{aligned}$$

where

$$0 \le \varphi < 2\pi, \qquad 0 \le \theta < \pi, \qquad r = \sqrt{x^2+y^2+z^2}.$$

We now repeat the same steps: A ball in four dimensions is constructed from two identical three-balls. We sit these balls one on top of the other and glue their boundaries together (these boundaries are actually spheres). We then 'inflate' this object so that the two ball-interiors are pushed in opposite directions into the fourth dimension. In coordinate terms, we have

$$\begin{aligned} w &= r\cos\psi, \\ z &= \sin\psi\ r\cos\theta, \\ y &= \sin\psi\ r\sin\theta\sin\varphi, \\ x &= \sin\psi\ r\sin\theta\cos\varphi, \end{aligned}$$

where

$$0 \le \varphi < 2\pi, \qquad 0 \le \theta < \pi, \qquad 0 \le \psi < \pi, \qquad r = \sqrt{x^2+y^2+z^2+w^2}.$$

For notational convenience, we re-write this system as

$$\begin{aligned} x_1 &= r\cos\psi, \\ x_2 &= r\sin\psi\cos\theta, \\ x_3 &= r\sin\psi\sin\theta\sin\varphi, \\ x_4 &= r\sin\psi\sin\theta\cos\varphi. \end{aligned}$$

Now a general vector $\boldsymbol{x}$ in $\mathbb{R}^4$ is written as

$$\boldsymbol{x} = \hat{\boldsymbol{e}}_1 x_1 + \hat{\boldsymbol{e}}_2 x_2 + \hat{\boldsymbol{e}}_3 x_3 + \hat{\boldsymbol{e}}_4 x_4,$$

where

$$\begin{aligned} \hat{\boldsymbol{e}}_1 &= (1,0,0,0)\,, \\ \hat{\boldsymbol{e}}_2 &= (0,1,0,0)\,, \\ \hat{\boldsymbol{e}}_3 &= (0,0,1,0)\,, \\ \hat{\boldsymbol{e}}_4 &= (0,0,0,1)\,. \end{aligned}$$

Hence,

$$\boldsymbol{x} = \hat{\boldsymbol{e}}_1 r\cos\psi + \hat{\boldsymbol{e}}_2 r\sin\psi\cos\theta + \hat{\boldsymbol{e}}_3 r\sin\psi\sin\theta\cos\varphi + \hat{\boldsymbol{e}}_4 r\sin\psi\sin\theta\sin\varphi.$$

Now, we can compute tangent vectors.

Clearly,

$$\hat{\boldsymbol{r}} = \frac{\partial \boldsymbol{x}}{\partial r} = \hat{\boldsymbol{e}}_1\cos\psi + \hat{\boldsymbol{e}}_2\sin\psi\cos\theta + \hat{\boldsymbol{e}}_3\sin\psi\sin\theta\cos\varphi + \hat{\boldsymbol{e}}_4\sin\psi\sin\theta\sin\varphi.$$

is the radial tangent vector with unit norm. Next,

$$\frac{\partial \boldsymbol{x}}{\partial \psi} = -\hat{\boldsymbol{e}}_1 r \sin\psi + \hat{\boldsymbol{e}}_2 r \cos\psi \cos\theta + \hat{\boldsymbol{e}}_3 r \cos\psi \sin\theta \cos\varphi + \hat{\boldsymbol{e}}_4 r \cos\psi \sin\theta \sin\varphi.$$

with norm r, hence

$$\hat{\boldsymbol{\psi}} = -\hat{\boldsymbol{e}}_1 \sin\psi + \hat{\boldsymbol{e}}_2 \cos\psi \cos\theta + \hat{\boldsymbol{e}}_3 \cos\psi \sin\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \cos\psi \sin\theta \sin\varphi.$$

Again,

$$\frac{\partial \boldsymbol{x}}{\partial \theta} = \hat{\boldsymbol{e}}_1 0 + r \sin\psi \left[-\hat{\boldsymbol{e}}_2 \sin\theta + \hat{\boldsymbol{e}}_3 \cos\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \cos\theta \sin\varphi\right],$$

with norm $r\sin\psi$, hence

$$\hat{\boldsymbol{\theta}} = -\hat{\boldsymbol{e}}_2 \sin\theta + \hat{\boldsymbol{e}}_3 \cos\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \cos\theta \sin\varphi.$$

Finally,

$$\frac{\partial \boldsymbol{x}}{\partial \varphi} = \hat{\boldsymbol{e}}_1 0 + \hat{\boldsymbol{e}}_2 0 + r \sin\psi \sin\theta \left[-\hat{\boldsymbol{e}}_3 \sin\varphi + \hat{\boldsymbol{e}}_4 \cos\varphi\right],$$

with norm $r \sin\psi \sin\theta$, hence

$$\hat{\boldsymbol{\varphi}} = -\hat{\boldsymbol{e}}_3 \sin\varphi + \hat{\boldsymbol{e}}_4 \cos\varphi.$$

Let's assemble these results.

- Tangent vectors:

$$\begin{aligned}
\hat{\boldsymbol{r}} &= \hat{\boldsymbol{e}}_1 \cos\psi + \hat{\boldsymbol{e}}_2 \sin\psi \cos\theta + \hat{\boldsymbol{e}}_3 \sin\psi \sin\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \sin\psi \sin\theta \sin\varphi, \\
\hat{\boldsymbol{\psi}} &= -\hat{\boldsymbol{e}}_1 \sin\psi + \hat{\boldsymbol{e}}_2 \cos\psi \cos\theta + \hat{\boldsymbol{e}}_3 \cos\psi \sin\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \cos\psi \sin\theta \sin\varphi, \\
\hat{\boldsymbol{\theta}} &= -\hat{\boldsymbol{e}}_2 \sin\theta + \hat{\boldsymbol{e}}_3 \cos\theta \cos\varphi + \hat{\boldsymbol{e}}_4 \cos\theta \sin\varphi, \\
\hat{\boldsymbol{\varphi}} &= -\hat{\boldsymbol{e}}_3 \sin\varphi + \hat{\boldsymbol{e}}_4 \cos\varphi.
\end{aligned}$$

- Scale factors:

$$\begin{aligned}
h_r &= 1, \\
h_\psi &= r, \\
h_\theta &= r \sin\psi, \\
h_\varphi &= r \sin\psi \sin\theta.
\end{aligned}$$

It is straightforward to check that these vectors are orthogonal: there are $(4-1)! = 6$ relations to check. For example,

$$\begin{aligned}
\hat{\boldsymbol{r}}\cdot\hat{\boldsymbol{\psi}} &= \left[\hat{\boldsymbol{e}}_1\cos\psi+\hat{\boldsymbol{e}}_2\sin\psi\cos\theta+\hat{\boldsymbol{e}}_3\sin\psi\sin\theta\cos\varphi+\hat{\boldsymbol{e}}_4\sin\psi\sin\theta\sin\varphi\right]\\
&\quad\cdot\left[-\hat{\boldsymbol{e}}_1\sin\psi+\hat{\boldsymbol{e}}_2\cos\psi\cos\theta+\hat{\boldsymbol{e}}_3\cos\psi\sin\theta\cos\varphi+\hat{\boldsymbol{e}}_4\cos\psi\sin\theta\sin\varphi\right],\\
&= -\cos\psi\sin\psi+\sin\psi\cos\psi\left[\cos^2\theta+\sin^2\theta\left(\cos^2\varphi+\sin^2\varphi\right)\right],\\
&= -\cos\psi\sin\psi+\sin\psi\cos\psi=0.
\end{aligned}$$

Now let's compute the volume of the four-ball:

$$\begin{aligned}
V_4 &= \int_0^R \mathrm{d}r\int_0^\pi \mathrm{d}\psi\int_0^\pi \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\, h_r h_\psi h_\theta h_\varphi,\\
&= \int_0^R \mathrm{d}r\int_0^\pi \mathrm{d}\psi\int_0^\pi \mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\, r^3\sin\psi\sin^2\theta,\\
&= \left(\int_0^R r^3\mathrm{d}r\right)\left(\int_0^\pi \mathrm{d}\psi\,\sin^2\psi\right)\left(\int_0^\pi \mathrm{d}\theta\,\sin\theta\right)\left(\int_0^{2\pi}\mathrm{d}\varphi\right),\\
&= \left(\frac{1}{4}r^4\right)\left[\tfrac{1}{2}\left(\psi-\sin\psi\cos\psi\right)_0^\pi\right]\left(-\cos\pi+\cos 0\right)2\pi,\\
&= \tfrac{1}{2}\pi^2 r^4.
\end{aligned}$$

Check against the general formula:

$$\begin{aligned}
V_n &= \frac{2\pi^{n/2}}{n\Gamma(n/2)}r^4,\\
&= \frac{2\pi^2}{4\Gamma(2)}r^4, \qquad n=4,\\
&= \frac{2\pi^2}{4\cdot 1!}r^4,\\
&= \tfrac{1}{2}\pi^2 r^4.
\end{aligned}$$

19.6 One more integral

The last integral in this chapter is the following one:

$$I(\boldsymbol{x})=\int_{-\infty}^{\infty}\mathrm{d}k_x\int_{-\infty}^{\infty}\mathrm{d}k_y\int_{-\infty}^{\infty}\mathrm{d}k_z\,\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}}{1+\boldsymbol{k}^2},\qquad \boldsymbol{k}=(k_x,k_y,k_z).$$

First, let us re-write this in a more suggestive form:

$$I(\boldsymbol{x})=\int \mathrm{d}^3k\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}}{1+k^2},$$

where the range is implicit and is equal to the whole of $\mathbb{R}^3$.

To do this integral, we go over to polar coordinates in $\boldsymbol{k}$:

$$
\begin{aligned}
k_z &= k\cos\theta, \\
k_y &= k\sin\theta\sin\varphi, \\
k_x &= k\sin\theta\cos\varphi, \qquad k = \sqrt{k_x^2+k_y^2+k_z^2}.
\end{aligned}
$$

As usual,

$$\mathrm{d}^3k = k^2\sin\theta\,\mathrm{d}k\mathrm{d}\theta\mathrm{d}\varphi.$$

Hence,

$$I = \int_0^\infty k^2\mathrm{d}k\int_0^\pi \sin\theta\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\,\frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}}{1+k^2}.$$

We choose a coordinate system in $\boldsymbol{x}$-space such that $\boldsymbol{x}$ aligns with the k_z-axis. Then,

$$\boldsymbol{k}\cdot\boldsymbol{x} = k|\boldsymbol{x}|\cos\theta,$$

and

$$
\begin{aligned}
I(\boldsymbol{x}) &= \int_0^\infty k^2\mathrm{d}k\int_0^\pi \sin\theta\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{x}|\cos\theta}}{1+k^2}, \\
&= 2\pi\int_0^\infty \frac{k^2}{1+k^2}\mathrm{d}k\int_0^\pi \sin\theta\mathrm{d}\theta\,\mathrm{e}^{\mathrm{i}k|\boldsymbol{x}|\cos\theta}
\end{aligned}
$$

Now we use a neat trick:

$$\sin\theta\mathrm{e}^{\mathrm{i}kx\cos\theta} = -\frac{1}{\mathrm{i}kx}\frac{d}{d\theta}\mathrm{e}^{\mathrm{i}kx\cos\theta}.$$

Hence,

$$
\begin{aligned}
I(\boldsymbol{x}) &= 2\pi\int_0^\infty \frac{k^2}{1+k^2}\mathrm{d}k\int_0^\pi \sin\theta\mathrm{d}\theta\,\mathrm{e}^{\mathrm{i}k|\boldsymbol{x}|\cos\theta}, \\
&= 2\pi\int_0^\infty \mathrm{d}k\frac{k^2}{1+k^2}\frac{\mathrm{i}}{kx}\int_0^\pi \mathrm{d}\theta\frac{d}{d\theta}\mathrm{e}^{\mathrm{i}kx\cos\theta}, \\
&= 2\pi\int_0^\infty \mathrm{d}k\,\frac{k^2}{1+k^2}\frac{\mathrm{i}}{kx}\left[\mathrm{e}^{-\mathrm{i}kx}-\mathrm{e}^{\mathrm{i}kx}\right], \\
&= \frac{4\pi}{x}\int_0^\infty \mathrm{d}k\frac{k\sin(kx)}{1+k^2}, \\
&= \frac{2\pi}{x}\int_{-\infty}^\infty \mathrm{d}k\frac{k\sin(kx)}{1+k^2}.
\end{aligned}
$$

In another course, you will hopefully be exposed to complex-variable theory, which determines this integral through Cauchy's residue theorem: $\int_{-\infty}^\infty \mathrm{d}k\cdots = \pi\mathrm{e}^{-x}$, hence

$$I(\boldsymbol{x}) = \frac{2\pi}{x}\left(2\pi\frac{\mathrm{e}^{-x}}{2}\right) = 2\pi^2\frac{\mathrm{e}^{-x}}{x},$$

and the final answer is a function of the *scalar* $x = |\boldsymbol{x}|$.

This completes the chapter about special integrals.

Chapter 20

Fin

Vector calculus was invented by mathematical physicists to formulate Electromagnetism.[1] It is thus the mathematical basis of Electromagnetism, and it also provides the mathematical key to understanding fluid mechanics, quantum mechanics, heat and mass transfer, and partial differential equations. When combined with geometry, such that differential laws can be formulated in non-flat spaces, one has the mathemtical tools at hand to study Relativity and Quantum Field Theory. It is thus indispensable in mathematical physics. I hope this module has succeeded in creating a foundation for you to study these topics in more detail in later years.

[1] *Vector analysis, a text-book for the use of students of mathematics and physics, founded upon the lectures of J. Willard Gibbs*, E. B. Wilson and J. W. Gibbs (1902)

Appendix A

The ratio test – proof

One way to test if a power series converges is the **ratio test**:

Theorem A.1 (Ratio Test) *Let*

$$b_0 + b_1 + b_2 + \cdots$$

be an infinite series. Form the ratio

$$\rho = \lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right|.$$

Then

- *If* $\rho < 1$ *the series is convergent.*
- *If* $\rho > 1$ *the series is divergent.*
- *If* $\rho = 1$ *the ratio test is inconclusive.*

Proof: The limit

$$\rho = \lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right|$$

means that there exists a positive integer N such that

$$\left| \frac{b_{n+1}}{b_n} \right| < \rho$$

for all $n > N$. Hence,

$$|b_{n+1}| < |b_n|\rho \qquad \forall n > N.$$

Telescope this result:

$$\begin{aligned}
|b_{N+2}| &< |b_{N+1}|\rho, \\
|b_{N+3}| &< |b_{N+2}|\rho < |b_{N+1}|\rho^2, \\
|b_{N+4}| &< |b_{N+3}|\rho < |b_{N+1}|\rho^3, \\
\vdots \quad & \quad \vdots \\
|b_{N+j}| &< |b_{N+1}|\rho^{j-1}, \qquad j = 2, 3 \cdots, \\
|b_{N+j}| &\leq |b_{N+1}|\rho^{j-1}, \qquad j = 1, 2, 3 \cdots,
\end{aligned}$$

Now, split the infinite series $b_0 + b_1 + \cdots$ into two:

$$b_0 + b_1 + b_2 + \cdots = \sum_{n=0}^{N} b_i + \sum_{n=N+1}^{\infty} b_i.$$

By repeated application of the triangle inequality $|x + y| \leq |x| + |y|$ for all real numbers x, y, and z, we have

$$\begin{aligned}
|b_0 + b_1 + b_2 + \cdots| &= \left|\sum_{j=0}^{N} b_j\right| + \left|\sum_{j=N+1}^{\infty} b_j\right|, \\
&\leq \left|\sum_{j=0}^{N} b_j\right| + \sum_{j=N+1}^{\infty} |b_j|, \\
&= \left|\sum_{j=0}^{N} b_j\right| + \sum_{j=N+1}^{\infty} |b_j|, \\
&= \left|\sum_{j=0}^{N} b_j\right| + \sum_{j=1,2,3,\cdots}^{\infty} |b_{N+j}|, \\
&\leq \left|\sum_{j=0}^{N} b_j\right| + \sum_{j=1,2,3,\cdots}^{\infty} |b_{N+1}|\,\rho^{j-1}.
\end{aligned}$$

We can now use our knowledge of geometric progressions to add up the infinite series here:

- For $\rho < 1$ we obtain
$$|b_0 + b_1 + \cdots| < \left|\sum_{j=0}^{N} b_j\right| + \frac{|b_{N+1}|}{1-\rho} < \infty.$$

- For $\rho > 1$ we obtain
$$|b_0 + b_1 + \cdots| = \infty.$$

- For $\rho = 1$ we cannot sum the geometric series and the test is inconclusive.