Mathematical Methods in Physics (Part I)[1]

Dr Cristina Zambon

13 April 2020

[1] *Despite the effort to eliminate all typographic errors, some of them could still be present. Hence be careful. Note that this summary is intended as a guideline for the materials covered in lectures and it is not supposed to replace the textbook.*

Introduction

The aim of the course is to teach you the mathematical tools required in order to tackle successfully future physics courses. The techniques acquired during this course will allow you to solve more diverse and difficult problems in physics, which will increase your understanding of physics itself.

Required textbook: *Mathematical Methods for Physics and Engineering*, K. F. Riley, M. P. Hobson, S. J. Bence.

Note that the content of this course is pretty standard for a module on Mathematical Methods in Physics and all the material can be found in any book on Mathematical Methods or Mathematical Techniques for Physicists that you can find in the library. Feel free to use any books that suit your style.

Syllabus:

- Vector algebra - Lectures 1,2.
- Vector spaces - Lectures 2,3.
- Matrices - Lectures 4,5,6.
- Fourier series - Lectures 7,8.
- Integral transforms (Fourier and Laplace transforms) - Lectures 9,10,11,12.
- Dirac δ function - Lectures: 10,11.
- Vector calculus (scalar and vector functions, *del* operator) - Lectures 13,14.
- Curves and line integrals - Lectures 14, 15.
- Surfaces, surface integrals and volume integrals - Lectures 16,17.
- Gauss' and Stokes' theorems - Lectures 17, 18.
- Orthogonal curvilinear coordinates - Lecture 19.

Typed **lecture notes** will be available on DUO. Note that they do **NOT** contain the solution for the examples provided in classes, so please take notes.

The emphasis of the course is on being able to solve exercises not on learning rigorous mathematical proofs. Therefore, the key for completing the course successfully is **practice**. In order to practise you have

i) Self-assessed formative **weekly problems** (8).

ii) **Workshops** (9). Workshop exercises will appear on DUO before workshops. Solutions will also appear on DUO, after workshops.

iii) **Quizzes** on **Jupyter Notebooks** to test your understanding of each single lecture.

The quizzes are available at https://dmaitre.phyip3.dur.ac.uk/notebooks/quiz/mmp/login/. To access the quizzes use your CIS login details.

iv) **Midterm formative test:** Thursday 14 November (week six of term).

Feel free to contact me via email, **cristina.zambon@durham.ac.uk**. I am in Ph210. An appointment can be arranged via email. I will also attend all workshops. They provide an excellent opportunity for asking questions.

Before starting lectures, **ensure that you are familiar with the mathematical concepts that follows**. This material will not be covered in any details during lectures.

Vector Algebra (Chapter 7 in Riley)

Consider the three dimensional space with basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ (Standard basis).

- **Scalar (or dot) product**:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

where θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$.

i) $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$.

ii) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.

iii) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}, \quad \mathbf{a} \cdot (\beta\, \mathbf{b}) = \beta\, \mathbf{a} \cdot \mathbf{b}$, where β is a scalar.

iv) If the scalar product of two vectors is zero, then the vectors are perpendicular.

- **Vector (or cross) product**:

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2)\,\mathbf{i} + (a_3b_1 - a_1b_3)\,\mathbf{j} + (a_1b_2 - a_2b_1)\,\mathbf{k}, \quad |\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta,$$
$$(\mathbf{a}\times\mathbf{b})_i$$

where θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$.

i) $|\mathbf{a}\times\mathbf{b}|$ is the area of the parallelogram with sides $\mathbf{a}$ and $\mathbf{b}$.

ii) $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$.

iii) $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b} + \mathbf{a}\times\mathbf{c}, \quad \mathbf{a}\times(\beta\,\mathbf{b}) = \beta\,\mathbf{a}\times\mathbf{b}$.

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

v) If the vector product of two vectors is zero, then the vectors are parallel or antiparallel.

- **Scalar triple product:**

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

i) $|\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})|$ is the volume of the parallelepiped with sides $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.

ii) $[\alpha\,\mathbf{a}+\beta\,\mathbf{b},\mathbf{c},\mathbf{d}] = \alpha\,[\mathbf{a},\mathbf{c},\mathbf{d}] + \beta\,[\mathbf{b},\mathbf{c},\mathbf{d}]$.

iii) $[\mathbf{a},\mathbf{b},\mathbf{c}] = [\mathbf{b},\mathbf{c},\mathbf{a}] = [\mathbf{c},\mathbf{a},\mathbf{b}] = -[\mathbf{a},\mathbf{c},\mathbf{b}] = -[\mathbf{c},\mathbf{b},\mathbf{a}] = -[\mathbf{b},\mathbf{a},\mathbf{c}]$

iv) If the scalar triple product of three vectors is zero, then the vectors are coplanar.

- **Vector triple product:**

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

i) The vector triple product is not associative.

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

Matrices (Chapter 8 in Riley)

- **Matrix operations:**

i) Matrix addition: $(A+B)_{ij} = A_{ij} + B_{ij}$.

ii) Multiplication by a scalar: $(\alpha A)_{ij} = \alpha A_{ij}$.

iii) Multiplication of matrices: $(AB)_{ij} = A_{ik}B_{kj}$, with $AB \neq BA$.

iv) Transposition: $(A^T)_{ij} = A_{ji}$, with $(ABC\ldots F)^T = F^T\ldots C^TB^TA^T$.

v) Complex conjugation: $(A^*)_{ij} = (A_{ij})^*$.

vi) Hermitian conjugation (adjoint): $(A^\dagger)_{ij} = (A_{ji})^*$, with
$(ABC\dots F)^\dagger = F^\dagger \dots C^\dagger B^\dagger A^\dagger$.

- **The determinant of a square matrix:**

$$|A| = A_{jk}C_{jk}, \quad \text{for any row } j, \qquad |A| = A_{kj}C_{kj}, \quad \text{for any column } j,$$

where $C_{mn} = (-1)^{m+n}|A_{mn}|$ is the *cofactor* associated to the matrix element A_{mn}. In turn, $|A_{mn}|$ is the *minor* associated to the matrix element A_{mn}. The minor is the determinant of the matrix obtained by removing the m-th row and n-th column from the matrix A.

Properties:

i) $|AB\dots F| = |A||B|\dots|F|$.

ii) $|A^T| = |A|, \quad |A^*| = |A|^*, \quad |A^\dagger| = |A|^*, \quad |A^{-1}| = |A|^{-1}$.

iii) If the rows (or the columns) are linearly dependent, then $|A| = 0$.

iv) If B is obtained from A by multiplying the elements of any row (or column) by a factor α, then $|B| = \alpha\,|A$.

v) If B is obtained from A by interchanging two rows (or columns), then $|B| = -|A|$.

vi) If B is obtained from A by adding k times one row (or column) to the other row (or column), then $|A| = |B|$.

- **Elementary row operations (on matrices):**

i) Multiply any row by a non zero constant.

ii) Interchange any two rows.

iii) Add some multiple of one row to any other row.

- **The inverse of a square matrix:**

$$A^{-1} = \frac{C^T}{|A|}, \quad \text{that is} \quad A^{-1}_{ij} = \frac{C_{ji}}{|A|}, \qquad A^{-1}A = AA^{-1} = I,$$

where C is the cofactor matrix and I the identity matrix ($I_{ij} = \delta_{ij}$). If $|A| = 0$ the inverse does not exist and the matrix A is said to be *singular*.

Note that in order to find the inverse of a matrix, you can also use the Gauss-Jordan method shown in the lectures, which makes use of the elementary row operations.

Properties:

i) $(AB\dots F)^{-1} = F^{-1}\dots B^{-1}A^{-1}$.

ii) $(A^T)^{-1} = (A^{-1})^T, \quad (A^\dagger)^{-1} = (A^{-1})^\dagger$.

Contents

Chapter 1

Vector Algebra in $\mathbb{R}^3$

- **Orthonormal vectors:** The vectors $\mathbf{a_1}, \mathbf{a_2}, \cdots, \mathbf{a_n}$ are said to be orthonormal if they are unit vectors mutually orthogonal. A good example of a set or orthonormal vectors is the standard basis in $\mathbb{R}^3$ i.e. $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$.

- **Summation convention:** An index that appears twice in a given term is understood to be summed over all values that the index can take. In the case of the space $\mathbb{R}^3$, the indices can take the values ,1,2,3. The summed over indices are called called *dummy indices* and the other *free indices*.

Example 1 Expand the following expressions

(i) $a_{ij}b_{jk} \equiv \sum_{j}^{3} a_{ij}b_{jk} = a_{i1}b_{1k} + a_{i2}b_{2k} + a_{i3}b_{3k}$.

(ii) $a_{ij}b_{jk}c_k \equiv \sum_{j}^{3} \sum_{k}^{3} a_{ij}b_{jk}c_k = \sum_{j}^{3} (a_{ij}b_{j1}c_1 + a_{ij}b_{j2}c_2 + a_{ij}b_{j3}c_3)$

$$
\begin{aligned}
&= (a_{i1}b_{11}c_1 + a_{i1}b_{12}c_2 + a_{i1}b_{13}c_3) + (a_{i2}b_{21}c_1 + a_{i2}b_{22}c_2 + a_{i2}b_{23}c_3) \\
&+ (a_{i3}b_{31}c_1 + a_{i3}b_{32}c_2 + a_{i3}b_{33}c_3).
\end{aligned}
$$

Let us introduce two mathematical objects that can be used in the context of the summation convention:

- **Kronecker delta:**

$$
\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \qquad i, j = 1, 2, 3
$$

Note that this object is symmetric.

Example 2 Use the index notation to rewrite the following expressions

(i) $b_i\delta_{ij} = b_1\delta_{1j} + b_2\delta_{2j} + b_3\delta_{3j} = b_j$.

(ii) $\delta_{jj} = \delta_{11} + \delta_{22} + \delta_{33} = 3$.

(iii) $\mathbf{a}\cdot\mathbf{b} = a_ib_i = \delta_{ij}a_ib_j$.

- **Levi-Civita symbol:**

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } (i,j,k) = (1,2,3) = (2,3,1) = (3,1,2), \\ -1 & \text{if } (i,j,k) = (1,3,2) = (3,2,1) = (2,1,3), \\ 0 & \text{otherwise.} \end{cases}$$

Note that this object is totally antisymmetric.

Example 3 Use the index notation to rewrite the following expressions

(i) $(\mathbf{a}\times\mathbf{b})_i = \epsilon_{ijk}a_jb_k$. For instance:
$(\mathbf{a}\times\mathbf{b})_1 = \epsilon_{1jk}a_jb_k = \epsilon_{123}a_2b_3 + \epsilon_{132}a_3b_2 = a_2b_3 - a_3b_2$.

(ii) $(\mathbf{b}\times\mathbf{a})_i = \epsilon_{1jk}b_ja_k = \epsilon_{1jk}a_kb_j = -\epsilon_{1kj}a_kb_j = -(\mathbf{a}\times\mathbf{b})_i$.

(iii) $[\mathbf{a},\mathbf{b},\mathbf{c}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = a_i(\mathbf{b}\times\mathbf{c})_i = a_i\epsilon_{ijk}b_jc_k = \epsilon_{ijk}a_ib_jc_k$

$$\begin{aligned} &= a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_3b_2c_1 - a_2b_1c_3 \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \end{aligned}$$

- **Scalar (or dot) product**:

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta = a_1b_1 + a_2b_2 + a_3b_3 = a_ib_i.$$

where θ is the angle between the vectors $\mathbf{a}$ and $\mathbf{b}$.

i) $|\mathbf{a}|^2 = \mathbf{a}\cdot\mathbf{a}$.

ii) $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$.

iii) $\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) = \mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}, \quad \mathbf{a}\cdot(\beta\,\mathbf{b}) = \beta\,\mathbf{a}\cdot\mathbf{b}$, where β is a scalar.

- **Orthogonal projection of b onto the direction of a ($\mathbf{b}_\parallel$)**:

$$\mathbf{b} = \mathbf{b}_\parallel + \mathbf{b}_\perp, \quad \mathbf{b}_\parallel = (|\mathbf{b}|\cos\theta)\,\frac{\mathbf{a}}{|\mathbf{a}|} = \frac{(\mathbf{a}\cdot\mathbf{b})}{|\mathbf{a}|}\,\frac{\mathbf{a}}{|\mathbf{a}|}.$$

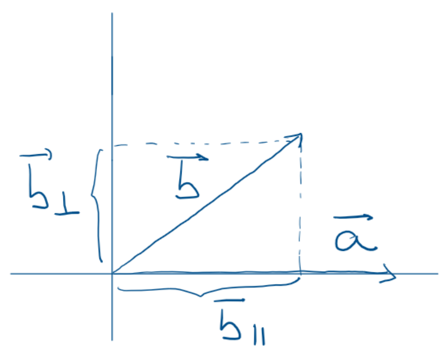

Figure 1.1: Orthogonal projection

- **Vector (or cross) product**:

$$\mathbf{a}\times\mathbf{b} = (a_2b_3 - a_3b_2)\,\mathbf{i} + (a_3b_1 - a_1b_3)\,\mathbf{j} + (a_1b_2 - a_2b_1)\,\mathbf{k}, \quad |\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta,$$
$$(\mathbf{a}\times\mathbf{b})_i = \epsilon_{ijk}a_jb_k.$$

where θ is the angle between the vectors **a** and **b**.

i) $|\mathbf{a}\times\mathbf{b}|$ is the area of the parallelogram with sides **a** and **b**.

ii) $\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a}$.

iii) $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}, \quad \mathbf{a}\times(\beta\,\mathbf{b}) = \beta\,\mathbf{a}\times\mathbf{b}$.

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

- **Scalar triple product:**

$$[\mathbf{a},\mathbf{b},\mathbf{c}] = \mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \epsilon_{ijk}a_ib_jc_k.$$

i) $|\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})|$ is the volume of the parallelepiped with sides **a**, **b** and **c**.

ii) $[\alpha\,\mathbf{a}+\beta\,\mathbf{b},\mathbf{c},\mathbf{d}] = \alpha\,[\mathbf{a},\mathbf{c},\mathbf{d}] + \beta\,[\mathbf{b},\mathbf{c},\mathbf{d}]$.

iii) $[\mathbf{a},\mathbf{b},\mathbf{c}] = [\mathbf{b},\mathbf{c},\mathbf{a}] = [\mathbf{c},\mathbf{a},\mathbf{b}] = -[\mathbf{a},\mathbf{c},\mathbf{b}] = -[\mathbf{c},\mathbf{b},\mathbf{a}] = -[\mathbf{b},\mathbf{a},\mathbf{c}]$

- **Vector triple product:**

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

  i) The vector triple product is not associative.

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

- **Equation of a line.**

  Given a point A with position vector $\mathbf{a}$ located on a line having a direction $\hat{\mathbf{b}}$, a generic point R on the same line with position vector $\mathbf{r}$ is given by

$$\mathbf{r} = \mathbf{a} + \lambda \hat{\mathbf{b}}, \quad \mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

  where λ is a scalar. Note that the same equation can be also written as follows

$$(\mathbf{r} - \mathbf{a}) \times \hat{\mathbf{b}} = 0.$$

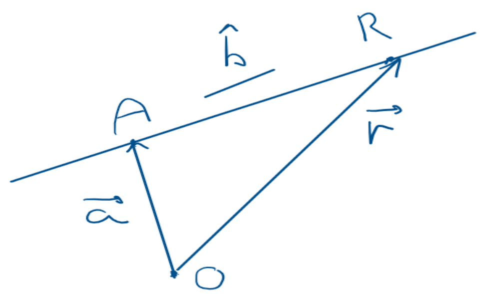

Figure 1.2: Line passing through the point A and having a direction $\hat{\mathbf{b}}$.

- **Equation of a plane.**

  i) A point R on a plane perpendicular to the unit vector $\hat{\mathbf{n}}$ and passing through the point A is

  $$(\mathbf{r}-\mathbf{a})\cdot\hat{\mathbf{n}}=0,$$

  where $\mathbf{a}$ and $\mathbf{r}$ are the position vectors of A and R, respectively.

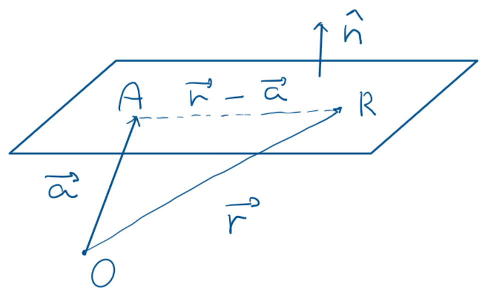

Figure 1.3: Plane perpendicular to the unit vector $\hat{\mathbf{n}}$ and passing through the point A.

  ii) A point R on a plane passing through the points A, B and C is

  $$\mathbf{r}=\mathbf{a}+\lambda(\mathbf{b}-\mathbf{a})+\mu(\mathbf{c}-\mathbf{a}),$$

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

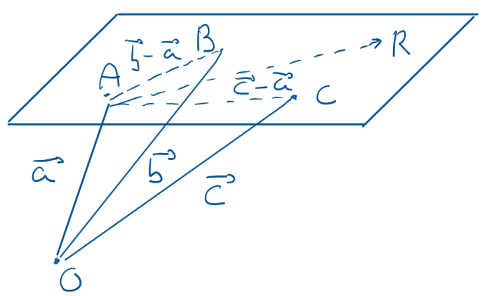

Figure 1.4: Plane passing through the points A, B and C.

Chapter 2

Vector Spaces

Linear vector space. A set of objects called *vectors* forms a vector space V if there are two operations defined on the elements of the set called *addition* and *multiplication by scalars*, which obey the following simple rules (the axioms of the vector space):

i) If $\mathbf{u}$ and $\mathbf{v}$ are in V then $\mathbf{u} + \mathbf{v}$ (addition) is in V. If $\mathbf{v}$ is in V then $\alpha\mathbf{v}$ (scalar multiplication) is in V.

   We say that the vector space V is *closed* with respect to *addition* and *scalar multiplication*.

ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), \quad (\alpha\,\beta)\mathbf{v} = \alpha(\beta\,\mathbf{v})$.

iii) There exists a *neutral element* $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v}$.

iv) There exists an *inverse element* $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v}$.

v) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

vi) $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\,\mathbf{u} + \alpha\,\mathbf{v}, \quad (\alpha + \beta)\mathbf{v} = \alpha\,\mathbf{v} + \beta\,\mathbf{v}$.

vii) $1\,\mathbf{v} = \mathbf{v}$ for all $\mathbf{v}$,

where $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are vectors and α, β are scalars. If the scalars α are real V is called a real vector space, otherwise V is called a complex vector space.

Example 1 Indicate which if the following set form a vector space

(i) $\mathbb{R}^3$. *Yes.*

(ii) $\mathbb{R}^n$. *Yes (Euclidean vector spaces.)*

(iii) $\mathbb{C}^n$. *Yes (Examples of complex vector spaces.)*

(iv) The set of real functions $f(x)$ with no restriction on the values of x and with the usual addition and scalar multiplication. *Yes.*

(v) The set of matrices of size $(n \times m)$ with real entries and with the usual addition and scalar multiplication. *Yes.*

(vi) The set of 2-dimensional vectors with real entries and the usual addition but the following definition of scalar multiplication

$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha x \\ 0 \end{pmatrix},$$

where α is a scalar. *No.*

(vii) The set of solutions of the following second order, linear, homogeneous differential equation

$$p(x)\, \frac{d^2 f}{dx^2} + q(x)\, \frac{df}{dx} + r(x)\, f = 0,$$

where p, q, r are given functions. *Yes.*

(viii) The set of vector $\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in the 3-dimensional space for which

$$2x - 3y + 11z + 2 = 0.$$

No.

- **Linear combinations:**

$$\alpha_1\, \mathbf{v}_1 + \alpha_2\, \mathbf{v}_2 + \cdots \alpha_k\, \mathbf{v}_k = \alpha_i\, \mathbf{v}_i,$$

where $\mathbf{v}_1$, $\mathbf{v}_2, \cdots \mathbf{v}_k$, are k vectors in V.

The set of all linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2, \cdots \mathbf{v}_k$, is called a span of $\mathbf{v}_1$, $\mathbf{v}_2, \cdots \mathbf{v}_k$, and denoted $\mathrm{Span}(\mathbf{v}_1, \, \mathbf{v}_2, \cdots \mathbf{v}_k)$.

Example 2 Answer the following questions

(i) What is a span of a single vector in $\mathbb{R}^3$? *It is the set of all scalar multiples of this vector. It is a line in the direction of the vector.*

(ii) What is a span of two non coplanar vectors in $\mathbb{R}^3$? *It is a plane through zero containing these two vectors.*

- **Linearly independent vectors:** k vectors $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ are said to be *linearly independent* if the equation

$$\alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \cdots + \alpha_k\mathbf{v_k} = 0$$

is satisfied if and only if all $\alpha_i = 0$. Otherwise, the vectors are said to be *linearly dependent*. That is, they are *linearly dependent* if the expression

$$\alpha_1\mathbf{v_1} + \alpha_2\mathbf{v_2} + \cdots + \alpha_k\mathbf{v_k} = 0$$

with at least one $\alpha_i \neq 0$. In other words, the vectors $\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}$ are *linearly dependent* if and only if one vector $\mathbf{v}_i$ can be written as a linear combination of the others.

Example 3 Indicate whether the following sets of *vectors* are linearly dependent or independent.

(i)

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

By definition: $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \begin{pmatrix} \alpha_3 \\ \alpha_+\alpha_2 + \alpha_3 \\ \alpha_1 + 2\alpha_2 - \alpha_3 \end{pmatrix} = 0,$

This implies, $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Hence the vectors are linearly independent.

(ii)

$$\mathbf{v}_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}.$$

We can see that $\mathbf{v}_3 = \mathbf{v}_1 + 2\mathbf{v}_2$. Hence the vectors are linearly dependent.

(iiii)

$$\{1 + x + x^2, \quad 1 - x + 3x^2, \quad 1 + 3x - x^2\}.$$

By definition: $\alpha_1(1 + x + x^2) + \alpha_2(1 - x + 3x^2) + \alpha_3(1 + 3x - x^2) = x^2(\alpha_1 + 3\alpha_2 - \alpha_3) + x(\alpha_1 - \alpha_2 + 3\alpha_3) + (\alpha_1 + \alpha_2 + \alpha_3) = 0$.

It follows that $\alpha_1 = -2\alpha_2$, $\alpha_2 = \alpha_3$. Hence the 'vectors' are linearly dependents.

- **Basis:** A basis is a minimal set of vectors that span a vector space. In other words, a set of vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_k$, in V is called a basis of V if and only if $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_k$, are linearly independent and $V = \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_k)$. Then

  i) The numbers of vector in a basis is called the dimension of the space V (dim V).

  ii) If the set $\{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_k}\}$ is a basis of the vector space V, then any vector $\mathbf{v}$ in V can be written as a unique linear combination of the basis vectors and the coefficients of the unique linear combination are called the components of $\mathbf{v}$ with respect to that basis.

Example 4 Write down a basis for the following vector spaces

(i) $\mathbb{R}^3$.

Basis: the set of vectors in Example 3 (i). Dim 3.

(ii) The set of (2×3) matrices with real entries.

Basis:

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \cdots, \quad E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dim 6.

(iii) The set of polynomials of degree two or less with real coefficients.
Basis: $\{1, \quad x, \quad x^2\}$. Dim 3.

- **Inner (or scalar) product:** Consider a vector space V. The inner product between two elements of V is a scalar function denoted $\langle \mathbf{u}|\mathbf{v}\rangle$ that satisfies the following properties

  i) $\langle \mathbf{u}|\mathbf{v}\rangle = \langle \mathbf{v}|\mathbf{u}\rangle^*$.

  ii) $\langle \mathbf{u}|(\lambda\mathbf{v} + \mu\mathbf{w})\rangle = \lambda\langle \mathbf{u}|\mathbf{v}\rangle + \mu\langle \mathbf{u}|\mathbf{w}\rangle, \qquad \lambda, \mu$ scalars.

  iii) $\langle \mathbf{u}|\mathbf{u}\rangle > 0$ if $\mathbf{u} \neq 0$.

  The length of a vector (*norm*) is $|\mathbf{u}| = \sqrt{\langle \mathbf{u}|\mathbf{u}\rangle}$.

  two vectors are *orthogonal* if $\langle \mathbf{u}|\mathbf{w}\rangle = 0$.

Example 5 Check the properties for the following scalar products

(i) In $\mathbb{R}^3$:
$$\langle \mathbf{u}|\mathbf{v}\rangle = \mathbf{u}^T \cdot \mathbf{v}.$$
Take $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$
For the first property: $\langle \mathbf{u}|\mathbf{v}\rangle = \mathbf{u}^T \cdot \mathbf{v} = (u_1v_1 + u_2v_2 + u_3v_3) = \mathbf{v}^T \cdot \mathbf{u} = (\mathbf{v}^T \cdot \mathbf{u})^* = \langle \mathbf{v}|\mathbf{u}\rangle^*$. Similar procedure for the other properties.

(ii) In $\mathbb{C}^3$:
$$\langle \mathbf{u}|\mathbf{v}\rangle = \mathbf{u}^\dagger \cdot \mathbf{v}.$$
Take $\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$
For the first property: $\langle \mathbf{u}|\mathbf{v}\rangle = \mathbf{u}^\dagger\cdot\mathbf{v} = (u_1^*v_1+u_2^*v_2+u_3^*v_3) = (u_1v_1^*+u_2v_2^*+u_3v_3^*)^* = (\mathbf{v}^\dagger \cdot \mathbf{u})^* = \langle \mathbf{v}|\mathbf{u}\rangle^*$. Similar procedure for the other properties.

Chapter 3

Matrices

From now on in the course, we will be in $\mathbb{R}^3/\mathbb{C}^3$

- **Linear operators:** An object $\mathcal{A}$ is called a linear operator if its action on vectors $\mathbf{u}$ and $\mathbf{v}$ is as follows
$$\mathcal{A}\left(\alpha\mathbf{u}+\beta\mathbf{v}\right)=\alpha\mathcal{A}\,\mathbf{u}+\beta\mathcal{A}\,\mathbf{v},$$
where α and β are scalars. Matrices are examples of operators.

- **Matrix operations:**

  i) Matrix addition: $(A+B)_{ij}=A_{ij}+B_{ij}$.
  ii) Multiplication by a scalar: $(\alpha A)_{ij}=\alpha A_{ij}$.
  iii) Multiplication of matrices: $(AB)_{ij}=A_{ik}B_{kj}$, with $AB\neq BA$.
  iv) Transposition: $(A^T)_{ij}=A_{ji}$, with $(ABC\dots F)^T=F^T\dots C^TB^TA^T$.
  v) Complex conjugation: $(A^*)_{ij}=(A_{ij})^*$.
  vi) Hermitian conjugation (adjoint): $(A^\dagger)_{ij}=(A_{ji})^*$, with
  $(ABC\dots F)^\dagger=F^\dagger\dots C^\dagger B^\dagger A^\dagger$.

- **The determinant of a square matrix:**
$$|A|=A_{jk}C_{jk},\quad \text{for any row } j,\qquad |A|=A_{kj}C_{kj},\quad \text{for any column } j,$$
where $C_{mn}=(-1)^{m+n}|A_{mn}|$ is the *cofactor* associated to the matrix element A_{mn}. In turn, $|A_{mn}|$ is the *minor* associated to the matrix element A_{mn}. The minor is the determinant of the matrix obtained by removing the m-th row and n-th column from the matrix A.

  Properties:

  i) $|AB\dots F|=|A||B|\dots|F|$.

ii) $|A^T| = |A|, \quad |A^*| = |A|^*, \quad |A^\dagger| = |A|^*, \quad |A^{-1}| = |A|^{-1}$.

iii) If the rows (or the columns) are linearly dependent, then $|A| = 0$.

iv) If B is obtained from A by multiplying the elements of any row (or column) by a factor α, then $|B| = \alpha\,|A|$.

v) If B is obtained from A by interchanging two rows (or columns), then $|B| = -|A|$.

vi) If B is obtained from A by adding k times one row (or column) to the other row (or column), then $|A| = |B|$.

- **Elementary row operations (on matrices):**

  i) Multiply any row by a non zero constant.

  ii) Interchange any two rows.

  iii) Add some multiple of one row to any other row.

- **The inverse of a square matrix:**

$$A^{-1} = \frac{C^T}{|A|}, \qquad \text{that is} \quad A^{-1}_{ij} = \frac{C_{ji}}{|A|}, \qquad A^{-1}A = AA^{-1} = I,$$

where C is the cofactor matrix and I the identity matrix ($I_{ij} = \delta_{ij}$). If $|A| = 0$ the inverse does not exist and the matrix A is said to be *singular*.

Note that in order to find the inverse of a matrix, you can also use the *Gauss-Jordan method*, which makes use of the elementary row operations.

Properties:

i) $(AB\ldots F)^{-1} = F^{-1}\ldots B^{-1}A^{-1}$.

ii) $(A^T)^{-1} = (A^{-1})^T, \quad (A^\dagger)^{-1} = (A^{-1})^\dagger$.

- **The trace of a square matrix:** It is the sum of the diagonal elements of the matrix, i.e.

$$\mathrm{Tr}\,A = \sum_k A_{kk} \equiv A_{kk}.$$

Properties:

i) The trace is a linear operation.

ii) $\mathrm{Tr}\,A^T = \mathrm{Tr}\,A$; $\mathrm{Tr}\,A^\dagger = (\mathrm{Tr}\,A)^*$.

iii) $\mathrm{Tr}(AB) = (AB)_{ii} = A_{ij}B_{ji} = B_{ji}A_{ij} = (BA)_{jj} = \mathrm{Tr}(BA)$.

iv) $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB)$, i.e. the trace is invariant under cyclic permutations of the matrices.

Example 1 A and B are two anticommuting matrices, i.e. $AB = -BA$ and $A^2 = B^2 = I$. Show that $\operatorname{Tr} A = \operatorname{Tr} B = 0$.

Multiply the relation given by A on the left hand side: $AAB = -ABA$.
Apply the trace and its properties: $\operatorname{Tr} B = -\operatorname{Tr}(ABA) = -\operatorname{Tr}(BAA) = -\operatorname{Tr}(B)$.
This implies that $\operatorname{Tr} B = 0$. Similar procedure can be applied in order to show that $\operatorname{Tr} A = 0$.

- **Special types of square matrices:**

  i) *Symmetric matrices:* $A^T = A$.

  Anti-symmetric or skew-symmetric matrices: $A^T = -A$.

  Any matrix can be written as the sum of a symmetric matrix and an antisymmetric matrix:

  $$A = \frac{1}{2}\left(A + A^T\right) + \frac{1}{2}\left(A - A^T\right)).$$

  ii) *Hermitian matrices:* $A^\dagger = A$.

  Anti-hermitian matrices: $A^\dagger = -A$.

  Any matrix can be written as the sum of an hermitian matrix and an anti-hermitian matrix:

  $$A = \frac{1}{2}\left(A + A^\dagger\right) + \frac{1}{2}\left(A - A^\dagger\right).$$

  iii) *Unitary matrices:* $A^\dagger = A^{-1}$

  It follows that $(A^\dagger A)_{ij} = \delta_{ij}$.

  Note that they unitary matrices preserve the length of vectors. In fact

  $$\mathbf{v} = A\mathbf{u} \to |\mathbf{v}|^2 = \mathbf{v}^\dagger \cdot \mathbf{v} = (\mathbf{u}^\dagger A^\dagger) \cdot (A\mathbf{v}) = |\mathbf{v}|^2.$$

  iv) *Orthogonal matrices:* $A^T = A^{-1}$

  They also preserve the length of vectors since they correspond to real unitary matrices ($A^T \equiv A^\dagger$.)

Example 2 Consider the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Show that the Pauli matrices, together with the identity matrix, I, form a basis for the vector space of the (2×2) hermitian matrices.

A general element of the vector space is:

$$\alpha\sigma_1 + \beta\sigma_2 + \gamma\sigma_3 + \delta I = \begin{pmatrix} \gamma + \delta & \alpha - i\beta \\ \alpha + i\beta & -\gamma + \delta \end{pmatrix},$$

where α, β, γ, δ are real. This is the general form of an hermitian matrix since it can be rewritten as follows

$$\begin{pmatrix} a & c \\ c^* & b \end{pmatrix},$$

where a b are real.

Chapter 4

The eigenvalue problem

Consider a $(n \times n)$ matrix. We want to answer the following question:

Are there any vectors $\mathbf{x} \neq 0$ which are transformed by a matrix A into multiple of themselves?

In other words: For which vectors $\mathbf{x}$ and scalar λ is the following *eigenvalue equation*

$$A\mathbf{x} = \lambda\mathbf{x}$$

satisfied?

i) The vector $\mathbf{x}$ is called *eigenvector* and λ is called the corresponding *eigenvalue*.

ii) The determinant $|A - \lambda I|$ is called the *characteristic polynomial* of degree n.

iii) The eigenvalue equation, being a set of non homogeneous linear equations, has a non trivial solution if and only if $|A - \lambda I| = 0$.

iv) The eigenvalues of the matrix A are the roots of the characteristic polynomial.

v) The eigenvectors associated to the eigenvalue μ are the vectors $\mathbf{x}$ such that

$$(A - \mu\, I)\mathbf{x} = 0.$$

Example 1 Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 1 \\ 2 & 1-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{vmatrix} = (1-\lambda)^2(2-\lambda) - 2(1-\lambda) - 4(1-\lambda) = 0.$$

Hence $\lambda_1 = 1$, $\lambda_2 = 4$, $\lambda_3 = -1$.
For $\lambda_1 = 1$:

$$(A - I)\mathbf{x}_1 = \begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_2 + x_3 \\ 2x_1 + x_3 \\ x_1 + x_2 + x_3 \end{pmatrix} = 0.$$

Hence $\mathbf{x}_1 = \begin{pmatrix} x_1 \\ x_1 \\ -2x_1 \end{pmatrix}$. A possible eigenvector is: $\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$.

For $\lambda_1 = 4$ solve $(A - 4I)\mathbf{x}_2 = 0 \quad \rightarrow \quad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

For $\lambda_1 = -1$ solve $(A + I)\mathbf{x}_3 = 0 \quad \rightarrow \quad \mathbf{x}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$.

- Eigenvectors associated to different eigenvalues are linearly independent.
- If a $n \times n$ matrix A has n distinct eigenvalues, then the set of corresponding eigenvectors represents a basis in the vector space on which the matrix acts. If the eigenvalues are not all distinct, it may or it may not exist a basis of eigenvectors.
- If a matrix has an eigenvalue equal to zero, then the matrix is singular since its determinant is zero.

Example 2 Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}.$$

The eigenvalues are: $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$. Hence one of the eigenvalues is degenerate.

For $\lambda_1 = 5$ solve $(A - 5I)\mathbf{x}_1 = 0 \quad \rightarrow \quad \mathbf{x}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.

For $\lambda_2 = \lambda_3 = -3$:

$$(A + 3I)\mathbf{x} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 - 3x_3 \\ 2x_1 + 4x_2 - 6x_3 \\ -x_1 - 2x_2 + 3x_3 \end{pmatrix} = 0.$$

Hence $\mathbf{x} = \begin{pmatrix} -2x_2 + 3x_3 \\ x_2 \\ x_3 \end{pmatrix}$.

Two linearly independent eigenvectors are: $\mathbf{x}_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$.

- **On the special types of square matrices:**

  i) *Hermitian matrix* and *symmetric matrix:*
  Their eigenvalues are real.
  Proof: Consider the expression $A\mathbf{x} = \lambda\mathbf{x}$. Take the adjoint of this expression, i.e. $\mathbf{x}^\dagger A^\dagger = \lambda^* \mathbf{x}^\dagger$. Multiply both expressions on the left by $\mathbf{x}^\dagger$ and compare them. You get $\lambda \mathbf{x}^\dagger \mathbf{x} = \lambda^* \mathbf{x}^\dagger \mathbf{x}$, which implies $\lambda = \lambda^*$.

  ii) *Anti-Hermitian matrix* and *antisymmetric matrix:*
  Their eigenvalues are purely imaginary ore zero.

  iii) *Unitary matrix* and *orthogonal matrix:*
  Their eigenvalues have absolute value equal to one.
  Proof: Consider the expression $A\mathbf{x} = \lambda\mathbf{x}$. Take the adjoint of this expression, i.e. $\mathbf{x}^\dagger A^\dagger = \lambda^* \mathbf{x}^\dagger$. Take the scalar product between these two expressions. You get $\mathbf{x}^\dagger (A^\dagger A)\mathbf{x} = \lambda^* \lambda (\mathbf{x}^\dagger \mathbf{x})$, which implies $|\lambda|^2 = 1$.

  iv) Theorem: The eigenvectors of all special matrices are linearly independent. In addition, they can always be chosen in such a way that they form a mutually orthogonal set.

- **Similar matrices:** Two $(n \times n)$ matrices A and B are said to be similar if it exist a matrix S such that

$$B = S^{-1}AS.$$

The two matrices represent the same linear operator in different basis. the two basis are related by the matrix S. The two matrices A and B have the same determinant, the same trace and the same set of eigenvalues.

Theorem: **Diagonalisation of a matrix:** If the new basis is a set of eigenvectors of A then $B \equiv D$ is diagonal with

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad S = \begin{pmatrix} | & | & \dots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix}$$

where λ_i are the eigenvalues and $\mathbf{x}_i$ the eigenvectors.

Proof:

$$AS = A \begin{pmatrix} | & | & \dots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = \begin{pmatrix} | & | & \dots & | \\ A\mathbf{x}_1 & A\mathbf{x}_2 & \dots & A\mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix}$$
$$= \begin{pmatrix} | & | & \dots & | \\ \lambda_1\mathbf{x}_1 & \lambda_2\mathbf{x}_2 & \dots & \lambda_n\mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = \begin{pmatrix} | & | & \dots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} D = SD.$$

Note that $|A| = \prod_i \lambda_i$ and $\mathrm{Tr}A = \sum_i \lambda_i$.

If A is a special matrix, then A is diagonalisable since it is always possible to find a basis of eigenvectors. Moreover, since the basis of eigenvectors can be chosen to be orthonormal, the matrix S is unitary, i.e.

$$D = S^\dagger AS.$$

In fact

$$S^\dagger S = \begin{pmatrix} - & \mathbf{x}_1^* & \rightarrow \\ - & \mathbf{x}_2^* & \rightarrow \\ \dots & \dots & \dots \\ - & \mathbf{x}_n^* & \rightarrow \end{pmatrix} \begin{pmatrix} | & | & \dots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^\dagger\mathbf{x}_1 & \mathbf{x}_1^\dagger\mathbf{x}_2 & \dots & \mathbf{x}_1^\dagger\mathbf{x}_n \\ \mathbf{x}_2^\dagger\mathbf{x}_1 & \dots & \dots & \mathbf{x}_2^\dagger\mathbf{x}_n \\ \dots & \dots & \dots & \dots \\ \mathbf{x}_n^\dagger\mathbf{x}_1 & \dots & \dots & \mathbf{x}_n^\dagger\mathbf{x}_n \end{pmatrix} = I.$$

Example 3 Diagonalise the matrix

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

The eigenvalues are: $\lambda_1 = 4$, $\lambda_2 = \lambda_3 = -2$.

The general forms of the eigenvectors are: $\mathbf{x}_1 = \begin{pmatrix} a \\ 0 \\ a \end{pmatrix}$, $\quad \mathbf{x}_{2/3} = \begin{pmatrix} b \\ c \\ -b \end{pmatrix}$.

Hence, a set of orthonormal eigenvectors is: $\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\quad \mathbf{x}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Hence $D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$

$$= \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -1 & 1 \end{pmatrix} = S^\dagger A S.$$

Applications: consider a square matrix A that can be diagonalised, then

i) n-power of A :

$$A^n = AA \dots A = (SDS^{-1})(SDS^{-1}) \dots (SDS^{-1}) = SD^n S^{-1}.$$

ii) Exponential of A :

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!},$$

then

$$e^A = e^{(SDS^{-1})} = \sum_{n=0}^{\infty} \frac{(SDS^{-1})^n}{n!} = S e^D S^{-1}.$$

Example 4 Consider a unitary matrix U.

i) Show that U has the form $U = e^{iH}$ for some hermitian matrix H.

Since U is matrix, it can be diagonalised: $U = SDS^\dagger$ with

$$D = \begin{pmatrix} e^{i\theta_1} & 0 & 0 \\ 0 & e^{i\theta_2} & 0 \\ \dots & \dots & \dots \\ 0 & 0 & e^{i\theta_n} \end{pmatrix} \quad \text{and} \quad e^{i\theta_i} = \lambda_i.$$

Then $U = Se^{i\Lambda}S^\dagger$ with $\Lambda = \begin{pmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ \dots & \dots & \dots \\ 0 & 0 & \theta_n \end{pmatrix}$.

It follows that $U = e^{iS\Lambda S^\dagger}$ where $S\Lambda S^\dagger \equiv H$ is an hermitian matrix.

ii) Show that $|U| = |e^{iH}| = e^{i\operatorname{Tr} H}$.

Since H is matrix, it can be diagonalised. Then

$$|U| = |e^{iH}| = |e^{iSDS^\dagger}| = |Se^{iD}S^\dagger| = |S||e^{iD}||S^\dagger| = |e^{iD}|$$

with

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ \dots & \dots & \dots \\ 0 & 0 & \lambda_n \end{pmatrix}, \text{ and } e^{iD} = \begin{pmatrix} e^{i\lambda_1} & 0 & 0 \\ 0 & e^{i\lambda_2} & 0 \\ \dots & \dots & \dots \\ 0 & 0 & e^{i\lambda_n} \end{pmatrix}.$$

Then

$$|e^{iD}| = \Pi_{j=1}^n e^{i\lambda_j} = e^{i\sum_j \lambda_j} = e^{i\operatorname{Tr} D} = e^{i\operatorname{Tr} H}.$$

Chapter 5

Fourier Series

The Fourier series of a periodic function $f(x)$ with period L is a representation of the function $f(x)$ as an infinite series of cosine and sine functions

$$f(x) = \frac{a_0}{2} + \sum_{r=1}^{\infty} \left[a_r \cos\left(\frac{2\pi r x}{L}\right) + b_r \sin\left(\frac{2\pi r x}{L}\right) \right].$$

The Fourier coefficients a_0, a_r and b_r are:

$$\begin{aligned}
a_0 &= \frac{2}{L} \int_{x_0}^{x_0+L} f(x)\, dx, \\
a_r &= \frac{2}{L} \int_{x_0}^{x_0+L} f(x) \cos\left(\frac{2\pi r x}{L}\right) dx \qquad r = 1, 2, 3, \ldots, \\
b_r &= \frac{2}{L} \int_{x_0}^{x_0+L} f(x) \sin\left(\frac{2\pi r x}{L}\right) dx \qquad r = 1, 2, 3, \ldots,
\end{aligned}$$

where x_0 is an arbitrary point along the x-axis. In order to guarantee that the series converges, the function $f(x)$ must satisfy the Dirichlet conditions in the interval L:

i) The function is single-valued.

ii) The function $f(x)$ has a finite number of extreme points (maxima and minima).

iii) The function $f(x)$ has a finite number of finite discontinuities.

Counterexample: the function $\sin(1/x)$ can not be represented by means of a Fourier series.

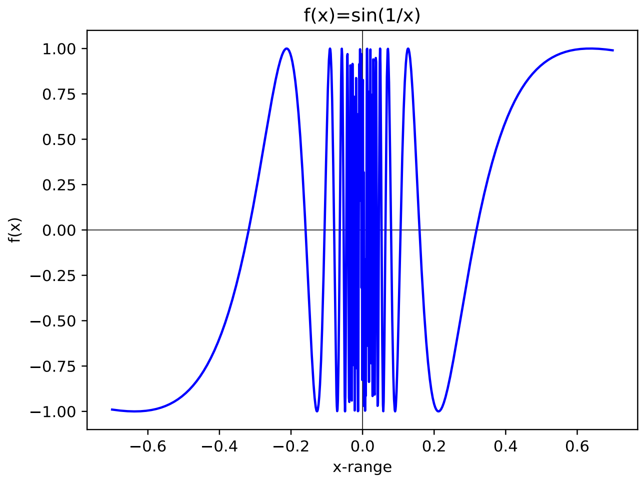

Figure 5.1: Counterexample.

- The set of all periodic functions on the interval L that can be represented by Fourier series forms a vector space:

  i) Basis:

$$\cos\left(\frac{2\pi r x}{L}\right), \quad r = 0, 1, 2, 3, \ldots; \qquad \sin\left(\frac{2\pi r x}{L}\right), \quad r = 1, 2, 3, \ldots$$

  ii) A general element of the space is:

$$f(x) = \frac{a_0}{2} \cdot 1 + \sum_{r=1}^{\infty} \left[a_r \cos\left(\frac{2\pi r x}{L}\right) + b_r \sin\left(\frac{2\pi r x}{L}\right) \right].$$

  iii) Operations: standard addition and multiplication by a scalar.

  iv) Inner product:

$$\langle f | g \rangle = \frac{2}{L} \int_0^L f(x) g(x) dx.$$

  v) The basis is orthogonal (check with the inner product above.)

- **Properties of the Fourier series:**

  i) If the function $f(x)$ is even, that is $f(-x) = f(x)$, $b_r = 0$ for $r = 1, 2, 3, \ldots$

  ii) If the function $f(x)$ is odd, that is $f(-x) = -f(x)$, $a_r = 0$ for $r = 0, 1, 2, 3, \ldots$

iii) If x_1 is a point of discontinuity for the function $f(x)$ in the interval L, then the value of the Fourier series at that point is:

$$f(x_1) = \frac{f(x_1^-) + f(x_1^+)}{2},$$

where $f(x_1^-)$ and $f(x_1^+)$ are the left and right limits of the function at x_1, respectively.

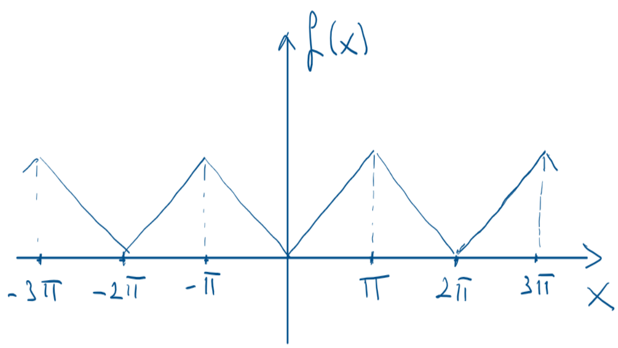

Figure 5.2: Function for *Example 1*.

Example 1 Calculate the Fourier series for the function sketched in the figure above.

The function is even, hence the coefficients $b_r = 0$. The interval $L = 2\pi$. Consider the function between $-\pi$ and π then

$$f(x) = \begin{cases} -x & \text{if} \quad -\pi \leq x \leq 0 \\ x & \text{if} \quad 0 \leq x \leq \pi. \end{cases},$$

and the Fourier coefficients are:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{0} (-x)\, dx + \int_{0}^{\pi} x\, dx = \frac{2}{\pi} \int_{0}^{\pi} x\, dx = \pi,$$

$$a_r = \frac{1}{\pi} \int_{-\pi}^{0} (-x) \cos rx\, dx + \int_{0}^{\pi} x\, \cos rx\, dx = \frac{2}{\pi} \int_{0}^{\pi} x\, \cos rx\, dx = \frac{2}{\pi}\, \frac{(-1)^r - 1}{r^2}.$$

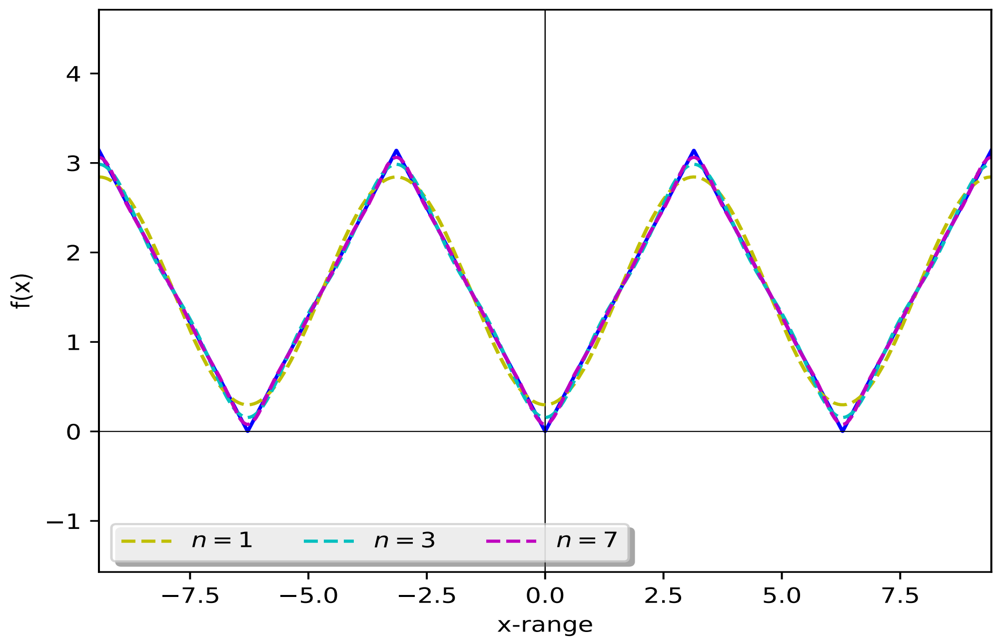

Figure 5.3: Fourier series for the function in *Example 1*

Example 2 Calculate the Fourier series for the function

$$f(x) = \begin{cases} -1 & \text{if} \quad -\pi < x < 0 \\ 0 & \text{if} \quad 0 < x < \pi. \end{cases}$$

The function is odd, hence the coefficients $a_r = 0$. The interval $L = 2\pi$. The Fourier coefficients are:

$$b_r = \frac{1}{\pi} \int_{-\pi}^{0} (-1)\, \sin rx\, dx + \int_{0}^{\pi} \sin rx\, dx = \frac{2}{\pi} \int_{0}^{\pi} \sin rx\, dx = \frac{2}{\pi}\, \frac{1-(-1)^r}{r\,\pi}.$$

The Fourier series is:

$$f(x) = \frac{4}{\pi} \sum_{r=1}^{\infty} \frac{\sin(2r-1)}{(2r-1)}\, x.$$

The function is discontinuous at $x = 0, \pm\pi, \pm 2\pi, \ldots$ and its value, at these points, is zero.

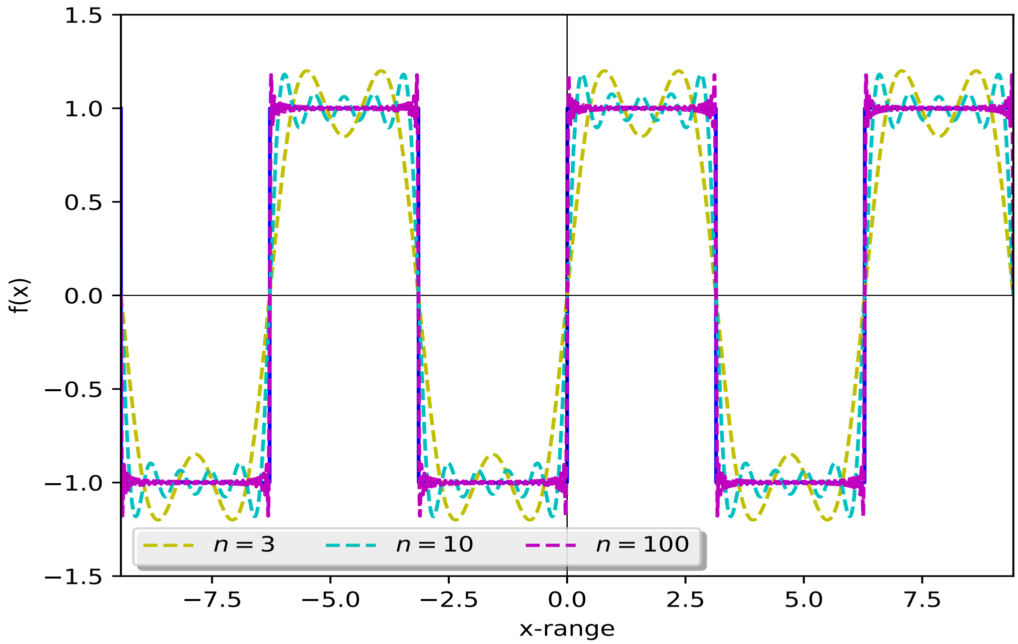

Figure 5.4: Fourier series for the function in *Example 2*

- Consider a function defined only on a finite interval L. Then, in order to calculate its Fourier series we need to extend the function over the whole x-axis. In other words we need to consider a periodic extension of the original function. The Fourier series of any extension is a representation of the original function on the finite interval L. However, normally continuous extension are preferable because they allow us to avoid the Gibbs's phenomenon at the points of discontinuity (see page 421 in Riley).

Example 3 Sketch possible extensions for the function

$$f(x) = \begin{cases} x^2 & \text{if} \quad 0 < x < 2 \\ 0 & \quad\text{otherwise.} \end{cases}$$

All extensions below provide good representations of the function $f(x)$ in the interval $0 \leq x \leq 2$.

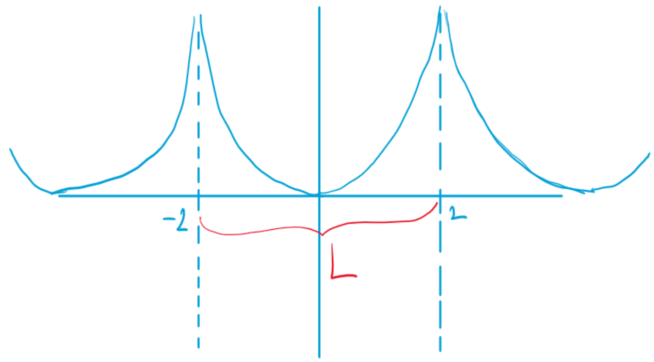

Figure 5.5: (i)*Example 3:* even and continuous extension.

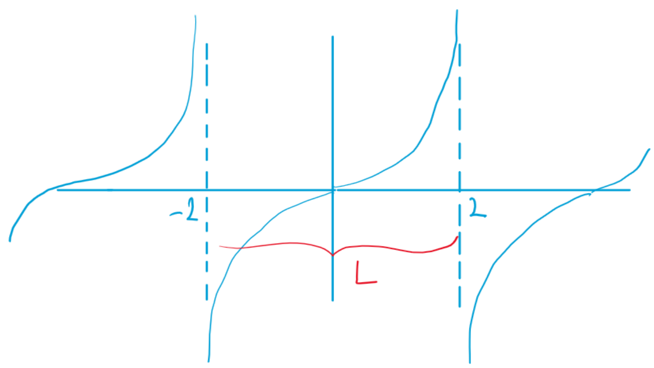

Figure 5.6: (ii) *Example 3:* odd and non-continuous extension.

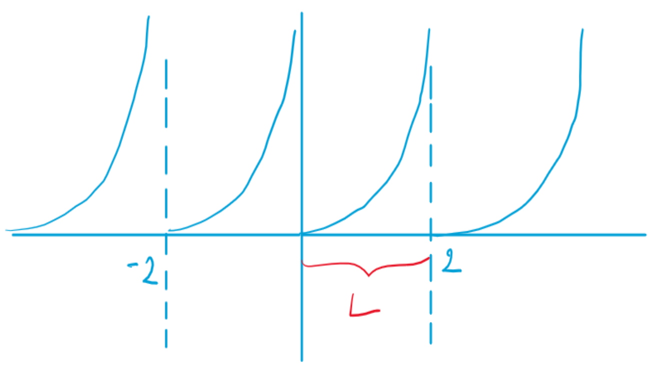

Figure 5.7: (iii) *Example 3:* non-continuous extension.

- Fourier series evaluated at specific points can be used to calculate series of constant terms. Consider the function in *Example 2* and its Fourier series

$$f(x) = \frac{4}{\pi}\left(\sin x + \sin\frac{3x}{3} + \sin\frac{5x}{5} + \cdots\right).$$

It follows that $f(\pi/2) = 1$. On the other hand, it is also true that

$$f(\pi/2) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

This expression is the *Leibniz formula for* π.

- Given a Fourier series, integration and differentiation can be used to obtain Fourier series for other functions. However, while integration is always a safe operation in the sense that convergence of the new series is always guaranteed, differentiation is not since an additional power of r at the numerator reduces the rate of convergence of the new series.

Example 4 Consider the even extension for the function

$$f(x) = \begin{cases} x^2 & \text{if} \quad 0 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

Write its Fourier series and by using the operations of integration and differentiation find the Fourier series of the function

$$g(x) = \begin{cases} x^3 & \text{if} \quad 0 \le x \le 2 \\ 0 & \text{otherwise.} \end{cases}$$

We choose the extension (i) in Example 3. Its Fourier series is:

$$f(x) = \frac{4}{3} + 16\sum_{r=1}^{\infty} \frac{(-1)^r}{(\pi r)^2} \cos\left(\frac{\pi r x}{2}\right) = x^2 \ (0 \le x \le 2).$$

Take the derivative of the expression above, $f(x),$:

$$-8\sum_{r=1}^{\infty} \frac{(-1)^r}{\pi r} \sin\left(\frac{\pi r x}{2}\right) = 2x \ (0 \le x \le 2) \quad \rightarrow \quad x = -4\sum_{r=1}^{\infty} \frac{(-1)^r}{\pi r} \sin\left(\frac{\pi r x}{2}\right).$$

Integrate $f(x)$:

$$\frac{4}{3}x + 32\sum_{r=1}^{\infty}\frac{(-1)^r}{(\pi r)^3}\sin\left(\frac{\pi r x}{2}\right) + c = \frac{x^3}{3} \ (0 \leq x \leq 2),$$

where c is the constant of integration. We can replace the result for x into this expression, then we get

$$x^3 = -16\sum_{r=1}^{\infty}\frac{(-1)^r}{\pi r}\sin\left(\frac{\pi r x}{2}\right) + 96\sum_{r=1}^{\infty}\frac{(-1)^r}{(\pi r)^3}\sin\left(\frac{\pi r x}{2}\right) + c' \ (0 \leq x \leq 2).$$

Since $g(0) = 0$, $c' = 0$ and the expression above becomes the Fourier series of the function $g(x)$.

Complex Fourier series: Using some manipulations, Fourier series can be written in a complex form. Writing trigonometric functions by means of exponentials we have

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{r=1}^{\infty}\left[a_r\cos\left(\frac{2\pi r x}{L}\right) + b_r\sin\left(\frac{2\pi r x}{L}\right)\right] \\ &= \frac{a_0}{2} + \sum_{r=1}^{\infty}\left[e^{i2\pi r x/L}\left(\frac{a_r}{2} + \frac{b_r}{2i}\right) + e^{-i2\pi r x/L}\left(\frac{a_r}{2} - \frac{b_r}{2i}\right)\right]. \end{aligned}$$

Set $(a_r - ib_r)/2 \equiv c_r$. Then, since $a_r = a_{-r}$ and $b_r = -b_{-r}$ we get $(a_r + ib_r)/2 \equiv c_{-r}$. It follows that

$$f(x) = \frac{a_0}{2} + \sum_{r=1}^{\infty}c_r\left(e^{i2\pi r x/L} + c_{-r}\,e^{-i2\pi r x/L}\right) = \sum_{r=-\infty}^{\infty}c_r\,e^{i2\pi r x/L}.$$

A similar manipulation allows us to find

$$c_r = \frac{1}{L}\int_{x_0}^{x_0+L}f(x)\,e^{-2\pi i r x/L}\,dx.$$

Chapter 6

Integral Transforms

An integral transform is a function g that can be expressed as an integral of another function f in the form

$$\mathcal{I}[f(x)](y) \equiv g(y) = \int_{-\infty}^{\infty} K(x,y) f(x) dx,$$

where $K(x,y)$ is called the kernel of the transform.

- $\mathcal{I}$ is a linear operator, that is:

$$\mathcal{I}[c_1 f_1 + c_2 f_2] = c_1 \mathcal{I}[f_1] + c_2 \mathcal{I}[f_2], \qquad c_1, c_2 \quad \text{constants.}$$

- Given $\mathcal{I}$ such that $\mathcal{I}[f] = g$, the inverse operator $\mathcal{I}^{-1}$ is also a linear operator and $\mathcal{I}^{-1}[g] = f$.

There are several types of integral functions. We are going to discuss the Fourier and the Laplace transforms.

6.1 Fourier Transforms

The Fourier Transform of the function $f(t)$ is:

$$\mathcal{F}[f(t)](\omega) \equiv \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t} dt.$$

The integral exists if:

- The function $f(x)$ has a finite number of finite discontinuities.

- $\int_{-\infty}^{\infty} |f(t)| dt$ is finite.

Its f is continuous the inverse is:

$$\mathcal{F}^{-1}[\hat{f}(\omega)](t) = f(t) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega t} d\omega.$$

There different ways to write Fourier transforms. We will stick to the notation above. However, be aware that you could encounter the following forms as well

$$\text{i)}\ \hat{f}(\omega) = \int\limits_{-\infty}^{\infty} f(t)\, e^{-i\omega t} dt, \qquad f(t) = \tfrac{1}{2\pi} \int\limits_{-\infty}^{\infty} \hat{f}(\omega)\, e^{i\omega t} d\omega,$$

$$\text{ii)}\ \hat{f}(\omega) = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} f(t)\, e^{i\omega t} dt, \qquad f(t) = \tfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \hat{f}(\omega)\, e^{-i\omega t} d\omega,$$

$$\text{iii)}\ \hat{f}(\nu) = \int\limits_{-\infty}^{\infty} f(t)\, e^{-i2\pi\nu t} dt, \qquad f(t) = \int\limits_{-\infty}^{\infty} \hat{f}(\nu)\, e^{ii2\pi\nu t} d\nu.$$

There are functions that are not periodic. Hence, we cannot use Fourier series. However, we could use Fourier transforms and think of these functions as periodic functions with a period that is infinite. Consider the complex Fourier series with period L

$$f(t) = \sum_{n=-\infty}^{\infty} c_n\, e^{i2\pi nt/L} = \sum_{n=-\infty}^{\infty} \left(\frac{1}{L} \int\limits_{-L/2}^{L/2} f(t)\, e^{-2\pi int/L}\, dt \right) e^{i2\pi nt/L}.$$

Set $2\pi n/L \equiv \omega_n$. This implies $\Delta\omega = \omega_{n+1} - \omega_n = 2\pi/L$. Then

$$f(t) = \sum_{n=-\infty}^{\infty} \left(\frac{\Delta\omega}{2\pi} \int\limits_{-L/2}^{L/2} f(t)\, e^{-it\omega_n}\, dt \right) e^{it\omega_n}.$$

When $L \longrightarrow \infty$ the expression above becomes

$$f(t) = \int\limits_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} \left(\int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} f(t)\, e^{-it\omega}\, dt \right) e^{it\omega} = \int\limits_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi}} \hat{f}(\omega)\, e^{it\omega}.$$

<u>*Example 1*</u> Calculate the complex Fourier series for the following periodic function

$$f(t) = \begin{cases} 0 & -L/2 < t < -a/2 \\ 1 & -a/2 < t < a/2 \\ 0 & a/2 < t < -L/2 \end{cases}, \qquad L > a,$$

and the Fourier transform for the following non periodic function

$$g(t) = \begin{cases} 1 & -a/2 < t < a/2 \\ 0 & \text{otherwise} \end{cases}$$

Then compare $|c_n|$ (from the Fourier series of $f(t)$) and $|\hat{g}(\omega)|$ by sketching them on the same x-axis.

The coefficients for the complex Fourier series of $f(t)$ are: $c_n = \frac{a}{L}\frac{\sin(n\pi a/L)}{n\pi a/L}$, for $n \neq 0$ and $c_0 = a/L$. Then, the Fourier series is:

$$f(t) = \sum_{n=-\infty}^{\infty} \frac{a}{L}\frac{\sin(n\pi a/L)}{n\pi a/L} e^{i2\pi nt/L}.$$

On the other hand, the Fourier transform of $g(t)$ is:

$$\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}}\int_{-a/2}^{a/2} e^{i\omega t}dt = \frac{1}{\sqrt{2\pi}}\frac{e^{i\omega t}}{-i\omega}\bigg|_{-a/2}^{a/2} = \frac{a}{\sqrt{2\pi}}\frac{\sin(a\omega/2)}{a\omega/2}.$$

$|c_n|$ represents a discrete spectrum while $|\hat{g}(\omega)|$ represents a continuum spectrum. $|\hat{g}(\omega)|$ is the envelope of $|c_n|$.

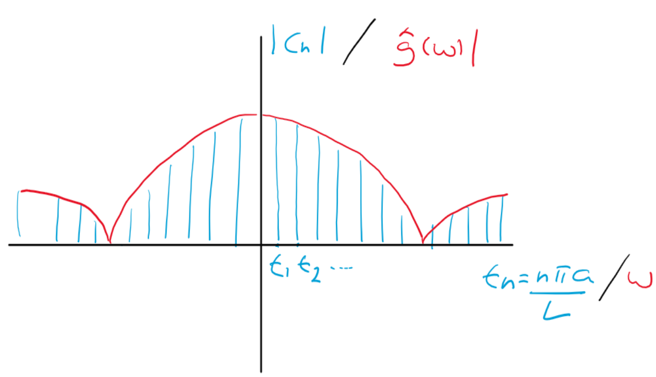

Figure 6.1: *Example 1:* Comparison between the discrete and continuum spectrum.

Properties of the Fourier transform

- **Scaling:**

$$\mathcal{F}[f(at)](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(at)\, e^{-i\omega t} dt,$$

where a is a constant. Set $at = t'$, then

$$\mathcal{F}[f(at)](\omega) = \frac{1}{|a|\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t')\, e^{-i\omega t'/a} dt = \frac{1}{a} \hat{f}\left(\frac{\omega}{a}\right),$$

where $a \neq 0$.

- **Translation:**

$$\mathcal{F}[f(t+a)](\omega) = e^{ia\omega} \hat{f}(\omega), \qquad a \quad \text{constant.}$$

- **Exponential multiplication:**

$$\mathcal{F}[e^{\alpha t} f(t)](\omega) = \hat{f}(\omega + i\alpha), \qquad \alpha \quad \text{constant.}$$

Example 2 Calculate the Fourier transform of $f(t/2)\cos\alpha t$ in terms of the Fourier transform of $f(t)$. The symbol α denotes a constant.

$$\mathcal{F}[f(t/2)\cos\alpha t](\omega) = \mathcal{F}[f(t/2)\, e^{i\alpha t}](\omega)/2 + \mathcal{F}[f(t/2)\, e^{-i\alpha t}](\omega)/2.$$

By using scaling and exponential multiplication in this order, we get

$$\begin{aligned} \mathcal{F}[f(t/2)\cos\alpha t](\omega) &= \mathcal{F}[f(t)\, e^{i\alpha t}](2\omega) + \mathcal{F}[f(t)\, e^{-i\alpha t}](2\omega) \\ &= \mathcal{F}[f(t)](2(\omega - \alpha)) + \mathcal{F}[f(t)](2(\omega + \alpha)). \end{aligned}$$

- **The Fourier transform of a derivative:**

$$\mathcal{F}[f'(t)](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\, e^{-i\omega t} dt$$

$$\left.\frac{f(t)}{\sqrt{2\pi}}\right|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (i\omega)\, f(t)\, e^{-i\omega t} dt = (i\omega)\hat{f}(\omega).$$

In fact, since $\int_{-\infty}^{\infty} |f(t)| dt$ is finite, $f(t) \longrightarrow 0$ when $t \to \pm\infty$.

For a derivative of order n we have

$$\mathcal{F}[f^{(n)}(t)](\omega) = (i\omega)^n \hat{f}(\omega).$$

We are now going to introduce a new operation between two functions, called *convolution*. This is used, for instance, in digital signal processing where two signals are combined to form a third signal. The Fourier transforms provide a way to analyse the spectrum of the signal involved.

The convolution of two functions f and g over the interval $(-\infty, \infty)$ is a function h defined as follows:

$$h(y) = \int_{-\infty}^{\infty} f(x)g(y-x)\,dx \equiv (f * g)(y) = (g * f)(y).$$

The convolution theorem for Fourier transforms.
The Fourier transform of the convolution $h(y)$ is:

$$\mathcal{F}[h(z)](\omega) = \sqrt{2\pi}\hat{f}(\omega)\hat{g}(\omega).$$

Proof:
Starting with the left hand side

$$\hat{h}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} h(y)\, e^{-iyk} dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy \left(\int_{-\infty}^{\infty} f(x)g(y-x)\,dx \right) e^{-iyk}$$

Swap the order of integration and set $y - x = z$, then

$$\hat{h}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, f(x) \int_{-\infty}^{\infty} dz\, g(z)\, e^{-i(z+x)k} = \sqrt{2\pi}\hat{f}(k)\hat{g}(k).$$

Starting with the right hand side

$$\sqrt{2\pi}\hat{f}(\omega)\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, f(x)\; e^{-ikx} \int_{-\infty}^{\infty} dz\, g(z)\, e^{-ikz}.$$

Taking $\exp(-ikx)$ into the second integral and setting $(z + x) = y$ we get

$$\sqrt{2\pi}\hat{f}(\omega)\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, f(x) \int_{-\infty}^{\infty} dy\, g(y-x)\, e^{-iky} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy\, e^{-iky}\, h(y) = \hat{h}(k).$$

Example 3 Use the Fourier transform in order to find a solution, i.e. $f(t)$, for the following ODE

$$\frac{d^2 f}{dt^2} + 2\,\frac{df}{dt} + f(t) = h(t),$$

where $h(t)$ is a known function. Start by taking the Fourier transform of the ODE. We get

$$\mathcal{F}\left[\frac{d^2 f}{dt^2}\right](\omega) + 2\,\mathcal{F}\left[\frac{df}{dt}\right](\omega) + \mathcal{F}\left[f(t)\right](\omega) = \mathcal{F}\left[h(t)\right](\omega),$$

which becomes: $-\omega^2 \hat{f}(\omega) + 2i\omega \hat{f}(\omega) + \hat{f}(\omega) = \hat{h}(\omega) \quad \rightarrow \quad \hat{f}(\omega) = \hat{h}(\omega)/(1 + 2i\omega - \omega^2)$. We have two possibilities

i) Take the inverse Fourier transform, i.e. $f(t) = \mathcal{F}^{-1}\left[\frac{\hat{h}(\omega)}{(1+2i\omega-\omega^2)}\right](t)$.

ii) Use the convolution theorem, i.e. $\hat{f}(\omega) = \hat{h}(\omega)\hat{g}(\omega)$ with $\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}}\,\frac{1}{(1+2i\omega-\omega^2)}$. Then $f(t) = \int_{-\infty}^{\infty} g(t')h(t-t')\,dt'$, where the functions $h(t)$ and $g(t)$ can be found by using the inverse Fourier transform.

6.2 Laplace transforms

The Laplace transform of the function $f(t)$ is:

$$\mathcal{L}[f(t)](s) \equiv \bar{f}(s) = \int_0^\infty f(t)\, e^{-st} dt,$$

where s is taken to be real. Note that, sometimes a constrain on the variable s should be imposed in order for the integral to exist.

Example 4 Calculate the Laplace transform of the following functions

i) t.

$\mathcal{L}[t](s) = \int_0^\infty t\, e^{-st} dt = \frac{e^{-st}}{-s^2}\Big|_0^\infty = \frac{1}{s^2}$ for $s > 0$.

What happens if you calculate the Fourier transform of the same function?

ii) $\cosh(kt)$, where k is a constant.

$\mathcal{L}[\cosh(kt)](s) = \int_0^\infty \cosh(kt)\, e^{-st} dt = \frac{1}{2}\left(\frac{e^{(k-s)}}{k-s}\Big|_0^\infty - \frac{e^{-(k+s)}}{k+s}\Big|_0^\infty\right) = \frac{s}{s^2-k^2}$ for $s > |k|$.

iii) $H(t-a)$.

$\mathcal{L}[H(t-a)](s) = \int_a^\infty e^{-st} dt = \frac{e^{-st}}{-s}\Big|_0^\infty = \frac{e^{-st}}{s}$ for $s > a$.

Properties of the Laplace transform.

- **Delay rule:**

$$\mathcal{L}[H(t-a)f(t-a)](s) = e^{-sa}\bar{f}(s), \qquad a \quad \text{constant.}$$

- **Exponential multiplication:**

$$\mathcal{L}[e^{at}f(t)](s) = \bar{f}(s-a), \qquad a \quad \text{constant.}$$

- **Scaling:**

$$\mathcal{L}[f(at)](s) = \frac{1}{|a|}\bar{f}\left(\frac{s}{a}\right), \qquad a \neq 0 \quad \text{constant.}$$

- **Polynomial multiplication:**

$$\mathcal{L}[t^n f(t)](s) = (-1)^n \frac{d^n \bar{f}(s)}{ds^n}, \quad n = 1, 2, 3 \ldots.$$

- **The Laplace transform of a derivative:**

For the derivative of order one we have

$$\mathcal{L}[f'(t)](s) = \int\limits_0^\infty f'(t)\, e^{-st}\, dt = f(t)\, e^{-st}\Big|_0^\infty + s \int\limits_0^\infty f(t)\, e^{-st}\, dt = -f(0) + s\bar{f}(s), \quad s > 0.$$

For a derivative of order n

$$\mathcal{L}[f^{(n)}(t)](s) = s^n \bar{f}(s) - s^{n-1} f(0) - s^{n-2} f^{(1)}(0) - \cdots - f^{(n-1)}(0), \quad s > 0,$$

where $f^{(n)}$ is the n^{th} derivative of the function f.

- **The Laplace transform for integration:**

$$\mathcal{L}\left[\int\limits_0^t f(u) du\right] = \frac{\bar{f}(s)}{s}.$$

Example 5 Using the properties of the Laplace transforms and the result $L[\cosh(kt)](s) = s/(s^2 - k^2)$ with $s > |k|$, calculate the Laplace transform of the following functions

i) $\sinh(kt)$.

Use the result of Example 4 (ii). Then

$\mathcal{L}[\sinh(kt)](s) = \mathcal{L}\left[\frac{d}{dt}\left(\frac{\cosh(kt)}{k}\right)\right](s) = -\frac{1}{k} + \frac{s}{k}\frac{s}{s^2-k^2} = \frac{k}{s^2-k^2}$ for $s > |k|$.

ii) $t\sinh(kt)$.

$\mathcal{L}[t\sinh(kt)](s) = (-1)\frac{d}{ds}\left(\frac{k}{s^2-k^2}\right) = \frac{2ks}{(s^2-k^2)^2}$ for $s > |k|$.

The convolution theorem for a Laplace transform:
If the functions f and g have Laplace transforms $\bar{f}$ and $\bar{g}$, then:

$$\mathcal{L}[(f * g)](s) = \mathcal{L}[(g * f)](s) = \mathcal{L}\left[\int\limits_0^t f(u) g(t-u)\, du\right](s) = \bar{f}(s)\bar{g}(s).$$

Note as the definition of convolution appears to be different. Let us calculate explicitly what the function $\bar{f}(s)\bar{g}(s)$ is.

$$\bar{f}(s)\bar{g}(s) = \int_0^\infty du\, f(u) \int_0^\infty dv\, e^{-s(u+v)}\, g(v).$$

Set $(u+v) = t$, then

$$\bar{f}(s)\bar{g}(s) = \int_0^\infty du\, f(u) \int_u^\infty dt\, e^{-st}\, g(t-u).$$

Swapping the order of integration and being careful with the new limits of integration, we have

$$\bar{f}(s)\bar{g}(s) = \int_0^\infty dt\, e^{-st} \left(\int_0^t du\, f(u)g(t-u) \right) = \mathcal{L}\left[\int_0^t f(u)g(t-u)\, du \right](s),$$

which implies

$$(f * g)(t) = \int_0^t f(u)g(t-u)\, du.$$

The inverse of a Laplace transform: $\mathcal{L}^{-1}[\bar{f}(s)] = f(t)$.
The general method for calculating inverse Laplace transforms requires notions of complex analysis. Nevertheless, in some cases it is possible to calculate an inverse Laplace transforms by means of

- partial fraction decomposition,
- convolution theorem,

together with the Laplace transform properties and tables of known Laplace transforms (see table on page 455 in Riley.) In this course we are going to limit ourselves to the use of these two techniques.

Example 6 Use partial fraction decomposition and the table at page 455 in order to calculate $f(t)$ given that

$$\bar{f}(s) = \frac{s+3}{s(s+1)}.$$

$$\bar{f}(s) = \frac{3}{s} - \frac{2}{s+1} = \bar{f}_1(s) + \bar{f}_2(s).$$

Using the tables we have $\mathcal{L}^{-1}\left[\bar{f}_1(s)\right](t) = 3$ for $s > 0$ and $\mathcal{L}^{-1}\left[\bar{f}_2(s)\right](t) = -2\,e^{-t}$ for $s > -1$ hence

$$\mathcal{L}^{-1}\left[\bar{f}(s)\right](t) = 3 - 2\,e^{t}$$

for $s > 0$.

Example 7 Use the convolution theorem and the table at page 455 in order to calculate $f(t)$ given that

$$\bar{f}(s) = \frac{2}{s^2(s-1)^2}.$$

$$\bar{f}(s) = \frac{2}{s^2}\frac{1}{(s-1)^2} = \bar{f}_1(s)\bar{f}_2(s).$$

Using the tables we have $\mathcal{L}^{-1}\left[\bar{f}_1(s)\right](t) = 2t$ for $s > 0$ and $\mathcal{L}^{-1}\left[\bar{f}_2(s)\right](t) = t\,e^{t}$ for $s > 1$ hence

$$\mathcal{L}^{-1}\left[\bar{f}(s)\right](t) = \int_0^\infty 2(t-u)\,u\,e^{u}\,du = 2\,e^{t}\,(t-2) + 2\,(t+2),$$

for $s > 1$.

Example 8 Use the Laplace transform in order to find a solution, i.e. $f(t)$, for the following ODE

$$\frac{df}{dt} + 2\,f(t) = e^{-t}, \quad f(0) = 3.$$

Start by taking the Laplace transform of the ODE.

$$\mathcal{L}\left[\frac{df}{dt}\right](s) + 2\,\mathcal{L}\left[f\right](s) = \mathcal{L}\left[e^t\right](s),$$

which becomes: $\quad -f(0) + s\,\bar{f}(s) + 2\,\bar{f}(s) = 1/(s+1) \quad \rightarrow$

$$\bar{f}(s) = \frac{3s+4}{(s+2)(s+1)} = \frac{1}{s+1}\frac{2}{s+2}.$$

Hence $f(t) = e^{-t} + 2\,e^{-2t}$.

Example 9 Consider the classical hamiltonian (energy) for a harmonic oscillator

$$H(p,x) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 = E, \quad \omega = \sqrt{\frac{k}{m}}$$

and the Schrödinger equation associated with it i.e

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\,\psi(x) + \frac{1}{2}m\omega^2 x^2\,\psi(x) = E\,\psi(x),$$

where $\psi(x)$ represents the wave function of the harmonic oscillator in coordinate space. The solution for the ground state is:

$$\psi_0(x) = e^{-(m\omega/3\hbar)x^2}, \quad E_0 = \frac{\hbar\omega}{2}.$$

This is a Gaussian with width $\Delta x = \sqrt{\hbar/m\omega}$. We want to find out the ground state wave function in momentum space. In order to do so, calculate the Fourier transform of $\psi_0(x)$. The variable in Fourier space is $k = p/\hbar$ (see workshops for Fourier transform of a Gaussian.) Then, calculate Δp. What is the meaning of the quantity $\Delta x \Delta p$?

The wave function in momentum space is:

$$\mathcal{F}[\psi(x)](k) = \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty} e^{-(m\omega/3\hbar)x^2}\, e^{-ikx}\, dx = \sqrt{\frac{\hbar}{m\omega}}\, e^{-k^2\hbar/(2m\omega)} = \sqrt{\frac{\hbar}{m\omega}}\, e^{-p^2/(2\hbar m\omega)}.$$

This is a Gaussian with width $\Delta p = \sqrt{\hbar m\omega}$. It follows that $\Delta x \Delta p = \hbar$, which codifies the uncertainty principle in QM.

Chapter 7

The Dirac delta function

Consider a pulse

$$\delta_n(x) = \begin{cases} n & -\frac{1}{2n} < x < \frac{1}{2n} \\ 0 & \text{otherwise} \end{cases}$$

If we take the duration of the pulse to decrease, while retaining a unit area, then, in the limit, we are led to the notion of the Dirac δ-function, i.e.

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x)\, dx = \int_{-\infty}^{\infty} \delta(x)\, dx = 1,$$

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x) f(x)\, dx = \int_{-\infty}^{\infty} \delta(x) f(x), dx = f(0),$$

where f is a well behaved function.

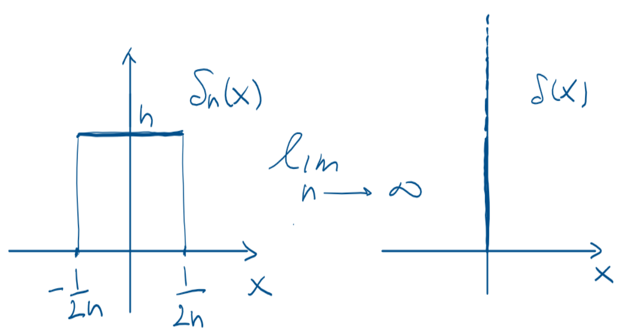

Figure 7.1: Dirac δ-function.

The Dirac delta function $\delta(x-a)$ - with a a constant - is a generalised function (or distribution) and it is defined as the limit of a sequence (not unique) of functions. Its defining properties are:

$$\delta(x-a)=0 \ \text{ for } \ x\neq a, \qquad \int_{\alpha}^{\beta} f(x)\delta(x-a)\,dx = \begin{cases} f(a) & \alpha<a<\beta \\ 0 & \text{otherwise} \end{cases}$$

Example 1 Calculate the following integrals

i) $\int_{-4}^{4} \delta(x-\pi)\,\cos x\,dx = \cos\pi = -1.$

ii) $\hat{\delta}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(t)\,e^{-i\omega t}dt = \frac{1}{\sqrt{2\pi}}.$

Integral representation of the Dirac delta function: Consider the following δ-function sequence

$$\delta_n(x) = \frac{\sin nx}{\pi x} = \frac{1}{2\pi}\int_{-n}^{n} e^{i\omega x}\,d\omega.$$

Then

$$\begin{aligned} f(x) &= \int_{\infty}^{\infty} \delta(t-x)\,f(t)\,dt = \lim_{n\to infty} \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)\left(\int_{-n}^{n} e^{i\omega(t-x)}\,d\omega\right)\,dt \\ &= \int_{-\infty}^{\infty} dt\,f(t)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(t-x)}\,d\omega\right), \end{aligned}$$

which implies

$$\delta(t-x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(t-x)}d\omega$$

Example 2 Calculate the inverse Fourier transform of $1/\sqrt{2\pi}$.

$$\mathcal{F}^{-1}\left[\frac{1}{\sqrt{2\pi}}\right](t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\,e^{i\omega t}d\omega = \delta(t).$$

Properties of the Dirac delta function.

- $\delta(x) = \delta(-x)$.
- $\delta(g(x)) = \sum\limits_{a} \delta(x-a)/|g'(a)|$.

  where a are the roots of the function $g(x)$ i.e. $g(a) = 0$ and $g'(a) \neq 0$

> *Example 3* Calculate $I = \int\limits_{-\infty}^{\infty} dt\, \delta(x^2 - b^2)\, f(x)\, dx$, where b is a constant.
>
> First, simplify the Dirac delta-function
>
> $$\delta(x^2 - b^2) = \frac{\delta(x-b)}{|2b|} + \frac{\delta(x+b)}{|-2b|},$$
>
> then
>
> $$I = \frac{1}{2b} \int\limits_{-\infty}^{\infty} dt\, \delta(x-b)\, f(x) + \frac{1}{2b} \int\limits_{-\infty}^{\infty} dt\, \delta(x+b)\, f(x) = \frac{1}{2b} \left(f(b) + f(-b) \right).$$

- $\int\limits_{-\infty}^{\infty} f(x)\delta'(x-a)dx = -f'(a)$.
- $H'(x) = \delta(x)$

  where $H(x)$ is the Heaviside step function defined as follows

  $$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}.$$

  In fact

  $$\int\limits_{-\infty}^{\infty} f(x)\, H'(x)\, dx = f(x)\, H(x)|_{-\infty}^{\infty} - \int\limits_{-\infty}^{\infty} f'(x)\, H(x)\, dx = f(\infty) - f(\infty) + f(0).$$

  Since $f(0) = \int_{-\infty}^{\infty} f(x)\, \delta(x)\, dx$ the property is proved.

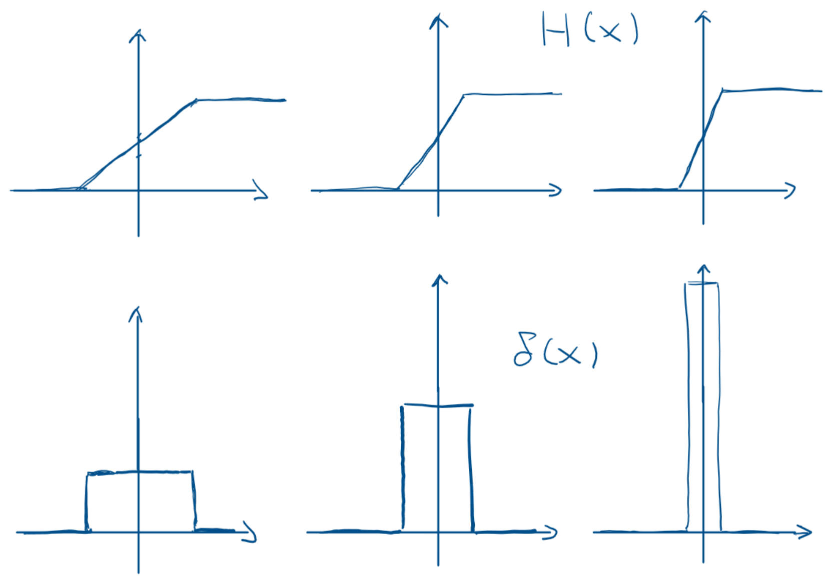

Figure 7.2: Dirac δ-function as derivative of the Heaviside step function.

Chapter 8

Vector Calculus

8.1 The del operator

Vector functions are vectors whose components are functions of one or more variables e.g.

$$\mathbf{a}(u, v, \dots) = a_x(u, v, \dots)\,\mathbf{i} + a_y(u, v, \dots)\,\mathbf{j} + a_z(u, v, \dots)\,\mathbf{k}.$$

A vector function defines a vector field.

- **Differentiation of vector functions:**

$$\frac{\partial \mathbf{a}}{\partial u} = \frac{\partial a_x}{\partial u}\,\mathbf{i} + \frac{\partial a_y}{\partial u}\,\mathbf{j} + \frac{\partial a_z}{\partial u}\,\mathbf{k}.$$

  Note that in cartesian coordinates $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, are constants.

  Differentiation rules:

  i)

$$\frac{\partial}{\partial u}(\phi \mathbf{a}) = \frac{\partial \phi}{\partial u}\mathbf{a} + \phi \frac{\partial \mathbf{a}}{\partial u}, \qquad \frac{\partial}{\partial u}\mathbf{a}(\phi(u, v, \dots)) = \frac{d\mathbf{a}}{d\phi}\frac{\partial \phi}{\partial u},$$

  ii)

$$\frac{\partial}{\partial u}(\mathbf{a} \cdot \mathbf{b}) = \frac{\partial \mathbf{a}}{\partial u} \cdot \mathbf{b} + \mathbf{a} \cdot \frac{\partial \mathbf{b}}{\partial u}, \qquad \frac{\partial}{\partial u}(\mathbf{a} \times \mathbf{b}) = \frac{\partial \mathbf{a}}{\partial u} \times \mathbf{b} + \mathbf{a} \times \frac{\partial \mathbf{b}}{\partial u},$$

  where $\mathbf{a}$, $\mathbf{b}$ are vector functions and ϕ, is a scalar function.

- **Differential of a vector function:**

$$d\mathbf{a} = \frac{\partial \mathbf{a}}{\partial u}\,du + \frac{\partial \mathbf{a}}{\partial v}\,dv + \cdots .$$

Example 1 Calculate the differential of the position vector $\mathbf{r} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$.

$$\frac{\partial \mathbf{r}}{\partial x} = \mathbf{i}, \quad \frac{\partial \mathbf{r}}{\partial y} = \mathbf{j}, \quad \frac{\partial \mathbf{r}}{\partial z} = \mathbf{k},$$

hence $d\mathbf{r} = \mathbf{i}\,dx + \mathbf{j}\,dy + \mathbf{k}\,dz$.

We define the linear vector differential operator**del (or nabla)** in cartesian coordinates as follows

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

Let us apply such an operator to scalar and vector functions.

- **The gradient** of a scalar field ϕ:

$$\mathrm{grad}\,\phi = \nabla\phi = \frac{\partial \phi}{\partial x}\,\mathbf{i} + \frac{\partial \phi}{\partial y}\,\mathbf{j} + \frac{\partial \phi}{\partial z}\,\mathbf{k}.$$

This is a vector field and some useful rules are:

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

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

iii) $\nabla(\psi(\phi)) = \psi'(\phi)\nabla\phi$

iv) Special cases: $\nabla r = \mathbf{r}/r$, $\nabla(1/r) = -\mathbf{r}/r^3$, $\nabla\phi(r) = \phi'\mathbf{r}/r$, where r is the modulus of the position vector $\mathbf{r}$, i.e. $r = \sqrt{x^2 + y^2 + z^2}$.

Consider a surface $\phi(x, y, z) = c$ where c is a constant. By definition

$$\begin{aligned} d\phi &= \frac{\partial \phi}{\partial x}\,dx + \frac{\partial \phi}{\partial y}\,dy + \frac{\partial \phi}{\partial z}\,dz \\ &= \left(\frac{\partial \phi}{\partial x}\,\mathbf{i} + \frac{\partial \phi}{\partial y}\,\mathbf{j} + \frac{\partial \phi}{\partial z}\,\mathbf{k}\right) \cdot (dx\,\mathbf{i} + dy\,\mathbf{j} + dz\,\mathbf{k}) = \nabla\phi \cdot d\mathbf{r} = 0. \end{aligned}$$

It follows that $\nabla\phi$ is perpendicular to the surface since $d\mathbf{r}$ is the tangent vector.

Example 2 Consider the surface $\phi(x, y, z) = x^2 + y^2 + z^2 = c$. Calculate $\nabla\phi$ and verify that it is perpendicular to the surface.

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

This vector is proportional to $\mathbf{r}$, hence it is clearly perpendicular to the surface.

- **The divergence** of a vector field $\mathbf{a}$:

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

This is a scalar field and some useful rules are:

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

ii) $\nabla\cdot(\phi\,\mathbf{a}) = \nabla\phi\cdot\mathbf{a} + \phi(\nabla\cdot\mathbf{a})$, $\quad \nabla\cdot(\mathbf{a}\times\mathbf{b}) = \mathbf{b}\cdot(\nabla\times\mathbf{a}) - \mathbf{a}\cdot(\nabla\times\mathbf{b})$

iii) Special case: $\nabla\cdot\mathbf{r} = 3$

iv) If $\nabla\cdot\mathbf{a} = 0$, $\mathbf{a}$ is said to be *solenoidal*.

Example 3 Use the index notation to show that

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

Note that this is a scalar triple product.

$$\begin{aligned}\nabla\cdot(\mathbf{a}\times\mathbf{b}) &= \epsilon_{ijk}\,\nabla_i(a_j b_k) = \epsilon_{ijk}\,(\nabla_i a_j)\,b_k + \epsilon_{ijk}\,a_j\,(\nabla_i b_k)\\ &= \epsilon_{kij}\,b_k\,(\nabla_i a_j) - \epsilon_{jik}\,a_j\,(\nabla_i b_k) = \mathbf{b}\cdot(\nabla\times\mathbf{a}) - \mathbf{a}\cdot(\nabla\times\mathbf{b})\end{aligned}$$

- **The curl** of a vector field $\mathbf{a}$:

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

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

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

iii) Special case: $\nabla\times\mathbf{r} = 0$

iv) If $\nabla\times\mathbf{a} = 0$, $\mathbf{a}$ is said to be *irrotational*.

Note that because ∇ is a differential operator the order matter i.e.

$$\nabla\cdot\mathbf{a} \neq \mathbf{a}\cdot\nabla, \quad \nabla\times\mathbf{a} \neq \mathbf{a}\times\nabla.$$

Let us apply the *nabla* operator on gradient, divergence and curl. We got five possible combination quite common in physics, for instance in electromagnetism. They are:

i) The divergence of a gradient is called **the Laplacian** of the scalar function

$$\nabla\cdot(\nabla\phi) = \nabla^2\phi = \frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2},$$

where ∇^2 is a scalar differential operator and it is called the Lapalcian.

ii) $\nabla \times (\nabla\phi) = 0$, all gradient are irrotational.

iii) $\nabla \cdot (\nabla \times \mathbf{a}) = 0$. All curls are solenoidal.

iv) $\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2\mathbf{a}$.

Example 4 Consider the Maxwell's equations in vacuum

$$\begin{aligned}(a) \quad \nabla \cdot \mathbf{B} &= 0, \quad (b) \quad \nabla \cdot \mathbf{E} = 0,\\ (c) \quad \nabla \times \mathbf{B} &= \epsilon_0\mu_0 \frac{\partial \mathbf{E}}{\partial t}, \quad (d) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.\end{aligned}$$

i) Derive the Laplace equation of electrostatic.

Since $\mathbf{E} = -\nabla\phi, \quad \nabla \cdot \mathbf{E} = -\nabla \cdot (\nabla\phi) = \nabla^2\phi = 0.$

ii) Derive the electromagnetic wave equation.

Take the curl of (d):

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial(\nabla \times \mathbf{B})}{\partial t}.$$

Take the time derivative of (c):

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

Combine the two results

$$\nabla^2\mathbf{E} = \epsilon_0\mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (\epsilon_0\mu_0) = c^{-2} \quad \to \quad \left(\frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} - \nabla^2\right)\mathbf{E}.$$

8.2 Curves and surfaces

- **Curves:** A curve C can be represented by a vector function $\mathbf{r}$ that depends on one parameter u (*parametric representation*)

$$\mathbf{r}(u) = x(u)\,\mathbf{i} + y(u)\,\mathbf{j} + z(u)\,\mathbf{k}.$$

Example 5 Provide a parametric representation for the following curves

i) The curve $y = -x$ with $-1 \le x \le 1$.

$\mathbf{r}(u) = u\,\mathbf{i} - u\,\mathbf{j}$.

ii) The curve $x^2/4 + y^2 = 1$ with $y \ge 0$ and $z = 3$.

$\mathbf{r}(u) = 2\,\cos u\,\mathbf{i} + \sin u\,\mathbf{j} + 3\,\mathbf{k}$.

i) The derivative $\mathbf{r}'(u) \equiv \mathbf{t}(u)$ is a vector tangent to the curve at each point.

ii) The *arch length* s measured along the curve satisfies:

$$\left(\frac{ds}{du}\right)^2 = \frac{d\mathbf{r}}{du} \cdot \frac{d\mathbf{r}}{du} = \left|\frac{d\mathbf{r}}{du}\right|^2, \quad ds = \pm\sqrt{\frac{d\mathbf{r}}{du} \cdot \frac{d\mathbf{r}}{du}}\,du,$$

where the sign fixes the direction of measuring s, for increasing or decreasing u. Note that ds is the *line element* of the curve.

iii) *Unit tangent vector* to the curve:

$$\hat{\mathbf{t}} = \frac{d\mathbf{r}}{du}\Big/\left|\frac{d\mathbf{r}}{du}\right| = \frac{d\mathbf{r}}{du}\frac{du}{ds} = \frac{d\mathbf{r}}{ds}.$$

Note the use of the chain rule.

iv) The derivative of the unit tangent vector with respect to the arch length defines the *radius of curvature* ρ:

$$\rho = \frac{1}{|\mathbf{n}|}, \quad \text{with} \quad \mathbf{n} = \frac{d^2\mathbf{r}}{ds^2} = \frac{d\hat{\mathbf{t}}}{ds}.$$

Note that $\hat{\mathbf{t}}$ is perpendicular to $\mathbf{n}$ since starting with $\hat{\mathbf{t}}^2 = 1$ and applying the derivative, we get $2\hat{\mathbf{t}} \cdot (d\hat{\mathbf{t}}/ds) = 0$.

Example 6 Consider the curve $\mathbf{r}(u) = 3\,\cos u\,\mathbf{i} + 3\,\cos u\,\mathbf{j} + 4u\,\mathbf{k}$. find its radius of curvature.

$$\frac{d\mathbf{r}}{du} = \mathbf{t} = -3\,\sin u\,\mathbf{i} + 3\,\cos u\,\mathbf{j} + 4\,\mathbf{k}.$$

The modulus is: $\left|\frac{d\mathbf{r}}{du}\right| = \frac{ds}{du} = 5$, hence $\hat{\mathbf{t}} = \frac{1}{5}\left(-3\,\sin u\,\mathbf{i} + 3\,\cos u\,\mathbf{j} + 4\,\mathbf{k}\right)$.
Then

$$\mathbf{n} = \frac{d\hat{\mathbf{t}}}{du}\frac{du}{ds} = \frac{1}{25} = \left(-3\,\cos u\,\mathbf{i} - 3\,\sin u\,\mathbf{j}\right).$$

It follows that the radius of curvature is $\rho = 25/3$.

- **Surfaces:** A surface S can be represented by a vector function $\mathbf{r}$ that depends on two parameters u and v (*parametric representation*)

$$\mathbf{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k}$$

i) The vectors $\partial\mathbf{r}/\partial u$, $\partial\mathbf{r}/\partial v$ are linear independent and tangent to the curve S.

ii) A vector normal to the surface is:

$$\mathbf{n} = \frac{\partial\mathbf{r}}{\partial u} \times \frac{\partial\mathbf{r}}{\partial v}.$$

iii) *Vector area element:*

$$d\mathbf{S} = \left(\frac{\partial\mathbf{r}}{\partial u} \times \frac{\partial\mathbf{r}}{\partial v}\right) du\,dv = \mathbf{n}\,du\,dv.$$

iv) *Scalar area element:*

$$dS = \left|\frac{\partial\mathbf{r}}{\partial u} \times \frac{\partial\mathbf{r}}{\partial v}\right| du\,dv = |\mathbf{n}|\,du\,dv.$$

dS represents a small area of the surface S.

v) The orientation of the surface S is determined by the sign of $\mathbf{n}$.

vi) A surface S is *orientable* if the vector $\mathbf{n}$ can be determined everywhere by a choice of sign.

vii) A surface is *bounded* if it can be contained within some shpere. A bounded surface can have a boundary, ∂S, consisting of a smooth closed curve. A bounded surface with no boundary is *closed*.

Chapter 9

Integrals

9.1 Line integrals

The line integral (or path integral) of a vector field $\mathbf{a}(\mathbf{r})$ along the curve C is:

$$\int_C \mathbf{a}(\mathbf{r}) \cdot d\mathbf{r} = \int_{u_{min}}^{u_{max}} \mathbf{a}(\mathbf{r}(u)) \cdot \frac{d\mathbf{r}}{du}\, du,$$

where C is a smooth *oriented* (a direction along C must be specified) curve defined by the equation $\mathbf{r}(u)$ with endpoints $\mathbf{A} = \mathbf{r}(u_{min})$ and $\mathbf{B} = \mathbf{r}(u_{max})$.

Example 1 Consider the vector function $\mathbf{a}(\mathbf{r}) = x\, e^y\, \mathbf{i} + z^2\, \mathbf{j} + xy\, \mathbf{k}$. Evaluate the integral $\int_C \mathbf{a} \cdot d\mathbf{r}$ along the following curves, with the same end points $\mathbf{A} = (0, 0, 0)$ and $\mathbf{B} = (1, 1, 1)$.

i) $C_1 : \ \mathbf{r}(u) = u\, \mathbf{i} + u\, \mathbf{j} + u\, \mathbf{k}, \quad 0 \leq u \leq 1.$

The parametrisation tells us that $x = u$, $y = u$, $z = u$. Hence

$$\mathbf{a}(\mathbf{r}(u)) = u\, e^u\, \mathbf{i} + u^2\, \mathbf{j} + u^2\, \mathbf{k}, \quad \mathbf{r}'(u) = \mathbf{i} + \mathbf{j} + \mathbf{k}, \quad \mathbf{a}(\mathbf{r}(u)) \cdot \mathbf{r}'(u) = u\, e^u + 2\, u^2,$$

hence

$$\int_{C_1} \mathbf{a} \cdot d\mathbf{r} = \int_0^1 (u\, e^u + 2\, u^2)\, du = \frac{5}{3}.$$

ii) $C_2: \ \mathbf{r}(u) = u\,\mathbf{i} + u^2\,\mathbf{j} + u^3\,\mathbf{k}, \quad 0 \le u \le 1.$

The parametrisation tells us that $x = u$, $y = u^2$, $z = u^3$. Hence

$$\begin{aligned} \mathbf{a}(\mathbf{r}(u)) &= u\,e^{u^2}\,\mathbf{i} + u^6\,\mathbf{j} + u^3\,\mathbf{k}, \quad \mathbf{r}'(u) = \mathbf{i} + 2u\,\mathbf{j} + 3u^2\mathbf{k}, \\ \mathbf{a}(\mathbf{r}(u)) \cdot \mathbf{r}'(u) &= u\,e^{u^2} + 2\,u^7 + 3u^5, \end{aligned}$$

hence

$$\int_{C_2} \mathbf{a} \cdot d\mathbf{r} = \int_0^1 (u\,e^{u2} + 2\,u^7 + 3u^5)\,du = \frac{e}{2} + \frac{1}{4}.$$

- Properties/Observations.

  i) In general the integral depends on the endpoints and the path C.

  ii) $\int_C \mathbf{a} \cdot d\mathbf{r} = -\int_{-C} \mathbf{a} \cdot d\mathbf{r}$, where C is a curve with orientation $A \longrightarrow B$ and $-C$ is a curve with orientation $B \longrightarrow A$.

  iii) If $C = C_1 + C_2 + \cdots + C_n$ then $\int_C \mathbf{a} \cdot d\mathbf{r} = \int_{C_1} \mathbf{a} \cdot d\mathbf{r} + \int_{C_2} \mathbf{a} \cdot d\mathbf{r}, \cdots + \int_{C_n} \mathbf{a} \cdot d\mathbf{r}$. Note that it must be a set of compatible choices amongst the different segments.

  iv) Other kinds of line integrals are possible. For instance:

$$\int_C \phi\,d\mathbf{r}, \qquad \int_C \mathbf{a} \times d\mathbf{r}, \qquad \int_C \phi\,ds, \qquad \int_C \mathbf{a}\,ds.$$

Example 2 Evaluate $\int_C \phi\,ds$ where $\phi = (x-y)^2$ and $\mathbf{r}(u) = a\,\cos u\,\mathbf{i} + a\,\sin u\,\mathbf{j}, \quad 0 \le u \le \pi$, a constant.
$ds = (\sqrt{d\mathbf{r}/du \cdot d\mathbf{r}/du})\,du = a\,du$, then

$$\int_C \phi\,ds = \int_0^\pi (a\,\cos u - a\,\sin u)^2\,a\,du = \pi\,a^3.$$

- **Simply connected region.** A region D is simply connected if every closed path within D can be shrunk to a point without leaving the region.

- **Theorem:** Consider the integral $I = \int_C \mathbf{a} \cdot d\mathbf{r}$, where the path C is in a simply connected region D. Then, the following statements are equivalent:

i) The line integral I is independent of the path C. It only depends on the end-points of the path C.

ii) It exists a scalar function ϕ (a potential) such that $\mathbf{a} = \nabla\phi$.

Notice that the sign is a convention. In physics, because of the meaning of potential in association with forces, we use $\mathbf{a} = -\nabla\phi$.

iii) $\nabla \times \mathbf{a} = 0$.

The vector field $\mathbf{a}$ is said to be *conservative* (or *irrotational*) and ϕ is its *potential*. In addition:

i) $I = \int\limits_C \nabla\phi \cdot d\mathbf{r} = \phi(\mathbf{B}) - \phi(\mathbf{A})$ where $\mathbf{A}$ and $\mathbf{B}$ are the endpoints of the path C.

Notice that if you use $\mathbf{a} = -\nabla\phi$, then $I = \phi(\mathbf{A}) - \phi(\mathbf{B})$.

ii) The line integral I along any closed path C in D is zero.

Example 3 Consider the vector function $\mathbf{a}(\mathbf{r}) = (xy^2 + z)\,\mathbf{i} + (x^2y + 2)\,\mathbf{j} + x\,\mathbf{k}$.

i) Show that the field $\mathbf{a}$ is conservative and find its potential ϕ.

Since $\nabla \times \mathbf{a} = 0$, the filed is conservative. Hence $\mathbf{a} = \nabla\phi = \frac{\partial\phi}{\partial x}\,\mathbf{i} + \frac{\partial\phi}{\partial y}\,\mathbf{j} + \frac{\partial\phi}{\partial z}\,\mathbf{k}$.

$$\begin{aligned}
\frac{\partial\phi}{\partial x} &= a_x = xy^2 + z \quad \rightarrow \quad \phi = \frac{x^2y^2}{2} + zx + f(y,z), \\
\frac{\partial\phi}{\partial y} &= a_y = x^2y + 2 = x^2y + \frac{\partial f}{\partial y} \quad \rightarrow \quad f = 2y + g(z), \\
\frac{\partial\phi}{\partial z} &= a_z = x = x + \frac{dh}{dz} \quad \rightarrow \quad h = c.
\end{aligned}$$

It follows that $\phi = (xy)^2/2 + xz + 2y + c$.

ii) Evaluate the integral $\int_C \mathbf{a} \cdot d\mathbf{r}$ along the curve $\mathbf{r}(u) = u\,\mathbf{i} + 1/u\,\mathbf{j} + \mathbf{k}$ with end points $\mathbf{A} = (1,1,1)$ and $\mathbf{B} = (3, 1/3, 1)$.

$$\int\limits_C \mathbf{a} \cdot d\mathbf{r} = \phi(\mathbf{B}) - \phi(\mathbf{A}) = \frac{2}{3}.$$

9.2 Surface integrals

The surface integrals of vector functions $\mathbf{a}(\mathbf{r})$ over a smooth surface S, defined by $\mathbf{r}(u, v)$ with orientation given by the normal $\hat{\mathbf{n}}$, is:

$$\int_S \mathbf{a}(\mathbf{r}) \cdot d\mathbf{S} = \int_S \mathbf{a}(\mathbf{r}) \cdot \hat{\mathbf{n}} dS = \int_{u_{min}}^{u_{max}}\int_{v_{min}}^{v_{max}} \mathbf{a}(\mathbf{r}(u.v)) \cdot \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) du\, dv.$$

<u>*Example 4*</u> Evaluate the integral $\int_S \mathbf{a} \cdot d\mathbf{S}$ where $\mathbf{a} = x\,\mathbf{i}$ and S is the surface of the hemisphere $x^2 + y^2 + z^2 = a^2$ with $z \geq 0$ and a constant. Use spherical polar coordinates to parametrise the surface (see further down in the notes.)

$\mathbf{r}(\theta, \phi) = a \sin\theta \cos\phi\, \mathbf{i} + a \sin\theta \sin\phi\, \mathbf{j} + a \cos\theta\, \mathbf{k}$ and

$$\frac{\partial \mathbf{r}}{\partial \theta} = a\cos\theta\cos\phi\,\mathbf{i} + a\cos\theta\sin\phi\,\mathbf{j} - a\sin\theta\,\mathbf{k}, \quad \frac{\partial \mathbf{r}}{\partial \theta} = -a\sin\theta\sin\phi\,\mathbf{i} + a\sin\theta\cos\phi\,\mathbf{j}.$$

Hence

$$d\mathbf{S} = \left(\frac{\partial \mathbf{r}}{\partial \theta} \times \frac{\partial \mathbf{r}}{\partial \phi}\right) d\theta d\phi = a\sin\theta\mathbf{r}\, d\theta d\phi \quad \mathbf{a}(\mathbf{r}(\theta, \phi)) = a\sin\theta\cos\phi\,\mathbf{i},$$

and $\mathbf{a} \cdot d\mathbf{S} = a^3 \sin^3\theta \cos^2\phi\, d\theta d\phi$. It follows that

$$\int_S \mathbf{a} \cdot d\mathbf{S} = a^2 \int_0^{2\pi} d\phi\, \cos^2\phi \int_0^{\pi/2} d\theta\, \sin^3\theta = \frac{2}{3}\, a^3\pi.$$

- Observation.
  - i) The integral depends on the orientation of the surface S since the sign of $d\mathbf{S}$ depends on the orientation of S.
  - ii) If the surface is closed, by convention, the vector $\mathbf{n}$ is pointed outwards the volume enclosed.
  - iii) In order to parametrise the surface is often useful to use alternative coordinates systems. For instance:
    - a) **Cylindrical polar coordinates:**

$$\begin{cases} x = \rho\cos\phi \\ y = \rho\sin\phi \\ z = z \end{cases}, \qquad \rho \geq 0, \quad 0 \leq \phi < 2\pi, \quad -\infty < z < \infty$$

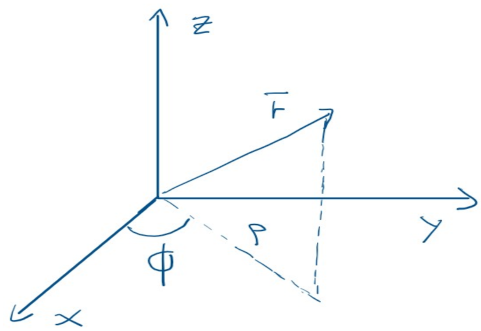

Figure 9.1: Cylindrical polar coordinates.

b) **Spherical polar coordinates:**

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases}, \qquad r \geq 0, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \theta \leq \pi$$

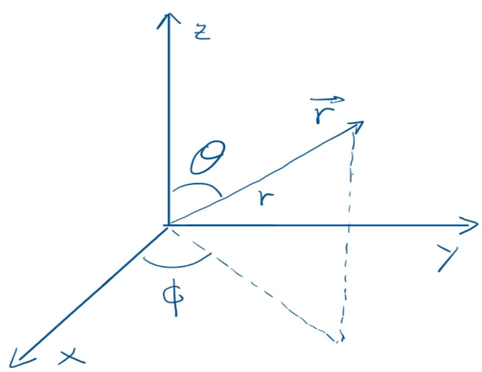

Figure 9.2: Spherical polar coordinates.

iv) Other kinds of integrals are possible. For instance:

$$\int_S \phi\, d\mathbf{S}, \qquad \int_S \mathbf{a} \times d\mathbf{S}, \qquad \int_S \phi\, dS, \qquad \int_S \mathbf{a}\, dS.$$

Example 5 Evaluate the integral $\int_S dS$ where S is the surface of the hemisphere $x^2 + y^2 + z^2 = a^2$ with $z \geq 0$ and a constant.

$dS = |d\mathbf{S}| = a^2 \sin\theta\, d\theta d\phi \quad \rightarrow \quad \int_S dS = \int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta\, a^2 \sin\theta = 2\pi a^2.$

9.3 Volume integrals

Volume integrals of a function $\phi(\mathbf{r})$ (or $\mathbf{a}(\mathbf{r})$) over a volume V described by $\mathbf{r}(u, v, w)$ is:

$$\int_V \phi(\mathbf{r})\, dV = \int_{u_{min}}^{u_{max}}\int_{v_{min}}^{v_{max}}\int_{w_{min}}^{w_{max}} \phi(\mathbf{r}(u,v,w)) \left|\frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right)\right| dudvdw.$$

Observation

- dV The order to parametrisation does not matter.

- Given a parametrisation $\mathbf{r}(u, v, w)$

$$dV = \frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right) dudvdw = \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| dudvdw,$$

  where the last expression highlights the 3-dimensional Jacobian.

  i) For cylindrical polar coordinates: $dV = \rho\, d\rho d\phi dz$.

  ii) For spherical polar coordinates: $dV = r^2 \sin\theta\, dr d\theta d\phi$.

9.4 Theorems on integrals

- **Divergence theorem (Gauss' theorem):**

$$\iiint_V (\nabla \cdot \mathbf{a})\, dV = \iint_S \mathbf{a} \cdot d\mathbf{S},$$

  where V is a bounded volume with boundary $\partial V = S$ and S is a closed surface $\hat{\mathbf{n}}(d\mathbf{S})$ with normal pointing outwards.

Example 6 Use the divergence theorem to show that the Gauss's law of electrostatic for a point-like charge q is equivalent to the Maxwell's equation $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$.

Consider the point-like charge at the origin. Then

$$\mathbf{E} = \frac{q\mathbf{r}}{4\pi\epsilon_0 r^3}, \quad \text{with} \quad q = \iiint\limits_V \rho \, dV,$$

where ρ is the charge density. Gauss' law states: $\iint_S = \mathbf{E} \cdot d\mathbf{S} = q/\epsilon_0$. By the divergence theorem

$$\iint\limits_S \mathbf{E} \cdot d\mathbf{S} = \iiint\limits_V (\nabla \cdot \mathbf{E}), dV = \frac{1}{\epsilon_0} \iiint\limits_V \rho \, dV.$$

Hence the initial statement is proved.

Example 7 Take V to be the solid hemisphere $x^2 + y^2 + z^2 \leq a^2$ with $z \geq 0$ and a constant and $\mathbf{a}$ the vector function $\mathbf{a} = (z + a)\, \mathbf{k}$. Verify the divergence theorem.

$$\iiint\limits_V (\nabla \cdot \mathbf{a}) \, dV = \iint\limits_{S_1} \mathbf{a} \cdot d\mathbf{S}_1 + \iint\limits_{S_2} \mathbf{a} \cdot d\mathbf{S}_2.$$

For the integral on the left hand side:

$$\nabla \cdot \mathbf{a} = 1 \quad \rightarrow \quad \iiint\limits_V (\nabla \cdot \mathbf{a}) \, dV = \frac{2}{3}\pi a^3.$$

For the integrals on the right hand side use spherical polar coordinates.
From *Example 4* we know that $d\mathbf{S}_1 = a \sin\theta \, \mathbf{r} \, d\theta d\phi$. Also

$$\mathbf{a}(\mathbf{r}_1(\theta, \phi)) = (a\cos\theta + a)\, \mathbf{k}, \quad \mathbf{a} \cdot d\mathbf{S}_1 = a^3(\cos\theta + 1)\sin\theta\cos\theta \, d\theta d\phi,$$

hence

$$\iint\limits_{S_1} \mathbf{a} \cdot d\mathbf{S}_1 = a^3 2\pi \int\limits_0^{\pi/2} (\cos^2\theta + \sin\theta\cos\theta) \, d\theta = \frac{5}{3}\pi a^3.$$

For the second integral on the right hand side a suitable parametrisation is $\mathbf{r}_2(\phi) = r\,\cos\phi\,\mathbf{i} + r\,\sin\phi\,\mathbf{j}$. Hence

$$d\mathbf{S}_1 = \left(\frac{\partial \mathbf{r}}{\partial \phi} \times \frac{\partial \mathbf{r}}{\partial r}\right) dr d\phi = -r\,\mathbf{k}\,dr d\phi \quad \to \quad \iint_{S_2} \mathbf{a}\cdot d\mathbf{S}_2 = -a2\pi \int_0^a r\,dr = -\pi a^3.$$

It follows that the divergence theorem reads: $\quad 2/3\,\pi a^3 = 5/3\pi a^3 - \pi a^3.$

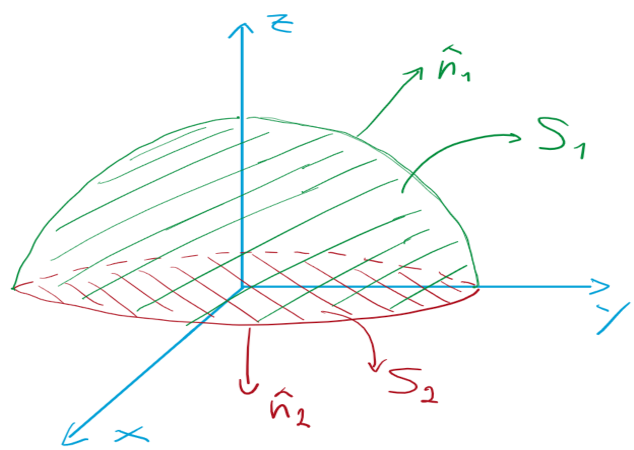

Figure 9.3: *Example 7.*

<u>*Example 8*</u> Derive Gauss's law for a general surface S. Then use the divergence theorem to show that

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

where $\delta(\mathbf{r})$ is the 3-dimensional delta function

$$\iiint_V f(\mathbf{r})\delta(\mathbf{r} - \mathbf{a})\,dV = \begin{cases} f(\mathbf{r}) & \mathbf{a} \in V \\ 0 & \text{otherwise} \end{cases}.$$

Consider a point-like charge at the origin. Hence, applying the divergence theorem

$$\iint_S \mathbf{E} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} \iint_S \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} \iiint_V \nabla \cdot \left(\frac{\mathbf{r}}{r^3}\right) dV = 0 \quad \text{if} \quad r \neq 0.$$

In fact, $\nabla \cdot (\mathbf{r}/r^3) = 0$ for $r \neq 0$, i.e. if the origin is not inside the surface S. Then consider the volume between the surface S and a small sphere around the origin S'. Because of the divergence theorem

$$\iint_S \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S} - \iint_{S'} \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S}' = 0.$$

We can calculate easily the second integral, since $d\mathbf{S}' = r^2\,\hat{\mathbf{r}} \sin\theta d\theta d\phi$, hence

$$\iint_{S'} \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S}' = 2\pi \int_0^\pi \sin\theta d\theta = 4\pi.$$

It follows that

$$\iint_S \mathbf{E} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} \iint_S \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S} = \frac{q}{4\pi\epsilon_0} 4\pi = \frac{q}{\epsilon_0}.$$

Since

$$\iint_S \frac{\mathbf{r}}{r^3} \cdot d\mathbf{S} = \begin{cases} 4\pi & r = 0 \\ 0 & r \neq 0 \end{cases},$$

it follows that

$$\iiint_V \nabla \cdot \left(\frac{\mathbf{r}}{r^3}\right) dV = -\iiint_V \nabla^2 \left(\frac{1}{r}\right) dV = \iiint_V 4\pi\delta(\mathbf{r})\, dV.$$

The final result follows.

- **Stokes' theorem:**

$$\iint_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_C \mathbf{a} \cdot d\mathbf{r},$$

where S is a bounded smooth surface with boundary $\partial S = C$ and S is a piecewise smooth curve. C and S must have compatible orientation.

Compatible orientation: Imagine you are walking on the surface (side with the normal pointing out). If you walk near the edge of the surface in the direction corresponding to the orientation of C, then the surface must be to your left.

Example 9 Take $\mathbf{a} = xz\,\mathbf{j}$ and S be the section of the cone $x^2 + y^2 = z^2$, with $a \le z \le b$, $b > a > 0$. Verify Stokes' theorem. Hint: Use cylindrical polar coordinates.

$$\iint_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_{C_a} \mathbf{a} \cdot d\mathbf{r}_a + \int_{C_b} \mathbf{a} \cdot d\mathbf{r}_b$$

On the left hand side:
A suitable parametrisation is: $\mathbf{r}(\rho, \phi) = \rho\,\cos\phi\,\mathbf{i} + \rho\,\sin\phi\,\mathbf{j} + \rho\,\mathbf{k}$, hence

$$d\mathbf{S} = \left(\frac{\partial \mathbf{r}}{\partial \rho} \times \frac{\partial \mathbf{r}}{\partial \phi} \right) d\rho d\phi = \rho(-\cos\phi\,\mathbf{i} - \sin\phi\,\mathbf{j} + \mathbf{k})\,d\rho d\phi.$$

The field evaluated on the surface is: $\nabla \times \mathbf{a} = \rho\,(-\cos\phi\,\mathbf{i} + \mathbf{k})$. It follows that

$$\iint_S (\nabla \times \mathbf{a}) \cdot d\mathbf{S} = \int_0^{2\pi} d\phi \int_a^b (\rho^2 \cos^2\phi + \rho^2)\,d\rho = \pi(b^3 - a^3).$$

On the right hand side:
A suitable parametrisation for the circle with radius b is:

$$\mathbf{r}(\phi) = b\,\cos\phi\,\mathbf{i} + b\,\sin\phi\,\mathbf{j} + b\,\mathbf{k}, \quad \mathbf{r}'(\phi) = -b\,\sin\phi\,\mathbf{i} + b\,\cos\phi\,\mathbf{j}, \quad \mathbf{a} = b^2\,\cos\phi\,\mathbf{j},$$

then

$$\int_{C_b} \mathbf{a} \cdot d\mathbf{r}_b = b^3 \int_0^{2\pi} \cos^2\phi\,d\phi = b^3\pi.$$

The circle with radius a has a clockwise parametrisation, hence

$$\int_{C_a} \mathbf{a} \cdot d\mathbf{r}_a = -a^3 \int_0^{2\pi} \cos^2\phi\,d\phi = -a^3\pi.$$

It follows that the Stokes' theorem reads $\pi(b^3 - a^3) = -\pi\,a^3 + \pi\,b^3$.

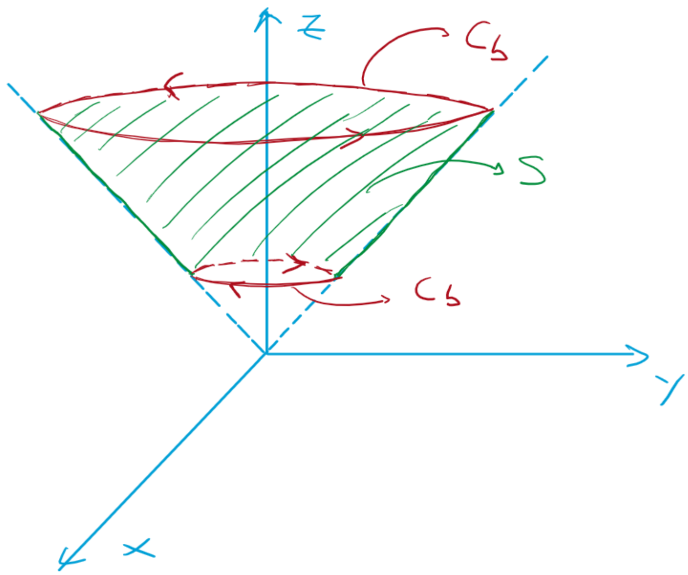

Figure 9.4: *Example 9.*

Chapter 10

Change of variables: orthogonal curvilinear coordinates

Given the position vector $\mathbf{r}$ expressed in cartesian coordinates x, y, z we can use a change of variable to express this vector in terms of a new set of coordinates u, v, w

$$\mathbf{r}(u,v,w) = x(u,v,w)\,\mathbf{i} + y(u,v,w)\,\mathbf{j} + z(u,v,w)\,\mathbf{k},$$

where x, y, z are continuous and differentiable functions.
The line element is:

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial u}\,du + \frac{\partial \mathbf{r}}{\partial v}\,dv + \frac{\partial \mathbf{r}}{\partial w}\,dw.$$

where the vectors $\partial\mathbf{r}/\partial u$, $\partial\mathbf{r}/\partial v$, $\partial\mathbf{r}/\partial w$ are linearly independent. If these vectors are orthogonal, then the coordinates u, v, w are said to be **orthogonal curvilinear coordinates**.

Properties

- **New basis:**
  $$\frac{\partial \mathbf{r}}{\partial u} = h_u\,\hat{\mathbf{e}}_u, \quad \frac{\partial \mathbf{r}}{\partial v} = h_v\,\hat{\mathbf{e}}_v, \quad \frac{\partial \mathbf{r}}{\partial w} = h_w\,\hat{\mathbf{e}}_w,$$
  where h_u, h_v h_w positive and called *scale factors*. In an orthogonal curvilinear coordinate system these vectors are orthogonal and $\hat{\mathbf{e}}_u$, $\hat{\mathbf{e}}_v$, $\hat{\mathbf{e}}w$ form an orthonormal basis of the three dimensional vector space $\mathbb{R}^3$.

- **Line element:**
  $$d\mathbf{r} = h_u\,\hat{\mathbf{e}}_u\,du + h_v\,\hat{\mathbf{e}}_v\,dv + h_w\,\hat{\mathbf{e}}_w\,dw.$$
  The scale factors determine the changes in length along each orthogonal direction resulting from changes in u, v, w.

- **Arc length:**
  $$ds^2 = d\mathbf{r}\cdot d\mathbf{r} = h_u^2\,(du)^2 + h_v^2\,(dv)^2 + h_w^2\,(dw)^2$$

- **Vector area element (surface of constant w, parametrised by u and v):**

$$d\mathbf{S} = \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) du\, dv = h_u h_v\, du dv.$$

- **Volume element:**

$$dV = \left|\frac{\partial \mathbf{r}}{\partial u} \cdot \left(\frac{\partial \mathbf{r}}{\partial v} \times \frac{\partial \mathbf{r}}{\partial w}\right)\right| du\, dv\, dw = h_u h_v h_w\, du dv dw.$$

- Note that vector algebra is the same in orthogonal curvilinear coordinates as in cartesian coordinates.

Example 1 Derive the the scale factors, basis vector and volume elements for

1) Cartesian coordinates.

   $\mathbf{r}(x, y, z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$, hence

   $$\frac{\partial \mathbf{r}}{\partial x} = \mathbf{i}, \quad \frac{\partial \mathbf{r}}{\partial y} = \mathbf{j}, \quad \frac{\partial \mathbf{r}}{\partial z} = \mathbf{k} \quad \to \quad h_x = h_y = h_z = 1, \quad \hat{\mathbf{e}}_x = \mathbf{i}, \quad \hat{\mathbf{e}}_y = \mathbf{j}, \quad \hat{\mathbf{e}}_z = \mathbf{k},$$

   and $dV = dxdydz$.

2) Cylindrical polar coordinates.

   $\mathbf{r}(\rho, \phi, z) = \rho\,\cos\phi\,\mathbf{i} + \rho\,\sin\phi\,\mathbf{j} + z\,\mathbf{k}$, hence

   $$\frac{\partial \mathbf{r}}{\partial \rho} = \cos\phi\,\mathbf{i} + \sin\phi\,\mathbf{j}, \quad \frac{\partial \mathbf{r}}{\partial \phi} = -\rho\,\sin\phi\,\mathbf{i} + \rho\,\cos\phi\,\mathbf{j}, \quad \frac{\partial \mathbf{r}}{\partial z} = \mathbf{k}, \quad \to$$

   $h_\rho = 1, \quad h_\phi = \rho, \quad h_z = 1, \quad \hat{\mathbf{e}}_\rho = \cos\phi\,\mathbf{i}+\sin\phi\,\mathbf{j}, \quad \hat{\mathbf{e}}_\phi = -\sin\phi\,\mathbf{i}+\cos\phi\,\mathbf{j}, \quad \hat{\mathbf{e}}_z = \mathbf{k}$
   and $dV = \rho\, d\rho d\phi dz$.

Gradient, divergence and curl in orthogonal curvilinear coordinates.

Consider a scalar function $f(u, v, w)$. Then

$$df = \frac{\partial f}{\partial u}\, du + \frac{\partial f}{\partial v}\, dv + \frac{\partial f}{\partial w}\, dw = \nabla f \cdot d\mathbf{r}.$$

In cartesian coordinates this becomes

$$\nabla f \cdot d\mathbf{r} = \left(\mathbf{i}\,\frac{\partial}{\partial x} + \mathbf{j}\,\frac{\partial}{\partial y} + \mathbf{k}\,\frac{\partial}{\partial z}\right) \cdot (\mathbf{i}\, dx + \mathbf{j}\, dy + \mathbf{k}\, dz)\,.$$

On the other hand, in general curvilinear coordinates it is:

$$\nabla f \cdot d\mathbf{r} = \nabla f \cdot \left(h_u\, \mathbf{e}_u\, du + h_v\, \mathbf{e}_v\, dv + h_w\, \mathbf{e}_w\, dw\right),$$

which implies

$$\nabla f = \frac{\hat{\mathbf{e}}_u}{h_u}\frac{\partial f}{\partial u} + \frac{\hat{\mathbf{e}}_v}{h_v}\frac{\partial f}{\partial v} + \frac{\hat{\mathbf{e}}_w}{h_w}\frac{\partial f}{\partial w}.$$

This is the **gradient** of the function f in general curvilinear coordinates. It follows that the **del operator** is:

$$\nabla = \frac{\hat{\mathbf{e}}_u}{h_u}\frac{\partial}{\partial u} + \frac{\hat{\mathbf{e}}_v}{h_v}\frac{\partial}{\partial v} + \frac{\hat{\mathbf{e}}_w}{h_w}\frac{\partial}{\partial w}.$$

Without derivation, we also have

- **Divergence:**

$$\nabla \cdot \mathbf{a} = \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}(h_v h_w a_u) + \frac{\partial}{\partial v}(h_w h_u a_v) + \frac{\partial}{\partial w}(h_u h_v a_w)\right],$$

  where $\mathbf{a} = a_u\, \hat{\mathbf{e}}_u + a_v\, \hat{\mathbf{e}}_v + a_w\, \hat{\mathbf{e}}_w$.

- **Curl:**

$$\nabla \times \mathbf{a} = \frac{1}{h_u h_v h_w}\begin{vmatrix} h_u\, \hat{\mathbf{e}}_u & h_v\, \hat{\mathbf{e}}_v & h_w\, \hat{\mathbf{e}}_w \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial w} \\ h_u\, a_u & h_v\, a_v & h_w\, a_w. \end{vmatrix}$$

- **Laplacian:**

$$\nabla^2 \phi = \frac{1}{h_u h_v h_w}\left[\frac{\partial}{\partial u}\left(\frac{h_v h_w}{h_u}\frac{\partial \phi}{\partial u}\right) + \frac{\partial}{\partial v}\left(\frac{h_w h_u}{h_v}\frac{\partial \phi}{\partial v}\right) + \frac{\partial}{\partial w}\left(\frac{h_u h_v}{h_w}\frac{\partial \phi}{\partial w}\right)\right].$$

Example 2 Find the position vector $\mathbf{r}$ in cylindrical polar coordinates and verify that $\nabla \cdot \mathbf{r} = 3$.

From Example 1, we have the unit vectors for the cylindrical polar coordinates. By inverting those relations we obtain:

$$\mathbf{i} = \cos\phi\, \hat{\mathbf{e}}_\rho - \sin\phi\, \hat{\mathbf{e}}_\phi, \quad \mathbf{j} = \sin\phi\, \hat{\mathbf{e}}_\rho + \cos\phi\, \hat{\mathbf{e}}_\phi, \quad \mathbf{k} = \hat{\mathbf{e}}_z.$$

Then

$$\mathbf{r} = \rho\, \cos\phi(\cos\phi\, \hat{\mathbf{e}}_\rho - \sin\phi\, \hat{\mathbf{e}}_\phi) + \rho\, \sin\phi(\sin\phi\, \hat{\mathbf{e}}_\rho + \cos\phi\, \hat{\mathbf{e}}_\phi) + z\, \hat{\mathbf{e}}_z = \rho\, \hat{\mathbf{e}}_\rho + z\, \hat{\mathbf{e}}_z$$

and

$$\nabla \cdot \mathbf{r} = \frac{1}{\rho}\left[\frac{\partial}{\partial \rho}\left(\rho^2\right) + \frac{\partial}{\partial z}\left(\rho\, z\right)\right] = 3.$$

Example 3 A rigid body is rotating about a fixed axis with a constant angular velocity $\boldsymbol{\omega}$. Take $\boldsymbol{\omega}$ to lie along the z-axis. Use cylindrical polar coordinates to compute

1) $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$.

The position vector has been found in Example 2. Then $\boldsymbol{\omega} = \omega\,\hat{\mathbf{e}}_z$. Then

$$\mathbf{v} = \begin{vmatrix} \hat{\mathbf{e}}_\rho & \hat{\mathbf{e}}_\phi & \hat{\mathbf{e}}_z \\ 0 & 0 & \omega \\ \rho & 0 & z \end{vmatrix} = \omega\rho\,\hat{\mathbf{e}}_\phi.$$

2) $\nabla \times \mathbf{v}$.

$$\nabla \times \mathbf{v} = \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\,\hat{\mathbf{e}}_\phi & \hat{\mathbf{e}}_z \\ \frac{\partial}{\partial\rho} & \frac{\partial}{\partial\phi} & \frac{\partial}{\partial z} \\ 0 & \omega\rho^2 & 0 \end{vmatrix} = 2\omega\,\hat{\mathbf{e}}_z = 2\,\boldsymbol{\omega}.$$

Summary of common orthogonal curvilinear coordinates:

- **Cylindrical polar coordinates:**

$$\mathbf{r}(\rho,\phi,z) = \rho\cos\phi\,\mathbf{i} + \rho\sin\phi\,\mathbf{j} + z\,\mathbf{k}, \qquad \rho \geq 0, \quad 0 \leq \phi < 2\pi, \quad -\infty < z < \infty$$

i) $$\begin{aligned} &h_\rho = 1, \quad \hat{\mathbf{e}}_\rho = \cos\phi\,\mathbf{i} + \sin\phi\,\mathbf{j}, \\ &h_\phi = \rho, \quad \hat{\mathbf{e}}_\phi = -\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j}, \\ &h_z = 1, \quad \hat{\mathbf{e}}_z = \mathbf{k}. \end{aligned}$$

ii) $$d\mathbf{S} = \begin{cases} \hat{\mathbf{e}}_\rho\,\rho\,d\phi dz & (\rho = \text{const}) \\ \hat{\mathbf{e}}_\phi\,d\rho dz & (\phi = \text{const}) \\ \hat{\mathbf{e}}_z\,\rho\,d\rho d\phi & (z = \text{const}). \end{cases}$$

iii) $dV = \rho\,d\rho d\phi dz$.

- **Spherical polar coordinates:**

$$\mathbf{r}(r,\theta,\phi) = r\sin\theta\cos\phi\,\mathbf{i} + r\sin\theta\sin\phi\,\mathbf{j} + r\cos\theta\,\mathbf{k}, \qquad r \geq 0, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \theta \leq \pi$$

i) $$\begin{aligned} &h_r = 1, && \hat{\mathbf{e}}_r = \sin\theta\cos\phi\,\mathbf{i} + \sin\theta\sin\phi\,\mathbf{j} + \cos\theta\,\mathbf{k}, \\ &h_\theta = r, && \hat{\mathbf{e}}_\theta = \cos\theta\cos\phi\,\mathbf{i} + \cos\theta\sin\phi\,\mathbf{j} - \sin\theta\,\mathbf{k}, \\ &h_\phi = r\sin\theta, && \hat{\mathbf{e}}_\phi = -\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j}. \end{aligned}$$

ii) $$d\mathbf{S} = \begin{cases} \hat{\mathbf{e}}_r\,r^2\sin\theta\,d\theta d\phi & (r = \text{const}) \\ \hat{\mathbf{e}}_\theta\,r\sin\theta\,dr d\phi & (\theta = \text{const}) \\ \hat{\mathbf{e}}_\phi\,r\,dr d\theta & (\phi = \text{const}). \end{cases}$$

iii) $dV = r^2\sin\theta\,dr d\theta d\phi$.