Natural Sciences Tripos: IB Mathematical Methods II

Contents

0 Introduction

0.1 Schedule

The schedules, or syllabuses, are determined by a committee which has input from all the Physical Science subjects in the Natural Sciences and from Computer Science and is agreed by the Faculty of Mathematics. The schedules are minimal for lecturing and maximal for examining; that is to say, all the material in the schedules will be lectured and only this material will be examined.

Below is a copy from the booklet of schedules.[1] The numbers in square brackets at the end of paragraphs of the schedules indicate roughly (I emphasise *roughly*) the number of lectures that will be devoted to the material in the paragraph.

Part IB: Mathematics

This course comprises Mathematical Methods I, Mathematical Methods II, Mathematical Methods III and six Computer Practicals. The material in Course A from Part IA will be assumed in the lectures for this course. Topics marked with asterisks should be lectured, but questions will not be set on them in examinations.

The material in the course will be as well illustrated as time allows with examples and applications of Mathematical Methods to the Physical Sciences.[2]

Mathematical Methods II **24 lectures, Lent term**

Sturm-Liouville theory
Self-adjoint operators, eigenfunctions and eigenvalues, reality of eigenvalues and orthogonality of eigenfunctions. Eigenfunction expansions and determination of coefficients. Legendre polynomials; orthogonality. [**3**]

Conditional stationary values and the calculus of variations
Lagrange multipliers, examples with two or three variables. Euler-Lagrange equations and examples.

Variational principles; Fermat's principle; Hamilton's principle and deduction of Lagrange's equation, illustrated by a system with:

$$L = \tfrac{1}{2}\, m_1 \dot{x}_1^2 + \tfrac{1}{2}\, m_2 \dot{x}_2^2 - V(x_1 - x_2)\,.$$

Variational principle for the lowest eigenvalue *and for higher eigenvalues* (Rayleigh-Ritz). [**6**]

Laplace and Poisson's equations
Solution by separation of variables of Laplace's equation in plane polar coordinates, and spherical polar coordinates (axisymmetric case); Legendre polynomials again.

Solution of Poisson's equation as an integral. Uniqueness for Poisson's equation with Dirichlet boundary conditions. Green's identity. Green's function for Laplace's equation with simple boundary conditions using the method of images. Applications to electrostatic fields and steady heat flow. [**5**]

Cartesian tensors
Transformation laws, addition, multiplication, contraction. Isotropic tensors, symmetric and anti-symmetric tensors. Principal axes and diagonalisation. Tensor fields, e.g. conductivity, polarizability, elasticity. [**4**]

Contour integration
Integration along a path; elementary properties. Cauchy's theorem; proof by Cauchy-Riemann equations and divergence theorem in 2–D. Integral of $f'(z)$; Cauchy's formula for $f(z)$. Calculus

[1] See https://www.maths.cam.ac.uk/undergradnst/files/misc/NSTschedules.pdf.
[2] Time is always short.

of residues; examples of contour integration; point at infinity; multi-valued functions, branch points, log (z). [**4**]

Transform methods
Fourier inversion by contour integration. Examples of simple linear differential equations, including diffusion equation. [**2**]

0.2 Books

As noted in the schedules, there are very many books which cover the sort of mathematics required by Natural Scientists. The following should be helpful as general reference. Books which can reasonably be used as principal texts for the course are marked with a dagger.

† G. Arfken & H.J. Weber, *Mathematical Methods for Physicists, 6th edition.* Elsevier, 2005.

D.E. Bourne and P.C. Kendall, *Vector Analysis and Cartesian Tensors, 3rd edition.* Nelson Thornes, 1992.

† J.W. Dettman, *Mathematical Methods in Physics and Engineering.* Dover, 1988.

E. Kreyszig, *Advanced Engineering Mathematics, 8th edition.* Wiley, 1999.

† J. Mathews & R.L. Walker, *Mathematical Methods of Physics, 2nd edition.* Pearson/Benjamin Cummings, 1970.

H.A. Priestley, *Introduction to Complex Analysis, 2nd edition.* Oxford University Press, 2003.

† K.F. Riley, M.P. Hobson & S.J. Bence, *Mathematical Methods for Physics and Engineering, 3rd edition.* Cambridge University Press, 2006 (*available online via the University Library website*).

R.N. Snieder, *A guided tour of mathematical methods for the physical sciences.* Cambridge University Press, 2004.

There is likely to be a resemblance between my notes and Riley, Hobson & Bence. This is because we both used the same source, i.e. previous Cambridge lecture notes.[3]

Of the other books, I like Mathews & Walker, but it might be a little mathematical for some. Also, the first time I gave a 'service' mathematics course (over 30 years ago to aeronautics students at Imperial), my notes bore a resemblance to Kreyszig . . . and that was not because we were using a common source!

0.3 Course Website

See `NST Part IB: Mathematics` on Moodle at

https://www.vle.cam.ac.uk/course/view.php?id=78772

The direct link to this term's section is

https://www.vle.cam.ac.uk/course/view.php?id=78772&sectionid=365782#section-19

but this might break if someone changes the number of sections!

Copies of the previous years' lecture notes (this year's will be remarkably close to last year's version) are available if you look.

0.4 Lectures

- This is a very different year to normal. If I am particularly unclear then, if you are watching the webcast, I welcome *constructive* heckling (although the 40-second lag will not help); hence, if you spot a bad typo, or I use jargon you do not understand, or I am just plain wrong, then please speak up.

[3] When I last lectured NST IB over 15 years ago, a student hoped that Riley *et al.* were getting royalties from my lecture notes; my hope is that my lecturers from 45 years ago are getting royalties from Riley *et al.*!

- If it is a minor typo, and/or if I am inaudible and/or illegible, and/or if you think that I could have been clearer, then
  - ◦ if you are watching the webcast, you can ask questions at the end of the lecture via Zoom (since I will endeavour to stay around for a few minutes);
  - ◦ alternatively, questions (and comments) can be emailed to me at `S.J.Cowley@maths.cam.ac.uk`.
- Comments and administration/organisational queries can also be sent to: `nst@maths.cam.ac.uk`.
- This is a 'service' course, so you will not get pure mathematical levels of rigour. However, I will give some justification for a method, rather than just a recipe, because if you are to use a method efficiently and effectively, or extend it as might be necessary in research, you need to understand *why* a method works, as well as *how* to apply it.
- If anyone is colour blind please tell me which colours you cannot read.

0.5 Example Sheets

- There will be four Example Sheets. They are all available on Moodle now.
- You should be able to do Example Sheets 1/2/3/4 after lectures 9/14/18/24 respectively. Please bear this in mind when arranging supervisions.
- There are answers to the sheets. I will make these available to you on Moodle at the end of weeks 3, 5, 7 and 9 (where I count weeks from 0, with week 0 starting on the Sunday before the Tuesday on which Full Term starts). If I forget to do this, please remind me by email.
- The good news for supervisors is that the sheets are the same as last year (and that they can have access to the answers immediately, as indicated on the Moodle site).

0.6 Examples Classes

There will be Examples Classes on Wednesday 17 February and Wednesday 10 March at 14:00. More details closer to the time.

0.7 Acknowledgements

These lecture notes were developed from the notes of Paul Townsend, Robert Hunt, Christopher Thomas and other previous lecturers, though I am responsible for any errors.

1 Sturm-Liouville Theory

1.0 Why Study This?

Numerous scientific and other phenomena are described by differential equations. You have already encountered a number of different methods for solving linear ordinary differential equations (ODEs), e.g. by writing the solution as a linear combination of solutions (with coefficients determined by the boundary conditions), Green's functions and finding the solution in terms of a power series.

This section is about extending your armoury for solving an important sub-class of ODEs, i.e. those of 'Sturm-Liouville' type, such as those that often arise in quantum mechanics and electrodynamics. In addition we will consider eigenvalue problems for 'Sturm-Liouville' operators. In subsequent courses, if you have not already, you will learn that such eigenvalues fix, say, the angular momentum of electrons in an atom.

1.1 Introduction

Notation. Let $\widetilde{\mathcal{L}}$ be a linear differential operator, e.g.

$$\widetilde{\mathcal{L}} = p(x)\frac{d^2}{dx^2} + r(x)\frac{d}{dx} + s(x)\,,$$

where p, r and s are real functions.

Suppose we want to solve an inhomogeneous ordinary differential equation (ODE) of the form,

$$\widetilde{\mathcal{L}}\,y(x) = F(x)\,, \tag{1.1a}$$

where F is a real function, and also suppose there are known boundary conditions on $y(x)$, say, at $x = \alpha$ and $x = \beta$.

Example. A forced damped oscillator equation is of this form; for instance

$$-\frac{d^2y}{dx^2} - \frac{dy}{dx} - \tfrac{1}{4}y = e^{-x/2}\,, \tag{1.1b}$$

subject to boundary conditions

$$y = 0 \quad \text{at} \quad x = 0\,, \quad \text{and} \quad \frac{dy}{dx} + \tfrac{1}{2}y = 0 \quad \text{at} \quad x = 1. \tag{1.1c}$$

Except for simple $F(x)$, it is usually not possible to find a particular integral in closed form even if we can find solutions of the homogeneous equation with $F(x) = 0$. Our aim to to exploit the *linearity* of $\widetilde{\mathcal{L}}$ to find solutions in terms of a *superposition* of a set of 'basis' functions, cf. the sine and cosine basis functions in the case of Fourier series (see page 13 for the solution to (1.1b) and (1.1c) by this method).

We shall see that a particularly convenient choice for the set of 'basis' functions is the set of *eigenfunctions* of $\widetilde{\mathcal{L}}$, say $y_i(x)$, which satisfy the boundary conditions and,

$$\widetilde{\mathcal{L}}\,y_i(x) = \lambda_i\,y_i(x)\,, \tag{1.1d}$$

where the constants λ_i are the *eigenvalues* of $\widetilde{\mathcal{L}}$ (cf. the *eigenvectors* and *eigenvalues* of a matrix).

However, before using eigenfunctions to find solutions to differential equations, we first need to develop some formalism and explore properties of differential operators and eigenfunctions.

1.2 Inner Products

Analogy between matrices and differential operators. There is a close analogy between matrices and differential operators. Matrices act on vectors in a finite-dimensional vector space, taking one vector to another vector. Differential operators act on functions in some (often infinite-dimensional) space of functions, taking one function to another.

In what follows we must specify the space of functions on which a differential operator acts. Although *we only consider real differential operators*, we will allow for complex functions and, in principle, complex eigenvalues (cf. the possibility of polynomials with real coefficients having complex roots).

Definition of an inner product for a n-dimensional vector space (revision). Let V be an n-dimensional vector space over the complex numbers $\mathbb{C}$, and let vectors $\mathbf{u}$, $\mathbf{v} \in V$. Then recall that a *scalar product* $\mathbf{u} \cdot \mathbf{v} \in \mathbb{C}$ or, in an alternative notation, an *inner product* $\langle \mathbf{u} \mid \mathbf{v} \rangle \in \mathbb{C}$, has the following properties.

(i) If we denote a complex conjugate with * then

$$\langle \mathbf{u} \mid \mathbf{v} \rangle = \langle \mathbf{v} \mid \mathbf{u} \rangle^* . \tag{1.2a}$$

Implicit in this equation is the statement that for a complex vector space the *ordering* of the vectors in the inner product is important. Further, if we let $\mathbf{u} = \mathbf{v}$, then this implies that

$$\langle \mathbf{v} \mid \mathbf{v} \rangle = \langle \mathbf{v} \mid \mathbf{v} \rangle^* , \tag{1.2b}$$

i.e. $\langle \mathbf{v} \mid \mathbf{v} \rangle$ is *real*.

(ii) The inner product is *linear in its second argument*, i.e. for $A, B \in \mathbb{C}$

$$\langle \mathbf{u} \mid A\mathbf{v}_1 + B\mathbf{v}_2 \rangle = A \langle \mathbf{u} \mid \mathbf{v}_1 \rangle + B \langle \mathbf{u} \mid \mathbf{v}_2 \rangle . \tag{1.2c}$$

(iii) The inner product of a vector with itself is positive, i.e.

$$\langle \mathbf{v} \mid \mathbf{v} \rangle \geqslant 0 . \tag{1.2d}$$

This allows us to write $\langle \mathbf{v} \mid \mathbf{v} \rangle = \|\mathbf{v}\|^2$, where the real positive number $\|\mathbf{v}\|$ is the *norm* (cf. length) of the vector $\mathbf{v}$.

(iv) The only vector of zero norm should be the zero vector, i.e.

$$\|\mathbf{v}\| = 0 \quad \Rightarrow \quad \mathbf{v} = 0 . \tag{1.2e}$$

Definition of an inner product of functions.

Given two piecewise continuous functions $u(x)$ and $v(x)$, defined on $\alpha \leqslant x \leqslant \beta$ in a space of complex functions, define an *inner product* (a complex number), $\langle u \mid v \rangle$, by

$$\langle u \mid v \rangle = \int_\alpha^\beta u^*(x) v(x)\, w(x)\, \mathrm{d}x , \tag{1.3a}$$

where $w(x)$ is a real *weight function* such that

$$w(x) > 0 \quad \text{for} \quad \alpha < x < \beta . \tag{1.3b}$$

Notation. When $w \neq 1$, we will sometimes include a subscript w (a non-standard notation) in the inner product, as in $\langle u \mid v \rangle_w$, in order to clarify the role of the weight function w.

For (1.3a) and (1.3b) to define an inner product, we need to confirm that the key properties of an inner product are satisfied. To this end we first note that for piecewise continuous functions u, v, v_1 and v_2, and complex constants A and B, properties (1.2a), (1.2c) and (1.2d) are satisfied by (1.3a):

$$\langle u \mid v \rangle = \langle v \mid u \rangle^* , \tag{1.4a}$$
$$\langle u \mid Av_1 + Bv_2 \rangle = A \langle u \mid v_1 \rangle + B \langle u \mid v_2 \rangle , \tag{1.4b}$$
$$\|v\|^2 \equiv \langle v \mid v \rangle \geqslant 0 , \tag{1.4c}$$

where (1.4c) defines the *norm* of v. Further, it is possible to show for a piecewise continuous linear function v, that (1.2e) is satisfied in the case of an inner product defined by (1.3a) and (1.3b), i.e.

$$\|v\| = 0 \quad \Rightarrow \quad v = 0 . \qquad \square \tag{1.4d}$$

Unlectured remark. However, $\|v\| = 0$ does *not* imply that $v = 0$ for *all* possible functions. For example, a function $v(x)$ that is unity when x is rational and zero otherwise will not contribute to the integral that defines the norm. The problem of how to make the idea of 'reasonably well-behaved', or 'nice', functions mathematically rigorous is left to the pure mathematicians.

1.3 Adjoint and Self-Adjoint Operators

Definition. For a general differential operator $\widetilde{\mathcal{L}}$, and a given inner product $\langle\, u \,|\, v \,\rangle$, the *adjoint* operator $\widetilde{\mathcal{L}}^\dagger$, is *defined* to be that operator such that

$$\left\langle\, u \,\middle|\, \widetilde{\mathcal{L}} v \,\right\rangle = \left\langle\, \widetilde{\mathcal{L}}^\dagger u \,\middle|\, v \,\right\rangle . \tag{1.5a}$$

Definition. A general differential operator $\widetilde{\mathcal{L}}$ is said to be *self-adjoint* if

$$\left\langle\, u \,\middle|\, \widetilde{\mathcal{L}} v \,\right\rangle = \left\langle\, \widetilde{\mathcal{L}} u \,\middle|\, v \,\right\rangle . \tag{1.5b}$$

Remarks.

(i) Self-adjoint operators have many analogous properties to Hermitian matrices, i.e. matrices such that $\mathsf{H}^\dagger = \mathsf{H}$, where in this case † indicates the complex conjugate of the transpose.

For example, suppose that an inner product for column vectors u and v is defined by

$$\langle\, \mathsf{u} \,|\, \mathsf{v} \,\rangle = \mathsf{u}^\dagger\, \mathsf{v} . \tag{1.5c}$$

Then for a Hermitian matrix H

$$\begin{aligned} \langle\, \mathsf{u} \,|\, \mathsf{H}\mathsf{v} \,\rangle &= \mathsf{u}^\dagger \mathsf{H}\, \mathsf{v} = \mathsf{u}^\dagger \mathsf{H}^\dagger\, \mathsf{v} && \text{since } \mathsf{H} \text{ is Hermitian} \\ &= (\mathsf{H}\mathsf{u})^\dagger\, \mathsf{v} = \langle\, \mathsf{H}\mathsf{u} \,|\, \mathsf{v} \,\rangle && \text{since } (\mathsf{A}\mathsf{B})^\dagger = \mathsf{B}^\dagger \mathsf{A}^\dagger . \end{aligned} \tag{1.5d}$$

A comparison of (1.5b) and (1.5d) suggests that self-adjoint operators are to general operators what Hermitian matrices are to general matrices.

(ii) Whether or not an operator $\widetilde{\mathcal{L}}$ is self-adjoint with respect to an inner product depends on the defintion of the inner product, e.g. the choice of weight function.

Example. If $\widetilde{\mathcal{L}}$ is a nth order differential operator, then for an inner product defined by (1.3a), integrate by parts n times to obtain

$$\begin{aligned} \langle u | \widetilde{\mathcal{L}} v \rangle &= \int_\alpha^\beta u^*(x) \widetilde{\mathcal{L}} v(x)\, w(x)\, dx . \\ &= \int_\alpha^\beta \left[\widetilde{\mathcal{L}}^\dagger u(x) \right]^* v(x)\, w(x)\, dx + \text{ b.t.} \\ &= \langle \widetilde{\mathcal{L}}^\dagger u | v \rangle + \text{ b.t.} , \end{aligned} \tag{1.5e}$$

where b.t. stands for the boundary terms arising from the integration by parts; the operator arising from this manipulation *defines* $\widetilde{\mathcal{L}}^\dagger$. If $\widetilde{\mathcal{L}}^\dagger = \widetilde{\mathcal{L}}$, and the boundary terms are zero, then the operator is self-adjoint.

1.4 The Sturm-Liouville Operator

Definition. A second-order linear differential operator $\mathcal{L}$ is said to be of *Sturm-Liouville* type if

$$\mathcal{L} = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho(x) \frac{\mathrm{d}}{\mathrm{d}x} \right) - q(x) , \tag{1.6a}$$

where $\rho(x)$ and $q(x)$ are real functions defined for $\alpha \leqslant x \leqslant \beta$, with

$$\rho(x) > 0 \quad \text{for} \quad \alpha < x < \beta . \tag{1.6b}$$

The Sturm-Liouville operator is self-adjoint for suitable boundary conditions. Consider the Sturm-Liouville operator (1.6a) in conjunction with the identity weight function $w = 1$, then

$$\begin{aligned}
\langle\, u \mid \mathcal{L}v\, \rangle &= \int_\alpha^\beta \mathrm{d}x\, u^* \mathcal{L}v && \text{from (1.3a)}\\
&= -\int_\alpha^\beta \mathrm{d}x\, u^* \left(\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho\frac{\mathrm{d}v}{\mathrm{d}x}\right) + qv\right) && \text{from (1.6a)}\\
&= -\left[u^*\rho\frac{\mathrm{d}v}{\mathrm{d}x}\right]_\alpha^\beta + \int_\alpha^\beta \mathrm{d}x\, \rho\,\frac{\mathrm{d}u^*}{\mathrm{d}x}\frac{\mathrm{d}v}{\mathrm{d}x} - \int_\alpha^\beta \mathrm{d}x\, qu^*v && \text{integrate by parts}\\
&= \left[-u^*\rho\frac{\mathrm{d}v}{\mathrm{d}x} + \rho\frac{\mathrm{d}u^*}{\mathrm{d}x}v\right]_\alpha^\beta - \int_\alpha^\beta \mathrm{d}x\, v\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho\frac{\mathrm{d}u^*}{\mathrm{d}x}\right) - \int_\alpha^\beta \mathrm{d}x\, v\, qu^* && \text{integrate by parts}\\
&= \left[\rho\left(v\frac{\mathrm{d}u^*}{\mathrm{d}x} - u^*\frac{\mathrm{d}v}{\mathrm{d}x}\right)\right]_\alpha^\beta + \int_\alpha^\beta \mathrm{d}x\, v\,\mathcal{L}u^* && \text{from (1.6a)}\\
&= \left[\rho\left(v\frac{\mathrm{d}u^*}{\mathrm{d}x} - u^*\frac{\mathrm{d}v}{\mathrm{d}x}\right)\right]_\alpha^\beta + \int_\alpha^\beta \mathrm{d}x\, (\mathcal{L}u)^* v && \text{since } \mathcal{L} \text{ real}\\
&= \left[\rho\left(v\frac{\mathrm{d}u^*}{\mathrm{d}x} - u^*\frac{\mathrm{d}v}{\mathrm{d}x}\right)\right]_\alpha^\beta + \langle\, \mathcal{L}u \mid v\, \rangle && \text{from (1.3a).} \qquad (1.7a)
\end{aligned}$$

Suppose we now insist that u and v be such that

$$\left[\rho\left(v\frac{\mathrm{d}u^*}{\mathrm{d}x} - u^*\frac{\mathrm{d}v}{\mathrm{d}x}\right)\right]_\alpha^\beta = 0\,, \tag{1.7b}$$

then (1.5b) is satisfied. We conclude that the Sturm-Liouville differential operator

$$\mathcal{L} = -\frac{\mathrm{d}}{\mathrm{d}x}\left(\rho(x)\frac{\mathrm{d}}{\mathrm{d}x}\right) - q(x)$$

acting on piecewise linear functions, say u or v, which satisfy homogeneous boundary conditions at $x = \alpha$ and $x = \beta$ (e.g. $u(\alpha) = 0$, $v(\alpha) = 0$ and $u(\beta) = 0$, $v(\beta) = 0$), is *self-adjoint* with respect to the inner product with $w = 1$.

Remark. The boundary conditions are part of the conditions for an operator to be self-adjoint.

1.5 The Rôle of the Weight Function

Not all second-order linear differential operators have the Sturm-Liouville form (1.6a). However, suppose that $\widetilde{\mathcal{L}}$ is a second-order linear differential operator *not* of Sturm-Liouville form, then we claim that, subject to the restriction (1.9b), there exists a function $w(x)$ so that

$$\mathcal{L} = w\widetilde{\mathcal{L}} \tag{1.8}$$

is of Sturm-Liouville form.

Proof. The general second-order linear differential operator acting on functions defined for $\alpha \leqslant x \leqslant \beta$ can be written in the form

$$\widetilde{\mathcal{L}} = -P(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} - R(x)\frac{\mathrm{d}}{\mathrm{d}x} - Q(x)\,, \tag{1.9a}$$

where P, Q and R are real functions; we shall assume that

$$P(x) > 0 \quad \text{for} \quad \alpha < x < \beta\,. \tag{1.9b}$$

Hence for $\mathcal{L}$ defined by (1.8)

$$\begin{aligned}
\mathcal{L} &= -wP\frac{\mathrm{d}^2}{\mathrm{d}x^2} - wR\frac{\mathrm{d}}{\mathrm{d}x} - wQ\\
&= -\frac{\mathrm{d}}{\mathrm{d}x}\left(wP\frac{\mathrm{d}}{\mathrm{d}x}\right) + \left(\frac{\mathrm{d}}{\mathrm{d}x}(wP) - wR\right)\frac{\mathrm{d}}{\mathrm{d}x} - wQ\,.
\end{aligned} \tag{1.10}$$

The operator $\mathcal{L}$ in (1.10) is of Sturm-Liouville form (1.6a) if we choose our *integrating factor* w so that

$$P\frac{\mathrm{d}w}{\mathrm{d}x} + \left(\frac{\mathrm{d}P}{\mathrm{d}x} - R\right) w = 0\,, \tag{1.11a}$$

and let

$$\rho = wP \quad \text{and} \quad q = wQ\,. \tag{1.11b}$$

On solving (1.11a), and on choosing the constant of integration so that $w(\alpha) = 1$, we obtain

$$w = \exp\int_\alpha^x \frac{1}{P(\zeta)}\left(R(\zeta) - \frac{\mathrm{d}P}{\mathrm{d}x}(\zeta)\right)\mathrm{d}\zeta\,. \tag{1.12}$$

Remark. It follows from (1.12) that $w > 0$, and hence from (1.9b) and (1.11b) that $\rho > 0$ for $\alpha < x < \beta$ (cf. (1.6b)).

Example. Bessel's operator is defined as

$$\widetilde{\mathcal{L}} = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x} + \lambda\,. \tag{1.13a}$$

Hence, from comparison with (1.9a), $P = -1$, $R = -\frac{1}{x}$ and $Q = -\lambda$, where we have relaxed the requirement that P is positive, but not the fact that P is single signed. From (1.12) we deduce the the weight function, w, should be (with a judicious choice of $\alpha = 1$),

$$w = \exp\int_1^x \frac{1}{\zeta}\,\mathrm{d}\zeta = x\,. \tag{1.13b}$$

As expected

$$\begin{aligned}\mathcal{L} = x\widetilde{\mathcal{L}} &= x\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{\mathrm{d}}{\mathrm{d}x} + x\lambda \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right) + x\lambda\,,\end{aligned} \tag{1.13c}$$

01/21 is of Sturm-Liouville form.

Is $\widetilde{\mathcal{L}}$ self-adjoint? We have seen that the general second-order linear differential operator $\widetilde{\mathcal{L}}$ can be transformed into Sturm-Liouville form by multiplication by a weight function w. It follows from § 1.4 that, subject to the boundary conditions (1.7b) being satisfied, $w\widetilde{\mathcal{L}} = \mathcal{L}$ is self-adjoint with respect to an inner product with the identity weight function, i.e.

$$\int_\alpha^\beta u^*\,(\mathcal{L}\,v)\,\mathrm{d}x = \int_\alpha^\beta (\mathcal{L}\,u)^*\,v\,\mathrm{d}x\,. \tag{1.14a}$$

However suppose that we slightly rearrange this equation to

$$\int_\alpha^\beta u^*\,(\widetilde{\mathcal{L}}\,v)\,w\,\mathrm{d}x = \int_\alpha^\beta (\widetilde{\mathcal{L}}\,u)^*\,v\,w\,\mathrm{d}x\,. \tag{1.14b}$$

Then from reference to the definition of an inner product with weight function w, i.e. (1.3a), we see that, subject to appropriate boundary conditions being satisfied, i.e.

$$\left[wP\left(v\frac{\mathrm{d}u^*}{\mathrm{d}x} - u^*\frac{\mathrm{d}v}{\mathrm{d}x}\right)\right]_\alpha^\beta = 0\,, \tag{1.14c}$$

01/20 $\widetilde{\mathcal{L}}$ is self-adjoint with respect to an inner product with the weight function w.

1.6 Eigenvalues and Eigenfunctions

The equation $\widetilde{\mathcal{L}}y = f$ is analogous to the matrix equation $\mathsf{M}\mathsf{x} = \mathsf{b}$. This analogy suggests that it might be profitable to consider the *eigenvalue equation*

$$\widetilde{\mathcal{L}}y = \lambda y\,, \tag{1.15}$$

where λ is the, possibly complex, *eigenvalue* associated with the *eigenfunction* $y \neq 0$.

Example. The Schrödinger equation for a one-dimensional quantum harmonic oscillator is

$$\left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \tfrac{1}{2}k^2x^2\right)\psi = E\psi\,.$$

This is an eigenvalue equation where the eigenvalue E is the energy level of the oscillator.

Remark. If $\widetilde{\mathcal{L}}$ is not in Sturm-Liouville form we can multiply by w to get the equivalent eigenvalue equation,

$$\mathcal{L}y = \lambda w y\,, \tag{1.16}$$

where $\mathcal{L}$ is in Sturm-Liouville form.

The Claim. We claim, but do not prove, that if the functions on which $\widetilde{\mathcal{L}}$ [or equivalently $\mathcal{L}$] acts are such that the boundary conditions (1.14c) [or equivalently (1.7b)] are satisfied, then it is generally the case that (1.15) [or equivalently (1.16)] has solutions only for a discrete, but infinite, set of values of λ:

$$\{\lambda_n, n = 1, 2, 3, \ldots\} \tag{1.17}$$

These are the eigenvalues of $\widetilde{\mathcal{L}}$ [or equivalently $\mathcal{L}$]. The corresponding solutions $\{y_n(x), n = 1, 2, 3, \ldots\}$ are the eigenfunctions.

Example: Harmonic Functions. Find the eigenvalues and eigenfunctions for the operator

$$\mathcal{L} = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,, \tag{1.18}$$

on the assumption that $\mathcal{L}$ acts on functions defined on $0 \leqslant x \leqslant \pi$ that vanish at end-points $x = 0$ and $x = \pi$.

Answer. $\mathcal{L}$ is in Sturm-Liouville form, i.e. (1.6a), with $\rho = 1$ and $q = 0$. Further, the boundary conditions ensure that (1.7b) is satisfied. Hence $\mathcal{L}$ is self-adjoint. The eigenvalue equation is

$$y'' + \lambda y = 0\,, \tag{1.19a}$$

with general solution

$$y = A\cos\lambda^{\frac{1}{2}}x + B\sin\lambda^{\frac{1}{2}}x\,. \tag{1.19b}$$

Non-zero solutions exist with $y(0) = y(\pi) = 0$ only if

$$y(0) = A = 0 \quad \text{and} \quad y(\pi) = B\sin\lambda^{\frac{1}{2}}\pi = 0\,. \tag{1.20}$$

For non-zero eigenfunctions we require $B \neq 0$. It then follows that $\lambda = n^2$ for integer n, and that the corresponding eigenfunctions are

$$y_n(x) = B\sin nx\,. \tag{1.21}$$

Remarks.

(i) The eigenvalues $\lambda_n = n^2$ are real (cf. the eigenvalues of an Hermitian matrix.).

(ii) It is conventional to normalize eigenfunctions to have unit norm. Recall from (1.4c) that the norm of $y_n(x)$ is given by (with $w = 1$)

$$\|y_n\|^2 = \int_\alpha^\beta |y_n|^2\,\mathrm{d}x\,. \tag{1.22a}$$

For our example (1.21), the unit-norm eigenfunctions are

$$y_n = \left(\frac{2}{\pi}\right)^{\frac{1}{2}}\sin nx\,. \tag{1.22b}$$

Unlectured example: Hermite polynomials. Hermite's equation,

$$y'' - 2xy' + 2ny = 0\,, \tag{1.23a}$$

arises when solving the harmonic oscillator in quantum mechanics. (1.23a) can be written as an eigenvalue equation, $\widetilde{\mathcal{L}}\,y = \lambda\,y$, with,

$$\widetilde{\mathcal{L}} = -\frac{d^2}{dx^2} + 2x\frac{d}{dx} = -e^{x^2}\frac{d}{dx}\left[e^{-x^2}\frac{d}{dx}\right], \qquad \lambda = 2n\,. \tag{1.23b}$$

$\widetilde{\mathcal{L}}$ is *not* of Sturm-Liouville form. However, multiplying by $w(x) = e^{-x^2}$ yields

$$\mathcal{L} = e^{-x^2}\widetilde{\mathcal{L}} = -\frac{d}{dx}\left[e^{-x^2}\frac{d}{dx}\right], \tag{1.23c}$$

which *is* of Sturm-Liouville form ($\rho = e^{-x^2}$, $q = 0$). Hence Hermite's equation is equivalent to

$$\mathcal{L}\,y = \lambda\,w(x)\,y\,. \tag{1.23d}$$

Remark. Hermite's equation can be viewed either as (1.23d), where $\mathcal{L}$ is self-adjoint with respect to a unit-weight inner product, *or equivalently* as

$$\widetilde{\mathcal{L}}\,y = \lambda\,y\,, \tag{1.23e}$$

where $\widetilde{\mathcal{L}}$ is self-adjoint with respect to a weight-w inner product.

You saw last term that if the boundary condition to be applied to solutions of (1.23a) is that the norm $\|y\|_w$ is finite, where

$$\|y\|_w^2 = \int_{-\infty}^{\infty} e^{-x^2}\,|y|^2\,dx\,, \tag{1.23f}$$

then non-zero solutions only exist when n is a non-negative integer; hence the eigenvalues, $\lambda = 2n$, are the even non-negative integers. The eigenfunctions are the nth order *Hermite polynomials*. In the context of a quantum harmonic oscillator, the wavefunction is $\psi(x) = y(x)e^{-x^2/2}$, and so it makes sense physically to require the norm of ψ to be finite.

1.6.1 Eigenvalues of a self-adjoint operator are real

Let $\widetilde{\mathcal{L}}$ be a self-adjoint operator with respect to an inner product (1.3a) with weight w, and suppose that y is a non-zero eigenfunction with eigenvalue λ satisfying

$$\widetilde{\mathcal{L}}y = \lambda y\,. \tag{1.24a}$$

Take the complex conjugate of this equation, remembering that $\widetilde{\mathcal{L}}$ is real, to obtain

$$\widetilde{\mathcal{L}}y^* = \lambda^* y^*\,. \tag{1.24b}$$

Consider

$$\begin{aligned}
\langle\, y \,|\, \widetilde{\mathcal{L}}y \,\rangle - \langle\, \widetilde{\mathcal{L}}y \,|\, y \,\rangle &= \int_\alpha^\beta \left(y^*\,\widetilde{\mathcal{L}}y - y\,\widetilde{\mathcal{L}}y^*\right) w\,\mathrm{d}x && \text{from (1.3a) and } \widetilde{\mathcal{L}} \text{ real}\\
&= \int_\alpha^\beta \left(y^*\lambda y - y\lambda^* y^*\right) w\,\mathrm{d}x && \text{from (1.24a) and (1.24b)}\\
&= (\lambda - \lambda^*) \int_\alpha^\beta |y|^2 w\,\mathrm{d}x\,.\\
&= (\lambda - \lambda^*)\|y\|_w^2\,.
\end{aligned} \tag{1.25}$$

But $\widetilde{\mathcal{L}}$ is self-adjoint with respect to an inner product with weight w, and hence the left hand side of (1.25) is zero (see (1.5b) with $u = v = y$). It follows that

$$(\lambda - \lambda^*)\|y\|_w^2 = 0\,, \tag{1.26}$$

But $\|y\|_w^2 > 0$ from (1.4d) since y has been assumed to be a non-zero eigenfunction. Hence

$$\lambda = \lambda^*\,, \quad \text{i.e. } \lambda \text{ is real.} \tag{1.27}$$

Remark. As for the eigenvalues of Hermitian matrices, this result can also be obtained, arguably in a more elegant fashion, using inner product notation:

$$\begin{aligned}
\lambda\langle\, y \mid y\,\rangle_w &= \langle\, y \mid \lambda y\,\rangle_w && \text{from (1.4b)}\\
&= \langle\, y \,|\, \widetilde{\mathcal{L}} y\,\rangle_w && \text{from (1.15)}\\
&= \langle\, \widetilde{\mathcal{L}} y \,|\, y\,\rangle_w && \text{since } \widetilde{\mathcal{L}} \text{ is self-adjoint wrt weight } w\\
&= \langle\, \lambda y \mid y\,\rangle_w && \text{from (1.15)}\\
&= \lambda^*\langle\, y \mid y\,\rangle_w && \text{from (1.4a) and (1.4b).}
\end{aligned} \tag{1.28}$$

This is essentially (1.26), and hence (1.27) follows as above.

1.6.2 Eigenfunctions of a self-adjoint operator with distinct eigenvalues are orthogonal

Definition. Two functions u and v are said to be *orthogonal* with respect to a given inner product, if

$$\langle\, u \mid v\,\rangle_w = 0\,. \tag{1.29}$$

As before let $\widetilde{\mathcal{L}}$ be a general second-order linear differential operator that is self-adjoint with respect to an inner product with weight w. Suppose that y_1 and y_2 are eigenfunctions of $\widetilde{\mathcal{L}}$, with *distinct* eigenvalues λ_1 and λ_2 respectively. Then from the definition (1.15)

$$\widetilde{\mathcal{L}} y_1 = \lambda_1 y_1\,, \tag{1.30a}$$

$$\widetilde{\mathcal{L}} y_2 = \lambda_2 y_2\,. \tag{1.30b}$$

This time use inner product notation; then, by analogy with the proof for eigenvectors for Hermitian matrices:

$$\begin{aligned}
\lambda_2\langle\, y_1 \mid y_2\,\rangle_w &= \langle\, y_1 \mid \lambda_2 y_2\,\rangle_w && \text{from (1.4b)}\\
&= \langle\, y_1 \,|\, \widetilde{\mathcal{L}} y_2\,\rangle_w && \text{from (1.30b)}\\
&= \langle\, \widetilde{\mathcal{L}} y_1 \,|\, y_2\,\rangle_w && \text{since } \widetilde{\mathcal{L}} \text{ is self-adjoint wrt weight } w\\
&= \langle\, \lambda_1 y_1 \mid y_2\,\rangle_w && \text{from (1.30a)}\\
&= \lambda_1^*\langle\, y_1 \mid y_2\,\rangle_w && \text{from (1.4a) and (1.4b)}\\
&= \lambda_1\langle\, y_1 \mid y_2\,\rangle_w && \text{from (1.27).}
\end{aligned} \tag{1.31a}$$

It follows that

$$(\lambda_2 - \lambda_1)\,\langle\, y_1 \mid y_2\,\rangle_w = 0\,. \tag{1.31b}$$

Hence if $\lambda_1 \neq \lambda_2$ then the eigenfunctions are orthogonal:

$$\langle\, y_1 \mid y_2\,\rangle_w = 0\,. \tag{1.31c}$$

Unlectured remark. Alternatively the same result can be obtained using the defintion of the inner product in terms of integrals. First, from taking the complex conjugate of (1.30a), we have that

$$\widetilde{\mathcal{L}} y_1^* = \lambda_1 y_1^*\,, \tag{1.32a}$$

since $\widetilde{\mathcal{L}}$ and λ_1 are real. Hence

$$\begin{aligned}
\int_\alpha^\beta \left(y_1^*\, \widetilde{\mathcal{L}} y_2 - y_2\, \widetilde{\mathcal{L}} y_1^*\right) w\,\mathrm{d}x &= \int_\alpha^\beta (y_1^*\, \lambda_2 y_2 - y_2\, \lambda_1 y_1^*)\, w\,\mathrm{d}x \qquad \text{from (1.30b) and (1.32a)}\\
&= (\lambda_2 - \lambda_1) \int_\alpha^\beta y_1^* y_2\, w\,\mathrm{d}x\\
&= (\lambda_2 - \lambda_1)\,\langle\, y_1 \,|\, y_2\,\rangle_w\,.
\end{aligned} \tag{1.32b}$$

But $\widetilde{\mathcal{L}}$ is self-adjoint, and hence the left hand side of (1.32b) is zero (e.g. (1.14b) with $u = y_1$ and $v = y_2$). It follows that, as in (1.31b),

$$(\lambda_2 - \lambda_1)\,\langle\, y_1 \mid y_2\,\rangle_w = 0\,. \tag{1.32c}$$

Hence if $\lambda_1 \neq \lambda_2$ then the eigenfunctions are orthogonal.

Orthonormal set. We have seen that eigenfunctions with different eigenvalues are mutually orthogonal. We claim, but do not prove, that mutually orthogonal eigenfunctions can always be constructed, even for repeated eigenvalues. Further, if we normalize all eigenfunctions to have unit norm then we have an *orthonormal* set of eigenfunctions, i.e.

$$\langle\, y_n \mid y_m \,\rangle_w = \int_\alpha^\beta w\, y_n^* y_m \,\mathrm{d}x = \delta_{mn}\,. \tag{1.33}$$

Example: Harmonic Functions. Return to the earlier example for the Sturm-Liouville operator (1.18), i.e.

$$\mathcal{L} = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,, \tag{1.34a}$$

acting on functions that vanish at end-points $x = 0$ and $x = \pi$. We note that

(i) the *eigenvalues* are n^2, and hence real (cf. (1.27));

(ii) the unit norm *eigenfunctions* (1.22b), i.e.

$$y_n = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \sin nx\,, \tag{1.34b}$$

satisfy

$$\begin{aligned}\int_0^\pi y_n^* y_m \,\mathrm{d}x &= \frac{2}{\pi}\int_0^\pi \sin nx \sin mx \,\mathrm{d}x \\ &= \frac{1}{\pi}\int_0^\pi \big(\cos(n-m)x - \cos(n+m)x\big)\,\mathrm{d}x \\ &= 0 \quad \text{if } n \neq m;\end{aligned} \tag{1.34c}$$

and are thus orthogonal (cf. (1.33)).

Example: Legendre Polynomials. Legendre's equation,

$$(1-x^2)y'' - 2xy' + \ell(\ell+1)y = 0\,, \tag{1.35a}$$

arises, *inter alia*, in solutions of Laplace's equation with axial symmetry, and Schrödinger's equation in three dimensions with a central potential. Equation (1.35a) can be written in Sturm-Liouville eigenvalue form, $\mathcal{L}y = \lambda y$, with

$$\mathcal{L} = -\frac{d}{dx}\left[\left(1-x^2\right)\frac{d}{dx}\right] \quad \text{and} \quad \lambda = \ell\,(\ell+1)\,, \tag{1.35b}$$

where, in terms of our standard notation,

$$\rho = 1 - x^2 \quad \text{and} \quad q = 0\,. \tag{1.35c}$$

In question 3 on Example Sheet 1 you will show that this operator is self-adjoint when acting on functions $y(x)$ that satisfy the boundary conditions that they are *finite* at $x = \pm 1$.

Last term, using series solutions, you found that the only non-zero solutions for which y is finite at $x = \pm 1$ are polynomials (the series terminates), and that this only happens if ℓ is an integer. Further, since negative integers just repeat the set of possible solutions, we may restrict to $\ell \geqslant 0$. These polynomials, are known as the *Legendre polynomials*, $P_\ell(x)$. Hence we can identify the eigenvalues and eigenfunctions as

$$\lambda_\ell = \ell(\ell+1) \quad \text{and} \quad y_\ell(x) = P_\ell(x) \quad \text{for} \quad \ell = 0, 1, 2, \ldots\,. \tag{1.35d}$$

Remark. In the context of the Schrödinger equation applied to the hydrogen atom, ℓ is the (quantized) orbital angular momentum of the electron, and the lowest energy state has $\ell = 0$.

A conventional normalization is $P_\ell(1) = 1$, in which case the first few polynomials are

$$P_0 = 1\,, \qquad P_1 = x\,, \qquad P_2 = \tfrac{1}{2}\left(3x^2 - 1\right)\,, \qquad P_3 = \tfrac{1}{2}\left(5x^3 - 3x\right)\,. \tag{1.35e}$$

It follows from our general theory that the Legendre polynomials are orthogonal (see also question 3 on Example Sheet 1). However, with the conventional normalisation they are *not* orthonormal; in fact

$$\int_{-1}^1 P_\ell(x) P_k(x) dx = \frac{2}{2\ell+1}\,\delta_{\ell k}\,. \tag{1.35f}$$

Unlectured remark. As a check on (1.35f) we note that if ℓ is odd and k is even, then $P_\ell P_k$ is an odd polynomial, and hence a symmetric integral like (1.35f) must be zero.

02/19
02/20

1.7 Eigenfunction Expansions

Let $\{y_n, n = 1, 2, \dots\}$ be an *orthonormal* set of eigenfunctions of a self-adjoint operator. Then we claim that any function $f(x)$ with the same boundary conditions as the eigenfunctions can be expressed as an eigenfunction expansion

$$f(x) = \sum_{n=1}^{\infty} a_n y_n(x)\,, \tag{1.36a}$$

where the coefficients a_n are given by

$$a_n = \langle\, y_n \mid f\, \rangle_w\,, \tag{1.36b}$$

i.e. we claim that the eigenfunctions form a basis. A set of eigenfunctions that has this property is said to be complete.

Consistency. We will not prove the existence of the expansion (1.36a). However, if we assume such an expansion does exist, the coefficients must be given by (1.36b) because

$$\begin{aligned}
\langle\, y_n \mid f\, \rangle_w &= \langle\, y_n \mid \textstyle\sum_{m=1}^{\infty} a_m y_m\, \rangle_w && \text{from (1.36a)}\\
&= \sum_{m=1}^{\infty} a_m \langle\, y_n \mid y_m\, \rangle_w && \text{from inner product property (1.2c)}\\
&= \sum_{m=1}^{\infty} a_m \delta_{nm} && \text{from (1.33) since the } y_n \text{ are orthonormal}\\
&= a_n && \text{as required.}
\end{aligned}$$

The completeness relation. Further, it follows from (1.36a) and (1.36b) that

$$\begin{aligned}
f(x) &= \sum_{n=1}^{\infty} \langle\, y_n \mid f\, \rangle_w\, y_n(x)\\
&= \sum_{n=1}^{\infty} y_n(x) \int_\alpha^\beta w(\zeta)\, y_n^*(\zeta) f(\zeta)\,\mathrm{d}\zeta && \text{from the definition of the inner product (1.3a)}\\
&= \int_\alpha^\beta f(\zeta)\left(w(\zeta) \sum_{n=1}^{\infty} y_n(x) y_n^*(\zeta)\right)\mathrm{d}\zeta && \text{interchange sum and integral.}
\end{aligned} \tag{1.37}$$

If this expression holds for all functions f satisfying the appropriate homogeneous boundary conditions, then from the defintion of the delta function

$$w(\zeta) \sum_{n=1}^{\infty} y_n(x) y_n^*(\zeta) = \delta(x-\zeta)\,. \tag{1.38a}$$

This is the completeness relation.

Remark. If x and ζ are exchanged in the complex conjugate of (1.38a), then since the weight function is real and the delta function is real and symmetric, it follows that

$$w(x) \sum_{n=1}^{\infty} y_n(x) y_n^*(\zeta) = \delta(x-\zeta)\,. \tag{1.38b}$$

Example: Fourier series. Again consider the Sturm-Liouville operator

$$\mathcal{L} = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,. \tag{1.39a}$$

In this case assume that the operator acts on functions that are 2π-periodic, i.e. $y(x) = y(x+2\pi)$. $\mathcal{L}$ is still self-adjoint with weight function $w = 1$ since the periodicity ensures that the boundary conditions

(1.7b) are satisfied if, say, $\alpha = 0$ and $\beta = 2\pi$. This time we choose to write the general solution of the eigenvalue equation (1.16), $\mathcal{L}y = \lambda y$, as

$$y = Ae^{\imath\lambda^{\frac{1}{2}}x} + Be^{-\imath\lambda^{\frac{1}{2}}x}\,, \tag{1.39b}$$

where A and B are constants. This solution is 2π-periodic if $\lambda = n^2$ for integer n (as before). Label the eigenfunctions by y_n for $n = \dots, -1, 0, 1, \dots$, with corresponding eigenvalues $\lambda_n = n^2$. Although there are repeated eigenvalues (there are two eigenfunctions for each eigenvalue except for $n = 0$), there still exists an orthonormal set of eigenfunctions (as claimed), i.e.

$$y_n = \frac{1}{\sqrt{2\pi}} e^{\imath nx} \quad \text{for } n \in \mathbb{Z}. \tag{1.39c}$$

Hence, from (1.36a), a 2π-periodic function f has an eigenfunction expansion

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{\infty} a_n e^{\imath nx}\,, \tag{1.39d}$$

where the a_n are given by (1.36b). Thus the Fourier series representation of f is an example of an expansion in terms of the eigenfunctions of a self-adjoint operator. The completeness relation (1.38a) is

$$\frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{\imath n(x-\zeta)} = \delta(x-\zeta)\,. \tag{1.39e}$$

1.8 Solution of Differential Equations

Finally we consider how eigenfunction expansions can be used to solve differential equations of the form

$$\mathcal{L}\, y(x) = f(x)\,, \tag{1.40}$$

for some *forcing function* $f(x)$ and homogeneous boundary conditions. We assume that

(i) the equation has been rewritten so that $\mathcal{L}$ is in Sturm-Liouville form;

(ii) $\mathcal{L}$ has a complete orthonormal set of eigenfunctions $\{y_n\}$ with eigenvalues $\{\lambda_n\}$ such that

$$\mathcal{L}y_n = \lambda_n w y_n\,, \qquad \langle y_n | y_m \rangle_w = \delta_{nm}\,, \qquad n, m = 1, 2, 3, \dots\,. \tag{1.41}$$

In order to solve (1.40) we introduce the *Green's function*, $G(x, \zeta)$, that

(i) is the response to a point-like source at $x = \zeta$, i.e. satisfies

$$\mathcal{L}_x\, G(x, \zeta) = \delta(x - \zeta)\,, \tag{1.42a}$$

where the subscript x on $\mathcal{L}$ is to emphasise that $\mathcal{L}$ contains x, $\frac{d}{dx}$ etc.;

(ii) is required to satisfy the boundary conditions when considered *both* as a function of x and as a function of ζ.

Given G, the formal solution to (1.40) is then

$$y(x) = \int_\alpha^\beta G(x, \zeta) f(\zeta) d\zeta\,, \tag{1.42b}$$

since

$$\begin{aligned} \mathcal{L}\, y(x) &= \mathcal{L} \int_\alpha^\beta G(x, \zeta) f(\zeta) d\zeta && \text{from (1.42b)} \\ &= \int_\alpha^\beta \delta(x - \zeta) f(\zeta) d\zeta && \text{from (1.42a)} \\ &= f(x)\,. \end{aligned} \tag{1.42c}$$

One method to find the Green's function is to observe that, provided none of the λ_n vanish, the function

$$G(x,\zeta) = \sum_{n=1}^{\infty} \frac{1}{\lambda_n}\, y_n(x) y_n^*(\zeta)\,, \tag{1.43a}$$

satisfies the required boundary conditions, and

$$\begin{aligned}
\mathcal{L}_x\, G(x,\zeta) &= \sum_{n=1}^{\infty} \frac{y_n^*(\zeta)}{\lambda_n}\, \mathcal{L}_x y_n(x) && \text{from (1.43a)}\\
&= \sum_{n=1}^{\infty} w(x) y_n^*(\zeta)\, y_n(x) && \text{from (1.41)}\\
&= \frac{w(x)}{w(\zeta)}\, \delta(x-\zeta) = \delta(x-\zeta) && \text{from (1.38b).}
\end{aligned} \tag{1.43b}$$

Remark. From the form of the Green's function (1.43a), we deduce that

$$G(x,\zeta) = G^*(\zeta, x)\,. \tag{1.43c}$$

The solution as an eigenfunction expansion. Substitute (1.43a) into (1.42b) to obtain

$$\begin{aligned}
y(x) &= \int_\alpha^\beta \sum_{n=1}^{\infty} \frac{1}{\lambda_n}\, y_n(x) y_n^*(\zeta) f(\zeta) d\zeta\\
&= \sum_{n=1}^{\infty} \frac{1}{\lambda_n}\, y_n(x) \int_\alpha^\beta y_n^*(\zeta) f(\zeta) d\zeta\\
&= \sum_{n=1}^{\infty} \frac{1}{\lambda_n}\, y_n(x) \langle\, y_n \mid f\,\rangle && \text{(recall that } w=1\text{).}
\end{aligned} \tag{1.44a}$$

In summary,

$$y(x) = \sum_{n=1}^{\infty} b_n\, y_n(x)\,, \tag{1.44b}$$

where

$$b_n = \frac{a_n}{\lambda_n}\,, \quad \text{and} \quad a_n = \langle\, y_n \mid f\,\rangle = \int_\alpha^\beta y_n^*(\zeta) f(\zeta) d\zeta\,. \tag{1.44c}$$

Resonance. If $\lambda_n = 0$ for some n, say $\lambda_1 = 0$, then $G(x,\zeta)$ does not exist and there is no [finite] solution for y for general f (see (1.44a)); in other words, there is no solution to the forced problem if there is a solution to the homogeneous equation, $\mathcal{L}y = 0$, satisfying the boundary conditions ($y_1(x)$ is precisely such a solution). The vanishing of one or more of the eigenvalues is related to the phenomenon of *resonance*. so that if a solution to the problem (including the boundary conditions) exists in the *absence* of the forcing f (i.e. if there exists a zero eigenvalue of $\mathcal{L}$), then any non-zero force elicits an infinite response.

If, instead, one of the eigenvalues, again say λ_1, is very small compared to the others, then from (1.44a), and assuming that $\langle y_1 | f\rangle$ is not too small,

$$y(x) \approx \frac{y_1(x)}{\lambda_1} \langle y_1 | f\rangle\,.$$

Hence, any forcing function with non-zero y_1-component causes a large 'resonant' response proportional to $y_1(x)$.

Example. Find the solution of $\mathcal{L}y = \sin^3 x$ subject to the boundary conditions $y(0) = y(\pi) = 0$, where $\mathcal{L} = -\frac{d^2}{dx^2} + 2$.

The unit-norm eigenfunctions of $\mathcal{L}$, which is of Sturm-Liouville form, are

$$y_n(x) = \sqrt{\frac{2}{\pi}}\sin(nx) \qquad \text{where} \qquad n = 1,2,3,\ldots\,,$$

with eigenvalues, $\lambda_n = n^2 + 2$. We can write $f(x) = \sin^3 x = \frac{3}{4}\sin x - \frac{1}{4}\sin 3x$ and so,

$$f(x) = \sum_{n=1}^{\infty} a_n\, y_n(x)\,, \quad \text{with} \quad a_1 = \tfrac{3}{4}\sqrt{\frac{\pi}{2}}\,, \quad a_3 = -\tfrac{1}{4}\sqrt{\frac{\pi}{2}} \quad \text{and the other } a_n \text{ zero.}$$

Writing $y(x) = \sum_n b_n y_n(x)$, and using $\mathcal{L} y_n = \lambda_n y_n$, gives

$$\mathcal{L} y = \sum_n \lambda_n b_n y_n(x) = f(x) = \sum_n a_n y_n(x) \quad \Rightarrow \quad b_n = a_n/\lambda_n\,,$$

as in (1.44c). It follows that

$$y(x) = \frac{3}{4\lambda_1}\sin x - \frac{1}{4\lambda_3}\sin 3x = \tfrac{1}{4}\sin x - \tfrac{1}{44}\sin 3x\,.$$

Remarks.

(i) Alternatively, we could obtain the same result using the general solution (1.42b), together with the Green's function from the general expression (1.43a):

$$G(x,\zeta) = \frac{2}{\pi}\sum_{n=1}^{\infty} \frac{\sin nx\, \sin n\zeta}{n^2+2}\,.$$

(ii) In this example we were able to write the solution in terms of a finite number of eigenfunctions, but in general we'll require a sum over an infinite number of eigenfunctions.

Example. Solve the problem specified by (1.1b) and (1.1c), i.e.

$$\widetilde{\mathcal{L}} y \equiv -\frac{d^2y}{dx^2} - \frac{dy}{dx} - \tfrac{1}{4}y = e^{-x/2}\,, \tag{1.1b}$$

with

$$y = 0 \quad \text{at} \quad x = 0\,, \quad \text{and} \quad \frac{dy}{dx} + \tfrac{1}{2}y = 0 \quad \text{at} \quad x = 1. \tag{1.1c}$$

(i) *First put the differential operator $\widetilde{\mathcal{L}} = -\frac{d^2}{dx^2} - \frac{d}{dx} - \frac{1}{4}$ in Sturm-Liouville form.*

A suitable weight function for $\widetilde{\mathcal{L}}$ is, from (1.9a) with with $P = R = 1$, and from using (1.12),

$$w(x) = e^{\int^x dx} = e^x\,.$$

Therefore the Sturm-Liouville form of (1.1b) is, from (1.8),

$$\mathcal{L} y = w(x)\,\widetilde{\mathcal{L}} y = -\frac{d}{dx}\left(e^x \frac{dy}{dx}\right) - \tfrac{1}{4}e^x\, y = e^{x/2}\,.$$

03/19

(ii) *Next determine the eigenvalues and orthonormal eigenfunctions for $0 \leqslant x \leqslant 1$ with boundary conditions* (1.1c)*, i.e.* $y(0) = 0$ *and* $\frac{dy}{dx} + \frac{1}{2}y = 0$ *at* $x = 1$.

We need to solve $\widetilde{\mathcal{L}} y_n = \lambda_n y_n$ or equivalently $\mathcal{L} y_n = \lambda_n w y_n$. Using the first form and trying $y = e^{kx}$ we deduce that

$$\left(-k^2 - k - \tfrac{1}{4}\right) e^{kx} = \lambda e^{kx}\,,$$

and hence that

$$k = -\tfrac{1}{2} \pm i\sqrt{\lambda}\,.$$

Therefore

$$y(x) = Ae^{-x/2}\cos\sqrt{\lambda}x + Be^{-x/2}\sin\sqrt{\lambda}x\,,$$

where A and B are constants. Applying the boundary condition (1.1c) at $x = 0$, i.e. $y(0) = 0$, yields $A = 0$, while applying the boundary condition at $x = 1$ yields, after cancellation,

$$B\sqrt{\lambda}\,e^{-1/2}\cos\sqrt{\lambda} = 0\,.$$

Hence, for a non-zero eigenfunction, we require that

$$\sqrt{\lambda} = \left(n + \tfrac{1}{2}\right)\pi \quad \text{for} \quad n = 0, 1, 2, 3, \ldots,$$

and so

$$\lambda_n = \left(n + \tfrac{1}{2}\right)^2 \pi^2\,, \quad y_n(x) = B_n\, e^{-x/2} \sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right] \quad \text{for} \quad n = 0, 1, 2, 3, \ldots.$$

It is left as an exercise to check that, with these eigenfunctions, the boundary terms (1.7b) vanish so that $\mathcal{L}$ is self-adjoint. For an orthonormal basis we require that

$$\langle y_n | y_m \rangle_w = \int_0^1 y_n^*(x)\, y_m(x)\, e^x\, dx = \delta_{nm}\,.$$

For the $y_n(x)$ to have unit norm we therefore require that

$$|B_n|^2 \int_0^1 \sin^2\left[\left(n + \tfrac{1}{2}\right)\pi x\right]\, dx = 1$$

An orthonormal basis is thus

$$y_n(x) = \sqrt{2}\, e^{-x/2} \sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right]\,, \qquad n = 0, 1, 2, 3, \ldots.$$

03/20
03/21

(iii) *Hence solve $\widetilde{\mathcal{L}} y = e^{-x/2}$, subject to the above boundary conditions.*

Our aim is to solve, in terms of the notation of (1.40),

$$\mathcal{L} y = w\, \widetilde{\mathcal{L}} y = e^{x/2} \equiv f\,,$$

using the orthonormal eigenfunctions, y_n, that satisfy $\mathcal{L} y_n = \lambda_n e^x y_n$. Hence, writing $y = \sum_n b_n y_n$ as in (1.44b), we require that

$$\mathcal{L}\, y = \sum_n b_n\, \lambda_n\, e^x\, y_n = e^{x/2} \equiv f\,.$$

Forming an inner product by multiplying by $y_m^*(x)$ and integrating, cf. (1.36b), we obtain

$$\sum_n \lambda_n\, b_n \int_0^1 y_m^*(x)\, e^x\, y_n(x)\, dx = \sum_n \lambda_n\, b_n\, \delta_{mn} = \int_0^1 e^{x/2}\, y_m^*(x)\, dx \equiv \langle\, y_m \mid f\, \rangle\,,$$

from which we deduce that, cf. (1.44c),

$$\begin{aligned} b_m = \frac{\langle\, y_m \mid f\, \rangle}{\lambda_m} &= \frac{\sqrt{2}}{\lambda_m} \int_0^1 \sin\left[\left(m + \tfrac{1}{2}\right)\pi x\right] dx\,, \\ &= -\frac{\sqrt{2}}{\lambda_m} \frac{1}{\left(m + \frac{1}{2}\right)\pi} \Big[\cos\left[\left(m + \tfrac{1}{2}\right)\pi x\right]\Big]_0^1 = \frac{\sqrt{2}}{\lambda_m} \frac{1}{\left(m + \frac{1}{2}\right)\pi}\,. \end{aligned}$$

Hence

$$y(x) = \frac{2}{\pi^3} \sum_{n=0}^{\infty} \frac{1}{\left(n + \frac{1}{2}\right)^3}\, e^{-x/2}\, \sin\left[\left(n + \tfrac{1}{2}\right)\pi x\right]\,.$$

Remark. There is a better way to solve this problem: by inspection

$$y = x e^{-x/2} - \tfrac{1}{2} x^2 e^{-x/2}\,.$$

1.9 Approximation via Eigenfunction Expansions

It is often useful, e.g. in a numerical method (since a computer cannot store an infinite number of terms), to approximate a function with Sturm-Liouville boundary conditions by a *finite* linear combination of Sturm-Liouville eigenfunctions, i.e.

$$f(x) \approx \sum_{n=1}^{N} a_n y_n(x)\,. \tag{1.45}$$

The question then arises as to how to optimize the choice of the $\{a_n\}$; should one continue to use the a_n for the infinite series as given by (1.36b)?

To this end, define the error of the approximation, i.e. the "distance" between f and the approximation, to be

$$\Sigma_N(a_1, a_2, \ldots, a_N) = \left\| f(x) - \sum_{n=1}^{N} a_n y_n(x) \right\|_w^2 . \tag{1.46}$$

Then, one definition of the 'best' approximation is that the error (1.46) should be minimized with respect to the coefficients $a_1, a_2, \ldots, a_N$. By expanding (1.46) we have, assuming that the y_n are an orthonormal set,

$$\begin{aligned}
\Sigma_N &= \left\langle f(x) - \sum_{n=1}^{N} a_n y_n(x) \;\middle|\; f(x) - \sum_{m=1}^{N} a_m y_m(x) \right\rangle_w \\
&= \langle\, f \mid f \,\rangle_w - \sum_{n=1}^{N} a_n^* \langle\, y_n \mid f \,\rangle_w - \sum_{m=1}^{N} a_m \langle\, f \mid y_m \,\rangle_w + \sum_{n=1}^{N}\sum_{m=1}^{N} a_n^* a_m \langle\, y_n \mid y_m \,\rangle_w \\
&= \|f\|_w^2 - \sum_{n=1}^{N} \Big(a_n^* \langle\, y_n \mid f \,\rangle_w + a_n \langle\, y_n \mid f \,\rangle_w^* \Big) + \sum_{n=1}^{N} a_n a_n^* .
\end{aligned} \tag{1.47}$$

To find the minimum, differentiate with respect to the a_k and a_k^* (viewed as independent variables) to obtain

$$\frac{\partial \Sigma_N}{\partial a_k} = -\langle y_k | f \rangle_w^* + a_k^* , \quad \text{and its complex conjugate} \quad \frac{\partial \Sigma_N}{\partial a_k^*} = -\langle y_k | f \rangle_w + a_k . \tag{1.48}$$

Thus the error Σ_N is minimized when

$$a_k = \langle\, y_k \mid f \,\rangle_w , \quad \text{or equivalently} \quad a_k^* = \langle\, y_k \mid f \,\rangle_w^* . \tag{1.49}$$

This is the same value for a_k as (1.36b), i.e. when the expansion (1.45) has an infinite number of terms. The value of the error, Σ_N, is then, from (1.47) and (1.49),

$$\Sigma_N = \|f\|_w^2 - \sum_{n=1}^{N} |a_n|^2 . \tag{1.50}$$

Since $\Sigma_N \geqslant 0$ from (1.46), we arrive at *Bessel's inequality*

$$\|f\|_w^2 \geqslant \sum_{n=1}^{N} |a_n|^2 . \tag{1.51}$$

It is possible to show, but not here, that this inequality becomes an equality when $N \to \infty$, and hence

$$\|f\|_w^2 = \sum_{n=1}^{\infty} |a_n|^2 \tag{1.52}$$

which is a generalization of *Parseval's theorem*.

Unlectured remark. While it is not strictly true that any function satisfying the Sturm-Liouville boundary conditions can be expressed as an eigenfunction expansion (1.36a) (since there are restrictions such as continuity), it is true that $\Sigma_\infty = 0$ for such functions, i.e.

$$\left\| f(x) - \sum_{n=1}^{\infty} \langle\, y_n \mid f \,\rangle_w \, y_n(x) \right\|_w^2 = 0 . \tag{1.53}$$

2 Calculus of Variations[4]

2.1 Functionals

A real function of many variables $\{y_k; k = 1, 2, \dots N\}$ maps the $\{y_k\}$ to a real number, i.e.

$$f : \{y_k\} \to f(\{y_k\}) \in \mathbb{R}\,.$$

Our aim is to generalise this idea to a continuous infinity of variables.

For instance, consider the following definite integral involving a real function $y(x)$

$$F[y] = \int_\alpha^\beta \left(y'(x)^2 - y(x)^2\right) dx\,. \tag{2.1}$$

This is a simple example of a *functional*, where a continuous infinity of variables $\{y(x); x \in \mathbb{R}\}$ is mapped to a real number, F, that is independent of x but depends on $y(x)$.

Definition. A *real functional* takes a function $y(x)$ and yields a real number as output

$$F : y(x) \to F[y] \in \mathbb{R}\,. \tag{2.2}$$

Remarks.

(i) In the particular example (2.1), the integrand of $F[y]$ *implicitly* depends on x through y and its derivatives; however, the integrand may also *explicitly* depend on x:

$$F[y] = \int_\alpha^\beta f(y, y', y'', \dots; x)\, dx\,. \tag{2.3a}$$

(ii) More generally, there may be more dependent variables $\{y_i\}$ and a multiple integral over a number of independent variables $\{x_j\}$.

(iii) We shall usually be concerned with functionals of the form,

$$F[y] = \int_\alpha^\beta f(y, y'; x)\, dx\,. \tag{2.3b}$$

The calculus of variations. The calculus of variations extends the calculus of functions to functionals. It aims to answer questions such as: what functions $y(x)$ extremize the functional $F[y]$?

Remarks

(i) It will usually be obvious from the problem whether a given extremum is a maximum, a minimum or something else – there's no equivalent of the Hessian criteria for functions (or at least one that's practical to use).

(ii) We must also keep in mind that, as with ordinary calculus, an extremum we find may be only a *local* extremum and not a *global* extremum.

(iii) Functionals are useful because many problems can be formulated as a *variational principle*, the extremization of some functional. For instance, a chain suspended between two fixed points hangs in equilibrium such that its total potential energy is minimized; an extension of this idea (incorporating chemical potential energy) can be applied to chemical reactions. Two well-known examples that we will examine are

- *Fermat's principle* in optics,
- *Hamilton's principle of least action* in mechanics.

[4] You are warned: this section is probably conceptually harder than other parts of the course.

2.2 Functional Derivatives

Consider the effect of changing a function $y(x)$ to a nearby function $y(x) + \delta y(x)$, assuming that the endpoints at A and B are fixed.

The *variation of* F, say as given by (2.3b), is defined by,

$$\begin{aligned}
\delta F &= F\,[y+\delta y] - F[y] && && \text{(2.4a)}\\
&= \int_\alpha^\beta f\left(y+\delta y, y' + (\delta y)'; x\right) dx - \int_\alpha^\beta f(y, y'; x) dx && \text{from (2.3b)} &&\\
&= \int_\alpha^\beta \left[\delta y \frac{\partial f}{\partial y} + (\delta y)'\,\frac{\partial f}{\partial y'}\right] dx \;+\; \dots && \text{cf. Taylor's Theorem} &&\\
&= \left[\delta y \frac{\partial f}{\partial y'}\right]_\alpha^\beta + \int_\alpha^\beta \delta y \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] dx \;+\; \dots && \text{by integrating by parts,} && \text{(2.4b)}
\end{aligned}$$

where we have omitted terms of order $(\delta y)^2$. If the boundary term in (2.4b) is zero (e.g. if, as above, y is fixed on the boundaries), equation (2.4b) can be written as

$$\delta F = \int_\alpha^\beta \delta y \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right] dx \;+\; \dots\,, . \tag{2.4c}$$

Definition. Define the *functional derivative* of F at the point x with respect to the function y to be the function

$$\frac{\delta F}{\delta y(x)} \qquad \text{as defined by the expression} \qquad \delta F = \int_\alpha^\beta \delta y(x) \frac{\delta F}{\delta y(x)}\, dx \;+\; \dots\,. \tag{2.5a}$$

In the case of F given by (2.3b), it follows from (2.4c) that

$$\frac{\delta F}{\delta y(x)} = \left.\frac{\partial f}{\partial y}\right|_{y',x} - \frac{d}{dx}\left(\left.\frac{\partial f}{\partial y'}\right|_{y,x}\right). \tag{2.5b}$$

Remarks.

(i) The value of the functional derivative depends on the point x.

(ii) The notation $\left.\frac{\partial f}{\partial y'}\right|_{y,x}$ may look strange since it seems impossible for y' to change if y does not. However, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial y'}$ are *formal* derivatives in which it is supposed that y and y' are unconnected. By contrast, $\frac{d}{dx}$ in (2.5b) is the usual full derivative with respect to x.

(iii) Compare with the variation of a function:

$$f(\{y_i\}) : \delta f = \sum_i \delta y_i \frac{\partial f}{\partial y_i}.$$

2.2.1 The Euler-Lagrange equation

Definition. The functional F is said to be *stationary* when $\frac{\delta F}{\delta y(x)} = 0$, i.e. when f satisfies the *Euler-Lagrange equation*

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial y}\,. \tag{2.6}$$

04/20
04/21

Example: Geodesics of the plane. What is the shortest distance between two points on the Euclidean plane (supposing that the Romans did not know the answer)? The distance L between points A and B on the curve given by $y(x)$ is given by

$$L = \int_A^B dl = \int_A^B \sqrt{dx^2 + dy^2} = \int_{x_A}^{x_B} \sqrt{1 + (y')^2}\, dx\,. \tag{2.7a}$$

We can view L as a functional of $y(x)$

$$L[y] = \int_{x_A}^{x_B} f(y')\, dx\,, \quad \text{where} \quad f(y') = \sqrt{1 + (y')^2}\,. \tag{2.7b}$$

For simplicity we assume that $y(x)$ is a single-valued function of x (so that the curve does not "double-back" on itself). The Euler-Lagrange equation (2.6) is thus

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + (y')^2}}\right) = 0\,. \tag{2.7c}$$

From integrating once

$$\frac{y'}{\sqrt{1 + (y')^2}} = c\,, \tag{2.7d}$$

where c is a constant. Hence y' is a constant, and from integrating again

$$y = ax + b\,, \tag{2.7e}$$

which is a straight line. The integration constants a and b must be chosen such that the line passes through the points A and B.

2.2.2 A first integral of the Euler-Lagrange equation

In the example above we reduced the second-order Euler-Lagrange equation to a first-order equation, (2.7d). This was because the function $f(y, y'; x)$ had no *explicit* dependence on y, i.e. $\frac{\partial f}{\partial y} = 0$, and so a *first integral* of the Euler-Lagrange equation (2.6) gave $\frac{\partial f}{\partial y'} = \text{const.}$

It is also possible to reduce the Euler-Lagrange equation to a first integral if $\frac{\partial f}{\partial x} = 0$, i.e. if f has no *explicit* dependence on x. To see this, first assume that f has an *explicit* dependence on x, i.e. $f \equiv f(y, y'; x)$, then we have from the chain rule that

$$\begin{aligned}\frac{df}{dx} &= \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial y'}\frac{d^2y}{dx^2} + \frac{\partial f}{\partial x} \\ &= \frac{\partial f}{\partial x} + y'\,\frac{\partial f}{\partial y} + y''\,\frac{\partial f}{\partial y'}\,. \end{aligned} \tag{2.8a}$$

Using the Euler-Lagrange equation (2.6) gives,

$$\begin{aligned}\frac{df}{dx} &= \frac{\partial f}{\partial x} + y'\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) + y''\frac{\partial f}{\partial y'} \\ &= \frac{\partial f}{\partial x} + \frac{d}{dx}\left(y'\frac{\partial f}{\partial y'}\right)\end{aligned} \tag{2.8b}$$

and hence

$$\frac{d}{dx}\left(f - y'\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial x}\,. \tag{2.8c}$$

When $f(y, y'; x)$ has no *explicit* dependence on x, i.e. when $\frac{\partial f}{\partial x} = 0$, then we have the first integral

$$f - y'\frac{\partial f}{\partial y'} = \text{const.} \tag{2.8d}$$

We will use this result in the next example, but it could also be applied to the above example for geodesics of the plane (an exercise for the keen).

2.2.3 Examples

The Brachistochrone. A bead slides down a frictionless wire from A to B. What shape must the wire have for the bead to fall from A to B in the shortest time?

The calculus of variations makes this difficult problem (relatively) easy. Take A to be the origin in the vertical plane, x to be the horizontal distance from the origin, and y to be the distance *below* the origin. Suppose that the bead starts with speed $v = 0$; then from conservation of energy we have that

$$\tfrac{1}{2}mv^2 = mgy \quad \Rightarrow \quad v = \sqrt{2gy}\,. \tag{2.9a}$$

We also have that

$$v = \sqrt{\dot{x}^2 + \dot{y}^2} = \dot{x}\sqrt{1 + (y')^2} \quad \Rightarrow \quad dt = \frac{dx}{v}\sqrt{1 + (y')^2}\,, \tag{2.9b}$$

where $y' = dy/dx$. The total time $T[y]$ is, from (2.9a) and (2.9b),

$$T[y] = \int_{t_A}^{t_B} dt = \frac{1}{\sqrt{2g}}\int_{x_A}^{x_B} f(y, y')dx\,, \quad \text{where} \quad f(y, y') = \sqrt{\frac{1 + (y')^2}{y}}\,. \tag{2.9c}$$

Here $\frac{\partial f}{\partial x} = 0$, and so the first integral (2.8d) implies that

$$f - y'\frac{\partial f}{\partial y'} = \sqrt{\frac{1 + (y')^2}{y}} - y'\frac{y'}{\sqrt{y\,(1 + (y')^2)}} = \frac{1}{\sqrt{y\,(1 + (y')^2)}} = \frac{1}{\sqrt{2c}}. \tag{2.9d}$$

where c is a conveniently chosen constant. Hence

$$y\left(1 + (y')^2\right) = 2c \qquad \Rightarrow \qquad y' = \sqrt{\frac{2c - y}{y}}\,. \tag{2.9e}$$

Substituting $y = c(1 - \cos\theta) = 2c\sin^2(\theta/2)$, we obtain

$$\sqrt{\frac{2c - y}{y}} = \frac{\cos(\theta/2)}{\sin(\theta/2)}\,, \quad y' = \frac{dy}{d\theta}\frac{d\theta}{dx} = 2c\,\sin(\theta/2)\cos(\theta/2)\,\frac{d\theta}{dx}\,, \tag{2.9f}$$

and so from (2.9e) and (2.9f)

$$dx = 2c\sin^2(\theta/2)\,d\theta = c(1 - \cos\theta)\,d\theta\,. \tag{2.9g}$$

Taking $y = 0$ at $\theta = x = 0$, the solution is given parametrically by

$$x = c\,(\theta - \sin\theta)\,, \qquad y = c\,(1 - \cos\theta)\,. \tag{2.9h}$$

Requiring $y = y_B$ when $x = x_B$ fixes c and the value of $\theta = \theta_B$ at point B. This is an inverted cycloid (i.e. the curve that a point on the rim of a wheel rolling along a flat plane traces out).

Remark. The calculus of variations is not restricted to functionals of the form considered so far. We can also apply it to functions of functionals, e.g. ratios, as in the following example.

Eigenfunctions as extremals of a functional. Consider the Sturm-Liouville operator $\mathcal{L}$, and the following real functionals of the real function y,

$$F[y] \equiv \langle\, y \mid \mathcal{L}y\,\rangle = \int_\alpha^\beta y\left(-\frac{d}{dx}\left(\rho(x)\frac{dy}{dx}\right) - q(x)\,y\right)dx \qquad \text{from (1.6a),} \tag{2.10a}$$

$$G[y] \equiv \langle\, y \mid y\,\rangle_w = \int_\alpha^\beta w(x)\,y^2\,dx\,, \tag{2.10b}$$

where

(i) as before, $\rho(x) > 0$ and $w(x) > 0$ for $\alpha < x < \beta$,

(ii) $F[y]$ is the inner product $\langle\, y \mid \mathcal{L}y\,\rangle$ with unit weight,

(iii) $G[y]$ is the weight-w norm $\langle\, y \mid y\,\rangle_w$ of y.

If we consider small variations in y analogous to those in (2.4b), then we have that

$$\begin{aligned} \delta F &= \int_\alpha^\beta \Big(-\delta y\,(\rho y')' - y\,(\rho(\delta y)')' - 2q\,y\,\delta y\Big)\,dx \\ &= 2\int_\alpha^\beta \delta y\Big\{-(\rho y')' - qy\Big\}dx + \Big[\rho\,(y'\,\delta y - y(\delta y)')\Big]_\alpha^\beta \quad \text{from integrating by parts twice,} \qquad (2.10c) \\ \delta G &= 2\int_\alpha^\beta \delta y\, wy\, dx \qquad (2.10d) \end{aligned}$$

From (1.6a) we see that the expression inside {} is $\mathcal{L}y$. Hence, if the boundary conditions are such that the boundary terms in (2.10c) vanish, e.g. if y satisfies homogeneous conditions $y = 0$ or $y' = 0$ at $x = \alpha$ and $x = \beta$ (cf. (1.7b)), then from the definition (2.5a) of a functional derivative, i.e. $\delta F = \int_\alpha^\beta \delta y(x) \frac{\delta F}{\delta y(x)}\,dx$, it follows that

$$\frac{\delta F}{\delta y} = 2\mathcal{L}y\,, \qquad \frac{\delta G}{\delta y} = 2wy\,. \qquad (2.10e)$$

Now consider the ratio,

$$\Lambda[y] = \frac{F[y]}{G[y]} \equiv \frac{\int_\alpha^\beta f(y,y';x)dx}{\int_\alpha^\beta g(y,y';x)dx}\,. \qquad (2.10f)$$

05/19

Then, keeping only first-order terms,

$$\begin{aligned} \delta\Lambda \equiv \Lambda[y+\delta y] - \Lambda[y] &= \frac{F[y+\delta y]}{G[y+\delta y]} - \frac{F[y]}{G[y]} \\ &= \frac{F+\delta F}{G+\delta G} - \frac{F}{G} \qquad \text{from (2.4a)} \\ &= \frac{1}{G}\Big[\Big(F+\delta F\Big)\Big(1-\frac{\delta G}{G}\Big) - F\Big] \\ &= \frac{1}{G}\Big[\delta F - \frac{F}{G}\,\delta G\Big]\,. \qquad (2.10g) \end{aligned}$$

From definition (2.5a) and (2.10e), this can be written in terms of functional derivatives as,

$$\frac{\delta\Lambda}{\delta y} = \frac{1}{G}\left[\frac{\delta F}{\delta y} - \Lambda\,\frac{\delta G}{\delta y}\right] = \frac{2}{G}\Big[\mathcal{L}\,y - \Lambda\,w\,y\Big]\,.$$

Therefore $\Lambda[y]$, which we recall is a real number, is extremized by solutions of

$$\mathcal{L}\,y = \lambda\,w\,y\,,$$

where the λ are the extremal values of Λ. This is the Sturm-Liouville eigenvalue problem with weight function w.

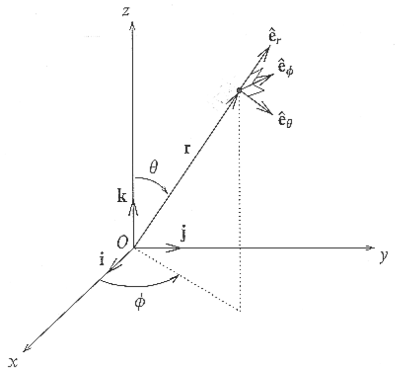
Interpretation. $\Lambda[y]$ is extremized by eigenfunctions of $\widetilde{\mathcal{L}} = w^{-1}\mathcal{L}$, and the eigenvalues are the extremal values of Λ.

Remark. A similar result holds for complex $y(x)$ if the definitions of $F[y]$ and $G[y]$ are modified to

$$\begin{aligned} F[y] &= \int_\alpha^\beta y^*\mathcal{L}y\,dx = \int_\alpha^\beta \left\{\rho(x)|y'|^2 - q(x)|y|^2\right\}dx\,, \\ G[y] &= \int_\alpha^\beta w(x)|y|^2 dx\,. \end{aligned}$$

05/20

Geodesics on the surface of a sphere.



We work in spherical polar coordinates (r, θ, ϕ), where $0 \leqslant r < \infty$, $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$. In such coordinates, the increment vector position is given by $d\mathbf{r} = dr\,\widehat{\mathbf{e}}_r + rd\theta\,\widehat{\mathbf{e}}_\theta + r\sin\theta\,d\phi\,\widehat{\mathbf{e}}_\phi$.

On the surface of a sphere $dr = 0$, and so the length of a path from A to B described by $\phi(\theta)$ on the surface of a sphere is given by

$$\begin{aligned} L &= \int_A^B |d\mathbf{r}| = \int_A^B (d\mathbf{r}\cdot d\mathbf{r})^{\frac{1}{2}}\,, \\ &= r\int_A^B \sqrt{d\theta^2 + \sin^2\theta\,d\phi^2}\,, \\ &= r\int_{\theta_A}^{\theta_B} \sqrt{1+\sin^2\theta\,(\phi')^2}\,d\theta\,, \quad \text{where} \quad \phi' = d\phi/d\theta\,. \end{aligned} \tag{2.11a}$$

05/21

By substituting θ, ϕ and $r\sqrt{1+\sin^2\theta\,(\phi')^2}$ for x, y and $f(y, y'; x)$, respectively, in the Euler-Lagrange equation (2.6), it follows that

$$\begin{aligned} \frac{d}{d\theta}\left(r\frac{\partial}{\partial\phi'}\sqrt{1+\sin^2\theta(\phi')^2}\right) &= r\frac{\partial}{\partial\phi}\sqrt{1+\sin^2\theta\,(\phi')^2} \\ &= 0\,. \end{aligned} \tag{2.11b}$$

Hence a first integral is

$$\frac{\sin^2\theta\,\phi'}{\sqrt{1+\sin^2\theta\,(\phi')^2}} = c\,. \tag{2.11c}$$

where c is a constant. This equation can be rearranged (after choosing the positive sign) to

$$\phi' = \frac{c}{\sin\theta\sqrt{\sin^2\theta - c^2}} = \frac{c}{\sin^2\theta\sqrt{1-c^2\operatorname{cosec}^2\theta}}\,. \tag{2.11d}$$

To integrate this substitute $u = \cot\theta$, so $du/d\theta = -\operatorname{cosec}^2\theta$ and,

$$\begin{aligned} \phi &= \int \frac{-c\,du}{\operatorname{cosec}^2\theta\sin^2\theta\sqrt{1-c^2\operatorname{cosec}^2\theta}} \\ &= \int \frac{-c\,du}{\sqrt{1-c^2(1+u^2)}} \\ &= \int \frac{-du}{\sqrt{k^2-u^2}} \qquad\qquad \text{where } k \equiv \frac{\sqrt{1-c^2}}{c}\,, \\ &= \cos^{-1}(u/k) + \phi_0\,, \end{aligned} \tag{2.11e}$$

where ϕ_0 is a constant. The path is therefore given by

$$\cot\theta = k\cos(\phi-\phi_0)\,, \tag{2.11f}$$

where the constants k and ϕ_0 are fixed so that the path passes through the end points.

Remark. This path is an arc of a great circle (i.e. a circle which has a centre that coincides with the centre of the sphere). For instance:

- if $\theta_A = \theta_B = \frac{\pi}{2}$, then $k = 0$ (and ϕ_0 is arbitrary), with the solution to (2.11f) being $\theta = \frac{\pi}{2}$ and ϕ taking values between ϕ_A and ϕ_B;
- if $\phi_A = \phi_B$, let $\kappa = 1/k$ and rewrite (2.11f) as

$$\kappa\cot\theta = \cos(\phi - \phi_0)\,, \tag{2.11g}$$

then $\kappa = 0$ and $\phi_0 = \phi_A - \frac{\pi}{2}$, with the solution to (2.11f) being $\phi = \phi_A = \phi_B$ and θ being able to take any value between θ_A and θ_B.

2.3 Variational Principles

Many problems can be formulated as the extremization of some functional. Two important examples are *Fermat's principle* in optics and *Hamilton's principle of least action* in mechanics.

2.3.1 Fermat's principle

Consider a material with variable *refractive index*, i.e. a material where the speed of propagation of light is c/μ, where c is speed of light in a vacuum. Then, *Fermat's principle* of geometric optics states that the path taken by a light ray from point A to point B is that which makes stationary the *optical path length*, P, where

$$P = \int_A^B \mu(\mathbf{r})\, dl\,, \tag{2.12a}$$

and dl is the three-dimensional length element, i.e.

$$dl = \sqrt{dx^2 + dy^2 + dz^2}\,. \tag{2.12b}$$

Remarks.

(i) When μ is a constant, compare with (2.7a) for a plane.

(ii) Water has $\mu \approx 1.33$, while a vacuum has $\mu = 1$ by definition.

(iii) Fermat's principle only applies in the geometric optics approximation, i.e. when the wavelength of light is small compared to the physical dimensions of the system. It does not apply when this approximation is not appropriate, e.g. as in the case of diffraction.

Using the x-coordinate to parameterize position along the path, and assuming there is no doubling back, the optical path length is a functional of $y(x)$ and $z(x)$, i.e.

$$P[y, z] = \int_{x_A}^{x_B} \mu(x, y, z)\sqrt{1 + (y')^2 + (z')^2}\, dx\,. \tag{2.12c}$$

Looking for stationary points of $P[y, z]$ with respect to variations of $y(x)$ and $z(x)$ gives the two simultaneous equations, cf. the definition (2.5a),

$$\frac{\delta P}{\delta y(x)} = 0\,, \qquad \frac{\delta P}{\delta z(x)} = 0\,. \tag{2.12d}$$

By independently varying the functions y and z, the generalisation of the Euler-Lagrange equation (2.6) is

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial y}\,, \tag{2.12e}$$

$$\frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right) = \frac{\partial f}{\partial z}\,, \tag{2.12f}$$

where $f = \mu(x, y, z)\sqrt{1 + (y')^2 + (z')^2}$.

Remarks.

(i) If μ is a constant, then these imply that y' and z' are constants (cf. geodesics of the plane). The path is thus the intersection of two planes, i.e. a straight line.

(ii) In simple applications the problem can often be reduced to a path in a plane and so the variation of a functional of one function.

Example: Snell's law. We can use Fermat's principle to derive Snell's law which gives the angle by which a light ray is bent on passing from one material to another.

We suppose that $\mu \equiv \mu(y)$, that the path is in the x-y plane, and that

$$\mu = \begin{cases} \mu_1 & y < y_1 \\ \mu_2 & y > y_2 \end{cases} . \tag{2.13a}$$

From (2.12c), the integrand to make stationary is

$$f(y, y') = \mu(y)\sqrt{1 + (y')^2} . \tag{2.13b}$$

This has no explicit x-dependence and so the first-integral (2.8d) gives

$$\mu(y)\sqrt{1 + y'^2} - \frac{\mu(y)\, y'^2}{\sqrt{1 + y'^2}} = k , \tag{2.13c}$$

where k is a constant. After a little manipulation, this yields

$$\frac{\mu(y)}{\sqrt{1 + (y')^2}} = k . \tag{2.13d}$$

Let $y' = -\cot\theta$ (as in the diagram), then the solution (2.13d) to the path can alternatively be written as

$$\mu \sin\theta = k . \tag{2.13e}$$

For regions where μ is constant this implies that the path is a straight line, and since k is the same constant for the entire path, we deduce Snell's law that

$$\mu_1 \sin\theta_1 = \mu_2 \sin\theta_2 . \tag{2.13f}$$

Remark. This result remains true if the interface is 'quite' sharp, i.e. if $y_1 \approx y_2$, although in order to be consistent with the geometric optics approximation, we need to assume that the interface still extends over a number of wavelengths of light.

Remark. Fermat's original formulation was slightly different. It was based on minimizing the time, say τ, for a light ray to go from A to B, i.e. on minimizing

$$\tau = \int_A^B \frac{dl}{v} = \int_A^B \frac{\mu\, dl}{c} \equiv \frac{P}{c} ,$$

where light travels at a speed $v(\mathbf{r}) = c/\mu(\mathbf{r})$. This is equivalent to minimizing the optical path length P specified in (2.12a). The modern version is to find *stationary* paths rather than the *minimal* paths. This has advantages because μ is directly measurable, while the speed of light in a medium is ambiguous (Fermat's velocity must be interpreted as the phase velocity, rather than the group velocity).

Sound waves. There is an analogous (approximate) principle for sound waves where the optical path length is replaced by the acoustic path length. This can be used to explain why distant sounds are heard better at night.

The speed of sound v depends on the density of air, and thence the absolute air temperature T such that, if the air is modelled by a perfect gas, $v = \sqrt{\gamma R T}$, where γ is the ratio of specific heats and R is specific gas constant. After sunset, the ground cools faster than the air setting up a temperature gradient. If we assume $T = T_0 + \alpha z$, where z is the height *above* ground and α is a positive constant, then $v \propto \sqrt{T_0 + \alpha z}$.

This leads to a variational problem for

$$P[z] = \int_{x_A}^{x_B} f(z, z')dx , \quad \text{where} \quad f(z, z') = \sqrt{\frac{1 + (z')^2}{T_0 + \alpha z}} , \tag{2.14}$$

06/19 that is equivalent to the Brachistochrone problem (cf. (2.9c), noting that y in the Brachistochrone problem points in the opposite direction to z here).

2.3.2 Hamilton's principle

Lagrange and Hamilton developed a powerful reformulation of Newtonian mechanics in terms of a '*principle of least action*' based on energy rather than force. The time evolution of a system is viewed as the motion of a point in a multi-dimensional *configuration space* described by some *generalised coordinates* $\{q_i\}$. For example

(i) A system of n particles (in $3n$-dimensional coordinate space) can be described by the 3 coordinates for each of n positions.

(ii) A rigid pendulum swinging in a vertical plane requires one generalised coordinate, the angle to the vertical, say, $\theta = q_1$.

(iii) A top spinning on its axis on a smooth plane requires five generalised coordinates:

- two, $(x, y) = (q_1, q_2)$, to describe the position of the point of contact,
- one, $\theta = q_3$, for the angle of the axis to the vertical,
- one, $\phi = q_4$, for the rotation of the axis about the vertical,
- one, $\psi = q_5$ [a], for the rotation of the top about its axis.

[a] Nearly as in Gangnam Style.

06/20
06/21

Remark. Problems can often be simplified by a convenient choice of generalised coordinates; this is part of the power of these methods.

Definition: Lagrangian. The *Lagrangian* is defined as

$$L = T - V \tag{2.15a}$$

where T is the kinetic energy and V is the potential energy.

Definition: Action. The *action* of a path, starting at time t_i and ending at t_f, is given by,

$$S[\{q_i\}] = \int_{t_i}^{t_f} L\left(\{q_i(t)\}, \{\dot{q}_i(t)\}, \ldots; t\right)\, dt\,. \tag{2.15b}$$

Hamilton's principle. **Hamilton's principle** states that the motion in configuration space extremizes the action functional S.

Lagrange's equations. For a Lagrangian $L\left(\{q_i\}, \{\dot{q}_i\}; t\right)$, with $i = 1, \ldots, N$ generalised coordinates, and for a motion with fixed start and end points, it follows from independently varying the coordinates that, cf. (2.6),

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0\,, \qquad i = 1, \ldots, N\,. \tag{2.15c}$$

These are *Lagrange's equations.*

A first integral. If the Lagrangian $L\left(\{q_i\}, \{\dot{q}_i\}; t\right)$ *has no explicit dependence on* t, then by generalizing the derivation of the first integral (2.8d), we can find a constant of the motion. The chain rule and Lagrange's equations (2.15c) give

$$\begin{aligned}
\frac{dL}{dt} &= \frac{\partial L}{\partial t} + \sum_{i=1}^{N}\left(\dot{q}_i \frac{\partial L}{\partial q_i} + \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) \\
&= \frac{\partial L}{\partial t} + \sum_{i=1}^{N}\left(\dot{q}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) + \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) \\
&= \frac{\partial L}{\partial t} + \frac{d}{dt}\left(\sum_{i=1}^{N} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right)\,,
\end{aligned} \tag{2.15d}$$

and hence

$$\frac{d}{dt}\left[L-\sum_{i=1}^{N}\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}\right]=\frac{\partial L}{\partial t}\,. \tag{2.15e}$$

If the Lagrangian has no explicit dependence on t, i.e. $\frac{\partial L}{\partial t}=0$, then we have the first integral

$$\sum_{i=1}^{N}\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}-L=\text{const.} \tag{2.15f}$$

Remark. Often, when the Lagrangian does not explicitly depend on time,

(i) the kinetic energy T is a homogeneous quadratic in the generalised velocities $\{\dot{q}_i\}$, i.e.

$$T=\sum_i\sum_j a_{ij}(q_1,\ldots,q_N)\,\dot{q}_i\dot{q}_j\,; \tag{2.15g}$$

(ii) the potential energy V does not depend on the generalised velocities, i.e.

$$V\equiv V(q_1,\ldots,q_N)\,. \tag{2.15h}$$

In such cases it follows that

$$\sum_{i=1}^{N}\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}-L=T+V=\text{const.} \tag{2.15i}$$

i.e. the total energy $E=T+V$ is conserved.

Unlectured proof. From (2.15a), (2.15g) and (2.15h), and careful use of dummy variables,

$$\begin{aligned}\sum_{i=1}^{N}\dot{q}_i\frac{\partial L}{\partial \dot{q}_i}-L&=\sum_{i=1}^{N}\dot{q}_i\frac{\partial}{\partial \dot{q}_i}\left(\sum_{k=1}^{N}\sum_{j=1}^{N}a_{kj}(q_1,\ldots,q_N)\,\dot{q}_k\dot{q}_j-V(q_1,\ldots,q_N)\right)-T+V\\&=\sum_{i=1}^{N}\dot{q}_i\left(\sum_{k=1}^{N}\sum_{j=1}^{N}a_{kj}\,\left(\delta_{ik}\dot{q}_j+\dot{q}_k\delta_{ij}\right)\right)-T+V\qquad\text{since }\frac{\partial \dot{q}_k}{\partial \dot{q}_i}=\delta_{ik}\\&=\sum_{i=1}^{N}\sum_{j=1}^{N}a_{ij}\,\dot{q}_i\dot{q}_j+\sum_{i=1}^{N}\sum_{k=1}^{N}a_{ki}\,\dot{q}_k\dot{q}_i-T+V\\&=T+V\,.\end{aligned}$$

□

Example. Consider a particle of mass m subject to a *conservative* force field $\mathbf{F}(\mathbf{r})=-\nabla V(\mathbf{r})$. From the definition of the Lagrangian (2.15a) and using Cartesian coordinates, we have that

$$\begin{aligned}L=T-V&=\tfrac{1}{2}m|\dot{\mathbf{r}}|^2-V(\mathbf{r})\\&=\tfrac{1}{2}m\sum_{i=1}^{3}\dot{x}_i^2-V(x_1,x_2,x_3)\,.\end{aligned} \tag{2.16a}$$

Thus

$$\frac{\partial L}{\partial \dot{x}_i}=m\dot{x}_i\,,\quad \frac{\partial L}{\partial x_i}=-\frac{\partial V}{\partial x_i}\,, \tag{2.16b}$$

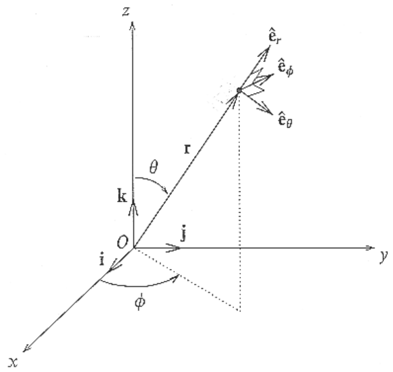
and so the Lagrange's equations (2.15c) yield Newton's second law

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_i}\right)-\frac{\partial L}{\partial x_i}=m\ddot{x}_i+\frac{\partial V}{\partial x_i}=0\,,\quad\text{i.e.}\quad m\ddot{\mathbf{r}}=-\nabla V=\mathbf{F}\,. \tag{2.16c}$$

Energy conservation. We also find using the first integral (2.15f), the definition of the Lagrangian (2.16a), and the partial derivatives (2.16b) that, consistent with (2.15i), the energy, $E=T+V$, is conserved:

$$\begin{aligned}\sum_{i=1}^{N}\dot{x}_i\frac{\partial L}{\partial \dot{x}_i}-L&=\sum_{i=1}^{N}\dot{x}_i\frac{\partial}{\partial \dot{x}_i}\left(\tfrac{1}{2}m\sum_{j=1}^{3}\dot{x}_j^2-V(x_1,x_2,x_3)\right)-\tfrac{1}{2}m\sum_{i=1}^{3}\dot{x}_i^2+V(x_1,x_2,x_3)\\&=\tfrac{1}{2}m\sum_{i=1}^{3}\dot{x}_i^2+V(x_1,x_2,x_3)=E=\text{const.}\end{aligned} \tag{2.16d}$$

Example. Reconsider the above example, but this time for a central force field V that depends only on $r = |\mathbf{r}|$.



Then the natural choice of coordinates are spherical polars $(q_1, q_2, q_3) \equiv (r, \theta, \phi)$, where $0 \leqslant r < \infty$, $0 \leqslant \theta \leqslant \pi$ and $0 \leqslant \phi \leqslant 2\pi$. From dynamics we know that the motion is planar, i.e. in the plane normal to the constant angular momentum vector. Hence, without loss of generality, we can orientate the axes so that the motion is in a plane $\phi =$ constant. It then follows that the Lagrangian is given by, cf. (2.16a),

$$L = \tfrac{1}{2} m\dot{r}^2 + \tfrac{1}{2} m r^2 \dot{\theta}^2 - V(r)\,. \tag{2.17a}$$

Hence Lagrange's equations, (2.15c), are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} = m\ddot{r} - mr\dot{\theta}^2 + V' = 0\,, \tag{2.17b}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = \frac{d}{dt}\left(r^2\dot{\theta}\right) = 0\,. \tag{2.17c}$$

The second equation, (2.17c), yields conservation of the angular momentum, J:

$$J/m = r^2\dot{\theta} = \text{constant} = h\,. \tag{2.17d}$$

Using this result the first equation, (2.17b), reduces to

$$m\ddot{r} = -\frac{dV}{dr} + \frac{mh^2}{r^3} = -\frac{dV_{\text{eff}}}{dr}\,, \quad \text{where} \quad V_{\text{eff}}(r) = V(r) + \frac{mh^2}{2r^2}\,. \tag{2.17e}$$

For $h \neq 0$ the *effective potential* $V_{\text{eff}}(r)$ has a *centrifugal barrier*. For instance, if $V = -\frac{GMm}{r}$, where G is Newton's gravitational constant and M the mass of, say, the sun, then we get

$$V_{\text{eff}}(r) = m\left(-\frac{GM}{r} + \frac{h^2}{2r^2}\right) \tag{2.17f}$$

Because $V_{\text{eff}} \propto m$, the acceleration will be independent of the mass m. When $h \neq 0$, the centrifugal barrier prevents an approach to $r = 0$; in this case a sketch of V_{eff} shows there will be stable orbits.

Example. Consider two particles, masses m_1 and m_2, interacting via a potential $V(\mathbf{r}_1 - \mathbf{r}_2)$.

A point in configuration space can be specified by the two position vectors $\mathbf{r}_1$, $\mathbf{r}_2$. Alternatively, we can use the centre of mass $\mathbf{R}$ and the relative position $\mathbf{r}$, where

$$\mathbf{q}_1 \equiv \mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{M} \quad \text{and} \quad M = m_1 + m_2\,, \tag{2.18a}$$

$$\mathbf{q}_2 \equiv \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2\,. \tag{2.18b}$$

07/19

Then the kinetic energy can be shown to be given by (after some manipulation)

$$T = \tfrac{1}{2} m_1 |\dot{\mathbf{r}}_1|^2 + \tfrac{1}{2} m_2 |\dot{\mathbf{r}}_2|^2 = \tfrac{1}{2} M |\dot{\mathbf{R}}|^2 + \tfrac{1}{2}\mu|\dot{\mathbf{r}}|^2 \quad \text{and} \quad \mu = \frac{m_1 m_2}{m_1 + m_2}\,, \tag{2.18c}$$

where μ is the *reduced mass*. The Lagrangian is therefore

$$L = T - V = \tfrac{1}{2} M \dot{\mathbf{R}} \cdot \dot{\mathbf{R}} + \tfrac{1}{2}\mu\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - V(\mathbf{r})\,. \tag{2.18d}$$

The Lagrange equations for the components of $\mathbf{R}$ can be written as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{R}_i}\right) - \frac{\partial L}{\partial R_i} = \frac{d}{dt}(M\dot{R}_i) = 0\,, \quad \Rightarrow \quad \ddot{\mathbf{R}} = 0 \quad \text{and} \quad \dot{\mathbf{R}} = \text{const.} \tag{2.18e}$$

Hence the centre of mass moves with constant velocity. The Lagrange equations for the components of $\mathbf{r}$ yield

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}_i}\right) - \frac{\partial L}{\partial r_i} = \mu \ddot{r}_i + \frac{\partial V}{\partial r_i} = 0\,, \quad \Rightarrow \quad \mu \ddot{\mathbf{r}} = -\boldsymbol{\nabla} V\,. \tag{2.18f}$$

Further, because

- T is a homogeneous quadratic in the generalised velocities,
- V doesn't depend on the velocities,
- L contains no explicit t-dependence,

we can also conclude that $E = T + V$ is constant.

2.4 Constrained Variation and Lagrange Multipliers

Recall from Taylor's theorem that in $\mathbb{R}^3$

$$\delta f = f(\mathbf{x} + \delta\mathbf{x}) - f(\mathbf{x}) = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z + \ldots = \boldsymbol{\nabla} f \cdot \delta\mathbf{x} + \ldots\,. \tag{2.19a}$$

Further, in the limit $|\delta\mathbf{x}| \to 0$ we have, in any number of dimensions,

$$df = \boldsymbol{\nabla} f \cdot d\mathbf{x}\,. \tag{2.19b}$$

Suppose that $f(x, y)$ is the height of a function above the x-y plane, and suppose that there is a *hilltop*. Consider lines of constant f, i.e. contours around the hilltop, together with a path specified by $p(x, y) = 0$ that does not reach the hilltop.

07/17

Hilltop maximum. At the hilltop (a maximum) a small change in position, say $d\mathbf{l}$, does not change f. Then

$$0 = df = d\mathbf{l} \cdot \nabla f\,, \tag{2.20}$$

i.e. ∇f is orthogonal to all possible displacement vectors and is zero there. To find the position of the hilltop we need to solve two equations $\nabla f = 0$ for (x, y); note that these criteria are the same for
07/20 minima and saddle points, so it may be necessary to check for a maximum.

07/21 *Path maximum.* Suppose instead we want to find the position of the highest point on the path. We still require $0 = df = d\mathbf{l} \cdot \nabla f$ but $d\mathbf{l}$ is no longer arbitrary; it must lie on the path. The constraint for $d\mathbf{l}$ to remain on the path $p(x, y) = 0$ is, using (2.19b), that

$$0 = dp = d\mathbf{l} \cdot \nabla p\,. \tag{2.21a}$$

From (2.20) and (2.21a), at the highest point on the path, ∇f will be orthogonal to all $d\mathbf{l}$ that are orthogonal to ∇p. Therefore ∇f and ∇p are parallel/anti-parallel, i.e. $\nabla f = \lambda \nabla p$ for some λ. Thus, the maximisation problem is to solve

$$\nabla f - \lambda \nabla p = 0 \quad \text{with} \quad p = 0\,. \tag{2.21b}$$

Remark. There is both an additional equation, $p = 0$, and an additional variable, λ.

A Lagrange multiplier. An alternative approach to these equations is to note that they also arise from extremization *without constraint* of the following function of three variables

$$\varphi(x, y; \lambda) = f(x, y) - \lambda\, p(x, y)\,. \tag{2.22}$$

- Extremization with respect to the *Lagrange multiplier* λ gives the constraint $p = 0$.
- Extremization with respect to x, y gives the other equations.

Remarks.

(i) By introducing a Lagrange multiplier we have turned a constrained variation problem into an unconstrained variation problem.

(ii) Actually we replaced maximization by extremization and so we could end up finding a minimum rather than a maximum, but this is usually quite easy to sort out.

(iii) The Lagrange multiplier sometimes has some significance to the problem.

Extension to functions of N variables. We can extend the method to find stationary points of a function $f(\boldsymbol{\xi})$ of N variables $(\xi_1, \ldots, \xi_N)$ subject to $k < N$ constraints $p_i(\boldsymbol{\xi}) = 0$ $(i = 1, \ldots, k)$. We now need k Lagrange multipliers and we have to extremize a generalised φ with respect to the $N + k$ variables:

$$\varphi(\boldsymbol{\xi}; \lambda_1, \ldots, \lambda_k) = f(\boldsymbol{\xi}) - \sum_{i=1}^{k} \lambda_i p_i(\boldsymbol{\xi}). \tag{2.23}$$

Extension to functionals. There is also a generalization to functionals ($N = \infty$). For example, to extremize $F[y]$ subject to the constraint $P[y] = 0$, we may extremize *without constraint*,

$$\Phi_\lambda[y] = F[y] - \lambda\, P[y]\,, \tag{2.24a}$$

with respect to the function y and the variable λ. Assuming the boundary terms are zero we obtain, cf. (2.21b),

$$\frac{\delta F}{\delta y(x)} - \lambda \frac{\delta P}{\delta y(x)} = 0\,, \quad P[y] = 0\,. \tag{2.24b}$$

Example: Catenary. What is the shape of the curve described by a uniform chain hanging under its own weight from two fixed points?

Let the two fixed points be at $x = \pm L$. Assume that the chain has a fixed length $\ell_0 > 2L$ and a constant mass per unit length ρ. The potential energy of an element dl is $dV = (\rho\, dl) g y$ where $y(x)$ is the height of the chain above the ground. Therefore,

$$\begin{aligned}\frac{V}{\rho g} &= \int_{\text{chain}} y\, dl = \int_{x=-L}^{x=L} y\sqrt{dx^2 + dy^2} \\ &= \int_{-L}^{L} y\sqrt{1 + (y')^2}\, dx\,. \end{aligned} \tag{2.25a}$$

We minimize $V/\rho g$ subject to the constraint

$$\ell_0 = \int_{\text{chain}} dl = \int_{-L}^{L} \sqrt{1 + (y')^2}\, dx\,. \tag{2.25b}$$

This is equivalent to extremizing *without constraint*,

$$\begin{aligned}\Phi_\lambda[y] &= \int_{-L}^{L} y\sqrt{1 + (y')^2}\, dx - \lambda\left(\int_{-L}^{L} \sqrt{1 + (y')^2}\, dx - \ell_0\right)\,, \\ &= \int_{-L}^{L} f(y, y'; \lambda)\, dx + \lambda \ell_0 \quad \text{with} \quad f(y, y'; \lambda) = (y - \lambda)\sqrt{1 + (y')^2}\,, \end{aligned} \tag{2.25c}$$

where λ is a Lagrange multiplier. Because there is no explicit dependence on x in the integrand, the first integral (2.8d) yields

$$y'\frac{\partial f}{\partial y'} - f = \frac{\lambda - y}{\sqrt{1 + (y')^2}} = c\,, \tag{2.25d}$$

where c is a constant. After rearrangement

$$y' = c^{-1}\sqrt{(y-\lambda)^2 - c^2}\,. \tag{2.25e}$$

08/19

On making the substitution $y - \lambda = c\cosh(u/c)$, we find that $u' = 1$. Hence $u = x + a$ for some constant a, and thus

$$y(x) = \lambda + c\cosh\left(\tfrac{x+a}{c}\right)\,. \tag{2.25f}$$

We can fix the constants (c, a) by requiring the chain to pass through the fixed end points. For simplicity, suppose both ends are at height h above the ground, then $a = 0$ from symmetry, and

$$h = \lambda + c\cosh(L/c)\,.$$

It can be shown that the total length of the curve is $2c\sinh(L/c)$, and so the catenary is specified by

$$y(x) = c\cosh(x/c)\,, \quad 2c\sinh(L/c) = \ell_0\,, \tag{2.25g}$$

where, because the shape is independent of h, we have chosen h so that the lowest point of the chain is at $y = c$ and $\lambda = 0$.

Example: Isoperimetric problem. What closed curve of fixed length L in a plane maximizes an enclosed area A?

We will assume that the curve does not double back (other than in the obvious way). The area of the strip is $dA = y\,dx$, so the total area is

$$A = \oint_{C_-+C_+} y(x)\,dx = \oint_C y(x)\,dx\,. \tag{2.26a}$$

We need to maximize A subject to the constraint

$$\begin{aligned} L = \oint_C dl &= \oint_C \sqrt{dx^2 + dy^2} \\ &= \oint_C \sqrt{1 + (y')^2}\,dx\,. \end{aligned} \tag{2.26b}$$

Therefore we have to extremize *without constraint*

$$\begin{aligned} \Phi_\lambda[y] &= \oint_C y dx - \lambda\left(\oint_C \sqrt{1+(y')^2}\,dx\, - L\right) \\ &= \oint_C f(y, y'; \lambda)\,dx + \lambda L \quad \text{where} \quad f(y, y'; \lambda) = y - \lambda\sqrt{1+(y')^2}\,, \end{aligned} \tag{2.26c}$$

with respect to the function y and the real variable λ.

The boundary terms are effectively periodic, and $f(y, y'; \lambda)$ has no explicit x-dependence. Hence the first integral, (2.8d), gives

$$f - y'\frac{\partial f}{\partial y'} = y - \frac{\lambda}{\sqrt{1+(y')^2}} = y_0\,, \tag{2.26d}$$

where y_0 is a constant. After some manipulation, this is equivalent to

$$y' = \left(\frac{\lambda^2}{(y-y_0)^2} - 1\right)^{\frac{1}{2}}\,. \tag{2.26e}$$

This separable ODE has solution

$$y = y_0 \pm \sqrt{\lambda^2 - (x-x_0)^2}\,, \tag{2.26f}$$

for some constant x_0. Hence the maximising curve is a circle of radius λ:

$$(x-x_0)^2 + (y-y_0)^2 = \lambda^2\,. \tag{2.26g}$$

Varying Φ_λ with respect to λ gives $2\pi\lambda = L$.

Example. Sturm-Liouville eigenfunctions. Consider a formulation of the Sturm-Liouville problem as an extremization of F subject to a normalisation constraint $G = 1$, where, as in (2.10a) and (2.10b),

$$F[y] = \langle\, y \mid \mathcal{L}y \,\rangle = \int_\alpha^\beta y \left(-\left(\rho(x)y'\right)' - q(x)\, y \right) dx \tag{2.27a}$$

$$G[y] = \langle\, y \mid y \,\rangle_w = \int_\alpha^\beta w(x)\, y^2 \, dx \,. \tag{2.27b}$$

08/20

This is equivalent to extremizing *without constraint*,

$$\Phi_\lambda[y] = F[y] - \lambda\left(G[y] - 1\right) = \left(F[y] - \lambda G[y]\right) + \lambda \,, \tag{2.27c}$$

with respect to the function y and the real variable λ. Assuming again that the boundary terms in (2.10c) are zero, then from (2.10e) we have that

$$\frac{\delta F}{\delta y} = 2\mathcal{L}y \,, \qquad \frac{\delta G}{\delta y} = 2wy \,. \tag{2.27d}$$

Extremizing $\Phi_\lambda[y]$ with respect to y we therefore obtain that

$$\mathcal{L}y - \lambda w y = 0 \,, \tag{2.27e}$$

i.e. the Sturm-Liouville eigenvalue equation, where the Lagrange multiplier λ is the eigenvalue.

08/21

Remark. This is the same result as we found on page 20 by extremizing the ratio $\Lambda = F/G$, where the extremal values of Λ are the eigenvalues.

2.5 Estimating Eigenvalues: The Rayleigh-Ritz Method

In §2.2.3 we have seen that the eigenvalues of a Sturm-Liouville problem are the extremal values of $\Lambda = F/G$ where, from (2.10a) and (2.10b), and assuming that the boundary terms after integrating by parts are zero,

$$\begin{aligned} F[y] = \langle\, y \mid \mathcal{L}y \,\rangle &= \int_\alpha^\beta y \left(-\left(\rho(x)y'\right)' - q(x)\, y \right) dx \\ &= \int_\alpha^\beta \left(\rho(x)\,(y')^2 - q(x)\, y^2 \right) dx \,, \qquad \text{from integrating by parts,} \end{aligned} \tag{2.28a}$$

$$G[y] = \langle\, y \mid y \,\rangle_w = \int_\alpha^\beta w(x)\, y^2 \, dx \,. \tag{2.28b}$$

Suppose that $q \leqslant 0$, in addition to $\rho > 0$, so that $F \geqslant 0$. In this case, recalling that $w(x) > 0$ and hence that $G > 0$ for non-zero y, it follows that $\Lambda \geqslant 0$. Further, this implies that one of the extremal values, λ_0, is an *absolute minimum*. Suppose that y_0 is the eigenfunction corresponding to λ_0 (and for simplicity we assume that there is no degeneracy). Then we have the inequalities

$$\Lambda[y] \geqslant \Lambda[y_0] = \lambda_0 \geqslant 0 \,, \tag{2.28c}$$

with $\Lambda[y] = \Lambda[y_0]$ if and only if $y = y_0$.

Rayleigh-Ritz method. This provides us with a way to find an upper bound on λ_0. Suppose we make an [educated] guess, y_{trial}, for y_0 and evaluate $\Lambda[y_{\text{trial}}]$, then

$$\Lambda[y_{\text{trial}}] \geqslant \lambda_0 \,. \tag{2.29a}$$

The better y_{trial} is, the closer $\Lambda[y_{trial}]$ will be to λ_0. Moreover, because $\Lambda[y]$ is stationary at $y = y_0$, a moderately-good guess should yield a reasonable approximation to λ_0.

Taking this further, we may choose y_{trial} to depend on one or more parameters $(\gamma_1, \gamma_2, \ldots)$. We know $\lambda_0 \leqslant \Lambda(\{\gamma_i\})$ for *all choices of the parameters* and so we get the best (lowest) upper bound by minimizing $\Lambda(\{\gamma_i\})$ with respect to $\{\gamma_i\}$, i.e.

$$\lambda_0 \leqslant \min_{\{\gamma_i\}} \Lambda(\{\gamma_i\}) \,. \tag{2.29b}$$

This is the *Rayleigh-Ritz method.*

Alternative derivation of (2.29a). Since the eigenfunctions of the Sturm-Liouville operator form a basis, we can express y_{trial} as

$$y_{\text{trial}} = \sum_{n=0}^{\infty} b_n y_n \,, \tag{2.29c}$$

for some coefficients b_n, where we do not need to know what the eigenfunctions y_n are, just that they are a complete set. It follows that

$$\begin{aligned}
\Lambda[y_{\text{trial}}] &= \frac{\langle \sum_{m=0}^{\infty} b_m y_m \mid \mathcal{L} \sum_{n=0}^{\infty} b_n y_n \rangle}{\langle \sum_{m=0}^{\infty} b_m y_m \mid \sum_{n=0}^{\infty} b_n y_n \rangle_w} && \text{with a careful choice of dummy variables} \\
&= \frac{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} b_m^* b_n \langle y_m \mid \lambda_n w y_n \rangle}{\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} b_m^* b_n \langle y_m \mid y_n \rangle_w} && \text{since } \mathcal{L} y_n = \lambda_n w y_n \\
&= \frac{\sum_{n=0}^{\infty} |b_n|^2 \lambda_n}{\sum_{n=0}^{\infty} |b_n|^2} && \text{using } \langle u \mid wv \rangle = \langle u \mid v \rangle_w \text{ \& orthonormality} \quad (2.29\text{d}) \\
&\geqslant \frac{\sum_{n=0}^{\infty} |b_n|^2 \lambda_0}{\sum_{n=0}^{\infty} |b_n|^2} = \lambda_0 && \text{since } \lambda_n \geqslant \lambda_0. \quad (2.29\text{e})
\end{aligned}$$

Remark. Applications include finding wavefunctions/energies in quantum mechanics, and in identifying resonances, as illustrated in the following examples.

Example: Quantum harmonic oscillator. Let

$$F[\psi] = \int_{-\infty}^{\infty} \left[(\psi')^2 + x^2\psi^2\right] dx \,, \quad G[\psi] = \int_{-\infty}^{\infty} \psi^2 dx \,. \tag{2.30a}$$

Then from above, the functions that extremize $\Lambda = F/G$ are the solutions to

$$\mathcal{L}\,\psi = 2E\,\psi \,, \quad \mathcal{L} = -\frac{d^2}{dx^2} + x^2 \,, \tag{2.30b}$$

subject to the boundary conditions that $\psi \to 0$ as $|x| \to \infty$ (these boundary conditions ensure that the Sturm-Liouville operator $\mathcal{L}$ is self-adjoint). By denoting the eigenvalues by $2E$, for a suitable choice of units, (2.30b) is the Schrödinger equation for a particle of energy E in a harmonic oscillator potential. Suppose we try

$$\psi_{\text{trial}}(x) = \exp\left(-\tfrac{1}{2}\alpha x^2\right) \,, \tag{2.31a}$$

with parameter $\alpha > 0$ to satisfy the boundary conditions. Then

$$F = \int_{-\infty}^{\infty} (\alpha^2+1) x^2 \exp\left(-\alpha x^2\right) dx \,, \quad G = \int_{-\infty}^{\infty} \exp\left(-\alpha x^2\right) dx \,. \tag{2.31b}$$

Using the result that

$$\int_{-\infty}^{\infty} x^{2n} e^{-\alpha x^2} dx = \frac{(2n)!}{2^{2n} n!}\sqrt{\frac{\pi}{\alpha^{2n+1}}} \,, \tag{2.31c}$$

we find that

$$\Lambda\left[\psi_{\text{trial}}\right] = \frac{\alpha^2+1}{2\alpha} \,. \tag{2.31d}$$

Since

$$\frac{\partial \Lambda\left[\psi_{\text{trial}}\right]}{\partial \alpha} = 1 - \frac{\alpha^2+1}{2\alpha^2} = \frac{\alpha^2-1}{2\alpha^2} \,, \tag{2.31e}$$

this is a minimum for $\alpha > 0$ when $\alpha = 1$, in which case

$$\psi_{\text{trial}} = \exp\left(-\tfrac{1}{2}x^2\right) \,, \quad \Lambda\left[\psi_{\text{trial}}\right] = 1 \,. \tag{2.31f}$$

We deduce that $\lambda_0 \leqslant 1$.

Remark. In fact this is the exact answer, because it so happens that we have included the exact eigenfunction corresponding to the lowest eigenvalue in the family of trial functions considered.

A less-inspired guess might be

$$\psi_{\text{trial}} = \left(1 + ax^2\right) e^{-x^2}\,, \tag{2.32a}$$

which yields after some manipulation

$$\Lambda\left[\psi_{\text{trial}}\right] = \frac{1}{4}\left(\frac{43a^2 - 8a + 80}{3a^2 + 8a + 16}\right) \equiv \Lambda(a)\,. \tag{2.32b}$$

The minimum of $\Lambda(a)$ is at $a_{\text{min}} = \frac{4}{23}\left(-7 + \sqrt{118}\right)$ and so

$$\lambda_0 \leqslant \Lambda\left(a_{\text{min}}\right) = \frac{2537 - 162\sqrt{118}}{708 + 4\sqrt{118}} \approx 1.03\,. \tag{2.32c}$$

Remark: estimating the error. This estimate is within 3% of the exact answer, but we would not know this if we did not already know the answer! We can, however, improve our bound by including more parameters. As the integrand depends on x^2, we might guess that $\psi_0(x)$ is an even function of x, so an improved trial function might be,

$$\psi_{\text{trial}} = \left(1 + ax^2 + bx^4\right) e^{-x^2}\,. \tag{2.32d}$$

This new estimate could then be systematically improved by including an x^6 term etc. An estimate of the [percentage] error might then be obtained by seeing how much each improvement changes the result.

Example: Circularly symmetric vibrations of a circular drum. Consider a drum with unit radius fixed at $r = 1$. Then the amplitude $y(r)$ of small-amplitude vibrations satisfies Bessel's equation

$$\frac{d^2y}{dr^2} + \frac{1}{r}\frac{dy}{dr} + \lambda y = 0\,, \tag{2.33a}$$

subject to $y(1) = 0$ and $y(0)$ being finite. The eigenvalue λ scales like the square of the angular frequency; the dominant sound is thus from the lowest frequency. Suppose that we want to estimate this, i.e. to estimate λ_0.

There are a number of steps.

(i) From (1.13b), first multiply (2.33a) by $w = r$ to put the operator in Sturm-Liouville form:

$$(ry')' + \lambda r y = 0\,. \tag{2.33b}$$

Then, from (2.28a) and (2.28b),

$$F[y] = \int_0^1 r\left(y'\right)^2\,dr\,,\quad G[y] = \int_0^1 r y^2\,dr\,. \tag{2.33c}$$

(ii) Try $y_{\text{trial}} = a + br^2 + cr^4$, where $a + b + c = 0$ to satisfy the boundary condition $y(1) = 0$. We include only even powers because the equation (2.33a) has the same form when $r \to -r$.

(iii) Next compute $F[y_{\text{trial}}]$ and $G[y_{\text{trial}}]$. Using $a = -b - c$ we get,

$$F[y_{\text{trial}}] = b^2 + \tfrac{8}{3}bc + 2c^2 = f(b,c)\,, \tag{2.33d}$$

$$G[y_{\text{trial}}] = \tfrac{1}{6}b^2 + \tfrac{5}{12}bc + \tfrac{4}{15}c^2 = g(b,c)\,. \tag{2.33e}$$

(iv) Now minimize $\Lambda(b,c) = f(b,c)/g(b,c)$ with respect to b, c to deduce that at the minimum

$$4\left(\Lambda - 6\right)b = \left(32 - 5\Lambda\right)c \quad\text{and}\quad 5\left(32 - 5\Lambda\right)b = 16\left(2\Lambda - 15\right)c\,. \tag{2.33f}$$

These imply that

$$3\Lambda^2 - 128\Lambda + 640 = 0 \quad\Rightarrow\quad \Lambda = \tfrac{8}{3}\left(8 \pm \sqrt{34}\right)\,. \tag{2.33g}$$

Choosing the minus sign to get the lower value of Λ gives an approximation to λ_0 of $\lambda \approx 5.784$. This is close to the exact value $\lambda = 5.7832\ldots$.

Remarks.

(i) If we used the simplified trial function $y_{\text{trial}} = a + br^2$, then putting $c = 0$ in (2.33d) and (2.33e) we deduce that $\lambda \approx 6$, which suggests a 4% error in λ_0 (in fact, as noted above, the error is much smaller than this).

(ii) Using the estimated value of λ we get a relation between b and c, Then, along with the normalization condition $G = 1$, this determines the *optimal* trial function of the form chosen.

2.5.1 Extension to higher eigenvalues (non-examinable)

Suppose we already have a good approximation to λ_0 and y_0, i.e. the *ground state* in quantum mechanics, and suppose that we want an approximation to the next-lowest eigenvalue λ_1. We know that the eigenfunction y_1 is orthogonal to y_0, which suggests we should consider a trial function $y_{\text{trial}}^{(1)}$ with this property. We can write such a function as, cf. (2.29c),

$$y_{\text{trial}}^{(1)} = \sum_{n=1}^{\infty} b_n y_n \tag{2.34a}$$

for some coefficients b_n. The $b_0 y_0$ term is missing because of the required orthogonality. We do not need to know what the functions y_n are, just that they are a complete set of eigenfunctions. Since

$$\mathcal{L} y_{\text{trial}}^{(1)} = \sum_{n=1}^{\infty} \lambda_n b_n y_n \,, \tag{2.34b}$$

it follows by the same method leading to (2.29d), that

$$\Lambda\left[y_{\text{trial}}^{(1)}\right] = \frac{\langle\, y \mid \mathcal{L} y\,\rangle}{\langle\, y \mid y\,\rangle_w} = \frac{\sum_{n=1}^{\infty} \lambda_n |b_n|^2}{\sum_{n=1}^{\infty} |b_n|^2}\,. \tag{2.34c}$$

Since $|b_n|^2 \geqslant 0$ and $\lambda_1 \leqslant \lambda_n$ for $n = 2, \ldots$, we have that

$$\sum_{n=1}^{\infty} \lambda_n |b_n|^2 \geqslant \lambda_1 \sum_{n=1}^{\infty} |b_n|^2\,, \tag{2.34d}$$

and hence $\Lambda\left[y_{\text{trial}}^{(1)}\right] \geqslant \lambda_1$. Hence we have an upper bound on λ_1 from any trial function orthogonal to y_0, and a reasonably good guess will give a good estimate for λ_1.

However, there is an obvious problem: how do we find a trial function orthogonal to y_0 if we only have an approximation to y_0? In general we cannot, in which case we are reduced to trying trial functions that are orthogonal to the approximation.

Nevertheless, there are exceptions. For example, a theorem in quantum mechanics states that the ground-state wavefunction of a particle in a symmetric potential, $V(x) = V(-x)$, is a symmetric function. Since any antisymmetric function is orthogonal to any symmetric function, we can find a [true] bound on λ_1 by
09/19 choosing any antisymmetric trial function.
09/20
09/21

3 Laplace's and Poisson's Equations

3.1 Physical Origins

3.1.1 Poisson's equation

The second-order partial differential equation

$$\nabla^2\Psi = \rho(\mathbf{x})\,, \tag{3.1}$$

is known as *Poisson's equation.* It arises in many different physical contexts.

Remark. There are different sign conventions, so do not be surprised to encounter the alternative $\nabla^2\Psi = -\rho$.

3.1.2 Laplace's equation

The special case where the 'source term', $\rho(\mathbf{x})$, in (3.1) is zero everywhere, or everywhere except in some specific regions or at some particular points, is known as *Laplace's equation*:

$$\nabla^2\Psi = 0\,. \tag{3.2}$$

3.1.3 Diffusion equation

Let $u(\mathbf{x}, t)$ be a scalar quantity that diffuses (e.g. the concentration of a solute in a solution, or the temperature in a heat-conducting medium). The flux, $\mathbf{F}$, of u is given by $\mathbf{F} = -\kappa\boldsymbol{\nabla}u$, where $\kappa > 0$ is the *diffusivity* or diffusion constant, where the minus sign arises because solute diffuses from high to low concentrations, and heat flows from hot to cold. When u is the temperature, κ is also known as the coefficient of heat conductivity.

In many simple cases the diffusion process is governed by the *diffusion equation* (as derived last term):

$$\frac{\partial u}{\partial t} = \kappa\nabla^2 u\,. \tag{3.3a}$$

More generally, if there are sources or sinks $S(\mathbf{x})$, e.g. sources of solute or heat,

$$\frac{\partial u}{\partial t} = \kappa\nabla^2 u + S(\mathbf{x})\,. \tag{3.3b}$$

Steady States. If the distribution is in steady-state, i.e. if $\frac{\partial u}{\partial t} = 0$, then (3.3b) reduces to Poisson's equation:

$$\nabla^2 u = -\frac{S(\mathbf{x})}{\kappa}\,. \tag{3.4}$$

If $S(\mathbf{x}) = 0$, this simplifies further to Laplace's equation.

3.1.4 Electrostatics

In the absence of a magnetic field, Maxwell's equations for a static electric field $\mathbf{E}(\mathbf{x})$ state that

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \rho_q(\mathbf{x})/\epsilon_0\,, \qquad \boldsymbol{\nabla}\times\mathbf{E} = 0\,, \tag{3.5a}$$

where $\rho_q(\mathbf{x})$ is the charge density distribution,[5] and ϵ_0 is the constant permittivity.

Because $\mathbf{E}$ is irrotational, i.e. $\boldsymbol{\nabla}\times\mathbf{E} = 0$, we can write $\mathbf{E} = -\boldsymbol{\nabla}\Phi$, where $\Phi(\mathbf{x})$ is the electric potential. It follows from (3.5a) that the potential satisfies Poisson's equation

$$\nabla^2\Phi = -\rho_q(\mathbf{x})/\epsilon_0\,, \tag{3.5b}$$

where, by convention, there is a *minus sign.* In a region where there is no electric charge, this reduces to Laplace's equation, $\nabla^2\Phi = 0$.

Remark. In the absence of currents, a static magnetic field $\mathbf{B}(\mathbf{x})$ satisfies $\boldsymbol{\nabla}\times\mathbf{B} = 0$ and $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$, and so there is also a magneto-static potential ψ satisfying $\nabla^2\psi = 0$.

[5] There are very many different physical quantities conventionally labelled by ρ, so we include the subscript q for clarity.

3.1.5 Gravitation

In a similar way it can be shown, using Gauss's law for gravity, that the gravitational potential $\Phi(\mathbf{x})$ satisfies Poisson's equation,

$$\nabla^2\Phi = 4\pi G\rho_m(\mathbf{x})\,, \tag{3.6}$$

where $\rho_m(\mathbf{x})$ is the mass density distribution and G is the gravitational constant. Note that by convention there is *no minus sign* in this equation.

3.1.6 Schrödinger's equation

Schrödinger's equation in quantum mechanics has a similar form and reduces to Poisson's equation in certain circumstances.

3.1.7 Ideal fluid flow

The flow of a fluid can be described by a vector field for the fluid's velocity $\mathbf{u}(\mathbf{x}, t)$. Suppose that we make the simplification (that is more honoured in the breach than in the observance) that

(i) the flow is not subject to viscous forces (i.e. is 'inviscid' or 'ideal');

(ii) the inviscid flow is irrotational, i.e. $\boldsymbol{\nabla}\times\mathbf{u} = 0$; in which case the flow can be described by a velocity potential Φ, where

$$\mathbf{u} = \boldsymbol{\nabla}\Phi\,. \tag{3.7a}$$

The flow must also satisfy the conservation of mass equation, or 'continuity' equation,

$$\frac{D\rho}{Dt} = -\rho\boldsymbol{\nabla}\cdot\mathbf{u}\,, \tag{3.7b}$$

where ρ is the fluid density and $\frac{D}{Dt}$ indicates the rate of change following a fluid particle.[6] If we make the further assumption that the fluid is incompressible, so that the density of a fluid particle does not change, i.e.

$$\frac{D\rho}{Dt} = 0\,, \tag{3.7c}$$

then the conservation of mass equation, (3.7b), reduces to

$$\boldsymbol{\nabla}\cdot\mathbf{u} = 0\,. \tag{3.7d}$$

From (3.7a) it follows that the velocity potential for the irrotational flow of an ideal incompressible fluid satisfies Laplace's equation

$$\nabla^2\Phi = 0\,. \tag{3.7e}$$

Such flow is known as *potential flow.*

3.2 Separation of Variables for Laplace's Equation

Superposition. First recall that, because Laplace's equation $\nabla^2\Psi = 0$ is linear in Ψ, the superposition of any two (or more) solutions is another solution; i.e. if $\nabla^2\Psi_1 = 0 = \nabla^2\Psi_2$ then $\nabla^2\Psi_3 = 0$ if $\Psi_3 = \alpha_1\Psi_1 + \alpha_2\Psi_2$ for constants α_1 and α_2

General solution. We then *claim* that the general solution can be written as a linear combination of some set of *basis* solutions (cf. the infinite set of *basis* eigenfunctions of a Sturm-Liouville operator).

Remark. The number of solutions is infinite and so the space of solutions can be viewed as an infinite-dimensional vector space.

[6] By convention we write $\frac{D}{Dt}$, but if you feel happier, then imagine the derivative as $\frac{d}{dt}$.

Separation of variables. In some orthogonal systems of coordinates, separation of variables can provide a method to find a useful set of basis solutions. For instance, you have previously used separation of variables in Cartesian coordinates to solve Laplace's equation. In particular, in Cartesian coordinates the Laplacian is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,, \tag{3.8}$$

and you have considered solutions of the factorized form

$$\Psi(x, y, z) = X(x)Y(y)Z(z)\,. \tag{3.9}$$

The general solution is then be written as a linear superposition of these solutions. Question 1 on Example Sheet 2 provides some revision of this.

The 'right' co-ordinate system. For any given problem, choosing a basis set appropriately, according to the symmetry of the problem, can often lead to the solution in a simpler form, e.g. only a few of the basis set may be needed. This involves choosing an appropriate coordinate system. For example, for a spherically-symmetric source in infinite space we expect spherical polar coordinates to be most useful, whereas for the flow of air around a very long cylinder, cylindrical polar coordinates may be more appropriate.

We will consider plane polar coordinates (equivalent to cylindrical polars with no z-dependence) and spherical polar coordinates with cylindrical symmetry.

3.2.1 Plane polar coordinates

Let (r, ϕ) denote plane polar coordinates, where $x = r\cos\phi$ and $y = r\sin\phi$. The expression for ∇^2 in plane polar coordinates was covered last term, from which it follows that Laplace's equation is given by

$$\nabla^2\Psi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \Psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\Psi}{\partial\phi^2} = 0\,, \tag{3.10}$$

where $\Psi(r, \phi)$ is a scalar field.

Remark. This is the same equation as for cylindrical polar coordinates (r, ϕ, z) when $\frac{\partial\Psi}{\partial z} = 0$.

Consider separable solutions of the form $\Psi(r, \phi) = R(r)\Phi(\phi)$ then from (3.10)

$$\frac{\Phi}{r}\frac{d}{dr}\left(r\frac{dR}{dr}\right) + \frac{R}{r^2}\frac{d^2\Phi}{d\phi^2} = 0\,. \tag{3.11a}$$

Rearranging we obtain

$$\underbrace{\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right)}_{\text{A function of } r \text{ only}} = \underbrace{-\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2}}_{\text{A function of } \phi \text{ only}}\,. \tag{3.11b}$$

The only way that a function of r can equal a function of ϕ is if they are both a constant, say λ. Then the equation for $\Phi(\phi)$ yields

$$\Phi'' = -\lambda\Phi\,, \tag{3.12a}$$

with solution

$$\Phi = \begin{cases} a_0 + b_0\,\phi & \text{for } \lambda = 0 \\ a_\lambda \cos\sqrt{\lambda}\phi + b_\lambda \sin\sqrt{\lambda}\phi & \text{for } \lambda \neq 0 \end{cases}\,. \tag{3.12b}$$

In many cases Ψ corresponds to some physical quantity, e.g. the concentration of solute or the temperature, and must be periodic, i.e. $\Psi(r, \phi) = \Psi(r, \phi + 2\pi)$. However, in other situations, e.g. electrostatics, Ψ is not a physical quantity just a potential and $\boldsymbol{\nabla}\Psi$ must be periodic. We will allow for the more general case by requiring $\Phi'(\phi) = \Phi'(\phi + 2\pi)$. It follows from (3.12b) that

$$2\pi\sqrt{\lambda} = 2\pi n\,, \quad \text{i.e.} \quad \lambda = n^2\,, \quad n \in \mathbb{Z}\,. \tag{3.12c}$$

Hence

$$\Phi = \begin{cases} a_0 + b_0\phi & \text{for } n = 0 \\ a_n \cos n\phi + b_n \sin n\phi & \text{for } n \neq 0 \end{cases}. \tag{3.12d}$$

Returning to the equation for $R(r)$, and using $\lambda = n^2$,

$$\frac{r}{R}\frac{d}{dr}\left(r\frac{dR}{dr}\right) = n^2$$

i.e.

$$r^2 R'' + rR' - n^2 R = 0\,. \tag{3.13a}$$

This is a dimensionally-homogeneous second-order ODE with solution (e.g. try $R = r^k$ or substitute $v = \ln r$)

$$R = \begin{cases} c_0 \ln r + d_0 & \text{for } n = 0 \\ c_n r^n + d_n r^{-n} & \text{for } n \neq 0 \end{cases}. \tag{3.13b}$$

Combining R and Φ from (3.13b) and (3.12b) respectively, we obtain

$$\Psi = R\Phi = \begin{cases} (c_0 \ln r + d_0)(a_0 + b_0\phi) & \text{for } n = 0 \\ (c_n r^n + d_n r^{-n})(a_n \cos n\phi + b_n \sin n\phi) & \text{for } n \neq 0 \end{cases}. \tag{3.14a}$$

At this point it is convenient to relabel the arbitrary constants, e.g. $A_0 = a_0 d_0$, $B_0 = b_0 d_0$, $C_0 = c_0 a_0$, $D_0 = c_0 b_0$, etc., and re-write the general solution as

$$\Psi = A_0 + B_0\phi + C_0 \ln r + \sum_{n=1}^{\infty}\left(A_n r^n + C_n r^{-n}\right)\cos n\phi + \sum_{n=1}^{\infty}\left(B_n r^n + D_n r^{-n}\right)\sin n\phi\,. \tag{3.14b}$$

We have set $D_0 = 0$, *after* relabelling, in order to *exclude* the $\phi \ln r$ combination because it does not satisfy the periodicity requirement on $\boldsymbol{\nabla}\Psi$. With some further relabelling this can be rewritten more compactly as

$$\Psi = A_0 + B_0\phi + C_0 \ln r + \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} r^n \left(\alpha_n \cos n\phi + \beta_n \sin n\phi\right)\,. \tag{3.14c}$$

Remark. Note: a common mistake is to retain too many arbitrary constants, e.g. in

$$\sum c_n r^n (a_n \cos n\phi + b_n \sin n\phi)\,, \tag{3.15}$$

it might seem that there are three arbitrary constants for each value of n, but in reality there are two independent arbitrary quantities, i.e. $c_n a_n$ and $c_n b_n$.

We now consider two examples (see also questions 2, 3 and 4 on Example Sheet 2).

Steady-state temperature distribution in a cylinder.

An infinitely-long cylinder of radius a is heated on its boundary as illustrated. The steady-state temperature $T(r,\phi)$ for $r < a$ satisfies

$$\nabla^2 T = 0\,, \tag{3.16a}$$

with boundary conditions,

$$T(a,\phi) = \begin{cases} -T_0 & \text{for } -\pi < \phi < 0 \\ +T_0 & \text{for } \quad 0 < \phi < \pi \end{cases}. \tag{3.16b}$$

Physics tells us that the temperature T must be single-valued and finite at $r = 0$, hence, considering the first form of the general solution, (3.14b),

$$B_0 = C_0 = 0\,, \quad C_n = D_n = 0 \quad \text{for } n \geqslant 1. \tag{3.16c}$$

It follows that

$$T(r,\phi) = A_0 + \sum_{n=1}^{\infty} r^n \left(A_n \cos n\phi + B_n \sin n\phi\right) , \tag{3.17a}$$

in which case at $r = a$,

$$T(a,\phi) = A_0 + \sum_{n=1}^{\infty} a^n \left(A_n \cos n\phi + B_n \sin n\phi\right) . \tag{3.17b}$$

This is just a Fourier series and so, applying the boundary conditions, the coefficients are given by

$$\begin{aligned}
A_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} T(a,\phi)\, d\phi = 0 && \text{by anti-symmetry of } T,\\
a^n A_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} T(a,\phi)\, \cos n\phi\, d\phi = 0 && \text{by anti-symmetry of } T,\\
a^n B_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} T(a,\phi)\, \sin n\phi\, d\phi \\
&= \frac{1}{\pi}\int_{0}^{\pi} T_0\, \sin n\phi\, d\phi - \frac{1}{\pi}\int_{-\pi}^{0} T_0\, \sin n\phi\, d\phi \\
&= \begin{cases} 4T_0/(n\pi) & n \text{ odd} \\ 0 & n \text{ even (by anti-symmetry about } \pm\pi/2) \end{cases} .
\end{aligned}$$

The final solution is therefore (with $n = 2m-1$, since n is odd)

$$T(r,\phi) = \frac{4T_0}{\pi} \sum_{m=1}^{\infty} \frac{r^{2m-1}}{(2m-1)a^{2m-1}} \sin\left((2m-1)\phi\right) . \tag{3.18}$$

10/19
10/20
10/21 *Two-dimensional fluid flow past a circular cylinder.*

Consider the two-dimensional steady flow of an incompressible, ideal (i.e. irrotational and non-viscous) fluid past a circular barrier of radius a. Assume that the fluid has constant velocity $\mathbf{U} = U\hat{\mathbf{x}}$ at infinity.

As discussed in § 3.1.7 on page 35, the velocity field $\mathbf{u}(r,\phi)$ of such a flow can be described by a velocity potential Φ, where $\mathbf{u} = \boldsymbol{\nabla}\Phi$ and $\nabla^2\Phi = 0$.

As $|\mathbf{x}| \to \infty$, we require $\boldsymbol{\nabla}\Phi \to U\hat{\mathbf{x}}$ and so,

$$\Phi \to Ux = Ur\cos\phi\,. \tag{3.19}$$

Considering the general solution (3.14b), since the physical quantity is $\boldsymbol{\nabla}\Phi$, without loss of generality we can assume $A_0 = 0$. Then the boundary condition at infinity, (3.19), implies[7] that in (3.14b)

$$A_1 = U\,, \quad A_n = 0 \quad \text{for } n \geqslant 2, \tag{3.20a}$$

$$B_n = 0 \quad \text{for } n \geqslant 0, \tag{3.20b}$$

$$C_0 = 0\,. \tag{3.20c}$$

Therefore

$$\Phi(r,\phi) = Ur\cos\phi + \sum_{n=1}^{\infty} r^{-n} \left(C_n \cos n\phi + D_n \sin n\phi\right) . \tag{3.20d}$$

On the surface of the cylinder, the flow must be in a tangential direction, i.e. there must be no radial component of velocity into the cylinder; hence

$$0 = \mathbf{u}(a,\phi)\cdot\widehat{\mathbf{r}} = \frac{\partial\Phi}{\partial r}(a,\phi) = U\cos\phi - \sum_{n=1}^{\infty} n a^{-(n+1)} \left(C_n \cos n\phi + D_n \sin n\phi\right) . \tag{3.21a}$$

[7] The reason that B_0 and C_0 are zero is, in reality, more subtle, in that we are also requiring that there is no swirl/circulation and that there is no net flux of fluid in to or out of a closed surface surrounding the cylinder.

This must be true for all ϕ, and we know that $\cos n\phi$ and $\sin n\phi$ are linearly independent, i.e. if $\sum(\alpha_n \cos n\phi + \beta_n \sin n\phi) = 0$ then $\alpha_n = \beta_n = 0$ for all n. Hence, from equating coefficients of $\cos n\phi$ and $\sin n\phi$ it follows that[8]

$$C_1 = Ua^2, \quad D_1 = 0, \quad \text{and} \quad C_n = D_n = 0 \quad \text{for } n \geqslant 2. \tag{3.21b}$$

Therefore

$$\Phi(r,\phi) = U\cos\phi\left(r + \frac{a^2}{r}\right) = \mathbf{U}\cdot\mathbf{r}\left(1 + \frac{a^2}{r^2}\right). \tag{3.22a}$$

Using $\boldsymbol{\nabla} r = \mathbf{r}/r$ and $\boldsymbol{\nabla}(\mathbf{U}\cdot\mathbf{r}) = \mathbf{U}$,

$$\mathbf{u}(r,\phi) = \left(1 + \frac{a^2}{r^2}\right)\mathbf{U} - \frac{2a^2}{r^4}(\mathbf{U}\cdot\mathbf{r})\mathbf{r}. \tag{3.22b}$$

3.2.2 Spherical polar coordinates (axisymmetric case)

Consider spherical polar coordinates, (r, θ, ϕ), where

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta. \tag{3.23}$$

If a function $\Psi(r,\theta,\phi)$ is axisymmetric, then it is independent of ϕ and $\frac{\partial\Psi}{\partial\phi} = 0$. In this case Laplace's equation (3.2) reduces to

$$\nabla^2\Psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right) = 0. \tag{3.24a}$$

Again we seek separable solutions, this time of the form $\Psi(r,\theta) = R(r)T(\theta)$. Then

$$\underbrace{\frac{1}{R}\left(r^2R'\right)'}_{\text{A function of } r \text{ only}} = \underbrace{-\frac{1}{T\sin\theta}\left(T'\sin\theta\right)'}_{\text{A function of } \theta \text{ only}} \tag{3.24b}$$

The only way that a function of r can equal a function of θ is if they are both a constant, say λ. Then the equation for $T(\theta)$ yields

$$\left(T'\sin\theta\right)' = -\lambda T\sin\theta. \tag{3.25a}$$

Setting $u = \cos\theta$, we have that $\frac{d}{d\theta} = -\sin\theta\frac{d}{du}$ and so

$$-\sin\theta\frac{d}{du}\left(-\sin^2\theta\frac{dT}{du}\right) = -\lambda T\sin\theta. \tag{3.25b}$$

Hence T satisfies *Legendre's equation*, as in (1.35a) and (1.35b),

$$\frac{d}{du}\left((1-u^2)\frac{dT}{du}\right) + \lambda T = 0. \tag{3.25c}$$

For well-behaved solutions at $u = \pm 1$, i.e. at the poles $\theta = 0$ and $\theta = \pi$, we require that

$$\lambda = \ell(\ell+1) \quad \text{where } \ell = 0, 1, 2, \ldots. \tag{3.25d}$$

With these values of ℓ, finite series solutions can be found, i.e. the Legendre polynomials $P_\ell(u)$. Recall that these polynomials are normalised so that $P_\ell(1) = 1$ (cf. (1.35e)).

Returning to (3.24b), the equation for $R(r)$ becomes

$$\left(r^2R'\right)' = \lambda R, \tag{3.26a}$$

or equivalently

$$r^2R'' + 2rR' - \ell(\ell+1)R = 0. \tag{3.26b}$$

[8] Alternatively, multiply (3.21a) by $\cos m\phi$ or $\sin m\phi$, integrate ϕ over $[0, 2\pi]$, and use the orthogonality properties of sine and cosine.

This is a dimensionally-homogeneous second-order ODE with solution (e.g. try $R = r^k$ or substitute $v = \ln r$)

$$R = Ar^{\ell} + Br^{-\ell-1}\,. \tag{3.26c}$$

The general regular solution to Laplace's equation in spherical polar coordinates in the axisymmetric case is therefore, after redefining the arbitrary constants,

$$\Psi(r,\theta) = \sum_{\ell=0}^{\infty} \left(A_\ell\, r^{\ell} + B_\ell\, r^{-\ell-1}\right) P_\ell(\cos\theta)\,. \tag{3.27}$$

Remark. In the non-axisymmetric case, a similar analysis gives an extra equation involving ϕ, and the Legendre polynomials are replaced by associated Legendre polynomials (solutions of the associated Legendre equation).

We now consider two examples (see also questions 5 and 6 on Example Sheet 2).

Diffusion of a solute past a solid sphere. Consider an impermeable fixed solid sphere of radius a surrounded by fluid. Suppose that solute diffuses through the fluid, and that at large distances there is a constant flux of solute of magnitude F parallel to the z-axis. What is the steady-state concentration $\Phi(r,\theta,\phi)$ of solute?

The problem is axisymmetric and so

$$\Phi(r,\theta,\phi) \equiv \Phi(r,\theta)\,. \tag{3.28a}$$

The flux of solute is $-\kappa\boldsymbol{\nabla}\Phi$, and far from the sphere we require

$$-\,\kappa\boldsymbol{\nabla}\Phi \to F\hat{\mathbf{z}} \quad \text{as } r\to\infty. \tag{3.28b}$$

Hence as $r\to\infty$, after using (1.35e),

$$\Phi \to -\frac{F}{\kappa}z = -\frac{F}{\kappa}r\cos\theta = -\frac{F}{\kappa}\,r\,P_1(\cos\theta)\,. \tag{3.28c}$$

Considering the general solution (3.27), the boundary condition at infinity implies

$$A_1 = -F/\kappa\,, \quad A_n = 0 \quad \text{for } n \geqslant 2, \tag{3.29a}$$

and thus

$$\Phi(r,\theta) = A_0 - \frac{F}{\kappa}rP_1(\cos\theta) + \sum_{n=0}^{\infty}\frac{B_n}{r^{n+1}}P_n(\cos\theta)\,. \tag{3.29b}$$

Since the sphere is impermeable, there is no radial flux of solute at $r=a$. Hence we require that

$$0 = \hat{\mathbf{r}}\cdot\boldsymbol{\nabla}\Phi\big|_{r=a} = \left.\frac{\partial\Phi}{\partial r}\right|_{r=a} = -\frac{F}{\kappa}P_1(\cos\theta) - \sum_{n=0}^{\infty}\frac{(n+1)B_n}{a^{n+2}}P_n(\cos\theta)\,. \tag{3.29c}$$

This must be true for all θ, and we know that the P_n are linearly independent, i.e. if $\sum \alpha_n P_n(\cos\theta) = 0$ then $\alpha_n = 0$ for all n. It follows that by equating the coefficients of the P_n that

$$B_1 = -\frac{Fa^3}{2\kappa}\,, \quad B_n = 0 \quad \text{for } n=0 \text{ and } n\geqslant 2. \tag{3.29d}$$

Hence the solution for the concentration of solute is

$$\Phi(r,\theta) = A_0 - \frac{F}{\kappa}\left(r + \frac{a^3}{2r^2}\right)\cos\theta\,, \tag{3.29e}$$

where A_0 is still an arbitrary constant (since only the flux, not the concentration, was specified).

The electric potential, Φ, inside and outside a sphere.

Consider a hollow conducting sphere of radius a that has its top hemisphere held at electric potential $\Phi = V_0$, while the bottom hemisphere, separated from the top by an insulating layer, is earthed so that $\Phi = 0$. What is the electric potential, Φ, inside and outside the sphere, assuming that $\Phi \to 0$ as $r \to \infty$.

The problem is axisymmetric with boundary conditions

$$\Phi(a,\theta) = \begin{cases} V_0 & \text{for } 0 < \theta < \pi/2 \\ 0 & \text{for } \pi/2 < \theta < \pi \end{cases} . \tag{3.30}$$

$r \leqslant a$. *Inside* the sphere Φ must be finite at $r = 0$. This implies that in the general solution, (3.27), $B_\ell = 0$ for all $\ell \geqslant 0$, so yielding at $r = a-$

$$\Phi(a,\theta) = \sum_{\ell=0}^{\infty} A_\ell \, a^\ell P_\ell(\cos\theta) . \tag{3.31a}$$

The A_ℓ can be determined using the orthogonality of Legendre polynomials, i.e. from (1.35f),

$$\int_{-1}^{1} P_m(u)\, P_n(u)\, du = \frac{2}{2m+1}\, \delta_{mn} . \tag{3.31b}$$

Hence from multiplying (3.31a) by $P_m(\cos\theta)$, writing $u = \cos\theta$, integrating u from -1 to $+1$, substituting the boundary conditions (3.30), and using (3.31b),

$$\begin{aligned} \tfrac{2}{2m+1} A_m\, a^m &= \int_{-1}^{1} \Phi(a,u)\, P_m(u)\, d\cos\theta \\ &= V_0 \int_0^1 P_m(u)\, du . \end{aligned} \tag{3.31c}$$

The integrals can be evaluated to give,

$$A_0 = \frac{V_0}{2}\,, \quad A_1 = \frac{3V_0}{4a}\,, \quad A_2 = 0\,, \quad A_3 = -\frac{7V_0}{16a^3}\,, \quad \dots\,, \tag{3.31d}$$

and so inside the sphere,

$$\Phi(r,\theta) = \frac{V_0}{2}\left[1 + \frac{3r}{2a} P_1(\cos\theta) - \frac{7r^3}{8a^3} P_3(\cos\theta) + \dots\right] . \tag{3.31e}$$

11/19

11/20 $r \geqslant a$. *Outside* the sphere we require that $\Phi \to 0$ as $r \to \infty$, and so $A_\ell = 0$ for all $\ell \geqslant 0$. Using a similar argument to the above, the boundary conditions imply that,

$$\tfrac{2}{2m+1} B_m\, a^{-m-1} = V_0 \int_0^1 P_m(u)\, du\,, \tag{3.32a}$$

and so outside the sphere,

$$\Phi(r,\theta) = \frac{V_0\, a}{2r}\left[1 + \frac{3a}{2r} P_1(\cos\theta) - \frac{7a^3}{8r^3} P_3(\cos\theta) + \dots\right] . \tag{3.32b}$$

11/21

3.3 Uniqueness of Solutions of Poisson's Equation

If we have a solution to Poisson's equation (or Laplace's equation) how do we know whether it is the only solution, i.e. the unique, non-fake, solution?

Consider solutions to Poisson's equation,

$$\nabla^2 \Phi = \rho(\mathbf{r})\,, \tag{3.33}$$

in a volume V with boundary surface S. In order to find a [unique] solution we require boundary conditions.

Dirichlet boundary conditions. One possible choice of boundary conditions is to specify $\Phi(\mathbf{r})$ on S, e.g.

$$\Phi(\mathbf{r}) = f(\mathbf{r}) \quad \text{for some } f \text{ defined on } S. \tag{3.34a}$$

Such boundary conditions on Φ, rather than $\boldsymbol{\nabla}\Phi$, are referred to as *Dirichlet boundary conditions*. Physically we then have a well-defined problem. For instance:

(i) in electrostatics the surface of a conductor might be specified to be at a constant potential (cf. the example of the potential specified on the surface of a sphere on page 41);

(ii) if the source of heat in some volume and the temperature on the boundary is specified, this should uniquely determine the temperature in the volume.

However, can we prove that the problem is well-defined? To this end, suppose that it is not the case and that there are two solutions $\Phi_1(\mathbf{r})$ and $\Phi_2(\mathbf{r})$. Let $\Psi \equiv \Phi_1 - \Phi_2$ then from (3.33)

$$\nabla^2\Psi = \nabla^2\Phi_1 - \nabla^2\Phi_2 = \rho - \rho = 0 \quad \text{in } V, \tag{3.34b}$$

and from (3.34a)

$$\Psi = f - f = 0 \quad \text{on } S. \tag{3.34c}$$

Now consider

$$\begin{aligned}\boldsymbol{\nabla}\cdot(\Psi\boldsymbol{\nabla}\Psi) &= \boldsymbol{\nabla}\Psi\cdot\boldsymbol{\nabla}\Psi + \Psi\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\Psi)\\ &= |\boldsymbol{\nabla}\Psi|^2 + \Psi\nabla^2\Psi\\ &= |\boldsymbol{\nabla}\Psi|^2 \qquad\qquad \text{using (3.34b).}\end{aligned}$$

Therefore, using the divergence theorem and the boundary condition (3.34c) that $\Psi = 0$ on S,

$$\int_V |\boldsymbol{\nabla}\Psi|^2\, dV = \int_V \boldsymbol{\nabla}\cdot(\Psi\boldsymbol{\nabla}\Psi)dV = \oint_S \Psi(\boldsymbol{\nabla}\Psi\cdot\mathbf{n})\, dS = 0\,, \tag{3.34d}$$

where $\mathbf{n}$ is the unit outward normal to S. On the assumption that Ψ is a continuous function, the integral on the left-hand side can only be zero if $\boldsymbol{\nabla}\Psi = 0$ everywhere, i.e. if Ψ is a constant in V. However, from (3.34c), $\Psi = 0$ on S and so $\Psi = 0$ throughout V. This means that $\Phi_1 = \Phi_2$, i.e. *there is a unique solution.*

Neumann boundary conditions. A similar theorem holds for *Neumann boundary conditions*, where we specify the normal gradient of Φ on S, i.e.

$$\frac{\partial\Phi}{\partial n} \equiv \mathbf{n}\cdot\boldsymbol{\nabla}\Phi = f(\mathbf{r}) \quad \text{on } S. \tag{3.35}$$

In this case solutions are unique up to a constant (see also question 7 on Example Sheet 2).

Remark. As a physical example, in electrostatics we could specify $\mathbf{E} = -\boldsymbol{\nabla}\Phi$ on S.

3.4 The Green's Function and the Fundamental Solution

The aim of the next few sections is to derive a general solution to Poisson's equation.

Dirichlet boundary conditions. We start by considering solutions of Poisson's equation with Dirichlet boundary conditions.

Definition. Define the *Green's function*, $G(\mathbf{r},\mathbf{r}')$, for Poisson's equation in a volume V with Dirichlet boundary conditions given on a boundary surface S, to be the solution to

$$\nabla^2_{\mathbf{r}} G(\mathbf{r},\mathbf{r}') = \delta^{(3)}(\mathbf{r}-\mathbf{r}') \qquad \text{for } \mathbf{r} \text{ in } V, \tag{3.36a}$$

$$G(\mathbf{r},\mathbf{r}') = 0 \qquad \text{for } \mathbf{r} \text{ on } S, \tag{3.36b}$$

where $\delta^{(3)}(\mathbf{r}-\mathbf{r}') = \delta(x-x')\delta(y-y')\delta(z-z')$ is the three-dimensional Dirac delta function satisfying

$$\int_V f(\mathbf{r})\,\delta^{(3)}(\mathbf{r}-\mathbf{r}')\, dV = \begin{cases} f(\mathbf{r}') & \text{for } \mathbf{r}' \in V,\\ 0 & \text{for } \mathbf{r}' \notin V.\end{cases} \tag{3.36c}$$

Definition. If V is all of space (the limit of a sphere with radius $\to \infty$), the Green's function is known as the *fundamental solution.*

Symmetry. It is possible to prove that a real Green's function is symmetric, i.e.

$$G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r}', \mathbf{r})\,. \tag{3.36d}$$

We saw this in (1.43c) for a one-dimensional Green's function written as a sum over eigenfunctions, and can check it is true in the examples we consider.

Interpretation. We can think of G as the potential due to a point charge at $\mathbf{r}'$. The physical interpretation of (3.36d) is that the potential at $\mathbf{r}$ due to a source at $\mathbf{r}'$ is the same as the potential at $\mathbf{r}'$ due to a source at $\mathbf{r}$.

Neumann boundary conditions. If instead the solution to Poisson's equation is required to satisfy Neumann boundary conditions on S, instead of $G = 0$ on S, i.e. (3.36b), we require,

$$\frac{\partial G}{\partial n} = \frac{1}{A} \qquad \text{on } S, \tag{3.36e}$$

where $A = \oint_S dS$ is the surface area of the boundary. When $A \to \infty$ the condition becomes $\frac{\partial G}{\partial n} = 0$ on S.

Remark. One might have expected the Neumann boundary condition for G to be $\frac{\partial G}{\partial n} = 0$ on S, however there is a necessary 'compatibility' condition on the surface average $\frac{\partial G}{\partial n}$. Specifically, from using the divergence theorem and the governing equation for G, (3.36a), it follows that

$$\begin{aligned} \int_S \frac{\partial G}{\partial n} dS \equiv \int_S \boldsymbol{\nabla} G \cdot \mathbf{n}\, dS &= \int_V \nabla^2 G\, dV \\ &= \int_V \delta^{(3)}(\mathbf{r} - \mathbf{r}')\, dV \\ &= 1\,. \end{aligned} \tag{3.36f}$$

The boundary condition (3.36e) is consistent with the compatibility condition (3.36f).

3.4.1 The fundamental solution in three dimensions

First consider the case when $\mathbf{r}' = 0$, so that there is a point source at the origin, and consider Dirichlet boundary conditions. Then

$$\nabla^2 G = \delta^{(3)}(\mathbf{r}) \quad \text{with} \quad G \to 0 \text{ as } |\mathbf{r}| \to \infty\,. \tag{3.37a}$$

The problem has spherical symmetry and so we will assume that G is a function of r only; if we can find a solution then it will be *the* unique solution from § 3.3.

Using the expression for ∇^2 in spherical polar coordinates, it follows from (3.24a) that for $r \neq 0$

$$\left(r^2 G'\right)' = 0 \quad \Rightarrow \quad G = A + \frac{C}{r}\,, \tag{3.37b}$$

for some constants A and C. From the boundary condition (3.36b) applied at infinity, we require $A = 0$. To determine C we integrate $\nabla^2 G$ over a sphere of radius ϵ centred on the origin. Using the divergence theorem,

$$\int_{r<\epsilon} \nabla^2 G\, dV = \oint_{r=\epsilon} \boldsymbol{\nabla} G \cdot \mathbf{n}\, dS = \oint_{r=\epsilon} \frac{\partial G}{\partial r}\, dS = -\frac{C}{\epsilon^2} \oint_{r=\epsilon} dS = -4\pi C\,. \tag{3.37c}$$

However, no matter how small ε, from (3.37a)

$$\int_{r<\varepsilon} \nabla^2 G\, dV = \int_{r<\varepsilon} \delta^{(3)}(\mathbf{r})\, dV = 1\,. \tag{3.37d}$$

Hence we require $C = -1/4\pi$, so giving

$$G = -\frac{1}{4\pi |\mathbf{r}|}\,. \tag{3.37e}$$

Shifting the origin to $\mathbf{r}'$, we obtain the solution

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|}\,. \tag{3.37f}$$

Remark. This is also the fundamental solution for Neumann boundary conditions.

Example. An electron (a point source with charge $-e$) is located at $\mathbf{r}_0$ and $\Phi \to 0$ as $|\mathbf{r}-\mathbf{r}_0| \to \infty$. What is the electrostatic potential?

The charge distribution is $\rho_q(\mathbf{r}) = -e\,\delta^{(3)}(\mathbf{r}-\mathbf{r}_0)$. So from (3.5b), the potential satisfies

$$\nabla^2\Phi = -\frac{\rho_q}{\epsilon_0} = \left(\frac{e}{\epsilon_0}\right)\delta^{(3)}(\mathbf{r}-\mathbf{r}_0)\,. \tag{3.38a}$$

The solution is therefore the fundamental solution multiplied by e/ϵ_0:

$$\Phi(\mathbf{r}) = -\frac{e}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}_0|}\,. \tag{3.38b}$$

3.4.2 The fundamental solution in two dimensions

Again, we start by considering the case $\mathbf{r}'=0$, so that

$$\nabla^2 G = \delta^{(2)}(\mathbf{r})\,. \tag{3.39a}$$

The problem has circular symmetry and so we assume that G is a function of r only. Then, using the expression for ∇^2 in plane polar coordinates, i.e. (3.10), it follows that for $r \neq 0$

$$(rG')' = 0 \quad \Rightarrow \quad G = A + C\ln r\,. \tag{3.39b}$$

for some constants A and C. Unlike the case of three dimensions it is now not possible to apply the Dirichlet boundary condition (3.36b) at infinity, i.e. the only solution that has $G \to 0$ as $|\mathbf{r}| \to \infty$ is the trivial solution $G=0$. Instead, we require that G vanishes on some circle of radius R, or $|\boldsymbol{\nabla} G| \to 0$ as $|\mathbf{r}| \to \infty$ (which fixes neither A nor C).

To determine C we integrate $\nabla^2 G$ over a circle of radius $\varepsilon > 0$ centred on the origin. Using the [two-dimensional] divergence theorem and (3.39b), it follows that

$$\int_{r<\varepsilon} \nabla^2 G\, dA = \oint_{r=\varepsilon} \boldsymbol{\nabla} G\cdot\mathbf{n}\, dl = \oint_{r=\varepsilon} \frac{\partial G}{\partial r}\, dl = \frac{C}{\varepsilon}\oint_{r=\varepsilon} dl = 2\pi C\,. \tag{3.39c}$$

However, no matter how small ε, from (3.39a)

$$\int_{r<\varepsilon} \nabla^2 G\, dA = \int_{r<\varepsilon} \delta^{(2)}(\mathbf{r})\, dA = 1\,. \tag{3.39d}$$

Hence

$$C = \tfrac{1}{2\pi}\,, \tag{3.39e}$$

giving

$$G = \tfrac{1}{2\pi}\ln|\mathbf{r}| + A\,. \tag{3.39f}$$

Reassuringly, as $r \to \infty$, $G'(r) \to 0$ although $G(r) \to \infty$. Shifting the origin to $\mathbf{r}'$ we obtain,

$$G(\mathbf{r},\mathbf{r}') = \tfrac{1}{2\pi}\ln|\mathbf{r}-\mathbf{r}'| + A\,. \tag{3.39g}$$

Example. Consider an infinite line of charges lying along the z-axis with infinitesimal width and a charge density of μ per unit length. What is the electrostatic potential?

This is, *de facto*, a two-dimensional problem with a charge distribution is $\rho_q(\mathbf{r}) = \mu\,\delta(x)\delta(y)$. We have from (3.5b) that the potential satisfies

$$\nabla^2\Phi = -\frac{\mu}{\epsilon_0}\delta(x)\delta(y)\,. \tag{3.40a}$$

Hence the solution, up to an arbitrary constant, is

12/19
12/20
12/21

$$\Phi(\mathbf{r}) = -\frac{\mu}{2\pi\epsilon_0}\ln\sqrt{x^2+y^2}\,. \tag{3.40b}$$

3.5 The Method of Images

So far we have sought fundamental solutions, i.e. Green's functions, G, when V is all of space. We can use the *method of images* to find G in some other simple geometries. We illustrate this method by a number of examples.

Uniqueness. Recall that *if* we find a solution that satisfies Poisson's equation and the boundary conditions, from the uniqueness of solutions it must be *the* solution (up to a constant if we have purely Neumann boundary conditions).

3.5.1 Three-dimensional half-space

Dirichlet boundary conditions. What is the Green's function for a domain D with *Dirichlet boundary conditions*, where D is the half-space of $\mathbb{R}^3$ with $z > 0$?

The Green's function satisfies,

$$\nabla^2 G = \delta^{(3)}(\mathbf{r}-\mathbf{r}') \qquad \mathbf{r}\in D, \tag{3.41a}$$

$$G = 0 \qquad \text{on } z = 0\,, \tag{3.41b}$$

$$G \to 0 \qquad \text{as } |\mathbf{r}|\to\infty,\ \mathbf{r}\in D. \tag{3.41c}$$

The uniqueness of solutions allows us to solve using a trick: remove the boundary at $z = 0$, consider all of space and add a point source of *opposite sign*, an '*image source*', at the image point $\mathbf{r}'' = (x', y', -z')$. The new Green's function is then required to satisfy, throughout $\mathbb{R}^3$,

$$\nabla^2 G = \delta^{(3)}(\mathbf{r}-\mathbf{r}') - \delta^{(3)}(\mathbf{r}-\mathbf{r}'')\,. \tag{3.41d}$$

The solution to (3.41d), by superposition of two fundamental solutions of the form (3.37f), gives

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi|\mathbf{r}-\mathbf{r}'|} + \frac{1}{4\pi|\mathbf{r}-\mathbf{r}''|}\,. \tag{3.41e}$$

This satisfies the boundary condition $G = 0$ on $z = 0$, i.e. (3.41b), either by geometry, or because when $z = 0$

$$|\mathbf{r}-\mathbf{r}'| = \sqrt{(x-x')^2 + (y-y')^2 + z'^2} = |\mathbf{r}-\mathbf{r}''|\,. \tag{3.41f}$$

The solution (3.41e) also satisfies the other two requirements when $\mathbf{r} \in D$, i.e. (3.41a) and (3.41c). Therefore, by uniqueness,

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}''|}\right)\,. \tag{3.41g}$$

Neumann boundary conditions. Suppose instead we impose *Neumann boundary conditions* at $z = 0$, i.e.

$$\frac{\partial G}{\partial n} = -\frac{\partial G}{\partial z} = 0 \quad \text{on } z = 0\,. \tag{3.42a}$$

Suppose that we all still require $G \to 0$ as $|\mathbf{r}| \to \infty$ for $\mathbf{r} \in D$. Then in order to satisfy (3.42a), we need a point charge of the *same sign* at the image point so that the Green's function is

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{1}{|\mathbf{r}-\mathbf{r}''|}\right)\,. \tag{3.42b}$$

As a check we note that when $\mathbf{r} \in D$, (3.42b) satisfies (3.41a) and (3.41c). It also satisfies (3.42a) because of a change in sign of z' between $\mathbf{r}'$ and $\mathbf{r}''$:

$$\begin{aligned}\frac{\partial}{\partial z}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right)_{z=0} &= \frac{z'}{[(x-x')^2+(y-y')^2+z'^2]^{3/2}}\\ &= -\frac{\partial}{\partial z}\left(\frac{1}{|\mathbf{r}-\mathbf{r}''|}\right)_{z=0}.\end{aligned}$$

3.5.2 Two-dimensional quarter-plane

What is the Green's function, say $G(\mathbf{r},\mathbf{r}_0)$, when the domain D is the quarter plane of $\mathbb{R}^2$ with $x>0$ and $y>0$, and G satisfies Dirichlet boundary conditions?

We require that

$$\begin{aligned}
\nabla^2 G &= \delta^{(2)}(\mathbf{r}-\mathbf{r}_0) && \mathbf{r}\in D, && (3.43a)\\
G &= 0 && \text{on } x=0 \text{ and } y=0, && (3.43b)\\
G &\to 0 && \text{as } |\mathbf{r}|\to\infty,\ \mathbf{r}\in D. && (3.43c)
\end{aligned}$$

In this case it turns out we need three image charges as shown in the figure, $\mathbf{r}_1$ and $\mathbf{r}_2$ with strength -1 and $\mathbf{r}_3$ with strength $+1$. This gives, from (3.39g),

$$\begin{aligned}
G(\mathbf{r},\mathbf{r}_0) &= +\tfrac{1}{2\pi}\ln|\mathbf{r}-\mathbf{r}_0| - \tfrac{1}{2\pi}\ln|\mathbf{r}-\mathbf{r}_1| \\
&\quad -\tfrac{1}{2\pi}\ln|\mathbf{r}-\mathbf{r}_2| + \tfrac{1}{2\pi}\ln|\mathbf{r}-\mathbf{r}_3| + B\\
&= \tfrac{1}{2\pi}\ln\frac{|\mathbf{r}-\mathbf{r}_0|\,|\mathbf{r}-\mathbf{r}_3|}{|\mathbf{r}-\mathbf{r}_1|\,|\mathbf{r}-\mathbf{r}_2|}\,, \qquad (3.43d)
\end{aligned}$$

where the constant B is zero from the boundary conditions on $x=0$ and on $y=0$, since by geometry

$$\begin{aligned}
|\mathbf{r}-\mathbf{r}_0| &= |\mathbf{r}-\mathbf{r}_1| \quad\text{and}\quad |\mathbf{r}-\mathbf{r}_2| = |\mathbf{r}-\mathbf{r}_3| \quad\text{on}\quad x=0\,,\\
|\mathbf{r}-\mathbf{r}_0| &= |\mathbf{r}-\mathbf{r}_2| \quad\text{and}\quad |\mathbf{r}-\mathbf{r}_1| = |\mathbf{r}-\mathbf{r}_3| \quad\text{on}\quad y=0.
\end{aligned}$$

3.5.3 Heat source in a three-dimensional half-space bounded by a constant temperature wall

Suppose that T_0 is both the temperature of the boundary, S, at $z=0$, and the ambient temperature as $|r|\to\infty$. Suppose also that there is a point heat source of strength Q at $\mathbf{r}_0$. What is the total heat flux across the wall?

For $z>0$, the temperature T satisfies, from (3.4),

$$\nabla^2 T = -\frac{Q}{\kappa}\,\delta(\mathbf{r}-\mathbf{r}_0)\,. \qquad (3.44a)$$

Hence, the solution that satisfies Poisson's equation and the boundary conditions is

$$T(\mathbf{r}) = T_0 - \frac{Q}{\kappa}\,G(\mathbf{r},\mathbf{r}_0)\,, \qquad (3.44b)$$

where G is the Green's function, (3.41g), for a three-dimensional half-space with Dirichlet boundary conditions.

From (3.44b), the total heat flux across the wall, say S, is

$$F = -\kappa\int_S \boldsymbol{\nabla}T\cdot\mathbf{n}\,dS = Q\int_S \boldsymbol{\nabla}G\cdot\mathbf{n}\,dS\,. \qquad (3.44c)$$

This could be evaluated directly, but there is a neater way. As $|r|\to\infty$, it follows from (3.41g) that $G\propto r^{-2}$, and $\boldsymbol{\nabla}G\propto r^{-3}$. Hence $\int_H \boldsymbol{\nabla}G\cdot\mathbf{n}\,dS$, where the surface, H, is a hemisphere of radius R, scales as R^{-1}, i.e. tends to zero as $R\to\infty$. It follows that, as expected on physical grounds,

$$\begin{aligned}
F &= Q\int_{S+H}\boldsymbol{\nabla}G\cdot\mathbf{n}\,dS = Q\oint\boldsymbol{\nabla}G\cdot\mathbf{n}\,dS && \text{by taking the limit } R\to\infty \text{ and adding zero}\\
&= Q\int_V \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}G\,dV = Q\int_V \nabla^2 G\,dV && \text{by using the divergence theorem}\\
&= Q && \text{by using (3.41a).} \qquad (3.44d)
\end{aligned}$$

Remark. The three-dimensional problem in $z > 0$, where we have a point electric charge Q at $\mathbf{r}_0$ near an earthed plate at $z = 0$ on which $\Phi = 0$, is very similar to the above; in this case

$$\Phi(\mathbf{r}) = -\frac{Q}{\epsilon_0} G(\mathbf{r}, \mathbf{r}_0)\,, \tag{3.44e}$$

13/19 where G is again the Green's function with Dirichlet boundary conditions for the three-dimensional half-space $z > 0$.

3.5.4 Heat source near an insulated wall

Suppose instead that a point heat source of strength Q is located at $\mathbf{r}_0$ near an *insulated* wall at $z = 0$ (i.e. a wall through which no heat can pass), and that $T \to 0$ as $|\mathbf{r}| \to \infty$. What is the temperature distribution for $z > 0$?

As before, the temperature T satisfies

$$\nabla^2 T = -(Q/\kappa)\,\delta(\mathbf{r} - \mathbf{r}_0) \quad \text{for } z > 0. \tag{3.45a}$$

The heat flux condition on $z = 0$,

$$F = -\kappa \boldsymbol{\nabla} T \cdot \mathbf{n} = \kappa \frac{\partial T}{\partial z} = 0\,, \tag{3.45b}$$

is a Neumann boundary condition, and so we introduce an image charge of strength $+Q$ at $\mathbf{r}_1 = (x_0, y_0, -z_0)$. We then find, from (3.42b), that

$$T(\mathbf{r}) = \frac{Q}{4\pi\kappa}\left(\frac{1}{|\mathbf{r} - \mathbf{r}_0|} + \frac{1}{|\mathbf{r} - \mathbf{r}_1|}\right), \tag{3.45c}$$

which satisfies both the boundary condition at $z = 0$ and the requirement that $T \to 0$ as $|\mathbf{r}| \to \infty$.

Exercise. What is the total heat flux across the hemisphere, H, at infinity?

Remarks.

(i) The equivalent problem in electrostatics has

$$E_z = -\frac{\partial \Phi}{\partial z} = 0 \quad \text{at } z = 0. \tag{3.45d}$$

13/20
13/21 (ii) See also question 8 on Example Sheet 2.

3.5.5 Images in a sphere

What is the Green's function, with Dirichlet boundary conditions, for a domain D which is $r < a$ in $\mathbb{R}^3$?

The Green's function satisfies

$$\nabla^2 G = \delta^{(3)}(\mathbf{r} - \mathbf{r}') \qquad \text{for } r < a, \tag{3.46a}$$

$$G = 0 \qquad \text{on } r = a. \tag{3.46b}$$

Claim. The image point, $\mathbf{r}''$, is the [classical] inverse point

$$r'\,r'' = a^2\,, \quad \text{i.e. } \mathbf{r}'' = \frac{a^2}{r'^2}\mathbf{r}'\,. \tag{3.46c}$$

What then needs to be fixed is the strength of the image source, and the strength that works, see (3.46e), is $-\frac{a}{r'}$. Hence, from the fundamental solution (3.37f), the required Green's function is

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi}\left(\frac{1}{|\mathbf{r} - \mathbf{r}'|} - \frac{a}{r'}\frac{1}{|\mathbf{r} - \mathbf{r}''|}\right). \tag{3.46d}$$

Check. The solution (3.46d) satisfies equation (3.46a) for $\mathbf{r} \in D$. It also satisfies that boundary condition (3.46b) because when $\mathbf{r}$ lies on the surface of the sphere, $|\mathbf{r}| = a$,

$$\begin{aligned}
-4\pi G(\mathbf{r}, \mathbf{r}') &= \frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{a}{r'}\frac{1}{|\mathbf{r}-\mathbf{r}''|} && \text{from (3.46d)} \\
&= \frac{1}{\sqrt{a^2 + r'^2 - 2\mathbf{r}\cdot\mathbf{r}'}} - \frac{a}{r'}\frac{1}{\sqrt{a^2 + r''^2 - 2\mathbf{r}\cdot\mathbf{r}''}} && \text{using the cosine rule} \\
&= \frac{1}{\sqrt{a^2 + r'^2 - 2ar'\cos\theta}} - \frac{1}{\frac{r'}{a}\sqrt{a^2 + \frac{a^4}{r'^2} - 2\frac{a^3}{r'}\cos\theta}} && \text{from (3.46c)} \\
&= 0\,. && (3.46e)
\end{aligned}$$

Remark. By symmetry, the same result holds if the domain is instead $r > a$.

3.5.6 Images in a circle

What is the Green's function, with Dirichlet boundary conditions, for a domain D which is $r < a$ in $\mathbb{R}^2$?

The image point is again the inverse point

$$\mathbf{r}'' = \frac{a^2}{r'^2}\mathbf{r}'\,, \tag{3.47a}$$

but in this case we require the image source to have strength -1. From the two-dimensional fundamental solution, (3.39g), it follows that the Green's function is

$$G(\mathbf{r}, \mathbf{r}') = \frac{1}{2\pi}\ln\frac{|\mathbf{r}-\mathbf{r}'|}{|\mathbf{r}-\mathbf{r}''|} + B\,, \quad \text{where} \quad B = -\frac{1}{2\pi}\ln\frac{r'}{a}\,. \tag{3.47b}$$

Here the constant B has been chosen to ensure that $G = 0$ on the circle $r = a$ (see question 9 on Example Sheet 2 for the calculation).

Remark. Further applications of the method of images are possible. For example, to planes inclined at 60° to each other, or to a pair of parallel planes. In the latter example there are an infinite number of image points.

3.6 The Integral Solution of Poisson's Equation

So far we have found solutions of Poisson's equation with point sources. However, Green's functions can be used to find the solution for an arbitrary source distribution. For this we will need *Green's identity.*

Definition. For any smooth functions Φ and Ψ defined in a volume V with surface S, *Green's identity* states that

$$\int_V \left(\Phi\nabla^2\Psi - \Psi\nabla^2\Phi\right)\,dV = \oint_S (\Phi\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Phi)\cdot\mathbf{n}\,dS \equiv \oint_S \left(\Phi\frac{\partial\Psi}{\partial n} - \Psi\frac{\partial\Phi}{\partial n}\right)\,dS\,. \tag{3.48a}$$

Proof. Using the divergence theorem

$$\begin{aligned}
\oint_S (\Phi\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Phi)\cdot\mathbf{n}\,dS &= \int_V \boldsymbol{\nabla}\cdot(\Phi\boldsymbol{\nabla}\Psi - \Psi\boldsymbol{\nabla}\Phi)\,dV \\
&= \int_V \left(\boldsymbol{\nabla}\Phi\cdot\boldsymbol{\nabla}\Psi + \Phi\boldsymbol{\nabla}^2\Psi - \boldsymbol{\nabla}\Psi\cdot\boldsymbol{\nabla}\Phi - \Psi\boldsymbol{\nabla}^2\Phi\right)dV \\
&= \int_V \left(\Phi\nabla^2\Psi - \Psi\nabla^2\Phi\right)dV\,.
\end{aligned}$$

Two dimensions. In two-dimensions, *Green's identity* for a plane surface S bounded by a curve C is

$$\int_S \left(\Phi\nabla^2\Psi - \Psi\nabla^2\Phi\right)\,dA = \oint_C \left(\Phi\frac{\partial\Psi}{\partial n} - \Psi\frac{\partial\Phi}{\partial n}\right)\,dl\,. \tag{3.48b}$$

Integral solution of Poisson's equation. Consider Poisson's equation with Dirichlet boundary conditions:

$$\nabla^2\Phi = \rho(\mathbf{r}) \qquad \text{for } \mathbf{r} \text{ in } V\,, \tag{3.49a}$$
$$\Phi(\mathbf{r}) = f(\mathbf{r}) \qquad \text{for } \mathbf{r} \text{ on } S\,. \tag{3.49b}$$

Into Green's identity, (3.48a), we substitute $\Psi = G$, and use

- the equation (3.36a), $\nabla^2 G = \delta^{(3)}(\mathbf{r}-\mathbf{r}')$ for $\mathbf{r} \in V$
- the boundary condition (3.36b), $G = 0$ for $\mathbf{r} \in S$,
- the equation (3.49a), $\nabla^2\Phi = \rho$ for $\mathbf{r} \in V$,
- the boundary condition (3.49b), $\Phi = f$ for $\mathbf{r} \in S$,

to obtain

$$\int_V \left(\Phi\,\delta^{(3)}(\mathbf{r}-\mathbf{r}') - G\rho\right) dV = \oint_S f\,\boldsymbol{\nabla} G\cdot\mathbf{n}\,dS\,. \tag{3.49c}$$

After rearrangement this yields

$$\Phi(\mathbf{r}') = \int_V \rho(\mathbf{r})\,G(\mathbf{r},\mathbf{r}')\,dV + \oint_S f(\mathbf{r})\,\frac{\partial G}{\partial n}\,dS\,, \tag{3.49d}$$

where the right-hand side consists of known quantities. This is the *integral solution of Poisson's equation* with Dirichlet boundary conditions.

Remark. This expression can also be used to solve Laplace's equation by setting $\rho(\mathbf{r}) = 0$.

All space. If V is all space we can use the fundamental solution, (3.37f), for G; however, we need to check the convergence of the integrals. In the case of $\int_V \rho\,G\,dV$, we will assume, say, that $\rho(\mathbf{r}) = 0$ for all $|\mathbf{r}| > R$ for some R. In the case of $\oint_S G\,\boldsymbol{\nabla}\Phi\cdot\mathbf{n}\,dS$, we no longer have $G = 0$ on S, hence we need to check that

$$\oint_S G\,\boldsymbol{\nabla}\Phi\cdot\mathbf{n}\,dS \;\to\; 0 \quad \text{as} \quad |\mathbf{r}| \to \infty. \tag{3.50a}$$

Since $G \propto r^{-1}$ as $r \to \infty$, this is true if, say, $|\Phi| \propto r^{-1}$ and $|\boldsymbol{\nabla}\Phi| \propto r^{-2}$ as $r \to \infty$, which we can check, *a posteriori*, is consistent with (3.50b). Hence, from (3.37f),

$$\Phi(\mathbf{r}') = \int_{\mathbb{R}^3} \rho(\mathbf{r})\,G(\mathbf{r},\mathbf{r}')\,dV = -\int_{\mathbb{R}^3} \frac{\rho(\mathbf{r})}{4\pi|\mathbf{r}-\mathbf{r}'|}\,dV\,. \tag{3.50b}$$

Electrostatics in $\mathbb{R}^3$. For example, consider a charge distribution, $\rho_q(\mathbf{r})$, that decays rapidly far from the origin. Then, from (3.5b), and accounting for the sign convention,

$$\Phi(\mathbf{r}') = \int_{\mathbb{R}^3} \frac{\rho_q(\mathbf{r})}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}\,dV\,. \tag{3.50c}$$

This can be understood as the superposition of many infinitesimal charge elements.

Remark. For $|\mathbf{r}'| \gg 1$, it follows from (3.50c) that $|\Phi(\mathbf{r}')| \propto 1/r'$ for $r' \gg 1$, and hence that (3.50a) is satisfied.

Neumann boundary conditions. An integral solution of Poisson's equation can also be derived for Neumann boundary conditions. Into Green's identity, (3.48a), we substitute $\Psi = G$, and use

- the equation (3.36a), $\nabla^2 G = \delta^{(3)}(\mathbf{r}-\mathbf{r}')$ for $\mathbf{r} \in V$
- the boundary condition (3.36e), $\frac{\partial G}{\partial n} = \frac{1}{A}$ for $\mathbf{r} \in S$,
- the equation (3.49a), $\nabla^2\Phi = \rho$ for $\mathbf{r} \in V$,
- the Neumann boundary condition $\frac{\partial\Phi}{\partial n} = g(\mathbf{r})$ for $\mathbf{r} \in S$,

to obtain

$$\Phi(\mathbf{r}') = \int_V \rho(\mathbf{r})\,G(\mathbf{r},\mathbf{r}')\,dV + \frac{1}{A}\oint_S \Phi(\mathbf{r})\,dS - \oint_S g(\mathbf{r})\,G(\mathbf{r},\mathbf{r}')\,dS\,. \tag{3.51a}$$

Whoops? Since Φ appears on the right-hand side of (3.51a), this appears *not* to be a solution. However, in the case of Neumann boundary conditions, Φ is only determined up to a constant. Define the mean of Φ on S as

$$\bar{\Phi} = \frac{1}{A}\oint_S \Phi(\mathbf{r})\, dS\,, \tag{3.51b}$$

where $\bar{\Phi}$ is a constant, albeit unknown. It follows from (3.51a) that

$$\varphi(\mathbf{r}') = \Phi(\mathbf{r}') - \bar{\Phi} = \int_V \rho(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, dV - \oint_S g(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, dS\,, \tag{3.51c}$$

is a solution with a known right-hand side; indeed, it is the [unique] solution with a zero mean over the boundary, i.e. $\bar{\varphi} = 0$.

Unbounded space. If V is unbounded as $|\mathbf{r}| \to \infty$, but with a finite interior surface S, then $A \to \infty$. Hence, as long as the surface integral of Φ at 'infinity' is finite, we again have, cf. (3.51c),

$$\Phi(\mathbf{r}') = \int_V \rho(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, dV - \oint_S g(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, dS\,. \tag{3.51d}$$

The electric potential generated by a wire of finite length.

Consider a wire of length $2L$ with charge density μ per unit length, lying along the z-axis from $z = -L$ to $+L$. What is the electric potential Φ?

We have that $\rho_q(\mathbf{r}) = \mu\delta(x)\delta(y)$ for $-L \leqslant z \leqslant L$ and zero otherwise. So from (3.50c), and by making use of the substitution $z - z' = \sqrt{x'^2 + y'^2}\sinh u$,

$$\begin{aligned}\Phi(\mathbf{r}') &= \int_{\mathbb{R}^3} \frac{\rho_q(\mathbf{r})}{4\pi\epsilon_0|\mathbf{r}-\mathbf{r}'|}\, dV \\ &= \int_{-L}^{+L} \frac{\mu}{4\pi\epsilon_0\sqrt{x'^2+y'^2+(z-z')^2}}\, dz \\ &= \frac{\mu}{4\pi\epsilon_0}\left(\sinh^{-1}\frac{L-z'}{\sqrt{x'^2+y'^2}} - \sinh^{-1}\frac{-L-z'}{\sqrt{x'^2+y'^2}}\right).\end{aligned} \tag{3.52a}$$

Relabelling to eliminate the $'$, we have that

$$\Phi(\mathbf{r}) = \frac{\mu}{4\pi\epsilon_0}\left[\sinh^{-1}\frac{L-z}{\sqrt{x^2+y^2}} + \sinh^{-1}\frac{L+z}{\sqrt{x^2+y^2}}\right]. \tag{3.52b}$$

Remark. As $L \to \infty$, we can use the result that $\sinh^{-1} Z \to \ln Z$ as $Z \to \infty$, to recover the expression for the two-dimensional potential around an infinitely long wire, i.e. (3.40b),

$$\Phi \to -\frac{\mu}{2\pi\epsilon_0}\ln\sqrt{x^2+y^2} + \text{const.} \tag{3.52c}$$

Solution of Laplace's equation in three-dimensional half-space.

What is the solution of Laplace's equation in the three-dimensional half-space with $z > 0$, subject to $\Phi = f(x,y)$ on $z = 0$?

We use the integral solution (3.49d) with $\rho = 0$, where V is the half-space and S is the $z = 0$ plane plus the hemisphere at ∞. We neglect the hemisphere at ∞, on the assumption that $\Phi \to 0$ there (an assumption that can be confirmed *a posteriori*). Hence,

$$\Phi(\mathbf{r}') = \oint_S f(\mathbf{r})\frac{\partial G}{\partial n}\, dS = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\frac{\partial G}{\partial z}\, dx\, dy\,. \tag{3.53a}$$

We derived the required Green's function earlier using the method of images; so from (3.41g)

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi}\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} - \frac{1}{|\mathbf{r}-\mathbf{r}''|}\right). \tag{3.53b}$$

We need,

$$\left.\frac{\partial G}{\partial z}\right|_{z=0} = -\frac{1}{4\pi}\frac{\partial}{\partial z}\left\{\frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} - \frac{1}{\sqrt{(x-x')^2+(y-y')^2+(z+z')^2}}\right\}\right|_{z=0}.$$

After some algebra this reduces to

$$\left.\frac{\partial G}{\partial z}\right|_{z=0} = -\frac{z'}{2\pi\left[(x-x')^2+(y-y')^2+z'^2\right]^{3/2}}. \tag{3.53c}$$

Hence

$$\Phi(\mathbf{r}') = \frac{z'}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(x,y)}{\left[(x-x')^2+(y-y')^2+z'^2\right]^{3/2}}\,dx\,dy\,. \tag{3.53d}$$

Remark. This example is relevant for a range of different problems, e.g. the steady-state temperature distribution with a wall at a specified temperature, the electric potential with a conducting wall held at a given potential, and the steady-state concentration of a solute with a wall kept at a specified concentration.

More example. See questions 10-13 on Example Sheet 2 for some more examples.

14/19
14/20
14/21

4 Cartesian Tensors

4.1 Vectors

We start be reviewing vectors since a *vector* is a particular example of a tensor, a *first-order* tensor.

Remark: number of dimensions. We will consider tensors in three dimensions, although most of the discussion carries over straightforwardly to n dimensions. Hence, free suffices, as in v_i, are assumed to range through $(1, 2, 3)$ without the need to explicitly say so.

A vector has a physical meaning (direction and magnitude) *independent* of the coordinate system used. However, we often think of a vector as a set of components (v_1, v_2, v_3) with respect to some coordinate system. The components will in general be different in other coordinate systems, but the vector will be the same.

For any coordinate system with a basis set of unit vectors $\mathbf{e}_i$, where $i = 1, 2, 3$, we can write a vector $\mathbf{v}$ as

$$\mathbf{v} = v_1\,\mathbf{e}_1 + v_2\,\mathbf{e}_2 + v_3\,\mathbf{e}_3 = \sum_i v_i\,\mathbf{e}_i\,. \tag{4.1a}$$

Notation. We often refer to 'the vector v_i', by which we mean 'the vector with components v_i' for a [known] basis set of unit vectors.

Orthonormal basis. In an *orthonormal coordinate system*, $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, and so, from (4.1a),

$$\mathbf{v}\cdot\mathbf{e}_j = v_j\,. \tag{4.1b}$$

Cartesian Coordinates. We shall consider only *Cartesian coordinate systems*, i.e. orthonormal coordinate systems where the $\mathbf{e}_i$ are *independent of position* (compare with, for instance, cylindrical or spherical polar coordinates).

Summation Convention. Recall that under *Einstein's summation convention* the explicit sum, $\sum$, can be omitted for *dummy suffices*. In particular

- if a suffix appears once it is taken to be a *free suffix* and ranged through, and in the case of an equation the free suffices should be identical on each side of an equation;
- if a suffix appears twice it is taken to be a *dummy suffix* and summed over;
- if a suffix appears more than twice *in one term* of an equation, ***something has gone wrong*** (unless there is an explicit sum).

Examples. Under suffix notation and the summation convention

$$\begin{array}{rcl}
\mathbf{a}+\mathbf{b}=\mathbf{c} & \text{can be written as} & a_i + b_i = c_i\,,\\
(\mathbf{a}\cdot\mathbf{b})\mathbf{c}=\mathbf{d} & \text{can be written as} & a_i\,b_i\,c_j = d_j\,,\\
\left((\mathbf{a}\cdot\mathbf{b})\mathbf{c}-(\mathbf{a}\cdot\mathbf{c})\mathbf{b}\right)_j & \text{can be written as} & a_i\,b_i\,c_j - a_k\,c_k\,b_j\,,\\
 & \text{or can be written as} & a_i\,b_i\,c_j - a_i\,c_i\,b_j\,,\\
 & \text{or can be written as} & a_i(b_i\,c_j - c_i\,b_j)\,.
\end{array}$$

Under suffix notation the following equations make no sense

$$\begin{array}{ll}
a_k = b_j & \text{because the free suffices are different,}\\
\left((\mathbf{a}\cdot\mathbf{b})\mathbf{c}\right)_i = a_i\,b_i\,c_i & \text{because } i \text{ appears more than twice on the right-hand side.}
\end{array}$$

Under suffix notation the following equation is problematical (and probably best avoided unless you will always remember to double count the i on the right-hand side)

$n_i\,n_i = n_i^2$ because i occurs twice on the left-hand side and only once on the right-hand side.

We shall use the summation convention unless otherwise stated.

Transforming between bases. Using the summation convention we can rewrite (4.1a) as

$$\mathbf{v} = v_j\,\mathbf{e}_j\,. \tag{4.2a}$$

If we use a different set of Cartesian coordinates with unit vectors $\mathbf{e}'_i$, then from (4.1a) and (4.1b)

$$\mathbf{v} = v'_i\,\mathbf{e}'_i\,, \qquad v'_i = \mathbf{e}'_i\cdot\mathbf{v}\,. \tag{4.2b}$$

But we also have from (4.2a) that

$$\mathbf{e}'_i\cdot\mathbf{v} = \mathbf{e}'_i\cdot\mathbf{e}_j\,v_j = L_{ij}v_j\,, \tag{4.2c}$$

where we define the *transformation matrix* L by

$$L_{ij} \equiv \mathbf{e}'_i\cdot\mathbf{e}_j\,. \tag{4.2d}$$

Hence, from (4.2b) and (4.2c),

$$v'_i = L_{ij}v_j \qquad \text{or} \qquad v' = \mathsf{L}v\,, \tag{4.2e}$$

where v' and v are column vectors with components v'_i and v_i respectively.

Remark. It may look like we are rotating the vector but remember that v' and v both represent the same vector.

Relationship between basis vectors. Suppose that we consider the components of the [new] $\mathbf{e}'_i$ basis vectors with respect to the [old] $\mathbf{e}_i$ basis vectors, then from (4.1a) or equivalently (4.2a), (4.1b) and (4.2d),

$$\mathbf{e}'_i = (\mathbf{e}'_i\cdot\mathbf{e}_j)\,\mathbf{e}_j = L_{ij}\mathbf{e}_j\,. \tag{4.3}$$

Reversing the argument, i.e. by interchanging the primed and non-primed bases and components

$$\begin{aligned} v_j &= \mathbf{e}_j\cdot\mathbf{v} && \text{from (4.1b)}\\ &= \mathbf{e}_j\cdot(v'_i\mathbf{e}'_i) && \text{from (4.2b)}\\ &= L_{ij}v'_i && \text{from (4.2d)}, \end{aligned} \tag{4.4a}$$

i.e.

$$v_j = L^{\mathrm{T}}_{ji}v'_i \qquad \text{or} \qquad v = \mathsf{L}^{\mathrm{T}}v'\,. \tag{4.4b}$$

Combining with (4.2e) we have that

$$v = \mathsf{L}^{\mathrm{T}}v' = \mathsf{L}^{\mathrm{T}}\mathsf{L}v \qquad \text{or} \qquad v' = \mathsf{L}v = \mathsf{L}\mathsf{L}^{\mathrm{T}}v'\,, \tag{4.4c}$$

and so

$$\mathsf{L}^{\mathrm{T}}\mathsf{L} = \mathsf{I} = \mathsf{L}\mathsf{L}^{\mathrm{T}}\,, \tag{4.4d}$$

where I is the identity matrix/tensor. Hence the transformation matrix L is an *orthogonal matrix*.

Definition. A Cartesian *vector* $\mathbf{v}$ is a set of coefficients v_i, defined with respect to a set of orthonormal basis vectors $\mathbf{e}_i$, such that the coefficients v'_i with respect to another orthonormal basis $\mathbf{e}'_i$ are given an orthogonal transformation of the form (4.2d) and (4.2e), i.e.

$$v'_i = L_{ij}v_j \quad \text{where} \quad L_{ij} = \mathbf{e}'_i\cdot\mathbf{e}_j\,. \tag{4.5}$$

Example: $\boldsymbol{\nabla}$. Consider the differential operator,

$$\boldsymbol{\nabla} = \mathbf{e}_i\,\partial_i\,, \quad \text{where} \quad \partial_i \equiv \frac{\partial}{\partial x_i}\,. \tag{4.6a}$$

Is this a vector? Since from (4.4a), $x_j = L_{kj}\,x'_k$, and L does not depend on $\mathbf{x}'$, we have that

$$\frac{\partial x_j}{\partial x'_i} = \frac{\partial(L_{kj}x'_k)}{\partial x'_i} = L_{kj}\frac{\partial x'_k}{\partial x'_i} = L_{kj}\delta_{ki} = L_{ij}\,. \tag{4.6b}$$

Thus, using the chain rule,

$$\frac{\partial}{\partial x'_i} = \frac{\partial x_j}{\partial x'_i}\frac{\partial}{\partial x_j} = L_{ij}\frac{\partial}{\partial x_j}\,. \tag{4.6c}$$

Therefore, from the definition (4.5), $\boldsymbol{\nabla}$ is a vector.

Remark. For more general straight-line coordinate systems, $\mathsf{L}^{\mathrm{T}} \neq \mathsf{L}^{-1}$, and then $\boldsymbol{\nabla}$ is *not* a vector; one has to distinguish between vectors and 'co-vectors', but there is no such distinction for Cartesian coordinates.

4.1.1 Axial-vectors

An orthogonal matrix has determinant ± 1:

- those with $\det(\mathsf{L}) = 1$ are rotation matrices ('*proper rotations*'),
- those with $\det(\mathsf{L}) = -1$ are the composition of a rotation with a reflection in some plane ('*improper rotations*').

For example the matrices

$$\begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{4.7}$$

are a rotation about the z-axis, and a reflection in the origin, and have determinant $+1$ and -1 respectively.

Consider a transformation from basis $\mathbf{e}_i$ to $\mathbf{e}'_i$ and then to $\mathbf{e}''_i$. The components of a vector transform as

$$v'_i = \mathsf{L}^{(1)}_{ij} v_j , \qquad v''_i = \mathsf{L}^{(2)}_{ij} v'_j \qquad \Rightarrow \qquad v''_i = \mathsf{L}^{(2)}_{ij} \mathsf{L}^{(1)}_{jk} v_k . \tag{4.8}$$

Therefore, if $\mathsf{L}^{(1)}$ and $\mathsf{L}^{(2)}$ are proper rotations, so is the composite transformation $\mathsf{L}^{(2)}\mathsf{L}^{(1)}$, since rotations form a subgroup. This is *not* true for improper rotations, since two reflections yield a rotation.

Definition. A Cartesian *axial-vector* (or *pseudo-vector*), $\boldsymbol{a}$, is a set of coefficients a_i, defined with respect to a set of orthonormal basis vectors $\mathbf{e}_i$, such that the coefficients a'_i with respect to another orthonormal basis $\mathbf{e}'_i$ are given by, cf. (4.5),

$$a'_i = \det(\mathsf{L})\, L_{ij}\, a_j . \tag{4.9}$$

When $\det(\mathsf{L}) = 1$, i.e. we do not change the handedness of the coordinate system, this is the same as for a vector, but it differs in sign when $\det(\mathsf{L}) = -1$.

Example. An example of an axial-vector is the angular momentum, $\mathbf{h} = \mathbf{r} \times \mathbf{p}$, of a particle with momentum $\mathbf{p}$ at position $\mathbf{r}$.[a] Suppose that in the $\mathbf{e}_i$ basis

$$\mathbf{r} = (r, 0, 0) , \quad \mathbf{p} = (0, p, 0) , \quad \mathbf{h} = (0, 0, rp) . \tag{4.10a}$$

Consider a reflectional transformation in the origin as given by the second matrix in (4.7); then $\mathbf{e}'_1 = -\mathbf{e}_1$, $\mathbf{e}'_2 = -\mathbf{e}_2$ and $\mathbf{e}'_3 = -\mathbf{e}_3$. In the $\mathbf{e}'_i$ basis

$$\mathbf{r} = (-r, 0, 0) , \quad \mathbf{p} = (0, -p, 0) , \tag{4.10b}$$

and, because of the *change in handedness*,

$$\mathbf{h} = (0, 0, rp) . \tag{4.10c}$$

Hence, from (4.9), $\mathbf{h}$ is an axial-vector.

[a] We use $\mathbf{h}$ for angular momentum, rather than $\mathbf{L}$, to avoid confusion with the transformation matrix L.

Passive and active transformations. So far we have considered *passive* transformations, i.e. a change in the coordinate system, without a physical change to the vector. An *active* transformation is a transformation which makes a physical change to the vector, e.g. an active rotation or reflection. Active transformations provide an alternative way of looking at axial-vectors.

Example. Returning to the angular momentum example, consider a *physical* reflectional transformation in the origin (but keeping the same right-handed basis) so that

$$\mathbf{r}' = -\mathbf{r} , \quad \mathbf{p}' = -\mathbf{p} . \tag{4.11}$$

Although $\mathbf{r}$ and $\mathbf{p}$ change sign, the angular momentum does not, consistent with $\mathbf{h}$ being an axial-vector.

Remark. We will see later that the cross product of two vectors is always an axial-vector. Hence, because $\boldsymbol{\nabla}$ is a vector, the curl of a vector is an axial-vector (e.g. the magnetic field $\mathbf{B}$ is an axial-vector since $\partial_t \mathbf{B} = -\boldsymbol{\nabla} \times \mathbf{E}$).

4.2 Tensors

Tensors are a generalisation of vectors. We can think of them as having some physical meaning independent of the coordinate basis and we can measure their components in some coordinate system.

Suppose we have a relation between two vectors, for example the electric current density $\mathbf{J}$ arising when an electric field $\mathbf{E}$ is applied to a material. For an isotropic material where the conductivity is the same in all directions,

$$\mathbf{J} = \sigma \mathbf{E} \tag{4.12a}$$

where σ is the conductivity. However, if the conductivity is instead anisotropic, for example high in one direction and low in the other directions, the generalised relation between current and field is

$$J_i = \sigma_{ij} E_j \,, \tag{4.12b}$$

where the σ_{ij} are the components of the *conductivity tensor* for a given coordinate basis, e.g. for conductor laminated so as to be an insulator in the z-direction

$$\sigma_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{4.12c}$$

In a different basis the components of $\mathbf{J}$, $\mathbf{E}$ and $\mathsf{\sigma}$ will change, say

$$J'_i = \sigma'_{ij} E'_j \,. \tag{4.12d}$$

From the transformation law for vectors, (4.5),

$$J'_i = L_{il} J_l = L_{il} \sigma_{lm} E_m = L_{il} \sigma_{lm} L^{\mathrm{T}}_{mj} E'_j \,, \tag{4.12e}$$

and hence

$$\left(\sigma'_{ij} - L_{il} \sigma_{lm} L^{\mathrm{T}}_{mj} \right) E'_j = 0 \,. \tag{4.12f}$$

Since L and $\mathbf{E}$ are arbitrary, we deduce that, cf. (4.5),

$$\sigma'_{ij} = L_{il} L_{jm} \sigma_{lm} \,, \quad \text{i.e.} \quad \mathsf{\sigma}' = \mathsf{L} \mathsf{\sigma} \mathsf{L}^{\mathrm{T}} \,. \tag{4.12g}$$

15/19 *Remark.* Having two indices, $\mathsf{\sigma}$ is a tensor of order two.

15/20
15/21 *Definition.* A Cartesian *tensor*, T, of order n (also known as a tensor of rank n) is a set of coefficients $T_{i_1 \dots i_n}$, labelled by n indices, defined with respect to a set of orthonormal basis vectors $\mathbf{e}_i$, and such that the coefficients with respect to another orthonormal basis $\mathbf{e}'_i = L_{ij} \mathbf{e}_j$ are given by the *transformation law*,

$$T'_{i_i \dots i_n} = L_{i_1 j_1} \cdots L_{i_n j_n} \, T_{j_1 \dots j_n} \,. \tag{4.13a}$$

Remark. A tensor of order zero is a *scalar*, i.e. a number. A tensor of order one is a *vector*.

Definition. A Cartesian *pseudo-tensor* E of order n is defined in a similar way to a tensor but there is an additional $\det(\mathsf{L})$ factor in the transformation law:

$$E'_{i_i \dots i_n} = \det(\mathsf{L}) \, L_{i_1 j_1} \cdots L_{i_n j_n} \, E_{j_1 \dots j_n} \,. \tag{4.13b}$$

Remarks

(i) When $\det(\mathsf{L}) = 1$, i.e. when there is no change in handedness of the coordinate system, this is the same as for a tensor and so there is no distinction. However, it differs in sign when $\det(\mathsf{L}) = -1$, i.e. for reflections.

(ii) Pseudo-tensors of first order, i.e. axial-vectors, have already been discussed in § 4.1.1.

(iii) A pseudo-scalar, i.e. an order-zero pseudo-tensor, is like a scalar but changes sign under reflections.

4.2.1 The Kronecker delta δ_{ij} and the Levi-Civita symbol ϵ_{ijk}

Kronecker delta. The Kronecker delta, δ_{ij}, has been defined without reference to a frame; i.e. it has been assumed that its components are the same in all frames: $\delta'_{ij} = \delta_{ij}$. We can check that the Kronecker delta is a second-order tensor using the definition (4.13a) and the orthogonality of L from (4.4d):

$$\delta'_{ij} = L_{ip}L_{jq}\delta_{pq} = L_{ip}L_{jp} = \delta_{ij}\,. \tag{4.14}$$

Levi-Civita symbol. Likewise, the Levi-Civita symbol, ϵ_{ijk}, should be the same in all Cartesian coordinate systems. While it has 27 components, it is sufficient to check one of the non-zero components, say $\epsilon_{123} = 1$. If it transformed as a tensor then, from (4.13a) and the definition of the determinant of a 3×3 matrix,

$$\epsilon'_{123} = L_{1p}L_{2q}L_{3r}\epsilon_{pqr} = \det \mathsf{L}\,, \tag{4.15a}$$

However, this would imply that $\epsilon'_{123} = -1$ under a reflection. Instead, the Levi-Civita symbol transforms as a pseudo-tensor since, using the definition (4.13b),

$$\epsilon'_{123} = \det(\mathsf{L})\, L_{1p}L_{2q}L_{3r}\epsilon_{pqr} = (\det \mathsf{L})^2 = 1\,. \tag{4.15b}$$

With this definition, $\epsilon'_{123} = 1$ in all frames as required. Therefore ϵ_{ijk} is an pseudo-tensor of order 3.

Remark. δ_{ij} and ϵ_{ijk} are examples of *isotropic tensors* (see below).

4.2.2 Inertia tensors

Consider a rigid body of variable mass density $\rho(\mathbf{x})$ within a volume V rotating with angular velocity $\boldsymbol{\omega}$.[9] Then the angular momentum of an infinitesimal mass element $dm = \rho(\mathbf{x})dV$ is

$$dm\,\mathbf{x} \times \mathbf{v} = dm\,\mathbf{x} \times (\boldsymbol{\omega} \times \mathbf{x}) = dm\left(|\mathbf{x}|^2\boldsymbol{\omega} - (\boldsymbol{\omega}\cdot\mathbf{x})\mathbf{x}\right)\,. \tag{4.16a}$$

It follows that the total angular momentum $\mathbf{J}$ is given by

$$J_i = \int_V \rho(\mathbf{x})\left(x_kx_k\omega_i - \omega_jx_jx_i\right)\,dV \;=\int_V \rho(\mathbf{x})\left(x_kx_k\delta_{ij} - x_jx_i\right)\omega_j\,dV \;= I_{ij}\omega_j\,, \tag{4.16b}$$

where I is here the *inertia tensor* of the rigid body (not the identity tensor) given by

$$I_{ij} = \int_V \rho(\mathbf{x})\left(x_kx_k\delta_{ij} - x_ix_j\right)dV\,. \tag{4.16c}$$

It can be confirmed that I is a second-order tensor by applying the same arguments that were used to show that the conductivity is a tensor.

Remark. The inertia tensor depends only on the properties of the rigid body, not on $\boldsymbol{\omega}$.

4.2.3 Electric and magnetic susceptibility tensors

Consider an electrical insulator (dielectric) in an external electric field $\mathbf{E}$. No current flows because the charges are not free to move. However, the field does induce an electric polarisation density (dipole moment density), $\mathbf{P}$, given by

$$P_i = \epsilon_0\,\chi_{ij}\,E_j\,, \tag{4.17a}$$

where χ is the *electric susceptibility* tensor. A related quantity is the *molecular polarizability*, α, that gives the dipole moment of a molecule induced by a local electric field:

$$p_i = \epsilon_0\,\alpha_{ij}\,E_j^{\text{local}}\,. \tag{4.17b}$$

The *magnetic susceptibility*, χ^M, is defined in a similar way:

$$M_i = \chi^M_{ij}H_j\,, \tag{4.17c}$$

where $\mathbf{M}$ is the magnetisation (magnetic dipole moment per unit volume) and $\mathbf{H}$ is the magnetic field.

Again, the same arguments that were used to show that the conductivity is a tensor confirm that the *electric susceptibility*, *molecular polarizability* and *magnetic susceptibility* are tensors (in the latter case it is necessary to take into account that both the magnetisation and the magnetic field are pseudo-vectors).

[9] Angular velocities are another example of axial-vectors.

4.2.4 Stress and elastic strain tensors

In an elastic body, a local deformation due to applied forces (stresses) can be described by an elastic *strain tensor*

$$e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \tag{4.18a}$$

where $\mathbf{u}(\mathbf{x})$ is the displacement vector of a small volume element whose unstrained position is $\mathbf{x}$.

The elements σ_{ij} of the *stress tensor*, σ, are defined as the x_j component of forces acting on a plane perpendicular to the x_i axis. A generalisation of Hooke's law (stress $\propto$ strain) gives, for certain materials,

$$\sigma_{ij} = c_{ijkl}e_{kl}\,, \tag{4.18b}$$

where c_{ijkl} is the (fourth-order) stiffness tensor.

4.2.5 Piezo-electric strain tensor

The application of a stress to certain materials produces an electric polarisation that results in an electric field. The electric polarisation density $\mathbf{P}$ is related to the applied stress σ by,

$$P_i = d_{ijk}\sigma_{jk}\,, \tag{4.19}$$

where d_{ijk} is the (third-order) piezo-electric strain tensor.

4.3 Properties of Tensors

Addition. If $A_{i_1\ldots i_n}$ and $B_{i_1\ldots i_n}$ are tensors of order n, and α and β are scalars, then

$$C_{i_1\ldots i_n} = \alpha A_{i_1\ldots i_n} + \beta B_{i_1\ldots i_n} \tag{4.20a}$$

is an order-n tensor.

Proof.

$$\begin{aligned} C'_{i_1\ldots i_n} &= \alpha' A'_{i_1\ldots i_n} + \beta' B'_{i_1\ldots i_n} \\ &= \alpha L_{i_1j_1}\ldots L_{i_nj_n}A_{j_1\ldots j_n} + \beta L_{i_1j_1}\ldots L_{i_nj_n}B_{j_1\ldots j_n} \\ &= L_{i_1j_1}\ldots L_{i_nj_n}\left(\alpha A_{j_1\ldots j_n} + \beta B_{j_1\ldots j_n}\right) \\ &= L_{i_1j_1}\ldots L_{i_nj_n}C_{j_1\ldots j_n}\,. \end{aligned} \tag{4.20b}$$

Outer (or tensor) product. If $A_{i_1\ldots i_n}$ and $B_{i_1\ldots i_m}$ are tensors of order n and m respectively, then

$$C_{i_1\ldots i_n i_{n+1}\ldots i_{n+m}} = A_{i_1\ldots i_n}B_{i_{n+1}\ldots i_{n+m}} \tag{4.21}$$

is a tensor of order $n+m$. The proof is a straightforward application of the transformation law.

Definition. This is called the *outer (or tensor) product* and written $\mathsf{C} = \mathsf{A}\otimes\mathsf{B}$.

Remark. We can write a tensor as $\mathsf{T} = T_{i_1i_2\ldots i_n}\,\mathbf{e}_{i_1}\otimes\mathbf{e}_{i_2}\ldots\otimes\mathbf{e}_{i_n}$.

Pseudo-tensors (again). If instead A is a tensor of order n but B is a pseudo-tensor of order m, then $\mathsf{C} = \mathsf{A}\otimes\mathsf{B}$ is a pseudo-tensor of order $n+m$. The proof is again a straightforward application of the transformation laws.

Contraction (or inner product). If $T_{i_1j\ldots i_m}$ is a tensor of order m, then if we contract two indices (set equal and sum over) we get a tensor of order $m-2$, e.g.

$$S_{i_1\ldots i_{k-1}i_{k+1}\ldots i_{\ell-1}i_{\ell+1}\ldots i_m} = T_{i_1\ldots i_{k-1}i_ki_{k+1}\ldots i_{\ell-1}i_ki_{\ell+1}\ldots i_m}\,. \tag{4.22a}$$

Proof.

$$\begin{aligned} S'_{\alpha_1\ldots\alpha_{k-1}\alpha_{k+1}\ldots\alpha_{\ell-1}\alpha_{\ell+1}\ldots\alpha_m} &= T'_{\alpha_1\ldots\alpha_{k-1}\alpha_k\alpha_{k+1}\ldots\alpha_{\ell-1}\alpha_k\alpha_{\ell+1}\ldots\alpha_m}\,. \\ &= L_{\alpha_1 i_1}\ldots L_{\alpha_k i_k}\ldots L_{\alpha_k i_\ell}\ldots L_{\alpha_m i_m}\; T_{i_1\ldots i_k\ldots i_\ell\ldots i_m} \\ &= (L_{\alpha_k i_k}L_{\alpha_k i_\ell})\,L_{\alpha_1 i_1}\ldots L_{\alpha_m i_m}\; T_{i_1\ldots i_k\ldots i_\ell\ldots i_m} \\ &= \delta_{i_k i_\ell} L_{\alpha_1 i_1}\ldots L_{\alpha_m i_m}\; T_{i_1\ldots i_k\ldots i_\ell\ldots i_m} \\ &= L_{\alpha_1 i_1}\ldots\ldots L_{\alpha_m i_m}\; S_{i_1\ldots i_m} \end{aligned} \tag{4.22b}$$

16/19

Examples.

(i) If T is a tensor of order two then, from the definition of the inner product (4.22a), $T_{ii} = \operatorname{tr}(\mathsf{T})$ is a scalar.

(ii) Suppose $\mathbf{u}$ and $\mathbf{v}$ are vectors, then from the definition of the outer product (4.21), $T_{ij} = u_i v_j$ is a tensor of order two. It follows from (4.22a) that $T_{ii} = u_i v_i = \mathbf{u}\cdot\mathbf{v}$ is a tensor of order zero; hence $\mathbf{u}\cdot\mathbf{v}$ is scalar. ☺

(iii) If A is an order-two tensor and $\mathbf{u}$ is a vector then, from (4.21) and (4.22a), $v_i = A_{ij}u_j$ is a vector, since '`(2+1)-2=1`'.

(iv) If A and B are tensors of order two then, from (4.21) and (4.22a), $A_{ij}B_{jk}$ is a tensor of order two, since '`(2+2)-2=2`'.

(v) The Levi-Civita symbol is used to define the cross product of two vectors:

$$(\mathbf{u}\times\mathbf{v})_i = \epsilon_{ijk}u_j v_k\,. \tag{4.23}$$

16/20 Hence, because ϵ_{ijk} is a third-order pseudo-tensor, this is an axial-vector, since '`(3+1+1)-4=1`'.

16/21

4.4 Symmetric and Antisymmetric Tensors

Definitions. A tensor $T_{ijk\ldots}$ is *symmetric* in a pair of indices α and β if

$$T_{\ldots\alpha\ldots\beta\ldots} = T_{\ldots\beta\ldots\alpha\ldots}\,, \tag{4.24a}$$

and it is *antisymmetric* in α and β if

$$T_{\ldots\alpha\ldots\beta\ldots} = -T_{\ldots\beta\ldots\alpha\ldots}\,. \tag{4.24b}$$

Symmetry/Antisymmetry is invariant. The (anti)symmetry property of pair of indices of a tensor is invariant under a change of coordinate system.

Proof. If $T_{ijk\ldots}$ is symmetric in, say, i and j then

$$\begin{aligned} T'_{ijk\ldots} &= L_{ip}L_{jq}L_{kr}\ldots T_{pqr\ldots} && \text{from (4.13a)} \\ &= L_{ip}L_{jq}L_{kr}\ldots T_{qpr\ldots} && \text{from (4.24a)} \\ &= L_{jq}L_{ip}L_{kr}\ldots T_{qpr\ldots} && \text{rearrange} \\ &= T'_{jik\ldots} && \text{from (4.13a).} \end{aligned} \tag{4.24c}$$

The proof for antisymmetry is similar.

Symmetric/antisymmetric contraction. If $S_{ijk\ldots}$ is symmetric in, say, i and j and $A_{pqr\ldots}$ is antisymmetric in, say, p and q, then

$$S_{ijk\ldots}A_{ijr\ldots} = 0\,. \tag{4.24d}$$

Proof.

$$\begin{aligned} S_{ijk\ldots}A_{ijr\ldots} &= -S_{jik\ldots}A_{jir\ldots} && \text{from (anti-)symmetry} \\ &= -S_{ijk\ldots}A_{ijr\ldots} && \text{from swapping dummy indices.} \end{aligned}$$

Hence

$$2S_{ijk\ldots}A_{ijr\ldots} = 0\,. \tag{4.24e}$$

Remarks.

- The Kronecker delta δ_{ij} is symmetric, and the Levi-Civita symbol ϵ_{ijk} is antisymmetric, in any pair of indices.
- The inertia tensor and strain tensors are symmetric from their definitions (4.16c) and (4.18a) respectively.
- In most situations, but not all, the stress tensor (4.18b) is also symmetric. The conductivity tensor, (4.12b), and susceptibility tensors, e.g. (4.17a), are usually symmetric.

4.5 Second-Order Tensors

Since a second-order tensor only has one pair of indices, if it is symmetric/antisymmetric in these indices we can refer to the tensor as symmetric/antisymmetric. The matrices corresponding to symmetric and antisymmetric second-order tensors are symmetric and antisymmetric respectively:

$$\mathsf{S}^{\mathrm{T}} = \mathsf{S} \qquad \text{if } \mathsf{S} \text{ is symmetric,} \tag{4.25a}$$

$$\mathsf{A}^{\mathrm{T}} = -\mathsf{A} \qquad \text{if } \mathsf{A} \text{ is antisymmetric.} \tag{4.25b}$$

Symmetric/antisymmetric decomposition. Any second-order tensor T_{ij} can be uniquely decomposed into the sum of a symmetric and an antisymmetric tensor:

$$T_{ij} = S_{ij} + A_{ij}\,, \quad \text{where} \quad S_{ij} = \tfrac{1}{2}\left(T_{ij} + T_{ji}\right) \quad \text{and} \quad A_{ij} = \tfrac{1}{2}\left(T_{ij} - T_{ji}\right)\,. \tag{4.26}$$

The equivalence of antisymmetric second-order tensors with axial-vectors. An antisymmetric second-order tensor has only three independent components, and can be written as

$$A_{ij} = -\epsilon_{ijk}\omega_k = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}\,. \tag{4.27a}$$

where $\boldsymbol{\omega}$ is known as the *dual vector*.[10] We can check this by defining

$$\omega_k = -\tfrac{1}{2}\epsilon_{klm}A_{lm}\,, \tag{4.27b}$$

and then noting that

$$\begin{aligned} -\epsilon_{ijk}\omega_k &= \tfrac{1}{2}\epsilon_{ijk}\epsilon_{klm}A_{lm} \\ &= \tfrac{1}{2}\left(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}\right)A_{lm} \\ &= \tfrac{1}{2}\left(A_{ij} - A_{ji}\right) \\ &= A_{ij}\,. \end{aligned} \tag{4.27c}$$

Since ϵ_{ijk} is a pseudo-tensor, $\boldsymbol{\omega}$ must be an axial-vector. So axial-vectors/pseudo-vectors are equivalent to second-order antisymmetric tensors.

Symmetric tensor decomposition. Any symmetric second-order tensor, S, has a unique decomposition in terms of a symmetric traceless tensor, $\widetilde{\mathsf{S}}$, and a scalar multiple of the identity I, i.e.

$$\mathsf{S} = \widetilde{\mathsf{S}} + \tfrac{1}{3}\,\mathrm{tr}(\mathsf{S})\,\mathsf{I}\,, \quad \text{where} \quad \widetilde{\mathsf{S}} = \mathsf{S} - \tfrac{1}{3}\,\mathrm{tr}(\mathsf{S})\,\mathsf{I}, \tag{4.28a}$$

since

$$\mathrm{tr}(\widetilde{\mathsf{S}}) = \mathrm{tr}(\mathsf{S}) - \tfrac{1}{3}\,\mathrm{tr}(\mathsf{S})\ \mathrm{tr}(\mathsf{I}) = 0\,. \tag{4.28b}$$

Using the transformation law for tensors, (4.13a), it can be shown that $\widetilde{\mathsf{S}}$ is traceless in any Cartesian coordinate system.

Example. An elastic body is subject to a simple shear so that the displacement $\mathbf{u}(\mathbf{x})$ at position $\mathbf{x} = (x, y, z)$ is given by $\mathbf{u} = (\gamma y, 0, 0)$ for some constant γ.

We can decompose $\frac{\partial u_i}{\partial x_j}$ into symmetric and antisymmetric parts,

$$\frac{\partial u_i}{\partial x_j} = \begin{pmatrix} 0 & \gamma & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2}\gamma & 0 \\ \frac{1}{2}\gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{2}\gamma & 0 \\ -\frac{1}{2}\gamma & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.29a}$$

where from (4.18a) the symmetric part is just the strain tensor, e_{ij}, and from (4.27a) and (4.27b), the antisymmetric part can be written as $-\epsilon_{ijk}\omega_k$ where $\boldsymbol{\omega} = \left(0, 0, -\frac{1}{2}\gamma\right)$.

This corresponds to writing

$$\mathbf{u} = \left(\tfrac{1}{2}\gamma y, \tfrac{1}{2}\gamma x, 0\right) + \left(\tfrac{1}{2}\gamma y, -\tfrac{1}{2}\gamma x, 0\right)\,, \tag{4.29b}$$

where the first term corresponds to a stretch at 45° to the x and y axes, while the second term is a rotation.

[10] Sign in definition changed on 06/03/21 to be consistent with Question 5 on Example Sheet 3, and convention!

4.5.1 Diagonalisation of symmetric second-order tensors

Suppose that S_{ij} is a symmetric second-order tensor with components relative to a coordinate system with basis vectors $\mathbf{e}_i$ $(i = 1, 2, 3)$. The matrix of components, S, is symmetric, i.e. Hermitian. Hence from the theory of last term, we know that S has three real eigenvalues, λ_i $(i = 1, 2, 3)$, and three orthonormal eigenvectors, $\mathbf{e}'_i$ $(i = 1, 2, 3)$, which, without loss of generality, we arrange in a right-handed set.

Now transform from the coordinate system with basis vectors $\mathbf{e}_i$, to a coordinate system with basis vectors $\mathbf{e}'_i$. From (4.2d), L is the matrix with the components of the eigenvectors $\mathbf{e}'_i$ as rows, i.e.

$$L_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j \,. \tag{4.30a}$$

Hence

$$\mathsf{S}\,\mathsf{L}^{\mathrm{T}} = \mathsf{S}\,(\mathbf{e}'_1 \,|\, \mathbf{e}'_2 \,|\, \mathbf{e}'_3) = (\lambda_1\mathbf{e}'_1 \,|\, \lambda_2\mathbf{e}'_2 \,|\, \lambda_3\mathbf{e}'_3) \,. \tag{4.30b}$$

Further, from (4.13a) or (4.12g),

$$\mathsf{S}' = \mathsf{L}\mathsf{S}\mathsf{L}^{\mathrm{T}} = \begin{pmatrix} {\mathbf{e}'_1}^{\mathrm{T}} \\ \hline {\mathbf{e}'_2}^{\mathrm{T}} \\ \hline {\mathbf{e}'_3}^{\mathrm{T}} \end{pmatrix} \left(\lambda_1\mathbf{e}'_1 \,\middle|\, \lambda_2\mathbf{e}'_2 \,\middle|\, \lambda_3\mathbf{e}'_3 \right) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} , \tag{4.30c}$$

because, for example, the top LHS entry is given by $\mathbf{e}'_1 \cdot \lambda_1\mathbf{e}'_1$, and the top RHS entry is $\mathbf{e}'_1 \cdot \lambda_3\mathbf{e}'_3$. Therefore, we can *diagonalise* S_{ij} with some appropriate rotation of the coordinate axes.

Principal values. The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are known as the *principal values* of the tensor. The principal values do not depend on frame; they are properties of the tensor not the coordinate system.

Principal axes. The Cartesian coordinate axes for which S_{ij} is diagonal, i.e. the eigenvectors of S_{ij}, are known as *principal axes*.

Moments of inertia. The *moments of inertia* are the principal values of the inertia tensor (4.16c). If the origin is taken to be the centre of mass of the rigid body, these are a property of the body itself. All three are different for a generic irregular-shaped body. When two (or more) are equal, the body is said
17/19 to be 'symmetric'.

17/20

Example. Consider a thin circular disk of radius a and total mass M that is centred on the origin and lies in the $x_3 = 0$ plane. What are its moments of inertia?

From (4.16c)

$$\begin{aligned} I_{ij} &= \frac{M}{\pi a^2} \int_V (\delta_{ij} x_k x_k - x_i x_j)\, \delta(x_3)\, dV \\ &= \frac{M}{\pi a^2} \int_{\mathrm{Disk}} (\delta_{ij} x_k x_k - x_i x_j)\, dx_1\, dx_2 \,. \end{aligned} \tag{4.31a}$$

First consider the non-diagonal elements with $i \neq j$. Since $x_3 = 0$ in the disk, $I_{13} = I_{23} = I_{31} = I_{32} = 0$. Further, since by symmetry $\int_{-\alpha}^{\alpha} x_j \, dx_j = 0$ for all α, $I_{12} = I_{21} = 0$. Hence $I_{ij} = 0$ for $i \neq j$, i.e. I is diagonal. The diagonal elements are given by

$$\begin{aligned} I_{11} &= \frac{M}{\pi a^2} \int_{r=0}^{a} \int_{\phi=0}^{2\pi} \left(r^2 - r^2 \cos^2 \phi\right) r \, dr \, d\phi \\ &= \tfrac{1}{4} M a^2 = I_{22} \,, \qquad (4.31b) \\ I_{33} &= \frac{M}{\pi a^2} \int_{r=0}^{a} \int_{\phi=0}^{2\pi} r^3 \, dr \, d\phi = \tfrac{1}{2} M a^2 \,. \qquad (4.31c) \end{aligned}$$

17/21 This shows that the disk is a symmetric body with the x_3-axis as a principal axis. Because $I_{11} = I_{22}$, the other two principal axes may be chosen to be any orthogonal directions in the $x_1 x_2$-plane.

4.6 Isotropic Tensors

Definition. An *invariant tensor* or an *invariant pseudo-tensor* is one which has the same components in all frames, i.e.

$$T'_{ijk\ldots} = T_{ijk\ldots} \,. \qquad (4.32)$$

Invariant tensors and invariant pseudo-tensors are both called *isotropic tensors*.

Remarks.

(i) All scalars are isotropic (from the transformation law for an order-zero tensor).

(ii) There are no non-zero isotropic vectors or isotropic axial-vectors.

(iii) From (4.14) it follows that δ_{ij} is an isotropic tensor. Below we show that the most general second-order isotropic tensor is $\lambda \delta_{ij}$, where λ is a scalar.

(iv) The most general third-order isotropic tensor is $\lambda \epsilon_{ijk}$, where λ is a scalar (see also below).

(v) The most general fourth-order isotropic tensor is

$$\lambda \, \delta_{ij} \delta_{kl} + \mu \, \delta_{ik} \delta_{jl} + \nu \, \delta_{il} \delta_{jk} \,, \qquad (4.33a)$$

where λ, μ and ν are scalars.

(vi) Isotropic tensors do not have any 'preferred' direction. For example, consider the conductivity tensor σ_{ij} of an isotropic medium, i.e. a medium that is the same in all directions. This should be an isotropic tensor, and so $\sigma_{ij} = \lambda \delta_{ij}$. Therefore

$$J_i = \sigma_{ij} E_j = \lambda \delta_{ij} E_j = \lambda E_i \qquad \text{or} \qquad \mathbf{J} = \lambda \mathbf{E} \,. \qquad (4.33b)$$

(vii) Do not confuse *isotropic* (no preferred direction) with *homogeneous* (the same at all points in space, i.e. uniform).

4.6.1 Second-order isotropic tensors

Consider a general tensor T of rank two, with components T_{ij} with respect to some set of axes $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Suppose that T is isotropic. Its components should then be unaltered by rotations.

First consider a rotation, L_1, that sends the basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ to $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) = (\mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_1)$. Since, $(\mathsf{L}_1)_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j = \mathbf{e}_{i+1} \cdot \mathbf{e}_j$ from (4.5), where we identify $\mathbf{e}_4 = \mathbf{e}_1$, it follows that

$$\mathsf{L}_1 = \begin{pmatrix} \mathbf{e}_2 \cdot \mathbf{e}_1 & \mathbf{e}_2 \cdot \mathbf{e}_2 & \mathbf{e}_2 \cdot \mathbf{e}_3 \\ \mathbf{e}_3 \cdot \mathbf{e}_1 & \mathbf{e}_3 \cdot \mathbf{e}_2 & \mathbf{e}_3 \cdot \mathbf{e}_3 \\ \mathbf{e}_1 \cdot \mathbf{e}_1 & \mathbf{e}_1 \cdot \mathbf{e}_2 & \mathbf{e}_1 \cdot \mathbf{e}_3 \end{pmatrix} . = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} . \qquad (4.34a)$$

Then, from (4.13a) or (4.12g)

$$\mathsf{T}' = \mathsf{L}_1 \mathsf{T} \mathsf{L}_1^{\mathrm{T}} \tag{4.34b}$$

i.e.

$$\begin{aligned}\begin{pmatrix} T'_{11} & T'_{12} & T'_{13} \\ T'_{21} & T'_{22} & T'_{23} \\ T'_{31} & T'_{32} & T'_{33} \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} T_{22} & T_{23} & T_{21} \\ T_{32} & T_{33} & T_{31} \\ T_{12} & T_{13} & T_{11} \end{pmatrix}. \end{aligned} \tag{4.34c}$$

Remark. Note that elements of T'_{ij} can be obtained from T_{ij} by cyclically permutating the subscripts of T_{ij} by $1 \to 2 \to 3 \to 1$, as is the case for the basis vectors.

Because T is isotropic, $T'_{ij} = T_{ij}$, so from comparing matrix entries

$$T_{11} = T_{22} = T_{33} \tag{4.34d}$$
$$T_{12} = T_{23} = T_{31} \tag{4.34e}$$
$$T_{21} = T_{32} = T_{13}\,. \tag{4.34f}$$

Next consider a rotation of $\pi/2$ about the x_3-axis, so that $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3) = (\mathbf{e}_2, -\mathbf{e}_1, \mathbf{e}_3)$. The matrix L_2 for this rotation is

$$\mathsf{L}_2 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.34g}$$

T' is given by

$$\begin{aligned}\begin{pmatrix} T'_{11} & T'_{12} & T'_{13} \\ T'_{21} & T'_{22} & T'_{23} \\ T'_{31} & T'_{32} & T'_{33} \end{pmatrix} &= \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} T_{22} & -T_{21} & T_{23} \\ -T_{12} & T_{11} & -T_{13} \\ T_{32} & -T_{31} & T_{33} \end{pmatrix}. \end{aligned} \tag{4.34h}$$

Remark. Note that elements of T'_{ij} can be obtained from T_{ij} by permutating the subscripts such that $1 \to 2$, $2 \to -1$ and $3 \to 3$, as is the case for the basis vectors.

Hence, if $T'_{ij} = T_{ij}$, we have that

$$T_{11} = T_{22}; \tag{4.34i}$$
$$T_{13} = T_{23} = -T_{13} \quad \text{so that} \quad T_{13} = T_{23} = 0; \tag{4.34j}$$
$$T_{31} = T_{32} = -T_{31} \quad \text{so that} \quad T_{31} = T_{32} = 0. \tag{4.34k}$$

Combining the results from the two rotations we conclude that all off-diagonal elements of T are zero, and all diagonal elements are equal, i.e. $T_{ij} = \lambda\delta_{ij}$ for some scalar λ.

In summary, we have shown that *any* isotropic second rank tensor must be equal to $\lambda\delta_{ij}$ for some scalar λ.

4.6.2 Third-order isotropic tensors

Consider a general tensor T of rank three that is isotropic, i.e such that for all rotations

$$T'_{ijk} = L_{ip} L_{jq} L_{kr} T_{pqr} \tag{4.35a}$$

First consider the effect of the rotation matrix L_1 defined by (4.34a). Since the effect of L_1 is to cyclically permute $1 \to 2 \to 3 \to 1$, we conclude that if $T'_{ijk} = T_{ijk}$ then

$$T_{111} = T_{222} = T_{333}\,, \tag{4.35b}$$
$$T_{112} = T_{223} = T_{331}\,, \quad T_{113} = T_{221} = T_{332}, \tag{4.35c}$$
$$T_{122} = T_{233} = T_{311}\,, \quad T_{133} = T_{211} = T_{322}, \tag{4.35d}$$
$$T_{121} = T_{232} = T_{313}\,, \quad T_{131} = T_{212} = T_{323}, \tag{4.35e}$$
$$T_{123} = T_{231} = T_{312}\,, \quad T_{132} = T_{213} = T_{321}. \tag{4.35f}$$

Next consider the effect of the rotation matrix L_2 defined by (4.34g). Then we require for $T'_{ijk} = T_{ijk}$ that

$$T_{111} = T_{222} = -T_{111}\,, \tag{4.35g}$$
$$T_{112} = -T_{221} = -T_{112}\,, \tag{4.35h}$$
$$T_{211} = -T_{122} = -T_{211}\,, \tag{4.35i}$$
$$T_{121} = -T_{212} = -T_{121}\,, \tag{4.35j}$$
$$T_{123} = -T_{213}\,. \tag{4.35k}$$

Hence, from (4.35b) and (4.35g)

$$T_{111} = T_{222} = T_{333} = 0\,, \tag{4.35l}$$

and from (4.35c), (4.35d), (4.35h) and (4.35i)

$$T_{112} = T_{223} = T_{331} = T_{113} = T_{221} = T_{332} = 0\,, \tag{4.35m}$$
$$T_{122} = T_{233} = T_{311} = T_{133} = T_{211} = T_{322} = 0\,, \tag{4.35n}$$
$$T_{121} = T_{232} = T_{313} = T_{131} = T_{212} = T_{323} = 0\,, \tag{4.35o}$$

and from (4.35f) and (4.35k)

$$T_{123} = T_{231} = T_{312} = -T_{132} = -T_{213} = -T_{321}\,. \tag{4.35p}$$

Hence T_{ijk} differs from ϵ_{ijk} by at most a scalar factor, i.e. *any* isotropic third rank tensor must be equal to $\lambda\epsilon_{ijk}$ for some scalar λ.

4.6.3 Application to integrals

We consider four examples.

A vector integral over a sphere. Consider the integral over the sphere (which calculates the position of the centre of mass for a radially symmetric density distribution)

$$\mathbf{X} = \int_{r\leqslant a} \mathbf{x}\,\rho(r)\,dV \quad \text{or} \quad X_i = \int_{r\leqslant a} x_i\,\rho(r)\,dV\,, \quad r^2 = x_i x_i\,. \tag{4.36a}$$

Relabelling the integration variables we can write

$$X_i = \int_{r'\leqslant a} x'_i\,\rho(r')\,dV'\,. \tag{4.36b}$$

Now make the substitution $x'_i = R_{ij}x_j$ for a rotation matrix R. The integration volume and the function ρ are spherically symmetric, so $r' = r$ and $dV' = dV$, and since R is an orthogonal matrix,

$$X_i = R_{ij}\int_{r\leqslant a} x_j\rho(r)\,dV = R_{ij}X_j = X'_i\,. \tag{4.36c}$$

from the definition, (4.5), of a vector. This equation says that $X_i = X'_i$, i.e. that $\mathbf{X}$ is an isotropic vector. But the only isotropic vector is the zero vector, so we deduce that

$$\mathbf{X} = 0\,. \tag{4.36d}$$

Check. Rewrite (4.36c) in matrix notation as $(\mathsf{I} - \mathsf{R})\,X = 0$. This is valid for any rotation matrix, so choose a rotation matrix for which the matrix $(\mathsf{I} - \mathsf{R})$ is invertible; it follows that $\mathbf{X} = 0$.

A second-order tensor integral over a sphere. Consider the integral

$$K_{ij} = \int_{r\leqslant a} x_i x_j \,\rho(r)\, dV\,. \tag{4.37a}$$

A similar argument as above shows that, using the definition of a second-order tensor (4.13a),

$$K_{ij} = R_{ik}R_{jl}K_{kl} = K'_{ij}\,. \tag{4.37b}$$

This means that K is an isotropic tensor, and so

$$K_{ij} = \lambda\delta_{ij}\,, \tag{4.37c}$$

for some scalar λ. Take the trace to deduce that,

$$\lambda = \tfrac{1}{3}\,\mathrm{tr}(\mathsf{K}) = \tfrac{1}{3}\int_{r\leqslant a} r^2\,\rho(r)\,dV\,, \tag{4.37d}$$

and hence that,

$$K_{ij} = \left(\int_{r\leqslant a} \tfrac{1}{3}r^2\rho(r)\,dV\right)\delta_{ij}\,. \tag{4.37e}$$

A second-order tensor integral over all space: unlectured. Calculate the following integral, where V is all of space:

$$K_{ij} = \int_V x_i x_j e^{-r^2}\,dV\,. \tag{4.38a}$$

Using the result derived above

$$K_{ij} = \left(\int_V \tfrac{1}{3}r^2 e^{-r^2}\,dV\right)\delta_{ij}\,. \tag{4.38b}$$

By straightforward integration

$$\int_V r^2 e^{-r^2}\,dV = 4\pi\int_0^\infty r^4 e^{-r^2}\,dr = \tfrac{3}{2}\sqrt{\pi^3}\,. \tag{4.38c}$$

Therefore

$$K_{ij} = \tfrac{1}{2}\sqrt{\pi^3}\,\delta_{ij}\,. \tag{4.38d}$$

An inertia tensor. Another example is the inertia tensor of a sphere of radius a and mass M. The density per unit volume is $\rho = 3M/4\pi a^3$, and hence from (4.16c), (4.37c) and (4.37d)

$$\begin{aligned}
I_{ij} &= \rho\int_V (x_k x_k\delta_{ij} - x_i x_j)\,dV \\
&= \frac{3M}{4\pi a^3}\left(\int_V \left(r^2 - \tfrac{1}{3}r^2\right)dV\right)\delta_{ij} \\
&= \frac{M}{2\pi a^3}\left(\int_{r\leqslant a} r^2\,dV\right)\delta_{ij} \\
&= \frac{2M}{a^3}\int_0^a r^4\,dr\,\delta_{ij} \\
&= \tfrac{2}{5}Ma^2\,\delta_{ij}\,.
\end{aligned} \tag{4.39}$$

4.7 Tensor Fields

Definition. A *tensor field* is a tensor that depends on position $\mathbf{x}$.

Scalar fields. Scalar fields, i.e. zeroth-order tensor fields, include temperature $T(\mathbf{x})$ and concentration $\Phi(\mathbf{x})$; both are just ordinary functions of $\mathbf{x}$.

Vector fields. A vector field is an first-order tensor field, e.g. the electric field $\mathbf{E}(\mathbf{x})$.

Higher-order tensor fields. In a conducting material, where the conductivity varies with position, $\sigma(\mathbf{x})$ is a second-order tensor field.

Other examples. Many of the other examples discussed in § 4.2 can be tensor fields if they vary with position, e.g. susceptibilities, stress and strain tensors, the fourth-order stiffness tensor, and the third-order piezo-electric strain tensor.

4.7.1 Tensor differential operators

We have already seen, in (4.6c), that $\boldsymbol{\nabla}$, with components $\partial_i \equiv \frac{\partial}{\partial x_i}$, is a vector. Using the transformation laws and/or the general properties of tensors discussed in § 4.3, we can deduce the following.

Gradient of a scalar. The gradient of a scalar field, Φ, is a vector field, $\partial_i \Phi$.

Divergence. The divergence of a vector field $\mathbf{F}$ is a scalar field, $\partial_i F_i$, since it is the contraction of a second-order tensor formed from the outer product of two vectors.

Curl. The curl of a vector field $\mathbf{F}$ is an axial-vector field, $\epsilon_{ijk}\partial_j F_k$, since it is the double contraction of a fifth-order pseudo-tensor formed from the outer product of a third-order pseudo-tensor with two vectors.

Laplacian. The Laplacian acting on a scalar field Φ is a scalar field, $\boldsymbol{\nabla}\Phi = \partial_i \partial_i \Phi$, since it is the contraction of a second-order tensor formed from the outer product of two vectors.

Derivative of a second-order tensor. The derivative of a second-order tensor, σ, is a third-order tensor field, $\partial_i \sigma_{jk}$.

18/19 *Remark.* Remember that ∂_i is an operator, and so the *order is important.*
18/20
18/20

5 Contour Integration

This section explores properties of analytic functions and functions that contain singularities, and leads to a new method for doing integrals such as

$$\int_0^{2\pi} \frac{d\theta}{2(a-\cos\theta)} \qquad \text{and} \qquad \int_0^{\infty} \frac{dx}{1+x^4}\,. \tag{5.1}$$

In § 6 these methods will be applied to Fourier transforms.

5.1 Analytic Functions of a Complex Variable

Recall from last term that requiring a function of a complex variable, $z = x + iy$, to be differentiable imposes a strong constraint, i.e. that

$$\frac{df}{dz} \equiv f'(z) = \lim_{\delta z \to 0} \frac{f(z+\delta z) - f(z)}{\delta z}\,, \tag{5.2}$$

must be finite and the limit must be the *same* when $\delta z \to 0$ in *any* direction in the complex plane.

Definition. A function is *analytic* in some region of the complex plane if $f'(z)$ exists and is continuous in that region. If $f'(z)$ exists and is continuous at z_0, and is analytic in some region, however small, around z_0, then $f(z)$ is analytic at z_0 .

Definition. A function that is analytic everywhere in the complex plane is termed an *entire function*

Remark. We will see below that if $f(z)$ is analytic at $z = z_0$ then *all* of its derivatives exist at z_0, not just $f'(z)$, and it has a Taylor series,

$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n\,, \qquad a_n = \frac{f^{(n)}(z_0)}{n!}\,, \tag{5.3}$$

which converges within some neighbourhood of z_0.

Examples.

(i) $f(z) = e^z$ is analytic everywhere in the complex plane, i.e. it is an entire function.

(ii) $f(z) = 1/z$ is analytic everywhere except at $z = 0$ where there is a simple pole.

(iii) $f(z) = |z|^2$ is not analytic.

(iv) $f(z) = \sqrt{z}$ is not analytic at the origin since not all the derivatives exist at the origin; indeed,
18/18 $\sqrt{z}$ is not single-valued (see below).

Cauchy-Riemann equations. Separating the function $f(z)$ into real and imaginary parts,

$$f(z) = u(x,y) + i\,v(x,y)\,, \tag{5.4a}$$

the requirement that the limit in (5.2) must be the *same* when $\delta z \to 0$ in *any* direction in the complex plane, leads to the *Cauchy-Riemann equations*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}\,, \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}\,. \tag{5.4b}$$

Unlectured remark. Suppose that f is a [not necessarily analytic] function of the two variables x and y, i.e. $f \equiv f(x,y)$. We may, alternatively, view f as a function of the two variables $z = x+iy$ and $z^* = x-iy$, i.e. $f \equiv f(z,z^*)$. Further, since

$$x = \tfrac{1}{2}(z+z^*)\,, \quad y = -\tfrac{1}{2}i\,(z-z^*)\,, \tag{5.4c}$$

it follows from the chain rule that

$$\left.\frac{\partial}{\partial z}\right|_{z^*} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\,\frac{\partial}{\partial y}\right) \quad \text{and} \quad \left.\frac{\partial}{\partial z^*}\right|_{z} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\,\frac{\partial}{\partial y}\right)\,. \tag{5.4d}$$

If now we suppose that f is analytic, then it is only a function of z, and so

$$\frac{1}{2}\left(\frac{\partial f}{\partial x}+i\frac{\partial f}{\partial y}\right)=\left.\frac{\partial f}{\partial z^*}\right|_z=0\,. \tag{5.4e}$$

If $f=u+iv$, then taking the real and imaginary parts of (5.4e) recovers the Cauchy-Riemann equations (5.4b).

Definition. Last term you saw that if $f(z)=u+iv$ is analytic, the Cauchy-Riemann equations (5.4b) imply that both u and v satisfy Laplace's equation in two dimensions:

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0\,,\quad \frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0\,. \tag{5.4f}$$

Functions such as u and v are said to be *harmonic*. See also Question 1 on Example Sheet 4.

5.2 Contour Integrals

Consider an integral,

$$\int_C f(z)\,dz\,, \tag{5.5}$$

from $z=\alpha$ to β in the complex plane. We need to specify the path or *contour*, C, along which we integrate.

For example, consider the integral $\int \frac{1}{z}\,dz$ from $z=-1$ to $z=+1$ along paths around half the unit circle

(i) clockwise (above the real axis),

(ii) anticlockwise (below the real axis).

Making the substitution $z=e^{i\theta}$, $dz=ie^{i\theta}d\theta$,

$$I_1=\int_{C_1}\frac{1}{z}\,dz=\int_\pi^0\frac{ie^{i\theta}}{e^{i\theta}}d\theta=\int_\pi^0 i\,d\theta=-i\pi\,, \tag{5.6a}$$

and

$$I_2=\int_{C_2}\frac{1}{z}\,dz=\int_\pi^{2\pi} i\,d\theta=+i\pi\,. \tag{5.6b}$$

Important remark. The result of a contour integration *may depend on the contour.*

Definition. As for a standard line integral, a contour integral can be formally defined by introducing a partition of an interval $[\alpha,\beta]$ of the form $\alpha=z_0<z_1<z_2<\cdots<z_N=\beta$, that divides C into small intervals $\delta z_k=z_k-z_{k-1}$ for $k=1,\ldots,N$. Then we define,

$$\int_C f(z)\,dz=\lim_{\Delta\to 0}\sum_{k=0}^{N-1}f(z_k)\,\delta z_k\,, \tag{5.7}$$

where $\Delta=\max\limits_k|\delta z_k|$ and $N\to\infty$ as $\Delta\to 0$.

Properties. The following properties follow from this definition.

- If C_1 is a contour from α to β, C_2 is a contour from β to γ, and C is C_1 followed by C_2, then

$$\int_C f(z)\,dz=\int_{C_1}f(z)\,dz+\int_{C_2}f(z)\,dz\,. \tag{5.8a}$$

- If C_+ is a contour from α to β and C_- is the contour in reverse, then
$$\int_{C_+} f(z)\,dz = -\int_{C_-} f(z)\,dz\,. \tag{5.8b}$$
- If C is a closed contour, then it does not matter where we start on C. However, if we reverse the direction of C the integral changes sign, i.e.
$$\oint_{\text{anticlockwise}} f(z)\,dz = -\oint_{\text{clockwise}} f(z)\,dz\,. \tag{5.8c}$$
- From (5.8a), we can split a closed contour into two parts:
$$\oint_C f(z)\,dz = \int_{C_1} f(z)\,dz + \int_{C_2} f(z)\,dz\,. \tag{5.8d}$$
- If the contour has length L then,
$$\left|\int_C f(z)\,dz\right| \leqslant L\,\max_C |f(z)|\,. \tag{5.8e}$$
- We can use integration by parts and substitution.

5.3 Cauchy's Theorem

Definition. A *simply-connected domain* (SCD) is a region R of the complex plane without any holes; any closed curve in R encircles points which are only in R.

Definition. A *simple closed curve* (SCC) is a continuous closed curve of finite length that does not intersect itself; it divides the complex plane into an *interior* region and an *exterior* region.

Theorem. Cauchy's theorem states that if a function $f(z)$ is analytic in a simply-connected domain R, then for any simple closed curve C in R,
$$\oint_C f(z)\,dz = 0\,. \tag{5.9a}$$

Proof. First recall that two-dimensional version of the divergence theorem, i.e. Green's theorem in a plane, for a vector $\mathbf{p} = (p, q)$ states that
$$\begin{aligned}\oint_C (-q\,dx + p\,dy) &= \oint_C \mathbf{p}\cdot\mathbf{n}\,dl\\ &= \int_R \boldsymbol{\nabla}\cdot\mathbf{p}\,dS\\ &= \int_R \left(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}\right)dx\,dy\,,\end{aligned} \tag{5.9b}$$
where R is the region inside C, and $\mathbf{n}$ is the exterior normal.

Hence, by expanding in real and imaginary parts, and using the Cauchy-Riemann equations (5.4b),
$$\begin{aligned}\oint_C f(z)\,dz &= \oint_C (u+iv)\,(dx+idy)\\ &= \oint_C (u\,dx - v\,dy) + i\oint_C (v\,dx + u\,dy)\\ &= \int_R \left(-\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right)dx\,dy + i\int_R\left(-\frac{\partial v}{\partial y} + \frac{\partial u}{\partial x}\right)dx\,dy\\ &= 0\,. \qquad \square\end{aligned} \tag{5.9c}$$

5.3.1 Deforming contours

Consider two different contours, C_1 and C_2, from α to β. Then *if there are no singularities in the closed curve* $C = C_1 - C_2$, it follows from Cauchy's theorem, (5.9a), that

$$\oint_C f(z)\,dz = \int_{C_1} f(z)\,dz - \int_{C_2} f(z)\,dz = 0\,. \tag{5.10a}$$

Hence

$$\int_{C_1} f(z)\,dz = \int_{C_2} f(z)\,dz\,. \tag{5.10b}$$

Corollaries.

(i) We can deform a contour without changing the value of the integral *as long as we do not move the contour across a singularity.*

(ii) We can also deform a closed contour, if we can do so *without passing it through any singularities*, to find that, because the joins cancel,

$$\begin{aligned}\oint_C f(z)\,dz &= \oint_{C_1} f(z)\,dz - \oint_{C_2} f(z)\,dz\\ &= 0\,,\end{aligned} \tag{5.10c}$$

and hence that

19/19

$$\oint_{C_1} f(z)\,dz = \oint_{C_2} f(z)\,dz\,. \tag{5.10d}$$

Integration of the derivative. Suppose that f and f' are analytic in some simply-connected domain R. Consider the integral from $z = \alpha$ to β in the complex plane, where α, β and the integration contour lie in R, and parameterize the curve by s, where $0 \leqslant s \leqslant 1$. Then

$$\begin{aligned}\int_\alpha^\beta f'(z)\,dz &= \int_0^1 f'(z(s))\frac{dz}{ds}\,ds\\ &= \int_0^1 \frac{df(z(s))}{ds}\,ds\\ &= \Big[f(z(s))\Big]_0^1\\ &= f(\beta) - f(\alpha)\,.\end{aligned} \tag{5.11a}$$

Corollary (5.10d) confirms that this result is independent of the contour path *within* R, and hence integration is the inverse of differentiation in the same sense as for real functions of a real variable.

Remarks.

(i) If f is entire, i.e. has no singularities anywhere, then the integral $\int_\alpha^\beta f(z)\,dz$ does not depend at all on the path taken.

(ii) If f is analytic at z_0 it can be written as g' for some other function g that is analytic at z_0. For instance, using a Taylor series,

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n = g'\,, \qquad g(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{(n+1)!}(z-z_0)^{n+1}\,. \tag{5.11b}$$

19/21

Examples.

(i) If $f = z$, then f is entire, and independent of path

$$\int_0^i z\,dz = \tfrac{1}{2}(i^2 - 0^2) = -\tfrac{1}{2}\,. \tag{5.12a}$$

(ii) If $f = e^z$, then f is entire, and independent of path

$$\int_\alpha^\beta e^z\,dz = e^\beta - e^\alpha\,. \tag{5.12b}$$

(iii) If $f = z^{-2}$, then f is analytic except at $z=0$ where it has a singularity. it is now necessary to specify a contour so that it lies in a region where f is analytic, e.g. the straight line, C, connecting $1+i$ to $-1+i$

$$\int_C \frac{1}{z^2}\,dz = \left[-\frac{1}{z}\right]_{1+i}^{-1+i} = -\frac{1}{-1+i} + \frac{1}{1+i} = 1\,. \tag{5.12c}$$

19/20

(iv) Cauchy's theorem, (5.9a), does *not* hold if there is a singularity of f inside C, since $f(z)$ is then not analytic in R. Consider the case of $f(z) = 1/z$ when C is a contour anticlockwise around the unit circle. Then, from (5.6a) and (5.6b) with $C = C_2 - C_1$,

$$\begin{aligned}\oint_C f(z)\,dz &= \int_{C_2} \frac{1}{z}\,dz - \int_{C_1} \frac{1}{z}\,dz \\ &= i\int_0^{2\pi} d\theta = 2\pi i\,. \end{aligned}\tag{5.12d}$$

In this case the simple pole singularity of $f(z)$ at $z=0$ results in the non-zero answer.

19/18

5.4 Residues

Definition. It can be shown that any function, which is analytic and single-valued throughout an annulus $\alpha < |z-z_0| < \beta$ centred on $z = z_0$, has a unique *Laurent series* about $z=z_0$ which converges for all values of z within the annulus:

$$\begin{aligned} f(z) &= \sum_{n=-\infty}^{\infty} a_n (z-z_0)^n \\ &= \dots \frac{a_{-2}}{(z-z_0)^2} + \frac{a_{-1}}{z-z_0} + a_0 + a_1(z-z_0) + \dots\,. \end{aligned}\tag{5.13}$$

Remark. If $f(z)$ has a *single isolated* singularity at $z=z_0$, then $\alpha > 0$ can be made arbitrarily small.

Definition. If $a_{-n} = 0$ for $n > N$ but $a_{-N} \neq 0$, $f(z)$ is said to have a *pole of order N* at z_0.

Definitions. A pole of order one is a termed a *simple pole*, a pole of order two is termed a *double pole*, etc. If there is no such N (i.e. the expansion in negative powers of $z - z_0$ does not terminate) the point z_0 is an *essential singularity.*

Definition. The coefficient a_{-1} is called the *residue* of the pole.

- For a simple pole,

$$\operatorname*{res}_{z=z_0} f(z) \equiv a_{-1} = \lim_{z\to z_0}\{(z-z_0)f(z)\} = \lim_{z\to z_0}\left\{a_{-1} + a_0(z-z_o) + a_1(z-z_0)^2 + \dots\right\}. \tag{5.14a}$$

- For a pole of order N,

$$\begin{aligned}\operatorname*{res}_{z=z_0} f(z) \equiv a_{-1} &= \lim_{z\to z_0}\left\{\frac{1}{(N-1)!}\frac{d^{N-1}}{dz^{N-1}}\left[(z-z_0)^N f(z)\right]\right\}. \\ &= \lim_{z\to z_0}\left\{\frac{1}{(N-1)!}\frac{d^{N-1}}{dz^{N-1}}\left[a_{-N} + \dots + a_{-1}(z-z_0)^{N-1} + a_0(z-z_0)^N + \dots\right]\right\}\end{aligned}\tag{5.14b}$$

5.4.1 Calculating example residues

(i) The function $f(z) = e^z/z^3$ has a pole of order 3 at $z = 0$. The Laurent series about $z = 0$ is

$$f = \frac{e^z}{z^3} = \frac{1}{z^3}\sum_{r=0}^{\infty}\frac{z^r}{r!} = \frac{1}{0!\,z^3} + \frac{1}{1!\,z^2} + \frac{1}{2!\,z} + \frac{1}{3!} + \dots\,. \tag{5.15a}$$

Hence, the residue is $\frac{1}{2}$. Alternatively, we could use the formula, (5.14b), for the residue at $z = 0$ of a triple pole:

$$\lim_{z\to 0}\left\{\frac{1}{2!}\frac{d^2}{dz^2}\left[z^3 f(z)\right]\right\} = \frac{1}{2}\,. \tag{5.15b}$$

(ii) The function $f(z) = (z^8 - 1)^{-1}$ has 8 simple poles at $z = e^{ni\pi/4}$ where $n = 0, 1, 2, \dots, 7$. We could evaluate the residue at, say, $z = 1$, by factorising the denominator; however, using L'Hôpital's rule is more straightforward:

$$\operatorname*{res}_{z=1} f(z) = \lim_{z\to 1}\frac{z-1}{z^8-1} = \lim_{z\to 1}\frac{(z-1)'}{(z^8-1)'} = \lim_{z\to 1}\frac{1}{8z^7} = \frac{1}{8}\,. \tag{5.16}$$

(iii) The function $f(z) = 1/\sinh \pi z$ has simple poles at $z = in$ where n is any integer. Using L'Hôpital's rule again:

$$\begin{aligned}\operatorname*{res}_{z=ni} f(z) &= \lim_{z\to in}\frac{z-in}{\sinh \pi z}\\ &= \lim_{z\to in}\frac{1}{\pi\cosh \pi z}\\ &= \frac{1}{\pi \cosh ni\pi}\\ &= \frac{1}{\pi\cos n\pi} = \frac{(-1)^n}{\pi}\,.\end{aligned} \tag{5.17}$$

Alternatively, we could have used a Taylor series for $\sinh \pi z$.

Remark. When calculating residues it can be a case of *horses for courses*, i.e. some methods are easier than others depending on context.

Further examples can be found in Questions 2, 3 and 4 on Example Sheet 4.

5.5 Calculus of Residues

Consider the integral of a function, $\oint_C f(z)\,dz$, anticlockwise around a pole at $z = z_0$, where C is a *simple closed curve* in a *simply-connected domain*, within which the Laurent series converges.

Consider each term of the Laurent series, (5.13), separately.

For $n \geqslant 0$:

$$\oint_C a_n(z-z_0)^n\,dz = 0\,, \tag{5.18a}$$

from Cauchy's theorem for analytic functions.

For $n \leqslant -1$, shrink the contour to a circle of radius ε about z_0 and substitute $z = z_0 + \varepsilon e^{i\theta}$ to obtain

$$\begin{aligned}\oint_{C'} a_n(z-z_0)^n\,dz &= \int_0^{2\pi} a_n\varepsilon^n e^{in\theta}\,i\varepsilon e^{i\theta}\,d\theta = ia_n\varepsilon^{n+1}\int_0^{2\pi} e^{i(n+1)\theta}\,d\theta\\ &= \begin{cases} ia_n\varepsilon^{n+1}\left[\frac{e^{i(n+1)\theta}}{i(n+1)}\right]_0^{2\pi} & \text{if } n\leqslant -2\\ ia_n\varepsilon^{n+1}\,2\pi & \text{if } n=-1\end{cases}\\ &= \begin{cases} 0 & \text{if } n\leqslant -2\\ 2\pi i\,a_{-1} & \text{if } n=-1\end{cases}\,.\end{aligned} \tag{5.18b}$$

Therefore, from the form of the Laurent series (5.13), and playing fast and loose with re-ordering the infinite series and the integral,

$$\oint_C f(z)\,dz = \sum_{n=-\infty}^{\infty} \oint_C a_n (z-z_0)^n\,dz = 2\pi i\, a_{-1} = 2\pi i \operatorname*{res}_{z=z_0} f(z)\,. \tag{5.18c}$$

Residue theorem. This result leads to the *residue theorem* which states that if a function $f(z)$ is analytic in a simply-connected domain R except for a finite number of poles at $z = z_1, z_2, \ldots z_n$, and C is a simple closed curve that encircles the poles in an *anticlockwise* direction, then

$$\frac{1}{2\pi i}\oint_C f(z)\,dz = \sum_{k=1}^{n} \operatorname*{res}_{z=z_k} f(z)\,. \tag{5.19a}$$

Proof. To prove this, consider a new contour C' as shown. C' doesn't encircle any poles and so

$$\oint_{C'} f(z)\,dz = 0. \tag{5.19b}$$

We can express this integral as the sum of the integral around C and integrals encircling each of the poles *clockwise*, C_k, i.e. $C' = C + \sum_k C_k$:

$$\oint_{C'} f(z)\,dz = \oint_C f(z)\,dz + \sum_k \oint_{C_k} f(z)\,dz\,. \tag{5.19c}$$

The 'joining lines' give zero net contribution because the contribution going one way cancels with the contribution going the other way. Using (5.19b), and the result we derived for a contour encircling a single pole, (5.18c),

$$0 = \oint_{C'} f(z)\,dz = \oint_C f(z)\,dz - 2\pi i \sum_k \operatorname*{res}_{z=z_k} f(z)\,. \tag{5.19d}$$

20/19 The residue theorem, (5.19a), follows. □

5.6 Cauchy's Formula for $f(z)$

Cauchy's formula. If $f(z)$ is analytic in a region R containing z_0, then *Cauchy's formula* states that,

$$f(z_0) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-z_0}\,dz\,, \tag{5.20a}$$

where C is a simple closed curve in R encircling z_0 anticlockwise.

Proof. Since $f(z)/(z-z_0)$ is analytic except for a simple pole at $z = z_0$, from the residue theorem, (5.19a),

$$\frac{1}{2\pi i}\oint_C \frac{f(z)}{z-z_0}\,dz = \operatorname*{res}_{z=z_0}\left(\frac{f(z)}{z-z_0}\right) = \lim_{z\to z_0} f(z) = f(z_0)\,. \tag{5.20b}$$

Remarks.

(i) If we know $f(z)$ on C then, from Cauchy's formula (5.20a), we know $f(z)$ throughout the interior of C.

(ii) Since the real and imaginary parts, u and v, of an analytic function satisfy Laplace's equation, this is equivalent to the uniqueness theorem for solutions of Laplace's equation with Dirichlet boundary conditions. In particular, if we specify u and v on C, then there is a unique solution for u and v inside C; this is (3.49d), after relabelling, with $\rho = 0$ and with the Green's function G
20/20 given by (3.39g).

(iii) If we differentiate Cauchy's formula n times with respect to z_0 we obtain

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z)}{(z-z_0)^{n+1}} dz\,. \tag{5.21}$$

Therefore, at any point where f is analytic, all its derivatives exist and it is differentiable infinitely many times.

20/21 See also Question 5 on Example Sheet 4.

5.7 The Point at Infinity

Some functions, $f(z)$, tend to a definite limit as $z \to \infty$ *irrespective of the direction* from which infinity is approached; e.g. $f(z) = 1/z$ goes to zero as $|z| \to \infty$. Therefore, it sometimes makes sense to think of ∞ as a single point, as illustrated by the *stereographic projection* of the complex plane onto the *Riemann sphere*, a sphere with the south pole, say, at $z = 0$.

For any point on the plane we draw a line to the north pole of the sphere; the point where the line intersects the sphere is the equivalent point on the sphere. The south pole is projected onto the origin. Circles of fixed latitude are projected onto concentric circles centred on the origin.

Definition. The north pole is projected to all points on a circle of infinite radius: think of this as a single *point at infinity*.

We can study the behaviour of $f(z)$ near the point at infinity by defining a new complex variable

$$\zeta = \frac{1}{z}\,. \tag{5.22a}$$

The point at infinity in the z-plane is the origin in the ζ-plane, and vice-versa. Setting

$$g(\zeta) = f(1/\zeta)\,, \tag{5.22b}$$

we can find a Laurent expansion for g about $\zeta = 0$; if g has a singularity at $\zeta = 0$ then f has this singularity at infinity.

For example:

(i) if $f(z) = z^n$, then $g(\zeta) = \zeta^{-n}$ has a pole of order n at $\zeta = 0$, and f has a pole of order n at infinity.

(ii) if $f(z) = e^z$, then $g(\zeta) = e^{1/\zeta}$ has an essential singularity at $\zeta = 0$, and so f has an essential singularity at infinity.

(iii) if $f(z) = \frac{1}{z}$, then $g = \zeta$ has a simple zero at $\zeta = 0$, which we interpret to mean that $f(z)$ has a simple zero at infinity.

Non-examinable remark. Care must be taken when combining the idea of a point at infinity with the residue theorem because the residue is not strictly a property of f alone but of $f\,dz$. For example, if C is the *anticlockwise* unit circle in the z-plane, then from (5.12d)

$$\frac{1}{2\pi i} \oint_C \frac{dz}{z} = 1\,. \tag{5.23a}$$

But C is the *clockwise* unit circle in the ζ-plane, so C may also be viewed as a simple closed curve that encloses the point at infinity in the z-plane. The integral is therefore equal to minus the sum of the residues *outside* C.

As $1/z$ has no singularities in the complex plane away from $z = 0$, the residue at $z = \infty$ must be -1, even though the function $1/z$ is not singular there! This is indeed the case because

20/18

$$\frac{dz}{z} = -\frac{d\zeta}{\zeta}\,. \tag{5.23b}$$

5.8 Applications of the Calculus of Residues

We now consider some examples of using residue calculus. See also questions 6, 7 and 8 (parts i, ii and iii) on Example Sheet 4.

5.8.1 Integrals involving trigonometric functions

How might we calculate the integral,

$$I = \int_0^{2\pi} \frac{d\theta}{2(a - \cos\theta)}\,, \tag{5.24a}$$

where $a > 1$ is a real constant?

Consider the substitution

$$z = e^{i\theta}\,. \tag{5.24b}$$

This gives $dz = iz\,d\theta$ and $\cos\theta = \frac{1}{2}(z + z^{-1})$, while the integral between $\theta = 0$ and 2π corresponds to the integral over z around a unit circle C in the complex plane. Then, from (5.24a),

$$\begin{aligned} I &= \oint_C \frac{dz}{2iz\left(a - \frac{1}{2}(z + z^{-1})\right)} \\ &= i\oint_C \frac{dz}{z^2 - 2az + 1} \\ &= i\oint_C \frac{dz}{(z - z_+)(z - z_-)}\,, \end{aligned} \tag{5.24c}$$

where the integrand has simple poles at $z_\pm = a \pm \sqrt{a^2 - 1}$.

Because $a > 1$, it follows that $0 < z_- < 1$ and $z_+ > 1$. Hence the pole at z_- is inside C but that at z_+ is outside it. The residue at z_- is

$$\frac{i}{z_- - z_+} = -\frac{i}{2\sqrt{a^2 - 1}}\,. \tag{5.24d}$$

Therefore, from the residue theorem (5.19a),

$$I = \frac{\pi}{\sqrt{a^2 - 1}}\,. \tag{5.24e}$$

Remark. The same method can be used for other integrals involving trigonometric functions.

5.8.2 Closing a contour at infinity

Suppose that we wish to calculate the integral

$$I = \int_0^\infty \frac{dx}{x^2 + 1}\,. \tag{5.25a}$$

We already know how to do this using trigonometric substitutions, but it is also possible to calculate it using residue calculus. In particular, consider

$$\oint_C \frac{dz}{z^2 + 1} = \oint_{C_0 + C_R} \frac{dz}{(z + i)(z - i)}\,, \tag{5.25b}$$

where $C = C_0 + C_R$ consists of two parts: first a contour, C_0, from $-R$ to $+R$ along the real axis, and second a contour, C_R, clockwise along a semicircle of radius R in the upper half-plane.

The integrand has two simple poles, but only the one at $z = i$ is enclosed by C. Hence, from the residue theorem (5.19a),

$$\oint_C \frac{dz}{z^2+1} = \oint_{C_0+C_R} \frac{dz}{(z+i)(z-i)} = 2\pi i\, \frac{1}{2i} = \pi\,. \tag{5.25c}$$

We also have that, using the symmetry of the integrand,

$$\int_{C_0} \frac{dz}{z^2+1} \equiv \int_{-R}^{R} \frac{dz}{z^2+1} = 2\int_0^R \frac{dz}{z^2+1} \to 2\,I \quad \text{as } R\to\infty\,. \tag{5.25d}$$

Finally we need to consider the value of the integral along C_R. On this semicircle, the integrand is $O(R^{-2})$, while the contour has length πR. Hence,

$$\left|\int_{C_R} \frac{dz}{z^2+1}\right| \leqslant \int_{C_R} \frac{|dz|}{\min|z^2+1|} \leqslant \frac{\pi R}{R^2-1} \to 0 \quad \text{as } R\to\infty\,. \tag{5.25e}$$

Combining (5.25c), (5.25d) and (5.25e), and taking the limit $R\to\infty$, we conclude that

$$I = \frac{\pi}{2}\,. \tag{5.25f}$$

Generalisation to a double pole. Next consider the integral,

$$I = \int_0^\infty \frac{dx}{(x^2+a^2)^2}\,, \tag{5.26a}$$

where $a > 0$ is a real constant.

The analysis is very similar to the previous example. As before, expand the integral to be over a closed curve:

$$\oint_C \frac{dz}{(z^2+a^2)^2} = \oint_{C_0+C_R} \frac{dz}{(z+ia)^2(z-ia)^2}\,. \tag{5.26b}$$

The integrand has double poles at $z = \pm ia$ but the contour only encloses the pole at $z = +ia$. The residue there is, from (5.14b),

$$\lim_{z\to ia} \frac{d}{dz}\frac{1}{(z+ia)^2} = \lim_{z\to ia} -\frac{2}{(z+ia)^3} = -\frac{i}{4a^3}\,. \tag{5.26c}$$

Similarly, by using the symmetry of the integrand as in (5.25d),

$$\int_{C_0} \frac{dz}{(z^2+a^2)^2} = \int_{-R}^{R} \frac{dx}{(x^2+a^2)^2} \to 2I \quad \text{as } R\to\infty. \tag{5.26d}$$

Further, we can estimate the integral around the semicircle, C_R, similarly as in (5.25e):

$$\left|\int_{C_R} \frac{dz}{(z^2+a^2)^2}\right| \leqslant \frac{\pi R}{(R^2-a^2)^2} \to 0 \quad \text{as } R\to\infty\,. \tag{5.26e}$$

Therefore, we conclude that, on taking the limit $R\to\infty$,

$$2I = 2\pi i\left(-\frac{i}{4a^3}\right)\,, \quad \text{i.e.} \quad I = \frac{\pi}{4a^3}\,. \tag{5.26f}$$

21/19

Generalisation to multiple poles. Finally, consider the integral

$$I = \int_0^\infty \frac{dx}{x^4+1}\,. \tag{5.27a}$$

Again we adopt a similar approach and consider

$$\oint_C \frac{dz}{z^4+1} = \oint_{C_0+C_R} \frac{dz}{(z-z_1)(z-z_2)(z-z_3)(z-z_4)}\,, \tag{5.27b}$$

where the integrand has simple poles at $(-)^{\frac{1}{4}}$, i.e. at $z_1 = e^{i\pi/4}$, $z_2 = e^{3i\pi/4}$, $z_3 = e^{-3i\pi/4}$ and $z_4 = e^{-i\pi/4}$. Only two of those poles, i.e. those at z_1 and z_2, are enclosed by C. Using L'Hôpital's rule, these have residues

$$\lim_{z\to z_1} \frac{z-z_1}{z^4+1} = \lim_{z\to z_1} \frac{1}{4z^3} = \tfrac{1}{4}e^{-3i\pi/4} = -\tfrac{1}{4}e^{i\pi/4}\,, \tag{5.27c}$$

21/20

$$\lim_{z\to z_2} \frac{z-z_2}{z^4+1} = \tfrac{1}{4}e^{-i\pi/4}\,. \tag{5.27d}$$

The contribution from C_R can again be shown to tend to zero as $R \to \infty$, hence in a similar way to above

$$2I = 2\pi i \left(-\tfrac{1}{4}e^{i\pi/4} + \tfrac{1}{4}e^{-i\pi/4}\right) = \pi \sin\frac{\pi}{4}\,, \tag{5.27e}$$

i.e.

$$I = \frac{\pi}{2\sqrt{2}}\,. \tag{5.27f}$$

Alternative method. Alternatively, we could use a contour, $C' = C_1' + C_R' + C_2'$, that is confined to the first quadrant, and that only encloses the pole at z_1. Then, from (5.27c),

$$\oint_{C'} \frac{dz}{z^4+1} = 2\pi i \left(\operatorname*{res}_{z=z_1} \frac{1}{z^4+1}\right) = -\frac{i\pi}{2}e^{i\pi/4}\,. \tag{5.28a}$$

Again it is possible to show that the integral along C_R' tends to 0 as $R\to\infty$. It follows that

$$\begin{aligned}
-\frac{i\pi}{2}e^{i\pi/4} &= \int_{C_1'+C_R'+C_2'} \frac{dz}{z^4+1} \\
&= \lim_{R\to\infty}\left(\int_0^R \frac{dx}{x^4+1} + \int_R^0 \frac{i\,dy}{(iy)^4+1}\right) \qquad \text{substituting } z=x \text{ and } z=iy \\
&= (1-i)\int_0^\infty \frac{dx}{x^4+1} = (1-i)I\,,
\end{aligned} \tag{5.28b}$$

and so, as in (5.27f),

$$I = \frac{\pi}{2\sqrt{2}}\,. \tag{5.28c}$$

21/21

5.9 Multi-Valued Functions and Branch Cuts

$\ln z$. Not all complex functions have a single value for each complex point $z = re^{i\theta}$. For instance, the complex function $\ln z = \ln r + i\theta$ has infinitely many values, or *branches*, because θ can take infinitely many values, e.g.

$$\begin{aligned}
\ln i &= \ldots, -\tfrac{3}{2}i\pi, \tfrac{1}{2}i\pi,\ \tfrac{5}{2}i\pi,\ \tfrac{9}{2}i\pi,\ \ldots \\
&= \left(2n + \tfrac{1}{2}\right) i\pi \qquad \text{for } n \in \mathbb{Z}.
\end{aligned} \tag{5.29}$$

Consider the curves in the complex z-plane shown in the figure. On C_1 and C_2 we can always choose θ to be in some range no matter how many times we circle the curve, e.g. $0 < \theta < \frac{\pi}{2}$ for C_1 and $\frac{\pi}{2} < \theta < \frac{3\pi}{2}$ for C_2. Consequently, $\ln z$ is continuous and single-valued on such curves. However, this is not possible for C_3 which encircles the origin, and where 2π is added to θ each time the origin is circled. The consequence is that $\ln z$ must either be multi-valued or discontinuous.

Definition. A point that *cannot* be encircled by a curve on which the *function is continuous and single-valued*, is called a *branch point*; the function has a *branch point singularity* at that point. In the example of $\ln z$ there is a branch point at the origin.

Further examples of branch points.

(i) $\ln(z-a)$ has a branch point at $z=a$.

(ii) $\ln(z^2-1) = \ln(z-1) + \ln(z+1)$ has two branch points, i.e. at $z = \pm 1$.

(iii) $\sqrt{z} = \sqrt{r}e^{i\theta/2}$ is double-valued and has a branch point at the origin. Unlike for a real function, we can't turn this into a single-valued function by just choosing the sign of $\sqrt{z}$ because the sign changes if we circle the origin once. Any function $z^\alpha = r^\alpha e^{i\alpha\theta}$, where α is *not* an integer, has a branch point at the origin.

21/18 (iv) $\sqrt{z^2-a^2} = \sqrt{z-a}\sqrt{z+a}$ has two branch points, i.e. at $z = \pm a$.

Definition. In order to make a function with a branch point continuous and single valued *on a curve*, it is necessary that the curve does not encircle the branch point. To do this, introduce a *branch cut* that no curve is permitted to cross.

Definition. Having introduced a branch cut, a *branch* of a function can be defined, e.g. in the neighbourhood of the branch point by choosing values of θ in a 2π range around the branch point.

Example: $\ln z$. The *canonical* (standard) branch cut for $\ln z$ is along the real axis from $-\infty$ to the origin, so that $-\pi < \theta < \pi$. With this choice of branch cut, the value of $\ln z$ is called the *principal value* of the logarithm

Just above the cut, say at $z = x + i0^+$ with $x < 0$, the argument $\theta = \pi$ and $\ln z = \ln|x| + i\pi$. Just below the cut at $z = x + i0^-$, $\theta = -\pi$ and $\ln z = \ln|x| - i\pi$. If a curve did cross the cut, $\ln z$ would be discontinuous and so not analytic.

Remarks.

(i) There are, potentially, an infinite number of branches.

(ii) The function is analytic everywhere on each branch except on the branch cut.

(iii) The function is single-valued and continuous on any curve that does not cross the cut.

(iv) Branch cuts need not be straight lines. In the case of $\ln z$, *any* continuous non-intersecting curve from the branch point to infinity is acceptable. However, the position of the cut is an essential part of the definition of the function because the function is discontinuous across the cut.

Example: z^α, $\alpha \notin \mathbb{Z}$. The same branch cut choices as for $\ln z$ can be made for $z^\alpha = e^{\alpha \log z}$, where α is not an integer. When $\alpha \in \mathbb{Q}$, i.e. when α is a rational, there are a finite number of branches (e.g. $\sqrt{z}$ has two branches), while when α is irrational there are an infinite number of branches (as for $\ln z$).

Remark. Because of the branch cuts, neither $\ln z$ or z^α have Laurent expansions about the origin; any annulus $\alpha < |z| < \beta$ would be crossed by the branch cut and so the function would not be analytic in the annulus.

Example: $\sqrt{z^2-1}$. Consider

$$f(z) = \sqrt{z^2-1} = \sqrt{z-1}\sqrt{z+1}\,, \tag{5.30a}$$

a function that has branch points at $z = \pm 1$. Setting,

$$z - 1 = r_1 e^{i\theta_1} \quad \text{and} \quad z + 1 = r_2 e^{i\theta_2}\,, \tag{5.30b}$$

we see that

$$f(z) = \sqrt{r_1 r_2}\, e^{i(\theta_1+\theta_2)/2}\,. \tag{5.30c}$$

If $z = 1$ is encircled by a small curve C_1, then

$$\theta_1 \to \theta_1 + 2\pi\,, \quad \theta_2 \to \theta_2 \quad \text{and} \quad \tfrac{1}{2}(\theta_1 + \theta_2) \to \tfrac{1}{2}(\theta_1 + \theta_2) + \pi\,. \tag{5.30d}$$

Hence $f(z)$ changes sign because $e^{i\pi} = -1$. The same applies to a small curve C_2 encircling $z = -1$. However, going around a curve C_3 that encircles both branch points has the following effect:

$$\theta_1 \to \theta_1 + 2\pi\,, \quad \theta_2 \to \theta_2 + 2\pi \quad \text{and} \quad \tfrac{1}{2}(\theta_1 + \theta_2) \to \tfrac{1}{2}(\theta_1 + \theta_2) + 2\pi\,. \tag{5.30e}$$

Hence $f(z)$ does not change sign because $e^{2i\pi} = 1$.

This means that we could introduce a branch cut that goes from from $z = -1$ to $z = +1$; the simplest choice is to put the cut on the real axis. The two branches of $f(z)$ correspond to specifying that $f(z)$ is positive or negative on the positive real axis.

Alternatively, we could introduce 'two' separate branch cuts: one from each branch point to infinity.

Remark. These two branch cuts can be viewed as a single branch cut that happens to pass through the point at infinity. This is because the cut[s] can be smoothly deformed until they lie along the real axis.

Remark. In general, when there is more than one branch point we *may* need more than one branch cut.

See also Question 9 on Example Sheet 4.

5.9.1 Contour integration around a branch cut

We illustrate contour integration around a branch cut by evaluating the integral

$$I = \int_0^\infty \frac{x^\alpha}{1 + \sqrt{2}\,x + x^2}\,dx\,, \tag{5.31a}$$

22/19 where $-1 < \alpha < 1$.

To this end, consider the contour integral

$$\oint_C \frac{z^\alpha}{1 + \sqrt{2}\,z + z^2}\,dz\,. \tag{5.31b}$$

The integrand has a branch point at $z = 0$, and simple poles at $z_1 = e^{3i\pi/4}$ and $z_2 = e^{5i\pi/4}$.

Choice of branch cut. As a *rule of thumb*, it is appropriate to choose a branch cut along the integration range, i.e. in this case along the *positive* real axis; we then define the branch by choosing $0 \leqslant \theta < 2\pi$, where $z = re^{i\theta}$.

It is then necessary to use a *keyhole contour*, C, in order to avoid the branch point and the branch cut. We consider the individual contributions from each part of the contour, $C = C_1 + C_R + C_2 + C_\varepsilon$, in turn.

C_1. The contribution from the contour, C_1, *just above* the branch cut is

$$\int_\varepsilon^R \frac{x^\alpha}{1 + \sqrt{2}x + x^2}\,dx \;\to I\,, \tag{5.31c}$$

as $\varepsilon \to 0$ and $R \to \infty$.

C_2. Substituting $z = re^{2\pi i}$ into (5.31b), the contribution from the contour, C_2, *just below* the branch cut is (since $z = r$, $z^2 = r^2$ but $z^\alpha = r^\alpha e^{2\pi\alpha i}$),

$$\int_R^\varepsilon \frac{r^\alpha e^{2\pi\alpha i}}{1+\sqrt{2}r+r^2}\,dr \ \to -e^{2\pi\alpha i}\, I \quad \text{as } \varepsilon\to 0 \text{ and } R\to\infty. \tag{5.31d}$$

C_ε. Substituting $z = \varepsilon e^{i\theta}$ into (5.31b), we obtain, since $\alpha > -1$,

$$\int_{C_\varepsilon} \frac{z^\alpha}{1+\sqrt{2}\,z+z^2}\,dz = \varepsilon^{\alpha+1}\int_{2\pi}^{0} \frac{e^{i(\alpha+1)\theta}}{1+\sqrt{2}\,\varepsilon e^{i\theta}+\varepsilon^2 e^{2i\theta}}\, i\,d\theta \ \to 0 \quad \text{as } \varepsilon\to 0. \tag{5.31e}$$

C_R. Substituting $z = Re^{i\theta}$ into (5.31b), we obtain, since $\alpha < 1$,

$$\begin{aligned}\int_{C_R} \frac{z^\alpha}{1+\sqrt{2}\,z+z^2}\,dz &= R^{\alpha+1}\int_0^{2\pi} \frac{e^{i(\alpha+1)\theta}}{1+\sqrt{2}\,Re^{i\theta}+R^2e^{2i\theta}}\, i\,d\theta \\ &= R^{\alpha-1}\int_0^{2\pi} \frac{e^{i(\alpha+1)\theta}}{R^{-2}+\sqrt{2}\,R^{-1}e^{i\theta}+e^{2i\theta}}\, i\,d\theta \ \to 0 \quad \text{as } R\to\infty.\end{aligned} \tag{5.31f}$$

Therefore, as $\varepsilon\to 0$ and $R\to\infty$,

$$\oint_C \frac{z^\alpha}{1+\sqrt{2}z+z^2}\,dz = \oint_{C_1+C_R+C_2+C_\varepsilon} \frac{z^\alpha}{(z-z_1)(z-z_2)}\,dz \ \to (1-e^{2\pi\alpha i})I\,. \tag{5.31g}$$

Because both poles are inside C, the residue theorem gives

$$\operatorname*{res}_{z=z_1} \frac{z^\alpha}{(z-z_1)(z-z_2)} = \frac{z_1^\alpha}{z_1-z_2} = \frac{e^{3\pi\alpha i/4}}{e^{3i\pi/4}-e^{5i\pi/4}}\,, \tag{5.31h}$$

$$\operatorname*{res}_{z=z_2} \frac{z^\alpha}{(z-z_1)(z-z_2)} = \frac{z_2^\alpha}{z_2-z_1} = \frac{e^{5\pi\alpha i/4}}{e^{5i\pi/4}-e^{3i\pi/4}}\,. \tag{5.31i}$$

Hence, from (5.31g),

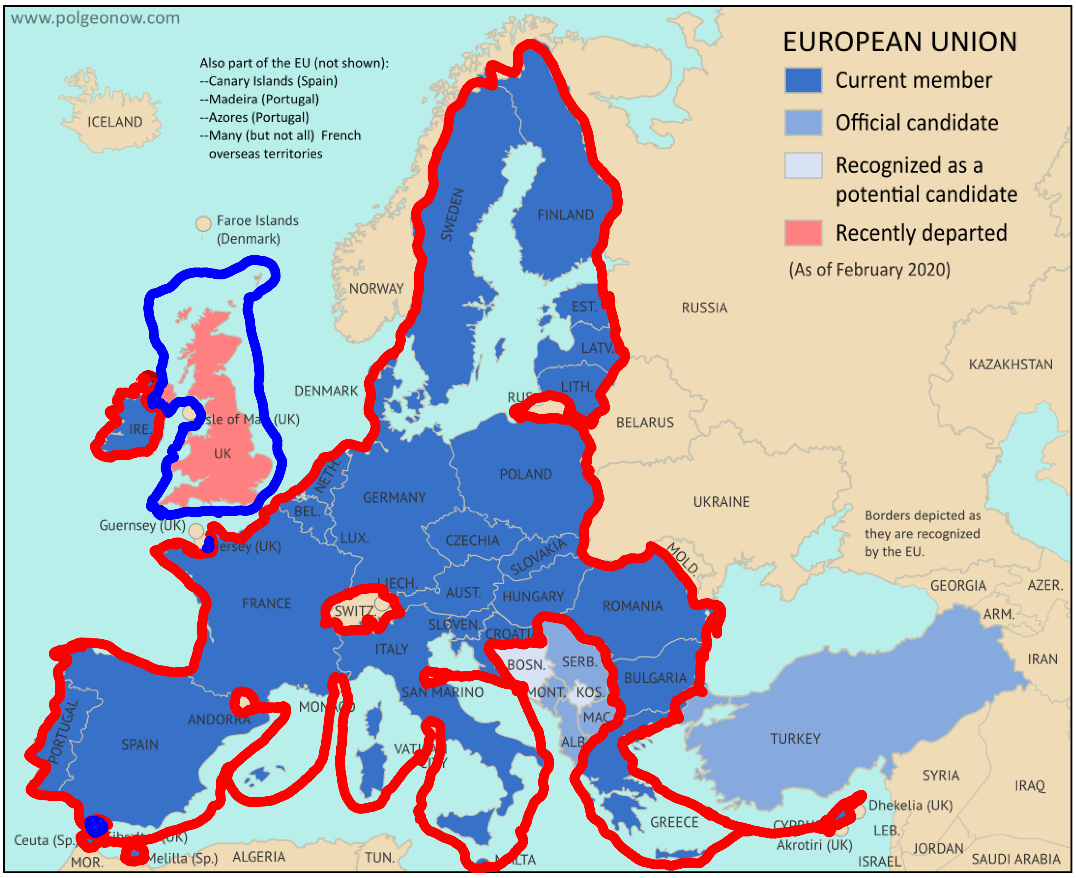
$$(1-e^{2\pi\alpha i})I = 2\pi i\left(\frac{e^{3\pi\alpha i/4}}{i\sqrt{2}} - \frac{e^{5\pi\alpha i/4}}{i\sqrt{2}}\right)\,, \tag{5.31j}$$

and so

$$\begin{aligned} I &= \sqrt{2}\pi\,\frac{e^{-\pi\alpha i/4}-e^{\pi\alpha i/4}}{e^{-\pi\alpha i}-e^{\pi\alpha i}} \\ &= \sqrt{2}\pi\,\frac{\sin(\alpha\pi/4)}{\sin(\alpha\pi)}\,. \end{aligned} \tag{5.31k}$$

22/20 Further examples can be found in Question 8 (parts iv and v) on Example Sheet 4.

5.9.2 Contour integral around the EU



Date	$\oint \mathrm{EU}\, dl$
Pre 23:00 GMT 31/01/20	$2\pi i\,(43.8)\,10^6$
Post 23:00 GMT 31/01/20	$2\pi i\,(43.0)\,10^6$

22/21

6 Transform Methods

6.0 Introduction

This section concerns the applications of contour integration to Fourier transform methods. However, first we need an important result.

6.1 Jordan's Lemma

Consider

$$\lim_{R\to\infty}\int_{C_R} g(z)e^{i\lambda z}\,dz\,, \tag{6.1a}$$

where

(i) $\lambda > 0$ is a real constant,

(ii) $g(z)$ is analytic in the upper half-plane, $\operatorname{Im} z > 0$, except possibly at a finite number of poles,

(iii) the contour C_R is a semicircle of radius R in the *upper half-plane.*

Jordan's Lemma. Jordan's lemma states that if $g(z) \to 0$ uniformly on C_R as $R \to \infty$, i.e. if, with $z = Re^{i\theta}$,

$$|g(Re^{i\theta})| \leqslant G(R) \text{ for } 0 < \theta < \pi\,, \tag{6.1b}$$

where $G(R) \to 0$ as $R \to \infty$, then

$$\lim_{R\to\infty}\int_{C_R} g(z)e^{i\lambda z}\,dz = 0\,. \tag{6.1c}$$

$\lambda < 0$. If instead $\lambda < 0$, with $g(z)$ again being analytic in the lower half-plane except possibly at a finite number of poles, then a similar result holds if the contour, C'_R, is taken to be a semicircle in the *lower half-plane.*

Proof using an extra restriction on g. Suppose that $g(z)$ satisfies the stronger bound

$$|g(Re^{i\theta})| \leqslant \frac{\mu}{R^2} \quad \text{as } R \to \infty\,, \tag{6.2a}$$

where μ is a positive constant. Then, since for $\lambda > 0$ and $y \geqslant 0$,

$$|e^{i\lambda z}| = |e^{i\lambda(x+iy)}| = e^{-\lambda y} \leqslant 1\,, \tag{6.2b}$$

we have, using (5.8e) and substituting $z = R\,e^{i\theta}$, $dz = iR\,e^{i\theta}d\theta$ and $|dz| = Rd\theta$, that

$$\begin{aligned}\left|\lim_{R\to\infty}\int_{C_R} g(z)e^{i\lambda z}\,dz\right| &\leqslant \lim_{R\to\infty}\int_{C_R}\frac{\mu}{R^2}R\,d\theta \\ &= \lim_{R\to\infty}\frac{2\pi\mu}{R} = 0\,. \qquad \square.\end{aligned} \tag{6.2c}$$

Remark: $\lambda < 0$. A similar argument applies for $\lambda < 0$ in which case we consider $y \leqslant 0$.

Proof using restriction (6.1b) *on g.* We start by observing that since the graph of $\sin\theta$ is concave on the interval $0 \leqslant \theta \leqslant \frac{\pi}{2}$, the graph of $\sin\theta$ lies above the straight line connecting its endpoints, i.e.

$$\tfrac{2}{\pi}\theta \leqslant \sin\theta \leqslant 1 \quad \text{for } 0 \leqslant \theta \leqslant \tfrac{\pi}{2}\,. \tag{6.3a}$$

Then, with $z = Re^{i\theta}$ and $dz = iRe^{i\theta}\,d\theta$,

$$
\begin{aligned}
\left|\int_{C_R} g(z)e^{i\lambda z}\,dz\right| &\leqslant \max_{C_R}|g(z)|\int_0^\pi \left|e^{i\lambda z}\right|\left|Re^{i\theta}\right|\,d\theta && \text{from (5.8e)}\\
&= \max_{C_R}|g(z)|\int_0^\pi \left|e^{i\lambda x}e^{-\lambda y}\right|\left|Re^{i\theta}\right|\,d\theta && \text{with } z = x+iy\\
&= R\,\max_{C_R}|g(z)|\int_0^\pi e^{-\lambda R\sin\theta}\,d\theta && \text{using } y = R\sin\theta\\
&= 2R\max_{C_R}|g(z)|\int_0^{\pi/2} e^{-\lambda R\sin\theta}\,d\theta && \text{using symmetry about } \theta = \tfrac{\pi}{2}\\
&\leqslant 2R\,\max_{C_R}|g(z)|\int_0^{\pi/2} e^{-2\lambda R\theta/\pi}\,d\theta && \text{using (6.3a)}\\
&= \frac{\pi}{\lambda}\left(1-e^{-\lambda R}\right)\max_{C_R}|g(z)|\\
&\to 0 \qquad \text{as } R\to\infty && \text{using (6.1b).} \quad \square \qquad (6.3b)
\end{aligned}
$$

See Question 10 on Example Sheet 4.

6.1.1 Example using Jordan's lemma

Consider the evaluation of

$$I = \int_{-\infty}^{+\infty}\frac{\sin x}{x}\,dx\,. \tag{6.4a}$$

First note that the integrand, $\sin x/x$, is well-behaved at the origin since, after extending into the complex plane,

$$\lim_{z\to 0}\frac{\sin z}{z} = 1\,. \tag{6.4b}$$

Second, since $\sin z = \frac{1}{2i}\left(e^{iz}-e^{-iz}\right)$, it should be possible to apply Jordan's lemma, but it will be necessary to close the contour in the upper/lower half-plane in the case of e^{iz}/e^{-iz} respectively.

Method 1. Because $\sin x/x$ is well-behaved at the origin, deform the integration contour near the origin into the lower half plane as illustrated, and then split the integral into two parts:

$$
\begin{aligned}
I &= \frac{1}{2i}\int_{-\infty}^{+\infty}\frac{e^{iz}-e^{-iz}}{z}\,dz\\
&= \frac{1}{2i}\int_{C_{L_\infty}}\frac{e^{iz}}{z}\,dz - \frac{1}{2i}\int_{C_{L_\infty}}\frac{e^{-iz}}{z}\,dz\,. \qquad (6.5a)
\end{aligned}
$$

In the case of the first term, which has a simple pole at $z=0$, close a [finite] contour, C_{L_R}, with a contour C_{R_U} in the upper half plane. Then, using the residue theorem, (5.19a),

$$\frac{1}{2\pi i}\oint_{C_{L_R}+C_{R_U}}\frac{e^{iz}}{z}\,dz = \operatorname*{res}_{z=0}\left(\frac{e^{iz}}{z}\right) = 1\,. \tag{6.5b}$$

From Jordan's lemma, (6.1c), the integral along C_{R_U} tends to zero as $R\to\infty$.

In the case of the second term in (6.5a), close the C_{L_R} contour in the lower half plane with C_{R_L}. Then, since no singularities are enclosed,

$$\frac{1}{2\pi i}\oint_{C_{L_R}+C_{R_L}}\frac{e^{-iz}}{z}dz = 0\,, \tag{6.5c}$$

where, from Jordan's lemma (6.1c), the integral along C_{R_L} tends to zero as $R \to \infty$. From (6.5b) and (6.5c) it follows from taking the limit $R \to \infty$, that

22/18
$$I = \frac{1}{2i} \int_{C_{L_\infty}} \left(\frac{e^{iz}}{z} - \frac{e^{-iz}}{z} \right) dz = \pi \,. \tag{6.5d}$$

Method 2. Similar to (6.5a) split up the integral, but this time on the real axis:

$$\begin{aligned} I &= \frac{1}{2i} \left(\int_{-\infty}^{+\infty} \frac{e^{iz}}{z}\, dz - \int_{-\infty}^{+\infty} \frac{e^{-iz}}{z} \right) dz \\ &= \operatorname{Im} \left(\int_{-\infty}^{+\infty} \frac{e^{iz}}{z}\, dz \right) \qquad \text{from complex conjugation.} \end{aligned} \tag{6.6a}$$

A difficulty now is that the contour passes through a pole, so instead consider the limit (actually a Cauchy principal value)

$$I = \operatorname{Im} \lim_{\varepsilon \to 0} \lim_{R \to \infty} \left(\int_{-R}^{-\varepsilon} \frac{e^{iz}}{z}\, dz + \int_{\varepsilon}^{R} \frac{e^{iz}}{z}\, dz \right) . \tag{6.6b}$$

Let C' be the contour as shown that encloses no poles. From Cauchy's Theorem, (5.9a),

$$\int_{-R}^{-\varepsilon} \frac{e^{iz}}{z}\, dz + \int_{C'_\varepsilon} \frac{e^{iz}}{z}\, dz + \int_{\varepsilon}^{R} \frac{e^{iz}}{z}\, dz + \int_{C'_R} \frac{e^{iz}}{z}\, dz = 0 \tag{6.6c}$$

On C'_ε, $z = \varepsilon e^{i\theta}$ and so,

$$\begin{aligned} \int_{C'_\varepsilon} \frac{e^{iz}}{z}\, dz &= \int_{\pi}^{0} \frac{\exp(i\varepsilon e^{i\theta})}{\varepsilon e^{i\theta}} i\varepsilon e^{i\theta}\, d\theta \\ &= -i \int_0^\pi \sum_{r=0}^{\infty} \frac{i^r \varepsilon^r e^{ir\theta}}{r!}\, d\theta \,. \end{aligned} \tag{6.6d}$$

Thence

$$\lim_{\varepsilon \to 0} \int_{C'_\varepsilon} \frac{e^{iz}}{z}\, dz = -i\pi \,. \tag{6.6e}$$

Further, from Jordan's lemma we know that the integral around C'_R vanishes as $R \to \infty$. Hence, taking the double limit $\varepsilon \to 0$ and $R \to \infty$, it follows from (6.6b), (6.6c) and (6.6e) that

$$I = \operatorname{Im}(i\pi) = \pi \,. \tag{6.6f}$$

Remark. Similar methods can be used to evaluate, say,

$$\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\, dx \,. \tag{6.7}$$

6.2 Fourier Transform Methods

Fourier transform. Recall that the *Fourier transform* of a function $f(x)$ of a real variable x is

$$\widetilde{f}(k) \equiv \mathcal{F}[f(x)] = \int_{-\infty}^{+\infty} f(x) e^{-ikx}\, dx \,. \tag{6.8a}$$

Inverse Fourier transform. The *inverse Fourier transform*,

$$f(x) \equiv \mathcal{F}^{-1}[\widetilde{f}(k)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widetilde{f}(k) e^{+ikx}\, dk \,, \tag{6.8b}$$

follows from the result that

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')}\, dk = \delta(x - x') \,. \tag{6.8c}$$

Warning. There are various definitions of Fourier transforms that differ in the signs of the exponents, and in where the 2π is, or $\sqrt{2\pi}$ are, placed.

Contour integrals. The integrals in (6.8a) and (6.8b) can be interpreted as contour integrals along the real
23/19 axis in the complex z-plane and complex k-plane respectively.

6.2.1 Damped harmonic oscillator

Consider the equation for the amplitude $x(t)$ of a driven damped harmonic oscillator,

$$\ddot{x}(t) + 2\gamma\dot{x}(t) + \omega_0^2 x(t) = f(t)\,, \tag{6.9}$$

where $f(t)$ is the forcing function, $\omega_0 > 0$ is real, and $\gamma > 0$ is real and represents the effects of damping due to, say, friction.

Assume that $x(t) \to 0$ as $|t| \to \infty$ so that we can define the Fourier transform, and its inverse, of $x(t)$ as

$$\widetilde{x}(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-i\omega t}\,dt \quad \text{and} \quad x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \widetilde{x}(\omega) e^{+i\omega t}\,d\omega\,, \tag{6.10a}$$

where, as is conventional, we use ω as the Fourier variable when the function depends on t.

Last term it was shown that the Fourier transforms of the first and second derivative of a function are related to the Fourier transform of the function by multiples of $i\omega$, i.e.

$$v(t) = \dot{x}(t) \quad \Leftrightarrow \quad \widetilde{v}(\omega) = i\omega\widetilde{x}(\omega) \quad \text{and} \quad a(t) = \ddot{x}(t) \quad \Leftrightarrow \quad \widetilde{a}(\omega) = -\omega^2\widetilde{x}(\omega)\,. \tag{6.10b}$$

Hence, by multiplying both sides of (6.9) by $e^{-i\omega t}$ and integrating over the real t axis, we deduce that

$$\left(-\omega^2 + 2i\gamma\omega + \omega_0^2\right)\widetilde{x}(\omega) = \widetilde{f}(\omega)\,. \tag{6.11a}$$

It follows that

$$\widetilde{x}(\omega) = \widetilde{f}(\omega)\,\widetilde{g}(\omega)\,, \tag{6.11b}$$

where

$$\widetilde{g}(\omega) = \frac{-1}{\omega^2 - 2i\gamma\omega - \omega_0^2} = \frac{-1}{(\omega - \omega_+)(\omega - \omega_-)} \qquad \text{and} \qquad \omega_\pm = i\gamma \pm \sqrt{\omega_0^2 - \gamma^2}\,. \tag{6.11c}$$

We can find $x(t)$ by taking the inverse Fourier transform:

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(\omega)\widetilde{g}(\omega)\,e^{i\omega t}\,d\omega\,. \tag{6.11d}$$

Recall that the convolution theorem states

$$h(t) = \int_{-\infty}^{\infty} f(s)g(t-s)\,ds \quad \Leftrightarrow \quad \widetilde{h}(\omega) = \widetilde{f}(\omega)\widetilde{g}(\omega)\,. \tag{6.12a}$$

Hence, we deduce that

$$x(t) = \int_{-\infty}^{\infty} f(s)g(t-s)\,ds\,, \tag{6.12b}$$

where, with $\tau = t - s$,

23/20

$$g(\tau) = \mathcal{F}^{-1}[\widetilde{g}(\omega)] = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega\tau}\,d\omega}{(\omega - \omega_+)(\omega - \omega_-)}\,. \tag{6.12c}$$

Remark. This is equivalent to a solution using the Green function $G(t,s) = g(t-s)$ of

$$\mathcal{L} = \frac{d^2}{dt^2} + 2\gamma\frac{d}{dt} + \omega_0^2\,. \tag{6.13}$$

To complete the solution to the problem we now have to determine $g(\tau)$ by integrating over ω. We will do
23/21 this by employing a contour integral in the complex ω-plane.

If $\tau < 0$, we choose a contour C that goes along the real axis and is closed with a semicircle in the *lower* half-plane. If $\tau > 0$ we instead close the contour with a semicircle in the *upper* half-plane. Since, from (6.11c), $g(\omega) \to 0$ as $|\omega| \to \infty$, Jordan's lemma, (6.1c), implies that in both cases the integral over the semicircle will vanish. It then follows that

$$\begin{aligned} g(\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{g}(\omega) e^{i\omega\tau}\, d\omega \\ &= \frac{1}{2\pi}\left(\int_{-\infty}^{\infty} + \int_{C_\infty}\right) \widetilde{g}(\omega) e^{i\omega\tau}\, d\omega \\ &= \frac{1}{2\pi}\oint_C \widetilde{g}(\omega) e^{i\omega\tau}\, d\omega\,. \end{aligned} \tag{6.14a}$$

As long as ω_0 is real, so that $\omega_0^2 > 0$, it follows from (6.11c) that the poles at $\omega = \omega_\pm$ are *both* in the upper half-plane. Therefore, from the residue theorem,

$$g(\tau) = 0 \quad \text{when} \quad \tau < 0\,. \tag{6.14b}$$

In other words, $g(t-s)$ is zero if $t < s$. Suppose that the forcing term is not switched on until $t = 0$, i.e. suppose that $f(t) = 0$ for $t < 0$, it follows from (6.12b) that

$$x(t) = \int_0^\infty f(s) g(t-s)\, ds = 0 \quad \text{for} \quad t < 0\,. \tag{6.14c}$$

Remark. This means that there is no response until the forcing term is switched on. This is *causal behaviour*, i.e. effect follows cause and not the other way around. The Green's function, $G(t,s) \equiv g(t-s)$, is said to be a *causal Green's function.*

For $\tau > 0$, there are two *simple* poles within C provided, as we [initially] assume, that $\gamma \neq \omega_0$. From (6.11c), the residues at $\omega = \omega_\pm$ are given by

$$\operatorname*{res}_{\omega=\omega_\pm}\left(\frac{1}{2\pi}\widetilde{g} e^{i\omega\tau}\right) = \frac{-e^{i\omega_\pm\tau}}{2\pi\,(\omega_\pm - \omega_\mp)} = \mp\frac{e^{-\gamma\tau} e^{\pm i\tau\sqrt{\omega_0^2-\gamma^2}}}{4\pi\sqrt{\omega_0^2-\gamma^2}}\,. \tag{6.14d}$$

Underdamped oscillator. For $\gamma < \omega_0$, the oscillator is said to be *underdamped.* We deduce from (6.14a) and the residue theorem that

$$g(\tau) = \frac{e^{-\gamma\tau}}{\sqrt{\omega_0^2-\gamma^2}} \sin\left(\tau\sqrt{\omega_0^2-\gamma^2}\right) \quad \text{when} \quad \tau > 0\,. \tag{6.14e}$$

Unit impulse. Suppose that there is a unit impulse at $t = 0$, i.e. $f(t) = \delta(t)$. It follows from (6.12b) that $x(t) = g(t)$, and hence from (6.14e) that the response to an impulsive force is oscillatory with an amplitude that dies away exponentially over a time of order $1/\gamma$. For $\gamma \ll \omega_0^2$ the main effect of the damping term is to cause the oscillations to slowly reduce in amplitude rather than change phase.

Overdamped oscillator. When $\gamma > \omega_0$ the oscillator is *overdamped*, and the damping prevents oscillation. See Question 12 on Example Sheet 4.

Critically damped oscillator. When $\gamma = \omega_0$, i.e. when there is *critical damping*, for $\tau > 0$ there is a *double* pole at $\omega = i\gamma$ inside the contour C. Expanding the integrand of (6.14a) about $\omega = i\gamma$ we obtain, using the residue formula (5.14b), that

$$\begin{aligned} \operatorname*{res}_{\omega=i\gamma}\left(\frac{\widetilde{g} e^{i\omega\tau}}{2\pi}\right) &= \operatorname*{res}_{\omega=i\gamma}\left(-\frac{e^{i\omega\tau}}{2\pi\,(\omega-i\gamma)^2}\right) \\ &= \lim_{\omega\to i\gamma}\left\{\frac{d}{d\omega}\left[-\frac{e^{i\omega\tau}}{2\pi}\right]\right\} \\ &= -\frac{i\tau e^{-\gamma\tau}}{2\pi}\,. \end{aligned} \tag{6.14f}$$

Hence the residue theorem yields

$$g(\tau) = \tau e^{-\gamma\tau} \quad \text{when} \quad \tau > 0\,. \tag{6.14g}$$

6.2.2 Gaussian integration lemma

The real integral

$$I = \int_{-\infty}^{\infty} e^{-(u+c)^2}\, du\,, \tag{6.15a}$$

where c is a *real* constant, can be evaluated by the substitution $x = u + c$ to obtain

$$I = \int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}\,. \tag{6.15b}$$

However, does this substitution continue to work if $c = a + ib$ is complex?

In this case extend the integral in to the complex plane, and define a new complex variable $z = u + c$. Then

$$I = \int_{-\infty}^{\infty} e^{-(u+c)^2}\, du = \int_{C_i} e^{-z^2}\, dz\,, \tag{6.15c}$$

where the contour C_i is the horizontal line in the complex z-plane with $\operatorname{Im} z = \operatorname{Im} c = b$, as illustrated (on the assumption, wlog, that $b > 0$).

The integrand e^{-z^2} is analytic everywhere and so the integral of e^{-z^2} around any closed contour is zero. In particular, consider the rectangular contour C_R with vertices at $\pm R$ and $\pm R + ib$.

Apply Cauchy's theorem to this contour to obtain

$$\begin{aligned}
0 &= \lim_{R\to\infty} \oint_{C_R} e^{-z^2}\, dz \\
&= \lim_{R\to\infty} \left\{ \int_{-R}^{R} e^{-z^2}\, dz + \int_0^b e^{-(R+iy)^2}\, i\, dy + \int_{R+ib}^{-R+ib} e^{-z^2}\, dz + \int_b^0 e^{-(-R+iy)^2}\, i\, dy \right\} \\
&= \sqrt{\pi} - I + \lim_{R\to\infty} 2e^{-R^2} \int_0^b e^{y^2} \sin(2Ry)\, dy\,.
\end{aligned} \tag{6.15d}$$

In the limit $R \to \infty$ the final term tends to zero, and so we deduce that, for any complex number c,

$$I = \int_{-\infty}^{\infty} e^{-(u+c)^2}\, du = \sqrt{\pi}\,. \tag{6.15e}$$

6.2.3 Solution to the diffusion equation

Consider the temperature $\theta(x,t)$ in an infinite one-dimensional bar with no source of heat. Then the temperature θ satisfies the diffusion equation, cf. (3.3b),

$$\frac{\partial \theta}{\partial t} = \lambda \frac{\partial^2 \theta}{\partial x^2}\,, \tag{6.16a}$$

where λ is the diffusion constant. Suppose that the initial temperature distribution at $t = 0$ is known,

$$\theta(x,0) = \theta_0(x)\,, \tag{6.16b}$$

and that we want to determine $\theta(x,t)$ at later times, $t > 0$.

This problem can be solved by taking the Fourier transform of θ with respect to x, where from (6.8a) and (6.8b),

$$\widetilde{\theta}(k,t) = \int_{-\infty}^{\infty} \theta(x,t)\, e^{-ikx}\, dx\,, \qquad \theta(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{\theta}(k,t)\, e^{ikx}\, dk\,. \tag{6.17a}$$

Recalling that Fourier transforms of the second derivative of a function is related to the Fourier transform of the function by a multiple of $-k^2$, cf. (6.10b), we find, on taking the Fourier transform of the diffusion equation (6.16a), that

$$\frac{\partial\widetilde{\theta}(k,t)}{\partial t} = -\lambda k^2\widetilde{\theta}(k,t)\,. \tag{6.17b}$$

The solution of this [ordinary] differential equation is

$$\widetilde{\theta}(k,t) = \widetilde{\theta}_0(k)e^{-\lambda k^2 t}\,, \tag{6.17c}$$

where

$$\widetilde{\theta}_0(k) = \widetilde{\theta}(k,0) = \int_{-\infty}^{\infty} e^{-ikx}\,\theta_0(x)\,dx\,. \tag{6.17d}$$

Rewrite equation (6.17c) as

$$\widetilde{\theta}(k,t) = \widetilde{\theta}_0(k)\widetilde{G}(k,t)\,, \tag{6.18a}$$

where

$$\widetilde{G}(k,t) = e^{-\lambda k^2 t}\,. \tag{6.18b}$$

Then, from the convolution theorem (6.12a), $\theta(x,t)$ is given by

$$\theta(x,t) = \int_{-\infty}^{\infty} \theta_0(y)G(x-y,t)\,dy\,, \tag{6.18c}$$

where, from the inverse Fourier Transform formula given in (6.17a),

$$G(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{+ikx-\lambda k^2 t}\,dk\,. \tag{6.18d}$$

To evaluate (6.18d), introduce a new integration variable $u = \sqrt{\lambda t}\,k$, complete the square in u, and use the Gaussian integration lemma to obtain

$$\begin{aligned} G(x,t) &= \frac{e^{-\frac{x^2}{4\lambda t}}}{2\pi\sqrt{\lambda t}}\int_{-\infty}^{\infty} e^{-\left(u-\frac{ix}{2\sqrt{\lambda t}}\right)^2}\,du \\ &= \frac{e^{-\frac{x^2}{4\lambda t}}}{\sqrt{4\pi\lambda t}} \qquad\qquad \text{from (6.15e)}\,. \end{aligned} \tag{6.18e}$$

Hence from (6.18c)

$$\theta(x,t) = \frac{1}{\sqrt{4\pi\lambda t}}\int_{-\infty}^{\infty} \theta_0(y)\,e^{-\frac{(x-y)^2}{4\lambda t}}\,dy\,. \tag{6.18f}$$

To proceed further we need $\theta_0(x)$. For example, suppose that,

$$\theta_0(x) = H(x) = \begin{cases} 1 & x>0 \\ 0 & x<0 \end{cases}\,. \tag{6.19a}$$

This might arise from an idealisation of two long bars, joined by an insulating layer, with each bar held at a constant temperature until the insulating layer is removed allowing heat to diffuse from one bar to the other.

Then, from (6.18f),

$$\begin{aligned} \theta(x,t) &= \frac{1}{\sqrt{4\pi\lambda t}}\int_0^{\infty} e^{-\frac{(x-y)^2}{4\lambda t}}\,dy = \frac{1}{\sqrt{\pi}}\int_{-\frac{x}{\sqrt{4\lambda t}}}^{\infty} e^{-v^2}\,dv \qquad \text{with } v=(y-x)/\sqrt{4\lambda t} \\ &= \frac{1}{\sqrt{\pi}}\left[\int_0^{\infty} e^{-v^2}\,dv + \int_0^{\frac{x}{\sqrt{4\lambda t}}} e^{-v^2}\,dv\right] \\ &= \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x}{\sqrt{4\lambda t}}\right)\right]\,, \end{aligned} \tag{6.19b}$$

where

$$\operatorname{erf}(y) \equiv \frac{2}{\sqrt{\pi}} \int_0^y e^{-v^2}\, dv\,, \tag{6.19c}$$

is the *error function.*

The error function

- is an odd function of y,
- has the asymptotic values $\operatorname{erf}(-\infty) = -1$ and $\operatorname{erf}(\infty) = 1$.

Checks. The initial conditions, (6.19a), are recovered in (6.19b) as $t \to 0$. Further, as might be expected, $\theta(x,t) \to \frac{1}{2}$ as $t \to \infty$ for fixed x.

24/19 For more examples see Questions 11-13 on Example Sheet 4.
24/20
24/21